Supplementary Appendix for "Informative Cheap Talk in Elections"

Navin Kartik[*] Richard Van Weelden[†]

October 19, 2017

Abstract

This Supplementary Appendix formalizes two extensions of the baseline model discussed in Section 5 of the main text: we allow candidates to have some private information about the state of the world when campaigning (Supplementary Appendix A); and we consider more than two policy-preference types and actions (Supplementary Appendix B).

A. Pre-election Private Information about the State

A.1. Model and result

In our baseline model, candidates have no private information about the policy-relevant state of the world when campaigning. We establish here that the nature of informative cheap talk extends when there is a limited amount of such private information.

Specifically, we extend the baseline model so that each candidate $i \in \{A, B\}$, in addition to privately knowing his policy-preference type, $\theta_i \in \{0, b\}$, observes a private signal $\beta_i \in [0, 1]$ about the state of the world, s. As before, the elected policymaker (PM) learns the state s after taking office. In any state s the candidates' signals, β_A and β_B, are drawn independently from a cumulative distribution $\Psi(\cdot, s)$. We assume that for all s, $\Psi(\cdot, s)$ admits a strictly positive and continuously differentiable density on $[0, 1]$ and that, if $s' > s$, then $\Psi(\beta_i, s) > \Psi(\beta_i, s')$ for all $\beta_i \in (0, 1)$. Hence a higher signal β_i indicates that the state s is likely to be higher. We define

[*]Department of Economics, Columbia University. Email: nkartik@columbia.edu.

[†]Department of Economics, University of Pittsburgh. Email: rmv22@pitt.edu.

$F(\cdot, \boldsymbol{\beta})$ and $f(\cdot, \boldsymbol{\beta})$ to be the posterior cumulative distribution and density, respectively, on the state s given signal $\boldsymbol{\beta} \equiv (\beta_A, \beta_B) \in [0,1]^2$. Our assumptions on Ψ ensure that for any $\boldsymbol{\beta}$, $F(\cdot, \boldsymbol{\beta})$ and $f(\cdot, \boldsymbol{\beta})$ are continuously differentiable; furthermore, if $\boldsymbol{\beta}' > \boldsymbol{\beta}$ (i.e., $\beta'_A \geq \beta_A$ and $\beta'_B \geq \beta_B$, with at least one strict inequality), then the conditional distribution of s after $\boldsymbol{\beta}'$ first-order stochastically dominates the distribution after $\boldsymbol{\beta}$.

To make precise the sense in which candidates have limited private information about the state, we adopt the following metric over (a class of) cumulative distributions of the state. For any distributions $F_1(\cdot)$ and $F_2(\cdot)$ that admit continuous densities $f_1(\cdot)$ and $f_2(\cdot)$ on $(\underline{s}, \infty)$, define

$$d(F_1(\cdot), F_2(\cdot)) := \sup_{s \in [\underline{s}, \infty)} \max\{|F_1(s) - F_2(s)|, |f_1(s) - f_2(s)|\}. \tag{A.1}$$

In words, the distance between two distributions at any given state is the maximum of the distance between the distributions and their densities at that state. The distance between the distributions is the supremum of their distance across all states. Note that in the above metric, if $\Psi(\beta_i, \cdot)$ approaches a constant function for all β_i, then $F(\cdot, \boldsymbol{\beta})$ approaches the prior $F(\cdot)$ for all $\boldsymbol{\beta}$.

As in the baseline model we allow candidates to simultaneously send one of two messages, $m_i \in \{0, b\}$. We restrict attention to a binary message space because we seek to establish that the same kind of equilibria as in the baseline model exist in this extension too. Our goal is not to characterize all equilibria with a general message space, which we believe is infeasible when there is rich pre-electoral private information about the state. The equilibria we construct with a binary message space would, however, continue to be equilibria even if candidates' had other messages available. We return to the issue of other equilibria and more general message spaces at the end of this subsection.

We focus on symmetric equilibria and seek an equilibrium in which congruent politicians always announce congruence, $m_i = 0$, but some non-congruent ones claim congruence as well. In such an equilibrium, $m_i = b$ reveals a candidate to be non-congruent, just as in the baseline model. Hence the elected PM's action does not alter the voter's belief if he announced $m_i = b$; but the voter will updated positively (resp., negatively) if the PM takes the low action $\underline{a}$ (resp., high action $\overline{a}$) after announcing $m_i = 0$. In such an equilibrium, a biased politician has a larger incentive to reveal his bias when he has observed a high β_i, because he believes he is more likely to lose reputation in office. We will identify an interior cutoff $\overline{\beta}$ such that a biased type chooses $m_i = b$ if and only if $\beta_i > \overline{\beta}$. A candidate's message thus reveals information both about his policy type and his signal about the state.[1] It is optimal for a congruent politician to always announce he is congruent if, regardless of his signal about the state, he is more likely to take the low action in office than a biased politician who receives the threshold signal $\overline{\beta}$. We will confirm that this limited single-crossing condition holds when the politician's signal of the state is not too informative.

The final piece of the equilibrium construction is that the voter must be indifferent between electing either candidate regardless of their messages. The behaviors supporting this equilibrium are analogous to those in the semi-separating equilibrium of the baseline model, with the caveat that the voter's belief about the state is now affected by the campaign messages. In particular, the voter must be indifferent between an uncertain candidate i and a known, biased candidate $-i$ (i.e., i's opponent) conditional on her belief about the state and anticipated behavior from the PM following the campaign announcements $m_i = 0$ and $m_{-i} = b$.

In reading the following result, recall that k^* and $p^*(k)$ were defined in Section 4.1 of the main text, and that $F(\cdot)$ is the prior distribution of the state.

Proposition A.1. *Assume $k > k^*$ and $p \in (0, p^*(k))$. There exists $\varepsilon(p,k) > 0$ such that, if $d(F(\cdot, \boldsymbol{\beta}), F(\cdot)) < \varepsilon(p,k)$ for all $\boldsymbol{\beta} \in [0,1]^2$, then there exists $\overline{\beta} \in (0,1)$ such that it is an equilibrium for each candidate $i \in \{A, B\}$ to announce $m_i = b$ if and only if $\theta_i = b$ and $\beta_i > \overline{\beta}$, and to announce $m_i = 0$ otherwise.*

(The proof is in Supplementary Appendix A.2.)

Proposition A.1 shows that a result similar to Proposition 3 of the main text holds when the candidates have some private information about the state of the world. Note that, as in the baseline model, message $m_i = b$ reveals both that the candidate is biased and that he is more likely to take the high action if elected.

The equilibrium construction here has some notable differences from the baseline model. First, the equilibrium is now in pure strategies; introducing private signals about the state "purifies" the mixing by biased candidates in the baseline model. Second, unlike in the baseline model, candidates' messages now reveal some information about the state of the world. As the voter updates about the state based on *both* candidates' messages, and the voter's belief influences the elected PM's pandering (so long as he was elected having claimed congruence, i.e., having announced $m_i = 0$), both candidates' messages influence the degree of

[1] Although we describe a message of $m_i = 0$ (resp., $m_i = b$) as a claim to be congruent (resp. biased), we could just as well interpret a message $m_i = 0$ (resp., $m_i = b$) as a claim that the state is likely to be low (resp., high) or that the candidate intends to take the low (resp., high) action.

post-electoral pandering even through the elected PM learns the state perfectly. In particular, the message of the losing candidate will affect the action taken by the elected PM, with the PM more likely to take the high action when his losing opponent revealed himself as biased.

Proposition A.1 characterizes one equilibrium with informative cheap talk; as in the baseline model (see Remark 2 of the main text) we do not characterize the entire equilibrium set. Nevertheless, communication about candidates' preferences is relevant regardless of which equilibria and what (non-singleton) message space one considers. We establish in Lemma A.1 below that the voter's welfare depends continuously on the distribution of the state of the world. Consequently, when the candidates' signals of the state are not very informative, the voter's welfare approaches $\mathcal{U}(p,k)$ in any equilibrium in which candidates' messages convey information only about the state and not their preference type. On the other hand, the voter's welfare approaches $\mathcal{U}(0,0)$ in the equilibrium characterized in Proposition A.1. It follows that our welfare conclusions about informative cheap talk about preference types, described in Corollary 1 and Proposition 4 of the main text, extend when candidates have limited private information about the state. In particular, when the strength of reputation concerns, k, is sufficiently large, any equilibrium in which candidates' messages do not provide information about their preference types yields lower voter welfare than the equilibrium of Proposition A.1.

A.2. Proof of Proposition A.1

The proof of Proposition A.1 requires a number of steps; many of them parallel and extend those in the main paper.

Before dealing with private signals about the state, we first prove a continuity property.

Lemma A.1. *There is a neighborhood of $F(\cdot)$ within which $s_0^*(p,k)$, $\mathcal{U}(p,k)-\mathcal{U}(0,k)$, and $W(\theta,p,k)$ are uniquely defined for each CDF with a continuous density and vary continuously with such CDFs.*

Proof. Recall that $s_0^*(p,k)$ is implicitly defined by Equation 6 of the main text; we have uniqueness if there is a unique solution to Equation 6. We must show that there is an open neighborhood of $F(\cdot)$ in which, for any distribution $\hat{F}(\cdot)$ that admits a continuous density $\hat{f}$ in this neighborhood, there is a unique solution to

$$\Gamma(s_0^*,\hat{F}(\cdot)) := s_0^* - \frac{\overline{a}+\underline{a}}{2} - \frac{k}{2(\overline{a}-\underline{a})}\left[V\left(\frac{p}{p+(1-p)\frac{\hat{F}(s_0^*-b)}{\hat{F}(s_0^*)}}\right) - V\left(\frac{p}{p+(1-p)\frac{1-\hat{F}(s_0^*-b)}{1-\hat{F}(s_0^*)}}\right)\right] = 0.$$

As $\Gamma(s,\hat{F}(\cdot))$ is continuous in s and $V(\cdot)$ is strictly increasing and bounded between 0 and 1, there exists a solution $s_0^* \in \left[\frac{\overline{a}+\underline{a}}{2}, \frac{\overline{a}+\underline{a}}{2} + \frac{k}{2(\overline{a}-\underline{a})}\right]$ and every solution must be in $\left[\frac{\overline{a}+\underline{a}}{2}, \frac{\overline{a}+\underline{a}}{2} + \frac{k}{2(\overline{a}-\underline{a})}\right]$. We now show that there there is a neighborhood around $F(\cdot)$ in which there cannot exist multiple solutions with $s_0^* \in \left[\frac{\overline{a}+\underline{a}}{2}, \frac{\overline{a}+\underline{a}}{2} + \frac{k}{2(\overline{a}-\underline{a})}\right]$ for any $\hat{F}(\cdot)$ in this neighborhood.

Repeating the steps of the proof of Proposition 1 of the main text using the shorthand $\alpha \equiv (1-p)/p$, $s \equiv s_0^*$, $G(s) \equiv \hat{F}(s-b)/\hat{F}(s)$, and $H(s) \equiv (1-\hat{F}(s-b))/(1-\hat{F}(s))$, it is sufficient to show that $\frac{\partial\Gamma(s,\hat{F}(\cdot))}{\partial s} = 1 - R(s,\hat{F}(\cdot)) > 0$ on $\left[\frac{\overline{a}+\underline{a}}{2}, \frac{\overline{a}-\underline{a}}{2} + \frac{k}{2(\overline{a}-\underline{a})}\right]$ for all $\hat{F}(\cdot)$ in a neighborhood of $F(\cdot)$, where

$$R(s,\hat{F}(\cdot)) := \frac{k\alpha}{2(\overline{a}-\underline{a})}\left[V'\left(\frac{1}{1+\alpha H(s)}\right)\frac{H'(s)}{(1+\alpha H(s))^2} - V'\left(\frac{1}{1+\alpha G(s)}\right)\frac{G'(s)}{(1+\alpha G(s))^2}\right],$$

and

$$G'(s) = \frac{\hat{F}(s)\hat{f}(s-b) - \hat{F}(s-b)\hat{f}(s)}{(\hat{F}(s))^2},$$
$$H'(s) = \frac{(1-\hat{F}(s-b))\hat{f}(s) - (1-\hat{F}(s))\hat{f}(s-b)}{(1-\hat{F}(s))^2}.$$

We established that $R(s,\hat{F}(\cdot)) \leq 0$ when $\hat{F}(\cdot) = F(\cdot)$ in the proof of Proposition 1. Furthermore, by inspection, $R(s,\hat{F}(\cdot))$ is continuous with respect to $\hat{F}(\cdot)$. Moreover, $G'(s)$ and $H'(s)$ vary continuously with $\hat{F}(\cdot)$ and, when $\hat{F}(\cdot) = F(\cdot)$, $G'(s)$ and $H'(s)$ are bounded on the interval $s \in \left[\frac{\overline{a}+\underline{a}}{2}, \frac{\overline{a}+\underline{a}}{2} + \frac{k}{2(\overline{a}-\underline{a})}\right]$. Finally, $V'(\cdot)$ is bounded by assumption. It follows that there is some $\delta > 0$ such that $1 - R(s,\hat{F}(\cdot)) > 0$ for all $s \in \left[\frac{\overline{a}+\underline{a}}{2}, \frac{\overline{a}+\underline{a}}{2} + \frac{k}{2(\overline{a}-\underline{a})}\right]$ when $d(\hat{F}(\cdot),F(\cdot)) < \delta$. As such, there can be at most one solution in this interval. As every solution must be in this interval, this proves there is a unique solution, denote it $s_0^*(p,k|\hat{F}(\cdot))$, for any $\hat{F}(\cdot)$ with $d(\hat{F}(\cdot),F(\cdot)) < \delta$. As $s_0^*(p,k|\hat{F}(\cdot))$ is uniquely determined, it is clear that $\mathcal{U}(p,k)$ and $W(\theta,p,k)$ are uniquely determined as well.

We next show that $s_0^*(p,k|\hat{F}(\cdot))$ is continuous in $\hat{F}(\cdot)$ within a neighborhood of $F(\cdot)$. Recall that $s_0^*(p,k|\hat{F}(\cdot))$ is defined by $\Gamma(s^*(p,k|\hat{F}(\cdot)),\hat{F}(\cdot)) = 0$. For any given s, $\Gamma(s,\hat{F}(\cdot))$ is continuous in $\hat{F}(\cdot)$. Since $\Gamma(s,F(\cdot))$ is strictly increasing in s, it follows that for any $\varepsilon > 0$, $\Gamma(s_0^*(p,k|F(\cdot)) - \varepsilon, F(\cdot)) < 0 < \Gamma(s_0^*(p,k|F(\cdot)) + \varepsilon, F(\cdot))$. Furthermore, since $\Gamma(s_0^*,\cdot)$ is continuous in its second argument, there exists $\delta > 0$ such that $\Gamma(s_0^*(p,k|F(\cdot)) - \varepsilon, \hat{F}(\cdot)) < 0 < \Gamma(s_0^*(p,k|F(\cdot)) + \varepsilon, \hat{F}(\cdot))$ whenever $d(\hat{F}(\cdot),F(\cdot)) < \delta$. We conclude that $s_0^*(p,k|\hat{F}(\cdot)) \in (s_0^*(p,k|F(\cdot)) - \varepsilon, s_0^*(p,k|F(\cdot)) + \varepsilon)$ for any $\hat{F}(\cdot)$ with $d(\hat{F}(\cdot),F(\cdot)) < \delta$. Therefore, $s_0^*(p,k)$ is continuous in $\hat{F}(\cdot)$.

Now consider $\mathcal{U}(p,k)-\mathcal{U}(0,k)$. Using Equation 7 from the main text and the definition of $U(\cdot)$ preceding it,

$$\mathcal{U}(p,k)-\mathcal{U}(0,k) = p\int_{\frac{\overline{a}+\underline{a}}{2}-b}^{s_0^*(p,k)}[(\overline{a}-s)^2-(s-\underline{a})^2]\hat{f}(s)\mathrm{d}s+(1-p)\int_{\frac{\overline{a}+\underline{a}}{2}-b}^{s_0^*(p,k)-b}[(\overline{a}-s)^2-(s-\underline{a})^2]\hat{f}(s)\mathrm{d}s.$$

As $s_0^*(p,k)$ is continuous in $\hat{F}(\cdot)$, the integrands are continuous and bounded, and the intervals of the integration are bounded, it follows that $\mathcal{U}(p,k)-\mathcal{U}(0,k)$ is continuous in $\hat{F}(\cdot)$ too.

Finally, consider $W(\theta,p,k)$. Using Equation 8 from the main text and the definition of v_θ in Equation 13 of the main Appendix,

$$W(\theta,p,k) = \int_{\frac{\overline{a}+\underline{a}}{2}-\theta}^{s_\theta^*(p,k)}[(\overline{a}-s-\theta)^2-(\underline{a}-s-\theta)^2]\hat{f}(s)\mathrm{d}s+k[\hat{F}(s_\theta^*(p,k))V(\overline{p}(p,k))+(1-\hat{F}(s_\theta^*(p,k)))V(\underline{p}(p,k))].$$

Recalling the definitions of $\overline{p}(p,k)$ and $\underline{p}(p,k)$ from Equations 2 and 3 of the main text, it follows by inspection that both are continuous in $\hat{F}(\cdot)$ and $s_0^*(p,k)$. Hence, as $s_0^*(p,k)$ is continuous in $\hat{F}(\cdot)$, the integrands are continuous and bounded, and the intervals of the integration are finite, it follows that $W(\theta,p,k)$ is continuous in $\hat{F}(\cdot)$ too. □

We now consider the model with private signals of the state. In the campaign stage the candidates' announcements determine not just the belief about the candidates' types but also about the state of the world. Let $\Phi(\cdot|\boldsymbol{m})$ denote the beliefs about the state given message pair $\boldsymbol{m}\equiv(m_A,m_B)\in\{0,b\}^2$. The next lemma establishes some properties of $\Phi(\cdot|\boldsymbol{m})$.

Lemma A.2. *Suppose that in the campaign stage a type-b candidate announces 0 if and only if $\beta_i\leq\overline{\beta}\in[0,1)$ and the type-0 candidate announces 0 for all β_i. Then, for all $\boldsymbol{m}\in\{0,b\}^2$,*

1. *$\Phi(\cdot|\boldsymbol{m})$ admits a continuous, strictly positive density $\phi(\cdot|\boldsymbol{m})$. Further, $\Phi(\cdot|(b,0))$ depends continuously on $\overline{\beta}\in[0,1)$.*
2. *$d(\Phi(\cdot|\boldsymbol{m}),F(\cdot))\leq\sup_{\boldsymbol{\beta}\in[0,1]^2}d(F(\cdot,\boldsymbol{\beta}),F(\cdot))$.*

Proof. First note that by symmetry, $\Phi(\cdot|(0,b))=\Phi(\cdot|(b,0))$, and so it is sufficient to calculate the distribution and density when $\boldsymbol{m}\in\{(0,0),(b,b),(0,b)\}$. These three message pair events correspond, respectively, to the following events:

$$\{(\theta_A,\beta_A),(\theta_B,\beta_B):\theta_A=\theta_B=b,\beta_A\geq\overline{\beta},\beta_B\geq\overline{\beta}\},$$
$$\{(\theta_A,\beta_A),(\theta_B,\beta_B):\theta_i=0 \text{ or } \beta_i\leq\overline{\beta} \text{ for } i=A,B\}, \text{ and}$$
$$\{(\theta_A,\beta_A):\theta_A=b,\beta_A\geq\overline{\beta}\}\cap\{(\theta_B,\beta_B):\theta_B=0 \text{ or } \beta_B\leq\overline{\beta}\}.$$

Note also that for any $\boldsymbol{m}\in\{(0,0),(b,b),(b,0)\}$, $\Phi(s|\boldsymbol{m})=\mathbb{E}[F(s,\boldsymbol{\beta})|\boldsymbol{m}]$ and $\phi(s|\boldsymbol{m})=\mathbb{E}[f(s,\boldsymbol{\beta})|\boldsymbol{m}]$.

Proof of Part 1: From the formula for $\phi(s|\boldsymbol{m})$ above it follows immediately that $\phi(s|\boldsymbol{m})$ is strictly positive for all s and depends continuously on s.

We now establish that $\Phi(\cdot|(b,0),\overline{\beta})$ and $\phi(\cdot|(b,0),\overline{\beta})$ depend continuously on $\overline{\beta}\in[0,1)$. To do so we first define, for any $\overline{\beta}_1\in[0,1)$ and $\overline{\beta}_2\in[0,1)$,

$$A(\overline{\beta}_1):=\{(\theta_A,\beta_A):\theta_A=b,\beta_A\geq\overline{\beta}_1\} \text{ and } B(\overline{\beta}_2):=\{(\theta_B,\beta_B):\theta_B=0 \text{ or } \beta_B\leq\overline{\beta}_2\}.$$

Note that the event $\boldsymbol{m}=(b,0)$ given cutoffs $(\overline{\beta}_1,\overline{\beta}_2)$ occurs when $A(\overline{\beta}_1)$ and $B(\overline{\beta}_2)$ both happen. The probability is

$$P(\overline{\beta}_1,\overline{\beta_2}):=\Pr(A(\overline{\beta}_1)\cap B(\overline{\beta}_2)),$$

which is strictly greater than 0 whenever $\overline{\beta}_1<1$, strictly decreasing in $\overline{\beta}_1$, strictly increasing in $\overline{\beta}_2$, and depends continuously on both $\overline{\beta}_1$ and $\overline{\beta}_2$.

Now fix $\overline{\beta}>0$ and let $\delta>0$. Our goal is to calculate how $\Phi(\cdot|(b,0),\overline{\beta})$ and $\phi(\cdot|(b,0),\overline{\beta})$ change when the cutoff changes to $\overline{\beta}+\delta$ or $\overline{\beta}-\delta$ for some $\delta>0$. We have

$$\begin{aligned}
&\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta}-\delta)\cap B(\overline{\beta})]P(\overline{\beta}-\delta,\overline{\beta})\\
&=\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta})\cap B(\overline{\beta})]P(\overline{\beta},\overline{\beta})\\
&\quad+\mathbb{E}[F(s,\boldsymbol{\beta})|(A(\overline{\beta}-\delta)\backslash A(\overline{\beta}))\cap B(\overline{\beta})][P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta},\overline{\beta})],
\end{aligned}$$

and also,

$$\begin{aligned}
&\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta}-\delta)\cap B(\overline{\beta})]P(\overline{\beta}-\delta,\overline{\beta})\\
&=\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta}-\delta)\cap B(\overline{\beta}-\delta)]P(\overline{\beta}-\delta,\overline{\beta}-\delta)\\
&\quad+\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta}-\delta)\cap(B(\overline{\beta})\backslash B(\overline{\beta}-\delta))][P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta}-\delta,\overline{\beta}-\delta)].
\end{aligned}$$

It follows that

$$
\begin{aligned}
&\Phi(s|(b,0),\overline{\beta}-\delta)-\Phi(s|(b,0),\overline{\beta})\\
&=\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta}-\delta)\cap B(\overline{\beta}-\delta)]-\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta})\cap B(\overline{\beta})]\\
&=\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta}-\delta)\cap B(\overline{\beta})]\left(\frac{P(\overline{\beta}-\delta,\overline{\beta})}{P(\overline{\beta}-\delta,\overline{\beta}-\delta)}-\frac{P(\overline{\beta}-\delta,\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\right)\\
&\quad+\mathbb{E}[F(s,\boldsymbol{\beta})|(A(\overline{\beta}-\delta)\backslash A(\overline{\beta}))\cap B(\overline{\beta})]\frac{P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta},\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\\
&\quad-\mathbb{E}[F(s,\boldsymbol{\beta})|A(\overline{\beta}-\delta)\cap(B(\overline{\beta})\backslash B(\overline{\beta}-\delta))]\frac{P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta}-\delta,\overline{\beta}-\delta)}{P(\overline{\beta}-\delta,\overline{\beta}-\delta)}.
\end{aligned}
$$

Therefore, as $F(s)\in[0,1]$ for all $s\in(\underline{s},\infty)$,

$$
\begin{aligned}
&|\Phi(s|(b,0),\overline{\beta}-\delta)-\Phi(s|(b,0),\overline{\beta})|\\
&\leq\left|\left(\frac{P(\overline{\beta}-\delta,\overline{\beta})}{P(\overline{\beta}-\delta,\overline{\beta}-\delta)}-\frac{P(\overline{\beta}-\delta,\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\right)\right|+\left|\frac{P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta},\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\right|+\left|\frac{P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta}-\delta,\overline{\beta}-\delta)}{P(\overline{\beta}-\delta,\overline{\beta}-\delta)}\right|.
\end{aligned}
$$

By a change of labels,

$$
\begin{aligned}
&|\Phi(s|(b,0),\overline{\beta})-\Phi(s|(b,0),\overline{\beta}+\delta)|\\
&\leq\left|\left(\frac{P(\overline{\beta},\overline{\beta}+\delta)}{P(\overline{\beta},\overline{\beta})}-\frac{P(\overline{\beta},\overline{\beta}+\delta)}{P(\overline{\beta}+\delta,\overline{\beta}+\delta)}\right)\right|+\left|\frac{P(\overline{\beta},\overline{\beta}+\delta)-P(\overline{\beta}+\delta,\overline{\beta}+\delta)}{P(\overline{\beta}+\delta,\overline{\beta}+\delta)}\right|+\left|\frac{P(\overline{\beta},\overline{\beta}+\delta)-P(\overline{\beta},\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\right|
\end{aligned}
$$

for all $s\in(\underline{s},\infty)$. Finally since $P(\overline{\beta}_1,\overline{\beta}_2)$ is continuous in $\overline{\beta}_1$ and $\overline{\beta}_2$ when $\overline{\beta}_1<1$, these bounds go to 0 as $\delta\to 0$ for any $\overline{\beta}\in[0,1)$.

Now consider the densities. We repeat the same calculations as above except conditioning on $f(\cdot)$ instead of $F(\cdot)$. Taking $\overline{f}=\sup_{s\in(\underline{s},\infty)}f(s)<\infty$, it follows that for all $s\in(\underline{s},\infty)$,

$$
\begin{aligned}
&|\phi(s|(b,0),\overline{\beta}-\delta)-\phi(s|(b,0),\overline{\beta})|\\
&\leq\overline{f}\left(\left|\left(\frac{P(\overline{\beta}-\delta,\overline{\beta})}{P(\overline{\beta}-\delta,\overline{\beta}-\delta)}-\frac{P(\overline{\beta}-\delta,\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\right)\right|+\left|\frac{P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta},\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\right|+\left|\frac{P(\overline{\beta}-\delta,\overline{\beta})-P(\overline{\beta}-\delta,\overline{\beta}-\delta)}{P(\overline{\beta}-\delta,\overline{\beta}-\delta)}\right|\right)
\end{aligned}
$$

and

$$
\begin{aligned}
&|\phi(s|(b,0),\overline{\beta})-\phi(s|(b,0),\overline{\beta}+\delta)|\\
&\leq\overline{f}\left(\left|\left(\frac{P(\overline{\beta},\overline{\beta}+\delta)}{P(\overline{\beta},\overline{\beta})}-\frac{P(\overline{\beta},\overline{\beta}+\delta)}{P(\overline{\beta}+\delta,\overline{\beta}+\delta)}\right)\right|+\left|\frac{P(\overline{\beta},\overline{\beta}+\delta)-P(\overline{\beta}+\delta,\overline{\beta}+\delta)}{P(\overline{\beta}+\delta,\overline{\beta}+\delta)}\right|+\left|\frac{P(\overline{\beta},\overline{\beta}+\delta)-P(\overline{\beta},\overline{\beta})}{P(\overline{\beta},\overline{\beta})}\right|\right).
\end{aligned}
$$

Again, these bounds go to 0 as $\delta \to 0$ when $\overline{\beta} < 1$ by the continuity of $P(\overline{\beta}_1, \overline{\beta}_2)$ when $\overline{\beta}_1 < 1$.

It follows that, for any $\overline{\beta} < 1$ and $\varepsilon > 0$, there exists a $\overline{\delta} > 0$ such that

$$d(\Phi(s|(b,0),\overline{\beta}), \Phi(s|(b,0),\beta)) < \varepsilon$$

for all $s \in (\underline{s}, \infty)$ if $\delta = |\overline{\beta} - \beta| < \overline{\delta}$. So $\Phi(\cdot|(b,0))$ is continuous in $\overline{\beta}$.

Proof of Part 2: As θ_i and β_i are independent for $i \in \{A, B\}$, it is immediate that for any $\boldsymbol{m} \in \{(0,0),(b,b),(b,0)\}$ and any $s \in (\underline{s}, \infty)$,

$$\inf_{\boldsymbol{\beta} \in [0,1]^2} F(s, \boldsymbol{\beta}) \leq \Phi(s|\boldsymbol{m}) = \mathbb{E}[F(s,\boldsymbol{\beta})|\boldsymbol{m}] \leq \sup_{\boldsymbol{\beta} \in [0,1]^2} F(s, \boldsymbol{\beta}),$$

and

$$\inf_{\boldsymbol{\beta} \in [0,1]^2} f(s, \boldsymbol{\beta}) \leq \phi(s|\boldsymbol{m}) = \mathbb{E}[f(s,\boldsymbol{\beta})|\boldsymbol{m}] \leq \sup_{\boldsymbol{\beta} \in [0,1]^2} f(s, \beta).$$

Therefore, it follows that $d(\Phi(\cdot|\boldsymbol{m}), F(\cdot)) \leq \sup_{\boldsymbol{\beta} \in [0,1]^2} d(F(\cdot, \boldsymbol{\beta}), F(\cdot))$. □

Given the beliefs induced by the messages we can study the policymaking "subgame". It proceeds exactly as in the baseline model, except that the belief about the state, $\Phi(\cdot|\boldsymbol{m})$, is now endogenously determined by the messages sent. At the policymaking stage we define $\mathcal{U}(p,k;\boldsymbol{m})$ and $W(\theta,p,k;\boldsymbol{m})$ to be the expected payoff for the voter and the elected PM given reputation p and the beliefs about the state induced by message pair $\boldsymbol{m}$.

The following lemma establishes that the cutoff in the policymaking stage is unique, and moreover, that the voter and candidate payoffs don't change much when the candidates' signals are not very informative.

Lemma A.3. *Suppose that in the campaign stage a type-b candidate announces 0 if and only if $\beta_i \leq \overline{\beta} \in [0,1)$ and the type-0 candidate announces 0 for all β_i. Let $\varepsilon_0 > 0$. There exists $\varepsilon_1(p,k,\varepsilon_0) > 0$ such that, for any $\overline{\beta} \in [0,1)$, if $d(F(\cdot,\boldsymbol{\beta}), F(\cdot)) < \varepsilon_1$ for all $\boldsymbol{\beta} \in [0,1]^2$, then for all $p \in [0,1]$ and $\boldsymbol{m} \in \{0,b\}^2$,*

1. *the equilibrium in the policymaking "subgame" is unique.*
2. $|\mathcal{U}(p,k;\boldsymbol{m}) - \mathcal{U}(0,k;\boldsymbol{m}) - \mathcal{U}(p,k) + \mathcal{U}(0,k)| < \varepsilon_0$.
3. $|W(\theta,p,k;\boldsymbol{m}) - W(\theta,p,k)| < \varepsilon_0$.

Proof. Without loss of generality, we consider a PM who was candidate A in the campaign stage. Any message pair $\boldsymbol{m} = (m_A, m_B) \in \{0,b\}^2$ induces distribution $\Phi(\cdot|\boldsymbol{m})$ over the state $s \in (\underline{s}, \infty)$ and belief $p_A(\boldsymbol{m}) \in [0,1]$ over the PM's type. By part 1 of Lemma A.2, $\Phi(\cdot|\boldsymbol{m})$ admits a continuous density. By Lemma A.1 it follows that there exists $\varepsilon_1^1(p,k) > 0$ such that, if $d(\Phi(\cdot|\boldsymbol{m}), F(\cdot)) < \varepsilon_1^1(p,k)$, then there is a unique equilibrium in the policymaking stage after A is elected. Furthermore, by part 2 of Lemma A.2, this follows if $d(F(\cdot,\boldsymbol{\beta}), F(\cdot)) < \varepsilon_1^1(p,k)$ for all $\boldsymbol{\beta} \in [0,1]^2$. Finally, having established that there is a unique equilibrium when $d(F(\cdot,\boldsymbol{\beta}), F(\cdot)) < \varepsilon_1^1(p,k)$, parts 2 and 3 follows from Lemma A.1 for some $\varepsilon_1(p,k,\varepsilon_0) \in (0, \varepsilon_1^1(p,k)]$. □

We next prove a limited single-crossing property: if a biased type announces $m_i = 0$ if and only if his signal is sufficiently low, then the congruent type optimizes by announcing $m_i = 0$ for all β_i.

Lemma A.4. *Let $p \in (0,1)$ and consider the following strategies in the campaign stage: type-b candidates announces 0 if and only if $\beta_i \le \overline{\beta}$ for some $\overline{\beta} \in (0,1)$ and type-0 candidates announce 0 for all β_i. There exists $\varepsilon_2(p,k,\overline{\beta}) > 0$ such that, if $d(F(\cdot,\boldsymbol{\beta}), F(\cdot)) < \varepsilon_2$ for all $\boldsymbol{\beta} \in [0,1]^2$, then, if type $\theta = b$ is optimizing, so is type $\theta = 0$.*

Proof. Without loss of generality, we consider candidate A. It is sufficient to show that there exists $\varepsilon_2(p,k,\overline{\beta}) > 0$ such that if type b is indifferent when $\beta_A = \overline{\beta}$ then type 0's best response is to announce 0 regardless of his signal of the state. To prove this we first calculate how much higher the payoff of a type-0 candidate would be from announcing $m_i = 0$ than a type b candidate under the specified strategies if β_i and s were uncorrelated. We then show that if the correlation isn't too strong the payoff of a type 0 candidate announcing $m_i = 0$ after observing $\beta_i = 1$ is higher than of a type b candidate who observed $\beta_i = \overline{\beta}$.

Unconditional on the state, the probability a candidate of type-b announces 0 is

$$\mu^b := \int_{\underline{s}}^{\infty} \Psi(\overline{\beta}, s) f(s) \mathrm{d}s,$$

and the probability a candidate of type 0 announces 0 is $\mu^0 := 1$. By Bayes' rule,

$$p^0 := \frac{p}{p + (1-p)\mu^b} \in (0,1)$$

is the voter's posterior on candidate i being congruent given the announcement $m_i = 0$. We define

$$\Delta(\overline{\beta}) := W(0, p^0, k) - W(b, p^0, k),$$

which is strictly greater than 0 by Lemma 2 in the main text. (Note that p^0 depends on $\overline{\beta}$, but we have suppressed the notation.) Let $\sigma \in [0,1]$ denote the probability with which the voter elects the candidate who announced $m_i = 0$ if both candidates make different announcements. This means that, if β and s were uncorrelated, then a type-0 candidate's payoff from announcing 0 would be

$$\left[(1-p)(1-\mu^b)\sigma + (1-(1-p)(1-\mu^b))\frac{1}{2}\right]\Delta(\overline{\beta}) > \frac{p\Delta(\overline{\beta})}{2},$$

higher than a type-b's payoff from announcing 0. On the other hand, both types would receive the same payoff from announcing b. Note that this payoff difference is bounded strictly above 0 for all σ.

However, since β_i and s are correlated, we now consider the above strategies with positive correlation. Since candidates only know their own signal, and not their rival's, we write $F(s,(\beta_A,\emptyset))$ and $f(s,(\beta_A,\emptyset))$ to denote the distribution and density of the state given candidate A's signal without knowing candidate B's signal. Note that as

$$f(s,(\beta_A,\emptyset)) = \int_0^1 f(s,(\beta_A,x))\Pr(\beta_B = x)\mathrm{d}x \;\text{ and }\; F(s,(\beta_A,\emptyset)) = \int_0^1 F(s,(\beta_A,x))\Pr(\beta_B = x)\mathrm{d}x,$$

it holds that $|f(s,(\beta_A,\emptyset)) - f(s)|$ and $|F(s,(\beta_A,\emptyset)) - F(s)|$ are both bounded above for all $\beta_A \in [0,1]$ by

$$\sup_{\boldsymbol{\beta}\in[0,1]^2} d(F(\cdot,\boldsymbol{\beta}),F(\cdot)).$$

We now consider the incentives of candidate A for different type and signal combinations. Consider first the case in which candidate A is type $\theta = b$. We require him to be indifferent between announcing b and 0 if $\beta_A = \overline{\beta}$. So define

$$w_0(\beta_A,\overline{\beta}) := (1-p)\int_{\underline{s}}^{\infty}(1-\Psi(\overline{\beta},s))f(s,(\beta_A,\emptyset))\mathrm{d}s \in [0,1]$$

as the probability, conditional on β_A, that B announces b given the specified strategies. Indifference requires that that the payoff from announcing b when $\beta_A = \overline{\beta}$,

$$\left[w_0(\overline{\beta},\overline{\beta})\frac{1}{2} + (1-w_0(\overline{\beta},\overline{\beta}))(1-\sigma)\right]c$$

is equal to the payoff from announcing 0 when $\beta_A = \overline{\beta}$,

$$w_0(\overline{\beta}, \overline{\beta})\sigma(c + W(b, p_A(0, b), k, (0, b))) + (1 - w_0(\overline{\beta}, \overline{\beta}))\frac{1}{2}(c + W(b, p_A(0, 0), k, (0, 0))).$$

Equivalently, letting $y^b(\sigma)$ reflect the net benefit of announcing 0 relative to b for a type-b candidate who observed signal $\overline{\beta}$, we must have

$$y^b(\sigma) := w_0(\overline{\beta}, \overline{\beta})\sigma W(b, p_A(0, b), k; (0, b)) + (1 - w_0(\overline{\beta}, \overline{\beta}))\frac{1}{2}W(b, p_A(0, 0), k; (0, 0)) - \left(\frac{1}{2} - \sigma\right)c = 0. \tag{A.2}$$

Now consider a type-0 candidate who observed signal $\beta_i = 1$. The difference between announcing 0 and b for this candidate is

$$y^0(\sigma) := w_0(1, \overline{\beta})\sigma W(0, p_A(0, b), k; (0, b)) + (1 - w_0(1, \overline{\beta}))\frac{1}{2}W(0, p_A(0, 0), k; (0, 0)) - \left(\frac{1}{2} - \sigma\right)c.$$

We must show that $y^0(\sigma) - y^b(\sigma) > 0$ for all $\sigma \in [0, 1]$. Observe that

$$\begin{aligned} y^0(\sigma) - y^b(\sigma) = \frac{1}{2}[&(1 - w_0(1, \overline{\beta}))W(0, p_A(0, 0), k; (0, 0)) - (1 - w_0(\overline{\beta}, \overline{\beta}))W(b, p_A(0, 0), k; (0, 0))] \\ &+ \sigma[w_0(1, \overline{\beta})W(0, p_A(0, b), k; (0, b)) - w_0(\overline{\beta}, \overline{\beta})W(b, p_A(0, b), k; (0, b))]. \end{aligned}$$

Let $\overline{\varepsilon} \in (0, 3)$ and define $s_1(\overline{\varepsilon})$ and $s_2(\overline{\varepsilon})$ as the unique solutions to $F(s_1) = \frac{\overline{\varepsilon}}{6}$ and $F(s_2) = 1 - \frac{\overline{\varepsilon}}{6}$. It is immediate that $s_1 < s_2$. Next, note that for all $\beta_A \in [0, 1]$,

$$|w_0(\beta_A, \overline{\beta}) - (1 - p)(1 - \mu^b)| \leq (1 - p)\int_{\underline{s}}^{\infty}(1 - \Psi(\overline{\beta}, s))|f(s, (\beta_A, \emptyset)) - f(s)|\mathrm{d}s.$$

We can break up the integral $\int_{\underline{s}}^{\infty}(1 - \Psi(\overline{\beta}, s))|f(s, (\beta_A, \emptyset)) - f(s)|\mathrm{d}s$ into the sum of integrals from $(\underline{s}, s_1)$, from (s_1, s_2), and (s_2, ∞). Moreover, we calculate that

$$\begin{aligned} \int_{\underline{s}}^{s_1}(1 - \Psi(\overline{\beta}, s))|f(s, (\beta_A, \emptyset)) - f(s)|\mathrm{d}s &\leq \int_{\underline{s}}^{s_1}(f(s, (\beta_A, \emptyset)) + f(s))\mathrm{d}s \\ &= F(s_1, (\beta_A, \emptyset)) + F(s_1) \\ &\leq \frac{\overline{\varepsilon}}{3} + \sup_{\boldsymbol{\beta} \in [0,1]^2} d(F(\cdot, \boldsymbol{\beta}), F(\cdot)). \end{aligned}$$

Similarly,

$$\int_{s_2}^{\infty}(1-\Psi(\overline{\beta},s))|f(s,(\beta_A,\emptyset))-f(s)|\mathrm{d}s \leq \int_{s_2}^{\infty}(f(s,(\beta_A,\emptyset))+f(s))\mathrm{d}s$$
$$= 1-F(s_2,(\beta_A,\emptyset))+1-F(s_2)$$
$$\leq \frac{\overline{\varepsilon}}{3}+\sup_{\boldsymbol{\beta}\in[0,1]^2} d(F(\cdot,\boldsymbol{\beta}),F(\cdot)).$$

Finally,

$$\int_{s_1}^{s_2}(1-\Psi(\overline{\beta},s))|f(s,(\beta_A,\emptyset))-f(s)|\mathrm{d}s \leq \int_{s_1}^{s_2}|f(s,(\beta_A,\emptyset))-f(s)|\mathrm{d}s \leq (s_2-s_1)\sup_{\boldsymbol{\beta}\in[0,1]^2} d(F(\cdot,\boldsymbol{\beta}),F(\cdot)).$$

Therefore, if

$$\sup_{\boldsymbol{\beta}\in[0,1]^2} d(F(\cdot,\boldsymbol{\beta}),F(\cdot)) < \frac{(1+2p)\overline{\varepsilon}}{3(1-p)(s_2-s_1+2)},$$

then it follows that

$$|w_0(\beta_A,\overline{\beta})-(1-p)(1-\mu^b)| < \overline{\varepsilon}.$$

As $\overline{\varepsilon}$ is arbitrary, we conclude that $w_0(\beta_A,\overline{\beta}) \to (1-p)(1-\mu^b)$ uniformly for all $\beta_A \in [0,1]$ as $d(F(\cdot,\boldsymbol{\beta}),F(\cdot)) \to 0$.

Given that $w_0(\beta_A,\overline{\beta}) \to (1-p)(1-\mu^b)$ and $p(0,m_B)$ is continuous in $w_0(\beta_A,\overline{\beta})$ for $m_B \in \{0,b\}$ it follows immediately that $p(0,m_B) \to p^0$ as $d(F(\cdot,\boldsymbol{\beta}),F(\cdot)) \to 0$. Moreover, by part (3) of Lemma A.3, $W(\theta,p_A(\boldsymbol{m}),k;\boldsymbol{m}) \to W(\theta,p_A(\boldsymbol{m}),k)$ for every $m \in \{0,b\}$. Hence, for every $\boldsymbol{\beta} \in [0,1]^2$,

$$\lim_{d(F(\cdot,\boldsymbol{\beta}),F(\cdot))\to 0} y^0(\sigma)-y^b(\sigma) = \left[(1-p)(1-\mu^b)\sigma+(1-(1-p)(1-\mu^b))\frac{1}{2}\right]\Delta(\overline{\beta}) > 0.$$

Therefore, there exists $\varepsilon_2(p,k,\overline{\beta}) > 0$ such that, if $d(F(\cdot,\boldsymbol{\beta}),F(\cdot)) < \varepsilon_2$ for all $\boldsymbol{\beta} \in [0,1]^2$, then for any σ for which type $\theta = b$ is optimizing, type 0 is optimizing by always announcing 0. □

Finally, we show there exists a cutoff strategy for the biased type that makes the voter willing to randomize after observing different messages, and a voter randomization against which that cutoff strategy is optimal for a biased candidate.

Lemma A.5. *Suppose $k > k^*$ and $p \in (0,p^*(k))$. Then there exists a $\varepsilon_3(p,k) > 0$ such that, if $d(F(\cdot,\boldsymbol{\beta}),F(\cdot)) < \varepsilon_3$ for all $\boldsymbol{\beta} \in [0,1]^2$, there exists a $\overline{\beta} \in (0,1)$ such that if type $\theta = b$ candidates* *announce* 0 *if and only if* $\beta_i \leq \overline{\beta}$ *and type* $\theta = 0$ *candidates announce* 0 *for all* $\beta_i \in [0,1]$*, then* $\mathcal{U}(p_i(m_i = 0, m_{-i} = b), k, (b,0)) = \mathcal{U}(0, k, (b,0))$.

Proof. Define

$$p^{\min} = \min \left\{ \underset{p' \in [p,1]}{\arg\min} \, \mathcal{U}(p', k) \right\}$$

and

$$\varepsilon_0(p,k) = \mathcal{U}(0,k) - \mathcal{U}(p^{\min}, k).$$

Note that because $p < p^*(k)$ it follows that $\varepsilon_0(p,k) > 0$ and that $p^{\min} \in [p, p^*(k))$. Now define $\beta^b \in (0,1]$ such that if type $\theta = b$ candidates announce 0 if and only if $\beta_i \leq \beta^b$ and type $\theta = 0$ candidates announce 0 for all $\beta_i \in [0,1]$, then

$$p_i(m_i = 0, m_{-i} = b) = p^{\min}.$$

Such a β^b must exist by the intermediate value theorem because $p_i(m_i = 0, m_{-i} = b) = p$ when $\beta^b = 1$ and $p_i(m_i = 0, m_{-i} = b) = 1 > p^*(k)$ if $\beta^b = 0$. When $\overline{\beta} = \beta^b$ it follows from Lemma A.3 that there exists a $\varepsilon_3(p,k) \equiv \varepsilon_1(p, k, \varepsilon_0(p,k)) > 0$ such that

$$\mathcal{U}(p_i(m_i = 0, m_{-i} = b), k; (b,0)) < \mathcal{U}(0, k; (b,0))$$

when $d(F(\cdot, \boldsymbol{\beta}), F(\cdot)) < \varepsilon_3(p,k)$ for all $\boldsymbol{\beta} \in [0,1]^2$. However, when $\overline{\beta} = 0$,

$$\mathcal{U}(p_i(m_i = 0, m_{-i} = b), k; (b,0)) = \mathcal{U}(1, k; (b,0)) > \mathcal{U}(0, k; (b,0)).$$

Hence, by Lemma A.1 and part 1 of Lemma A.2, it follows from the intermediate value theorem that there exists a $\overline{\beta} \in (0, \beta^b)$ such that if type $\theta = b$ candidates announce 0 if and only if $\beta_i \leq \overline{\beta}$,

$$\mathcal{U}(p_i(m_i = 0, m_{-i} = b), k; (b,0)) = \mathcal{U}(0, k; (b,0)).$$ □

Lemma A.6. *Let* $\overline{\beta}$ *be as defined in Lemma A.5, and suppose the voter elects each candidate with equal probability when* $m_A = m_B$ *and the candidate who announced* $m_i = 0$ *with probability* σ *when* $m_A \neq m_B$*. There exists a* $\sigma \in (0, 1/2)$ *such that, if type* $\theta = 0$ *is announcing* $m_i = 0$ *for all* β_i *then type* $\theta = b$ *is optimizing by announcing* $m_i = 0$ *if and only if* $\beta_i \leq \overline{\beta}$.

Proof. This is equivalent to showing that there exists a solution to $y^b(\sigma) = 0$ in $(0, 1/2)$, where $y^b(\cdot)$ is defined in Equation A.2. As $W(b, p, k; \boldsymbol{m}) \in (0, k)$ for all $p \in (0,1)$ and $\boldsymbol{m}$ and $c \geq k$, it follows immediately that $y^b(0) < 0$ and $y^b(1/2) > 0$. Hence, since $y^b(\cdot)$ is continuous, existence of a solution to $y^b(\sigma) = 0$ with $\sigma \in (0, 1/2)$ follows by the intermediate value theorem. □

Proof of Proposition A.1. Define

$$\varepsilon(p, k) = \min\{\varepsilon_2(p, k, \overline{\beta}), \varepsilon_3(p, k)\},$$

where $\overline{\beta}$ and $\varepsilon_3(p, k)$ are defined in Lemma A.5 and $\varepsilon_2(p, k, \overline{\beta})$ is defined in Lemma A.4. The result is then an immediate combination of Lemma A.4, Lemma A.5 and Lemma A.6. □

B. Arbitrary Number of Types and Actions

This section of the Supplementary Appendix generalizes our paper's main results to more than two types and actions. Supplementary Appendix B.1 presents a more general model. Supplementary Appendix B.2 provides results about how sufficiently strong reputation concerns make the voter worse off when she has uncertainty about the PM's type than under any known type of PM. Supplementary Appendix B.3 discusses informative cheap talk in the campaign stage. Proofs are in Supplementary Appendix B.4.

B.1. Model and preliminaries

The baseline model in the paper assumed binary type and action spaces and made some assumptions (e.g. related to log convexity) on the distribution of states. We relax these assumptions here.

Setting. Assume the set of possible types for the PM is $\Theta \equiv \{\theta_1, \ldots, \theta_T\} \subset \mathbb{R}$, where $T \geq 2$ ($T \in \mathbb{N}$), $\theta_i > \theta_{i-1}$ for all $i \in \{2, \ldots, T\}$, and $0 \in \Theta$. Denote the prior on a politician's type by the vector $\boldsymbol{p} \equiv (p_1, \ldots, p_T)$ where p_i is the probability of type θ_i. For convenience, we also write $p[\theta]$ for the probability that $\boldsymbol{p}$ assigns to type θ. For now we consider only the policymaking stage, assuming that the PM is drawn according to $\boldsymbol{p}$; the campaign stage is discussed in Supplementary Appendix B.3. The PM takes an action after observing the state of the world. The set of states is $S \equiv \mathbb{R}$ with cumulative distribution function $F(\cdot)$. We assume the distribution has support S and a strictly positive density, $f(\cdot)$.[2] The PM's action must be chosen from a finite set $A \equiv \{a_1, \ldots, a_N\} \subset \mathbb{R}$, where $\underline{a} \equiv a_1 < \cdots < a_N \equiv \overline{a}$ and $N \geq 2$

[2] Strictly speaking, this is not a generalization of the baseline model because there we allowed for a finite lower bound on states. As should be clear from the subsequent analysis, we could also allow that here if we restricted attention to $\theta \geq 0$ (as in the baseline model).

($N \in \mathbb{N}$). The PM's preferences over actions $a \in A$ for any $\theta \in \Theta$ and $s \in S$ are represented by the utility function $u(a, s, \theta)$. We assume that for all a, s, and θ,

$$u(a, s - \theta, \theta) = u(a, s, 0). \tag{B.1}$$

We also assume $u(a, s, \theta)$ is continuously differentiable in s for all a and θ, and that, for all a', a, s, and θ,

$$a' > a \implies \frac{\partial}{\partial s}[u(a', s, \theta) - u(a, s, \theta)] > 0, \tag{B.2}$$

$$a' > a \implies \begin{cases} \lim_{s \to -\infty}[u(a', s, \theta) - u(a, s, \theta)] = -\infty \\ \lim_{s \to \infty}[u(a', s, \theta) - u(a, s, \theta)] = \infty. \end{cases} \tag{B.3}$$

Requirement (B.1) says that the PM's type acts like a "location shift": type 0's preferences are sufficient to deduce those of all other types. Requirement (B.2) says that the PM's utility function is supermodular in a and s; while (B.3) says that in extreme states the payoff difference between different actions becomes arbitrarily large. An example that satisfies all the requirements is $u(a, s, \theta) = -(a - s - \theta)^2$, which was used in the baseline model.

We assume the voter shares the policy preferences of a type-0 PM: the voter's utility function is also $u(a, s, 0)$. As in the baseline model, the voter does not observe the state or her own utility, but updates her belief about the PM's preference type based on the action she observes. For reasons to do with avoiding off-path beliefs, we now allow the voter to observe the PM's action with noise. Specifically, when the PM takes action a, the voter observes a with probability $1 - \eta$; with probability $\eta \in [0, 1)$, she observes a draw from a uniform distribution over A.[3] We think of η as either being equal to or close to 0, although this is not formally necessary. We maintain throughout that either $N = 2$ or $\eta > 0$. This will ensure that all actions are observed by the voter on the equilibrium path. When $N = 2$ we can allow for (but do not require) $\eta = 0$ because there will be no off-path actions; indeed, the baseline model corresponds to $N = 2$ and $\eta = 0$. When $N > 2$ it will be possible that some actions are never taken by the PM; the assumption that $\eta > 0$ will ensure that the voter's beliefs can be determined by Bayes' rule.[4]

The PM cares about reputation in addition to policy. A posterior $\hat{\boldsymbol{p}}(a) \equiv (\hat{p}_1(a), \ldots, \hat{p}_T(a))$ provides reputational value $\tilde{k}V(\hat{\boldsymbol{p}}(a))$ when the voter observes a, where $\tilde{k} > 0$ is a scalar and

[3] Any full support distribution could be used for the noise.

[4] Moreover, the beliefs after observing any action will vary continuously in the PM's strategy, which will prove an important technical property.

$V(\cdot)$ is a continuously differentiable function that maps voter posteriors into real numbers. The PM's net payoff when he chooses action a and the voter observes action a' is given by $u(a, s, \theta) + \tilde{k}V(\hat{\boldsymbol{p}}(a'))$. Thus, the expected payoff to the PM from choosing action a is

$$u(a, s, \theta) + \tilde{k}\left[(1 - \eta)V(\hat{\boldsymbol{p}}(a)) + \eta\frac{1}{N}\sum_{a' \in A} V(\hat{\boldsymbol{p}}(a'))\right].$$

Defining $k := \tilde{k}(1 - \eta)$, and noting that $\tilde{k}\eta(1/N)\sum_{a' \in A} V(\hat{\boldsymbol{p}}(a'))$ is independent of the PM's action, the PM's objective reduces to choosing a to maximize

$$u(a, s, \theta) + kV(\hat{\boldsymbol{p}}(a)),$$

i.e. the PM behaves as if the voter observes his action without noise but with a different weight on reputation. We refer to k as the PM's strength of reputation concern. We assume, consistent with the baseline model, that $\min V(\cdot) = 0$ and $\max V(\cdot) = 1$.

Strategies and equilibria. With some abuse of notation, a pure strategy for the PM is a function $a(s, \theta)$. We say that a PM's pure strategy is *monotonic* if for all θ and $s < s'$, $a(s, \theta) \leq a(s', \theta)$. Let $\boldsymbol{\delta_\theta}$ denote the probability distribution that puts probability one on type θ. We say that a distribution over types $\boldsymbol{p}$ is non-degenerate if it has non-singleton support. Without any essential loss of generality (because of the properties of $u(\cdot)$), we focus on equilibria in pure strategies, and furthermore, on those equilibria in which any action that is taken by some type on the equilibrium path is taken by that type for a positive-probability set of states.

Define the PM's utility difference between taking actions a' and a in state s when the voter updates using $\hat{\boldsymbol{p}}(\cdot)$ as

$$D(a', a, s, \theta, \hat{\boldsymbol{p}}(\cdot)) := u(a', s, \theta) + kV(\hat{\boldsymbol{p}}(a')) - [u(a, s, \theta) + kV(\hat{\boldsymbol{p}}(a))].$$

Plainly, (B.2) and (B.3) imply that for any a', a, s, θ, and $\hat{\boldsymbol{p}}(\cdot)$,

$$a' > a \implies \frac{\partial}{\partial s}D(a', a, s, \theta, \hat{\boldsymbol{p}}(\cdot)) > 0, \tag{B.4}$$

$$a' > a \implies \begin{cases} \lim\limits_{s \to -\infty} D(a', a, s, \theta, \hat{\boldsymbol{p}}(\cdot)) = -\infty \\ \lim\limits_{s \to \infty} D(a', a, s, \theta, \hat{\boldsymbol{p}}(\cdot)) = \infty. \end{cases} \tag{B.5}$$

It follows from (B.4) and (B.5) that in any equilibrium the PM uses a monotonic strategy and every type takes actions $\overline{a}$ and $\underline{a}$ on path. Let

$$\mathcal{S} := \{\boldsymbol{s} : s_1 = -\infty < s_2 \leq \cdots \leq s_N < s_{N+1} = +\infty\}.$$

We can identify the PM's behavior in any equilibrium by a vector for each type, $\boldsymbol{s}(\theta, \boldsymbol{p}, k) \in \mathcal{S}$, with the understanding that $a(s,\theta) = a_n$ if (and only if, modulo behavior at the boundaries) $s \in (s_n(\theta, \boldsymbol{p}, k), s_{n+1}(\theta, \boldsymbol{p}, k))$.[5] If $s_n(\theta, \boldsymbol{p}, k) = s_{n+1}(\theta, \boldsymbol{p}, k)$ then type θ does not play a_n. Since both $\underline{a}$ and $\overline{a}$ are taken with positive probability by all types, it holds in any equilibrium and for any type that $s_1 < s_2 \leq s_N < s_{N+1}$. If a type uses $\boldsymbol{s}$ with $s_2 = s_N$ then that type only ever takes actions $\underline{a}$ and $\overline{a}$.

Letting $a^-(s,\theta) := \lim_{\tilde{s}\uparrow s} a(\tilde{s},\theta)$ and $a^+(s,\theta) := \lim_{\tilde{s}\downarrow s} a(\tilde{s},\theta)$, it must hold in any equilibrium that for all θ,

$$\forall n: \; D\left(a^-(s_n(\theta,\cdot),\theta), a^+(s_n(\theta,\cdot),\theta), s_n(\theta,\cdot), \theta, \hat{\boldsymbol{p}}(\cdot)\right) = 0. \tag{B.6}$$

Equation B.6 simply says that at any cutoff state, s_n, the PM is indifferent between the action that he would take if the state were marginally higher and the action he would take if the state were marginally lower.[6] An equilibrium can be viewed as a collection of vectors, $\{\boldsymbol{s}(\theta, \boldsymbol{p}, k)\}_{\theta\in\Theta}$, such that: (i) $D(a(s,\theta), a', s, \theta, \hat{\boldsymbol{p}}(\cdot)) \geq 0$ for all θ, s, and a'; (ii) Equation B.6 is satisfied for each θ; and (iii) for any action a, $\hat{\boldsymbol{p}}(a)$ is derived via Bayes' rule given $\boldsymbol{p}$, η, and $\{\boldsymbol{s}(\theta, \boldsymbol{p}, k)\}_{\theta\in\Theta}$. Since $N = 2$ or $\eta > 0$, all actions are observed by the voter on the equilibrium path and Bayes' rule pins down the voter's beliefs.

Our assumption in Equation B.1 permits a simplification: in any equilibrium, any type θ will use cutoffs that differ from type 0's cutoffs by θ. This property follows from the fact that, for any $a', a, s, \theta, \hat{\boldsymbol{p}}(\cdot)$, $D(a', a, s - \theta, \theta, \hat{\boldsymbol{p}}(\cdot)) = D(a', a, s, 0, \hat{\boldsymbol{p}}(\cdot))$, and so, the preferences of type θ at state $s - \theta$ are identical to those of type 0 at state s. Hence, in any equilibrium $\{\boldsymbol{s}(\theta, \boldsymbol{p}, k)\}_{\theta\in\Theta}$, it holds that for all θ and n:

$$s_n(\theta, \boldsymbol{p}, k) = s_n(0, \boldsymbol{p}, k) - \theta. \tag{B.7}$$

By Equation B.7 any equilibrium can be characterized by a single vector $\boldsymbol{s}(\boldsymbol{p}, k) := \boldsymbol{s}(0, \boldsymbol{p}, k) \in \mathcal{S}$, which we will also view and refer to as a partition of S. While one cannot generally expect

[5] To be clear: $\boldsymbol{p}$ and k are parameters; we include them as arguments of $\boldsymbol{s}(\cdot)$ because we will subsequently study how $\boldsymbol{s}(\cdot)$ changes when $\boldsymbol{p}$ and k change. While the equilibrium also depends on η, we suppress that dependence as we are not interested in changes in η.

[6] Note that $a^-(s_n(\theta, \boldsymbol{p}, k), \theta)$ need not equal a_{n-1} because it may be that $s_{n-1} = s_n$, in which case a_{n-1} is never taken by type θ.

equilibrium uniqueness, it holds that:

Lemma B.1. *For any $\boldsymbol{p}$ and k, an equilibrium $\boldsymbol{s}(\boldsymbol{p}, k)$ exists. Moreover, there exists an equilibrium selection that is continuous in $\boldsymbol{p}$.*

We define $\mathcal{U}(\boldsymbol{p}, k)$ to be the voter's expected utility when the PM is drawn according to the prior $\boldsymbol{p}$ and the reputation concern is k; it is important to note that left implicit is an equilibrium selection $\boldsymbol{s}(\boldsymbol{p}, k)$. If the equilibrium selection $\boldsymbol{s}(\boldsymbol{p}, k)$ is continuous in $\boldsymbol{p}$ then $\mathcal{U}(\boldsymbol{p}, k)$ is also continuous in $\boldsymbol{p}$. We will use this continuity property in Supplementary Appendix B.3.

If $k = 0$, the PM has no reputational concern and so chooses a to maximize $u(\theta, a, s)$. Define $\boldsymbol{s}_\theta := \boldsymbol{s}(\boldsymbol{p}, 0) - \theta$ as the partition that would be used by a PM of type θ in this case. If the PM's type is known then there is no updating by the voter, and hence $\boldsymbol{s}(\boldsymbol{\delta_\theta}, k) = \boldsymbol{s}_\theta$ for all k and θ. Hence the voter's expected utility from a known candidate of type θ is $\mathcal{U}(\boldsymbol{\delta_\theta}, k) = \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$.

Assumptions for large distortions. Our aim is to establish that large enough reputation concerns can create extreme distortions. We first require that the voter is better off when faced with any known non-congruent type who has all available actions than a known congruent type who is constrained to not take one of the extreme actions. To state the assumption formally, let $\mathcal{U}_{\tilde{A}}(\boldsymbol{\delta_0}, 0)$ denote the voter's ex-ante utility when he is faced with known congruent type who is constrained to only take actions in the set $\tilde{A}$, i.e. take $a \in \tilde{A} \subseteq A$.

Assumption 1. *For any $\theta \in \Theta \setminus \{0\}$, $\mathcal{U}(\boldsymbol{\delta_\theta}, 0) > \max\{\mathcal{U}_{A\setminus\{\overline{a}\}}(\boldsymbol{\delta_0}, 0), \mathcal{U}_{A\setminus\{\underline{a}\}}(\boldsymbol{\delta_0}, 0)\}$.*

To interpret Assumption 1, observe that the gain from choosing $\overline{a}$ (resp. $\underline{a}$) over any other action in state s becomes become arbitrarily large as $s \to +\infty$ (resp. $s \to -\infty$), by (B.3). Thus, Assumption 1 holds when there is enough weight on tail states relative to the other parameters. In particular, for any $F(\cdot)$ and A, it holds if $|\theta|$ is small enough for all $\theta \in \Theta$. Moreover, when the type and action space are binary and the utility functions are quadratic loss, Assumption 1 is equivalent to the assumption that $\mathbb{E}\left[s|s > \frac{\underline{a}+\overline{a}}{2} - b\right] > \frac{\underline{a}+\overline{a}}{2}$, which was a maintained assumption in the baseline model.

We also assume the PM receives a higher reputational payoff from being thought more moderate:

Assumption 2. *$V(\hat{\boldsymbol{p}}) > V(\hat{\boldsymbol{p}}')$ if*

(i) $\sum_{i=j}^{m} \hat{p}_i \geq \sum_{i=j}^{m} \hat{p}'_i$ *for all $j, m \in \{1, \ldots, T\}$ with $\theta_j \leq 0 \leq \theta_m$, and*

(ii) $\sum_{i=j}^{m} \hat{p}_i > \sum_{i=j}^{m} \hat{p}'_i$ *for some* $j, m \in \{1, \ldots, T\}$ *with* $\theta_j \leq 0 \leq \theta_m$.

In general, Assumption 2 does not require that $V(\hat{\boldsymbol{p}}) > V(\hat{\boldsymbol{p}}')$ when $\hat{p}[0] > \hat{p}'[0]$. But when $T = 2$, Assumption 2 simplifies to $\hat{p}[0] > \hat{p}'[0] \iff V(\hat{\boldsymbol{p}}) > V(\hat{\boldsymbol{p}}')$: the PM's reputational payoff is increasing in the perceived probability of being congruent, as in the baseline model.

B.2. Any known devil is better than an unknown angel

This section addresses when any known, biased PM is preferred to a PM with uncertain type in any equilibrium, which we will refer to as "any known devil is better than an unknown angel".

B.2.1. Asymmetric distribution of types

For any probability distribution over types $\boldsymbol{p}$ define

$$p^+(\boldsymbol{p}) := \sum_{\{i:\theta_i>0\}} p_i \text{ and } p^-(\boldsymbol{p}) := \sum_{\{i:\theta_i<0\}} p_i$$

as the probability assigned by $\boldsymbol{p}$ to (strictly) upward and downward biased types respectively.

The following result says that for any sufficiently asymmetric $\boldsymbol{p}$, any known devil is preferred to an unknown angel in every equilibrium.

Proposition B.1. *Suppose Assumption 1 and Assumption 2 are satisfied. For any* $\alpha, \beta > 0$, *there exists an* $\varepsilon > 0$ *such that, for all* $\boldsymbol{p}$ *with* $p[0] \geq \beta$ *and either*

(i) $p^+(\boldsymbol{p}) \geq \alpha$ *and* $p^-(\boldsymbol{p}) < \varepsilon$, *or*

(ii) $p^-(\boldsymbol{p}) \geq \alpha$ *and* $p^+(\boldsymbol{p}) < \varepsilon$,

there exists a $\hat{k}(\boldsymbol{p})$ *such that, for all* $k > \hat{k}(\boldsymbol{p})$ *and all* $\theta \in \Theta$, $\mathcal{U}(\boldsymbol{p}, k) < \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$ *in every equilibrium.*

Here is the intuition behind Proposition B.1. If the PM could only be biased in favor of high (resp., low) actions, then owing to Assumption 2, the PM's reputation would be higher (resp., lower) after choosing $\underline{a}$ as opposed to $\overline{a}$. Consequently, when the type distribution is sufficiently skewed toward the PM having a bias in only direction, one of the extreme actions $a \in \{\underline{a}, \overline{a}\}$ will only be taken in arbitrarily extreme states as reputational concerns become paramount. By Assumption 1, if the concern for reputation is sufficiently strong, the voter will prefer a PM of known type to an uncertain PM.

Proposition B.1 generalizes the message from our baseline model because, when $T = 2$, either $p^-(\cdot) = 0$ or $p^+(\cdot) = 0$, and hence Proposition B.1 simply concludes that for any non-degenerate prior, a sufficiently strong reputation concern causes the voter to prefer a known-type PM to the PM of unknown type. Note that even within the scope of $T = 2$, the conclusion here is more general than in the baseline model in multiple ways: the utility functions need not be quadratic loss; there can be more than two actions; and we allow for any full-support distribution of states. In particular, while the baseline model's assumption about the density of the state distribution being log-convex on a suitable domain is needed to ensure a unique equilibrium in the policymaking stage and to obtain clean comparative statics there, it is not needed for the voter's welfare to be non-monotonic in the prior.

B.2.2. Symmetric environments

We now provide a result for symmetric environments (in a sense made precise below). Intuitively, when there are more than two actions ($N > 2$), the PM has an incentive to choose moderate actions in order to signal that he is a moderate type, which he desires to according to Assumption 2. However, when k is large, it may be possible to support an equilibrium in which only the extreme actions $\underline{a}$ and $\overline{a}$ are taken by the PM. Our next assumption says that it is not in the voter's interest to have the PM only take actions $\underline{a}$ and $\overline{a}$. Specifically, the voter is better off under any known non-congruent PM who can choose any action than under a known congruent PM who is constrained to only choose an extreme action.

Assumption 3. *For all $\theta \in \Theta \setminus \{0\}$, $\mathcal{U}(\boldsymbol{\delta_\theta}, 0) > \mathcal{U}_{\{\underline{a},-\overline{a}\}}(\boldsymbol{\delta_0}, 0)$.*

Plainly, Assumption 3 can only be satisfied if $N > 2$. If $N > 2$, Assumption 3 holds when the highest and lowest actions are sufficiently far apart (so that intermediate actions are optimal for a broad range of states), or when $|\theta|$ is sufficiently small for all $\theta \in \Theta$. Recall that the latter is also sufficient to guarantee Assumption 1.

We say that the *type space is symmetric* if, for all $i = 1, \ldots, \frac{T-1}{2}$, $\theta_i = -\theta_{T+1-i}$. Further, we say that a *type distribution $\boldsymbol{p}$ is symmetric* if, for all $i = 1, \ldots, \frac{T-1}{2}$, $p_i = p_{T+1-i}$. Finally, the reputation function is symmetric if it only depends on the magnitude, not the direction, the PM is thought to be biased. More precisely, given a symmetric type space, *the reputation function is symmetric* if $V(\hat{\boldsymbol{p}}) = V(\hat{\boldsymbol{p}}')$ for all $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{p}}'$ such that $\hat{p}_i + \hat{p}_{T+1-i} = \hat{p}'_i + \hat{p}'_{T+1-i}$ for all $i = 1, \ldots, \frac{T-1}{2}$.

Assumption 4. *The type space, Θ, and reputation function, $V(\cdot)$, are symmetric. The state density, $f(\cdot)$, is symmetric around 0 and single-peaked.*

The following result says that under our assumptions, any known devil is preferred to an unknown angel in any equilibrium of a sufficiently symmetric environment with large enough reputation concerns. The intuition is that, in any equilibrium with large k, either only the extreme actions are ever taken (and the result follows from Assumption 3) or moderate actions are taken for all but the most extreme states (and the result follows from Assumption 1).

Proposition B.2. *Suppose Assumptions 1–4 are satisfied. For any symmetric prior $\boldsymbol{p}$ that has full support, there exists $\varepsilon > 0$ such that for any prior $\boldsymbol{q}$ with $|\boldsymbol{p} - \boldsymbol{q}| < \varepsilon$, there exists $\hat{k}(\boldsymbol{q})$ such that, for all $k > \hat{k}(\boldsymbol{q})$ and all $\theta \in \Theta$, in every equilibrium it holds that $\mathcal{U}(\boldsymbol{q}, k) < \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$.*

Combining Proposition B.1 and Proposition B.2 we see that the generalized distortion result holds for any number of types when either: (i) the distribution of types is sufficiently asymmetric, or (ii) the environment is close enough to symmetric. These conditions are sufficient, but not necessary, to generate any known devil being preferred to a unknown angel.

B.3. Informative cheap talk

In this section we study cheap-talk communication during an election between two candidates that determines which of them becomes the PM. Each candidate's type is drawn independently from the prior distribution $\boldsymbol{p}$; they make simultaneous costless and non-binding announcements; and the voter elects one of them.

B.3.1. A limiting case

Proposition B.1 and Proposition B.2 provide conditions under which the voter prefers a known, biased PM to a PM whose type is uncertain. As in the paper's baseline model this creates an avenue for informative cheap-talk communication in the campaign stage. However, a general analysis of candidates' incentives to reveal information in the campaign stage is complicated when we have more than two types and actions. For this reason we focus on a limiting case, based on Section 5.1 of the main text, in which candidates care only about winning office in the campaign stage. Once in office, the benefits from being elected are no longer relevant, and the PM's incentives are as in Supplementary Appendix B.1. While there may not be a unique equilibrium in the policymaking stage, we fix a selection of "subgame" equilibria $\boldsymbol{s}(\boldsymbol{p}, k)$ that is continuous in $\boldsymbol{p}$, as is feasible by Lemma B.1.

As in the main text's two-type model, we will focus on symmetric equilibria in the campaign stage: both candidates randomize with the same (type-dependent) probabilities, and the voter's election probability after each pair of messages doesn't depend on which candidate made the announcement. An equilibrium is *informative* if there are distinct messages, each sent with positive probability on the equilibrium path, that induce different beliefs for the voter.

Proposition B.3. *Fix $k > 0$ and a selection, $\boldsymbol{s}(\boldsymbol{p}, k)$, from the equilibrium correspondence of the policymaking stage that is continuous in $\boldsymbol{p}$. If candidates seek to maximize their probability of election during the campaign stage and if $\mathcal{U}(\boldsymbol{p}, k) < \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$ for some $\theta \neq 0$, then there is an equilibrium with informative cheap-talk campaigns.*

The requirement that $\mathcal{U}(\boldsymbol{p}, k) < \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$ for some $\theta \neq 0$ amounts to saying that *some* known devil is preferred to an unknown angel; it is weaker than requiring that *any* known devil is. Plainly, it is sufficient that the hypotheses of either Proposition B.1 or Proposition B.2 are satisfied.

Proposition B.3 derives from an application of the intermediate value theorem. Suppose there are two messages, 0 and θ. If only the type θ candidate ever announced θ (and did so with small enough probability), the voter's utility would be higher from the candidate who announced θ than one who announced 0. Conversely, if all types other than 0 announced θ, then the voter would get a higher payoff from the candidate who announced 0. So there exists some intermediate profile of announcements for which the voter is indifferent. When the voter is indifferent and randomizes with equal probability after observing different messages we have an equilibrium. This equilibrium is informative as the posterior is different across different messages: in particular, if the voter observes message θ she knows with certainty the candidate is not type 0.

B.3.2. Additional examples

When candidates are forward looking in the campaign stage, they are also concerned with the reputation with which they are elected. As in the baseline model the candidate's utility is

$$c + v_\theta + u(a, s, \theta) + kV(\hat{\boldsymbol{p}}(a)),$$

if elected, with $v_\theta := -\mathbb{E}[u(a, s, \theta)|\boldsymbol{s}_\theta]$, and 0 if not elected. While a general analysis appears intractable, there are salient cases with more than two types/actions for which informative communication exists in the campaign stage when c and k are sufficiently large. Below, we describe the intuitions for two of them; a formal analysis is available on request.

Asymmetric environment. Informative communication is possible when there are two actions and three candidate types with biases in opposite directions (recall that type 0 is congruent) and a sufficiently asymmetric type distribution. If the probability that candidates are biased towards low (high) actions is sufficiently small, an equilibrium exists in which type 0 and low (high) types announce they are centrist/congruent, while high (low) types randomize between revealing their type and claiming to be centrist. The randomization that makes the voter indifferent will reduce, but not eliminate, pandering in office after a claim of congruence, and so the voter will update negatively after observing the high (low) action. Consequently, candidates who are biased towards the high (low) action have less incentive to claim to be centrist, which allows for informative communication in the campaign stage. In this sense, the baseline model's results on informative communication are robust to introducing a small probability the PM could be biased in the opposite direction.

Symmetric environment. Another interesting case is a symmetric environment with three types and three actions, $A = \{-1, 0, 1\}$. Here, an equilibrium exists in which type 0 announces that it is centrist/congruent while biased types randomize between announcing their true type and claiming to be centrist. The biased types' randomization makes the voter indifferent between candidates who claim to be centrist and candidates who reveal themselves to be biased. The voter's belief about a candidate is symmetric after a claim of being centrist. Hence, in the policymaking stage, the voter updates positively after seeing centrist actions. This ensures that biased types have less incentive than the centrist type to claim to be centrist in the campaign stage. The interesting conclusion is that rather than simply announcing whether they are biased or not, it is possible that candidates sometimes reveal the direction of their bias in their campaigns.

B.4. Proofs

B.4.1. Proofs for Supplementary Appendix B.1

Proof of Lemma B.1. Throughout the proof, fix some k. We first show that for any $\boldsymbol{p}$, an equilibrium exists. Our assumptions on $u(\cdot)$ imply that there is some $M > 0$ such that no matter the posterior $\hat{\boldsymbol{p}}(\cdot)$, the PM's partition choice $\boldsymbol{s}$ will satisfy $s_2, s_N \in [-M, M]$. Plainly, $\mathcal{S}^M := \{\boldsymbol{s} \in \mathcal{S} : s_2, s_N \in [-M, M]\} \subset \mathcal{S}$ is a compact and convex set.[7] Given any prior $\boldsymbol{p}$ and partition $\boldsymbol{s}$, the voter's posterior $\hat{\boldsymbol{p}}(\cdot)$ is determined by Bayes' rule because all actions have a positive probability of being observed (since $N = 2$ or $\eta > 0$); furthermore, the posterior is continuous in $\boldsymbol{s}$. By (B.4), there is a unique partition that maximizes the PM's utility given any $\hat{\boldsymbol{p}}(\cdot)$. Moreover, as the PM's payoff is continuous in $\hat{\boldsymbol{p}}(\cdot)$, the theorem of the maximum implies

[7] To avoid any confusion: we view $\mathcal{S}$ as endowed with the metric $|\boldsymbol{s} - \boldsymbol{s}'| = \max_{n \in \{2,\dots,N\}} |s_n - s'_n|$.

that the PM's optimal partition is continuous in $\hat{\boldsymbol{p}}(\cdot)$.[8] A routine application of Brouwer's fixed point theorem now yields a pure-strategy equilibrium.

Next, we claim the equilibrium correspondence is upper hemi-continuous in $\boldsymbol{p}$. Since $\boldsymbol{s}(\cdot, k) \in \mathcal{S}^M$ it is sufficient to verify that given any sequence $j \to \infty$, with $\boldsymbol{p}^j \to \boldsymbol{p}$ and $\boldsymbol{s}^j \in \boldsymbol{s}(\boldsymbol{p}^j, k)$, it holds that $\lim_{j \to \infty} \boldsymbol{s}^j \in \boldsymbol{s}(\boldsymbol{p}, k)$. That this is true follows immediately from Equation B.6, the continuity of $D(a, a', s, 0, \hat{\boldsymbol{p}}(\cdot))$ in $\hat{\boldsymbol{p}}(\cdot)$ and s, and the continuity of $\hat{\boldsymbol{p}}(\cdot)$ in $\boldsymbol{p}$.

Finally, as $D(a, a', s, 0, \hat{\boldsymbol{p}}(\cdot))$ is continuously differentiable in s and $\hat{\boldsymbol{p}}(\cdot)$, and $\hat{\boldsymbol{p}}(\cdot)$ is continuously differentiable in $\boldsymbol{p}$, it follows from (B.4) that the implicit function theorem can be applied to conclude that for any $\boldsymbol{p}^* \in \operatorname{int}(\Delta\Theta)$ and $\boldsymbol{s}^* \in \mathcal{S}$ solving Equation B.6 (with $\theta = 0$), there is a neighborhood of $\boldsymbol{p}^*$ in which there is a continuous solution $\boldsymbol{s}(\boldsymbol{p}, k)$ to Equation B.6 with $\boldsymbol{s}(\boldsymbol{p}, k) = \boldsymbol{s}^*$. Hence, by the aforementioned upper hemi-continuity property, it follows that there exists an equilibrium selection, $\boldsymbol{s}(\boldsymbol{p}, k)$, that is continuous in $\boldsymbol{p}$ on $\Delta\Theta$. $\square$

B.4.2. Proofs for Supplementary Appendix B.2

To prove Proposition B.1 and Proposition B.2, we will first establish a result (Lemma B.2 below) that assures that if certain conditions are satisfied and reputational concerns are sufficiently strong, any known devil is preferred to an unknown angel. The conditions (Condition 1 and Condition 2 below) require that for a class of the PM's strategies, the voter's posterior should not be constant over all actions taken by the PM in equilibrium. We then prove the propositions by verifying that each proposition's hypotheses imply the conditions.

For any $L \geq 0$, let $\mathcal{U}^L$ be the maximum possible expected utility for the voter when she can choose any strategy for the PM subject to a constraint that for all $s \in [-L, L]$, either (i) action $\underline{a}$ must not be taken, or (ii) action $\overline{a}$ must not be taken. From Assumption 1, continuity of the voter's utility in $\boldsymbol{s}$ implies that there exists $L > 0$ such that the voter's payoff is higher with any known biased type than a known congruent type who is subject to the above constraint, i.e.

$$\forall \theta \in \Theta : \; \mathcal{U}(\boldsymbol{\delta_\theta}, 0) > \mathcal{U}^L. \tag{B.8}$$

We let $L > 0$ be some value that satisfies (B.8) and define

$$\mathcal{S}^L := \{\boldsymbol{s} \in \mathcal{S} : s_2, s_N \in [-L, L]\}.$$

Any strategy in $\mathcal{S}^L$ entails type 0 taking action $\underline{a}$ for all $s < -L$ and action $\overline{a}$ for all $s > L$.

[8] Here and in the rest of the proof, we view $\hat{\boldsymbol{p}}(\cdot)$ as an $(N \times T)$-dimensional vector.

Condition 1. *Fix a prior* $\boldsymbol{p}$*. Given any vector* $\boldsymbol{s} \in \mathcal{S}^L$*, were each type* θ *to take action* a_n *if and only if* $s \in (s_n - \theta, s_{n+1} - \theta)$*, the induced posterior,* $\hat{\boldsymbol{p}}(a)$*, would not lead to a constant* $V(\hat{\boldsymbol{p}}(a))$ *across the* a*'s the PM ever takes under* $\boldsymbol{s}$*.*

Condition 1 requires that $V(\cdot)$ be non-constant over the actions taken by the PM given any $\boldsymbol{s} \in \mathcal{S}^L$, including those with $s_2 = s_N$. The next condition weakens the requirement to only those $\boldsymbol{s} \in \mathcal{S}^L$ that have at least three actions taken by the PM (i.e. those with $s_2 < s_N$).

Condition 2. *Fix a prior* $\boldsymbol{p}$*. Given any vector* $\boldsymbol{s} \in \mathcal{S}^L$ *with* $s_2 < s_N$*, were each type* θ *to take action* a_n *if and only if* $s \in (s_n - \theta, s_{n+1} - \theta)$*, the induced posterior,* $\hat{\boldsymbol{p}}(a)$*, would not lead to a constant* $V(\hat{\boldsymbol{p}}(a))$ *across the* a*'s the PM ever takes under* $\boldsymbol{s}$*.*

The key result is:

Lemma B.2. *Fix any prior* $\boldsymbol{p}$*. Suppose either*

1. *Assumption 1 and Condition 1 are satisfied; or*
2. *Assumption 1, Assumption 3, and Condition 2 are satisfied.*

Then there exists a $\hat{k}(\boldsymbol{p})$ *such that* $k > \hat{k}(\boldsymbol{p}) \implies \mathcal{U}(\boldsymbol{p}, k) < \min_\theta \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$ *in any equilibrium.*

Proof. Throughout, we maintain Assumption 1.

Part 1: First consider the case in which Condition 1 is satisfied. Fix any sequence of equilibria as $k \to \infty$, $\boldsymbol{s}(\boldsymbol{p}, k) \in \mathcal{S}$. We claim there exists a $\hat{k}$ such that if $k > \hat{k}$ then either $s_2(\boldsymbol{p}, k) < -L + \theta_1$ or $s_N(\boldsymbol{p}, k) > L + \theta_T$. Suppose not, so there exists a sequence $k_m \to \infty$, with

$$\forall k_m : -L + \theta_1 \leq s_2(\boldsymbol{p}, k_m) \leq s_N(\boldsymbol{p}, k_m) \leq L + \theta_T. \tag{B.9}$$

Then, passing to a subsequence of (k_m) if needed and denoting $s_n^\infty := \lim_{k_m \to \infty} s_n(\boldsymbol{p}, k_m)$ for $n \in \{2, \ldots, N\}$, it follows that $-L + \theta_1 \leq s_2^\infty \leq \cdots \leq s_N^\infty \leq L + \theta_T$. By Condition 1, under $(s_n^\infty)_{n=2}^N$, there exist on-path actions a and a' such that $V(\hat{\boldsymbol{p}}(a)) > V(\hat{\boldsymbol{p}}(a'))$. Thus, by continuity of all the relevant objects, once k_m is large enough, given that each type θ uses cutoff $\boldsymbol{s} - \theta$, no type of the PM will play a' when $s \in (-L, L)$. But this contradicts (B.9). Hence we cannot have an equilibrium with $\boldsymbol{s} \in \mathcal{S}^L$.

Finally, since $\mathcal{U}^L$ provides an upper bound on the payoff the voter can receive if there is some action $a \in \{\underline{a}, \overline{a}\}$ that is never taken when $s \in [-L, L]$, by (B.8) we conclude that when $k > \hat{k}$, $\mathcal{U}(\boldsymbol{p}, k) < \min_\theta \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$.

Part 2: Now consider the case in which Assumption 3 and Condition 2 are satisfied. First note that by Assumption 1, Assumption 3 and continuity, there exists a $B > 0$ such that $\min_\theta \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$ is higher than the payoff from any $\boldsymbol{s} \notin \mathcal{S}_B^L := \{\boldsymbol{s} \in \mathcal{S}^L : s_N - s_2 \geq B\}$. Fix any sequence of equilibria as $k \to \infty$, $\boldsymbol{s}(\boldsymbol{p}, k) \in \mathcal{S}$. We claim there exists a $\hat{k}$ such that if $k > \hat{k}$ then either $\boldsymbol{s}_2(\boldsymbol{p}, k) < -L + \theta_1$, $\boldsymbol{s}_N(\boldsymbol{p}, k) > L + \theta_T$ or $\boldsymbol{s}_N(\boldsymbol{p}, k) - \boldsymbol{s}_2(\boldsymbol{p}, k) < B$. Suppose not, so there exists a sequence $k_m \to \infty$,

$$\forall k_m : -L + \theta_1 \leq s_2(\boldsymbol{p}, k_m) \leq s_N(\boldsymbol{p}, k_m) - B \leq L + \theta_T - B. \tag{B.10}$$

Then, passing to a subsequence of (k_m) if needed and denoting $s_n^\infty := \lim_{k_m \to \infty} s_n(\boldsymbol{p}, k_m)$ for $n \in \{1, \ldots, N+1\}$, it follows that $-L + \theta_1 \leq s_2^\infty \leq \cdots \leq s_N^\infty - B \leq L + \theta_T - B$. By Condition 2, under $(s_n^\infty)_{n=1}^{N+1}$, there exist on path actions a and a' such that $V(\hat{\boldsymbol{p}}(a)) > V(\hat{\boldsymbol{p}}(a'))$. Thus, by continuity of all the relevant objects, once k_m is large enough, given that each type θ uses cutoff $\boldsymbol{s} - \theta$, no type of the PM will play $\boldsymbol{s} \in \mathcal{S}_B^L$. But this contradicts (B.10). By the definition of $\mathcal{S}_B^L$ we conclude that when $k > \hat{k}$, $\mathcal{U}(\boldsymbol{p}, k) < \min_\theta \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$. □

To prove Proposition B.1, we will also need the following lemma, which demonstrates that if all biases are to the right (left), then the reputation from choosing $\underline{a}$ is higher (lower) than from taking action $\overline{a}$. Intuitively, if it is only possible for the PM to be biased in one direction, and if the prior on the PM is non-degenerate, the posterior on the PM's type after observing one extreme action must first order stochastically dominate the other.

Lemma B.3. *Suppose Assumption 2 holds and let $\boldsymbol{p}$ be any non-degenerate prior on Θ. For any $\boldsymbol{s} \in \mathcal{S}$, $V(\hat{\boldsymbol{p}}(\underline{a})) > V(\hat{\boldsymbol{p}}(\overline{a}))$ if $p[\theta] = 0$ for all $\theta < 0$, and $V(\hat{\boldsymbol{p}}(\underline{a})) < V(\hat{\boldsymbol{p}}(\overline{a}))$ if $p[\theta] = 0$ for all $\theta > 0$.*

Proof. We provide the argument for the case where $p[\theta] = 0$ for all $\theta < 0$; the other case is analogous. Let j be such that $\theta_j = 0$. When all types with positive prob. are non-negative, Assumption 2 implies that $V(\hat{\boldsymbol{p}}) > V(\hat{\boldsymbol{p}}')$ if, for all $m \in \{j, \ldots, T\}$, $\sum_{i=j}^m \hat{p}_i \geq \sum_{i=j}^m \hat{p}'_i$, with at least one inequality strict. Now note that for any $\boldsymbol{s} \in \mathcal{S}$ and any $m \in \{j, \ldots, T\}$,

$$\hat{p}_m(\underline{a}) = \frac{p_m((1-\eta)F(s_1 - \theta_m) + \eta/N)}{\sum_{i=j}^T p_i((1-\eta)F(s_1 - \theta_i) + \eta/N)},$$

and

$$\hat{p}_m(\overline{a}) = \frac{p_m((1-\eta)(1 - F(s_N - \theta_m)) + \eta/N)}{\sum_{i=j}^T p_i((1-\eta)(1 - F(s_N - \theta_i)) + \eta/N)}.$$

Since $F(s-\theta_i)$ is decreasing in i for any s, it holds that for all $\boldsymbol{s} \in \mathcal{S}$ and $m \in \{j, \ldots, T\}$,

$$\begin{aligned}\sum_{i=j}^{m} \hat{p}_i(\underline{a}) &= \frac{\sum_{i=j}^{m} p_i((1-\eta)F(s_1-\theta_i)+\eta/N)}{\sum_{i=j}^{T} p_i((1-\eta)F(s_1-\theta_i)+\eta/N)} \\ &\geq \frac{\sum_{i=j}^{m} p_i((1-\eta)(1-F(s_N-\theta_i))+\eta/N)}{\sum_{i=j}^{T} p_i((1-\eta)(1-F(s_N-\theta_i))+\eta/N)} \\ &= \sum_{i=j}^{m} \hat{p}'_i(\overline{a}),\end{aligned}$$

and, if $m < T$, we have equality only if $p_i \in \{0, 1\}$ for all $i \in \{j, \ldots, m\}$. As $\boldsymbol{p}$ is non-degenerate, this inequality is strict for some $m \in \{j, \ldots, T-1\}$. We can then conclude that $V(\hat{\boldsymbol{p}}'(\underline{a})) > V(\hat{\boldsymbol{p}}'(\overline{a}))$. $\square$

Proof of Proposition B.1. We prove the result for the case where $p[0] \geq \beta$, $p^+(\boldsymbol{p}) \geq \alpha$, and $p^-(\boldsymbol{p}) < \varepsilon$; the other case is analogous. By Lemma B.2, it is sufficient to show that Condition 1 is satisfied, which follows if we can show that $V(\hat{\boldsymbol{p}}(\overline{a})) \neq V(\hat{\boldsymbol{p}}(\underline{a}))$ for all $\boldsymbol{s} \in \mathcal{S}^L$.

Let $\alpha, \beta > 0$ satisfy $\alpha + \beta < 1$. Define P_α^β to be the set of all distributions such that $p[\theta] = 0$ for all $\theta < 0$, $p[0] \geq \beta$, and $p^+ \geq \alpha$. Note that, by Lemma B.3, for all $\boldsymbol{p} \in P_\alpha^\beta$ and $\boldsymbol{s} \in \mathcal{S}^L$, $V(\hat{\boldsymbol{p}}(\underline{a})) > V(\hat{\boldsymbol{p}}(\overline{a}))$. Moreover, since P_α^β and $\mathcal{S}^L$ are compact we can define

$$d_\alpha^\beta := \min_{\boldsymbol{p} \in P_\alpha^\beta, \boldsymbol{s} \in \mathcal{S}^L} V(\hat{\boldsymbol{p}}(\underline{a})) - V(\hat{\boldsymbol{p}}(\overline{a})) > 0.$$

Since $V(\cdot)$ is a continuous function on $\Delta\Theta$, and $\Delta\Theta$ is compact, it follows that $V(\cdot)$ is uniformly continuous. Moreover, since the set of priors is compact, $\hat{\boldsymbol{p}}(a)$ is uniformly continuous in the prior $\boldsymbol{p}$ for any a. Then, for any $\boldsymbol{s} \in \mathcal{S}^L$ and any $a \in A$, there exists an $\varepsilon_1(a, \boldsymbol{s}) > 0$ such that

$$|V(\hat{\boldsymbol{p}}(a)) - V(\hat{\boldsymbol{p}}'(a))| < \frac{d_\alpha^\beta}{2}$$

for any priors $\boldsymbol{p}$ and $\boldsymbol{p}'$ for which $|\boldsymbol{p} - \boldsymbol{p}'| < \varepsilon_1(a, \boldsymbol{s})$. Furthermore, since A is finite and $\mathcal{S}^L$ is compact,

$$\varepsilon := \min_{\boldsymbol{s} \in \mathcal{S}^L} \min_{a \in A} \varepsilon_1(a, \boldsymbol{s}) > 0.$$

Consequently, for any $\boldsymbol{p}$ and $\boldsymbol{p}'$ such that $|\boldsymbol{p} - \boldsymbol{p}'| < \varepsilon$, for all $\boldsymbol{s} \in \mathcal{S}^L$ and $a \in A$,

$$|V(\hat{\boldsymbol{p}}(a)) - V(\hat{\boldsymbol{p}}'(a))| < \frac{d_\alpha^\beta}{2}.$$

Now let $\boldsymbol{s} \in \mathcal{S}^L$ and $\boldsymbol{p}$ be any prior such that $p[0] \geq \beta$, $p^+(\boldsymbol{p}) \geq \alpha$ and $p^-(\boldsymbol{p}) < \varepsilon$. We define $\boldsymbol{p}'$ such that $p'[\theta] = 0$ if $\theta < 0$, $p'[\theta] = p[\theta]$ if $\theta > 0$, and $p'[\theta] = \sum_{\theta \leq 0} p[\theta]$. Then since $|\boldsymbol{p} - \boldsymbol{p}'| < \varepsilon$, and $\boldsymbol{p}' \in P_\alpha^\beta$, it holds that

$$V(\hat{\boldsymbol{p}}(\overline{a})) < V(\hat{\boldsymbol{p}}'(\overline{a})) + \frac{d_\alpha^\beta}{2} \leq V(\hat{\boldsymbol{p}}'(\underline{a})) - \frac{d_\alpha^\beta}{2} < V(\hat{\boldsymbol{p}}(\underline{a})).$$

Hence Condition 1 is satisfied, and so the result follows from Lemma B.2. □

The following Lemma will facilitate the proof of Proposition B.2.

Lemma B.4. *Under Assumption 2 and Assumption 4, Condition 2 is satisfied when the prior $\boldsymbol{p}$ is symmetric and has full support.*

Proof. For any s_2 and s_N, with $s_2 < s_N$, there are at least three actions $\underline{a}$, a_n, and $\overline{a}$ that are taken by the PM on the equilibrium path. Upon observing action $a \in \{\underline{a}, a_n, \overline{a}\}$ the updated beliefs are

$$\hat{p}_j(\underline{a}) = \frac{p_j((1-\eta)F(s_2 - \theta_j) + \eta/N)}{\sum_{i=1}^T p_i((1-\eta)F(s_2 - \theta_i) + \eta/N)},$$
$$\hat{p}_j(a_n) = \frac{p_j((1-\eta)(F(s_{n+1} - \theta_j) - F(s_n - \theta_j)) + \eta/N)}{\sum_{i=1}^T p_i((1-\eta)(F(s_{n+1} - \theta_i) - F(s_{n+1} - \theta_i)) + \eta/N)},$$
$$\hat{p}_j(\overline{a}) = \frac{p_j((1-\eta)(1 - F(s_N - \theta_j)) + \eta/N)}{1 - \sum_{i=1}^T p_i((1-\eta)F(s_N - \theta_i) + \eta/N)}.$$

As the distribution of states is symmetric about 0 and single-peaked, $F(s + |\theta|) + F(s - |\theta|)$ is increasing in $|\theta|$ when $s < 0$, decreasing in $|\theta|$ when $s > 0$, and constant in $|\theta|$ when $s = 0$.

Now note that, for any $j \leq \frac{T-1}{2}$, the probability of a PM who chose $\underline{a}$ having bias at least θ_j in magnitude is

$$\begin{aligned}
&\sum_{i=1}^{j} [\hat{p}_i(\underline{a}) + \hat{p}_{T+1-i}(\underline{a})] \\
&= \frac{\sum_{i=1}^j (p_i(1-\eta)(F(s_2 - \theta_i) + \eta/N) + p_{T+1-i}((1-\eta)F(s_2 - \theta_{T+1-i}) + \eta/N))}{\sum_{i=1}^T p_i((1-\eta)F(s_2 - \theta_i) + \eta/N)} \\
&= \frac{\sum_{i=1}^j p_i((1-\eta)(F(s_2 - \theta_i) + F(s_2 + \theta_i)) + 2\eta/N)}{\sum_{i=1}^{\frac{T-1}{2}} p_i((1-\eta)(F(s_2 - \theta_i) + F(s_2 + \theta_i)) + 2\eta/N) + p_{\frac{T+1}{2}}((1-\eta)F(s_2) + \eta/N)}.
\end{aligned}$$

Hence we can conclude that for all $j \leq \frac{T-1}{2}$,

$$s_2 < 0 \implies \sum_{i=1}^{j} (\hat{p}_i(\underline{a}) + \hat{p}_{T+1-i}(\underline{a})) > \sum_{i=1}^{j} 2p_i,$$
$$s_2 \leq 0 \implies \sum_{i=1}^{j} (\hat{p}_i(\underline{a}) + \hat{p}_{T+1-i}(\underline{a})) \leq \sum_{i=1}^{j} 2p_i.$$

Similarly, for any $j \leq \frac{T-1}{2}$, the probability of a PM who chose $\overline{a}$ having bias at least θ_j in magnitude is

$$\sum_{i=1}^{j} (\hat{p}_i(\overline{a}) + \hat{p}_{T+1-i}(\overline{a})) =$$
$$\frac{\sum_{i=1}^{j} p_i((1-\eta)((1-F(s_N-\theta_i)) + (1-F(s_N+\theta_i))) + 2\eta/N)}{\sum_{i=1}^{\frac{T-1}{2}} p_i((1-\eta)((1-F(s_N-\theta_i)) + (1-F(s_N+\theta_i))) + 2\eta/N) + p_{\frac{T+1}{2}}((1-\eta)(1-F(s_N)) + \eta/N)}.$$

Hence, for all $j \leq \frac{T-1}{2}$,

$$s_N > 0 \implies \sum_{i=1}^{j} (\hat{p}_i(\overline{a}) + \hat{p}_{T+1-i}(\overline{a})) > \sum_{i=1}^{j} 2p_i,$$
$$s_N \leq 0 \implies \sum_{i=1}^{j} (\hat{p}_i(\overline{a}) + \hat{p}_{T+1-i}(\overline{a})) \leq \sum_{i=1}^{j} 2p_j.$$

Assumption 2 and the above calculations imply that any partition with $V(\hat{p}(\overline{a})) = V(\hat{p}(\underline{a}))$ and $s_2 < s_N$ must involve $s_2 < 0 < s_N$. Thus, to verify Condition 2, we can restrict attention to partitions with $s_2 < 0 < s_N$.

Let a_n be an action taken with positive probability for which $0 \in [s_n, s_n + 1]$. Since conditional probabilities are martingales, the previous arguments imply that for all $j \leq \frac{T-1}{2}$,

$$\sum_{i=1}^{j} (\Pr[\theta = \theta_i | a < a_n] + \Pr[\theta = \theta_{T+1-i} | a < a_n]) \geq \sum_{i=1}^{j} 2p_i, \text{ and}$$
$$\sum_{i=1}^{j} (\Pr[\theta = \theta_i | a > a_n] + \Pr[\theta = \theta_{T+1-i} | a > a_n]) \geq \sum_{i=1}^{j} 2p_i,$$

with at least one inequality strict. Hence, the probability of a PM who chose a_n having bias at least $|\theta_j|$ in magnitude is

$$\sum_{i=1}^{j}(\hat{p}_i(a_n)+\hat{p}_{T+1-i}(a_n)) < \sum_{i=1}^{j} 2p_i < \sum_{i=1}^{j}(\hat{p}_i(\overline{a})+\hat{p}_{T+1-i}(\overline{a})).$$

Hence, by Assumption 2 and Assumption 4, $V(\hat{\boldsymbol{p}}(a_n)) > V(\hat{\boldsymbol{p}}(\overline{a}))$, as required. □

Proof of Proposition B.2. Assume the hypotheses. Let $A(\boldsymbol{s})$ denote the set of actions taken on the equilibrium path by the PM given $\boldsymbol{s}$. By Assumption 1, Assumption 3 and continuity, there exists $B > 0$ such that $\min_\theta \mathcal{U}(\boldsymbol{\delta_\theta}, 0)$ is higher than the payoff from any $\boldsymbol{s} \notin \mathcal{S}_B^L := \{\boldsymbol{s} \in \mathcal{S}^L : s_N - s_2 \geq B\}$. By Lemma B.2 it is sufficient to show that for any symmetric $\boldsymbol{p}$ there exists an $\varepsilon > 0$ such that, for all $\boldsymbol{q}$ with $|\boldsymbol{p}-\boldsymbol{q}| < \varepsilon$, there does not exist a $\boldsymbol{s} \in \mathcal{S}_B^L$ such that $V(\hat{\boldsymbol{q}}(a))$ is constant on $A(\boldsymbol{s})$. By Lemma B.4 and Condition 2, it follows that

$$d_{\boldsymbol{s}} := \max_{a\in A(\boldsymbol{s})} V(\hat{\boldsymbol{p}}(a)) - \min_{a\in A(\boldsymbol{s})} V(\hat{\boldsymbol{p}}(a)) > 0, \text{ and } d := \min_{s\in \boldsymbol{S}_B^L} d_{\boldsymbol{s}} > 0,$$

where the second part uses the fact that $\mathcal{S}_B^L$ is compact and $V(\cdot)$ is continuous.

Since $V(\cdot)$ is continuous on $\Delta\Theta$ and $\Delta\Theta$ is compact, $V(\cdot)$ is uniformly continuous. Moreover, since the set of priors is compact, for any a, $\hat{\boldsymbol{p}}(a)$ is uniformly continuous in $\boldsymbol{p}$ given $\boldsymbol{s}$. Hence, for any $\boldsymbol{s} \in \mathcal{S}_B^L$ and $a \in A$, there exists an $\varepsilon_1(a, \boldsymbol{s}) > 0$ such that $|V(\hat{\boldsymbol{p}}(a)) - V(\hat{\boldsymbol{p}}'(a))| < \frac{d}{2}$ for any priors $\boldsymbol{p}$ and $\boldsymbol{p}'$ for which $|\boldsymbol{p}-\boldsymbol{p}'| < \varepsilon_1(a, \boldsymbol{s})$. Furthermore, since A is finite and $\mathcal{S}_B^L$ is compact,

$$\varepsilon = \min_{\boldsymbol{s}\in\mathcal{S}_B^L} \min_{a\in A(\boldsymbol{s})} \varepsilon_1(a, \boldsymbol{s}) > 0.$$

Consequently, for any $\boldsymbol{q}, \boldsymbol{p}$ such that $|\boldsymbol{p}-\boldsymbol{q}| < \varepsilon$, for all $\boldsymbol{s} \in \mathcal{S}_B^L$ and all $a \in A$,

$$|V(\hat{\boldsymbol{q}}(a)) - V(\hat{\boldsymbol{p}}(a))| < \frac{d}{2}.$$

Now by the definition of d, for all $\boldsymbol{s} \in \mathcal{S}_B^L$, there exist $a, a' \in A(\boldsymbol{s})$ such that $V(\hat{\boldsymbol{p}}(a)) - V(\hat{\boldsymbol{p}}(a')) \geq d$. Hence, for any $\boldsymbol{q}$ such that $|\boldsymbol{p}-\boldsymbol{q}| < \varepsilon$,

$$V(\hat{\boldsymbol{q}}(a')) < V(\hat{\boldsymbol{p}}(a')) + \frac{d}{2} \leq V(\hat{\boldsymbol{p}}(a)) - \frac{d}{2} < V(\hat{\boldsymbol{q}}(a)).$$

The result follows from Lemma B.2. □

B.4.3. Proofs for Supplementary Appendix B.3

Proof of Proposition B.3. Let $\theta \neq 0$ be such that $\mathcal{U}(\boldsymbol{\delta_\theta}, 0) > \mathcal{U}(\boldsymbol{p}, k)$. We construct an informative equilibrium with two message, 0 and θ. For any $x \in [0, 2]$ define the following announcement strategy: type 0 announces message 0, type θ announces θ with probability $\min\{x, 1\}$ (and announces 0 with complementary probability), and all other types announce θ with probability $\max\{x - 1, 0\}$ (and announce 0 with complementary probability). Denote the voter's beliefs after hearing each message $m \in \{0, \theta\}$ as $\boldsymbol{p}(m, x)$; note that for each m, $p(m, x)$ is continuous in x. We seek an x such that $\mathcal{U}(\boldsymbol{p}(\theta, x), k) = \mathcal{U}(\boldsymbol{p}(0, x), k)$; the voter is then indifferent and we have an informative equilibrium where the candidates' announcement strategies are characterized by x and the voter randomizes with equal probability regardless of messages.

Since $\mathcal{U}(\boldsymbol{p}, k)$ is continuous in $\boldsymbol{p}$ (owing to a suitable selection of equilibria in the policy-making stage), it follows that $\mathcal{U}(\boldsymbol{p}(\theta, x), k)$ and $\mathcal{U}(\boldsymbol{p}(0, x), k)$ are continuous in x. Furthermore,

$$\lim_{x \to 0} \mathcal{U}(\boldsymbol{p}(\theta, x), k) = \mathcal{U}(\boldsymbol{\delta_\theta}, k) > \mathcal{U}(\boldsymbol{p}, k) = \mathcal{U}(\boldsymbol{p}(0, 0), k).$$

Conversely,

$$\mathcal{U}(\boldsymbol{p}(\theta, 2), k) < \mathcal{U}(\boldsymbol{\delta_0}, k) = \mathcal{U}(\boldsymbol{p}(0, 2), k).$$

By the intermediate value theorem, there exists $x \in (0, 2)$ such that $\mathcal{U}(\boldsymbol{p}(\theta, x), k) = \mathcal{U}(\boldsymbol{p}(0, x), k)$. □