Fake News, Voter Overconfidence, and the Quality of Democratic Choice

by Melis Kartal and Jean-Robert Tyran

ONLINE APPENDIX

A Proofs and Additional Results

A.1 Proofs of Results in Section IA

Proof of Lemma 1. The equilibrium characterization is shown in the main text in the discussion preceding Lemma 1. We now show that in an equilibrium where the correct policy is chosen with a probability (weakly) greater than 0.5 in both states, $q^a = \Pr(S = B|piv_a)$ and $q^b = \Pr(S = A|piv_b)$ must hold. Assume that the correct policy is chosen with a probability (weakly) greater than 0.5 in both states in equilibrium. This is true if and only if the relative turnout rate for the correct policy is weakly greater than 0.5. As a result, $\Pr(piv_b|A) \geq \Pr(piv_a|A)$ and $\Pr(piv_a|B) \geq \Pr(piv_b|B)$ must hold. Now, suppose towards a contradiction that $q^a \neq \Pr(S = B|piv_a)$. This implies that $\Pr(S = B|piv_b) > \Pr(S = B|piv_a)$ by (2), which in turn implies that $\Pr(S = A|piv_a) > \Pr(S = A|piv_b)$ because $\Pr(S = A|piv_j) = 1 - \Pr(S = B|piv_j)$ for $j \in \{a, b\}$. Thus, $q^b \neq \Pr(S = A|piv_b)$ by (3). More generally, $q^a \neq \Pr(S = B|piv_a)$ if and only if $q^b \neq \Pr(S = A|piv_b)$. Rewriting $\Pr(S = B|piv_b) > \Pr(S = B|piv_a)$ and using $\Pr(piv_b|A) \geq \Pr(piv_a|A)$, we obtain

$$\frac{(1-\pi)\Pr(piv_b|B)}{(1-\pi)\Pr(piv_b|B) + \pi\Pr(piv_a|A)} > \frac{(1-\pi)\Pr(piv_a|B)}{(1-\pi)\Pr(piv_a|B) + \pi\Pr(piv_a|A)},$$

which implies that $\Pr(piv_b|B) > \Pr(piv_a|B)$, a contradiction. Finally, as shown above $q^b = \Pr(S = A|piv_b)$ if and only if $q^a = \Pr(S = B|piv_a)$. Hence, the result is proved.

We next show that if $\pi = 0.5$, then $q^a = \Pr(S = B|piv_a)$ and $q^b = \Pr(S = A|piv_b)$ must hold. It is enough to show that $q^a = \Pr(S = B|piv_a)$ given that $q^b = \Pr(S = A|piv_b)$ if and only if $q^a = \Pr(S = B|piv_a)$. Assume towards a contradiction that $\Pr(S = B|piv_a) < \Pr(S = B|piv_b)$. Then, given the definitions of q^a and q^b in (2)-(3), it can be checked that there is full no abstention in equilibrium, and $q^a + q^b = 1$ must hold. Thus, either $q^a \geq 0.5$ or

$q^b \geq 0.5$ (or $q^a = q^b = 0.5$). Due to symmetry with $\pi = 0.5$, it is without loss of generality to only consider the case in which $q^a \geq 0.5$ and show that there is a contradiction. In particular, we will show that

$$q^a \Pr(piv_a|A) - (1-q^a)\Pr(piv_a|B) = (1-q^a)\Pr(piv_b|B) - q^a \Pr(piv_b|A)$$

cannot hold, but it is a necessary condition as an individual who obtains an α signal and has $q_i = q^a$ must be indifferent between voting for policy a and for policy b. There are two cases to consider: (i) $q^a > 0.5$ and (ii) $q^a = 0.5$. First, assume that $q^a > 0.5$. There are five cases to consider depending on the values of q^a, $\underline{q}$ and $\bar{q}$.

Case 1: $q^a \geq \bar{q}$, and $1 - q^a \leq \underline{q}$. Note that $1 - q^a \leq \underline{q}$ and $q^a \geq \bar{q}$ imply that $q^b \leq \underline{q}$ and $1 - q^b \geq \bar{q}$ since $q^a + q^b = 1$. Thus, this is analogous to a nonresponsive equilibrium in which no individual is pivotal since every individual votes for b regardless of their signal and accuracy, which we rule out.

Case 2: $\underline{q} < q^a < \bar{q}$, and $1 - q^a > \underline{q}$.[55] In this case, the benefit from voting for a for an individual who obtains an α signal with an accuracy of q^a equals

$$q^a \binom{N-1}{\lfloor \frac{N-1}{2} \rfloor} \lambda_A(a)^{\lfloor \frac{N-1}{2} \rfloor} \lambda_A(b)^{N-1-\lfloor \frac{N-1}{2} \rfloor} - (1-q^a) \binom{N-1}{\lfloor \frac{N-1}{2} \rfloor} \lambda_B(a)^{\lfloor \frac{N-1}{2} \rfloor} \lambda_B(b)^{N-1-\lfloor \frac{N-1}{2} \rfloor}, \quad (4)$$

where $\lambda_S(a)$ and $\lambda_S(b)$ represent the respective turnout rate for policy a and policy b in state S; i.e.,

$$\begin{aligned}
\lambda_A(a) &= \textstyle\int_{q^a}^{\bar{q}} q dF + \int_{\underline{q}}^{1-q^a} (1-q) dF \\
\lambda_A(b) &= \textstyle\int_{\underline{q}}^{q^a} q dF + \int_{1-q^a}^{\bar{q}} (1-q) dF \\
\lambda_B(a) &= \textstyle\int_{q^a}^{\bar{q}} (1-q) dF + \int_{\underline{q}}^{1-q^a} q dF \\
\lambda_B(b) &= \textstyle\int_{\underline{q}}^{q^a} (1-q) dF + \int_{1-q^a}^{\bar{q}} q dF
\end{aligned}$$

and the same individual's benefit from voting for b equals

$$(1-q^a) \binom{N-1}{\lfloor \frac{N-1}{2} \rfloor} \lambda_B(b)^{\lfloor \frac{N-1}{2} \rfloor} \lambda_B(a)^{N-1-\lfloor \frac{N-1}{2} \rfloor} - q^a \binom{N-1}{\lfloor \frac{N-1}{2} \rfloor} \lambda_A(b)^{\lfloor \frac{N-1}{2} \rfloor} \lambda_A(a)^{N-1-\lfloor \frac{N-1}{2} \rfloor}. \quad (5)$$

Recall that since $q^a + q^b = 1$, there is no abstention and thus $\lambda_S(a) + \lambda_S(b) = 1$ for $S \in \{A, B\}$. However, we will show that (4) is strictly greater than (5), which is a contradiction. To see

[55] Note that $1 - q^a \leq \bar{q}$ must always hold by the initial hypothesis that $q^a > 0.5$.

why, note that since $q^a > 0.5$, $\lambda_S(a) + \lambda_S(b) = 1$, and $\lambda_B(b)$ is the largest turnout term, $\lambda_A(a)\lambda_A(b) > \lambda_B(a)\lambda_B(b)$ must hold, and therefore,

$$q^a \binom{N-1}{\lfloor \frac{N-1}{2} \rfloor} \lambda_A(a)^{\lfloor \frac{N-1}{2} \rfloor} \lambda_A(b)^{\lfloor \frac{N-1}{2} \rfloor} (\lambda_A(a) + \lambda_A(b))$$

is greater than

$$(1 - q^a) \binom{N-1}{\lfloor \frac{N-1}{2} \rfloor} \lambda_B(b)^{\lfloor \frac{N-1}{2} \rfloor} \lambda_B(a)^{\lfloor \frac{N-1}{2} \rfloor} (\lambda_B(a) + \lambda_B(b)),$$

which ensures that the benefit from voting for a exceeds the benefit from voting for b with $s_i = \alpha$ and $q_i = q^a$, regardless of whether N is even or odd. Thus, there cannot be an equilibrium with $q^a < \bar{q}$, and $1 - q^a > \underline{q}$.

Case 3: $\underline{q} < q^a < \bar{q}$, and $1 - q^a \leq \underline{q}$. The steps in the proof of Case 2 still apply (with minor modifications in $\lambda_S(a)$ and $\lambda_S(b)$ due to $1 - q^a \leq \underline{q}$).

Case 4: $q^a \geq \bar{q}$, and $1 - q^a > \underline{q}$. The steps in the proof of Case 2 still apply (with minor modifications in $\lambda_S(a)$ and $\lambda_S(b)$ due to $q^a \geq \bar{q}$).

Case 5: $q^a \leq \underline{q}$. This case is possible only if $\underline{q} > 0.5$ since $q^a > 0.5$ by initial hypothesis. It also follows from $q^a + q^b = 1$ that $q^b < \underline{q}$ must hold. In this case, everyone votes and does so according to their signal. However, since we have a symmetric environment with $\pi = 0.5$, this implies that $q^a = q^b$. But then from $q^a + q^b = 1$, $q^a = 0.5$ must hold, a contradiction.

Hence, we have shown that $\Pr(S = B|piv_a) < \Pr(S = B|piv_b)$ cannot hold if $q^a > 0.5$ and $\pi = 0.5$. Next, assume that $q^a = 0.5$. From $q^a + q^b = 1$, $q^b = 0.5$ must hold. It is easy to see that this results in a case in which the correct policy is chosen with a probability greater than 0.5 in either state, and therefore, $\Pr(S = B|piv_a) < \Pr(S = B|piv_b)$ cannot hold given what we proved above.

Finally, we show that if $\pi = 0.5$ and $\underline{q} = 0.5$, then $q^a = q^b$. Suppose towards a contradiction (and without loss of generality) that $q^a > q^b$. There are three possibilities: either $q^a > q^b \geq 0.5$ or $q^a > 0.5 > q^b$ or $0.5 \geq q^a > q^b$.

(1) First, assume that $q^a > q^b \geq 0.5$.[56] In that case, voting is informative as no individual votes against their signal. In particular, by Lemma 1 every i with $s_i = \alpha$ and $q_i \geq q^a$ votes for policy a and every i with $s_i = \beta$ and $q_i \geq q^b$ votes for policy b. Consider the benefit from voting for a for an individual with $s_i = \alpha$ and $q_i = q^a$. This benefit, which we denote by

[56] Note that $q^a > q^b \geq \bar{q}$ can never hold as this implies no one votes in equilibrium, a contradiction. Our proof also works if $q^a > \bar{q}$ as pivotality conditions below adjust to that.

$\Pi^a(q^a, \alpha)$, equals

$$q^a \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_A)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} q dF \right)^{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^b}^{\bar{q}} (1-q) dF \right)^{t - \lfloor \frac{t}{2} \rfloor} -$$
$$(1-q^a) \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_B)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} (1-q) dF \right)^{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^b}^{\bar{q}} q dF \right)^{t - \lfloor \frac{t}{2} \rfloor},$$

where λ_S represents the turnout rate in state S; i.e., $\lambda_A = \int_{q^a}^{\bar{q}} q dF + \int_{q^b}^{\bar{q}} (1-q) dF$ and $\lambda_B = \int_{q^a}^{\bar{q}} (1-q) dF + \int_{q^b}^{\bar{q}} q dF$. Next, consider the benefit from voting for b for an individual with $s_i = \beta$ and $q_i = q^b$. This benefit, which we denote by $\Pi^b(q^b, \beta)$ equals

$$q^b \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_B)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} (1-q) dF \right)^{t - \lfloor \frac{t}{2} \rfloor} \left(\int_{q^b}^{\bar{q}} q dF \right)^{\lfloor \frac{t}{2} \rfloor} -$$
$$(1-q^b) \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_A)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} q dF \right)^{t - \lfloor \frac{t}{2} \rfloor} \left(\int_{q^b}^{\bar{q}} (1-q) dF \right)^{\lfloor \frac{t}{2} \rfloor}.$$

Note that an individual who obtains an α signal with an accuracy of q^a is indifferent between voting for a and abstaining, and an individual who obtains a β signal with an accuracy of $q^b \geq 0.5$ weakly prefers voting for b over abstaining; i.e., $\Pi^a(q^a, \alpha) = \Pi^b(q^b, \beta) = 0$ must hold. However, we will show that $\Pi^a(q^a, \alpha) > \Pi^b(q^b, \beta)$ resulting in a contradiction. To see why, first note that $\lambda_A < \lambda_B$ and thus, $1 - \lambda_A > 1 - \lambda_B$. To see why $\Pi^a(q^a, \alpha) > \Pi^b(q^b, \beta)$, first note that for realized turnout $t = 0$,[57] $q^a + 1 - q^b > q^b + 1 - q^a$, for $t > 0$ even,

$$\int_{q^a}^{\bar{q}} q dF \int_{0.5}^{q^b} (1-q) dF > \int_{q^a}^{\bar{q}} (1-q) dF \int_{q^b}^{\bar{q}} q dF, \tag{6}$$

and for $t > 0$ odd,

$$q^a \int_{q^b}^{\bar{q}} (1-q) dF + (1-q^b) \int_{q^a}^{\bar{q}} q dF > q^b \int_{q^a}^{\bar{q}} (1-q) dF + (1-q^a) \int_{q^b}^{\bar{q}} q dF. \tag{7}$$

It can be checked that (6) holds as $\int_{q^a}^{\bar{q}} q dF/(1 - F(q^a)) > \int_{0.5}^{q^b} q dF/(1 - F(q^b))$. Suppose towards a contradiction that (7) does not hold. But this implies that

$$q^a(1 - F(q^b)) - q^b(1 - F(q^a)) \leq \int_{q^b}^{q^a} q dF,$$

which cannot hold as $q^a > q^b$ and $q^a(F(q^a) - F(q^b)) > \int_{q^b}^{q^a} q dF$.

[57] If $q^a > \bar{q}$, only $t = 0$ and $t = 1$ are relevant.

(b) Next, assume that $q^a \geq 0.5 > q^b$. In that case, voting is informative only for those who obtain a β signal. In particular, since $0.5 > q^b$, by Lemma 1 every individual who obtains a β signal votes for policy b, every i who obtains an α signal but has $q_i \leq 1 - q^b$ votes for policy b and finally, every i who obtains an α signal and has $q_i \geq q^a$ votes for policy a (recall that $q^a \geq 1 - q^b$ holds by the definition of q^a and q^b). Consider the benefit from voting for a for an individual with $s_i = \alpha$ and $q_i = q^a$. This benefit, $\Pi^a(q^a, \alpha)$, equals

$$q^a \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_A)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} q dF \right)^{\lfloor \frac{t}{2} \rfloor} \left(\int_{0.5}^{1-q^b} q dF + \int_{0.5}^{\bar{q}} (1-q) dF \right)^{t - \lfloor \frac{t}{2} \rfloor} -$$
$$(1-q^a) \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_B)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} (1-q) dF \right)^{\lfloor \frac{t}{2} \rfloor} \left(\int_{0.5}^{1-q^b} (1-q) dF + \int_{0.5}^{\bar{q}} q dF \right)^{t - \lfloor \frac{t}{2} \rfloor},$$

where λ_S represents the turnout rate in state S; i.e., $\lambda_A = \int_{q^a}^{\bar{q}} q dF + \int_{0.5}^{1-q^b} q dF + \int_{0.5}^{\bar{q}} (1-q) dF$, and $\lambda_B = \int_{q^a}^{\bar{q}} (1-q) dF + \int_{0.5}^{1-q^b} (1-q) dF + \int_{0.5}^{\bar{q}} q dF$. Next, consider the benefit from voting for b for an individual with $s_i = \beta$ and $q_i = 1 - q^b$ (note that by hypothesis $1 - q^b > 0.5$).[58] This benefit, denoted by $\Pi^b(1 - q^b, \alpha)$, equals

$$q^b \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_B)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} (1-q) dF \right)^{t - \lfloor \frac{t}{2} \rfloor} \left(\int_{0.5}^{1-q^b} (1-q) dF + \int_{0.5}^{\bar{q}} q dF \right)^{\lfloor \frac{t}{2} \rfloor} -$$
$$(1-q^b) \sum_{t=0}^{N-1} \binom{N-1}{t} (1-\lambda_A)^{N-1-t} \binom{t}{\lfloor \frac{t}{2} \rfloor} \left(\int_{q^a}^{\bar{q}} q dF \right)^{t - \lfloor \frac{t}{2} \rfloor} \left(\int_{0.5}^{1-q^b} q dF + \int_{0.5}^{\bar{q}} (1-q) dF \right)^{\lfloor \frac{t}{2} \rfloor}.$$

First, assume that $q^a > 1 - q^b$. Then, an individual who obtains an α signal with an accuracy of q^a is indifferent between voting for a and abstaining, and an individual who obtains an α signal with an accuracy of $1 - q^b > 0.5$ is indifferent between voting for b and abstaining; i.e., $\Pi^a(q^a, \alpha) = \Pi^b(1 - q^b, \alpha) = 0$ must hold. Next, assume that $q^a = 1 - q^b$. In this case, there is no abstention in equilibrium and an individual who obtains an α signal with an accuracy of q^a is indifferent between voting for a and voting for b (and prefers either one over abstaining). In equilibrium, $\Pi^a(q^a, \alpha) = \Pi^b(1 - q^b, \alpha)$ must hold. However, we will now show that $\Pi^a(q^a, \alpha) > \Pi^b(1 - q^b, \alpha)$ holds, resulting in a contradiction. To see why, first note that $\lambda_A < \lambda_B$ and thus, $1 - \lambda_A > 1 - \lambda_B$ if $q^a > 1 - q^b$ and $\lambda_A = \lambda_B = 1$ if $q^a = 1 - q^b$. Then, to show $\Pi^a(q^a, \alpha) > \Pi^b(1 - q^b, \alpha)$, it is enough to note that $q^a + 1 - q^b > q^b + 1 - q^a$

[58] We assume that $1 - q^b \leq \bar{q}$ which is without loss of generality for the results.

for realized turnout $t = 0$, and for $t > 0$, it is enough to show that

$$\left(\int_{q^a}^{\bar{q}} qdF\right)\left(\int_{0.5}^{1-q^b} qdF + \int_{0.5}^{\bar{q}}(1-q)dF\right) > \left(\int_{q^a}^{\bar{q}}(1-q)dF\right)\left(\int_{0.5}^{1-q^b}(1-q)dF + \int_{0.5}^{\bar{q}} qdF\right), \tag{8}$$

and that

$$q^a\left(\int_{0.5}^{1-q^b} qdF + \int_{0.5}^{\bar{q}}(1-q)dF\right) + \left(1-q^b\right)\int_{q^a}^{\bar{q}} qdF > \tag{9}$$

$$q^b\int_{q^a}^{\bar{q}}(1-q)dF + (1-q^a)\left(\int_{0.5}^{1-q^b}(1-q)dF + \int_{0.5}^{\bar{q}} qdF\right).$$

Suppose towards a contradiction that (8) does not hold. This implies that

$$\int_{q^a}^{\bar{q}} qdF \le \int_{q^a}^{\bar{q}} dF \int_{0.5}^{1-q^b} dF - \int_{q^a}^{\bar{q}} qdF \int_{0.5}^{1-q^b} dF - \int_{q^a}^{\bar{q}} dF \int_{0.5}^{1-q^b} qdF + \int_{q^a}^{\bar{q}} dF \int_{0.5}^{\bar{q}} qdF$$

must hold. This inequality implies in turn that

$$\int_{q^a}^{\bar{q}} qdF(1+F(1-q^b)) + (1-F(q^a)\int_{0.5}^{1-q^b} qdF \le (1-F(q^a))(F(1-q^b) + \int_{0.5}^{\bar{q}} qdF).$$

Dividing both sides by $1-F(q^a)$ and then taking $\int_{0.5}^{1-q^b} qdF$ to the right-hand side, we obtain

$$\frac{\int_{q^a}^{\bar{q}} qdF}{1-F(q^a)}(1+F(1-q^b)) \le F(1-q^b) + \int_{1-q^b}^{\bar{q}} qdF.$$

However, writing $(1+F(1-q^b))$ above as $(1-F(1-q^b)+2F(1-q^b))$ and noting that $\frac{\int_{q^a}^{\bar{q}} qdF}{1-F(q^a)} > q^a$, it can be checked that the right-hand side of the inequality above is strictly greater than $\frac{\int_{q^a}^{\bar{q}} qdF}{1-F(q^a)}(1-F(1-q^b)) + 2q^aF(1-q^b)$, which in turn is greater than $F(1-q^b) + \int_{1-q^b}^{\bar{q}} qdF$ because $2q^a \ge 1$ and by $q^a \ge 1-q^b$, $\frac{\int_{q^a}^{\bar{q}} qdF}{1-F(q^a)} \ge \frac{\int_{1-q^b}^{\bar{q}} qdF}{1-F(1-q^b)}$ holds, which is a contradiction. Thus, (8) must hold. Next, suppose towards a contradiction that (9) does not hold. This implies that $q^a + \int_{q^a}^{\bar{q}} qdF \le q^b(1-F(q^a)) + \int_{1-q^b}^{\bar{q}} qdF + (1-q^a)F(1-q^b)$, and so, $q^a(1+F(1-q^b)) \le q^b(1-F(q^a)) + \int_{1-q^b}^{q^a} qdF + F(1-q^b)$. Noting that $\int_{1-q^b}^{q^a} qdF < q^a(F(q^a)-F(1-q^b))$, this implies $q^a(1+2F(1-q^b)-F(q^a)) < q^b(1-F(q^a)) + F(1-q^b)$. However, this cannot hold as $q^a > q^b$ and $2q^a > 1$.

(c) Finally, we rule out the case where $0.5 \ge q^a > q^b$. As mentioned above, it can be checked from the definition of q^a and q^b that $q^a \ge 1-q^b$ must hold. Thus, it is not possible to have $0.5 \ge q^a > q^b$.

Proof of Lemma 2. First, consider the case in which $\underline{q} < 0.5$. Let $\hat{q}$ be such that

$\hat{q} = \int_{\underline{q}}^{0.5}(1-q)dF + \int_{0.5}^{\bar{q}} q dF$ — this is the expected precision if i votes for (against) s_i for every $q_i > 0.5$ ($q_i < 0.5$). Obviously, $\hat{q} > 0.5$. Next, let π^* be such that

$$\pi^* = \sum_{t=\frac{N}{2}+1}^{N} \binom{N}{t} (\hat{q})^t (1-\hat{q})^{N-t} + 0.5 \binom{N}{\frac{N}{2}} (\hat{q})^{\frac{N}{2}} (1-\hat{q})^{\frac{N}{2}}$$

for N even, and for N odd, let $\pi^* = \sum_{t=\frac{N+1}{2}}^{N} \binom{N}{t} (\hat{q})^t (1-\hat{q})^{N-t}$. Since $\hat{q} > 0.5$, $\pi^* > 0.5$. In fact, $\pi^* > \hat{q}$ for $N \geq 3$ ($\pi^* = \hat{q}$ for $N = 1, 2$) by Lemma 4 below. By construction, the optimal symmetric equilibrium must be a responsive equilibrium for every $\pi \in (1-\pi^*\pi^*)$. This is because: (i) the efficiency of a nonresponsive equilibrium cannot be greater than $\max\{1-\pi, \pi\}$; (ii) there exists an optimal symmetric strategy by Lemma 3 below, which in turn is the optimal symmetric equilibrium; and (iii) by construction, the expected accuracy in the optimal symmetric equilibrium with $\pi \in (1-\pi^*\pi^*)$ must be higher than π^*. Next, we show that the optimal equilibrium must have an interior cutoff for every $\pi \in (1-\pi^*\pi^*)$. Suppose not. We can immediately rule out the case in which no individual votes as this is clearly suboptimal. Moreover, the case in which every i who receives an α signal (a β signal) votes for a (b) and every i who receives a β signal (an α signal) abstains cannot be an equilibrium. It is enough to consider the case where every i who receives an α signal votes for a and every i who receives a β signal abstains (the other case is symmetric). By Lemma 1, this implies that $q^a \leq \underline{q} < 0.5$ (this is necessary for i with $s_i = \alpha$ and $q_i \geq \underline{q}$ to prefer voting for a over abstention or voting for b), and as a result, $1 - q^a > 0.5$. Thus, every individual with $s_i = \beta$ and $q_i < 1 - q^a$ must strictly prefer voting for a, contradicting the strategy. In a similar vein, Lemma 1 rules out the case in which every i with $s_i = \alpha$ ($s_i = \beta$) votes for b (a) and every i with $s_i = \beta$ ($s_i = \alpha$) abstains cannot be an equilibrium. The remaining cases are the cases in which every i votes either always for or always against s_i regardless of q_i. Consider the former case. Again, by Lemma 1, this cannot be an equilibrium because if $q^a \leq \underline{q} < 0.5$ then $1-q^a > 0.5$, and every individual with $s_i = \beta$ and $q_i < 1-q^a$ must strictly prefer voting for a over abstention or voting for b, a contradiction. The latter strategy can also not be part of an equilibrium as, by Lemma 1, $\bar{q} \leq 1 - q^b$ must hold (this is necessary for every i with $s_i = \alpha$ to prefer voting for b over abstention or voting for a), but this implies that $q^b < 0.5$. Therefore, every i with $s_i = \beta$ and $q_i > q^b$ (for example $q_i = 0.5$) must strictly prefer voting for b over abstention or voting for a, which is a contradiction. Hence, the optimal equilibrium must have an interior cutoff for every $\pi \in (1-\pi^*, \pi^*)$.

Next, consider the case in which $\underline{q} = 0.5$. Let π^* be such that

$$\pi^* = \sum_{t=\frac{N}{2}+1}^{N} \binom{N}{t} (\mathbb{E}(q))^t (1-\mathbb{E}(q))^{N-t} + 0.5 \binom{N}{\frac{N}{2}} (\mathbb{E}(q))^{\frac{N}{2}} (1-\mathbb{E}(q))^{\frac{N}{2}}$$

for N even, and for N odd, let $\pi^* = \sum_{t=\frac{N+1}{2}}^{N} \binom{N}{t} (\mathbb{E}(q))^t (1-\mathbb{E}(q))^{N-t}$, where $\mathbb{E}(q) = \int_{0.5}^{\bar{q}} q dF$. By Lemma 4 below, $\pi^* > \mathbb{E}(q)$ if $N \geq 3$ (and $\pi^* = \mathbb{E}(q)$ if $N = 1, 2$). First, consider the case in which N is even. As shown above, by construction, the optimal equilibrium must be a responsive equilibrium for every $\pi \in (1-\pi^*, \pi^*)$. Next, we show that the optimal equilibrium must have an interior cutoff for $\pi \in (1-\pi^*, \pi^*)$. Suppose not. We can immediately rule out the case where no individual votes. Next, it can be shown that the case in which every i with $s_i = \alpha$ ($s_i = \beta$) votes for a (b) and every i with $s_i = \beta$ ($s_i = \alpha$) abstains cannot be an equilibrium (except possibly in one knife-edge case). It is enough to consider the case where every i with $s_i = \alpha$ votes for a and every i with $s_i = \beta$ abstains. In this case, $0.5 = \frac{(1-\pi)\Pr(piv_a|S=B)}{\pi \Pr(piv_a|S=A)+(1-\pi)\Pr(piv_a|S=B)}$ and $\bar{q} < 1$ must hold by Lemma 1. This equality requires π to satisfy $\pi = \frac{\left(\int_{0.5}^{\bar{q}} q dF\right)^{N-1}}{\left(\int_{0.5}^{\bar{q}} q dF\right)^{N-1} + \left(\int_{0.5}^{\bar{q}} (1-q) dF\right)^{N-1}}$ (as well as $\pi > \bar{q}$), a nongeneric case we rule out. In a similar vein, Lemma 1 rules out the case in which every i with $s_i = \alpha$ ($s_i = \beta$) votes for b (a) and every i with $s_i = \beta$ ($s_i = \alpha$) abstains. Remaining cases are the cases in which every i votes either always for or always against s_i regardless of q_i. Consider the former case. This cannot be an equilibrium because it can be shown that given the described strategy, $q^a = q^b > 0.5$ if $\pi = 0.5$ (since N is even), $q^a > 0.5$ if $\pi < 0.5$ and $q^b > 0.5$ if $\pi > 0.5$, but this contradicts with the hypothesized strategy since $\underline{q} = 0.5$. Now, consider the latter case. By Lemma 1, $\bar{q} \leq 1 - q^b$ must hold (this is necessary for i with α signal and $q_i \leq \bar{q}$ to prefer voting for b over voting for a), but this implies that $q^b < 0.5$, and therefore, every individual with a β signal must strictly prefer voting for b over abstention or voting for a, which is a contradiction. The proof for the case in which N is odd is analogous except that the optimal equilibrium with $\pi = \frac{1}{2}$ may not involve interior cutoffs.

Finally, consider the case in which $\underline{q} > 0.5$. First, let $\hat{\pi}$ be analogous to π^* as defined in the case with $\underline{q} = 0.5$ above. As shown in Lemma 4, $\hat{\pi} > \mathbb{E}(q)$ if $N \geq 3$ and $\hat{\pi} = \mathbb{E}(q)$ otherwise. We now define $\pi^* = \min\{\bar{q}, \hat{\pi}\}$. First, consider the case in which N is even. Similar to our discussion above with $\underline{q} \leq 0.5$, for every $\pi \in (1-\pi^*, \pi^*)$, the optimal equilibrium must be a responsive equilibrium (note that $\underline{q} < \mathbb{E}(q) < \pi^*$). We will now show that the optimal equilibrium must have an interior cutoff if $\pi \in (1-\pi^*, 1-\underline{q}] \cup [\underline{q}, \pi^*)$.

Suppose not. We can immediately rule out the case in which no individual votes. Moreover, by construction, the case in which every i with $s_i = \alpha$ ($s_i = \beta$) votes for a (b) and every i with $s_i = \beta$ ($s_i = \alpha$) abstains cannot be an equilibrium. It is enough to consider the case where every i with $s_i = \alpha$ votes for a and every i with $s_i = \beta$ abstains. Since $\pi^* \leq \bar{q}$ by definition, and since $\Pr(piv_b|S = B) > \Pr(piv_b|S = A)$ given the prescribed strategy, $q^b = \frac{\pi \Pr(piv_b|S=A)}{\pi \Pr(piv_b|S=A)+(1-\pi)\Pr(piv_b|S=B)} < \pi < \bar{q}$ holds for every $\pi \in (1-\pi^*, 1-\underline{q}] \cup [\underline{q}, \pi^*)$. Thus, abstention cannot be optimal for every i if $s_i = \beta$. As in the previous cases above, Lemma 1 rules out the case in which every i with $s_i = \alpha$ ($s_i = \beta$) votes for b (a) and every i with $s_i = \beta$ ($s_i = \alpha$) abstains. Remaining cases are the cases in which every i votes either always for or always against s_i regardless of q_i. Consider the former case. This cannot be an equilibrium because given the described strategy, $q^a > \underline{q}$ must hold if $\pi \leq 1 - \underline{q}$ (because $\Pr(piv_a|S = A) < \Pr(piv_a|S = B)$ and $q^a = \frac{(1-\pi)\Pr(piv_a|S=B)}{\pi \Pr(piv_a|S=A)+(1-\pi)\Pr(piv_a|S=B)} > 1 - \pi \geq \underline{q}$) and similarly $q^b > \underline{q}$ must hold if $\pi \geq \underline{q}$, which contradicts the voting strategy described. Now, consider the latter case. In this case, $\bar{q} \leq 1 - q^b$ must hold by Lemma 1, but this implies that $q^b < 0.5$, and therefore, an individual with a β signal must strictly prefer voting for b over abstention or voting for a, which is a contradiction. The proof for the case where N is odd and $\pi \in (1-\pi^*, 1-\underline{q}) \cup (\underline{q}, \pi^*)$ is analogous.

For the existence of a responsive equilibrium, see Corollary 1 below.

Proof of Proposition 1. We first prove the existence of the optimal symmetric strategy in Lemma 3. To do that, we first define a "cutoff strategy": a cutoff strategy consists of four cutoffs $\underline{q}^j \in [\underline{q}, \bar{q}]$ and $\bar{q}^j \in [\underline{q}, \bar{q}]$, $j \in \{a, b\}$ such that individual i votes for policy a if $s_i = \alpha$ and $q_i \geq \bar{q}^a$ or if $s_i = \beta$ and $q_i \leq \underline{q}^a$, and i votes for policy b if $s_i = \beta$ and $q_i \geq \bar{q}^b$ or if $s_i = \alpha$ and $q_i \leq \underline{q}^b$.

Lemma 3 *For every symmetric (measurable-)strategy that is not a cutoff strategy, there exists a cutoff strategy that strictly dominates it. As a result, by the Weierstrass theorem, there exists an optimal strategy among all symmetric measurable strategies.*

Proof. Let $\lambda : [\underline{q}, \bar{q}] \times \{\alpha, \beta\} \to [0, 1] \times [0, 1]$ represent a strategy that maps every q_i and s_i to a probability of voting for the policy that matches s_i and to a probability of voting for the opposite policy (i abstains with the remaining probability). Next, let $T_j(S)$ denote the expected turnout rate for policy $j \in \{a, b\}$ in state $S \in \{A, B\}$. Fix an arbitrary symmetric strategy that doesn't have the cutoff form. In particular, abusing notation let $\lambda_s^j(q)$ represent the probability with which i votes for $j \in \{a, b\}$ given $s_i = s$ and $q_i = q$, and assume that $\lambda_s^j(q)$ does not have a cutoff form for at least one (j, s) pair. Assuming that s is the signal consistent with state S (i.e., α for A and β for B) and $s \neq s'$,

$$T_j(S) = \int_{\underline{q}}^{\bar{q}} \left(\lambda_s^j(q)q + \lambda_{s'}^j(q)(1-q)\right) dF$$

for policy $j \in \{a, b\}$ in state $S \in \{A, B\}$. We assume without loss of generality that $\lambda_s^j(q)$ is such that the correct policy is chosen with a probability (weakly) greater than 0.5 in both states (see Footnote 59). We first look for $\underline{q}^a \in [\underline{q}, \bar{q}]$ and $\bar{q}^a \in [\underline{q}, \bar{q}]$ such that $T_a(B)$ remains constant as follows: $\underline{q}^a$ satisfies $\int_{\underline{q}}^{\underline{q}^a} qdF = \int_{\underline{q}}^{\bar{q}} \lambda_\beta^a(q)qdF$, and $\bar{q}^a$ satisfies $\int_{\bar{q}^a}^{\bar{q}}(1-q)dF = \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^a(q)(1-q)dF$. We will show that the former equality implies that

$$\int_{\underline{q}}^{\underline{q}^a}(1-q)dF \geq \int_{\underline{q}}^{\bar{q}} \lambda_\beta^a(q)(1-q)dF,$$

with strict inequality unless $\lambda_\beta^a(q) = \mathbf{1}_{q \leq x}$ for some $x \in [\underline{q}, \bar{q}]$. Similarly, we will show that $\int_{\bar{q}^a}^{\bar{q}}(1-q)dF = \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^a(q)(1-q)dF$ implies that

$$\int_{\bar{q}^a}^{\bar{q}} qdF \geq \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^a(q)qdF,$$

with strict inequality unless $\lambda_\alpha^a(q) = \mathbf{1}_{q \geq x}$ for some $x \in [\underline{q}, \bar{q}]$. We first prove the former claim. The proof is trivial if $\lambda_\beta^a(q) = \mathbf{1}_{q \leq x}$ for some $x \in [\underline{q}, \bar{q}]$, so assume that $\lambda_\beta^a(q) \neq \mathbf{1}_{q \leq x}$ for any $x \in [\underline{q}, \bar{q}]$. Thus, $\underline{q}^a \in (\underline{q}, \bar{q})$. Suppose towards a contradiction that $\int_{\underline{q}}^{\underline{q}^a}(1-q)dF \leq \int_{\underline{q}}^{\bar{q}} \lambda_\beta^a(q)(1-q)dF$. This and $\int_{\underline{q}}^{\underline{q}^a} qdF = \int_{\underline{q}}^{\bar{q}} \lambda_\beta^a(q)qdF$ imply that

$$\int_{\underline{q}}^{\underline{q}^a} dF \leq \int_{\underline{q}}^{\bar{q}} \lambda_\beta^a(q)dF = \int_{\underline{q}}^{\underline{q}^a} \lambda_\beta^a(q)dF + \int_{\underline{q}^a}^{\bar{q}} \lambda_\beta^a(q)dF,$$

and thus $\int_{\underline{q}}^{\underline{q}^a}(1-\lambda_\beta^a(q))dF \leq \int_{\underline{q}^a}^{\bar{q}} \lambda_\beta^a(q)dF$. Multiplying both sides of this inequality by $\underline{q}^a \in (\underline{q}, \bar{q})$, the right hand side is strictly lower than $\int_{\underline{q}^a}^{\bar{q}} \lambda_\beta^a(q)qdF$ (since $\underline{q}^a < \bar{q}$)), and the left hand side is strictly greater than $\int_{\underline{q}}^{\underline{q}^a}(1-\lambda_\beta^a(q))qdF$ (since $\underline{q}^a > \underline{q}$)), contradicting $\int_{\underline{q}}^{\underline{q}^a} qdF = \int_{\underline{q}}^{\bar{q}} \lambda_\beta^a(q)qdF$. Next, let $\bar{q}^a$ be such that $\int_{\bar{q}^a}^{\bar{q}}(1-q)dF = \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^a(q)(1-q)dF$, and assume towards a contradiction that $\int_{\bar{q}^a}^{\bar{q}} qdF \leq \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^a(q)qdF$, where $\lambda_\alpha^a(q) \neq \mathbf{1}_{q \geq x}$ for any $x \in [\underline{q}, \bar{q}]$. These imply that

$$\int_{\bar{q}^a}^{\bar{q}} dF \leq \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^a(q)dF = \int_{\underline{q}}^{\bar{q}^a} \lambda_\alpha^a(q)dF + \int_{\bar{q}^a}^{\bar{q}} \lambda_\alpha^a(q)dF,$$

and thus $\int_{\bar{q}^a}^{\bar{q}}(1-\lambda_\alpha^a(q))dF \leq \int_{\underline{q}}^{\bar{q}^a} \lambda_\alpha^a(q)dF$. Multiplying both sides of the inequality by $1-\bar{q}^a$, and rearranging we obtain a contradiction to $\int_{\bar{q}^a}^{\bar{q}}(1-q)dF = \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^a(q)(1-q)dF$. As a result, $T_a(A)$ strictly increases in our construction, unless $\lambda_\beta^a(q) = \mathbf{1}_{q \leq \underline{q}^a}$ and $\lambda_\alpha^a(q) = \mathbf{1}_{q \geq \bar{q}^a}$ both hold ($T_a(A)$ remains unchanged in that case).

Next, we look for $\underline{q}^b \in [\underline{q}, \bar{q}^a)$ such that $\int_{\underline{q}}^{\underline{q}^b} q dF = \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^b(q) q dF$.[59] If this equality cannot be satisfied for any $\underline{q}^b \in [\underline{q}, \bar{q}^a)$, then we set $\underline{q}^b = \bar{q}^a$ (in that case, there is no abstention after an α signal). Similarly, we look for $\bar{q}^b \in (\underline{q}^a, \bar{q}]$ such that $\int_{\bar{q}^b}^{\bar{q}} (1-q) dF = \int_{\underline{q}}^{\bar{q}} \lambda_\beta^b(q)(1-q) dF$, and if there exists no $\bar{q}^b \in (\underline{q}^a, \bar{q}]$ that satisfies the equality, then we set $\underline{q}^a = \bar{q}^b$ (in that case, there is no abstention after a β signal). Thus, $T_b(B)$ weakly increases in our construction (strictly if $\underline{q}^b < \bar{q}^a$ and $\lambda_\alpha^b(q) \neq \mathbf{1}_{q \leq x}$ for any $x \in [\underline{q}, \bar{q}]$ or if $\bar{q}^b > \underline{q}^a$ and $\lambda_\beta^b(q) \neq \mathbf{1}_{q \geq x}$ for any $x \in [\underline{q}, \bar{q}]$). While $T_b(A)$ may decrease due to having to set $\underline{q}^b = \bar{q}^a$ or $\underline{q}^a = \bar{q}^b$, $T_a(A) + T_b(A)$ cannot decrease in our construction since $\underline{q}^b = \bar{q}^a$ implies that there is no abstention after an α signal and $\underline{q}^a = \bar{q}^b$ implies that there is no abstention after a β signal. Thus, given these four cutoffs we construct from the initial strategy, the expected turnout rate $T_a(S) + T_b(S)$ weakly increases in both states, and it can be checked that the "relative turnout rate" for the correct policy weakly increases in both states; i.e., both $\frac{T_a(A)}{T_a(A)+T_b(A)}$ and $\frac{T_b(B)}{T_b(B)+T_a(B)}$ increase. In fact, at least one of these relative turnout rates must strictly increase by construction since $\lambda_s^j(q)$ does not have a cutoff form for at least one (j, s) pair.

To see why, first assume that $T_a(A)$ remains constant in our construction because the initial strategy is such that $\lambda_\beta^a(q) = \mathbf{1}_{q \leq \underline{q}^a}$ for $\underline{q}^a \in [\underline{q}, \bar{q}]$ and $\lambda_\alpha^a(q) = \mathbf{1}_{q \geq \bar{q}^a}$ for $\bar{q}^a \in [\underline{q}, \bar{q}]$ (recall that $T_a(B)$ is constant by construction). Then, it can be checked that $T_b(A)$ must be constant as well given our construction: there must exist $\underline{q}^b \in [\underline{q}, \bar{q}^a]$ such that $\int_{\underline{q}}^{\underline{q}^b} q dF = \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^b(q) q dF$, and $\bar{q}^b \in [\underline{q}^a, \bar{q}]$ such that $\int_{\bar{q}^b}^{\bar{q}} (1-q) dF = \int_{\underline{q}}^{\bar{q}} \lambda_\beta^b(q)(1-q) dF$ (otherwise, $T_a(A) + T_b(A) > 1$, which is a contradiction). If $\underline{q}^b < \bar{q}^a$ or if $\bar{q}^b > \underline{q}^a$, then $T_b(B)$ and $\frac{T_b(B)}{T_b(B)+T_a(B)}$ strictly increase as desired. Indeed one of the two ($\underline{q}^b < \bar{q}^a$ or $\bar{q}^b > \underline{q}^a$) must hold as otherwise $\lambda_s^j(q)$ *is* a cutoff strategy in contrast to our initial hypothesis. As a result, if $T_a(A)$ remains constant in our construction, then the expected turnout rate in state A, $T_a(A) + T_b(A)$, is constant, but $\frac{T_b(B)}{T_b(B)+T_a(B)}$ and $T_a(B) + T_b(B)$ are strictly higher. Next, assume that $T_a(A)$ strictly increases because $\lambda_\beta^a(q) \neq \mathbf{1}_{q \leq x}$ for any $x \in [\underline{q}, \bar{q}]$ and/or $\lambda_\alpha^a(q) \neq \mathbf{1}_{q \geq x}$ for any $x \in [\underline{q}, \bar{q}]$. As discussed above, by construction, $T_a(A) + T_b(A)$ cannot decrease, and $T_b(A)$ weakly decreases. As a result, and given that $T_a(A)$ is strictly higher,

[59] The proof slightly differs here if $\lambda_s^j(q)$ is such that the correct policy is chosen with a probability greater than 0.5 in only one state. Assume without loss of generality that $\pi \geq 0.5$ and that according to $\lambda_s^j(q)$, policy b is chosen with a probability strictly lower than 0.5 in state B. In that case, we look for $\underline{q}^b \in [\underline{q}, \bar{q}^a)$ such that $\int_{\underline{q}}^{\underline{q}^b} (1-q) dF = \int_{\underline{q}}^{\bar{q}} \lambda_\alpha^b(q)(1-q) dF$. It can be checked that such $\underline{q}^b$ must satisfy $\underline{q}^b \leq \bar{q}^a$. Similarly, we look for $\bar{q}^b$ such that $\int_{\bar{q}^b}^{\bar{q}} q dF = \int_{\underline{q}}^{\bar{q}} \lambda_\beta^b(q) q dF$. It can be checked that $\bar{q}^b$ must satisfy $\bar{q}^b \geq \underline{q}^a$. Thus, $T_b(B)$ and $T_a(B)$ remain identical, whereas $T_a(A)$ increases and $T_b(A)$ decreases in our construction. Then, the analogues of Lemma 4 and Lemma 5 with similar proofs suffice to obtain the desired result.

$\frac{T_a(A)}{T_a(A)+T_b(A)}$ strictly increases. Moreover, as discussed above, $T_b(B)$ and $T_a(B)+T_b(B)$ weakly increase by construction. To complete the proof, we need Lemma 4 and Lemma 5.

Lemma 4 *Let $W_t(q)$ be such that*

$$W_t(q) = \begin{cases} \sum\limits_{i=\frac{t+1}{2}}^{t} \binom{t}{i} q^i (1-q)^{t-i} & \text{if } t \text{ is odd} \\ \sum\limits_{i=\frac{t}{2}+1}^{t} \binom{t}{i} q^i (1-q)^{t-i} + \frac{1}{2}\binom{t}{\frac{t}{2}} q^{\frac{t}{2}} (1-q)^{\frac{t}{2}} & \text{if } t \text{ is even} \end{cases}$$

where $q \in (1/2, 1)$. Then, $W_t(q)$ is monotonic in t. In particular, $W_t(q) = W_{t+1}(q) < W_{t+2}(q)$ for every odd $t > 0$. Moreover, $W_1(q) = W_2(q) = q$ and thus $W_t(q) > q$ for every $t > 2$.

Proof. First, we show that $W_t(q) = W_{t+1}(q)$ if t is odd. Let $\Pr(X \leq k; t, q)$ denote the cumulative Binomial distribution function with t trials, X successes and success probability q; i.e.,

$$\Pr(X \leq k; t, q) = (t-k)\binom{t}{k} \int_0^{1-q} x^{t-k-1}(1-x)^k dx.$$

Using this formula and that $\Pr(X = \frac{t+1}{2}; t+1, q) = \Pr(X \leq \frac{t+1}{2}; t+1, q) - \Pr(X \leq \frac{t-1}{2}; t+1, q)$, it can be checked that $\Pr(X \leq \frac{t+1}{2}; t+1, q) - \frac{1}{2}\Pr(X = \frac{t+1}{2}; t+1, q) = \Pr(X \leq \frac{t-1}{2}; t, q)$ must hold. Thus, $W_t(q) = W_{t+1}(q)$ for t odd.

Next, we show that $W_t(q) < W_{t+2}(q)$ if t is odd. First, we prove that $W_t(q) < W_{t+2}(q)$ for $q \geq q^* > \frac{1}{2}$, where q^* is given by $q^*(1-q^*) = \frac{1}{4}\frac{t+1}{t+2}$. Then, for $t > 0$ odd, $W_{t+2}(q) - W_t(q)$ equals

$$\frac{t+1}{2}\binom{t}{\frac{t-1}{2}} \int_0^{1-q} x^{\frac{t-1}{2}}(1-x)^{\frac{t-1}{2}} dt - \frac{t+3}{2}\binom{t+2}{\frac{t+1}{2}} \int_0^{1-q} x^{\frac{t+1}{2}}(1-x)^{\frac{t+1}{2}} dx,$$

which equals

$$\int_0^{1-q} x^{\frac{t-1}{2}}(1-x)^{\frac{t-1}{2}} \left(\frac{t+1}{2}\binom{t}{\frac{t-1}{2}} - \frac{t+3}{2}\binom{t+2}{\frac{t+1}{2}} x(1-x) \right) dx.$$

It can be checked that this term is strictly greater than 0 for all $q \in [q^*, 1)$, where $q^* > \frac{1}{2}$ and satisfies $q^*(1-q^*) = \frac{1}{4}\frac{t+1}{t+2}$. We now show that $W_t(\frac{1}{2}) = W_{t+2}(\frac{1}{2})$ and that $\frac{d}{dq}(W_{t+2}(q) - W_t(q)) > 0$ for every $q \in [\frac{1}{2}, q^*)$, which will imply that $W_{t+2}(q) > W_t(q)$ for every $q \in (\frac{1}{2}, q^*)$ and complete the proof. It can be checked that $\frac{d}{dq}(W_{t+2}(q) - W_t(q))$ is equal to $q^{\frac{t-1}{2}}(1-q)^{\frac{t-1}{2}} \left(\frac{t+3}{2}\binom{t+2}{\frac{t+1}{2}} q(1-q) - \frac{t+1}{2}\binom{t}{\frac{t-1}{2}} \right)$, which is strictly positive for all $q \in [\frac{1}{2}, q^*)$. Moreover,

$W_t(\frac{1}{2}) = W_{t+2}(\frac{1}{2})$ because $\sum_{i=\frac{t+1}{2}}^{t} \binom{t}{i} \left(\frac{1}{2}\right)^t = \sum_{i=\frac{t+3}{2}}^{t+2} \binom{t+2}{i} \left(\frac{1}{2}\right)^{t+2}$ using the fact that $\sum_{i=0}^{t} \binom{t}{i} = 2 \sum_{i=\frac{t+1}{2}}^{t} \binom{t}{i} = 2^t$.

Lemma 5 *If p and q increase to $p' \geq p$ and $q' \geq q$ respectively (with at least one strict inequality), then*

$$\sum_{t=0}^{N} \binom{N}{t} p^t (1-p)^{N-t} W_t(q) < \sum_{t=0}^{N} \binom{N}{t} (p')^t (1-p')^{N-t} W_t(q').$$

Proof. First, we prove that $W_t(q)$ is strictly increasing in q for $t \geq 1$. To see why, first assume that t is odd, and note that $W_t(q') - W_t(q)$ equals

$$(t-k)\binom{t}{k} \int_{1-q'}^{1-q} x^{t-k-1}(1-x)^k dx,$$

which is strictly positive if $q' > q$. Next, assume that $t \geq 2$ is even. By Lemma 4, $W_{t-1}(q) = W_t(q)$, and it follows that $W_t(q') - W_t(q) = W_{t-1}(q') - W_{t-1}(q) > 0$ for $q' > q$. To complete the proof, it is enough to show that

$$\sum_{t=0}^{N} \binom{N}{t} p^t (1-p)^{N-t} W_t(q') < \sum_{t=0}^{N} \binom{N}{t} (p')^t (1-p')^{N-t} W_t(q')$$

for $p < p'$. But this is true because an increase in p results in a new t distribution that first order stochastically dominates the original, and by Lemma 4, $W_t(q')$ is monotonically increasing in t such that $W_t(q') \leq W_{t+1}(q')$.

Since our strategy construction in the proof of Lemma 3 weakly increases p and q (in the notation of Lemmas 4 and 5) in both states of the world and strictly in at least one state, our construction is strictly better than the non-cutoff strategy $\lambda_s^j(q)$. Thus, Lemma 3 is proved.

We can now prove Proposition 1. By what McLennan (1998) has shown, the optimal symmetric strategy, if it exists, must be an equilibrium strategy. In Lemma 3, we have shown the existence of an optimal strategy and that it must consist of cutoffs. Next, consider the outcome of the optimal equilibrium in a biased electorate. If the outcome is not consistent with a cutoff strategy (i.e., a positive measure of individuals with identical q and s behave in different ways due to perception biases), the outcome is inconsistent with the optimal strategy and thus worse than the optimal unbiased equilibrium outcome. Second, assume that the outcome with biased voting is consistent with a cutoff strategy outcome due to the particular form of overconfidence and underconfidence biases. If the outcome exhibits no interior cutoff, then this is suboptimal under the conditions stated in Lemma 2. If the outcome is consistent with an interior cutoff strategy and happens to coincide with another optimal unbiased equilibrium outcome, this is a nongeneric case because a small perturbation in, e.g., π or $\lambda_o(q)$ would rule it out. Thus, voter behavior with biased perceptions is generically inconsistent with the optimal unbiased equilibrium and thus suboptimal under the conditions stated in Lemma 2.

As mentioned in the main text in Footnote 6, the results are robust to allowing for different levels of overconfidence and underconfidence. Assume that for fixed q, there are finitely many overconfidence functions $p_o^j(q) \in (q, \bar{q}]$ and respective probabilities $\lambda_o^j(q)$ where $q < p_o^j(q) < p_o^{j+1}(q)$ for every q and $j \in \{1, ..., J-1\}$. Similarly, assume that there are finitely many underconfidence functions $p_u^k(q) \in [\underline{q}, q)$ and respective probabilities $\lambda_u^k(q)$ where $p_u^{k+1}(q) < p_u^k(q) < q$ for every q and $k \in \{1, ..., K-1\}$. None of the proofs above are affected by this extension.

Corollary 1 *Under conditions stated in Lemma 2, a responsive equilibrium always exists and has at least one interior cutoff.*

Proof. Given Lemma 3, an optimal symmetric strategy always exists, and it constitutes the optimal symmetric equilibrium. Under the conditions stated in Lemma 2, the optimal symmetric equilibrium involves an interior cutoff by construction. Hence, there is indeed a responsive equilibrium under the conditions stated in Lemma 2.

A.2 Results in Section IA under Awareness of Others' Perception Biases

Awareness of others' perception biases implies the following: every i knows that for every j and q_j, $p_j(q_j)$ takes one of three possible values: $p_o(q_j)$, $p_u(q_j)$, and q_j with respective probabilities $\lambda_o(q)$, $\lambda_u(q)$, and $1 - \lambda_o(q) - \lambda_u(q)$. Under this assumption, the equilibrium characterization stated in Lemma 1 naturally holds. Also, the proof for showing that $q^a = \Pr(S = B|piv_a)$ and $q^b = \Pr(S = A|piv_b)$ if the correct policy is chosen with a probability (weakly) greater than 0.5 in either state is unaffected. However, it may no longer be true that $q^a = q^b$ if $\pi = \underline{q} = 0.5$, or that $q^a = \Pr(S = B|piv_a)$ if $\pi = 0.5$. Awareness of others' perception biases does not affect Lemma 2 since Lemma 2 concerns unbiased equilibria.

Moreover, the steps used in the proof of Proposition 1 still apply, and the main results are robust to awareness regarding others' overconfidence and underconfidence biases.

A.3 Proof of Proposition 2

Before proving Proposition 2, we will state it formally.

Proposition 2 *Assume that π is bounded away from 0 and 1, and that $1 - \underline{q} < \bar{q} \leq 1$. If media veracity is low enough so that $\int_{\underline{q}}^{\bar{q}} q dF < 0.5$ holds, and the electorate is sufficiently overconfident, then the probability that the wrong policy is chosen goes to one in at least one state as N goes to infinity, whereas the correct policy is chosen with a probability that goes to one in both states in the optimal unbiased equilibrium.*

Proof. We first construct a strategy that fully aggregates information in unbiased electorates. The strategy is such that i votes for the policy that matches s_i if $q_i \geq q^*$ and abstains otherwise, where q^* satisfies $\int_{q^*}^{\bar{q}} q dF > \int_{q^*}^{\bar{q}} (1-q) dF$ (such q^* surely exists since $\bar{q} > 0.5$). This implies that the relative turnout share for the correct policy is strictly greater than 0.5 in both states (also in the limit), ensuring that the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$. By what we have shown above, the optimal symmetric strategy exists and is the optimal symmetric equilibrium in an unbiased electorate. Thus, optimal unbiased equilibria will generate the same outcome; i.e., the sequence of optimal unbiased equilibria will select the correct policy in both states with a probability that goes to one as N goes to infinity.

We will now prove the inefficiency of overconfidence assuming *awareness* regarding others' overconfidence since the proof naturally extends to the case with unawareness. Let q_N^a and q_N^b denote the equilibrium cutoff pair given electorate size N. We consider only those equilibria in which the correct policy is chosen with a probability that goes to one in both states. We assume without loss of generality that $p_o(q) < \bar{q}$ for $q < \bar{q}$, and thus, $p_o^{-1}(q) \to \bar{q}$ as $q \to \bar{q}$.[60]

First, we consider the case in which $\bar{q} < 1$. Assume towards a contradiction that $\limsup_{N\to\infty} q_N^a \geq \bar{q}$. $\limsup_{N\to\infty} q_N^a \geq \bar{q}$ and the assumption that $1 - \underline{q} < \bar{q}$ imply that $\limsup_{N\to\infty} q_N^b \geq \bar{q}$ because otherwise the same policy (policy b) will be chosen with a probability that goes to one in both states. More generally, $\limsup_{N\to\infty} q_N^a \geq \bar{q}$ if and only if $\limsup_{N\to\infty} q_N^b \geq \bar{q}$ focusing on those equilibria in which the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$. Also, a strict inequality cannot

[60]If we allow for $p_o(q) = \bar{q}$ for $q < \bar{q}$, then it is straightforward to construct overconfidence functions that prevent the aggregation of information in at least one state.

hold as that would mean that the probability of winning goes to either 0.5 or 0 for the correct policy in at least one state as $N \to \infty$. Hence, (with an abuse of notation) there exists a sequence of equilibrium cutoff pairs (q_N^a, q_N^b) such that $q_N^a \to \bar{q}$ and $q_N^b \to \bar{q}$ with $\lim_{N\to\infty} \frac{(1-\pi)\Pr(piv_a|B,N)}{\pi \Pr(piv_a|A,N)} = \frac{\bar{q}}{1-\bar{q}}$ and $\lim_{N\to\infty} \frac{(1-\pi)\Pr(piv_b|B,N)}{\pi \Pr(piv_b|A,N)} = \frac{1-\bar{q}}{\bar{q}}$ due to Lemma 1. Note that $\Pr(piv_a|B,N)$ and $\Pr(piv_b|B,N)$ have a common term (representing the tie events) and the other term differs by only a multiplicative term. That is, $\Pr(piv_a|B,N)$ and $\Pr(piv_b|B,N)$ can be written as $\Pr(piv_a|B,N) = x_N + y_N$, and $\Pr(piv_b|B,N) = x_N + y_N\gamma_N$, where

$$\gamma_N = \frac{\int_{q_N^a}^{\bar{q}}(1-q)dF + \int_{p_o^{-1}(q_N^a)}^{q_N^a}\lambda(q)(1-q)dF}{\int_{q_N^b}^{\bar{q}} qdF + \int_{p_o^{-1}(q_N^b)}^{q_N^b}\lambda(q)qdF},$$

and x_N and y_N represent the respective probabilities of a tie event and the case in which policy a is behind by one vote in state B. Similarly, $\Pr(piv_a|A,N)$ and $\Pr(piv_b|A,N)$ have a common term, and the other term differs by a multiplicative term. Thus, $\Pr(piv_a|A,N)$ and $\Pr(piv_b|A,N)$ can be written as $\Pr(piv_a|A,N) = w_N + z_N$, and $\Pr(piv_b|A,N) = w_N + z_N\xi_N$, where

$$\xi_N = \frac{\int_{q_N^a}^{\bar{q}} qdF + \int_{p_o^{-1}(q_N^a)}^{q_N^a}\lambda(q)qdF}{\int_{q_N^b}^{\bar{q}}(1-q)dF + \int_{p_o^{-1}(q_N^b)}^{q_N^b}\lambda(q)(1-q)dF},$$

and w_N and ξ_N represent the respective probabilities of a tie event and the case in which policy a is behind by one vote in state A.[61] Thus, we can write $\lim_{N\to\infty} \frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)} = \frac{\bar{q}}{1-\bar{q}}$ and $\lim_{N\to\infty} \frac{(1-\pi)(x_N+\gamma_N y_N)}{\pi(w_N+\xi_N z_N)} = \frac{1-\bar{q}}{\bar{q}}$. Note that $\gamma_N < 1 < \xi_N$ must hold as we focus on equilibria in which the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$. Moreover, the term ξ_N is bounded above by a finite number because

$$\gamma_N = \frac{\int_{q_N^a}^{\bar{q}}(1-q)dF + \int_{p_o^{-1}(q_N^a)}^{q_N^a}\lambda(q)(1-q)dF}{\int_{q_N^b}^{\bar{q}} qdF + \int_{p_o^{-1}(q_N^b)}^{q_N^b}\lambda(q)qdF} = \frac{1-\tilde{q}_N^a}{\tilde{q}_N^b}\frac{1-F(q_N^a) + \int_{p_o^{-1}(q_N^a)}^{q_N^a}\lambda(q)dF}{1-F(q_N^b) + \int_{p_o^{-1}(q_N^b)}^{q_N^b}\lambda(q)dF},$$

and

$$\xi_N = \frac{\int_{q_N^a}^{\bar{q}} qdF + \int_{p_o^{-1}(q_N^a)}^{q_N^a}\lambda(q)qdF}{\int_{q_N^b}^{\bar{q}}(1-q)dF + \int_{p_o^{-1}(q_N^b)}^{q_N^b}\lambda(q)(1-q)dF} = \frac{\tilde{q}_N^a}{1-\tilde{q}_N^b}\frac{1-F(q_N^a) + \int_{p_o^{-1}(q_N^a)}^{q_N^a}\lambda(q)dF}{1-F(q_N^b) + \int_{p_o^{-1}(q_N^b)}^{q_N^b}\lambda(q)dF},$$

where $\tilde{q}_N^a \in (p_o^{-1}(q_N^a), \bar{q})$ and $\tilde{q}_N^b \in (p_o^{-1}(q_N^b), \bar{q})$ are conditional expectations. It follows that

[61] By hypothesis, q_N^a and q_N^b are very close to $\bar{q}$ for large N. Therefore, $1-q_N^a$ and $1-q_N^b$ cutoffs are irrelevant by the assumption that $\bar{q} > 1-\underline{q}$, and voting against signal will not take place.

$\xi_N < \frac{\tilde{q}^a_N}{1-\tilde{q}^b_N}\frac{\tilde{q}^b_N}{1-\tilde{q}^a_N} < \left(\frac{\bar{q}}{1-\bar{q}}\right)^2$ given that $\gamma_N < 1$. Therefore, $\limsup_{N\to\infty} \xi_N < \infty$. Moreover, $\liminf_{N\to\infty} \gamma_N > 0$ because from $\xi_N > 1$ it follows that $\gamma_N > \frac{1-\tilde{q}^a_N}{\tilde{q}^b_N}\frac{1-\tilde{q}^b_N}{\tilde{q}^a_N} > \left(\frac{1-\bar{q}}{\bar{q}}\right)^2$. We now prove that the limits $\lim_{N\to\infty} \frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)} = \frac{\bar{q}}{1-\bar{q}}$ and $\lim_{N\to\infty} \frac{(1-\pi)(x_N+\gamma_N y_N)}{\pi(w_N+\xi_N z_N)} = \frac{1-\bar{q}}{\bar{q}}$ as $q^a_N \to \bar{q}$ and $q^b_N \to \bar{q}$ result in a contradiction. We can write (taking convergent subsequences of x_N/y_N and w_N/z_N if necessary)

$$\lim_{N\to\infty} \frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)} = \lim_{N\to\infty} \frac{(1-\pi)(x_N/y_N+1)}{\pi(w_N/z_N+1)} \lim_{N\to\infty} \frac{y_N}{z_N} = \frac{(1-\pi)(c+1)}{\pi(d+1)} \lim_{N\to\infty} \frac{y_N}{z_N},$$

where, abusing notation, $c = \lim_{N\to\infty} \frac{x_N}{y_N} \in (0,\infty)$ and $d = \lim_{N\to\infty} \frac{w_N}{z_N} \in (0,\infty)$ by Claim 2 below in the proof of Proposition 4. Thus, $\lim_{N\to\infty} y_N/z_N$ exists, and $\lim_{N\to\infty} y_N/z_N \in (0,\infty)$ since $y_N/z_N = \left(\frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)}\right) / \left(\frac{(1-\pi)(x_N/y_N+1)}{\pi(w_N/z_N+1)}\right)$, $\lim_{N\to\infty} \frac{(1-\pi)(x_N/y_N+1)}{\pi(w_N/z_N+1)} = \frac{(1-\pi)(c+1)}{\pi(d+1)} \in (0,\infty)$ and $\lim_{N\to\infty} \frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)} = \frac{\bar{q}}{1-\bar{q}} \in (0,\infty)$. As a result, we have that

$$\frac{\bar{q}}{1-\bar{q}} = \frac{(1-\pi)(c+1)}{\pi(d+1)} \lim_{N\to\infty} \frac{y_N}{z_N},$$

and from $\lim_{N\to\infty} \frac{(1-\pi)(x_N+\gamma_N y_N)}{\pi(w_N+\xi_N z_N)} = \frac{1-\bar{q}}{\bar{q}}$, we have that

$$\frac{1-\bar{q}}{\bar{q}} = \frac{(1-\pi)(c+\lim_{N\to\infty}\gamma_N)}{\pi(d+\lim_{N\to\infty}\xi_N)} \lim_{N\to\infty} \frac{y_N}{z_N}.$$

From these inequalities, it follows that

$$\left(\frac{\bar{q}}{1-\bar{q}}\right)^2 = \frac{(d+\lim_{N\to\infty}\xi_N)/(d+1)}{(c+\lim_{N\to\infty}\gamma_N)/(c+1)},$$

where $\lim_{N\to\infty}\xi_N \in [1,\infty)$ and $\lim_{N\to\infty}\gamma_N \in (0,1]$ (because $\gamma_N < 1 < \xi_N$, $\limsup_{N\to\infty}\xi_N < \infty$, and $\liminf_{N\to\infty}\gamma_N > 0$ as shown above). However, $\frac{(d+\lim_{N\to\infty}\xi_N)/(d+1)}{(c+\lim_{N\to\infty}\gamma_N)/(c+1)} < \frac{\lim_{N\to\infty}\xi_N}{\lim_{N\to\infty}\gamma_N}$ because $\lim_{N\to\infty} \frac{\xi_N}{\gamma_N} = \left(\frac{\bar{q}}{1-\bar{q}}\right)^2 = \frac{\lim_{N\to\infty}\xi_N}{\lim_{N\to\infty}\gamma_N}$, $c = (0,\infty)$ and $d = (0,\infty)$ by Claim 2. Thus, we have a contradiction. As a result, either $\limsup_{N\to\infty} q^a_N < \bar{q}$ or $\limsup_{N\to\infty} q^b_N < \bar{q}$ must hold. More generally, repeating the steps above, it can be shown that either q^a_N or q^b_N must be uniformly bounded above by $\tilde{q} < \bar{q}$ for fixed π and $F[\underline{q},\bar{q}]$ regardless of the form of overconfidence in the population if the sequence of equilibrium cutoff pairs (q^a_N, q^b_N) are such that the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$. This is because the method of proof in Claim 2 is general and extends to the case where x_N, y_N, w_N, z_N, ξ_N and γ_N are functions of $\lambda_{o,N}(q)$ and $p_{o,N}(q)$ with $p_{o,N}(q) < \bar{q}$ for

$q < \bar{q}$. To analyze the case in which $\bar{q} = 1$, we need the following assumption for tractability.

Assumption 1 *If $\bar{q} = 1$, a fraction $\frac{\epsilon}{2} > 0$ of the electorate votes for a, and a fraction $\frac{\epsilon}{2} > 0$ votes for b (i.e., they are partisans), where ϵ is possibly very small.*

Note that this assumption would not affect the result with $\bar{q} < 1$. Assumption 1 also eliminates unresponsive equilibria. Fix small $\epsilon > 0$, and let (q_N^a, q_N^b) be a sequence of equilibrium cutoff pairs such that the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$. If $q_N^a \to 1$ and $q_N^b \to 1$, we have that $\lim_{N\to\infty} \frac{x_N + y_N}{x_N + \gamma_N y_N} \frac{w_N + \xi_N z_N}{w_N + z_N} = \infty$. Note that

$$\lim_{N\to\infty} \frac{x_N + y_N}{x_N + \gamma_N y_N} \frac{w_N + \xi_N z_N}{w_N + z_N} \leq \lim \sup_{N\to\infty} \frac{\xi_N}{\gamma_N},$$

since $\gamma_N < 1 < \xi_N$ and Assumption 1 holds. However, $\limsup_{N\to\infty} \frac{\xi_N}{\gamma_N} < \infty$ again due to Assumption 1, resulting in a contradiction. Thus, either $\limsup_{N\to\infty} q_N^a < 1$ or $\limsup_{N\to\infty} q_N^b < 1$ must hold. In particular, either q_N^a or q_N^b must be uniformly bounded above by $\tilde{q} < 1$ for fixed π and $F[\underline{q}, \bar{q}]$ regardless of the form of overconfidence in the population if the sequence of equilibrium cutoff pairs (q_N^a, q_N^b) are such that the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$.

Finally, we show that a sufficiently high level of overconfidence will prevent information aggregation in at least one state and result in the wrong policy being chosen with a probability that goes to one. To show this, consider overconfidence functions $\lambda_o(q)$ and $p_o(q)$ such that

$$\textstyle\int_{\tilde{q}}^{\bar{q}} (1-q)dF + \int_{p_o^{-1}(\tilde{q})}^{\tilde{q}} \lambda_o(q)(1-q)dF + \int_{\underline{q}}^{p_o^{-1}(1-\tilde{q})} qdF + \int_{p_o^{-1}(1-\tilde{q})}^{1-\tilde{q}} (1-\lambda_o(q))qdF$$

is strictly higher than

$$\textstyle\int_{\tilde{q}}^{\bar{q}} qdF + \int_{p_o^{-1}(\tilde{q})}^{\tilde{q}} \lambda_o(q)qdF + \int_{\underline{q}}^{p_o^{-1}(1-\tilde{q})} (1-q)dF + \int_{p_o^{-1}(1-\tilde{q})}^{1-\tilde{q}} (1-\lambda_o(q))(1-q)dF.$$

To see why such overconfidence functions exist, first note that a high level of Dunning-Kruger effect reduces $p_o^{-1}(1-\tilde{q})$ and makes the cutoff $1-\tilde{q}$ trivial. Moreover, high levels of the Dunning-Kruger effect makes $p_o^{-1}(\tilde{q})$ closer and closer to $\underline{q}$, and $\lambda_o(q)$ higher and higher for $q \in (\underline{q}, \tilde{q})$. This coupled with $\mathbb{E}(q) < 0.5$ implies that there exist overconfidence functions $\lambda_o(q)$ and $p_o(q)$ that make the former term strictly higher than the latter. Assume without loss of generality that $\limsup_{N\to\infty} q_N^b \leq \limsup_{N\to\infty} q_N^a$ with such an overconfidence level. If $\limsup_{N\to\infty} q_N^b > \tilde{q}$, then by what we have shown above, this is an inefficient equilibrium sequence that prevents information aggregation in at least one state. Therefore, assume that $\limsup_{N\to\infty} q_N^b \leq \tilde{q}$. Again by what we have shown above, the sufficiently high levels of overconfidence we characterized above result in a vote share for the wrong policy that is strictly larger than the vote share for the correct policy in at least one state as N goes to infinity.

A.4 Proofs of Results in Section IC

We will invoke the following Lemma later.

Lemma 6 *In a symmetric equilibrium of every model discussed in the main text with independent signals, $q^a = \Pr(S = B|piv_a)$ and $q^b = \Pr(S = A|piv_b)$ must hold if the correct policy is chosen with a probability (weakly) greater than 0.5 in both states.*

Proof. None of the steps in the proof of Lemma 1 depends on the distribution of signal precisions being identical or the absence of partisans. In particular, the difference in the distributions of signal precisions across states or the presence of partisan voters are accounted for in the probabilities of piv_a and piv_b events. Thus, the proof presented in Lemma 1 applies to the two extension models with partisan voters and asymmetric media veracity.

A.4.1 Model with Partisan voters

Proposition 3 *Assume that π is bounded away from 0 and 1, and $p_a \leq p_b < 0.5$. Then, an unbiased electorate makes the correct decision with a probability that goes to one in both states as N goes to infinity. Next, let media veracity be such that*

$$p_b > p_a + (1 - p_a - p_b)\left(\int_{\underline{q}}^{\bar{q}} q dF - \int_{\underline{q}}^{\bar{q}} (1-q) dF\right).$$

If $\bar{q} = 1$, then a highly overconfident electorate will choose policy b with a probability that goes to one in both states. If $\bar{q} < 1$, the same result obtains if in addition $\bar{q} \geq 1 - \underline{q}$ and the media veracity is such that $p_a + (1 - p_a - p_b)\int_{0.5}^{\bar{q}} (1-q) dF > p_b$.

Proof. First, we show that the sequence of optimal unbiased equilibria is such that the correct policy is chosen with a probability that goes to one as $N \to \infty$. To show this, first note that the proof in Lemma 3 applies to show the existence of an optimal symmetric strategy and its cutoff form. Therefore, it is enough to show that there exists a symmetric strategy that results in the correct policy being chosen with a probability that goes to 1 as $N \to \infty$ in both states. To show the existence of such a strategy, we assume without loss of generality that the first inequality condition in the Proposition statement holds. Then, there exists a $q^* \in (\underline{q}, \bar{q})$ such that every nonpartisan with $s_i = \alpha$ or with $s_i = \beta$ and $q_i < q^*$ votes for a, and every nonpartisan with $s_i = \beta$ and $q_i > q^*$ votes for b satisfying the following: $p_a + (1 - p_a - p_b)\left(\int_{\underline{q}}^{\bar{q}} qdF + \int_{\underline{q}}^{q^*}(1-q)dF\right) > p_b + (1 - p_a - p_b)\int_{q^*}^{\bar{q}}(1-q)dF$, and $p_a + (1 - p_a - p_b)\left(\int_{\underline{q}}^{\bar{q}}(1-q)dF + \int_{\underline{q}}^{q^*} qdF\right) < p_b + (1 - p_a - p_b)\int_{q^*}^{\bar{q}} qdF$. The existence of this q^* follows from the fact that there exists a $\hat{q} \in (\underline{q}, \bar{q})$ such that $p_a + (1 - p_a - p_b)\left(\int_{\underline{q}}^{\bar{q}} qdF + \int_{\underline{q}}^{\hat{q}}(1-q)dF\right) = p_b + (1 - p_a - p_b)\int_{\hat{q}}^{\bar{q}}(1-q)dF$. Such $\hat{q} \in (\underline{q}, \bar{q})$ implies that $p_a + (1 - p_a - p_b)\left(\int_{\underline{q}}^{\bar{q}}(1-q)dF + \int_{\underline{q}}^{\hat{q}} qdF\right) < p_b + (1 - p_a - p_b)\int_{\hat{q}}^{\bar{q}} qdF$ because $\int_{\underline{q}}^{\bar{q}} q > 0.5$, as implied by the first inequality condition assumed in the proposition, and thus, $\int_{\hat{q}}^{\bar{q}} qdF > \int_{\hat{q}}^{\bar{q}}(1-q)dF$. As a result, the desired inequalities both hold if we set $q^* = \hat{q} + \epsilon$, where is $\epsilon > 0$ arbitrarily small. Given this cutoff strategy described above with q^*, the relative turnout share for the correct policy is strictly greater than 0.5 in both states (and also in the limit) ensuring that the correct policy is chosen with a probability that goes to one in both states.

Next, we characterize the equilibria in which the correct policy is chosen with a probability that goes to one as $N \to \infty$ in both states. We assume away perception biases for ease of notation, which is without loss of generality for the equilibrium characterization under unawareness. The case with overconfidence and awareness is analyzed in the final part of the proof. Let q_N^a and q_N^b denote the respective equilibrium cutoffs for electorate size N. We need the following claim.

Claim 1 *Assume that $\bar{q} \geq 1 - \underline{q}$ and consider those equilibria in which the correct policy is chosen with a probability that goes to one in both states. Then, $\limsup_{N\to\infty} q_N^b < \bar{q}$.*

Proof. First, consider the case in which $\bar{q} < 1$ and assume towards a contradiction that $\limsup_{N\to\infty} q_N^b \geq \bar{q}$. By Lemma 6, this implies that $\limsup_{N\to\infty} \frac{\pi \Pr(piv_b|A,N)}{(1-\pi)\Pr(piv_b|B,N)} \geq \frac{\bar{q}}{1-\bar{q}}$. Then, (with an abuse of notation) there exists an electorate size N and equilibrium sequence (q_N^a, q_N^b) such that $\frac{(1-\pi)\Pr(piv_b|B,N)}{\pi \Pr(piv_b|A,N)}$ goes to a number weakly smaller than $\frac{1-\bar{q}}{\bar{q}}$. Moreover, $\frac{(1-\pi)\Pr(piv_a|B,N)}{\pi \Pr(piv_a|A,N)} \to \frac{\widehat{q}^a}{1-\widehat{q}^a}$, where $\widehat{q}^a = \lim_{N\to\infty} q_N^a \geq 0$ (taking a convergent subsequence of the subsequence if necessary). Note that $\widehat{q}^a < \bar{q}$ as otherwise policy b would win in both states with a probability that goes to one as $N \to \infty$, which is suboptimal. $\Pr(piv_a|B,N)$ and $\Pr(piv_b|B,N)$ have a common term (representing the tie events) and the other term differs by only a multiplicative term. That is, $\Pr(piv_a|B,N)$ can be written as $\Pr(piv_a|B,N) = x_N + y_N$, and thus, $\Pr(piv_b|B,N)$ is equal to $x_N + y_N\gamma_N$, where x_N and y_N represent the respective probabilities of a tie event and the case in which policy a is behind by one vote in state B, and

$$\gamma_N = \frac{p_a + (1-p_a-p_b)\left(\int_{q_N^a}^{\bar{q}}(1-q)dF + \int_{\underline{q}}^{1-q_N^a} qdF\right)}{p_b + (1-p_a-p_b)\left(\int_{q_N^b}^{\bar{q}} qdF + \int_{\underline{q}}^{1-q_N^b}(1-q)dF\right)}.$$

We assume without loss of generality that $1-q_N^a > \underline{q}$ and $q_N^b < \bar{q}$ for large N (the proof is virtually unaffected if $1-q_N^a \leq \underline{q}$ or $q_N^b \geq \bar{q}$). However, since $q_N^b \to \bar{q}$ and $\bar{q} > 1-\underline{q}$, $1-q_N^b < \underline{q}$ must hold for all large N. $\Pr(piv_a|A,N)$ and $\Pr(piv_b|A,N)$ also have a common term, and the other term differs by a multiplicative term. Thus, $\Pr(piv_a|A,N)$ and $\Pr(piv_b|A,N)$ can be written as $\Pr(piv_a|A,N) = w_N + z_N$, and $\Pr(piv_b|A,N) = w_N + z_N\xi_N$, where w_N and z_N represent the respective probabilities of a tie event and the case in which policy a is behind by one vote in state A, and

$$\xi_N = \frac{p_a + (1-p_a-p_b)\left(\int_{q_N^a}^{\bar{q}} qdF + \int_{\underline{q}}^{1-q_N^a}(1-q)dF\right)}{p_b + (1-p_a-p_b)\left(\int_{q_N^b}^{\bar{q}}(1-q)dF + \int_{\underline{q}}^{1-q_N^b} qdF\right)}.$$

As the correct policy must be chosen with a probability that goes to one in both states, $\xi_N > 1$ and $\gamma_N < 1$ must hold. Moreover, if $\bar{q} < 1$, then $q_N^a > 0.5$ for all large N because otherwise policy a is chosen in both states with a probability that goes to one since $q_N^b \to \bar{q}$ and $p_a + (1-p_a-p_b)\int_{0.5}^{\bar{q}}(1-q)dF > p_b$ as assumed in Proposition 3. Thus, $\widehat{q}^a = \lim_{N\to\infty} q_N^a \in (0.5, \bar{q})$. As a result,

$$\begin{aligned}
\frac{\bar{q}}{1-\bar{q}}\frac{\widehat{q}^a}{1-\widehat{q}^a} &\leq \lim_{N\to\infty} \frac{\frac{1-\pi}{\pi}\frac{x_N+y_N}{w_N+z_N}}{\frac{1-\pi}{\pi}\frac{x_N+\gamma_N y_N}{w_N+\xi_N z_N}} = \lim_{N\to\infty} \frac{x_N+y_N}{x_N+\gamma_N y_N}\frac{w_N+\xi_N z_N}{w_N+z_N} \\
&\leq \limsup_{N\to\infty} \frac{\xi_N}{\gamma_N} = \frac{p_a + (1-p_a-p_b)\left(\int_{\widehat{q}^a}^{\bar{q}} qdF + \int_{\underline{q}}^{1-\widehat{q}^a}(1-q)dF\right)}{p_a + (1-p_a-p_b)\left(\int_{\widehat{q}^a}^{\bar{q}}(1-q)dF + \int_{\underline{q}}^{1-\widehat{q}^a} qdF\right)}.
\end{aligned}$$

($1-\widehat{q}^a$ is relevant as a cutoff if $\underline{q} < 0.5$ and $1-\widehat{q}^a > \underline{q}$.) However, the final term above is strictly smaller than $\frac{\int_{\widehat{q}^a}^{\bar{q}} qdF + \int_{\underline{q}}^{1-\widehat{q}^a}(1-q)dF}{\int_{\widehat{q}^a}^{\bar{q}}(1-q)dF + \int_{\underline{q}}^{1-\widehat{q}^a} qdF}$, which in turn is strictly smaller than $\frac{\bar{q}}{1-\bar{q}}$, resulting in a contradiction. Thus, $\limsup_{N\to\infty} q_N^b < \bar{q} < 1$. We now analyze the case in which $\bar{q} = 1$. Assume towards a contradiction that $\limsup_{N\to\infty} q_N^b \geq 1$. This implies that $\limsup_{N\to\infty} \frac{\pi \Pr(piv_b|A,N)}{(1-\pi)\Pr(piv_b|B,N)} = \infty$. Abusing notation, we have that $\frac{\Pr(piv_b|B,N)}{\Pr(piv_b|A,N)} \to 0$. However, if $\frac{\Pr(piv_b|B,N)}{\Pr(piv_b|A,N)} \to 0$, then $\frac{\Pr(piv_a|B,N)}{\Pr(piv_a|A,N)} \to 0$ must also hold. This is because $\liminf_{N\to\infty} \gamma_N > 0$ and $\limsup_{N\to\infty} \xi_N < \infty$, and $\frac{\Pr(piv_b|B,N)}{\Pr(piv_b|A,N)} = \frac{x_N+\gamma_N y_N}{w_N+\xi_N z_N} \geq \frac{\gamma_N}{\xi_N}\frac{x_N+y_N}{w_N+z_N} = \frac{\gamma_N}{\xi_N}\frac{\Pr(piv_a|B,N)}{\Pr(piv_a|A,N)}$. In turn, $q_N^a \to 0$. As a result, almost everyone but b-partisans votes for policy a in large elections, and a is chosen in both states with a probability that goes to one, a contradiction.

Analysis without awareness: By what we have shown above, $\limsup_{N\to\infty} q_N^b < \bar{q}$. Therefore, in large elections a sufficiently high level of the Dunning-Kruger effect implies that

$$p_a+(1-p_a-p_b)\left(\int_{q^a}^{\bar{q}} qdF + \int_{p_o^{-1}(q^a)}^{q^a} \lambda_o(q)qdF + \int_{\underline{q}}^{p_o^{-1}(1-q^a)}(1-q)dF + \int_{p_o^{-1}(1-q^a)}^{1-q^a}(1-\lambda_o(q))(1-q)dF\right)$$

is strictly lower than

$$p_b+(1-p_a-p_b)\left(\int_{q^b}^{\bar{q}} (1-q)dF + \int_{p_o^{-1}(q^b)}^{q^b} \lambda_o(q)(1-q)dF + \int_{\underline{q}}^{p_o^{-1}(1-q^b)}qdF + \int_{p_o^{-1}(1-q^b)}^{1-q^b}(1-\lambda_o(q))qdF\right).$$

To see why, note that if $1-q^j$ is relevant as a cutoff (because $\underline{q} < 0.5$ and $1-q^j > \underline{q}$ for $j \in \{a,b\}$), then a sufficiently high Dunning-Kruger effect makes it trivial. Moreover, a sufficiently high Dunning-Kruger effect makes $p_o^{-1}(q^b)$ closer and closer to $\underline{q}$ substantially increasing turnout from individuals with low q, and as a result, the condition in the statement of the Proposition ensures that policy b is chosen in both states with a probability that goes to one as $N \to \infty$ (i.e., if $p_o(\underline{q})$ is sufficiently close to or greater than $\limsup_{N\to\infty} q_N^b$, and $\lambda_o(q)$ is sufficiently high at every $q \in (\underline{q}, \limsup_{N\to\infty} q_N^b)$, then policy b is chosen in both states with a probability that goes to one).

Analysis with awareness: Equilibrium analysis is cumbersome if individuals are aware of others' overconfidence. Therefore, we consider a finite type space with $\{q_1, q_2, ..., q_T\}$ where $q_T = 1$. Obviously $q_N^b < q_T = 1$ for every N in any responsive equilibrium. As a result, a sufficiently high level of the Dunning-Kruger effect in the population will prevent information aggregation despite awareness: if for example $p_o(q_1) = q_T$, then the sufficient condition ensures that there exist high enough $\{\lambda_o(q_t)\}_{t\in\{1,2,...,T-1\}}$ such that

$$\frac{p_b - p_a}{1-p_a-p_b} \geq \sum_{t=1}^{T} q_t + (1-\lambda_o(q_t))(1-q_t)\Pr(q_t) - \sum_{t=1}^{T} \lambda_o(q_t)(1-q_t)\Pr(q_t),$$

(i.e., even if we assume every unbiased i with $q_i < q_T$ votes against b regardless of signal) and thus, the expected turnout for b exceeds the expected turnout for a in state A.

A.4.2 Model with Asymmetry in Media Veracity across States

The signal precision of each individual is a function of media veracity v_S in state $S \in \{A, B\}$ and individual competence q, which is an i.i.d. draw from distribution $F[\underline{q}, \bar{q}]$. The respective signal precision for individual i in state A and state B is given by $g(q_i, v_A) = q_i$

and $g(q_i, v_B) = m(q_i)$, where $m(q)$ is increasing in q. We assume that $m(q) \leq q$ for every q and $m(q) < q$ for some q representing the media influence of a third party that strictly prefers policy a. The function $m(q)$ is common knowledge. While individual i knows the realization of q_i (or rather perceives it as $p_i(q_i)$), i does not know with certainty whether the true precision is q_i or $m(q_i)$ as it depends on the state of the world. Consider as an example the case where $m(q) = q^x$ and $x > 1$. As x increases, the average signal precision drastically falls in state B, and eventually $\int_{\underline{q}}^{\bar{q}} m(q)dF < 0.5$ holds. Note that $\int_{\underline{q}}^{\bar{q}} m(q)dF < 0.5$ can hold even if $\pi\underline{q} + (1-\pi)m(\underline{q}) \geq 0.5$ (that is, on average across the two states every individual is more likely to be correctly informed than misinformed). This represents the generalization of the simple setting in Example 1 in the main text.

Proposition 4 *Assume that π is bounded away from 0 and that $m(\bar{q}) > 1 - m(\underline{q})$. If media veracity in state B is sufficiently low so that $\int_{\underline{q}}^{\bar{q}} m(q)dF < 0.5 < \int_{\underline{q}}^{\bar{q}} qdF$, and the electorate is sufficiently overconfident, then the probability that policy a is chosen goes to one in both states as N goes to infinity, whereas in an unbiased electorate the correct policy is chosen with a probability that goes to one in both states.*

Proof. Equilibrium characterization in this setting closely resembles that in Lemma 1. Let q^a and q^b be as defined in (2) and (3) respectively. In characterizing equilibria, we assume away perception biases (i.e., we assume that $p_i(q_i) = q_i$) for brevity in notation, which is without loss of generality. Extending the steps in Lemma 1 (i.e., the case where $m(q_i) = q_i$), it follows that in every Bayesian Nash equilibrium, individual i votes for a if either i's signal is α and $\frac{q_i}{q_i+1-m(q_i)} \geq q^a$ or i's signal is β and $\frac{m(q_i)}{m(q_i)+1-q_i} \leq 1 - q^a$. In a similar vein, i votes for b if either i's signal is β and $\frac{m(q_i)}{m(q_i)+1-q_i} \geq q^b$ or i's signal is α and $\frac{q_i}{q_i+1-m(q_i)} \leq 1 - q^b$. As $\frac{q}{q+1-m(q)}$ and $\frac{m(q)}{m(q)+1-q}$ are monotone increasing in q, equilibrium voting behavior is once again characterized by cutoffs. One issue is proving the existence of an optimal strategy, for which the proof of Lemma 3 does not suffice. We will either assume that one exists or assume that there is a finite set of types, in which case an optimal symmetric strategy always exists and must coincide with the optimal symmetric equilibrium. The equilibrium characterization and arguments below are virtually unaffected if there is a finite set of q types.[62]

We now construct a strategy that fully aggregates information in the limit in the absence of perception biases, under the conditions stated in the Proposition. The strategy

[62] To be more precise, if there is a type q (or rather $p(q)$) that exactly equals one of the equilibrium cutoffs, that type may be randomizing in equilibrium. For example, if i's signal is α and it turns out that $q_i = q^a$, then i may randomize in equilibrium. However, such randomization will be accounted for in the pivotality calculus, and the formal equilibrium characterization is unaffected.

is such that i votes for the policy that matches s_i if $q_i \geq q^*$ and abstains otherwise, where q^* satisfies $\int_{q^*}^{\bar{q}} m(q)dF > \int_{q^*}^{\bar{q}} (1-m(q))dF$ (such q^* exists because $m(\bar{q}) > 0.5$ must hold by hypothesis) and $\int_{q^*}^{\bar{q}} qdF > \int_{q^*}^{\bar{q}} (1-q)dF$. Given this strategy, the relative turnout share for the correct policy is strictly greater than 0.5 in both states (and in the limit) ensuring that the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$. Thus, the optimal equilibrium will result in the same outcome in the limit.

Below, we assume away perception biases for ease of notation, which is without loss of generality for the equilibrium characterization under unawareness. The case with overconfidence and awareness is analyzed in the final part of the proof. Let q_N^a and q_N^b denote the equilibrium cutoff pair given electorate size N. We analyze those equilibria in which the correct policy is chosen with a probability that goes to one in both states. First, we consider the case in which $\bar{q} < 1$. By the assumption that $m(\bar{q}) > 1 - m(\underline{q})$, we have that $\frac{m(\bar{q})}{m(\bar{q})+1-\bar{q}} > \frac{1-m(\underline{q})}{\underline{q}+1-m(\underline{q})}$ and that $\bar{q} > 1 - \underline{q}$. Furthermore, $\bar{q} > 1 - \underline{q}$ and $m(\bar{q}) > 1 - m(\underline{q})$ imply that $\frac{\bar{q}}{\bar{q}+1-m(\bar{q})} > \frac{1-\underline{q}}{m(\underline{q})+1-\underline{q}}$. We now show that $\limsup_{N\to\infty} q_N^a < \frac{\bar{q}}{\bar{q}+1-m(\bar{q})}$ and $\limsup_{N\to\infty} q_N^b < \frac{m(\bar{q})}{m(\bar{q})+1-\bar{q}}$ must hold. Assume towards a contradiction that $\limsup_{N\to\infty} q_N^b \geq \frac{m(\bar{q})}{m(\bar{q})+1-\bar{q}}$. By what we have shown above, $\frac{m(\bar{q})}{m(\bar{q})+1-\bar{q}} > \frac{1-m(\underline{q})}{\underline{q}+1-m(\underline{q})}$, and thus, the turnout rate for policy b goes to zero as $N \to \infty$. This implies that $\limsup_{N\to\infty} q_N^a \geq \frac{\bar{q}}{\bar{q}+1-m(\bar{q})}$ must hold because, otherwise, the relative turnout rate for policy b (hence, the probability that b is chosen) goes to zero in state B, which contradicts our initial hypothesis. More generally, $\limsup_{N\to\infty} q_N^b \geq \frac{m(\bar{q})}{m(\bar{q})+1-\bar{q}}$ if and only if $\limsup_{N\to\infty} q_N^a \geq \frac{\bar{q}}{\bar{q}+1-m(\bar{q})}$. A strict inequality (or strict inequalities) can however not hold, as that would mean at least one policy receives zero votes in both states as $N \to \infty$, contradicting our initial hypothesis that the correct policy is chosen in both states with a probability that goes to one. Hence, $\limsup_{N\to\infty} \frac{(1-\pi)\Pr(piv_a|B,N)}{\pi \Pr(piv_a|A,N)} = \frac{\bar{q}}{1-m(\bar{q})}$ and $\limsup_{N\to\infty} \frac{(1-\pi)\Pr(piv_b|B,N)}{\pi \Pr(piv_b|A,N)} = \frac{1-\bar{q}}{m(\bar{q})}$. Thus, (with an abuse of notation) there exists an electorate size N and equilibrium sequence $\{q_N^a, q_N^b\}$ such that $\frac{(1-\pi)\Pr(piv_a|B,N)}{\pi \Pr(piv_a|A,N)} \to \frac{\bar{q}}{1-m(\bar{q})}$ and $\frac{(1-\pi)\Pr(piv_b|B,N)}{\pi \Pr(piv_b|A,N)} \to \frac{1-\bar{q}}{m(\bar{q})} > 0$.

We will now show that these limits result in a contradiction. To see why, first note that $\Pr(piv_a|B,N)$ and $\Pr(piv_b|B,N)$ have a common term (representing the tie events) and the other term differs by only a multiplicative term. That is, $\Pr(piv_a|B,N)$ can be written as $\Pr(piv_a|B,N) = x_N + y_N$, and thus, $\Pr(piv_b|B,N)$ is equal to $x_N + y_N\gamma_N$ where $\gamma_N = \frac{\int_{\hat{q}_N^a}^{\bar{q}} (1-m(q))dF}{\int_{\hat{q}_N^b}^{\bar{q}} m(q)dF}$, $\hat{q}_N^a$ and $\hat{q}_N^b$ are the respective values that solve for $q_N^a = \frac{\hat{q}_N^a}{\hat{q}_N^a+1-m(\hat{q}_N^a)}$ and $q_N^b = \frac{m(\hat{q}_N^b)}{m(\hat{q}_N^b)+1-\hat{q}_N^b}$, and x_N and y_N represent the respective probabilities of a tie event and the case in which policy a is behind by one vote in state B.[63] In a similar vein, $\Pr(piv_a|A,N)$ and $\Pr(piv_b|A,N)$ have a common term, and the other term differs by a multiplicative term. Thus, $\Pr(piv_a|A,N)$ can be written as $\Pr(piv_a|A,N) = w_N + z_N$, and $\Pr(piv_b|A,N)$ is equal to $w_N + z_N\xi_N$, where $\xi_N = \frac{\int_{\hat{q}_N^a}^{\bar{q}} qdF}{\int_{\hat{q}_N^b}^{\bar{q}} (1-q)dF}$, and w_N and ξ_N represent the respective probabilities of a tie event and the case in which policy a is behind by one vote in state A.

Note that $\xi_N > 1$ and $\gamma_N < 1$ must hold as we focus on equilibria in which the correct policy is chosen in both states with a probability that goes to one as $N \to \infty$. Moreover, the term ξ_N is bounded above by a finite number because $\xi_N \leq \frac{\bar{q}}{1-\bar{q}} \frac{1-F(\hat{q}_N^a)}{1-F(\hat{q}_N^b)}$ and $1 > \gamma_N \geq \frac{1-m(\bar{q})}{m(\bar{q})} \frac{1-F(\hat{q}_N^a)}{1-F(\hat{q}_N^b)}$ imply that $\xi_N \leq \frac{\bar{q}}{1-\bar{q}} \frac{m(\bar{q})}{1-m(\bar{q})} < \infty$ as $\bar{q} < 1$. Therefore, $\limsup_{N\to\infty} \xi_N < \infty$. Moreover, $\liminf_{N\to\infty} \gamma_N = \frac{1-m(\bar{q})}{m(\bar{q})} \liminf_{N\to\infty} \frac{1-F(\hat{q}_N^a)}{1-F(\hat{q}_N^b)} > 0$ because $\frac{1-F(\hat{q}_N^a)}{1-F(\hat{q}_N^b)} > \frac{1-\bar{q}}{\bar{q}} > 0$ from $\xi_N > 1$ and $\bar{q} < 1$. Since $\frac{(1-\pi)\Pr(piv_a|B,N)}{\pi \Pr(piv_a|A,N)} \to \frac{\bar{q}}{1-m(\bar{q})}$ by hypothesis, $\lim_{N\to\infty} \frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)} = \frac{\bar{q}}{1-m(\bar{q})}$. We now need the following claim.

Claim 2 *Assume that $\bar{q} < 1$ and consider a sequence of N such that $\xi_N > 1 > \gamma_N$ with an equilibrium abstention rate bounded away from 0 (i.e., it is not necessary that $\hat{q}_N^a \to \bar{q}$ and $\hat{q}_N^b \to \bar{q}$ hold). Under these conditions,* $\liminf_{N\to\infty} x_N/y_N > 0$ *and* $\limsup_{N\to\infty} x_N/y_N < \infty$. *Similarly,* $\liminf_{N\to\infty} w_N/z_N > 0$ *and* $\limsup_{N\to\infty} w_N/z_N < \infty$ *must hold.*

Proof. It is enough to prove that $\liminf_{N\to\infty} x_N/y_N > 0$ and $\limsup_{N\to\infty} x_N/y_N < \infty$ as the proof of the other statement is analogous. First, we prove that $\liminf_{N\to\infty} x_N/y_N > 0$. Suppose towards a contradiction that there exists a sequence, which we again denote by N, such that x_N/y_N goes to 0. Thus, $x_N/\gamma_N y_N$ goes to 0 as well because $\liminf_{N\to\infty} \gamma_N > 0$ as we showed above. Let $\Pr(t_a, t_b|B,N)$ denote the probability that there are t_a votes for policy a and t_b votes for policy b in state B with electorate size N. It can be checked that $\Pr(t+1,t|B,N) = \Pr(t,t|B,N)\frac{N-2t}{t+1}\frac{T_a(B,N)}{1-(T_a(B,N)+T_b(B,N))}$, where $T_j(S,N)$ denotes the expected turnout rate for policy $j \in \{a,b\}$ in state $S \in \{A,B\}$ with electorate size N (thus, $\gamma_N = \frac{T_a(B,N)}{T_b(B,N)}$). If $NT_b(B,N)$ goes to a finite number as $N \to \infty$, then obviously $x_N/\gamma_N y_N$ cannot go to 0 because it is greater than $\frac{1-(T_a(B,N)+T_b(B,N))}{NT_a(B,N)}$, which is bounded below by a number strictly larger than zero for any N because by hypothesis $\gamma_N < 1$ and thus, $T_a(B,N) < T_b(B,N)$. Next, assume that $NT_b(B,N)$ goes to infinity as $N \to \infty$. Note that $\Pr(t,t|B,N) = \Pr(t,t-1|B,N)\frac{N-2t+1}{t}\frac{T_b(B,N)}{1-(T_a(B,N)+T_b(B,N))}$, and that since $NT_b(B,N)$ goes to infinity as $N \to \infty$, there must exist $t^*(N) > 0$ such that $\frac{N-2t^*(N)+1}{t^*(N)}\frac{T_b(B,N)}{1-(T_a(B,N)+T_b(B,N))} >$

[63] By hypothesis, $\hat{q}_N^a$ and $\hat{q}_N^b$ are very close to $\frac{\bar{q}}{\bar{q}+1-m(\bar{q})}$ and $\frac{m(\bar{q})}{m(\bar{q})+1-\bar{q}}$, respectively, for large N. Therefore, $1-\hat{q}_N^a$ and $1-\hat{q}_N^b$ cutoffs are irrelevant, and voting against signal will not take place because $1-\hat{q}_N^a < \frac{m(\underline{q})}{m(\underline{q})+1-\underline{q}}$ and $1-\hat{q}_N^b < \frac{\underline{q}}{\underline{q}+1-m(\underline{q})}$ hold by the inequalities we have shown above.

$1 \geq \frac{N-2t^*(N)-1}{t^*(N)+1} \frac{T_b(B,N)}{1-(T_a(B,N)+T_b(B,N))}$. In particular, for every $t \geq t^*(N)$, $\frac{N-2t}{t+1} \frac{T_a(B,N)}{1-(T_a(B,N)+T_b(B,N))}$ is bounded above by $1 + \frac{1}{t^*(N)+1} \frac{T_a(B,N)}{1-(T_a(B,N)+T_b(B,N))}$, which is also bounded above by a finite number because abstention rate is bounded away from 0 by hypothesis. Recalling that $\Pr(t+1, t|B, N) = \Pr(t, t|B, N) \frac{N-2t}{t+1} \frac{T_a(B,N)}{1-(T_a(B,N)+T_b(B,N))}$, x_N/y_N cannot go to 0, a contradiction.

We now show that $\limsup_{N\to\infty} x_N/y_N < \infty$. Suppose towards a contradiction that there exists a sequence which we again denote by N such that y_N/x_N goes to 0. Note that $\Pr(t, t+1|B, N) = \Pr(t, t|B, N) \frac{N-2t}{t+1} \frac{T_b(B,N)}{1-(T_a(B,N)+T_b(B,N))}$, and $\Pr(t+1, t+1|B, N) = \Pr(t, t+1|B, N) \frac{N-2t-1}{t+1} \frac{T_a(B,N)}{1-(T_a(B,N)+T_b(B,N))}$. If $NT_b(B, N)$ goes to a finite number as $N \to \infty$, then y_N/x_N cannot go to 0 because $\frac{y_N}{x_N} > \frac{1}{2} \min\{\frac{1-(T_a(B,N)+T_b(B,N))}{(N-1)T_a(B,N)}, \frac{NT_b(B,N)}{1-(T_a(B,N)+T_b(B,N))}\}$, which converges to a number strictly larger than zero.[64] Next, assume that $NT_b(B, N)$ goes to infinity as $N \to \infty$. Then, for large N there must exist $t^*(N) \geq 1$ such that $\frac{N-2t^*(N)}{t^*(N)+1} \frac{T_b(B,N)}{1-(T_a(B,N)+T_b(B,N))} > 1$ and $\frac{N-2t}{t+1} \frac{T_b(B,N)}{1-(T_a(B,N)+T_b(B,N))} \leq 1$ for $t > t^*(N)$. Thus, $\frac{N-2t-1}{t+1} \frac{T_a(B,N)}{1-(T_a(B,N)+T_b(B,N))} < 1$ for $t > t^*(N)$. This proves that y_N/x_N cannot go to 0, a contradiction. Hence, Claim 2 is proved.

We now prove that the limits $\frac{(1-\pi)\Pr(piv_a|B,N)}{\pi \Pr(piv_a|A,N)} \to \frac{\bar{q}}{1-m(\bar{q})}$ and $\frac{(1-\pi)\Pr(piv_b|B,N)}{\pi \Pr(piv_b|A,N)} \to \frac{1-\bar{q}}{m(\bar{q})}$ result in a contradiction. Taking a convergent subsequence of x_N/y_N (and a convergent subsubsequence w_N/z_N if necessary), we can write

$$\lim_{N\to\infty} \frac{(1-\pi)(x_N + y_N)}{\pi(w_N + z_N)} = \lim_{N\to\infty} \frac{(1-\pi)(x_N/y_N + 1)}{\pi(w_N/z_N + 1)} \lim_{N\to\infty} \frac{y_N}{z_N} = \frac{(1-\pi)(c+1)}{\pi(d+1)} \lim_{N\to\infty} \frac{y_N}{z_N},$$

where, abusing notation, $c = \lim_{N\to\infty} \frac{x_N}{y_N} \in (0, \infty)$ and $d = \lim_{N\to\infty} \frac{w_N}{z_N} \in (0, \infty)$ by Claim 2. Thus, $\lim_{N\to\infty} y_N/z_N$ exists, and $\lim_{N\to\infty} y_N/z_N \in (0, \infty)$ since $\lim_{N\to\infty} \frac{(1-\pi)(x_N/y_N+1)}{\pi(w_N/z_N+1)} = \frac{(1-\pi)(c+1)}{\pi(d+1)} \in (0, \infty)$ and $\lim_{N\to\infty} \frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)} = \frac{\bar{q}}{1-m(\bar{q})} \in (0, \infty)$. Thus, we have that

$$\frac{\bar{q}}{1-m(\bar{q})} = \frac{(1-\pi)(c+1)}{\pi(d+1)} \lim_{N\to\infty} \frac{y_N}{z_N},$$

and from $\lim_{N\to\infty} \frac{(1-\pi)(x_N+\gamma_N y_N)}{\pi(w_N+\xi_N z_N)} = \frac{1-\bar{q}}{m(\bar{q})}$, we have that

$$\frac{1-\bar{q}}{m(\bar{q})} = \frac{(1-\pi)(c + \lim_{N\to\infty} \gamma_N)}{\pi(d + \lim_{N\to\infty} \xi_N)} \lim_{N\to\infty} \frac{y_N}{z_N}.$$

[64] By initial hypothesis, $NT_b(B, N)$ cannot go to 0 as $N \to \infty$ because $NT_b(B, N) \to 0$ would in the limit result in a strictly positive probability that policy a is chosen in state B.

From these inequalities, it follows that

$$\frac{\bar{q}}{1-m(\bar{q})}\frac{m(\bar{q})}{1-\bar{q}}\frac{d+1}{c+1} = \frac{d+\lim_{N\to\infty}\xi_N}{c+\lim_{N\to\infty}\gamma_N}.$$

However, $\lim_{N\to\infty}\frac{\xi_N}{\gamma_N} = \frac{\bar{q}}{1-\bar{q}}\frac{m(\bar{q})}{1-m(\bar{q})} = \frac{\lim_{N\to\infty}\xi_N}{\lim_{N\to\infty}\gamma_N}$ with $\lim_{N\to\infty}\gamma_N \leq 1 \leq \lim_{N\to\infty}\xi_N$, $c = (0,\infty)$, and $d = (0,\infty)$. Thus, the equality above cannot hold. As a result, there must exist a $\hat{q} < \bar{q}$ such that $\hat{q}_N^a$ and $\hat{q}_N^b$ are smaller than $\hat{q}$ for all N. Hence, $\hat{q}_N^a$ and $\hat{q}_N^b$ are bounded above away from $\bar{q}$ for $\bar{q} < 1$ for any equilibrium sequence such that the correct policy is chosen with a probability that goes to one as $N \to \infty$ in both states.

To analyze the case in which $\bar{q} = 1$, we impose *Assumption 1* presented above: a small fraction $\frac{\epsilon}{2} > 0$ of the electorate always votes for a, and a fraction $\frac{\epsilon}{2} > 0$ always votes for b. Note that such an assumption does not affect our results with $\bar{q} < 1$. Assume towards a contradiction that $\limsup_{N\to\infty} q_N^b \geq \frac{m(\bar{q})}{m(\bar{q})+1-\bar{q}} = 1$. As in the case with $\bar{q} < 1$, $\limsup_{N\to\infty} q_N^b \geq 1$ if and only if $\limsup_{N\to\infty} q_N^a \geq \frac{\bar{q}}{\bar{q}+1-m(\bar{q})}$. In particular, $\limsup_{N\to\infty} q_N^a = \frac{\bar{q}}{\bar{q}+1-m(\bar{q})}$ and $\limsup_{N\to\infty} q_N^b = 1$ as strict inequalities cannot hold in a sequence of optimal equilibria. These imply that $\limsup_{N\to\infty}\frac{(1-\pi)\Pr(piv_a|B,N)}{\pi\Pr(piv_a|A,N)} = \frac{1}{1-m(\bar{q})}$ and $\limsup_{N\to\infty}\frac{(1-\pi)\Pr(piv_b|B,N)}{\pi\Pr(piv_b|A,N)} = 0$. Thus, (with an abuse of notation) there exists an equilibrium sequence q_N^a and q_N^b giving rise to $\frac{(1-\pi)\Pr(piv_a|B,N)}{\pi\Pr(piv_a|A,N)} \to \frac{1}{1-m(\bar{q})}$ and $\frac{(1-\pi)\Pr(piv_b|B,N)}{\pi\Pr(piv_b|A,N)} \to 0$ with $\hat{q}_N^a \to 1$ and $\hat{q}_N^b \to 1$. Using the previous notation introduced above, these limits translate to $\lim_{N\to\infty}\frac{(1-\pi)(x_N+y_N)}{\pi(w_N+z_N)} = \frac{1}{1-m(\bar{q})}$ and $\lim_{N\to\infty}\frac{(1-\pi)(x_N+\gamma_N y_N)}{\pi(w_N+\xi_N z_N)} = 0$. Thus, $\lim_{N\to\infty}\frac{x_N+y_N}{x_N+\gamma_N y_N}\frac{w_N+\xi_N z_N}{w_N+z_N} = \infty$. However, this is impossible because $\lim_{N\to\infty}\frac{x_N+y_N}{x_N+\gamma_N y_N}\frac{w_N+\xi_N z_N}{w_N+z_N} \leq \lim_{N\to\infty}\frac{\xi_N}{\gamma_N} = 1$. The equality $\lim_{N\to\infty}\frac{\xi_N}{\gamma_N} = 1$ holds since $\hat{q}_N^a \to 1$ and $\hat{q}_N^b \to 1$,

$$\gamma_N = \frac{\epsilon/2 + (1-\epsilon)\int_{\hat{q}_N^a}^1 (1-m(q))dF}{\epsilon/2 + (1-\epsilon)\int_{\hat{q}_N^b}^1 m(q)dF},$$

and

$$\xi_N = \frac{\epsilon/2 + (1-\epsilon)\int_{\hat{q}_N^a}^1 q dF}{\epsilon/2 + (1-\epsilon)\int_{\hat{q}_N^b}^1 (1-q)dF}.$$

Analysis without awareness: We have shown above that $\hat{q}_N^a$ is bounded above away from $\bar{q}$ for $\bar{q} \leq 1$ for any equilibrium sequence such that the correct policy is chosen with a probability that goes to one as $N \to \infty$ in both states. Given the bound on the equilibrium cutoff $\hat{q}_N^a$, we can construct an overconfidence function (as we did in the proof of Proposition 2 or Proposition 3) such that the probability that policy b is chosen in state B goes to zero because too many people follow their α signal (i.e., vote for policy a) due to the fact that

$\int_{\underline{q}}^{\bar{q}} m(q)dF < 0.5.$

Analysis with awareness: Equilibrium analysis is cumbersome if individuals are aware of others' overconfidence. Therefore, we consider a finite type space with $\{q_1, q_2, ..., q_T\}$ where $q_T = 1 = m(q_T)$. Obviously, $q_N^b < q_T = 1$ for every N in any responsive equilibrium. As a result, a sufficiently high level of the Dunning-Kruger effect in the population will prevent information aggregation despite awareness: if for example $p_o(q_1) = q_T$, then the sufficient condition ensures that there exist high enough $\{\lambda_o(q_t)\}_{t\in\{1,2,...,T-1\}}$ such that

$$\sum_{t=1}^{T} m(q_t) + (1 - \lambda_o(q_t))(1 - m(q_t)) \Pr(q_t) < \sum_{t=1}^{T} \lambda_o(q_t)(1 - m(q_t)) \Pr(q_t),$$

(i.e., even if we assume every unbiased i with $q_i < q_T$ votes against a regardless of s_i) and thus, the expected turnout for a exceeds the expected turnout for b in state B.

A.4.3 Model with Correlated News Signals

Proposition 5 *Assume that $\pi \in (1 - q_H, q_H)$ is bounded away from q_H and $1 - q_H$, and that m and n are odd numbers.* **(i)** *If $q_H = 1$ and $\mathbb{E}(q) < 0.5$, an unbiased electorate makes the correct decision with a probability that goes to one in both states as N goes to infinity, whereas in a sufficiently overconfident electorate the limiting probability is strictly lower than one.* **(ii)** *If $q_H < 1$, an unbiased electorate makes the correct decision with a probability that goes to $\sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q_H^i (1-q_H)^{n-i}$ in both states as N goes to infinity, whereas in a sufficiently overconfident electorate the limiting probability is strictly lower if $\bar{q} > \frac{n}{n+1}$ and $\mathbb{E}(q) < \frac{n}{n+1}$ hold.*

Proof. We first show that every responsive equilibrium consists of cutoffs $\underline{q}^a$, $\bar{q}^a$, $\underline{q}^b$ and $\bar{q}^b$ and derive Lemma 7 below. We start by assuming that the signal of individual i is α. In that case, i prefers voting for a over abstention if and only if

$$\Pr(piv_a \cap S = A | s_i = \alpha) - \Pr(piv_a \cap S = B | s_i = \alpha) \geq 0$$

as in the independent signal case. Different from the independent signal case, we must differentiate between the case where s_i comes from a high quality source and the case where it comes from a low quality source. Let $Q_i = H$ ($Q_i = L$) denote the case where s_i comes from a high quality (low quality) source. It can be checked that $\Pr(piv_a \cap S = A | s_i = \alpha)$ equals

$$\Pr(piv_a \cap S = A | \left(s_i = \alpha \cap Q_i = H\right) \cup \left(s_i = \alpha \cap Q_i = L\right)),$$

and thus, $\Pr(piv_a \cap S = A | s_i = \alpha)$ equals

$$\frac{\Pr(piv_a \cap S = A \cap s_i = \alpha \cap Q_i = H) + \Pr(piv_a \cap S = A \cap s_i = \alpha \cap Q_i = L)}{\Pr(s_i = \alpha)},$$

where, for example, $\Pr(piv_a \cap S = A \cap s_i = \alpha \cap Q_i = H) = q_i \pi \Pr(piv_a \cap s_i = \alpha | S = A \cap Q_i = H)$. Thus, we have that i weakly prefers voting for a over abstention with $s_i = \alpha$ if and only if

$$p_i(q_i) \geq \frac{x}{x+y},$$

where

$$\begin{aligned} x &= (1-\pi)\Pr(piv_a \cap s_i = \alpha | S = B \cap Q_i = L) - \pi \Pr(piv_a \cap s_i = \alpha | S = A \cap Q_i = L) \\ y &= \pi \Pr(piv_a \cap s_i = \alpha | S = A \cap Q_i = H) - (1-\pi)\Pr(piv_a \cap s_i = \alpha | S = B \cap Q_i = H). \end{aligned}$$

Next, we consider the (weak) preference of i with $s_i = \alpha$ for voting for a over voting for b. This will be the case if and only if $p_i(q_i) \geq \frac{x+w}{x+w+y+z}$, where x and y are as defined above, and

$$\begin{aligned} w &= (1-\pi)\Pr(piv_b \cap s_i = \alpha | S = B \cap Q_i = L) - \pi \Pr(piv_b \cap s_i = \alpha | S = A \cap Q_i = L) \\ z &= \pi \Pr(piv_b \cap s_i = \alpha | S = A \cap Q_i = H) - (1-\pi)\Pr(piv_b \cap s_i = \alpha | S = B \cap Q_i = H). \end{aligned}$$

Hence, we derive the cutoff $\bar{q}^a$ for voting for policy a conditional on $s_i = \alpha$: $\bar{q}^a = \max\{\frac{x}{x+y}, \frac{x+w}{x+w+y+z}\}$. In a similar vein, it can be shown that i with $s_i = \alpha$ prefers voting for b over abstention and voting for a if and only if $p_i(q_i) \leq \min\{\frac{w}{w+z}, \frac{x+w}{x+w+y+z}\}$. Hence, we derive the cutoff $\underline{q}^b$ for voting for b conditional on $s_i = \alpha$: $\underline{q}^b = \min\{\frac{w}{w+z}, \frac{x+w}{x+w+y+z}\}$. The derivation of $\underline{q}^a$ and $\bar{q}^b$ are analogous and therefore omitted. Thus, we obtain Lemma 7.

Lemma 7 *In the correlated signal model, every responsive and symmetric Bayesian Nash equilibrium consists of four cutoffs $\underline{q}^a$, $\bar{q}^a$, $\underline{q}^b$ and $\bar{q}^b$ such that* (1) *an individual votes for a if and only if either i's signal is α and $p_i(q_i) \geq \bar{q}^a$ or i's signal is β and $p_i(q_i) \leq \underline{q}^a$; and* (2) *an individual votes for b if and only if either i's signal is β and $p_i(q_i) \geq \bar{q}^b$ or i's signal is α and $p_i(q_i) \leq \underline{q}^b$.*

In order to characterize equilibria with $q_H < 1$ further, let x denote the realized number of high quality news outlets (out of a total of n) such that $s_h = \alpha$. In a similar vein, let y denote the realized number of low quality news outlets (out of a total of m) such that $s_l = \alpha$. One thing to note is that given the realized x and y values, the conditional turnout rate is exactly the same in the two states. Moreover, the conditional *relative* turnout rates are exactly the same. To see why, let $P^j_{x,y,A}$ and $P^j_{x,y,B}$ denote the respective turnout rate for policy $j \in \{a, b\}$ conditional on x and y realizations in state A and state B. Given equilibrium cutoffs $\underline{q}^a$, $\bar{q}^a$, $\underline{q}^b$ and $\bar{q}^b$, $P^a_{x,y,A}$ equals

$$\frac{x}{n}\int_{\bar{q}^a}^{\bar{q}} qdF + \frac{y}{m}\int_{\bar{q}^a}^{\bar{q}}(1-q)dF + \frac{n-x}{n}\int_{\underline{q}}^{\underline{q}^a} qdF + \frac{m-y}{m}\int_{\underline{q}}^{\underline{q}^a}(1-q)dF,$$

whereas $P^b_{x,y,A}$ equals

$$\frac{n-x}{n}\int_{\bar{q}^b}^{\bar{q}} qdF + \frac{m-y}{m}\int_{\bar{q}^b}^{\bar{q}}(1-q)dF + \frac{x}{n}\int_{\underline{q}}^{\underline{q}^b} qdF + \frac{y}{m}\int_{\underline{q}}^{\underline{q}^b}(1-q)dF.$$

It can be checked that $P^a_{x,y,B}$ is exactly the same as $P^a_{x,y,A}$; i.e., $P^a_{x,y,A} = P^a_{x,y,B}$. This is true also with perception biases (with and without awareness). Therefore, hereafter $P^a_{x,y}$ denotes the turnout rate for policy a conditional on x and y realizations in either state. In a similar vein, it can be checked that $P^b_{x,y,A} = P^b_{x,y,B}$. Thus, hereafter $P^b_{x,y}$ denotes the turnout rate for policy b conditional on x and y realizations in either state. Finally let $P_{x,y}$ denote the total turnout rate conditional on x and y realizations in either state; that is, $P_{x,y} = P^a_{x,y} + P^b_{x,y}$. Thus, the expected probability of selecting the correct policy equals

$$\sum_{x=0}^{n}\sum_{y=0}^{m}\sum_{T=0}^{N}\binom{N}{T}P^T_{x,y}(1-P_{x,y})^{N-T}(\pi M_A(x,y)P(a|x,y,T) + (1-\pi)M_B(x,y)P(b|x,y,T)),$$

where $M_S(x,y)$ denotes the probability of (x,y) realizations in state $S \in \{A, B\}$, and $P(j|x,y,T)$ denotes the probability that policy $j \in \{a, b\}$ wins given x and y realizations and realized turnout being equal to T. For example, for T odd $P(a|x,y,T)$ is equal to $\sum_{i=\frac{T+1}{2}}^{T}\binom{T}{i}\left(\frac{P^a_{x,y}}{P_{x,y}}\right)^i\left(\frac{P^b_{x,y}}{P_{x,y}}\right)^{T-i}$.

We can now prove the result below for unbiased electorates. Note that from here onward, we will either assume that an optimal strategy exists or that there is a finite set of types, in which case optimal symmetric strategy exists and coincides with the optimal symmetric equilibrium. The equilibrium characterization and our arguments are virtually unchanged if there is a finite set of q types.[65]

Lemma 8 *Let π be bounded above away from q_H and below away from $1 - q_H$. In large elections with $q_H < 1$, optimal equilibria are such that the realization of x determines whether*

[65] To be more precise, if there is a type q (or rather type $p(q)$) that exactly equals one of the equilibrium cutoffs, this type may be randomizing in equilibrium. However, this does not affect the formal characterization of equilibrium cutoffs above.

$\frac{P^a_{x,y}}{P_{x,y}} > \frac{1}{2}$ or $\frac{P^a_{x,y}}{P_{x,y}} < \frac{1}{2}$. More precisely, in every sufficiently large election, if $x \geq \frac{n+1}{2}$, then $\frac{P^a_{x,y}}{P_{x,y}} > \frac{1}{2}$, and if $x < \frac{n+1}{2}$, then $\frac{P^a_{x,y}}{P_{x,y}} < \frac{1}{2}$ (regardless of y). As a result, the probability of selecting the correct policy converges to $\sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q_H^i (1-q_H)^{n-i}$ as $N \to \infty$. If $q_H = 1$, the probability converges to 1 as $N \to \infty$.

Proof. We start with the case where $q_H = 1$. It is enough to note that there always exists $\hat{q} \in (\underline{q}, \bar{q})$ such that $\int_{\hat{q}}^{\bar{q}} q dF > \int_{\hat{q}}^{\bar{q}} (1-q) dF$ since $\bar{q} > 0.5$. Thus, the voting strategy with $\underline{q}^a = \underline{q}^b = \underline{q}$ and $\bar{q}^a = \bar{q}^b = \hat{q}$ ensures that the relative turnout share for the correct policy is strictly greater than 0.5 in either state, which results in the correct policy being chosen in both states with a probability that goes to one as $N \to \infty$. Therefore, the optimal equilibrium strategy will generate the same result in the limit. Next, assume that $q_H < 1$. We first construct the following strategy. There exists a $\hat{q} \in (\underline{q}, \bar{q})$ such that $\frac{1}{n} \int_{\hat{q}}^{\bar{q}} q dF > \int_{\hat{q}}^{\bar{q}} (1-q) dF$ since $\bar{q} > \frac{n}{n+1}$. Thus, the voting strategy with $\underline{q}^a = \underline{q}^b = \underline{q}$ and $\bar{q}^a = \bar{q}^b = \hat{q}$ ensures that whenever a majority of the high quality sources provide an α signal, the relative turnout rate for policy a is strictly greater than 0.5. As a result, the probability of selecting the correct policy goes to $\sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} q_H^k (1-q_H)^{n-k}$ in both states as N goes to infinity, and thus, the optimal equilibrium strategy can only improve it. The probability of selecting the correct policy which we denote by C equals

$$\sum_{x=0}^{n} \sum_{y=0}^{m} \sum_{T=0}^{N} \binom{N}{T} P_{x,y}^T (1-P_{x,y})^{N-T} (\pi M_A(x,y) P(a|x,y,T) + (1-\pi) M_B(x,y)(1-P(a|x,y,T))$$

by what we have shown above. Note that C is maximized if $P(a|x,y,T) = 1$ for $\pi M_A(x,y) > (1-\pi) M_B(x,y)$ and $P(a|x,y,T) = 0$ for $\pi M_A(x,y) < (1-\pi) M_B(x,y)$. Given that π is bounded away from q_H and $1-q_H$, we have that $\pi M_A(x,y) > (1-\pi) M_B(x,y)$ if and only if $x \geq \frac{n+1}{2}$, and $\pi M_A(x,y) < (1-\pi) M_B(x,y)$ if and only if $x < \frac{n+1}{2}$ (recall than n is odd). As a result, C is bounded above by $\sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q_H^i (1-q_H)^{n-i}$. However, since the probability of selecting the correct policy converges to $\sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q_H^i (1-q_H)^{n-i}$ as N goes to infinity in the strategy that we constructed above, it must also converge to $\sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q_H^i (1-q_H)^{n-i}$ in the optimal equilibrium. We can now finish the proof of the lemma. Suppose towards a contradiction that there exists an electorate size sequence for which the claim does not hold in the optimal equilibrium; e.g., there exists some $x \geq \frac{n+1}{2}$ such that $\frac{P^a_{x,y}}{P_{x,y}} \leq \frac{1}{2}$ for every element in that sequence. Then, it can be checked that the limit probability of selecting the correct policy is bounded above away from $\sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q_H^i (1-q_H)^{n-i}$ in that sequence, a contradiction. Hence, the lemma is proved.

To make the analysis of the effect of overconfidence tractable, we will assume a finite type space with $\{q_1, q_2, ..., q_T\}$ where $q_1 = \underline{q}$ and $q_T = \bar{q} \leq 1$. Note that none of the results we proved above rely on a continuum type space. We first assume that $q_H < 1$. Assume that $\limsup_{N\to\infty} \bar{q}_N^a \geq \bar{q}$ and that $\limsup_{N\to\infty} \bar{q}_N^b \geq \bar{q}$ focusing on optimal equilibria. It follows that $\limsup_{N\to\infty} \bar{q}_N^a = \limsup_{N\to\infty} \bar{q}_N^b = \bar{q}$. For example, $\limsup_{N\to\infty} \bar{q}_N^a > \bar{q}$ would violate Lemma 8 if $x = n$ and $y = m$. Let $\sigma_N(\bar{q}, s)$ denote the probability that individual i with precision $q_i = \bar{q}$ and signal s votes for the policy that matches s with electorate size N. Also, let $\sigma_N(\underline{q}, s)$ denote the probability that an individual with $q_i = \underline{q}$ and signal s votes against s. Assume without loss of generality that $\sigma_N(\bar{q}, \alpha) \geq \sigma_N(\bar{q}, \beta)$ for those equilibria in which $\bar{q}_N^a = \bar{q}_N^b = \bar{q}$. Note that $\liminf_{N\to\infty} \sigma_N(\bar{q}, \alpha) > 0$ must hold. Suppose not. Then, (taking a convergent subsequence if necessary) we have not only $\sigma_N(\bar{q}, \alpha) \to 0$ but also $\sigma_N(\bar{q}, \alpha) \geq \sigma_N(\bar{q}, \beta) \to 0$. These require $\underline{q}_N^a \leq \underline{q}$ and $\underline{q}_N^b \leq \underline{q}$ to hold for large N. To see why, note for example that $\frac{P_{x,y}^b}{P_{x,y}} > \frac{1}{2}$ fails to hold and Lemma 8 is violated if $\underline{q}_N^a > \underline{q}$, $x = 0$, and $y = 0$. In fact, by the same argument, $\sigma_N(\underline{q}, \alpha) \to 0$ and $\sigma_N(\underline{q}, \beta) \to 0$ must hold if $\underline{q}_N^a = \underline{q}$ and $\underline{q}_N^b = \underline{q}$. However, $\sigma_N(\bar{q}, s) \to 0$ and $\sigma_N(\underline{q}, s) \to 0$ give rise to a contradiction. To see why, note that if both $\sigma_N(\bar{q}, s)$ and $\sigma_N(\underline{q}, s)$ are doubled for both $s = \alpha$ and $s = \beta$, this doubles $P_{x,y}$ for every x and y realization (strictly for some x and y) without changing $\frac{P_{x,y}^b}{P_{x,y}}$ and $\frac{P_{x,y}^b}{P_{x,y}}$. This strictly increases the expected payoff by Lemma 5, a contradiction. Thus, $\liminf_{N\to\infty} \sigma_N(\bar{q}, \alpha) > 0$. In fact, using the same argument we can show that $\sigma_N(\bar{q}, \alpha) \geq \min\{1, \frac{f(\underline{q})}{f(\bar{q})}\}$ must hold in the optimal equilibrium if $\bar{q}_N^a = \bar{q}_N^b = \bar{q}$. Assume without loss of generality that $\sigma_N(\bar{q}, \alpha) < 1$. Then, if $p_o^{-1}(\bar{q}) = \underline{q}$ and $\{\lambda_o(q)\}_{q \geq \underline{q}}$ is sufficiently high,

$$\sigma_N(\bar{q}, \alpha)\left(\frac{n-1}{2n}\left(\sum_{\underline{q}\leq q<\bar{q}} \lambda_o(q) q f(q) + \bar{q} f(\bar{q})\right) + \sum_{\underline{q}\leq q<\bar{q}} \lambda_o(q)(1-q) f(q) + (1-\bar{q}) f(\bar{q})\right) +$$
$$\sigma_N(\underline{q}, \beta)(1-\lambda_o(\underline{q})) f(\underline{q}) \underline{q} \frac{n+1}{2n}$$

is greater than

$$\sigma_N(\bar{q}, \beta)\frac{n+1}{2n}\left(\sum_{\underline{q}\leq q<\bar{q}} \lambda_o(q) q f(q) + \bar{q} f(\bar{q}))\right) + \sigma_N(\underline{q}, \alpha)\left(1 - \underline{q} + \underline{q}\frac{n-1}{2n}\right)(1-\lambda_o(\underline{q})) f(\underline{q})$$

if the sufficient condition $\mathbb{E}(q) < \frac{n}{n+1}$ holds.[66] This implies that if $x = \frac{n-1}{2}$, and $y = m$, then $\frac{P_{x,y}^b}{P_{x,y}} > \frac{1}{2}$ fails to hold, violating Lemma 8. The proof for the case with awareness is similar

[66] The proof is unaffected if $\underline{q}_N^a > \underline{q}$ or $\underline{q}_N^b > \underline{q}$.

but more tedious. The proof for the case where $\limsup_{N\to\infty} \bar{q}_N^j < \bar{q}$ for $j \in \{a,b\}$ or $q_H = 1$ is also similar.

A.5 An Alternative Interpretation and Model

In a slightly different formulation of the model with $\underline{q} < 0.5$, we can interpret $F[\underline{q}, \bar{q}]$ as the distribution of the accuracy of individual opinions. Then, the case where $q_i < 0.5$ is associated with not only widespread misinformation but also overconfidence such that $p(q_i) \geq 0.5$ because an individual cannot rationally hold on to a belief or opinion that they objectively know is more likely to be false than correct. In this model, not only does $p_i(q_i)$ differ from q_i for biased individuals, but also the distribution of q_i differs between the biased electorate and its unbiased version. In this alternative model, q_i cannot fall below 0.5 in an electorate that consists of only unbiased individuals. As a result, this model involves two different q_i distributions. However, our equilibrium characterization and other main results also hold in this alternative model under mild assumptions. For example, one possible and intuitive assumption is that the distribution of q_i is identical in the biased electorate and its unbiased version except that for every i such that $q_i < 0.5$ in the biased electorate, $q_i = 0.5$ in the unbiased case.

B Additional Experimental Analysis

B.1 Pooling Baseline Sessions

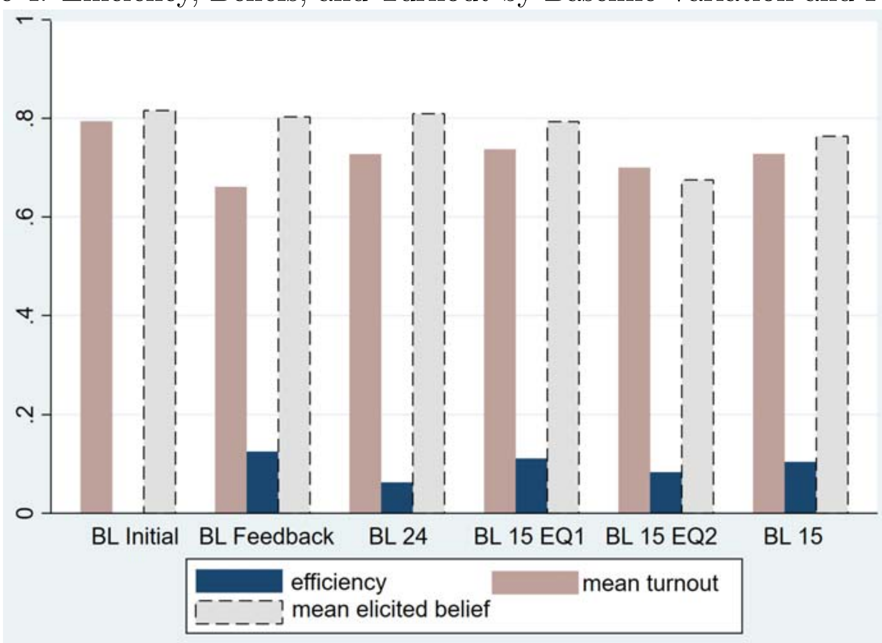
We first consider the twelve *Baseline* sessions with $N = 24$: six sessions of *Baseline Initial* (*BL Initial*) without feedback and six sessions of *Baseline Feedback* with aggregated feedback (*BL Feedback*). We test for differences at the session level using Mann-Whitney tests unless otherwise stated. *BL Initial* and *BL Feedback* show no or limited statistical variation in four key dimensions: efficiency, turnout, mean elicited belief, and quiz score. In particular, they do not statistically differ in mean elicited beliefs, quiz scores, or efficiency. The idea behind *BL Feedback* is to manipulate the confidence levels of subjects during the course of the experiment. Aggregated feedback resulted in lower turnout in *BL Feedback* as can be seen in Figure 4. However, the decrease in turnout is limited (about 13 percentage points). In particular, turnout did not sufficiently decline from those in the bottom 2/3. Thus, giving aggregated feedback was not successful as a debiasing mechanism and did not affect efficiency. Therefore, we pool the data from these twelve sessions with $N = 24$ and call them

Table 4: Features of Baseline Variations

	BL Initial	*BL Feedback*	*BL EQ1*	*BL 15 EQ2*
N	24	24	15	15
Aggregated Feedback	no	yes	no	no
Quiz difficulty	very easy (*EQ1*)	very easy (*EQ1*)	very easy (*EQ1*)	easy (*EQ2*)
# Sessions	6	6	6	2

Notes: This table shows the design features of our baseline variations.

Figure 4: Efficiency, Beliefs, and Turnout by Baseline Variation and Pooling



Notes: *BL 15 EQ1* refers to six sessions with $N = 15$ and the original very easy quiz ("*EQ1*") and *BL 15 EQ2* refers to two sessions with the slightly more difficult quiz ("*EQ2*").

Baseline 24 (*BL 24*).

Next, we consider the eight *Baseline* sessions with $N = 15$. Six of those exactly replicate *BL Initial* except that $N = 15$. In addition, two sessions have an easy quiz that is slightly more difficult than the original quiz to check the robustness of our inefficiency results to a modest increase in the quiz difficulty (see also Online Appendix B.4). This quiz variation is denoted by "*EQ2*" in Figure 4 ("*EQ1*" denotes the original "very easy" quiz). *EQ2* reduced the average quiz score (by less than 20%) and resulted in a lower mean $p_i(q_i)$ (see the quiz score distributions in Online Appendix B.4). However, the difference in mean $p_i(q_i)$ is limited at less than 12 percentage points, which is only weakly significant.[67] As a

[67]We can use a t-test since 90 subjects took *EQ1* and 30 subjects took *EQ2* in sessions with $N = 15$

Table 5: Pairwise Comparisons of Efficiency, Turnout, Beliefs, and Quiz Scores

	BL Initial	*BL Feedback*	*BL 24*	*BL EQ1*	*BL 15*
BL Initial	-	ns, ***, ns, ns	-	*, ns, ns, ns	*, *, ns, ns
BL Feedback	-	-	-	ns, ns, ns, ns	ns, *, ns, ns
BL 24	-	-	-	ns, ns, ns, ns	ns, ns, ns, ns
BL EQ1	-	-	-	-	-
BL 15	-	-	-	-	-

Notes: This table reports the test results for pairwise comparisons of the baseline variations. We use two-sided Mann-Whitney tests at the session level. *BL 24* is the merged *BL Initial* (6 sessions) and *BL Feedback* (6 sessions). *BL 15* is the merged *BL 15 EQ1* (6 sessions) and *BL 15 EQ2* (2 sessions). Test results are shown in the form a, b, c, d indicating the respective significance level for the pairwise comparisons of efficiency, turnout level, mean elicited belief, and quiz scores. ***, and * indicate significance at the 1% and 10% level, respectively. *ns* indicates no significance. Some cells are empty since, e.g., *BL 24* involves *BL Initial*, and thus, they cannot be compared.

result, the quiz variation had no discernible impact on turnout or efficiency as can be seen in Figure 4. Therefore, we pool the data from eight sessions with $N = 15$ and denote them by *BL 15* as mentioned in the main text.

Table 4 shows the features in which the baseline variations differ. Comparing *BL 24* and *BL 15*, we report that there is no statistically significant difference in any of the four key dimensions, namely efficiency, turnout, elicited beliefs, and quiz scores. Table 5 presents the test results for pairwise comparisons of the baseline variations in terms of efficiency, turnout, elicited beliefs, and quiz scores.

B.2 *OBJ* Analysis

As mentioned in the main text, the *OBJ* condition in the pre- and post-pandemic sessions differs in the number of rounds and N. However, this difference is inconsequential for the main variables of interest in *OBJ*, namely efficiency and the level of turnout. We test for differences at the session level using Mann-Whitney tests. In those sessions in which the total number of rounds is eight and N is either 15 or 16, the share of correct group decisions in *OBJ* is 93.5% with an average turnout rate of 44.9%. In sessions with 15 rounds and $N = 24$, the share of correct group decisions is 95%, and the average turnout rate is 46.5%. These differences are very small and statistically insignificant (the respective p-values for efficiency and turnout comparisons are 0.535 and 0.187).

Next, we analyze turnout behavior in *OBJ*. To that aim, we estimate a random

($p = 0.073$).

Table 6: Explaining Individual Turnout Decision in OBJ

	q_i	$p_i(vote_j = 1)$	Round#
Coefficient	0.968 (0.005)	0.129 (0.021)	-0.003 (0.001)

Notes: Coefficients show the average marginal effects in a random effects panel probit regression. The dependent variable is the subject's binary choice between voting (= 1) and abstaining/voting against s_i (= 0). Standard errors clustered by session are in parentheses.

effects panel probit model with the individual voting decision as the dependent variable. The independent variables are (i) the precision of subject i's signal ("q_i"); (ii) the elicited belief regarding other group members' likelihood of voting ("$p_i(vote_j = 1)$"); and (iii) a time trend. We cluster errors at the session level. Table 6 presents the average marginal effects in this regression. As predicted, q_i is economically large and statistically strongly significant ($p < 0.001$).

B.3 Omitted Wilcoxon Signed-rank Test Results

As mentioned in the main text, we use both the sign test and the Wilcoxon signed-rank test in paired-data and one-sample tests, and we report only the sign test p-value unless the two tests disagree. Below, we provide the Wilcoxon signed-rank test results that are not reported in the main text. Recall that tests are one sided in Section IIIA and two sided in Section IIIB.

Result 1 in Section IIIA Comparing efficiency in *BL* and *BL 15* against 50%: $p < 0.001$ and $p = 0.005$ respectively.

Result 2 in Section IIIA Comparing efficiency in *HQ* against 50%: $p = 0.124$.

Result 3 in Section IIIA Comparing efficiency in *SV* against 66%: $p = 0.013$.

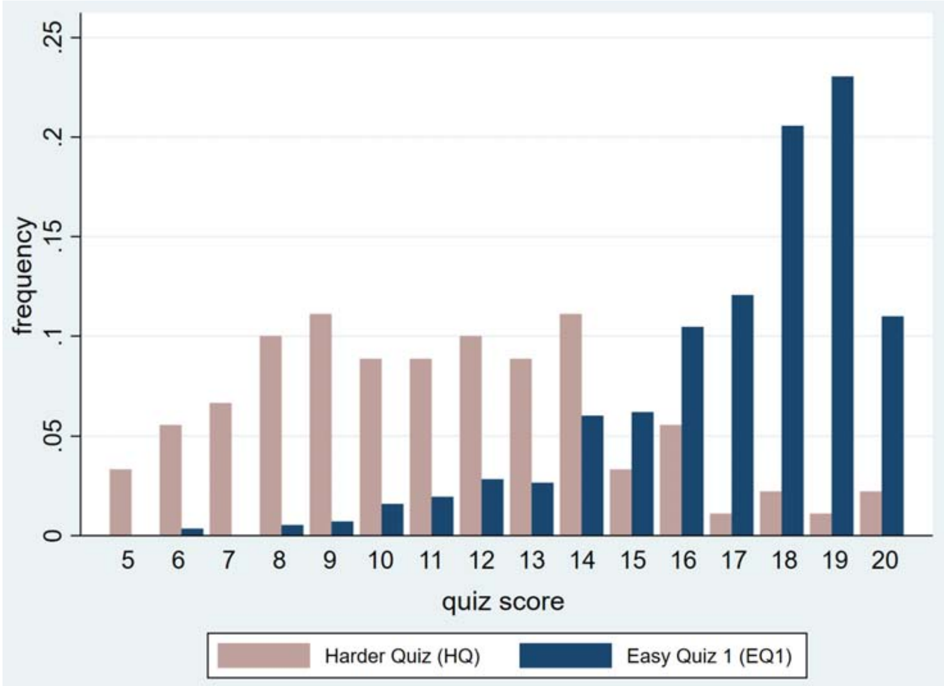
Table 2 in Section IIIB Comparing the belief gap among voters against 0: $p < 0.001$ in *BL* and $p = 0.028$ in *HQ*, *SV*, and *TH*.

Paragraph 3 in Section IIIB Comparing beliefs of subjects in the top 1/3 to beliefs of subjects in the bottom 2/3 in *BL*, *HQ*, and *SV*: $p < 0.001$.

Paragraph 3 in Section IIIB Comparing beliefs of subjects in the top half to beliefs of subjects in the bottom half *TH*: $p = 0.028$.

Paragraph 4 in Section IIIB Comparing mean $p_i(q_j|vote_j = 1)$ to the share of voters in the top 1/3 in the last round in treatments with $x = 3$ and no aggregated feedback: $p < 0.001$.

Figure 5: Quiz Score Distribution: *HQ* vs Easy Quiz *EQ1*



Notes: The graph shows the frequency distributions of *Harder Quiz* scores (90 subjects) and *EQ1* scores (564 subjects). See Figure 6 for *EQ2* scores (30 subjects).

Footnote 38 in Section IIIB Gender difference in belief gap among voters: $p = 0.067$ in *BL*, $p = 0.345$ in *HQ*, $p = 173$ in *SV*, and $p = 0.249$ in *TH*.

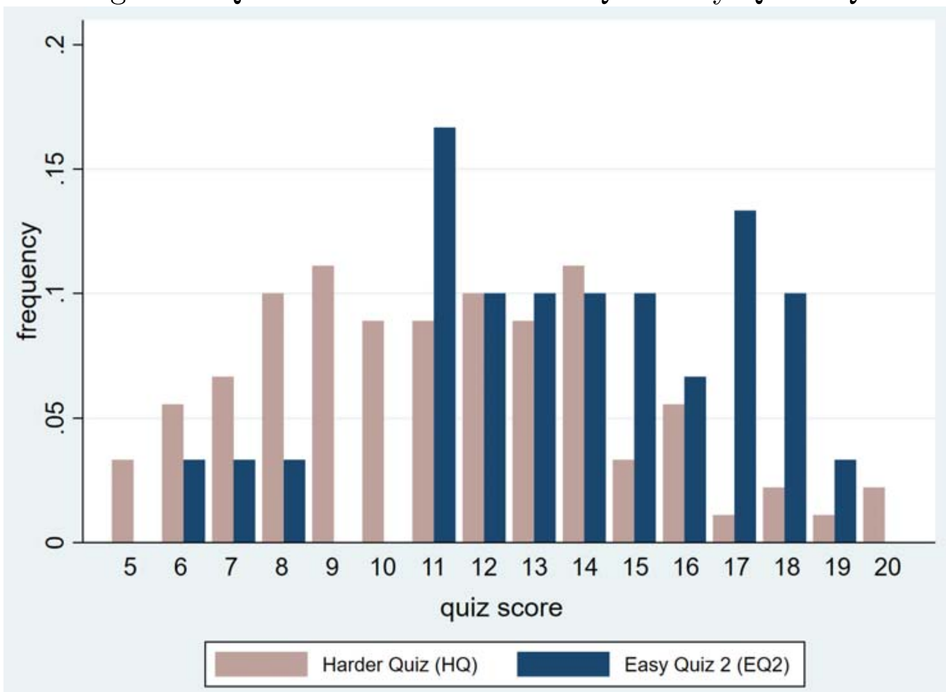
B.4 Quiz Scores

This section presents the frequency distributions of quiz scores. As discussed in the main text, *HQ* significantly increases quiz difficulty. We use *EQ1* (i.e., the original *Baseline* quiz) in *SV* and *TH*. As mentioned in Online Appendix B.1, *EQ1* is very easy, and two *BL 15* sessions have a slightly more difficult quiz, *EQ2*, as a robustness check. As a result, a total of 564 subjects took *EQ1*, 30 subjects took *EQ2*, and 90 subjects took the harder quiz in *HQ*.

Figure 5 presents jointly the frequency distribution of the quiz scores in *HQ* and *EQ1*. We observe a very striking contrast in the score distributions. In particular, there is a substantial share of very low scorers in *HQ* and a substantial share of very high scorers in *EQ1*.

Figure 6 also exhibits a prominent shift in the quiz score distribution across *HQ* and *EQ2*. This is particularly transparent at the lower and upper tails of the distributions.

Figure 6: Quiz Score Distribution: *HQ* vs Easy Quiz *EQ2*



Note: The graph shows the frequency distributions of *Harder Quiz* scores (90 subjects) and *EQ2* scores (30 subjects).

Almost 46% of subjects in *HQ* have 10 or fewer correct answers, but this is the case for only 10% of subjects who took *EQ2*. One third of the subjects who took *EQ2* have 16 or more correct answers, whereas the same holds for 12.2% of subjects in *HQ*.

Finally, we note that even in the harder quiz of *HQ*, every subject has at least 5 correct answers. This is not surprising given that every quiz has a multiple-choice format, which results in five correct answers in expectation if *all* answers are random choices. See Online Appendix C for the quiz format.

B.5 Regression Analysis: Overconfidence and Efficiency in *TH* and *SV*

As mentioned in the main text in Section IIIA, there is substantial variation in the proportion of correct group decisions across sessions of *HQ*, and this variation is nicely explained by the sizable differences in overconfidence at the session level. We run two ordered probit regressions in which the dependent variable is the group decision in a round in *HQ*. This variable takes one of three values, representing a wrong decision, a tie, and a correct decision, respectively. The main explanatory variable is overconfidence, which we proxy by mean $p_i(q_i)$

Table 7: Increase in overconfidence reduces efficiency in HQ and SV

		HQ			SV	
	(1)	(2)	(3)	(4)	(5)	(6)
	Wrong	Tie	Correct	Wrong	Tie	Correct
mean $p_i(q_i)$ in	1.631	-0.200	-1.431	3.943	-1.096	-2.847
bottom 2/3	(0.348)	(0.181)	(0.461)	(1.872)	(0.687)	(1.411)
Round#	0.006	-0.001	-0.005	-0.003	0.001	0.002
	(0.027)	(0.003)	(0.023)	(0.014)	(0.004)	(0.010)
belief gap	2.119	-0.361	-1.758	3.783	-1.221	-2.562
among voters	(0.330)	(0.189)	(0.474)	(0.685)	(0.429)	(0.484)
Round#	0.022	-0.004	-0.018	-0.021	0.007	0.014
	(0.020)	(0.005)	(0.016)	(0.018)	(0.006)	(0.012)
Observations	36	36	36	36	36	36

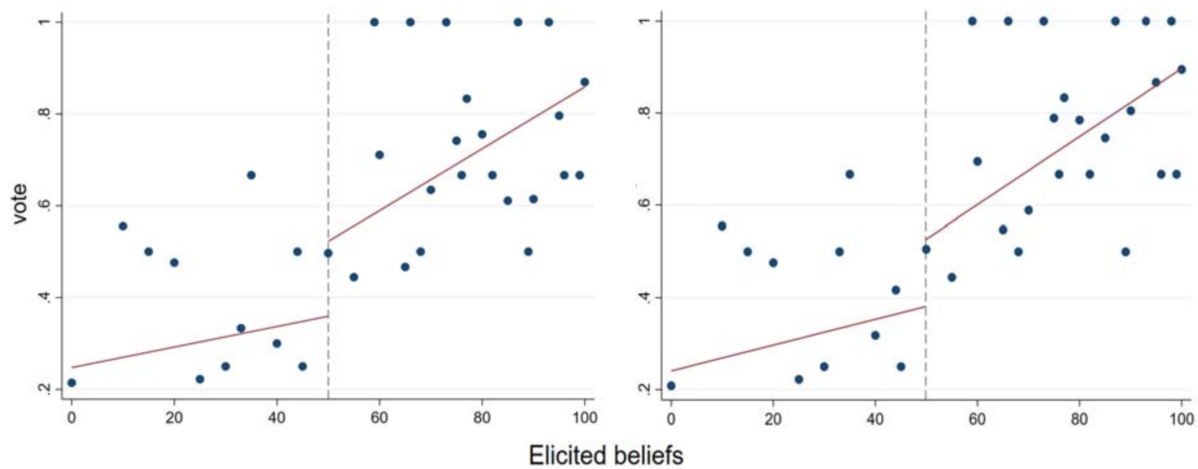
Notes: Coefficients show the average marginal effects in ordered probit regressions. The dependent variable is the group decision in a round and takes one of three values, representing a wrong decision, a tie, and a correct decision. Standard errors clustered by session are in parentheses.

in the bottom 2/3 in one regression and by the belief gap among voters in the other (see Section IIIB for the definition and discussion of the belief gap among voters). We also include a time trend. Errors are clustered at the session level. Table 7 presents the average marginal effects. The table shows that an increase in overconfidence is associated with a decrease in the probability of making the correct decision and an increase in the probability of making the wrong decision. For example, column (1) shows that a one-percent increase in mean $p_i(q_i)$ among the bottom 2/3 increases the probability of making the wrong decision by 1.6%, and a one-percent increase in the belief gap among voters increases the probability of making the wrong decision by 2.1%

In columns (4)-(6) of Table 7, we repeat the two regressions described above using SV data. When the belief gap among voters is the independent variable that proxies overconfidence, we rule out subjects who vote against s_i in the computation of the belief gap consistent with the analysis in the main text. Columns (4)-(6) show that an increase in overconfidence is associated with a decrease in the probability of making the correct decision and an increase in the probability of making the wrong decision. For example, column (4) shows that a one-percent increase in mean $p_i(q_i)$ among the bottom 2/3 increases the probability of making the wrong decision by 3.9%, and a one-percent increase in the belief gap among

Figure 7: Binscatter of Elicited Beliefs and Treatment Turnout Rate with a Discontinuity at 0.5

(a) Subjects below the top $1/x$ (b) All Data



Note: The panels illustrate the relationship between elicited beliefs and turnout rate with a discontinuity at 50%.

voters increases the probability of making the wrong decision by 3.8%.

B.6 Relationship between Elicited Beliefs and Turnout Rate with Discontinuity at 50%

The two panels in Figure 7 present binscatters of the relationship between elicited beliefs and turnout in the treatments with a discontinuity at 50% (i.e., $p_i(q_i) = 0.5$). The left panel plots the relationship using the data of only subjects below the top $1/x$, whereas the right panel uses all data. As mentioned in the main text, these plots suggest that 50% is empirically relevant for voting and abstention choices as a heuristic cutoff. They also replicate the monotone pattern observed in Figure 3.

B.7 Regression Analysis: Behavior of Low-confidence Subjects

We estimate a random effects panel probit model in which the dependent variable is the voting decision of subjects with $p_i(q_i) < 0.5$ in the treatments. One of the independent variables is the mistake rate in *OBJ*; i.e., turnout rate conditional on $q_i < 0.5$ ("mistake rate" in Table 8). The remaining independent variables are the same as in Table 3 in the main text: (i) placement in the top $1/x$ ("top $1/x$"); (ii) the elicited belief regarding own placement in the top $1/x$ ("$p_i(q_i)$"); (iii) the elicited belief regarding other group members'

Table 8: Explaining Individual Turnout Decision when $p_i(q_i) < 0.5$

	(1) All Rounds	(2) All Rounds	(3) Final round	(4) Final round
mistake rate	0.560 (0.144)	0.684 (0.096)	0.870 (0.168)	1.179 (0.161)
$p_i(q_i)$	0.343 (0.137)	0.291 (0.129)	0.243 (0.207)	0.130 (0.206)
$p_i(vote_j)$	0.261 (0.082)	0.291 (0.088)	0.469 (0.144)	0.339 (0.147)
$top\ 1/x$	0.148 (0.097)	0.162 (0.090)	0.145 (0.175)	0.157 (0.158)
$p_i(q_j\|vote_j = 1)$	-	-	0.146 (0.139)	0.073 (0.133)
Round#	0.002 (0.008)	0.002 (0.008)	-	-
Controls	No	Yes	No	Yes
Observations	492	492	82	82

Notes: Coefficients show the average marginal effects in random effects panel probit regressions described in the text. The dependent variable is the subject's binary choice between voting (= 1) and abstaining/voting against s_i in SV (= 0). Standard errors clustered by session are in parentheses.

likelihood of voting ("$p_i(vote_j = 1)$"); and (iv) a time trend. We cluster errors at the session level.

Table 8 presents the average marginal effects. Comparing specifications (1) and (2), we see that including dummy variables for treatment variations as regressors makes no difference regarding the effect of "mistake rate": in either case, its effect is economically large and statistically strongly significant ($p < 0.001$). We also run probit regressions using the turnout data of the final round with the variable "$p_i(q_j|vote_j = 1)$" as an additional regressor. Once again, the effect of the mistake rate in *OBJ* is economically large and statistically strongly significant with and without treatment controls.

B.8 Multiple Testing

The core elements of our experimental analysis are six pairwise treatment comparisons, i.e., the efficiency comparison of *BL* with *HQ*, *SV*, and *TH* as well as the comparison of *BL* and *HQ* in terms of quiz scores, mean $p_i(q_i)$, and turnout. Table 9 presents the original (one-sided) p-values, reported in the main text, as well as their adjusted values using Bonferroni, Holm, and Benjamini and Hochberg methods. Our main conclusions remain unchanged in each method.

Table 9: Adjusted p-values in Pairwise Treatment Comparisons

	(1) p-value	(2) Bonferroni	(3) Holm	(4) Benjamini & Hochberg
BL vs TH: efficiency	<0.001	<0.001	<0.001	<0.001
BL vs SV: efficiency	0.133	0.798	0.133	0.133
BL vs HQ: efficiency	0.004	0.022	0.011	0.006
BL vs HQ: quiz score	<0.001	<0.001	<0.001	<0.001
BL vs HQ: mean $p_i(q_i)$	<0.001	0.002	0.002	<0.001
BL vs HQ: turnout	0.008	0.049	0.016	0.010

Notes: The Bonferroni method multiplies the values in (1) by 6. The Holm method and The Benjamini and Hochberg method order the values in (1) from lowest to highest. The Holm method multiplies the ith lowest value with $7 - i$. The Benjamini and Hochberg multiplies the ith lowest value with $\frac{6}{i}$. The Bonferroni and Holm methods control the family-wise error rate, whereas the Benjamini and Hochberg method controls the false discovery rate.

B.9 Efficiency Benchmark in Treatment SV

Consider the heuristic strategy described for SV in the main text: every i votes for s_i if $p_i(q_i) \geq 0.5$ and against s_i otherwise. We will show that the probability that an unbiased subject who uses this strategy casts a correct vote is higher than 0.66 with every q distribution $F[\underline{q}, \bar{q}]$ that satisfies $\mathbb{E}(q) = 1/3$. We assume that the distribution of q is continuous, which is without loss of generality. If there is an atom at $q = 0.5$, this is inconsequential for the steps below. First assume that $\bar{q} \geq 0.5$. The probability that an unbiased subject who uses the above described strategy casts a correct vote is given by

$$\int_{\underline{q}}^{0.5}(1-q)dF + \int_{0.5}^{\bar{q}} q dF.$$

Since $\int_{\underline{q}}^{0.5}(1-q)dF + \int_{0.5}^{\bar{q}} q dF \geq \int_{\underline{q}}^{0.5}(1-q)dF + \int_{0.5}^{\bar{q}}(1-q)dF = 1 - \mathbb{E}(q) = 2/3$, the desired result follows. Next, assume that $\bar{q} < 0.5$. In that case, the probability that an unbiased subject who uses the above described strategy casts a correct vote is given by $\int_{\underline{q}}^{\bar{q}}(1-q)dF = 1 - \mathbb{E}(q) = 2/3$.

C Instructions

Instructions consist of five parts: four of them are paper based, and one part is computer based. Paper-based instructions are identical in *BL*, *BL 15* and *HQ* except the group size N. Paper-based instructions in *SV* are almost identical to those in *BL*, *BL 15*, and *HQ* except that (i) it is explained to subjects that they can vote for red or blue (regardless of their signal); and (ii) they are asked about the probability that a randomly selected (other) voter votes correctly.

We present the instructions used in *BL* with $N = 24$ below. After the first-paper based part, subjects receive computer-based instructions for the quiz and the subsequent belief elicitation task regarding their quiz performance. We transcribe these computerized instructions and present the screenshots of the quiz (*EQ1*). We reduce the original font size and spacing to conserve space. The final page shows the screenshots of the harder quiz in *HQ*.

PART 1

Welcome to the experiment! Please turn off your cell phones and do not communicate with other subjects during the experiment. You will be paid for your participation in this experiment. The amount of money you earn depends on your decisions and decisions of other participants. Your decisions will be treated anonymously. The money you earn will be paid to you in cash at the end of the experiment.

This experiment consists of several parts. We explain the details of Part 1 now, the details of other parts will be explained in a short time.

Part 1 involves a "**Guessing Task**". In this task, you will be presented with various statements. Consider for example the statement "it snowed in Amsterdam in April 1991". We know whether or not this statement is true but you may not know it for certain. We will ask you to report your "best guess" about the chances that such a statement is true.

You will report your guess by choosing a percentage between 0 and 100. The percentage that you choose indicates your "best guess". **The higher the accuracy of your best guess, the higher the payoff you get.** In order to maximize your payoff:

- If you are certain a statement is true then you should choose a percentage of 100, and if you are certain a statement is false you should choose 0.
- In many cases you do not know for certain whether a statement is true or not. If you think the statement is equally likely to be true or false, you should choose a percentage of 50.
- **More generally, the more confident you are a statement is "true" the higher the percentage you should assign**. If for example you are *very confident* a statement is true, you should choose a percentage close to 100.
- **Conversely, the more confident you are a statement is "false" the lower the percentage you should assign**. If for example you are *very confident* a statement is false, you should choose a percentage close to 0.

You will make guesses regarding 6 statements in total, and 2 statements will be randomly selected in order to determine your payoff in this part. You will earn an amount from €0 to €1 in each selected statement depending on the accuracy of your guess.

You will earn the most if you "honestly report your best guess" about the chances that a statement is true because your payoff increases in the accuracy of your guess. (Your exact payoff for each selected statement is calculated as follows. Suppose you assign a percentage of P to the statement being true. If the statement is **true** your payoff equals $1\text{-}(1\text{-}P\%)^2$, and your payoff equals $1\text{-}(P\%)^2$ if the statement is **false.**)

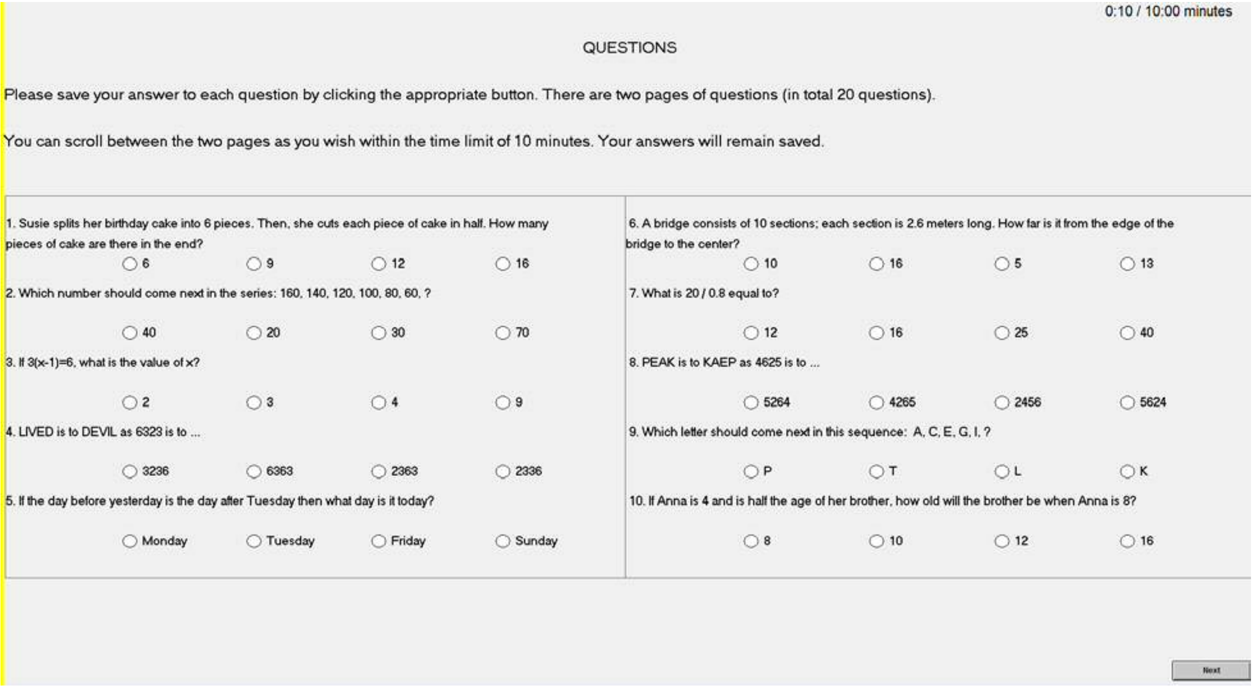
If you have any questions or need assistance of any kind please raise your hand and an experimenter will come to you. Please click OK on your screen when you are ready to start the experiment.

[*Part 1 was followed by the quiz and the subsequent belief elicitation task regarding quiz performance. The instructions for this part were only computerized. We transcribe the instruction screens below and present the screenshots of the quiz.*

Screen before the beginning of the quiz] In this part of the experiment, you will be taking a QUIZ on math and logic puzzles taken from various tests. The quiz involves **20 questions**. You will have **10 minutes** to correctly answer as many questions as you can. You will be paid **30 cents** for each correct answer. Your quiz score will also be relevant for later parts of the experiment. We will explain soon how exactly it will be relevant. Please click OK to continue.

[*Following screen before the beginning of the quiz*] You will next see the first page of the quiz. Please click the appropriate button to record your answer to a quiz question. When you want to see the second page of the quiz, click NEXT to continue. By clicking BACK on the second page you can go back to the first page of the quiz. You can go back and forth between the two pages as you wish within the time limit of 10 minutes. Answers that you have given will always remain saved when you move between the two pages. Please click START when you are ready to start the quiz.

[*Screenshots of the quiz: First screen*]



0:10 / 10:00 minutes

QUESTIONS

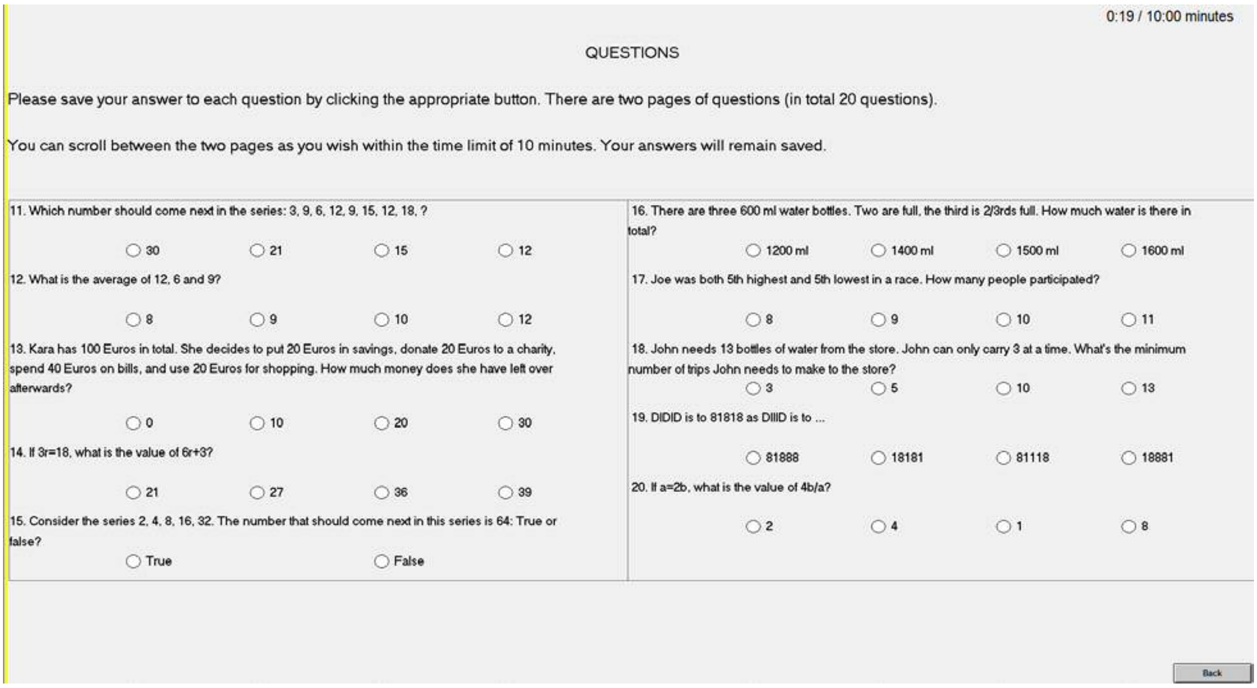
Please save your answer to each question by clicking the appropriate button. There are two pages of questions (in total 20 questions).

You can scroll between the two pages as you wish within the time limit of 10 minutes. Your answers will remain saved.

1. Susie splits her birthday cake into 6 pieces. Then, she cuts each piece of cake in half. How many pieces of cake are there in the end?
○ 6 ○ 9 ○ 12 ○ 16
2. Which number should come next in the series: 160, 140, 120, 100, 80, 60, ?
○ 40 ○ 20 ○ 30 ○ 70
3. If 3(x-1)=6, what is the value of x?
○ 2 ○ 3 ○ 4 ○ 9
4. LIVED is to DEVIL as 6323 is to ...
○ 3236 ○ 6363 ○ 2363 ○ 2336
5. If the day before yesterday is the day after Tuesday then what day is it today?
○ Monday ○ Tuesday ○ Friday ○ Sunday
6. A bridge consists of 10 sections; each section is 2.6 meters long. How far is it from the edge of the bridge to the center?
○ 10 ○ 16 ○ 5 ○ 13
7. What is 20 / 0.8 equal to?
○ 12 ○ 16 ○ 25 ○ 40
8. PEAK is to KAEP as 4625 is to ...
○ 5264 ○ 4265 ○ 2456 ○ 5624
9. Which letter should come next in this sequence: A, C, E, G, I, ?
○ P ○ T ○ L ○ K
10. If Anna is 4 and is half the age of her brother, how old will the brother be when Anna is 8?
○ 8 ○ 10 ○ 12 ○ 16

Next

[*Second quiz screen*]



0:19 / 10:00 minutes

QUESTIONS

Please save your answer to each question by clicking the appropriate button. There are two pages of questions (in total 20 questions).

You can scroll between the two pages as you wish within the time limit of 10 minutes. Your answers will remain saved.

11. Which number should come next in the series: 3, 9, 6, 12, 9, 15, 12, 18, ?
○ 30 ○ 21 ○ 15 ○ 12
12. What is the average of 12, 6 and 9?
○ 8 ○ 9 ○ 10 ○ 12
13. Kara has 100 Euros in total. She decides to put 20 Euros in savings, donate 20 Euros to a charity, spend 40 Euros on bills, and use 20 Euros for shopping. How much money does she have left over afterwards?
○ 0 ○ 10 ○ 20 ○ 30
14. If 3r=18, what is the value of 6r+3?
○ 21 ○ 27 ○ 36 ○ 39
15. Consider the series 2, 4, 8, 16, 32. The number that should come next in this series is 64: True or false?
○ True ○ False
16. There are three 600 ml water bottles. Two are full, the third is 2/3rds full. How much water is there in total?
○ 1200 ml ○ 1400 ml ○ 1500 ml ○ 1600 ml
17. Joe was both 5th highest and 5th lowest in a race. How many people participated?
○ 8 ○ 9 ○ 10 ○ 11
18. John needs 13 bottles of water from the store. John can only carry 3 at a time. What's the minimum number of trips John needs to make to the store?
○ 3 ○ 5 ○ 10 ○ 13
19. DIDID is to 81818 as DIIID is to ...
○ 81888 ○ 18181 ○ 81118 ○ 18881
20. If a=2b, what is the value of 4b/a?
○ 2 ○ 4 ○ 1 ○ 8

Back

[*Belief elicitation screen after the quiz ends*]

PLEASE READ CAREFULLY

We have now obtained the quiz scores of the 24 participants in this room, and ranked the participant scores from highest to lowest.

IMPORTANT: The quiz refers to the 20 math and logic puzzles that you have just answered, NOT the guessing tasks at the beginning of the experiment!

Your quiz score ranks in the **TOP 1/3 if at most 7 participants scored better than you in the quiz**. Exactly 8 out of 24 participants are in the top 1/3. For example, if you and another participant have the same quiz score and tie for the 8th place, then the tie is broken fairly and each of you is selected to be in the top 1/3 with equal chance.

We will now ask you to indicate your best guess about a statement regarding your score ranking. You will earn an amount from €0 to €3 depending on the accuracy of your guess. As before, you will indicate your guess choosing a percentage between 0 and 100, and as before, the higher the accuracy of your guess the higher the payoff you get; so you earn the most when you honestly report your best guess.

Please indicate your best guess about the statement below. What are the chances that it is true?

"My quiz score ranks in the top 1/3."

Please enter a percentage from 0 to 100 to indicate your best guess.

PART 2

*Part 2 and the following parts are on **group decision making**. From now on, you will be making choices in a **group**. In each round, the computer will randomly pick **RED** or **BLUE** as "**your group color**." You will not learn the color until the end of the round. **Your task as a group** is to try to **guess your group color correctly—based on information group members may receive**.*

Here is a detailed description of Part 2:

In each round, you will make choices in a **group of 24** (including you).

In each round, the computer will randomly pick either **RED** or **BLUE** as "**your group color**." There is a 50% chance **RED** will be picked and a 50% chance **BLUE** will be picked. In other words, **RED** and **BLUE** are *equally likely* to be your group color.

You will not learn your group color until the end of the round. Your task as a group is to try to guess the group color correctly. The group decision will be made by **voting**.

Before voting, each member of your group will be shown a "card", which may give information regarding your group color.

You will see only "**Your own Card"**. Similarly, each group member sees only his/her own card.

Each card is either red or blue. After you are shown **Your Card**, you will choose between voting for the color of Your Card and abstaining. In other words:

- If you are shown a **red** card then you choose between **voting for red** and **abstaining**.
- If you are shown a **blue** card then you choose between **voting for blue** and **abstaining.**

Cards are of two types: **informative** and **misleading**. A card is "**informative"** if its color is the **same** as your group color and it is "**misleading"** if it has the **opposite color**. In each round, Your Card is either informative or misleading.

*Since an "informative card" has the **same** color as your group, voting for the color of an informative card will result in a **CORRECT VOTE**. Since a "misleading card" has the **opposite** color, voting for the color of a misleading card will result in an **INCORRECT VOTE**.*

To repeat, Your Card is either **informative** or **misleading**. However, you **will NOT know** for certain whether Your Card is informative or misleading. This will be determined by **CHANCE** in each round. To be more precise, in each round, Your Card is

- an informative card with **X%** chance and a misleading card with **(100-X)%** chance.
- You will learn your X value before making your voting decision.
- At the beginning of every round, you will have a new X value that is randomly drawn from {1,2,3,…,99,100} by the computer. All possible values of X are equally likely.

Notice that:

- The closer X is to 100, the higher the chances that you have an informative card and observe your group's true color.
- The closer X is to 0, the higher the chances that you have a misleading card and observe the opposite color.

Here is an example: If your X value is exactly 50, then you are **equally likely to get an informative card as a misleading card.**

Another example: If your X value is 25, then you are **three times more likely to get a misleading card than an informative card,** and conversely, you are **three times more likely to get an informative card than a misleading card** if your X value is 75**.**

What about other members of your group? The rules for other members of your group are exactly the same as for you. Every member has his/her own X value that is randomly drawn from {1,2,3,…,99,100} by the computer. Every member observes his/her own Card, which is informative or misleading depending on the member's own X value. Note that you will NOT learn the X value or the card color of any other group member.

To summarize so far: After you learn your X value and the color of Your Card, you will choose between voting for the color of Your Card and abstaining. The same is true for every member of your group.

The color that receives a majority of the votes is the "group decision" and ties are broken fairly.[1] The group decision is "correct" if it is the same as your group color. You will earn **€4** if the group decision is correct and **€0** otherwise.

Reminder: When you observe Your Card's color, you will NOT know for certain whether or not Your Card is an informative card. However, you will know your X value, representing the chance with which Your Card is informative. Therefore, in your decision whether or not to vote, it is **important to weigh potential gains against potential losses GIVEN YOUR X VALUE.**

i. The more likely you are to have an informative card, the more likely you are to cast a correct vote and therefore the group decision is more likely to be correct if you vote. Hence, the higher the X value, the higher the **potential gains** from voting.
ii. Conversely, the more likely you are to have a misleading card, the more likely you are to cast a wrong vote and thus the group decision is more likely to be wrong if you vote. Hence, the lower the X value, the higher the **potential losses** from voting.

Therefore, if your X value is not sufficiently high, then **the potential loss** due to your voting is **higher** than **the potential gain** from your vote.

However, exactly which values of X allow your vote to generate higher potential gains than losses will depend on the behavior of other group members.

Because the precise value of X where the potential gains from your vote start dominating the potential losses depends on the voting behavior of other group members, we will also ask you to report your

[1] If Red and Blue receive the same number of votes then we will pick Red with 50% chance and Blue with 50% chance to determine the "group decision".

best guess about the chances that a randomly selected group member (other than you) chose to vote in each round.

As in the previous part, you will indicate your guess choosing a percentage between 0 and 100. As before, you earn the most if you honestly report your best guess. You will earn an amount from **€0** to **€1** for a guessing task depending on the accuracy of your guess.

You will play a total of **15 rounds** in this part and **2 rounds** will be randomly selected for payment. The amount you earn from the group decision and the guessing task in each of the selected rounds will be added to determine your payoff in this part. **Since these 2 rounds will be randomly selected, you should treat each round as a round you could be paid for.**

SUMMARY

In each round, you will make choices in a group of 24.

The computer will randomly pick **RED** or **BLUE** as your group color. Your task as a group is to try to guess your group color correctly. The group decision will be made by **voting**.

Each member of your group will be privately shown a "card". Each card is either red or blue.

You will only see the color of "**Your own Card"**.

Cards are of two types: **informative** and **misleading**. An "**informative"** card has the **same** color as your group, a "**misleading"** card has the **opposite color**.

You will NOT know for certain whether Your Card is informative or misleading. However, you will know your **X** value, representing the chance with which Your Card is informative.

After you observe the color of Your Card, you will choose between voting for the color of Your Card and abstaining.

The rules for other members of your group are *exactly the same as for you.*

The color that receives a majority of the votes is the "group decision." The group decision is "correct" if it is the same as your group color.

You will play a total of 15 rounds. 2 rounds will be randomly selected and the amount you earn from the group decision and the guessing task in the selected rounds will be added to determine your payoff in this part. Thus, you should treat each round as a round you could be paid for.

***

Please raise your hand if you have any questions. Please click OK when you are ready to start this part.

PART 3

This part is similar to Part 2. In Part 2, you did **NOT** know whether Your Card is informative or misleading with certainty as it was determined by **chance**. In this part, Your Card will be determined by **your quiz score** from Part 1, instead of being determined by chance.

You completed a quiz on math and logic puzzles in Part 1. As explained before, we ranked the quiz scores of the 24 participants in this room from highest to lowest. We know whether or not your quiz score is in the top 1/3 but you do **NOT** know it for certain.[2] In this part, Your Card will depend on your score as follows:

(1) If your quiz score is **in the top 1/3** then Your Card is an **informative card**

---

[2]Recall that exactly 8 out of 24 participants are in the top 1/3. If you and another participant tie for the 8th place, then each of you is selected to be in top 1/3 with equal chance.

(2) If your quiz score is **below the top 1/3** then Your Card is a **misleading card**[3]

After you took the quiz, we asked you to report your "best guess" about the chances that your quiz score ranks in the top 1/3—we will soon remind you of your guess. However, note that Your Card depends only on your true ranking, *not* on your guess.

What about other members of your group?

The rules for other members of your group are *exactly the same as for you*.

*To repeat, whether Your Card is informative or misleading depends on your quiz score. Thus, your belief regarding the chances that your quiz score ranks in the top 1/3 is analogous to your **X** value in Part 2.*

The group-decision making is the same as before. After you observe the color of Your Card, you will choose between voting for the color of Your Card and abstaining. The same is true for every member of your group. The color that receives a majority of the votes is the "group decision" and ties are broken randomly. The group decision is "correct" if it is the same as your group color. You will earn **€5** if the group decision is correct and **€0** otherwise.

Additionally, you will earn money in a Guessing Task just as in Part 2. In each round, you will report your best guess about the chances that a randomly selected member of your group (other than you) voted in that round. As before, you will indicate your guess choosing a percentage between 0 and 100. You can earn an amount from **€0** to **€1** in a guessing task depending on the accuracy of your guess.

You will play a total of **5 rounds** in this part. **1 round** will be randomly selected and the amount you earn from the group decision and the guessing task in that round will determine your payoff in this part.

Please raise your hand if you have any questions. Please click OK when you are ready to start this part.

PART 4

This is the final part of the experiment. This part is exactly the same as Part 3 **except that you will now also make a guess regarding the "competence of the average VOTER" in the room (other than you).** You will see the following information on your computer screen.

We have now randomly picked one member (other than you) that chose to **VOTE** in this round. Please indicate your best guess about the statement below: What are the chances it is true?

"The quiz score of this randomly selected VOTER is in the top 1/3."

You will indicate your guess choosing a percentage between 0 and 100, as before. As before, you will also make a guess about the chances that a randomly selected member of your group (other than you) voted. As before, you will earn an amount from **€0** to **€1** in each guessing task depending on the accuracy of your guess.

You will play 1 **round** in this part. You will earn **€5** if the group decision is correct and **€0** otherwise. Please raise your hand if you have any questions. Please click OK to start.

[3] Recall that an informative card has the same color as your group, and a misleading card has the opposite color.