Informative Cheap Talk in Elections*

Navin Kartik† Richard Van Weelden‡

September 29, 2014

Abstract

Why do political candidates who are largely office-motivated sometimes espouse views that are non-congruent with those of their electorate? Can non-congruent statements convey any information to voters about what a politician will do if elected, and if so, why would voters elect a politician who makes such statements? We develop a model of credible costless and non-binding communication in electoral campaigns about candidates' policy orientation. The foundation is an endogenous voter preference for a politician who is known to be non-congruent over one whose congruence is sufficiently uncertain. This preference arises because uncertainty about an elected official's policy orientation can generate large policymaking distortions due to politicians' reputation or career concerns. We find that informative cheap talk in elections can either increase or decrease voter welfare relative to uninformative communication. The scope for welfare benefits increases in the strength of politicians' reputation concerns.

*We are grateful to Wiola Dziuda, Alex Frankel, Massimo Morelli, Salvatore Nunnari, Ken Shotts, Stephane Wolton, and various conference and seminar audiences for helpful comments. Teck Yong Tan and Enrico Zanardo provided excellent research assistance. Kartik gratefully acknowledges financial support from the NSF.

†Department of Economics, Columbia University. Email: nkartik@columbia.edu.

‡Department of Economics, University of Chicago. Email: rvanweelden@uchicago.edu.

"I think the American people are looking at somebody running for office and they want to know what they believe . . . and do they really believe it."
— President George W. Bush

1 Introduction

Political candidates want to convince voters to elect them. While campaign strategies involve an array of different tactics, a central component is the discussion of policy-related issues. Through a candidate's speeches, writings, and advertisements, voters form beliefs about the kinds of policies he is likely to implement if elected. There is a significant obstacle, however, as candidates are not bound in any formal sense—e.g. by law—to uphold their campaign stances. It is also difficult to hold a candidate accountable for these stances for at least two reasons. First, policies must adapt to variable circumstances that are hard to monitor. Second, candidates rarely take precise policy positions during campaigns; at most they make broad claims about policy orientations: are they in favor of small government, hawkish on international policy, inclined toward stricter financial regulation, and so on.

The *cheap-talk* nature of electoral campaigns creates an obvious puzzle (Alesina, 1988; Harrington, 1992): wouldn't candidates tend to say whatever it is that is most likely to get them elected, and if so, how is it possible to glean any policy-relevant information from their messages? Notwithstanding, it seems clear empirically that candidates often try to convey different messages during elections; in particular, some candidates pronounce views that are not shared by (the median member of) their electorate.[1] Is all this just "babbling", i.e. uninformative communication that should be ignored by rational voters? And if so, how does it square with the evidence that campaigns do sometimes provide useful information about the kinds of policies candidates will pursue in office (see, e.g., Claibourn (2011) and references therein)?

This paper develops a novel rationale for informative cheap talk in elections about candidates' policy orientation. Section 2 lays out a stylized setting of representative democracy in which a (representative or median) voter elects a politician to whom policy decisions are then delegated. The voter's preferred policy depends on some "state of the world" that the elected politician learns after the election. Political candidates value holding office and also have policy preferences that may either be congruent or non-congruent with that of the voter. Due to career concerns—which may represent either future electoral concerns or concerns

[1] In the context of the 2006 U.S. House elections, Stone and Simas (2010) document substantial heterogeneity in how candidates are perceived relative their own district constituents' average ideology.

about post-political life—the elected politician also benefits from establishing a reputation for congruence through his actions in office.

In this setting, cheap talk in the election is about candidates' policy "types", viz. whether their policy preference is the same as the voter's or not. Casual intuition would suggest that since the voter always prefers to elect a congruent politician over a non-congruent one, cheap talk cannot be informative because every candidate would simply claim to be congruent.

This intuition is wrong. Our key insight, developed in Section 3, is that the voter's expected welfare from the elected politician can be non-monotonic in how likely the politician is to be congruent. Indeed, the voter may prefer to elect a politician who is known to be non-congruent than elect a politician who is congruent with some non-degenerate probability. To put it more colorfully: even though a known angel is always better than a known devil, a known devil may be better than an unknown angel.

Why? The action taken by a policymaker is guided by a combination of his policy preference and the action's reputational value, the latter being determined in equilibrium. As is now familiar (e.g. Canes-Wrone, Herron, and Shotts, 2001; Maskin and Tirole, 2004), reputation concerns generate *pandering*: relative to their own policy preferences, both types of a politician tilt their behavior in favor of actions that are more likely to be chosen by the congruent type. Crucially, the degree of pandering and its welfare consequences depend on the voter's belief about the politician's congruence when he takes office. We establish that, under appropriate conditions, for any non-degenerate such belief, a slight reputation concern generates an (expected) welfare benefit to the voter, but a strong-enough reputation concern induces policy distortions that are so severe that the voter would be better off by instead delegating decisions to a politician who is known to be non-congruent.

The logic underlying this result is simple: while a known non-congruent policymaker will sometimes take actions that the voter would prefer he doesn't, the associated welfare loss may be swamped by the welfare loss generated by a policymaker who has some chance of being congruent but distorts his actions significantly to enhance his reputation. To wit, on the policy issue of whether to go to China, voters can be better served by Richard Nixon (a known anti-communist) than by a president whose preferences may be more moderate, but who is concerned about being perceived as soft on communism.[2] Reputational pandering

[2]For related informational explanations of this episode, see Cukierman and Tommasi (1998), Cowen and Sutter (1998), and Moen and Riis (2010); our emphasis on voter welfare as a function of the belief about the politician is distinct. Note that it is not necessary for our point that the politician who is free from reputation concerns act against his policy bias. The record of Russ Feingold, a former U.S. Democratic senator recognized for being very liberal, provides a good illustration. Feingold was the only senator to vote against the 2001 USA

thus endogenously generates the phenomenon of "a known devil is better than an unknown angel." Of course, a known angel is always better than a known devil. It follows that the voter's welfare is non-monotonic in her belief about the policymaker's congruence.

This non-monotonicity opens an avenue for informative cheap talk during the election. We show in Section 4 that, under appropriate conditions, our model admits *semi-separating* equilibria of the following form: a congruent candidate always announces that he is congruent, whereas a non-congruent candidate sometimes announces congruence and sometimes admits non-congruence. Since a candidate always prefers to be elected with a belief that he may be congruent than with certainty that he is non-congruent, a candidate who reveals himself to be non-congruent must be compensated with a higher probability of winning the election.[3] This is possible in equilibrium because the candidates' behavior ensures that the voter is in fact indifferent between electing a candidate who reveals himself to be non-congruent and electing a candidate whose type she is unsure about.

We emphasize that even though communication is non-binding and costless, informative cheap talk endogenously ties candidates' post-election behavior to their electoral campaign. In a semi-separating equilibrium, a candidate's pronouncement of non-congruence acts as a credible commitment to not pander in his post-election policies, unlike a pronouncement of congruence. Candidates' equilibrium messages can be viewed as amounting to either "You may not (always) agree with me, but you'll know where I stand" or "I share your values." The former spiel has been used successfully by several politicians, perhaps most famously by John McCain who even labeled his presidential campaign bus "Straight Talk Express", and, as suggested by our epigraph, by George W. Bush.

Our model makes a distinction between three kinds of political motivations: office motivation (direct benefits of holding office, including "ego rents"), policy motivation (preferences about which policy is chosen as a function of the state of the world), and reputation motivation (officeholders also care about the electorate's inference about their preference type). The sufficient conditions we provide for informative cheap talk are that reputation motivation is high relative to policy motivation, and office motivation is high relative to reputation motivation. The former guarantees that candidates whose congruence is uncertain when elected will engage in sufficiently detrimental pandering; the latter ensures that non-congruent can-

Patriot Act, was in the minority to vote against authorizing the use of force against Iraq, and was the first senator to subsequently call for the withdrawal of troops; these were all actions in line with his bias. Yet he was also the only Democratic senator to vote against a motion to dismiss Congress' 1998–99 impeachment case against Bill Clinton, an action against his bias.

[3]As discussed in Section 5, the voter may favor a candidate who claims congruence if communication involves costly misrepresentation rather than cheap talk.

didates are willing to reveal their type if it sufficiently increases their election probability.

Since uninformative equilibria always exist in our setting (as in virtually any cheap-talk game), an important question is whether informative electoral campaigns provide higher voter welfare when they exist. Interestingly, the answer depends on the prior about candidates' congruence. For low priors, voter welfare is higher in uninformative equilibria than the aforementioned semi-separating equilibria. On the other hand, the comparison is reversed for moderately high priors.[4] An intuition is that the degree of pandering by the elected politician is non-monotonic—initially increasing and then decreasing—in the voter's belief about his congruence; hence, for low (resp. high) priors, a candidate who announces congruence in a semi-separating equilibrium will pander more (resp. less) if elected than he would in an uninformative equilibrium.

We find that semi-separating equilibria exist—and also benefit the electorate, relative to uninformative equilibria—for a larger set of priors when candidates are more concerned with their reputation. Intuitively, this is because greater reputation motivation induces more pandering by a politician who is elected with uncertainty about his type; consequently, a candidate benefits more from convincing the voter that he will not pander in office. If reputation motivation owes to re-election concerns, this comparative static can be interpreted as saying that (informative) divergence of messages is more likely when re-election concerns are greater. This contrasts with what one may intuit based on models such Wittman (1977) and Calvert (1985) that predict less scope for policy divergence when office motivation is larger.

Section 5 is the paper's conclusion. All formal proofs are contained in the Appendix; a Supplementary Appendix available at the authors' webpages contains additional material.

Related Literature

The benchmark theory of electoral competition, the Hotelling-Downs model (Downs, 1957; Hotelling, 1929), assumes that candidates can credibly commit to the policies they will implement if elected. A number of authors have subsequently questioned the assumption of commitment. In this paper, we take the antithetical approach of assuming that campaign announcements are entirely non-binding. Asymmetric information between candidates and the electorate seems important for non-binding communication to play an indispensable role.[5]

[4]Holding all other parameters fixed, semi-separating equilibria do not exist for sufficiently high priors.

[5]For this reason, symmetric-information models of elections without commitment justly ignore electoral announcements (e.g. Osborne and Slivinski, 1996; Besley and Coate, 1997). We note that even in these settings, non-

However, most existing electoral models with asymmetric information either preclude cheap-talk announcements on the basis that they would be uninformative or allow for it and argue that they cannot be informative in equilibrium.[6]

Harrington (1992) is perhaps the first formal model of informative cheap talk in one-shot elections. Roughly speaking, he assumes that candidates are uncertain about the electorate's preferences and finds that informative—indeed, fully separating—equilibria exist if and only if candidates would prefer to be in office when there is public support for their ideal policy. This mechanism is different from the one we focus on; in particular, the welfare of a representative voter in Harrington's (1992) framework is monotone in the probability that the elected candidate is congruent with the voter, and informative communication cannot arise when candidates are largely office-motivated. Harrington (1993) develops a similar idea to Harrington (1992) but in a setting with multiple elections.

Panova (2014) also studies a multiple-election model in which candidates can convey some information about their policy preferences through cheap talk. In broad strokes, the rationale for informative cheap talk in her setting is that there is no Condorcet winner, i.e. there is no median voter. Interestingly, she finds that informative equilibria can yield lower expected welfare than uninformative equilibria. This possibility also emerges in our setting, albeit through a distinct mechanism.

Kartik and McAfee (2007) develop a model in which some candidates have "character", which means they announce their true position even if that does not maximize their electoral prospects. In an extension, the authors consider the case where announcements are non-binding and costless (*de facto*, only for those office-motivated candidates who do not have character) and voters care solely about final policy. They derive informative equilibria under some conditions. Schnakenberg (2014) analyzes cheap talk in elections with multi-dimensional policy spaces and, under certain symmetry assumptions, constructs "directionally informative" equilibria (cf. Chakraborty and Harbaugh, 2010). The basis for informative communication in our setting is different from either of these papers: we rely on how post-election pandering can induce a voter preference for a politician who is known to be non-congruent over one who may or may not be congruent. In particular, a politician's post-

binding communication can be viewed as a useful device for coordination. However, the role of communication is murky because standard equilibrium analysis could generate the same outcomes without communication; this applies, for example, to the repeated-election model of Aragones, Palfrey, and Postlewaite (2007).

[6]For example, Banks and Duggan (2008) "view campaign promises as cheap talk (and therefore omit them from the[ir] model)"; similarly, Großer and Palfrey (2014) write that they "abstract from any policy promises…which would only result in cheap talk because of their incentive to misrepresent ideal points". Kartik, Squintani, and Tinn (2014) argue that cheap talk should not be informative in their setting.

election behavior is independent of the electoral campaign in both Kartik and McAfee (2007) and Schnakenberg (2014); this is crucially not the case in our analysis.

Naturally, non-binding electoral announcements can also be informative about future policies if the two are linked through direct costs, because announcements are then costly signals; Banks (1990), Callander and Wilkie (2007), and Huang (2010) study such models.

More broadly, our work connects to a number of papers on decision making in the presence of reputational incentives. The idea that reputational incentives can have perverse welfare implications is not new; early contributions such as Scharfstein and Stein (1990), Prendergast (1993), Prendergast and Stole (1996) and Canes-Wrone et al. (2001) focussed on unknown ability. With unknown preferences, as in the current paper, most existing models of "bad reputation" (e.g. Ely and Välimäki, 2003; Morris, 2001; Maskin and Tirole, 2004) focus on how the presence of "bad" types can reduce the welfare of both "good" types and the uninformed player(s). Our work highlights a more severe point: the uninformed player may prefer to face an agent who is known to be "bad" (but consequently has no reputational incentives) rather than face an agent who may be "good" but has reputation concerns.

This property—viz., that a known devil may be preferred to an unknown angel—can only obtain in settings in which reputationally-driven distortions can become sufficiently severe. While this need not always be possible,[7] it is quite natural in many contexts, particularly in delegated decision-making when there is some degree of common interest even among agents with different preferences. Acemoglu, Egorov, and Sonin (2013) have previously demonstrated that reputationally-induced distortions can lead to policy outcomes that are worse than those that would be chosen by a biased but reputationally-insulated politician; see also Fox and Stephenson (2013) and Morelli and van Weelden (2013a,b). Unlike us, these authors do not focus on the voter's welfare as a function of her belief and do not consider the issue of signaling (through cheap talk or otherwise) in electoral campaigns.

A general lesson from our work is that there can be benefits for agents from establishing themselves as "bad" types rather than uncertain types in reputational settings.[8] While we

[7] For example, in Morris's (2001) cheap-talk model, knowing that the agent is biased would lead to uninformative communication, which is clearly weakly worse for the decision-maker than any communication. In Ely and Välimäki (2003), knowing that the mechanic is bad would lead to market shutdown, which is also weakly worse for every (short-lived) consumer than any equilibrium when the mechanic may be good, because consumers always have the choice of taking their outside option. Similarly, in Maskin and Tirole (2004), without reputation concerns, a non-congruent policymaker always takes the worst possible action for the voter.

[8] Bar-Isaac and Deb (2014) discuss non-monotonic reward functions in reputational settings. To put it succinctly, their point is that it may be difficult to determine who the angel is and who the devil is, or that the ordering of angel and devil may be counterintuitive. By contrast, our point is that even when this relationship

focus here on the implications for information revelation in elections, we believe it would also be fruitful to study the phenomenon in other contexts.

2 The Model

We consider a model in which a representative (or median) voter elects a career-minded politician to take an action on her behalf. Candidates for office have policy preferences that may either be congruent or non-congruent with the voter, and candidates benefit from establishing a reputation for congruence.

In more detail: the voter's utility depends on a state of the world, $s \in \mathbb{R}$, and a policy action, $a \in \{\underline{a}, \overline{a}\} \subset \mathbb{R}$, where $\overline{a} > \underline{a}$. The action is chosen by a policymaker (PM, hereafter) who is elected in a manner described below. The elected policymaker chooses a after privately observing s. We assume the state s is drawn from a cumulative distribution F with support $[\underline{s}, \infty)$, where $\underline{s}$ can either be finite or $-\infty$, and that F admits a differentiable density f with $f(s) > 0$ on $(\underline{s}, \infty)$.[9] The voter's utility is maximized when the action matches the state of the world. For convenience, we assume the voter's von-Neumann Morgenstern utility is given by a standard quadratic loss function: $u(a, s) = -(a - s)^2$.

There are two candidates (synonymous with politicians) who compete for office. Each candidate may have one of two policy-preference *types*, denoted $\theta \in \{0, b\}$, where $b > 0$.[10] Each candidate's type is his private information, and each candidate is independently drawn as congruent with ex-ante probability $p \in (0, 1)$.[11] During the election, each candidate i simultaneously sends a cheap-talk (i.e., non-binding and payoff-irrelevant) message $m_i \in \{0, b\}$ about his type. That is, the candidate announces either that he is congruent or non-congruent, and this announcement is made before any information is obtained about the state of the world.The voter observes both messages, updates her beliefs about each candidate i's congruence based on his message to $p_i(m_i)$, and elects one candidate as the PM.

Whichever candidate is elected then observes the state s and chooses the policy action a. After observing the action taken (but before she learns her utility or anything else directly

is entirely intuitive, the known devil can be better than the unknown angel.

[9] We can also allow for the state to be bounded above, but this unnecessarily complicates later exposition.

[10] In the Supplementary Appendix we consider a more general setting that allows for an arbitrary (finite) number of types and policy actions.

[11] It is not important for our results that the ex-ante probability of each candidate being congruent is the same. Moreover, our main themes would be fundamentally unchanged if there were more than two candidates.

about the state), the voter updates her belief about the PM's congruence. Let $\hat{p}(a, p_i)$ denote the posterior on the PM's type after observing a if the PM is believed to be congruent with probability $p_i \in [0, 1]$ when elected. To keep matters simple, we assume that a candidate who is not elected into office receives a fixed payoff normalized to 0.[12] The elected politician derives utility from holding office, the policy he implements as a function of the state, and his final reputation for congruence. Specifically, the elected politician's payoff is

$$c + v_\theta - (a - s - \theta)^2 + kV(\hat{p}), \tag{1}$$

where $k > 0$, $c > 0$, and $v_\theta > 0$ are scalars, and $V : [0, 1] \to \mathbb{R}_+$ is a continuously differentiable and strictly increasing function. We normalize $V(0) = 0$ and $V(1) = 1$. The parameter $c > 0$ captures the direct benefits from holding office: salary, ego rents, etc. The quadratic loss policy-payoff component justifies why we refer to type $\theta = 0$ as congruent and type $\theta = b$ as non-congruent or biased toward action $\overline{a}$. We elaborate on the role of v_θ subsequently; we will use it to equate the payoff for both types of the PM in the absence of reputation concerns.

The function $V(\cdot)$ captures the reputational payoff, scaled by the parameter $k > 0$. The higher is k, the more a politician benefits from generating a better reputation. While politicians may have reputation concerns for a variety of reasons, including for legacy or post-political life, one obvious motive is re-election. Indeed, the reputation function $V(\cdot)$ can be micro-founded by a two-period model in which a second election takes place between the periods. Suppose the challenger in this second election has probability q of being a congruent type, where q is stochastic, drawn from a cumulative distribution V, and publicly observed after the first-period action is taken. Since the candidate who is elected in the second period is electorally unaccountable, the voter's expected payoff in the second period is higher from a candidate who is more likely to be congruent. Hence, she will (rationally) re-elect the PM if and only if $\hat{p} > q$, which implies the PM will be re-elected with probability $V(\hat{p})$. The parameter k would then represent the PM's value from being re-elected.

Figure 1 summarizes the game form. All aspects of the game except the realizations of each θ_i and s are common knowledge. Our solution concept is Perfect Bayesian Equilibrium (Fudenberg and Tirole, 1991), which we refer to as simply equilibrium hereafter. Loosely put, equilibrium requires the behavior of the politicians and the voter to be sequentially rational and beliefs to be calculated by Bayes' rule at any information set that occurs on the equilib-

[12] Analogous results to ours can also be obtained if the unelected candidate derives utility from policy and reputation when out of office, but the analysis becomes more cumbersome without adding commensurate insight.

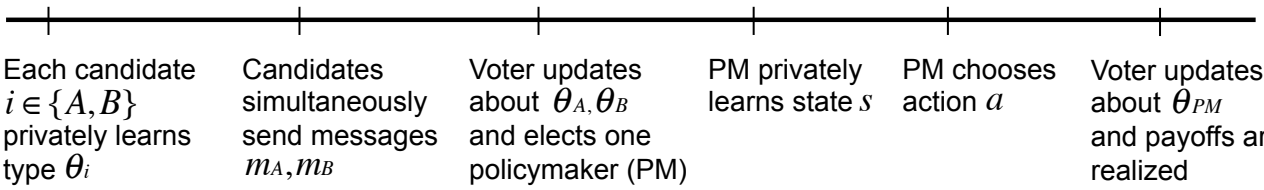

Figure 1 – Summary of the game form.

rium path.[13] As explained in more detail later, we will restrict attention to *symmetric equilibria*, which are equilibria in which both candidates use the same cheap-talk strategy and the voter treats candidates symmetrically in the election. We say that cheap talk is *informative* if there is some on-path message m_i such that $p_i(m_i)$, viz., the voter's belief about θ_i after observing m_i, is different from the prior p. Cheap talk is *uninformative* if it is not informative.

Some preliminaries. From the voter's perspective—which we equate with social welfare—it is optimal to take action $\overline{a}$ if and only if (modulo indifference) $s > s_{FB} := \frac{\overline{a}+\underline{a}}{2}$. In the absence of reputation concerns ($k = 0$), a PM of type $\theta \in \{0, b\}$ would take action $\overline{a}$ if and only if $s > s_\theta := \frac{\overline{a}+\underline{a}}{2} - \theta$. So, in the absence of reputation concerns, a congruent PM would use the first-best threshold whereas a non-congruent PM would take the higher action $\overline{a}$ for a strictly larger set of states.

To provide a cohesive exposition, we maintain throughout the following two assumptions. (Primes on functions denote derivatives, as usual.)

Assumption 1. *The distribution F and the bias b jointly satisfy:*

1. $\underline{s} < \frac{\overline{a}+\underline{a}}{2} - b$;
2. *On the domain* $\left[\frac{\overline{a}+\underline{a}}{2} - b, \infty\right)$, $f(\cdot)$ *is log-convex, i.e.* $\frac{f'(s)}{f(s)} \geq \frac{f'(t)}{f(t)}$ *if* $s > t \geq \frac{\overline{a}+\underline{a}}{2} - b$;
3. $\mathbb{E}\left[s \middle| s \geq \frac{\overline{a}+\underline{a}}{2} - b\right] > \frac{\overline{a}+\underline{a}}{2}$.

[13] More precisely, we will view an equilibrium as consisting of the following objects: (i) a pair of functions specifying each candidate's probability distribution over campaign announcements given his type; (ii) the voter's updating rule specifying a belief on each candidate's type as a function of the realized campaign announcements; (iii) the voter's election rule specifying the probability of electing each candidate as a function of the realized campaign announcements; (iv) the PM's action rule specifying which action he takes in office as a function of the belief with which he is elected, his type, and the state he observes; (v) the voter's updating rule specifying the posterior on the PM's type as a function of the belief he was elected with and the observed policy action.

Assumption 2. $c \geq k$.

Part 1 of Assumption 1 is fairly mild: it says that in the absence of reputation concerns, each action would be taken by both types of the PM. Part 2 is for technical convenience, as it will facilitate certain equilibrium uniqueness and comparative statics. A number of familiar distributions have log-convex densities on their entire domain; our leading example will be the exponential distribution.[14] Part 3 of Assumption 1 is substantive: it is equivalent to assuming that the voter prefers having action $\underline{a}$ taken if and only if $s < \frac{\overline{a}+\underline{a}}{2} - b$ (as done by a non-congruent PM without reputation concerns) to having action $\underline{a}$ for all s; see the proof of Proposition 2. Clearly, given any distribution F (whose support is unbounded above), part 3 of Assumption 1 holds if the bias b is small enough. Alternatively, given any b, part 3 of Assumption 1 holds if the distribution F has enough weight in the right-tail; in particular, it is sufficient that $\mathbb{E}[s] > \frac{\overline{a}+\underline{a}}{2}$. We elaborate on the role of Assumption 1 in Section 3. Assumption 2 says that the direct benefits from office-holding should be sufficiently large compared to reputational concerns; as this will only come into play in Section 4, we elaborate on it there.

Due to their different policy preferences, the two types of a candidate will generally have a different expected utility from holding office even in the absence of any reputation concern. We choose a value of v_θ in expression (1) to avoid this unappealing property; specifically, for each θ, we set v_θ so that type θ's expected payoff from holding office in the absence of reputation concerns $(k = 0)$ and ignoring officeholding benefits $(c = 0)$ would be zero.[15] Since $c > 0$, $k > 0$, and $V(\cdot) \geq 0$, this normalization also ensures that the expected payoff from holding office is strictly higher than from not holding office (which is zero) for either type $\theta \in \{0, b\}$.

Remark 1. Consider $k = 0$. Both types strictly prefer to be elected than not elected, independent of the voter's belief about their type. The elected candidate with type θ uses threshold s_θ

[14] Other well-known examples are the Pareto distribution, and, for suitable parameters, the Gamma and Weibull distributions (both of which subsume the exponential distribution); see Bagnoli and Bergstrom (2005).

[15] Formally, the expected payoff for type θ from holding office given $k = c = 0$ is

$$W_\theta^0 := v_\theta - \int_{\underline{s}}^{\frac{\overline{a}+\underline{a}}{2}-\theta} (\underline{a} - s - \theta)^2 f(s)ds - \int_{\frac{\overline{a}+\underline{a}}{2}-\theta}^{\infty} (\overline{a} - s - \theta)^2 f(s)ds,$$

because type θ uses threshold $s_\theta = \frac{\overline{a}+\underline{a}}{2} - \theta$. We set v_θ so that $W_\theta^0 = 0$, i.e. we set

$$v_\theta = \int_{\underline{s}}^{\frac{\overline{a}+\underline{a}}{2}-\theta} (\underline{a} - s - \theta)^2 f(s)ds + \int_{\frac{\overline{a}+\underline{a}}{2}-\theta}^{\infty} (\overline{a} - s - \theta)^2 f(s)ds. \tag{2}$$

to determine his policy action. Therefore, the voter strictly prefers to elect a candidate who is more likely to be congruent. It follows that electoral campaigns are uninformative. ‖

We will see that the effects of reputation concerns in the policymaking stage create the opportunity for informative cheap talk in the electoral stage.

3 Policymaking with Reputation Concerns

3.1 Equilibrium pandering

We begin by solving for an equilibrium of the policymaking stage. With an abuse of notation, in this section we use $p \in [0,1]$ to denote the probability that the elected PM is congruent. (This belief will eventually be determined as part of the equilibrium of the overall game.) We look for an *interior equilibrium*—hereafter, just equilibrium—of the policymaking "subgame", viz. an equilibrium in which both policy actions are taken with positive probability on the equilibrium path.[16]

Given any belief-updating rule for the voter, the PM's reputational payoff depends only on the action he takes (and not on the state, as this is not observed by the voter). Since the PM's policy utility is supermodular in a and s, any equilibrium involves the PM using a *threshold rule*: the PM of type θ takes action $\overline{a}$ if and only if the state s exceeds some cutoff s^*_θ. The necessary and sufficient conditions for a pair of thresholds $(s^*_0, s^*_b) \in (\underline{s}, \infty)^2$ to constitute an equilibrium are:[17]

$$\overline{p} := \frac{pF(s^*_0)}{pF(s^*_0) + (1-p)F(s^*_b)}, \tag{3}$$

$$\underline{p} := \frac{p(1-F(s^*_0))}{p(1-F(s^*_0)) + (1-p)(1-F(s^*_b))}, \tag{4}$$

$$-(\underline{a} - s^*_0)^2 + kV(\overline{p}) = -(\overline{a} - s^*_0)^2 + kV(\underline{p}), \tag{5}$$

[16] Since the state of the world is unbounded above, there cannot exist an equilibrium in which either PM type chooses action $\underline{a}$ in all states. However, if (and only if) $\underline{s} > -\infty$, a sufficiently large k allows for an equilibrium in which both types take action $\overline{a}$ regardless of the state; the equilibrium is supported by assigning a sufficiently high probability to the PM being non-congruent if he takes the off-path action $\underline{a}$. But such off-path beliefs are inconsistent with standard belief-based refinements in signaling games (Banks and Sobel, 1987; Cho and Kreps, 1987), as the congruent type has a larger incentive to take action $\underline{a}$ than the non-congruent type.

[17] Using part 1 of Assumption 1, it can be shown that in any interior equilibrium, each type must use a threshold in $(\underline{s}, \infty)$.

$$-(\underline{a} - s_b^* - b)^2 + kV(\overline{p}) = -(\overline{a} - s_b^* - b)^2 + kV(\underline{p}). \tag{6}$$

The first two equations represent Bayesian updating: the voter's posterior that the PM is congruent is $\overline{p}$ following action $\underline{a}$ and $\underline{p}$ following $\overline{a}$. (Our notational convention is to use an underlined variable to represent a lower value than the same variable with a bar.) The latter two conditions are the indifference conditions at each type's threshold.

Equations (5) and (6) imply that in any equilibrium, $s_b^* = s_0^* - b$. In other words, the non-congruent type's threshold is pinned down by the congruent type's, and is simply a shift down by the bias. Manipulating (3)–(6), an equilibrium can be succinctly characterized by a single equation of one variable, s_0^*:

$$s_0^* - \frac{\overline{a} + \underline{a}}{2} = \frac{k}{2(\overline{a} - \underline{a})} \left[V\left(\frac{p}{p + (1-p)\frac{F(s_0^* - b)}{F(s_0^*)}} \right) - V\left(\frac{p}{p + (1-p)\frac{1 - F(s_0^* - b)}{1 - F(s_0^*)}} \right) \right]. \tag{7}$$

When $p \in \{0, 1\}$, the right-hand side (RHS) above is zero and hence the unique solution to Equation 7 is $s_0^* = (\overline{a} + \underline{a})/2$. However, when $p \in (0, 1)$, the RHS is strictly positive because $s_b^* = s_0^* - b < s_0^*$. In words, there is a reputational payoff gain to taking action $\underline{a}$ because that action is more likely to come from the congruent type, as the congruent type uses a higher threshold than the non-congruent type.

Proposition 1. *The policymaking stage has a unique equilibrium. In this equilibrium, the congruent type uses a threshold $s_0^*(p, k)$ that solves Equation 7 and the non-congruent type uses a threshold $s_b^*(p, k) = s_0^*(p, k) - b$. Moreover:*

1. *If $p \in (0, 1)$, then*
$$s_0^*(p, k) > \frac{\overline{a} + \underline{a}}{2} = s_0^*(0, k) = s_0^*(1, k).$$
2. *For any $p \in (0, 1)$, $s_0^*(p, k)$ is strictly increasing in k, with*
$$\lim_{k \to 0} s_0^*(p, k) = \frac{\overline{a} + \underline{a}}{2} \text{ and } \lim_{k \to \infty} s_0^*(p, k) = \infty.$$

(All proofs are in the Appendix.)

The uniqueness of equilibrium owes to part 2 of Assumption 1, or more precisely, that the distribution of states, F, has a non-increasing hazard rate on the domain $s \geq \frac{\overline{a} + \underline{a}}{2} - b$.[18] Part 1

[18] Recall that the hazard rate is $f/(1 - F)$. Log-convexity of f on the relevant domain (part 2 of Assumption 1)

of Proposition 1 says that when there is any uncertainty about the PM's type, the equilibrium exhibits *pandering*: both PM types take action $\underline{a}$ for a strictly larger set of states than they would in the absence of reputation concerns, i.e. when $p \in \{0, 1\}$, or just as well, were $k = 0$ rather than our maintained $k > 0$. Part 2 establishes an intuitive monotonicity: the degree of pandering—measured by $s^*_\theta - s_\theta$ for either type θ—is increasing in the reputation concern, k; furthermore, pandering vanishes as $k \to 0$, whereas both types of the PM take action $\underline{a}$ with probability approaching one as $k \to \infty$.[19] In particular, for any $p \in (0, 1)$, once k is large enough, the equilibrium has *over-pandering* in the sense that both types use a threshold above the complete-information threshold of the congruent type, $(\overline{a} + \underline{a})/2$, even though the biased type prefers lower thresholds than the congruent type. This point is analogous to the "populist bias" in Acemoglu et al. (2013). Finally, on a technical note, the implicit function theorem ensures that $s^*_0(p, k)$ is continuously differentiable in both arguments; we will use this property subsequently.

3.2 The voter's welfare from the policymaker

We now study the effect of pandering on voter welfare, and how this depends both on the voter's belief about the PM's congruence and the strength of the PM's reputation concern.

Since the voter's welfare from any PM who uses a threshold rule depends solely on the threshold used and not directly on the PM's preferences, define $U(\tau)$ as the voter's expected payoff when the PM uses threshold τ:

$$U(\tau) := -\int_{\underline{s}}^{\tau} (\underline{a} - s)^2 f(s) ds - \int_{\tau}^{\infty} (\overline{a} - s)^2 f(s) ds.$$

Differentiating,

$$U'(\tau) = (\overline{a} - \underline{a})\,(\overline{a} + \underline{a} - 2\tau)\, f(\tau). \tag{8}$$

Hence, as is intuitive, $U(\cdot)$ is strictly quasi-concave with a unique maximum at $(\overline{a} + \underline{a})/2$, which is the first-best threshold the voter would use if she could observe the state and choose

implies that the hazard rate is non-increasing on this domain (An, 1998). We should note that equilibrium uniqueness is not essential for the rest of Proposition 1; interested readers are referred to the Supplementary Appendix for details.

[19] Pandering also increases in the degree of bias, i.e. s^*_0 is also increasing in b. The reason is that given any equilibrium threshold s^*_0, a higher b increases the difference between the reputations induced by actions $\underline{a}$ and $\overline{a}$: $\overline{p}$ in Equation 3 goes up while $\underline{p}$ in Equation 4 goes down. This is because $s^*_b = s^*_0 - b$, so holding fixed s^*_0, a higher b means more states in which the two types take different actions. Consequently, both types' reputational incentive to take action $\underline{a}$ increases.

policy actions directly.

It follows that when the PM is congruent with probability $p \in [0,1]$, has bias $b > 0$ when non-congruent, and has reputational-concern strength $k > 0$, the voter's expected payoff from having the PM make decisions is

$$\begin{aligned}\mathcal{U}(p,k) &:= pU(s_0^*(p,k)) + (1-p)U(s_b^*(p,k)) \\ &= p\left[U(s_0^*(p,k)) - U(s_0^*(p,k)-b)\right] + U(s_0^*(p,k)-b),\end{aligned} \tag{9}$$

where $s_0^*(p,k)$ is the equilibrium threshold used by the congruent type. We refer to $\mathcal{U}(\cdot)$ as the *voter's welfare* or just welfare, and use subscripts on $\mathcal{U}$ to denote partial derivatives. As $s^*(\cdot)$ is differentiable, $\mathcal{U}(\cdot)$ is also differentiable.[20]

We are interested in properties of the voter's welfare as k and p vary. We begin with the strength of the PM's reputation concern, k.

Lemma 1. *For any $p \in (0,1)$, there is some $\tilde{k}(p) > 0$ such that $\mathcal{U}(p,\cdot)$ is strictly increasing on $(0,\tilde{k}(p))$ and strictly decreasing on $(\tilde{k}(p),\infty)$.*

Lemma 1 implies that when there is uncertainty about the PM's type, a little reputation concern benefits voter welfare but too much harms it. This point is intuitive: were $k = 0$, neither type would distort policymaking, with the congruent type using the voter-optimal threshold and the non-congruent type using a threshold that is too low from the voter's point of view. A small reputation concern, $k \approx 0$, causes both types to increase their thresholds (Proposition 1), which has a first-order welfare benefit when the PM is non-congruent and only a second-order welfare loss when the PM is congruent. When k becomes large, however, pandering becomes extreme; indeed, Proposition 1 says that both types use an arbitrarily large threshold as $k \to \infty$, which is plainly detrimental to welfare. In addition to these limit cases, the strict quasi-concavity assured by Lemma 1 owes to part 2 of Assumption 1, viz. that $f(\cdot)$ is log-convex on the appropriate domain.[21]

Figure 2 depicts welfare as a function of the strength of reputation concern, computed for some representative parameters and three different values of p.[22] Besides illustrating

[20] Since we consider $p \in [0,1]$ and $k > 0$, the domain of $\mathcal{U}$ is $[0,1] \times \mathbb{R}_{++}$. However, in the obvious way we will write $\mathcal{U}(p,0) := \lim_{k\to 0} \mathcal{U}(p,k) = pU\left(\frac{\overline{a}+\underline{a}}{2}\right) + (1-p)U\left(\frac{\overline{a}+\underline{a}}{2} - b\right)$, since $\lim_{k\to 0} s_0^*(p,k) = \frac{\overline{a}+\underline{a}}{2}$. This extends the differentiability of $\mathcal{U}(\cdot)$ to $[0,1] \times \mathbb{R}_+$.

[21] If log-convexity is not assumed, then depending on parameters, some restrictions on the bias parameter b may be needed to assure quasi-concavity of $\mathcal{U}(p,\cdot)$.

[22] This and subsequent figures are computed with F being an exponential distribution with mean 10, $\underline{a} = 0$, $\overline{a} = 2$, and $b = 0.1$.

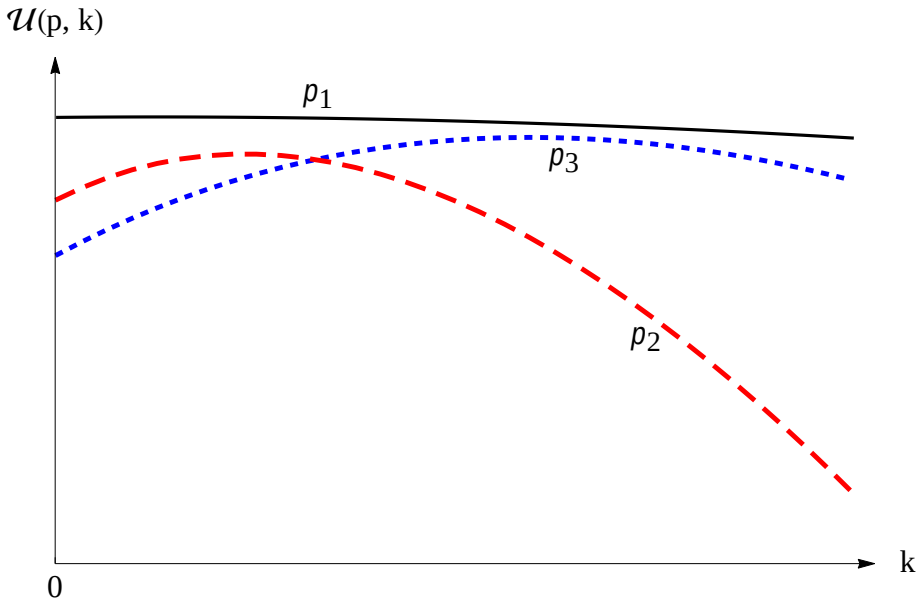

Figure 2 – Voter welfare as a function of PM's reputation concern, with $p_1 > p_2 > p_3$.

Lemma 1, the figure demonstrates another important point: the voter's welfare ranking between PMs with different probabilities of being congruent can turn on the value of k. When k is small, the voter would obviously prefer a PM who is more likely to be congruent: the figure's red (dashed) curve starts out above the blue curve. Once k is sufficiently large, however, welfare can—perhaps counterintuitively—be higher under a PM who is less likely to be congruent: the red (dashed) curve eventually drops below the blue (dotted) curve. The reason is that as $p \to 0$, pandering vanishes, which can be preferable to excess pandering. Of course, welfare approaches first-best as $p \to 1$, as pandering again vanishes but now the PM is very likely congruent: in Figure 2, the black (solid) curve is always above both other curves. Overall, for some values of k, welfare can be non-monotonic in p.

The next result develops the comparative statics of welfare in p and the interaction with k.

Proposition 2. *The voter's welfare, $\mathcal{U}(\cdot)$, has the following properties:*

1. *For all $k > 0$, $\mathcal{U}_p(0,k) > 0$ and $\mathcal{U}(1,k) > \mathcal{U}(p,k)$ for all $p \in [0,1)$.*
2. *For any $p \in (0,1)$, there is a unique $\hat{k}(p) > 0$ such that $\mathcal{U}(p,\hat{k}(p)) = \mathcal{U}(0,0)$. Furthermore, (i) $\mathcal{U}(p,k) < \mathcal{U}(0,0)$ if and only if $k > \hat{k}(p)$, and (ii) $\hat{k}(p) \to \infty$ as either $p \to 0$ or $p \to 1$.*

3. *Consequently, if* $k > \min_{p\in(0,1)} \hat{k}(p)$ *then* $\mathcal{U}(p,k) = \mathcal{U}(0,0)$ *for at least two values of* $p \in (0,1)$*; while if* $k < \min_{p\in(0,1)} \hat{k}(p)$ *then* $\mathcal{U}(p,k) > \mathcal{U}(0,0)$ *for all* $p > 0$.

Part 1 of Proposition 2 implies that $\mathcal{U}(\cdot,k)$ is strictly increasing when $p \approx 0$ and $p \approx 1$, with a global maximum at $p = 1$. The reasons are straightforward; we remark only that a small $p > 0$ yields higher welfare than $p = 0$ both because of a direct effect that the politician may be congruent, and an indirect effect of causing the non-congruent type to use a preferable threshold (recall that $U(\cdot)$ is single-peaked at $(\overline{a} + \underline{a})/2$).

Part 2 of the proposition shows that whenever the reputational incentive is sufficiently strong, the voter's welfare is higher with a PM who is known to be non-congruent $(p = 0)$ than with a PM whose type is uncertain.[23] This "known devil may be better than unknown angel" property is a consequence of the fact that, for any $p \in (0,1)$, pandering gets arbitrarily severe as $k \to \infty$ (Proposition 1, part 2) and that the voter prefers a non-congruent PM with no reputational incentive to a PM who always takes action $\underline{a}$ (Assumption 1, part 3).

Finally, part 3 of Proposition 2 follows from the earlier parts: for any k not too small, as p goes from 0 to 1, $\mathcal{U}(\cdot,k)$ is initially increasing, but then falls below $\mathcal{U}(0,0)$, and eventually increases again up to its maximum. This implies that for any k not too small, $\mathcal{U}(\cdot,k)$ intersects the welfare level provided by a PM who is known to be non-congruent at least twice on the domain $(0,1)$.[24]

Figure 3 illustrates Proposition 2 by graphing $\mathcal{U}(\cdot,k)$ for three different values of k. (The horizontal axis labels $p^*(\cdot)$ will be discussed later.)

It is interesting to note that whenever $\mathcal{U}(\cdot,k)$ is non-monotonic (i.e., once k is sufficiently large), an increase in p—which can be interpreted as an apparently better pool of policymakers, in the sense that a larger fraction of them is congruent—can reduce voter welfare. We will return to this issue after endogenizing campaign communication. Also noteworthy is that whenever $\mathcal{U}(p,k) < \mathcal{U}(0,0)$, it must hold that

$$U(s_0^*(p,k)) < U(s_b^*(p,k)) = U(s_0^*(p,k) - b),$$

or in words, that the voter prefers the equilibrium behavior of the non-congruent PM to that of the congruent PM! This owes to the single-peakedness of $U(\cdot)$: if, in equilibrium, the voter

[23] Note that while we write $\mathcal{U}(0,0)$ to denote the welfare from a PM who is known to be non-congruent, it clearly holds that $\mathcal{U}(0,0) = \mathcal{U}(0,k)$ for any $k > 0$, as there is no pandering no matter the value of k when $p = 0$.

[24] We have not ruled out that $\mathcal{U}(\cdot,k)$ may oscillate multiple times for intermediate p in a way that creates more than two intersections with $\mathcal{U}(0,0)$.

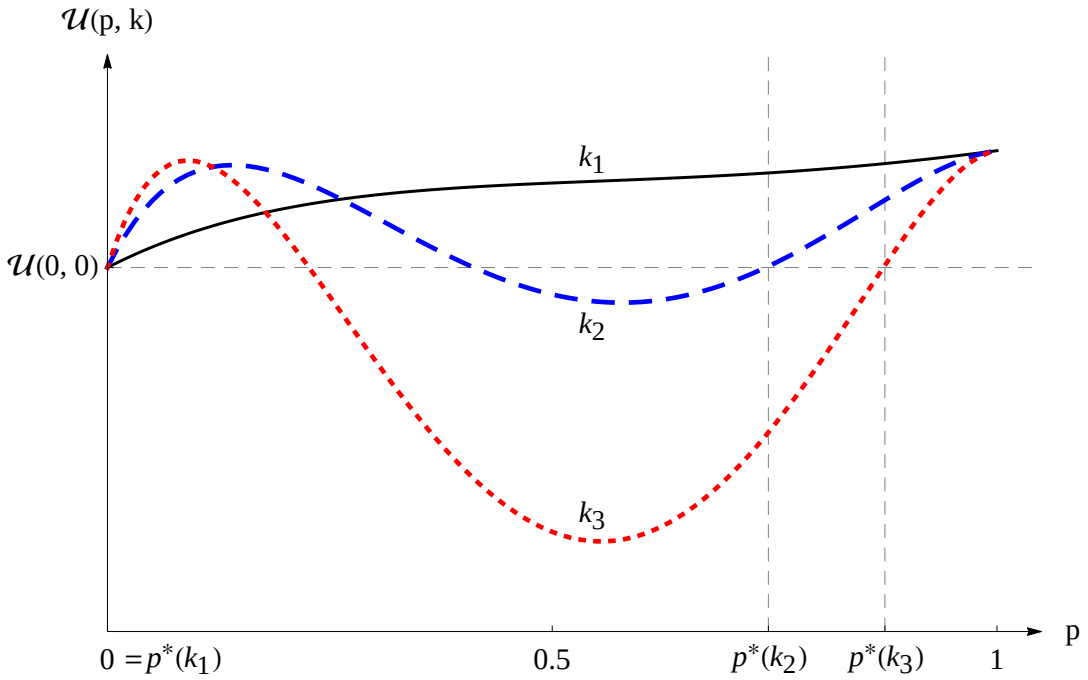

Figure 3 – Voter welfare as a function of her belief, with $k_1 < k_2 < k_3$.

prefers the threshold used by the congruent PM to that of the non-congruent PM, then the non-congruent PM must be using a threshold below the first-best threshold, $(\overline{a} + \underline{a})/2$, which implies that both thresholds are preferred by the voter to $(\overline{a} + \underline{a})/2 - b$, the threshold used by the non-congruent PM when $p = 0$. Proposition 2 thus implies that for any $p \in (0, 1)$, when reputation concerns are sufficiently strong, the voter endogenously—i.e., in equilibrium—prefers the non-congruent type's behavior to the congruent type's behavior, reversing her complete-information ranking over types.

3.3 The policymaker's expected utility

In addition to the voter's welfare, we will also need some properties of the PM's expected payoff. Ignoring the constant c that captures the direct benefits officeholding, a type-θ PM has expected payoff

$$W(\theta, p, k) := v_\theta - \int_{\underline{s}}^{s^*_\theta(p,k)} (\underline{a} - s - \theta)^2 f(s) ds - \int_{s^*_\theta(p,k)}^{\infty} (\overline{a} - s - \theta)^2 f(s) ds$$
$$+ k[F(s^*_\theta(p,k)) V(\overline{p}(p,k)) + (1 - F(s^*_\theta(p,k))) V(\underline{p}(p,k))], \tag{10}$$

where $s^*_\theta(\cdot)$ denotes equilibrium threshold used by type θ and $\overline{p}(\cdot)$ and $\underline{p}(\cdot)$ denote the voter's equilibrium beliefs after observing actions $\underline{a}$ and $\overline{a}$ respectively (see Equation 3 and Equation 4).

Lemma 2. *For any $\theta \in \{0, b\}$, $p \in (0,1)$, and $k > 0$,*

$$0 = W(\theta, 0, k) < W(\theta, p, k) < W(\theta, 1, k) = k.$$

Moreover, for all $p \in (0,1)$ and $k > 0$, $W(0, p, k) > W(b, p, k)$, and hence

$$W(0, p, k) - W(0, 0, k) > W(b, p, k) - W(b, 0, k).$$

The first part of Lemma 2 provides intuitive bounds on $W(\cdot)$. The inequalities say that, no matter his true type, the PM would least (resp., most) prefer the voter's belief putting probability zero (resp., one) on him being congruent.[25] The two equalities owe to $V(0) = 0$, $V(1) = 1$, and how we set v_θ (Equation 2).

The second part of Lemma 2 says that being thought of as non-congruent with some non-degenerate probability is less valuable to a non-congruent PM than to a congruent one, relative to being thought of as non-congruent for sure. The intuition is that, on expectation, for any $p \in (0,1)$, the ex-post reputation of a congruent PM will be higher than that of a non-congruent PM, whereas their reputation will be the same if the prior is 0 (as the voter would simply not update in this case). This limited "single-crossing property" will play an important role. Note that a global single-crossing property does not hold: the congruent type does not benefit more from an arbitrary increase in the voter's belief; to the contrary, Lemma 2 implies that for any $p \in (0,1)$, $W(0,1,k) - W(0,p,k) < W(b,1,k) - W(b,p,k)$.[26]

4 Informative Cheap-Talk Campaigns

We are now ready to study the cheap-talk campaign stage. We revert to using $p \in (0,1)$ for the ex-ante probability of a candidate being congruent. We will assume that if candidate

[25] It is natural to expect $W(\theta, p, k)$ to be increasing in p; while this is true in examples, we are unable to generally rule out non-monotonicities because of how changes in p affect the voter's equilibrium updating rule.

[26] The failure of a global single-crossing condition is related to Mailath and Samuelson's (2001) analysis of the demand for reputation. They find that more competent firms have a greater incentive to purchase an average reputation because they expect to build that reputation up, whereas less competent firms have a greater incentive to purchase either a low or a high reputation to dampen consumers' updating.

$i \in \{A, B\}$ is elected with a belief p_i, then the policymaking stage unfolds as described by the unique interior equilibrium characterized in Proposition 1, with belief p_i in place of p.

Our focus will be on *symmetric equilibria*, which are equilibria in which both candidates use the same strategy and the voter treats candidates symmetrically. More precisely, for $\theta \in \{0, b\}$, let $\mu^\theta \in [0, 1]$ be the probability with which a candidate of type θ sends message $m = 0$, which is interpreted as announcing that he is a congruent type (so he sends message $m = b$ or announces that he is non-congruent with probability $1 - \mu^\theta$). Let $\sigma \in [0, 1]$ denote the probability with which the voter elects the candidate who announces $m = 0$ when the candidates announce different messages. The voter randomizes uniformly over the two candidates when they announce the same message. Hereafter, equilibrium without qualifier refers to a symmetric equilibrium.

Candidate i's (expected) payoff from being elected with a belief $p_i \in [0, 1]$ when his type is θ and the reputation concern is k is given by $c + W(\theta, p_i, k)$, where $W(\cdot)$ was defined in Equation 10. Assumption 2, that $c \geq k$, ensures that office-motivation is sufficiently strong; while this may seem to stack the deck against informative communication, it will turn out to simplify our analysis. More precisely, since $W(\theta, 0, k) = 0 < k = W(\theta, 1, k)$ for either type θ (Lemma 2), Assumption 2 ensures that any candidate would rather be elected with probability one even if believed to be non-congruent than elected with probability one half and believed to be congruent.[27]

As messages are cheap talk, we can, without loss of generality, restrict attention to equilibria in which $\mu^0 \geq \mu^b$, so that a candidate announcing that he is congruent does not decrease the voter's belief about his congruence. An uninformative equilibrium has $\mu^0 = \mu^b$ and always exists. An informative equilibrium has $\mu^0 > \mu^b$. We say an equilibrium is *separating* if $\mu^b = 0$ and $\mu^0 = 1$; an informative equilibrium is *semi-separating* if $\mu^b = 0$ or $\mu^0 = 1$ but not both. We denote by p^m the voter's posterior belief about a candidate who announces message $m \in \{0, b\}$.

The following result establishes that a necessary condition for cheap talk to be informative is that voter welfare in the policymaking subgame cannot depend on which electoral message the PM was elected under.

Lemma 3. *In any informative equilibrium, $\mathcal{U}(p^0, k) = \mathcal{U}(p^b, k)$. Consequently, a separating equilibrium does not exist, and any semi-separating equilibrium has $1 = \mu^0 > \mu^b > 0$.*

[27] If one interprets k as the (discounted) value an incumbent places on re-election and $V(\cdot)$ the probability of re-election as a function of the voter's posterior after observing the policy action, then Assumption 2 says that direct officeholding benefits are larger than the maximum value of re-election. Versions of our results also hold without Assumption 2.

The intuition is straightforward: the voter will elect the candidate from whom she anticipates higher welfare. So if, say, $\mathcal{U}(p^0, k) > \mathcal{U}(p^b, k)$ and both messages are used in equilibrium, candidates would have a higher probability of winning with message 0 than message b. When candidates are sufficiently office motivated—which is ensured by Assumption 2—they would then never use message b, a contradiction. The requirement of voter indifference in an informative equilibrium implies that no message can reveal that a candidate is congruent, as $\mathcal{U}(\cdot, k)$ is uniquely maximized at $p = 1$ (Proposition 2).

Remark 2. Below, we will focus on semi-separating equilibria. We note, however, that in general we cannot rule out the possibility of informative equilibria that are not semi-separating. By Lemma 3, such equilibria must involve both types randomizing. We can establish that such equilibria do not exist when k is sufficiently high and p is sufficiently small, which is a parameter region in which semi-separating equilibria will be shown to exist.[28] Moreover, some of our substantive points below—such as the ambiguous welfare effects of informative communication, and that informative communication is only possible when k is sufficiently large—can be shown to apply to the set of all informative equilibria. ‖

4.1 Semi-separating equilibria

We now examine the conditions under where there is a semi-separating equilibrium with $1 = \mu^0 > \mu^b > 0$. In such an equilibrium, the voter's belief after messages 0 and b are respectively given by

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

Define $p^*(k)$ to be the largest p that makes the voter indifferent between electing a candidate with belief p and a known non-congruent candidate:

$$p^*(k) := \max\{p \in [0,1] : \mathcal{U}(p, k) = \mathcal{U}(0, 0)\}.$$

The function $p^*(\cdot)$ is well defined because $\mathcal{U}(\cdot, k)$ is continuous, $\mathcal{U}(0, k) = \mathcal{U}(0, 0)$, and $\mathcal{U}(1, k) > \mathcal{U}(0, 0)$. Note that for any $p > p^*(k)$, $\mathcal{U}(p, k) > \mathcal{U}(0, k)$. It is also useful to define

$$k^* := \max\{k > 0 : \mathcal{U}(p, k) \geq \mathcal{U}(0, 0) \text{ for all } p \in [0, 1]\}.$$

[28] In canonical signaling games, one proves that multiple types cannot be randomizing over the same set of messages because indifference of any type implies that a "higher" type strictly prefers the "higher" message. As noted in the discussion after Lemma 2, our setting does not have a standard single-crossing property, which is why it may be possible for some parameters to have both types randomizing.

In words, k^* is the largest reputation concern such that the voter cannot be made worse by the PM's pandering, no matter the belief he is elected with, as compared to a known non-congruent. Recalling the function $\hat{k}(\cdot)$ from Proposition 2, $k^* = \min\{\hat{k}(p) : p \in (0,1)\}$.

Lemma 4. *It holds that: (i)* $k^* \in (0,\infty)$*; (ii)* $p^*(k) = 0$ *for all* $k \in (0,k^*)$*; (iii)* $p^*(k^*) > 0$*; (iv)* $p^*(\cdot)$ *is strictly increasing on* $[k^*,\infty)$*; and (v)* $\lim_{k\to\infty} p^*(k) = 1$.

The logic behind the monotonicity of $p^*(\cdot)$ can be understood by comparing the k_2 and k_3 curves in Figure 3. As k increases, pandering becomes more severe, and so $\mathcal{U}(p,k) < \mathcal{U}(0,k)$ for a wider range of p. This property leads to our main result about informative cheap talk.

Proposition 3. *A semi-separating equilibrium exists if and only if* $k \geq k^*$ *and* $p \in (0,p^*(k))$*. In any such equilibrium,* $1 = \mu^0 > \mu^b > 0$*,* $\mathcal{U}(p^0,k) = \mathcal{U}(0,0)$*, and* $\sigma \in (0,1/2)$*. Moreover:*

1. *The larger is* k*, the larger the set (in set-inclusion sense) of priors for which a semi-separating equilibrium exists.*

2. *For any* p*, there is a semi-separating equilibrium if and only if* k *is sufficiently large.*

The logic underlying the characterization of semi-separating equilibria in Proposition 3 can be seen using Figure 3. When k is sufficiently small (k_1 in the figure), $\mathcal{U}(p,k)$ is always strictly above $\mathcal{U}(0,0)$ for all $p > 0$, hence there is no informative strategy of the candidate that can leave the voter indifferent after both messages. Once k is sufficiently large (k_2 or k_3 in the figure), for any prior $p \in (0,p^*(k))$, there is a (unique) semi-separating strategy that induces beliefs $p^b = 0$ and $p^0 = p^*(k) < 1$. The voter is then willing to randomize between the candidates when they make distinct announcements. Since a candidate prefers to be elected with uncertainty about his type rather than with the voter being sure that he is non-congruent, the mixing of a non-congruent candidate must be sustained by $\sigma < 1/2$, i.e. the voter must favor a candidate who pronounces non-congruence over a candidate who pronounces congruence when the two candidates make distinct announcements. Given that $p^b = 0 < p^0 < 1$, Lemma 2 assures that when the non-congruent type is willing to randomize, the congruent type has a strict incentive to announce congruence.

Figure 4 graphs $p^*(\cdot)$ and depicts the comparative statics noted in parts 1 and 2 of Proposition 3, both of which build on Lemma 4. Part 2 of the proposition represents our central conclusion: given any (non-degenerate) p, informative cheap talk is possible when reputation concerns are sufficiently strong. Intuitively, this owes to the fact that for any non-degenerate belief, a sufficiently large k results in such severe pandering by a PM who is elected with

that belief that the voter would prefer to have a known non-congruent PM in office. It bears emphasis that even as k increases, the office-motivation component continues to dominate candidates' preferences during the election, because c also increases by Assumption 2.

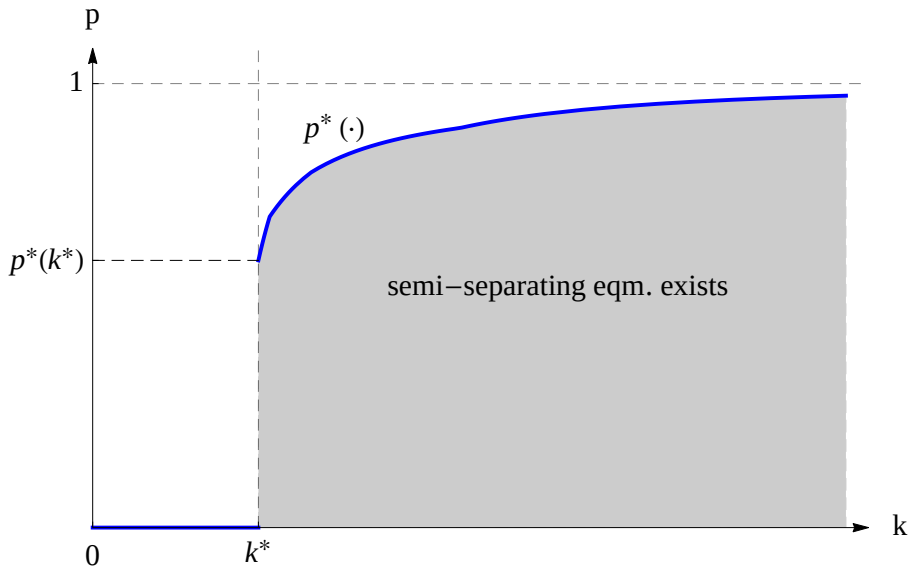

Figure 4 – Existence of semi-separating equilibrium.

It is noteworthy that in a semi-separating equilibrium, a non-congruent candidate is indifferent over announcements when he doesn't know his opponent's announcement, but he would not be indifferent after observing his opponent's announcement. In other words, the equilibrium has the realistic feature that a candidate's best response depends on his opponent's electoral message; given the voter's strategy, each candidate has a greater incentive to claim to be congruent if the other candidate is also claiming congruence.[29] This property is not shared by other models of informative cheap talk in elections (e.g. Kartik and McAfee, 2007; Schnakenberg, 2014).

We should also note that when $k > k^*$, there will be more than one semi-separating equilibrium for a range of priors, due to the multiple-intersection property established in Proposition 2 (part 3). For example, when $k = k_2$ or $k = k_3$ in Figure 3, there is a range of p, viz. those below the first positive intersection of the respective curve with $\mathcal{U}(0,0)$, in which there are exactly two semi-separating equilibria: p^0 can either be the belief corresponding to the lower

[29]This implies that assumptions about timing are important, as the prescribed strategies would not be an equilibrium if candidates made their announcements sequentially. Nonetheless, informative cheap talk is also possible under sequential communication; specifically, an equilibrium in which both candidates' play as in Proposition 3 can be sustained by having the voter treat the candidates asymmetrically (as is natural once timing creates an inherent asymmetry between candidates).

or the higher intersection. These equilibria are payoff equivalent for the voter, however, as the voter's expected payoff in any semi-separating equilibrium is simply $\mathcal{U}(0,0)$.[30]

In any semi-separating equilibrium, the voter's posterior after a candidate announces congruence, p^0, is not affected by small changes in prior, p — the only effect is to alter a non-congruent candidate's mixing probability, μ^b. An increase in p decreases the probability of observing an announcement of non-congruence not only because of the direct effect that a candidate is ex ante less likely to be congruent but also because μ^b is decreasing in p (to keep p^0 constant).

Importantly, the welfare effects of informative communication depend on the prior. In an uninformative equilibrium, voter welfare is $\mathcal{U}(p,k)$; in a semi-separating equilibrium it is $\mathcal{U}(0,0)$. When $k > k^*$, Proposition 2 implies that there necessarily exists a region of priors within $(0, p^*(k))$ where $\mathcal{U}(\cdot,k) > \mathcal{U}(0,0)$ and one where $\mathcal{U}(\cdot,k) < \mathcal{U}(0,0)$. Thus:

Corollary 1. *Cheap-talk campaigns have the following welfare properties:*

1. *Assume $k > k^*$, so that a semi-separating equilibrium exists. Relative to uninformative communication, there is a non-degenerate interval of priors in which any semi-separating equilibrium strictly improves voter welfare, and a non-degenerate interval of priors in which any semi-separating equilibrium strictly reduces voter welfare.*

2. *For any k and p, there is an equilibrium in which the voter's payoff is at least $\mathcal{U}(0,0)$.*

Part 1 of the result says that campaigns—in the sense of their semi-separating cheap-talk equilibria—can either be welfare enhancing or welfare decreasing.[31] As suggested by Figure 3, a typical pattern is that semi-separating equilibria are deleterious to welfare for low priors, beneficial for moderate priors, and non-existent for high-enough priors. More succinctly: campaigns (can) help the voter when there is sufficient uncertainty about the candidates.

[30] In general, the candidates' ranking across semi-separating equilibria is ambiguous. First, $W(\theta,p,k)$ may not be increasing in p, as discussed in fn. 25. But suppose it is for both types (as may be viewed as a "typical case") , and consider two semi-separating equilibria with respective beliefs after message 0, p^0 and $\tilde{p}^0 > p^0$. A congruent candidate faces a tradeoff between the two equilibria: on the one hand, he prefers being elected with belief $\tilde{p}^0$ than with belief p^0; on the other hand, his probability of election is necessarily smaller under $\tilde{p}^0$ because the non-congruent candidate must be compensated for sending message b in this equilibrium with a higher election probability, as he too prefers being elected with $\tilde{p}^0$ than with p^0.

[31] It is worth noting that for sufficiently low priors, any informative equilibrium—semi-separating or not (cf. Remark 2)—must decrease welfare relative to an uninformative equilibrium. To see this, recall that $\mathcal{U}(p,k)$ is strictly increasing for small p (Proposition 2). Since $p^b < p$ in an informative equilibrium, it holds for small p that $\mathcal{U}(p^0,k) = \mathcal{U}(p^b,k) < \mathcal{U}(p,k)$, where the equality is by Lemma 3.

The second part of Corollary 1 identifies a sense in which electoral campaigns can ensure that the voter is protected against too much policy pandering. Without informative cheap talk, the voter's welfare would be $\mathcal{U}(p,k)$, which can be much lower than $\mathcal{U}(0,0)$ due to acute pandering by the elected PM. But it is precisely in this parameter region that a semi-separating equilibrium exists in the election, which provides the voter with welfare $\mathcal{U}(0,0)$. Thus, while informative cheap talk quite crucially relies on the *possibility* of severe pandering, in (a semi-separating) equilibrium, the actual extent of pandering by the elected PM will be limited.

There is another sense in which electoral campaigns can protect the voter. Changes in p—either an increase or decrease—can reduce $\mathcal{U}(p,k)$, which harms the voter in the absence of cheap talk. Plainly, however, such changes do not affect voter welfare in semi-separating equilibria; they only alter the equilibrium mixing probability of non-congruent candidates. It follows that when $\mathcal{U}(p,k) < \mathcal{U}(0,0)$, semi-separating equilibria neutralize (small) adverse effects of changes in the pool of politicians. In particular, when $\mathcal{U}(p,k) < \mathcal{U}(0,0)$, cheap talk can nullify the "perverse" finding noted at the end of Subsection 3.2 that an apparently better pool of politicians (i.e., higher p) may reduce voter welfare. On the flip side, when $\mathcal{U}(p,k) > \mathcal{U}(0,0)$, semi-separating equilibria can also preclude harnessing the beneficial effects of changes in the politician pool.

We next relate the welfare effects of informative campaigns with the strength of reputation concerns. Define, for any $k > 0$,

$$P^k := \{p \in (0,1) : \mathcal{U}(p,k) < \mathcal{U}(0,0)\}$$

as the set of priors for which a semi-separating equilibrium exists that strictly improves voter welfare relative to uninformative communication. Corollary 1 assured that for $k > k^*$, $P^k \neq \emptyset$.

Proposition 4. *Cheap-talk campaigns have the following welfare comparative statics:*

1. *For any k_1, k_2 such that $k_2 > \max\{k^*, k_1\}$, $P^{k_1} \subsetneq P^{k_2}$.*
2. $\lim_{k\to\infty} P^k = (0,1)$.
3. *For any k_1, $p \in P^{k_1}$, and $k > k_1$, $\frac{\partial}{\partial k}\left[\mathcal{U}(0,0) - \mathcal{U}(p,k)\right] > 0$.*

The first part of the result says that the higher is k (above k^*) the larger is the set of priors for which semi-separating equilibria are welfare enhancing. In fact, for any prior $p \in (0,1)$, semi-separating equilibria exist and increase voter welfare (relative to uninformative communication) if k is large enough, because then $\mathcal{U}(p,k) < \mathcal{U}(0,0)$ (Proposition 2, part 2); this

explains the second part of Proposition 4. Finally, part 3 is because the voter's welfare is decreasing in k when $\mathcal{U}(p,k) < \mathcal{U}(0,0)$ (Lemma 1); thus, if semi-separating equilibria are welfare enhancing, then greater reputation concerns amplify their welfare gains.

4.2 A limiting case

Let us briefly consider what happens if candidates are so office-motivated that during the election they simply maximize the probability of getting elected. Loosely put, it is as if $c = \infty$ in our baseline. Of course, once elected, c is irrelevant, and so the behavior of the elected PM is unchanged.

Proposition 5. *Assume candidates maximize the probability of being elected, while still behaving as before in post-election policymaking. Then:*

1. *For any k and p, there is an informative cheap-talk equilibrium if and only if there are p' and p'' such that $p \in (p', p'')$ and $\mathcal{U}(p', k) = \mathcal{U}(p'', k)$.*
2. *For any p and any $\varepsilon > 0$, there is $\overline{k} > 0$ such that for all $k > \overline{k}$, there is an informative equilibrium in which voter welfare is larger than $\mathcal{U}(1,0) - \varepsilon$.*

To understand this result, first observe that Lemma 3 continues to apply, in particular $\mathcal{U}(p^b, k) = \mathcal{U}(p^0, k)$ in any informative equilibrium, because candidates' post-election behavior has not changed. The key difference with our earlier analysis is that both candidates are now willing to randomize over messages if (and only if) $\sigma = 1/2$, i.e. so long as electoral prospects don't depend on which message a candidate sends. Thus, a pair of beliefs (p^0, p^b) can be sustained in an informative equilibrium if and only $p^b < p < p^0$ and $\mathcal{U}(p^b, k) = \mathcal{U}(p^0, k)$, which explains part 1 of Proposition 5.

Part 2 of the proposition says that for any (non-degenerate) prior, when reputation concerns are sufficiently strong, there is an informative equilibrium that yields approximately first-best voter welfare. The reason is that as $k \to \infty$, there is $\hat{p}(k) \to 0$ such that $\hat{p}(k)$ is a local maximizer of $\mathcal{U}(\cdot, k)$ and $\mathcal{U}(\hat{p}(k), k) \to \mathcal{U}(1,0)$. This point can be seen in Figure 3 by comparing voter welfare at the local maximum with that at the global maximum for both the k_2 and k_3 curves. Intuitively, as $k \to \infty$, a PM who is elected with a suitably low belief is expected to deliver close to the first-best welfare because the reputational concern then disciplines a non-congruent PM into using the first-best threshold. Since, for any $p \in (0,1)$, $\mathcal{U}(p,k) < \mathcal{U}(0,0)$ for all large enough k, it follows that when k is large enough, candidates can suitably mix to generate $p^b < p < p^0$ with $\mathcal{U}(p^b, k) = \mathcal{U}(p^0, k) \approx \mathcal{U}(1,0)$.

We view Proposition 5 as reinforcing the message from our main analysis: when policy pandering can get severe due to reputation concerns, but office-motivation still looms large, cheap talk can not only be informative but also substantially improve voter welfare. Note that the equilibria of Proposition 5 can be viewed as ε-equilibria of our baseline model when c, the direct benefit from office, is sufficiently large.

5 Conclusion

Elections are often flush with candidates' talk about their general views, but short on concrete policy proposals. This makes it difficult for voters to hold politicians accountable for their electoral campaigns. Nevertheless, candidates' communications during major elections elicit a tremendous amount of attention. *Prima facie*, this appears puzzling: given the lack of accountability, wouldn't candidates tend to say whatever it is that would maximize their electoral prospects, resulting only in "babbling" or uninformative communication?

This paper has developed a simple rationale for why non-binding electoral communication can be informative. We have argued that while voters prefer candidates who are known to have preferences that match their own, they also dislike uncertainty about politicians' preferences, because uncertainty generates reputationally-motivated policy distortions in office no matter a policymaker's true preferences. Sufficiently severe distortions bear out the adage that a known devil is preferred to an unknown angel. Under suitable conditions, this phenomenon allows for informative communication: it becomes credible for a politician to sometimes reveal that he has different policy preferences from those of the (median or representative) voter, because this acts as an endogenous commitment to not pander if elected.

When reputation concerns stem from electoral accountability, this paper contributes to a literature highlighting how accountability can induce undesirable pandering by officeholders. Plainly, there are a number reasons outside our model that electoral accountability is desirable. The novel lesson from our analysis is that cheap talk in elections can mitigate the distortions induced by electoral accountability.

We close by mentioning some additional issues.

More types or policies. We have focussed on a simple model where the set of politicians' policy types and the policy space are both binary. In the Supplementary Appendix, we extend the analysis to more than two types and policies. The main insight is that under reasonably broad conditions, a voter will prefer certainty about the politician's type—regardless of what

that type is—to sufficient uncertainty whenever the politician's reputation concern is sufficiently strong. Although the analysis of communication appears intractable in general, we discuss some richer specifications that also confirm the possibility of informative cheap talk.

Costly signaling. The assumption that campaign communication is cheap talk stacks the deck against informative communication. Suppose instead that a candidate of type $\theta \in \{0, b\}$ bears a utility cost $\delta \geq 0$ if he sends message $b-\theta$. This cost could represent personal integrity, the difficulty of crafting a credible but insincere campaign stance, or a reduced-form expected cost of being caught in a "web of lies." When $\delta > 0$, messages are no longer cheap talk, but they remain non-binding. An interesting observation is that under our maintained assumptions, neither is the existence of a semi-separating equilibrium nor the corresponding voter welfare altered by small changes in δ. The reason is a familiar property of mixed-strategy equilibria: candidates' behavior in semi-separating equilibria are pinned down by voter indifference; the only effect of small changes in δ is to alter the voter's randomization probability (when the two candidates announce distinct messages) to preserve a non-congruent candidate's indifference. Notice, though, that when $\delta > 0$, a semi-separating equilibrium is compatible with $\sigma > 1/2$, viz. the voter can favor a candidate who claims to be congruent.

The reputation function. A common assumption, which we have also made, is that the reputational benefit for the policymaker, $V(\cdot)$, is strictly increasing in the voter's belief that the policymaker is congruent. However, we have seen that this can induce policymaking behavior which leads the voter to prefer a policymaker with a lower probability of being congruent. If $V(\cdot)$ represents post-political life benefits or is otherwise not tied to future policymaking, then there is no tension between the monotonicity assumption and the non-monotonicity conclusion. However, if $V(\cdot)$ represents a payoff from re-election, then can one square the assumption with its consequence? One micro-foundation is that politicians face a two-term limit and compete against a randomly-drawn challenger after their first term, in a manner similar to that described in Section 2. Then, even though the voter's welfare from electing a new policymaker may be non-monotone in the probability of his congruence, the voter's welfare from re-electing an incumbent is monotone in that probability. More generally, though, what if the voter's welfare from re-electing an incumbent is also non-monotone in the probability of congruence, e.g. because there are no term limits? In ongoing work, we are analyzing the implications of requiring, roughly speaking, a "functional fixed point" between a policymaker's reputational value and voter welfare.

Appendix: Proofs

Proof of Proposition 1. The discussion preceding the proposition explained why Equation 7 characterizes (interior) equilibria.

Step 1: We first establish that Equation 7 has a unique solution s_0^*. Since

$$\frac{1-F(s_0^*-b)}{1-F(s_0^*)} \geq 1 \geq \frac{F(s_0^*-b)}{F(s_0^*)}, \tag{11}$$

the right-hand side (RHS) of Equation 7 is non-negative for all s_0^*. The left-hand side (LHS) is non-negative if and only if $s_0^* \geq (\overline{a}+\underline{a})/2$. Hence, any solution has $s_0^* \geq (\overline{a}+\underline{a})/2$; we restrict attention in the remainder of the proof to this domain. Existence of a solution follows from continuity and the observations that the RHS of Equation 7 is bounded in s_0^* while the LHS tends to ∞ as $s_0^* \to \infty$. For uniqueness, it is sufficient to show that the RHS of Equation 7 is non-increasing, because the LHS is strictly increasing.

Differentiating the RHS of Equation 7 with respect to s_0^* and using the shorthand $\alpha \equiv (1-p)/p$, $s \equiv s_0^*$, $G(s) \equiv F(s-b)/F(s)$, and $H(s) \equiv (1-F(s-b))/(1-F(s))$ yields

$$\begin{aligned} RHS' &= V'\left(\frac{1}{1+\alpha G(s)}\right)\left[-1(1+G(s))^{-2}\alpha G'(s)\right] - V'\left(\frac{1}{1+\alpha H(s)}\right)\left[-1(1+H(s))^{-2}\alpha H'(s)\right] \\ &= \alpha\left[V'\left(\frac{1}{1+\alpha H(s)}\right)\frac{H'(s)}{(1+\alpha H(s))^2} - V'\left(\frac{1}{1+\alpha G(s)}\right)\frac{G'(s)}{(1+\alpha G(s))^2}\right], \end{aligned} \tag{12}$$

where

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

Since $V'(\cdot) > 0$, expression (12) is weakly negative if $G'(s) \geq 0 \geq H'(s)$, which is equivalent to

$$\min\left\{\frac{F(s)}{F(s-b)}, \frac{1-F(s)}{1-F(s-b)}\right\} \geq \frac{f(s)}{f(s-b)},$$

which, because of (11), simplifies to

$$\frac{f(s-b)}{1-F(s-b)} \geq \frac{f(s)}{1-F(s)}.$$

The above inequality holds for all $s \geq (\overline{a}+\underline{a})/2$ because f is log-convex on the domain $\left[\frac{\overline{a}+\underline{a}}{2}-b, \infty\right)$ (part 2 of Assumption 1) and hence has a non-increasing hazard rate on this domain (An, 1998).

Step 2: We now prove parts 1 and 2 of Proposition 1. Let the unique solution to Equation 7 be denoted $s_0^*(p,k)$. Since both sides of Equation 7 are continuously differentiable in all arguments, the

implicit function theorem ensures that $s_0^*(p,k)$ is continuously differentiable in p and k.

For part 1, note that because $k > 0$ and $V(\cdot)$ is strictly increasing, the RHS of Equation 7 is strictly positive for any $p \in (0,1)$. Therefore, $s_0^*(p,k) > (\overline{a}+\underline{a})/2$ for any $p \in (0,1)$. However, when $p \in \{0,1\}$ the RHS is equal to 0, and hence $s_0^*(0,k) = s_0^*(1,k) = (\overline{a}+\underline{a})/2$.

For part 2, fix an arbitrary $p \in (0,1)$. First note that $s_0^*(p,k)$ is increasing in k because the RHS of Equation 7 is non-increasing in s_0^* (by Step 1) and increasing in k. That $s_0^*(p,k) \to (\overline{a}+\underline{a})/2$ as $k \to 0$ follows from the fact that the RHS of Equation 7 tends to 0 as $k \to 0$. Conversely, that $s_0^*(p,k) \to \infty$ as $k \to \infty$ follows from the fact that, for any s_0^*, the RHS tends to ∞ as $k \to \infty$. □

Proof of Lemma 1. Partially differentiating Equation 9 and suppressing the arguments of $s_0^*(\cdot)$,

$$\begin{aligned}
\mathcal{U}_k(p,k) &= [pU'(s_0^*) + (1-p)U'(s_0^* - b)]\frac{\partial s_0^*}{\partial k} \\
&\propto pU'(s_0^*) + (1-p)U'(s_0^* - b) \\
&= (\overline{a}-\underline{a})[p(\overline{a}+\underline{a}-2s_0^*)f(s_0^*) + (1-p)(\overline{a}+\underline{a}-2s_0^*+2b)f(s_0^*-b)] \\
&\propto \left(\frac{\overline{a}+\underline{a}}{2} - s_0^*\right) + \frac{(1-p)bf(s_0^*-b)}{pf(s_0^*) + (1-p)f(s_0^*-b)},
\end{aligned} \tag{13}$$

where the first proportionality uses $\frac{\partial s_0^*}{\partial k} > 0$ (Proposition 1), the equality uses Equation 8, and the second proportionality obtains from a division by $2(\overline{a}-\underline{a})(pf(s_0^*) + (1-p)f(s_0^*-b)) > 0$.

Fix any $p \in (0,1)$. Expression (13) is strictly positive as $k \to 0$ because $s_0^* \to \frac{\overline{a}+\underline{a}}{2}$ as $k \to 0$ (Proposition 1), whereas the last fraction in (13) is strictly positive and bounded away from zero as $s_0^* \to \frac{\overline{a}+\underline{a}}{2}$. Analogously, (13) is strictly negative for large k because $s_0^* \to \infty$ as $k \to \infty$ while the last fraction is always less than one. Therefore, it suffices to show that expression (13) has a unique zero, i.e. that

$$s_0^* - \frac{\overline{a}+\underline{a}}{2} = \frac{(1-p)bf(s_0^*-b)}{pf(s_0^*) + (1-p)f(s_0^*-b)}$$

has a unique solution. The LHS is strictly increasing in s_0^*. It is straightforward to check by differentiation that the RHS is non-increasing in s_0^* if $f'(s_0^*)f(s_0^*-b) \geq f(s_0^*)f'(s_0^*-b)$, which is assured because $f(\cdot)$ is log-convex on $\left[\frac{\overline{a}+\underline{a}}{2} - b, \infty\right)$ (part 2 of Assumption 1), $s_0^* \geq (\overline{a}+\underline{a})/2$, and $b > 0$. □

Proof of Proposition 2. We prove each part of the result in sequence.

Part 1: Partially differentiating Equation 9 with respect to p yields

$$\begin{aligned}
\mathcal{U}_p(p,k) &= U(s_0^*(p,k)) - U(s_0^*(p,k)-b) + p\frac{\partial s_0^*(p,k)}{\partial p}\left[U'(s_0^*(p,k)) - U'(s_0^*(p,k)-b)\right] \\
&\quad + U'(s_0^*(p,k)-b)\frac{\partial s_0^*(p,k)}{\partial p} \\
&= \frac{\partial s_0^*(p,k)}{\partial p}\left[(\overline{a}-\underline{a})(\overline{a}+\underline{a}-2s_0^*(p,k)) + 2(1-p)(\overline{a}-\underline{a})b\right] + U(s_0^*(p,k)) - U(s_0^*(p,k)-b),
\end{aligned}$$

where the second equality uses Equation 8.

When $p = 0$, we use $s_0^*(0,k) = (\overline{a} + \underline{a})/2$ to obtain

$$\mathcal{U}_p(0,k) = \frac{\partial s_0^*(0,k)}{\partial p} 2(1-p)(\overline{a} - \underline{a})b + U\left(\frac{\overline{a}+\underline{a}}{2}\right) - U\left(\frac{\overline{a}+\underline{a}}{2} - b\right) > 0,$$

where the inequality is because $\frac{\partial s_0^*(0,k)}{\partial p} \geq 0$ (as a consequence of part 1 of Proposition 1) and $U(\cdot)$ is uniquely maximized at $(\overline{a} + \underline{a})/2$.

That $\mathcal{U}(\cdot, k)$ is uniquely maximized at $p = 1$ follows from Proposition 1 establishing that $s_0^*(1,k) = (\overline{a} + \underline{a})/2 = s_{FB}$, while for any $p < 1$ either $s_0^*(p,k) \neq s_{FB}$ or $s_b^*(p,k) \neq s_{FB}$. In words, only when $p = 1$ does the voter put probability one on the PM using the first-best threshold.

Part 2: Fix any $p \in (0,1)$. Since $s_0^*(p,0) = (\overline{a} + \underline{a})/2$,

$$\mathcal{U}(p,0) = pU\left(\frac{\overline{a}+\underline{a}}{2}\right) + (1-p)U\left(\frac{\overline{a}+\underline{a}}{2} - b\right) > U\left(\frac{\overline{a}+\underline{a}}{2} - b\right) = \mathcal{U}(0,0).$$

Since $\lim_{k\to\infty} s_0^*(p,k) = \infty$ (Proposition 1),

$$\lim_{k\to\infty} \mathcal{U}(p,k) = p \lim_{k\to\infty} U(s_0^*(p,k)) + (1-p) \lim_{k\to\infty} U\left(s_0^*(p,k) - b\right) = -\int_{\underline{s}}^{\infty} (\underline{a} - s)^2 f(s) ds.$$

Thus, $\lim_{k\to\infty} \mathcal{U}(p,k) < \mathcal{U}(0,0)$ if and only if

$$\int_{\underline{s}}^{\infty} (\underline{a} - s)^2 f(s) ds > \int_{\underline{s}}^{\frac{\overline{a}+\underline{a}}{2} - b} (\underline{a} - s)^2 f(s) ds + \int_{\frac{\overline{a}+\underline{a}}{2} - b}^{\infty} (\overline{a} - s)^2 f(s) ds,$$

or, equivalently, if and only if

$$\int_{\frac{\overline{a}+\underline{a}}{2} - b}^{\infty} (\underline{a} - s)^2 f(s) ds > \int_{\frac{\overline{a}+\underline{a}}{2} - b}^{\infty} (\overline{a} - s)^2 f(s) ds.$$

Expanding the quadratic term, dividing both sides by $2(\overline{a} - \underline{a})\left(1 - F\left(\frac{\overline{a}+\underline{a}}{2} - b\right)\right)$, and simplifying, the preceding inequality is equivalent to

$$\mathbb{E}\left[s \mid s \geq \frac{\overline{a}+\underline{a}}{2} - b\right] > \frac{\overline{a}+\underline{a}}{2},$$

which is precisely what was assumed in part 3 of Assumption 1 .

Therefore, $\mathcal{U}(p,0) > \mathcal{U}(0,0) > \lim_{k\to\infty} \mathcal{U}(p,k)$, and so the intermediate value theorem implies that there exists a $\hat{k}(p) > 0$ such that $\mathcal{U}(p, \hat{k}(p)) = \mathcal{U}(0,0)$. Since Lemma 1 established that $\mathcal{U}(p,k)$ is strictly quasi-concave in k, it follows that $\hat{k}(p)$ is unique, and that $\mathcal{U}(p,k) < \mathcal{U}(0,0)$ if and only if $k > \hat{k}(p)$. To see that $\hat{k}(p) \to \infty$ as $p \to 0$ or as $p \to 1$, suppose to the contrary that $\hat{k}(p)$ stays bounded. Then, using

the facts that (i) $U(\cdot)$ is strictly quasi-concave with a maximum at $(\overline{a}+\underline{a})/2$, (ii) for any k, $s_0^*(p,k) > (\overline{a}+\underline{a})/2$ for any $p \in (0,1)$ but $s_0^*(p,k) \to (\overline{a}+\underline{a})/2$ as $p \to 0$ or as $p \to 1$, and (iii) $\mathcal{U}(p,k)$ is given by expression Equation 9 whereas $\mathcal{U}(0,0) = U((\overline{a}+\underline{a})/2 - b)$, it follows that $\mathcal{U}(p, \hat{k}(p)) > \mathcal{U}(0,0)$ for all small or large enough $p \in (0,1)$, a contradiction.

Part 3: Follows immediately from the first two parts of the proposition. □

Proof of Lemma 2. In this proof, it will be convenient to denote the expected policy utility for a PM of type θ who uses a threshold τ as

$$\tilde{U}(\tau, \theta) := -\int_{\underline{s}}^{\tau} (\underline{a} - s - \theta)^2 f(s) ds - \int_{\tau}^{\infty} (\overline{a} - s - \theta)^2 f(s) ds.$$

Note that by Equation 2, $v_\theta = -\tilde{U}(s_\theta, \theta)$, where $s_\theta = (\overline{a}+\underline{a})/2 - \theta$ is the threshold type θ would use in the absence of reputation concern.

Fix any $p \in (0,1)$. We first show that for either type θ,

$$0 = W(\theta, 0, k) < W(\theta, p, k) < W(\theta, 1, k) = k. \tag{14}$$

The two equalities in (14) follow from the definition of $W(\cdot)$ in Equation 10, the fact that $v_\theta = \tilde{U}(s_\theta, \theta)$, and that $s_\theta^*(0,k) = s_\theta^*(1,k) = s_\theta$ (Proposition 1). The last inequality in (14) holds because

$$\begin{aligned}
W(\theta, p, k) &= v_\theta + \tilde{U}(s_\theta^*(p,k), \theta) + k[F(s_\theta^*(p,k))V(\overline{p}(p,k)) + (1 - F(s_\theta^*(p,k)))V(\underline{p}(p,k))] \\
&< v_\theta + \tilde{U}(s_\theta, \theta) + k[F(s_\theta^*(p,k))V(\overline{p}(p,k)) + (1 - F(s_\theta^*(p,k)))V(\underline{p}(p,k))] \\
&= k[F(s_\theta^*(p,k))V(\overline{p}(p,k)) + (1 - F(s_\theta^*(p,k)))V(\underline{p}(p,k))] \\
&< k,
\end{aligned}$$

where the first equality uses the definition of $W(\cdot)$ and $\tilde{U}(\cdot)$, the first inequality uses $s_\theta^*(\cdot) > s_\theta$, the second equality uses $v_\theta = -\tilde{U}(s_\theta, \theta)$, and the final inequality uses $V(\cdot) < 1$ for any interior belief.

To show the first inequality in (14), we observe that

$$\begin{aligned}
W(\theta, p, k) &\geq v_\theta + \tilde{U}(s_\theta, \theta) + k[F(s_\theta)V(\overline{p}(p,k)) + (1 - F(s_\theta))V(\underline{p}(p,k))] \\
&= k[F(s_\theta)V(\overline{p}(p,k)) + (1 - F(s_\theta))V(\underline{p}(p,k))] \\
&> 0,
\end{aligned}$$

where the first inequality is because type 0 uses threshold $s_\theta^*(\cdot)$ rather than deviating to threshold s_θ, and the last inequality is because $V(\cdot) > 0$ for any non-degenerate belief.

We now prove the second part of the lemma, which in light of (14) is equivalent to showing $W(0, p, k) > W(b, p, k)$. There are two exhaustive possibilities to cover:

Case 1: $s_b^*(p,k) \leq (\overline{a}+\underline{a})/2 = s_0$. Then we observe that

$$\begin{aligned}
W(0,p,k) &\geq v_0 + \tilde{U}(s_0,0) + k[F(s_0)V(\overline{p}(p,k)) + (1-F(s_0))V(\underline{p}(p,k))] \\
&= k[F(s_0)V(\overline{p}(p,k)) + (1-F(s_0))V(\underline{p}(p,k))] \\
&\geq k[F(s_b^*(p,k))V(\overline{p}(p,k)) + (1-F(s_b^*(p,k)))V(\underline{p}(p,k))] \\
&> k[F(s_b^*(p,k))V(\overline{p}(p,k)) + (1-F(s_b^*(p,k)))V(\underline{p}(p,k))] + v_b + \tilde{U}(s_b^*(p,k),b) \\
&= W(b,p,k),
\end{aligned}$$

where the first inequality is because type 0 uses threshold $s_0^*(\cdot)$ rather than deviating to threshold s_0, the first equality is because $v_0 = -\tilde{U}(s_0,0)$, the second inequality is because $s_b^*(\cdot) \leq s_0$ and $\overline{p}(p,k) > \underline{p}(p,k)$, and the final inequality is because $s_b^*(\cdot) > s_b$ implies $v_b = -\tilde{U}(s_b,b) < -\tilde{U}(s_b^*(\cdot),b)$.

Case 2: $s_b^*(p,k) > (\overline{a}+\underline{a})/2 = s_0$. Now we consider a deviation by type 0 to threshold $s_b^*(p,k)$. Notice that under the deviation, the expected reputational payoff for type 0 is the same as the equilibrium expected reputational payoff for type b. Consequently,

$$\begin{aligned}
W(0,p,k) - W(b,p,k) &\geq v_0 + \tilde{U}(s_b^*(p,k),0) - \left[v_b + \tilde{U}(s_b^*(p,k),b)\right] \\
&= \int_{s_0}^{s_b^*(p,k)} \left[(\overline{a}-s)^2 - (\underline{a}-s)^2\right] f(s)ds - \int_{s_0-b}^{s_b^*(p,k)} [(\overline{a}-s-b)^2 - (\underline{a}-s-b)^2] f(s)ds \\
&> 0,
\end{aligned}$$

where the first inequality is because type 0 uses threshold $s_0^*(\cdot)$ rather than deviating to threshold $s_b^*(\cdot)$ (and the identical expected reputational payoff for the two types under type 0's deviation); the equality follows from $v_\theta = -\tilde{U}(s_\theta,\theta)$, expanding $\tilde{U}(\cdot)$, and some algebraic manipulation; and the final inequality is because (i) $(\overline{a}-s-b)^2 < (\underline{a}-s-b)^2$ if $s > s_0 - b$ and (ii) $(\overline{a}-s-b)^2 - (\underline{a}-s-b)^2 < (\overline{a}-s)^2 - (\underline{a}-s)^2$ for any s. □

Proof of Lemma 3. Suppose, per contra, that there exists an informative (symmetric) equilibrium in which $\mathcal{U}(p^0,k) \neq \mathcal{U}(p^b,k)$. Let $j \in \{0,b\}$ be the message such that $\mathcal{U}(p^j,k) > \mathcal{U}(p^{b-j},k)$. Then, if $m_1 \neq m_2$ the voter must elect the candidate who announced j, and if $m_1 = m_2$ the voter randomizes with equal probability. Hence, no matter the opponent's announcement, a candidate at least doubles his probability of winning by announcing j rather than $b-j$.

Now consider a candidate i with type θ_i. Since a candidate's payoff is 0 if not elected, the expected utility from announcing message m is $\Pr(i \text{ being elected}|m_i = m)(c + W(\theta_i, p^m, k))$. Observe that

$$\begin{aligned}
&\Pr(i \text{ being elected}|m_i = j)(c + W(\theta_i,p^j,k)) - \Pr(i \text{ being elected}|m_i = b-j)(c + W(\theta_i,p^{b-j},k)) \\
\geq &\Pr(i \text{ being elected}|m_i = b-j)\left[2c + 2W(\theta_i,p^j,k) - c - W(\theta_i,p^{b-j},k))\right] \\
> &\Pr(i \text{ being elected}|m_i = b-j)\,[c-k] \\
\geq &0,
\end{aligned}$$

where the first inequality is because $m_i = j$ at least doubles the winning probability over $m_i = b - j$; the second inequality is due to Lemma 2 implying $0 \leq W(\theta_i, p^j, k)$ and $W(\theta_i, p^{b-j}, k) \leq k$ with one of these inequalities holding strictly because $p^{b-j} = 1$ and $p^j = 0$ is ruled out by $\mathcal{U}(p^j, k) > \mathcal{U}(p^{b-j}, k)$; and the final inequality follows from Assumption 2.

Hence, any candidate strictly prefers to send message j over message $b - j$, a contradiction with the equilibrium being informative. □

Proof of Lemma 4. First note that $k^* = \inf\{\hat{k}(p) : p \in (0,1)\}$, where $\hat{k}(\cdot)$ was defined in part 2 of Proposition 2 as the unique positive solution to $\mathcal{U}(p, \hat{k}(p)) = \mathcal{U}(0,0)$. Since $\hat{k}(p)$ is finite for all p, it follows that $k^* < \infty$. That $k^* > 0$ follows from the observations that $\hat{k}(p)$ is continuous and strictly positive for any $p \in (0,1)$, and does not tend to 0 as $p \to 0$ or as $p \to 1$ (Proposition 2). This establishes part (i) of the lemma. Parts (ii) and (iii) follow from the definition of $p^*(\cdot)$ and that $k^* = \inf\{\hat{k}(p) : p \in (0,1)\}$ and $\mathcal{U}_p(0,k) > 0$ for all k (part 1 of Proposition 2).

For part (iv): note that for any $k \geq k^*$, $\mathcal{U}(p^*(k), k) = \mathcal{U}(0,0)$ and $k = \hat{k}(p^*(k))$. Therefore, Proposition 2 implies that for all $k' > k$, $\mathcal{U}(p^*(k), k') < \mathcal{U}(0,0)$. By continuity, there exists $p' > p^*(k)$ such that $\mathcal{U}(p', k') < \mathcal{U}(0,0)$, and so $p^*(k') > p^*(k)$.

Finally, for part (v): since $k = \hat{k}(p^*(k))$ for $k \geq k^*$ and $\hat{k}(\cdot)$ is continuous and unbounded (Proposition 2), it follows that $p^*(k) \to 1$ as $k \to \infty$. □

Proof of Proposition 3. We show that a semi-separating equilibrium exists if and only if $p < p^*(k)$; note that this condition implies $k \geq k^*$. By Lemma 3, any semi-separating equilibrium has $1 = \mu^0 > \mu^b > 0$ and voter beliefs $p^0 > p > p^b = 0$ such that $\mathcal{U}(p^0, k) = \mathcal{U}(0, k)$. The "only if" direction of the result now follows from the fact that, by the definition of $p^*(\cdot)$, $\mathcal{U}(p^0, k) > \mathcal{U}(0,0)$ when $p^0 > p^*(k)$.

For the "if" direction, assume $p < p^*(k)$, and hence also $k \geq k^*$. We construct a semi-separating equilibrium where $p^0 = p^*(k)$ and $p^b = 0$. Let $\mu^0 = 1$ and $\mu^b \in (0,1)$ be the unique solution to

$$\frac{p}{p + (1-p)\mu^b} = p^*(k),$$

and let $w^0 := p + (1-p)\mu^b$ be the probability that a candidate announces message 0.

Plainly, given the candidates' strategies, any behavior is optimal for the voter (when the candidates send distinct messages), because $\mathcal{U}(p^0, k) = \mathcal{U}(p^*(k), k) = \mathcal{U}(0,0) = \mathcal{U}(p^b, k)$. For the candidates, it suffices to check that the non-congruent type is playing optimally by mixing, because the second part of Lemma 2 then ensures that it is (strictly) optimal for the congruent type to play $\mu^0 = 1$. Thus, we are left to construct the voter's strategy to generate indifference of the non-congruent type. The indifference condition for a non-congruent candidate i is

$$\Pr(i \text{ being elected} | m_i = 0)(c + W(b, p^0, k)) = \Pr(i \text{ being elected} | m_i = b)(c + W(b, 0, k)),$$

or, since the voter elects the candidate announcing message 0 with probability σ upon observing distinct messages and randomizes uniformly across candidates when they send the same message,

$$\left(\frac{1}{2}w^0 + (1-w^0)\sigma\right)(c + W(b, p^0, k)) = \left(\frac{1}{2}(1-w^0) + w^0(1-\sigma)\right)(c + W(b, 0, k)). \tag{15}$$

As the LHS of Equation 15 is increasing in σ while the RHS is decreasing in it, there is at most value of σ that solves Equation 15 . The argument given in the proof of Lemma 3 shows that the RHS of Equation 15 is strictly larger than the LHS when $\sigma = 0$; on the other hand, when $\sigma = 1/2$, the LHS is strictly larger than the RHS because $W(b, p^0, k) > W(b, 0, k)$ by Lemma 2. Continuity implies there is exactly one value of $\sigma \in (0, 1/2)$ that solves Equation 15 and hence constitutes an equilibrium. Note that this argument also implies that $\sigma \in (0, 1/2)$ in *any* semi-separating equilibrium, even if $p^0 \neq p^*(k)$.

The last two parts of the proposition follow immediately from the part we have just proved when combined with $p^*(\cdot)$ being strictly increasing on $[k^*, \infty)$ and $p^*(k) \to 1$ as $k \to \infty$ (Lemma 4). □

Proof of Corollary 1. As explained before the corollary, the result follows from Proposition 2. □

Proof of Proposition 4. First note using Proposition 2, which defined $\hat{k}(\cdot)$, that

$$P^k = \{p \in (0,1) : k > \hat{k}(p)\}. \tag{16}$$

Part 1: That $P^{k_1} \subseteq P^{k_2}$ for any $k_1 < k_2$ is immediate from Equation 16. When $k_2 > k^*$, the inclusion is strict because $\hat{k}(p) \to \infty$ as $p \to 1$ (Proposition 2) and the continuity of $\hat{k}(\cdot)$ together imply $P^{k_2} \setminus P^{k_1} \neq \emptyset$.

Part 2: Follows immediately from Equation 16.

Part 3: Since $\mathcal{U}(p, \hat{k}(p)) = \mathcal{U}(0,0)$, the strict quasi-concavity of $\mathcal{U}(p, \cdot)$ established in Lemma 1 implies that $\mathcal{U}(p, \cdot)$ is strictly decreasing on $[\hat{k}(p), \infty)$. Since $p \in P^{k_1}$ implies $k_1 > \hat{k}(p)$, it follows that for all $k > k_1$, $\frac{\partial \mathcal{U}(p,k)}{\partial k} < 0$. □

Proof of Proposition 5. We prove each part of the result in sequence.

Part 1: Since the PM's incentives in office are the same as in the baseline model, Lemma 3 applies: $\mathcal{U}(p^b, k) = \mathcal{U}(p^0, k)$ and $p^b < p < p^0$ in any informative equilibrium. This implies the "only if" portion of the result. For the "if" portion, note that if the voter always randomizes between both candidates with equal probability, candidates are indifferent over messages. A standard result concerning Bayesian updating implies that candidates' randomization can be chosen in a way to induce the voter's belief after observing messages b and 0 to respectively be any p' and p'' satisfying $p' < p < p''$.

Part 2: Fix any $\varepsilon > 0$ and $p \in (0,1)$, and recall that $\mathcal{U}(p,k) = pU(s_0^*(p,k)) + (1-p)U(s_b^*(p,k))$.

Assume k is large enough that $s_b^*(p,k) > (\overline{a}+\underline{a})/2$ and define

$$p^b(k) = \min\left\{p' \in (0,p) : s_b^*(p,k) = \frac{\overline{a}+\underline{a}}{2}\right\}.$$

(This is well-defined by Proposition 1.) Since $U(\cdot)$ is strictly decreasing above $(\overline{a}+\underline{a})/2$, it follows that $\mathcal{U}(p^b(k),k) > \mathcal{U}(p,k)$. Since $\mathcal{U}(\cdot,k)$ is continuous and uniquely maximized at 1, there exists $p^0(k) \in (p,1)$ such that $\mathcal{U}(p^b(k),k) = \mathcal{U}(p^0(k),k)$. By the first part of the proposition, there is an informative equilibrium in which the voter's expected utility is

$$\mathcal{U}(p^b(k),k) = p^b(k)U\left(\frac{\overline{a}+\underline{a}}{2}+b\right) + (1-p^b(k))U\left(\frac{\overline{a}+\underline{a}}{2}\right).$$

Since for all p', $\lim_{k\to\infty} s_b^*(p',k) = \infty$, it follows that $\lim_{k\to\infty} p^b(k) = 0$. Consequently,

$$\lim_{k\to\infty} \mathcal{U}(p^b(k),k) = U\left(\frac{\overline{a}+\underline{a}}{2}\right) = \mathcal{U}(1,0),$$

which implies that there is some $\overline{k}$ such that $\mathcal{U}(p^b(k),k) > \mathcal{U}(1,0) - \varepsilon$ for all $k > \overline{k}$. □

References

ACEMOGLU, D., G. EGOROV, AND K. SONIN (2013): "A Political Theory of Populism," *Quarterly Journal of Economics*, 128, 771–805.

ALESINA, A. (1988): "Credibility and Policy Convergence in a Two-Party System with Rational Voters," *American Economic Review*, 78, 796–805.

AN, M. Y. (1998): "Logconcavity versus Logconvexity: A Complete Characterization," *Journal of Economic Theory*, 80, 350–369.

ARAGONES, E., T. PALFREY, AND A. POSTLEWAITE (2007): "Reputatation and Rhetoric in Elections," *Journal of the European Economic Association*, 5, 846–884.

BAGNOLI, M. AND T. BERGSTROM (2005): "Log-concave Probability and its Applications," *Economic Theory*, 26, 445–469.

BANKS, J. S. (1990): "A Model of Electoral Competition with Incomplete Information," *Journal of Economic Theory*, 50, 309–325.

BANKS, J. S. AND J. DUGGAN (2008): "A Dynamic Model of Democratic Elections in Multidimensional Policy Spaces," *Quarterly Journal of Political Science*, 3, 269–299.

BANKS, J. S. AND J. SOBEL (1987): "Equilibrium Selection in Signaling Games," *Econometrica*, 55, 647–661.

BAR-ISAAC, H. AND J. DEB (2014): "What is a Good Reputation? Career Concerns with Heterogeneous Audiences," *International Journal of Industrial Organization*, 34, 44–50.

BESLEY, T. AND S. COATE (1997): "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 112, 85–114.

CALLANDER, S. AND S. WILKIE (2007): "Lies, Damned Lies, and Political Campaigns," *Games and Economic Behavior*, 60, 262–286.

CALVERT, R. L. (1985): "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence," *American Journal of Political Science*, 29, 69–95.

CANES-WRONE, B., M. HERRON, AND K. W. SHOTTS (2001): "Leadership and Pandering: A Theory of Executive Policymaking," *American Journal of Political Science*, 45, 532–550.

CHAKRABORTY, A. AND R. HARBAUGH (2010): "Persuasion by Cheap Talk," *American Economic Review*, 100, 2361–82.

CHO, I.-K. AND D. KREPS (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179–221.

CLAIBOURN, M. P. (2011): *Presidential Campaigns and Presidential Accountability*, University of Illinois Press.

COWEN, T. AND D. SUTTER (1998): "Why Only Nixon Could Go to China," *Public Choice*, 97, 605–15.

CUKIERMAN, A. AND M. TOMMASI (1998): "When Does it Take a Nixon to Go to China," *American Economic Review*, 88, 180–197.

DOWNS, A. (1957): *An Economic Theory of Democracy*, New York: Harper and Row.

ELY, J. C. AND J. VÄLIMÄKI (2003): "Bad Reputation," *Quarterly Journal of Economics*, 118, 785–814.

FOX, J. AND M. STEPHENSON (2013): "The Distributional Effects of Minority-Protective Judicial Review," Forthcoming in the *Journal of Theoretical Politics*.

FUDENBERG, D. AND J. TIROLE (1991): *Game Theory*, Cambridge, MA: MIT Press.

GROSSER, J. AND T. R. PALFREY (2014): "Candidate Entry and Political Polarization: An Antimedian Voter Theorem," *American Journal of Political Science*, 58, 127–143.

HARRINGTON, J. E. (1992): "The Revelation of Information through the Electoral Process: An Exploratory Analysis," *Economics & Politics*, 4, 255–276.

——— (1993): "The Impact of Reelection Pressures on the Fulfillment of Campaign Promises," *Games and Economic Behavior*, 5, 71–97.

HOTELLING, H. (1929): "Stability in Competition," *Economic Journal*, XXXIX, 41–57.

HUANG, H. (2010): "Electoral Competition When Some Candidates Lie and Others Pander," *Journal of Theoretical Politics*, 22, 333–358.

KARTIK, N. AND R. P. MCAFEE (2007): "Signaling Character in Electoral Competition," *American Economic Review*, 97, 852–870.

KARTIK, N., F. SQUINTANI, AND K. TINN (2014): "Information Revelation and Pandering in Elections," Unpublished.

MAILATH, G. J. AND L. SAMUELSON (2001): "Who Wants a Good Reputation?" *The Review of Economic Studies*, 68, 415–441.

MASKIN, E. AND J. TIROLE (2004): "The Politician and the Judge: Accountability in Government," *American Economic Review*, 94, 1034–1054.

MOEN, E. R. AND C. RIIS (2010): "Policy Reversal," *American Economic Review*, 100, 1261–1268.

MORELLI, M. AND R. VAN WEELDEN (2013a): "Ideology and Information in Policy Making," *Journal of Theoretical Politics*, 25, 412–439.

——— (2013b): "Re-election Through Division," Unpublished.

MORRIS, S. (2001): "Political Correctness," *Journal of Political Economy*, 109, 231–265.

OSBORNE, M. J. AND A. SLIVINSKI (1996): "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics*, 111, 65–96.

PANOVA, E. (2014): "Informative Campaign Promises," Unpublished.

PRENDERGAST, C. (1993): "A Theory of "Yes Men"," *American Economic Review*, 83, 757–70.

PRENDERGAST, C. AND L. STOLE (1996): "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy*, 104, 1105–34.

SCHARFSTEIN, D. S. AND J. C. STEIN (1990): "Herd Behavior and Investment," *American Economic Review*, 80, 465–79.

SCHNAKENBERG, K. (2014): "Directional Cheap Talk in Electoral Campaigns," Unpublished.

STONE, W. J. AND E. N. SIMAS (2010): "Candidate Valence and Ideological Positions in U.S. House Elections," *American Journal of Political Science*, 54, 371–388.

WITTMAN, D. (1977): "Candidates with Policy Preferences," *Journal of Economic Theory*, 14, 180–189.