

Informative Cheap Talk in Elections*

Navin Kartik[†]

Richard Van Weelden[‡]

October 19, 2017

Abstract

Why do office-motivated politicians sometimes espouse views that are non-congruent with their electorate's? Can non-congruent statements convey any information about what a politician will do if elected, and if so, why would voters elect a politician who makes such statements? Furthermore, can electoral campaigns also directly affect an elected official's behavior? We develop a model of credible "cheap talk"—costless and non-binding communication—in elections. 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 preferences generates policymaking distortions due to reputation/career concerns. We show that cheap talk can alter the electorate's beliefs about a politician's policy preferences and thereby affect the elected official's behavior. Informative cheap talk can increase or decrease voter welfare, with a greater scope for welfare benefits when reputation concerns are more important.

Keywords: Pandering, campaigns, reputational distortions, career concerns, voter learning.

JEL: D72, D83

*We are grateful to Sandeep Baliga, Odilon Câmara, Chris Cotton, Ernesto Dal Bó, Allan Drazen, Wiola Dziuda, Alex Frankel, Emir Kamenica, Massimo Morelli, Salvatore Nunnari, Ken Shotts, Stephane Wolton, the Editor (Botond Kőszegi), anonymous referees, and various conference and seminar audiences for helpful comments. Vinayak Iyer, Teck Yong Tan, Enrico Zanardo, and Weijie Zhong 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 Pittsburgh. Email: rmv22@pitt.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, 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 evidence that campaigns provide useful information about what candidates will do in office ([Sulkin, 2009](#); [Claibourn, 2011](#); [Bidwell, Casey, and Glennerster, 2016](#)), and furthermore, with the notion that a candidate’s post-election behavior may be affected by his campaign statements?

This paper develops a novel rationale for informative cheap talk in elections. We show how cheap-talk campaign statements can not only reveal information about candidates’ policy preferences, but also alter a candidate’s behavior if he is elected.

[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

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

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 about post-political life—the elected politician also benefits from establishing a reputation for congruence through his actions in office.²

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. Unguarded intuition would suggest that since the voter always prefers 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 may or may not be congruent. 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

² It is well recognized that reputational concerns affect policymaking. For example, many perceive President Obama's policy choices in his second term (but not in his first term) as "freed from the political constraints of an impending election" ([Davis, 2015](#)). Obama himself has said about his second term, "I'm just telling the truth now. I don't have to run for office again, so I can just, you know, let her rip" ([Obama, 2014](#)).

who is concerned about being perceived as soft on communism.³ Reputational pandering thus endogenously generates the phenomenon of “a known devil is better than an unknown angel.” But 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.

Accordingly, our analysis illuminates *why* voters benefit from knowing a policymaker’s preferences/values, and our framework can micro-found a dislike for “flip-floppers” even when voters care only that appropriate policies be chosen.⁴ Notably, voters’ aversion to politicians whose ideology is uncertain is not because of uncertainty regarding what such politicians would *do*—to the contrary, in our model there is greater uncertainty about the action taken when there is less uncertainty about a policymaker’s type because a policymaker will adjust policy to the state more when his type is known—but rather because of the policy distortions caused by subsequent pandering. This distinction may help rationalize recent empirical work. Rogowski and Tucker (2016) argue that, all else equal, support for a candidate decreases in the variance of their perceived ideology; however, there does not appear to be a similar effect when the uncertainty concerns what policies will be enacted (e.g., Tomz and Van Houweling, 2009).

The aforementioned welfare 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. We confirm a limited single-crossing property that sustains this structure; in equilibrium, candidates’ behavior is such that the voter is indifferent between electing a candidate who reveals himself to be non-congruent and electing a candidate whose type she is unsure about.

Informative communication in our model endogenously ties candidates’ post-election behavior to their electoral campaign, despite communication being non-binding and costless.

³ 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 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.

⁴ As a corollary, our analysis also explains why voters may value traits like “honesty” or “character” in politicians—a characteristic of voter preferences that is sometimes assumed in reduced form (e.g., Kartik and McAfee, 2007; Fernandez-Vasquez, 2014).

Put differently, our analysis explains how campaign pronouncements can influence post-election policymaking—controlling for a policymaker’s policy preference and the realized state of the world—even when such pronouncements are cheap talk. 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.⁵ 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 2000 presidential campaign bus the “Straight Talk Express”. Voters’ reluctance to support candidates whose policy preferences they are uncertain about is also illustrated in recent U.S. presidential elections. Al Gore in 2000 was described as “willing to say anything”, John Kerry in 2004 as a “flip-flopper”—perceptions which, as suggested by our epigraph, were exploited by George W. Bush’s campaigns—and Mitt Romney faced similar travails in 2012. Our theory attributes voters’ concerns with these candidates as (at least partly) stemming from apprehension about their post-electoral policy pandering. It is particularly interesting to contrast the Romney campaign with that of Michael Bloomberg, another businessman turned politician, who was elected mayor of New York city three times and praised for demonstrating “real leadership” by taking positions at odds with the majority of his electorate (e.g., [McGregor, 2010](#)).

An important question is whether equilibria with informative cheap talk generate higher voter welfare than uninformative equilibria (which always exist in virtually any cheap-talk game). As informative campaigns provide information about candidates’ preferences but also change the elected candidate’s behavior, their welfare effects turn out to depend on the prior about candidates’ congruence. For low priors, voter welfare is higher in uninformative equilibria than in the aforementioned semi-separating equilibria. The comparison is reversed for a range of higher priors. 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., moderate) 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. Our analysis thus yields the novel insights that informative electoral campaigns (or, indeed, any information about candidates’ preferences, even if from a third party like the media) can either mitigate or exacerbate policymaking distortions

⁵ In [Carrillo and Castanheira \(2008\)](#), candidates face moral hazard in investment on a vertical quality dimension, whose outcome is observed with some probability prior to the election. They discuss how committing to a non-centrist ideology can act as a credible commitment to invest in quality.

induced by reputation concerns and, consequently, improve or reduce voter welfare.⁶

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. 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 \(1983\)](#) and [Calvert \(1985\)](#) that predict less scope for policy divergence when office motivation is larger. While any empirical test of our theory would have to be carefully designed, our comparative-static prediction could be checked. For example, one might use political salary to proxy for office-holding benefits (e.g., [Hoffman and Lyons, 2017](#)) and the change in voters’ beliefs (with suitable controls) between the beginning and end of the campaign to proxy for informativeness.

[Section 5](#) contains some extensions of our main results, and [Section 6](#) 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.⁷ However, most existing electoral models with asymmetric information either preclude cheap-talk announcements on the basis that they would be uninformative (e.g., [Banks and Duggan,](#)

⁶ Cheap-talk campaigns cannot reduce voter welfare when one focuses on welfare-maximizing equilibria, but this may entail uninformative communication. Focussing on welfare-maximizing equilibria, our results have the interesting implication that cheap-talk campaigns provide a lower bound on voter welfare even as reputational concerns get arbitrary large.

⁷ 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-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 \(2007\)](#).

2008; [Großer and Palfrey, 2014](#)) or allow for it and argue that they should not be informative in equilibrium (e.g., [Kartik, Squintani, and Tinn, 2015](#)).

[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 monotonic 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 \(2017\)](#) 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 the final policy. They derive informative equilibria under some conditions. [Schnakenberg \(2016\)](#) 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-election behavior is independent of the electoral campaign in both [Kartik and McAfee \(2007\)](#) and [Schnakenberg \(2016\)](#); 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 \(2007\)](#), [Huang \(2010\)](#), and [Agranov \(2016\)](#) study such models. One can also appeal to “behavioral preferences” on the voter side ([Grillo, 2016](#)).

To our knowledge, this paper is the first to study the implications of reputational distortions in policymaking on electoral campaigns and the initial selection of policymakers. We build on 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, viz. that 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

The property 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,⁸ it is quite natural in many contexts, particularly in delegated decision-making when there is some degree of common interest. [Acemoglu, Egorov, and Sonin \(2013\)](#) have previously demonstrated that reputation concerns 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 \(2015\)](#), [Morelli and Van Weelden \(2013\)](#) and [Ash, Morelli, and Van Weelden \(2017\)](#). Unlike us, these authors do not focus on the voter’s welfare as a function of her belief nor do they consider how electoral campaigns interact with pandering in policymaking. Studying these issues are our central contributions.

2. The Model

We model a representative (or median) voter electing a politician to take a policy action on her behalf. Our model makes a distinction between three kinds of political motivations: office motivation (direct benefits of holding office, including salary and “ego rents”), policy motivation (preferences about which policy is chosen), 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

⁸For 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 known non-congruent policymaker always takes the worst possible action for the voter.

policy motivation and office motivation is high relative to reputation motivation. The former guarantees that politicians whose preferences are uncertain when elected will engage in sufficiently detrimental pandering; the latter ensures that politicians are willing to reveal their preference type if doing so sufficiently increases their probability of being elected.

In more detail: the voter’s utility depends on a state of the world, $s \in \mathbb{R}$, and a policy action, $a \in \{\underline{a}, \bar{a}\} \subset \mathbb{R}$, with $\bar{a} > \underline{a}$. 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 . The state s is drawn from a cumulative distribution F with support $[\underline{s}, \infty)$, where \underline{s} can either be finite or $-\infty$; the distribution F admits a differentiable and bounded density f with $f(s) > 0$ on (\underline{s}, ∞) . The voter’s utility is maximized when the action matches the state of the world. For simplicity, we assume the voter’s von-Neumann Morgenstern utility is given by a 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 $\theta \in \{0, b\}$, with $b > 0$. We call $\theta = 0$ the *congruent type* and $\theta = b$ the *non-congruent* or *biased type*. Each candidate’s type is his private information, and each candidate is independently drawn as congruent with ex-ante probability $p \in (0, 1)$.⁹ 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. (Subsection 5.5 considers an extension in which the candidates receive a noisy signal of the state prior to the election.) 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.

The elected politician learns the state s and chooses the policy action a . After observing the action taken—but before she learns her utility or anything else directly about the state—the voter updates her belief about the PM’s congruence. (Subsection 5.4 elaborates on how our results are qualitatively unchanged even if the voter’s posterior can depend on some direct information about the state.) 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

⁹ A number of modeling choices here are for simplicity only: (i) it is not important that the ex-ante probability of each candidate being congruent is the same; (ii) we could allow for the two candidates’ biases to be in opposite directions (to reflect party affiliation) subject to appropriate assumptions; and (iii) our main themes would be fundamentally unchanged if there were more than two candidates. Also, see the [Supplementary Appendix](#) for a more general setting that allows for an arbitrary (finite) number of types and policy actions.

payoff normalized to 0.¹⁰ 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

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

where $k \geq 0$, $c > 0$, and $v_\theta > 0$ are scalars, and $V : [0, 1] \rightarrow \mathbb{R}_+$ 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 \bar{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 $k \geq 0$. The higher k is, 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. See [Subsection 5.2](#) for an alternative micro-foundation using a richer dynamic model.

[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 (weak) 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 equilibrium path. As explained in more detail in [Section 4](#), 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*

¹⁰ Analogous results to ours can 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.

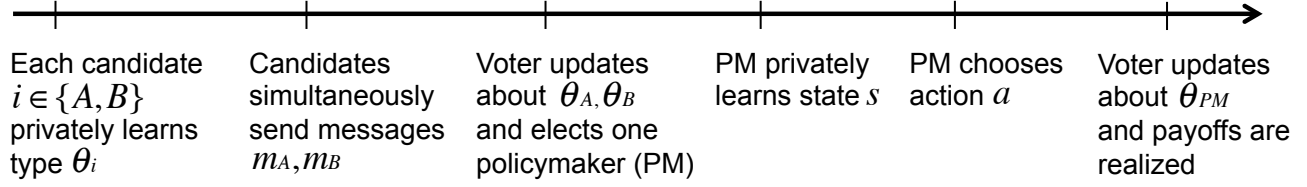


Figure 1 – Summary of the game form.

if there is some on-path message m_i such that $p_i(m_i)$, 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 \bar{a} if and only if (modulo indifference) $s > s_{FB} := (\bar{a} + \underline{a})/2$. In the absence of reputation concerns ($k = 0$), a PM of type $\theta \in \{0, b\}$ would take action \bar{a} if and only if $s > s_\theta := (\bar{a} + \underline{a})/2 - \theta$. 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 \bar{a} in 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. $\underline{s} < \frac{\bar{a} + \underline{a}}{2} - b$;
2. On the domain $[\frac{\bar{a} + \underline{a}}{2} - b, \infty)$, $f(\cdot)$ is log-convex, i.e., $\frac{f'(s)}{f(s)} \geq \frac{f'(t)}{f(t)}$ if $s > t \geq \frac{\bar{a} + \underline{a}}{2} - b$;
3. $\mathbb{E}[s | s \geq \frac{\bar{a} + \underline{a}}{2} - b] > \frac{\bar{a} + \underline{a}}{2}$.

Assumption 2. $c \geq k$.

Part 1 of [Assumption 1](#) is mild: it requires that in the absence of reputation concerns, each action would be taken by both types of the PM. Part 2 is not essential for our main points, but it will prove to be technically convenient by facilitating certain uniqueness results and comparative statics.¹¹ The [Supplementary Appendix](#) shows that our main results hold without

¹¹ A number of familiar distributions have log-convex densities on their entire domain; our leading example will be the exponential distribution. 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 \(2005\)](#).

part 2 of [Assumption 1](#). Part 3 of the assumption is substantive: it is equivalent to assuming that the voter is better off with a non-congruent PM who has no reputation concern than with a PM who always takes action \underline{a} . This equivalence is verified in the [proof of Proposition 2](#). Part 3 of [Assumption 1](#) holds if the distribution F has enough weight in the right-tail; in particular, no matter the bias b , it is sufficient that $\mathbb{E}[s] \geq \frac{\bar{a}+\underline{a}}{2}$. Alternatively, given any F (with support unbounded above), part 3 of [Assumption 1](#) holds if b is small enough. 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. Note that if k is interpreted as the value of re-election in the two period model described earlier, then [Assumption 2](#) is satisfied.

Due to their different policy preferences, the two types of a candidate will generally value holding office differently even in the absence of any reputation concerns. One may worry that this asymmetry by itself—as opposed to the effects of reputation concerns—creates an avenue for informative cheap talk in elections. Accordingly, we choose a value of v_θ in expression (1) to avoid this 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.¹² Since $c > 0$, $k \geq 0$, and $V(\cdot) \geq 0$, our choices of v_0 and v_b ensure that the expected payoff from holding office is strictly higher than from not holding office (which was normalized to zero) for both candidate types. Our choices of v_0 and v_b stack the deck against the possibility of informative cheap talk; our results are robust to other choices of v_0 and v_b , so long as the value of holding office is positive and not too asymmetric across types.

Remark 1. Consider $k = 0$. A policymaker with type θ uses threshold s_θ to determine his policy action. The voter thus prefers to elect a candidate who is more likely to be congruent. Since both types of a candidate prefer to be elected than not elected, independent of the voter's belief about the candidate's type, it follows that electoral campaigns are uninformative. \parallel

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

¹² Formally, the expected payoff for type θ from holding office given $k = c = 0$ is

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

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

3. Policymaking with Reputation Concerns

3.1. Equilibrium pandering

We begin by solving 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.¹³

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 \bar{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:¹⁴

$$\bar{p} := \frac{pF(s_0^*)}{pF(s_0^*) + (1-p)F(s_b^*)}, \quad (2)$$

$$\underline{p} := \frac{p(1 - F(s_0^*))}{p(1 - F(s_0^*)) + (1-p)(1 - F(s_b^*))}, \quad (3)$$

$$-(\underline{a} - s_0^*)^2 + kV(\underline{p}) = -(\bar{a} - s_0^*)^2 + kV(\bar{p}), \quad (4)$$

$$-(\underline{a} - s_b^* - b)^2 + kV(\underline{p}) = -(\bar{a} - s_b^* - b)^2 + kV(\bar{p}). \quad (5)$$

The first two equations above represent Bayesian updating: the voter’s posterior that the PM is congruent is \bar{p} following action \underline{a} and \underline{p} following \bar{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 equations are the indifference conditions at each type’s threshold.

Equation 4 and Equation 5 imply that $s_b^* = s_0^* - b$ in any equilibrium. In other words, the non-congruent type’s threshold is pinned down by the congruent type’s, and is simply a shift

¹³ For some parameters of our model, there can be an equilibrium in which both types take action \bar{a} regardless of the state; such equilibria are supported by assigning a sufficiently high probability to the PM being non-congruent if he takes the off-path action \underline{a} . But these 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.

¹⁴ Part 1 of Assumption 1 ensures that in any interior equilibrium, both types must use thresholds in (\underline{s}, ∞) .

down by the bias. Manipulating (2)–(5), an equilibrium can be succinctly characterized by a single equation of one variable, s_0^* :

$$s_0^* - \frac{\bar{a} + \underline{a}}{2} = \frac{k}{2(\bar{a} - \underline{a})} \left[V \left(\frac{p}{p + (1-p) \frac{F(s_0^* - b)}{F(s_0^*)}} \right) - V \left(\frac{p}{p + (1-p) \frac{1-F(s_0^* - b)}{1-F(s_0^*)}} \right) \right]. \quad (6)$$

When $p \in \{0, 1\}$ or $k = 0$, the right-hand side (RHS) above is zero and hence the unique solution to Equation 6 is $s_0^* = (\bar{a} + \underline{a})/2$. However, when $p \in (0, 1)$ and $k > 0$, 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.

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 6 and the non-congruent type uses a threshold $s_b^*(p, k) = s_0^*(p, k) - b$. Moreover, $s_0^*(p, k)$ is continuously differentiable in both arguments, and:*

1. If $p \in (0, 1)$ and $k > 0$, then

$$s_0^*(p, k) > \frac{\bar{a} + \underline{a}}{2} = s_0^*(0, k) = s_0^*(1, k).$$

2. For any $p \in (0, 1)$, $s_0^*(p, k)$ is strictly increasing in k , with range $[(\bar{a} + \underline{a})/2, \infty)$.

(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 $s \geq \frac{\bar{a} + \underline{a}}{2} - b$.¹⁵ Part 1 of [Proposition 1](#) says that when there is any uncertainty about the PM's type and the PM has reputation concerns, the equilibrium exhibits *pandering* in the sense that both PM types distort their behavior toward action \underline{a} , which the voter (correctly) believes is more likely to come from the congruent type.¹⁶ Part 2 establishes an intuitive monotonicity: the degree of pandering, measured by $s_0^* - s_0$, is increasing in the strength of the reputation concern, k ; furthermore, pandering vanishes as $k \rightarrow 0$, whereas both types of the PM take action \underline{a} with probability approaching one as $k \rightarrow \infty$.¹⁷ It follows that for any $p \in (0, 1)$, once k is

¹⁵ Recall that the hazard rate is $f/(1-F)$. Log-convexity of f on the relevant domain (part 2 of [Assumption 1](#)) implies that the hazard rate is non-increasing on this domain ([An, 1998](#)). Equilibrium uniqueness is not essential for the rest of [Proposition 1](#); interested readers are referred to the [Supplementary Appendix](#) for details.

¹⁶ Action \underline{a} may or may not be the ex-ante optimal action for the voter; this is immaterial to our analysis.

¹⁷ 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

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, $(\bar{a} + \underline{a})/2$, even though the biased type prefers lower thresholds than the congruent type. This point is analogous to the “populist bias” in [Acemoglu et al. \(2013\)](#).

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. Among other things, we will establish that the voter may prefer a PM who is known to be non-congruent over one who could be congruent or non-congruent.

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 :

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

This expected payoff function is strictly quasi-concave with a maximum at $(\bar{a} + \underline{a})/2$, which is the first-best threshold the voter would use if she could observe the state and choose 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

$$\mathcal{U}(p, k) := p[U(s_0^*(p, k)) - U(s_0^*(p, k) - b)] + U(s_0^*(p, k) - b), \quad (7)$$

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.

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 $(\tilde{k}(p), \infty)$.*

\bar{a} : \bar{p} in [Equation 2](#) goes up while \underline{p} in [Equation 3](#) goes down. Consequently, both types’ reputational incentive to take action \underline{a} increases.

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: if $k = 0$, neither type distorts its action, 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, $k \approx 0$ (but $k > 0$), 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 $k \rightarrow \infty$, 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.¹⁸

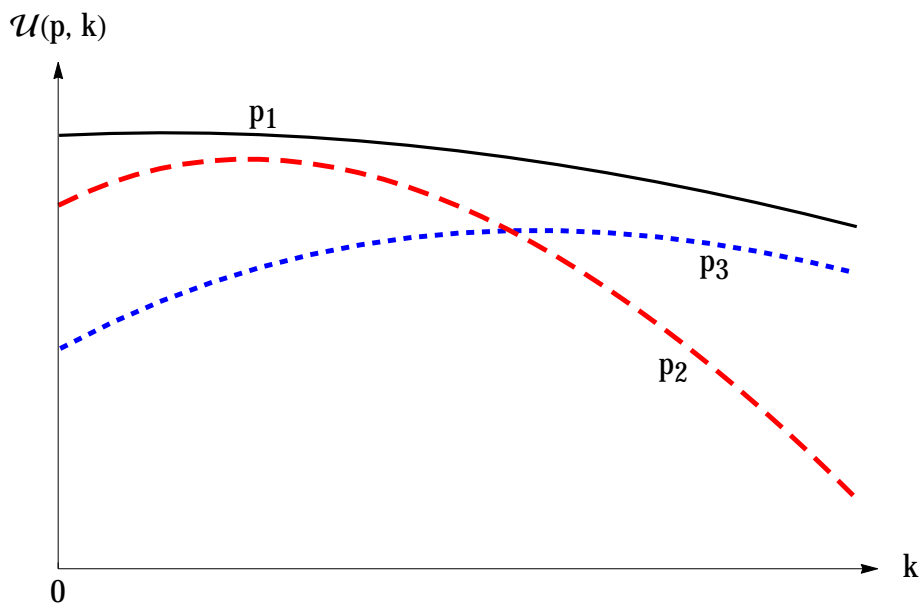


Figure 2 – Voter welfare as a function of PM’s reputation concern, with $p_1 > p_2 > p_3$.

Figure 2 depicts welfare as a function of the strength of reputation concern, computed for some representative parameters and three different values of p .¹⁹ Besides illustrating 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

¹⁸ 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 $\mathcal{U}(p, \cdot)$. Yet, as shown in the Supplementary Appendix, our main points continue to hold without the log-convexity assumption.

¹⁹ This and subsequent figures are computed with F being an exponential distribution with mean 10, $\underline{a} = 0$, $\bar{a} = 2$, and $b = 0.1$.

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 (dotted) 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 $p \rightarrow 0$, pandering vanishes, which can be preferable to excess pandering. Of course, welfare approaches the first-best as $p \rightarrow 1$, 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 any $k \geq 0$, $\mathcal{U}_p(0, k) > 0$ 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 $k > \hat{k}(p)$; (ii) $\hat{k}(p) \rightarrow \infty$ as either $p \rightarrow 0$ or $p \rightarrow 1$; and (iii) $\hat{k}(\cdot)$ is continuous.
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 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 $p = 0$ because of both a direct effect that the politician may be congruent, and, when $k > 0$, an indirect effect of causing the non-congruent type to use a preferable threshold.

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.²⁰ This “known devil may be better than unknown angel” property is a consequence of the facts that, for any $p \in (0, 1)$, pandering gets arbitrarily severe as $k \rightarrow \infty$ ([Proposition 1](#), part 2) and 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, then falls below the welfare level provided by a

²⁰ 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 $k \geq 0$, as there is no pandering no matter the value of k when $p = 0$.

PM who is known to be non-congruent (i.e., $\mathcal{U}(0, 0)$), and eventually increases again up to its maximum.

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 in Subsection 4.1.)

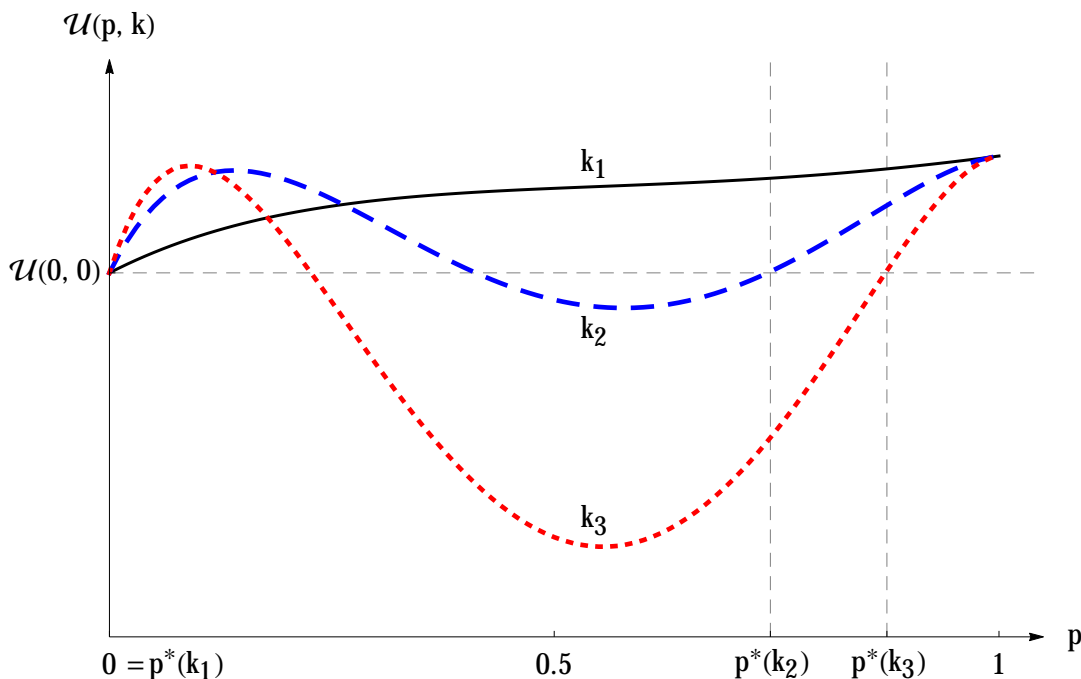


Figure 3 – Voter welfare as a function of her belief, with $k_1 < k_2 < k_3$.

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. The reason is simply that a higher p can exacerbate undesirable pandering. 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 property owes to the single-peakedness of $U(\cdot)$.²¹ Proposition 2

²¹ To see why, suppose (towards proving the contrapositive) the voter prefers the congruent PM's equilibrium threshold to that of the non-congruent PM. Then the non-congruent PM must be using a threshold below the first-best threshold, $(\bar{a} + \underline{a})/2$, which implies that both thresholds are preferred by the voter to $(\bar{a} + \underline{a})/2 - b$, the threshold used by the non-congruent PM when $p = 0$. Hence, $U(s_b^*(p, k)) \leq U(s_0^*(p, k)) \implies \mathcal{U}(p, k) \geq \mathcal{U}(0, 0)$.

thus implies that for any $p \in (0, 1)$, when reputation concerns are sufficiently strong, the voter prefers the non-congruent type's equilibrium behavior to the congruent type's equilibrium 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 to officeholding, a type- θ PM has expected payoff

$$W(\theta, p, k) := v_\theta - \int_{\underline{s}}^{s_\theta^*(p, k)} (\underline{a} - s - \theta)^2 f(s) ds - \int_{s_\theta^*(p, k)}^{\infty} (\bar{a} - s - \theta)^2 f(s) ds + k[F(s_\theta^*(p, k))V(\bar{p}(p, k)) + (1 - F(s_\theta^*(p, k)))V(\underline{p}(p, k))], \quad (8)$$

where $s_\theta^*(\cdot)$ denotes the equilibrium threshold used by type θ and $\bar{p}(\cdot)$ and $\underline{p}(\cdot)$ denote the voter's equilibrium beliefs after observing actions \underline{a} and \bar{a} respectively (see [Equation 2](#) and [Equation 3](#)).

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

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

Moreover, $W(0, p, k) > W(b, p, k)$, 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. The two equalities owe to $V(0) = 0$, $V(1) = 1$, and how we set v_θ ([fn. 12](#)).

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 for any $p \in (0, 1)$, the ex-post reputation of a congruent PM will on expectation be higher than that of a non-congruent PM, whereas their reputation will be the same if the prior is zero (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 $p \in (0, 1)$ and $k > 0$, $W(0, 1, k) - W(0, p, k) < W(b, 1, k) - W(b, p, k)$.²²

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 $i \in \{A, B\}$ 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 $1 - \mu^\theta$).²³ 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 8](#). [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 when $k > 0$ ([Lemma 2](#)), [Assumption 2](#) ensures that any reputationally-concerned 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.²⁴

²² 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.

²³ One can also interpret communication as being about what action a candidate would take if elected (as a function of the realized state). As we will see, in the relevant equilibria, candidates who announce they are biased will be more likely to take action \bar{a} .

²⁴ 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](#).

As messages are cheap talk, there is no loss of generality in restricting attention to equilibria in which $\mu^0 \geq \mu^b$. In words, a candidate’s announcement of congruence 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. Let p^m denote 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 $1 = \mu^0 > \mu^b > 0$.*

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 the voter’s welfare $\mathcal{U}(\cdot, k)$ is uniquely maximized at $p = 1$ ([Proposition 2](#)).

Remark 2. We will focus on semi-separating equilibria below. In general we cannot rule out the possibility of informative equilibria that are not semi-separating. [Lemma 3](#) implies that such equilibria must involve both types randomizing.²⁵ 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. 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 which 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

²⁵ 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.

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:

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

For any k , $p^*(k) < 1$ and $\mathcal{U}(p, k) > \mathcal{U}(0, k)$ for any $p > p^*(k)$. See [Figure 3](#), which indicates $p^*(\cdot)$ on the horizontal axis for different values of reputation concern. It is also useful to define

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

In words, k^* is the largest reputation concern such that the PM's pandering—no matter what belief he is elected with—cannot harm the voter relative to a known non-congruent PM. It follows from our earlier analysis ([Proposition 2](#)) that $k^* > 0$: every uncertain PM is preferred to a known non-congruent PM if and only if reputation concerns are not too strong.²⁶

Lemma 4. $p^*(k) = 0$ if and only if $k < k^*$, and $p^*(\cdot)$ is strictly increasing on $[k^*, \infty)$ with $\lim_{k \rightarrow \infty} p^*(k) = 1$.

The logic behind the monotonicity in [Lemma 4](#) 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

²⁶ Recalling the function $\hat{k}(\cdot)$ from part 2 of [Proposition 2](#), $k^* = \min\{\hat{k}(p) : p \in (0, 1)\} \in (0, \infty)$.

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 ensures 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.²⁷ 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.

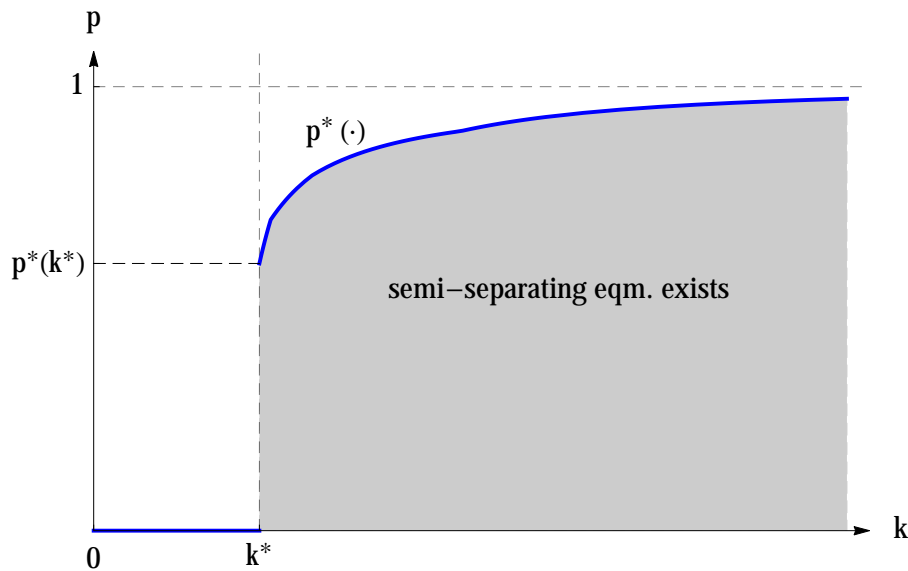


Figure 4 – Existence of semi-separating equilibrium.

Three points are noteworthy about a semi-separating equilibrium. First, the voter gets in-

²⁷ Recall that this property is assured by Assumption 1 (part 3), which may be violated if the bias parameter, b , is too large. In that case, semi-separating cheap-talk equilibria would not exist. But it is not always true that the scope for semi-separating equilibria decreases in b . Although the voter's utility from a known non-congruent candidate is lower when b is higher, a candidate of unknown type will also pander more in this case. Consequently, there are examples in which p^* is increasing in b for a range of parameters.

formation both about a candidate's type and about which action (contingent on the realized state) he will take in office; a candidate who reveals non-congruence reveals that he is more likely to take the high action if elected. Second, the electoral campaign alters a PM's behavior. The reason is that a PM of either type uses a policy threshold that depends on the voter's belief with which he is elected ([Proposition 1](#)). A non-congruent PM's behavior thus varies with his electoral announcement. Although a congruent PM always pronounces congruence, he is elected with a different (higher) belief than in the absence of communication, and in this sense his policymaking behavior is also affected by his announcement. Third, 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.²⁸ This property is not shared by other models of informative cheap talk in elections (e.g., [Kartik and McAfee, 2007](#); [Schnakenberg, 2016](#)).

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 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 $\mathcal{U}(0, 0)$.

In a semi-separating equilibrium, the voter's posterior when a candidate announces congruence, p^0 , is not affected by small changes in the prior, p ; rather, 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 a candidate is ex ante less likely to be congruent but also because μ^b is increasing 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:

²⁸ Timing assumptions are thus important: the prescribed strategies would not form an equilibrium if candidates' announcements were sequential. Nonetheless, informative cheap talk remains possible under sequential communication; both candidates' playing as in [Proposition 3](#) can be supported by having the voter treat the candidates asymmetrically, as is natural once timing creates an inherent asymmetry between candidates.

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 help or harm welfare.²⁹ 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.

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 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

²⁹ 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 for any k , $\mathcal{U}(p, k)$ is increasing in p for small p (Proposition 2, part 1). 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.

concerns. Define, for any $k \geq 0$,

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

as the set of priors for which a semi-separating equilibrium exists that 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. $\lim_{k \rightarrow \infty} P^k = (0, 1)$.
3. For any $k_1, p \in P^{k_1}$, and $k > k_1$, $\frac{\partial}{\partial k} [\mathcal{U}(0, 0) - \mathcal{U}(p, k)] > 0$.

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.

5. Extensions

5.1. 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 model. 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 $\bar{k} > 0$ such that for all $k > \bar{k}$, 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 if $p^b < p < p^0$ 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 $k \rightarrow \infty$, there is $\hat{p}(k) \rightarrow 0$ such that $\hat{p}(k)$ is a local maximizer of $\mathcal{U}(\cdot, k)$ and $\mathcal{U}(\hat{p}(k), k) \rightarrow \mathcal{U}(1, 0)$. 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 $k \rightarrow \infty$, 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.

5.2. Embedding in a dynamic model

We have studied a one-shot interaction between politicians and voters for simplicity. In follow-up work (with a different focus), [Kartik and Van Weelden \(2017\)](#) establish that our key reputational effects—the non-monotonicity of voter welfare in the belief about a PM's type, with a known devil sometimes preferred to an unknown angel—also emerge in an infinite-horizon model of repeated elections in which politicians are subject to a two-term limit. That framework micro-founds a first-term PM's reputation function, $V(\cdot)$, along the lines mentioned in [Section 2](#) wherein an incumbent runs for re-election against a random challenger. The resulting “overlapping generations” structure preserves a connection with the current paper despite the infinite horizon. While that paper does not study cheap talk, it is straightforward based on the current analysis that, for appropriate parameters, challengers can engage in informative cheap talk whereas incumbents who are re-running for office cannot (since a

PM's behavior in his second term is independent of voter beliefs). This asymmetry between challengers and incumbents is another potential empirical test of the theory.

5.3. More types or policies

We have focussed on a simple model in which 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, allowing for politicians who could be biased in either direction. 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 is more complicated, we discuss how informative cheap talk obtains in some richer specifications.

5.4. Observability of the state

We have assumed that the voter updates her belief about the PM's congruence by observing only his policy action, without any direct information about the state. This is an appropriate assumption for policies whose consequences are revealed with sufficient lags. Notwithstanding, our fundamental themes would be qualitatively unchanged even if the PM's reputation were influenced by some independent information about the state. Specifically, if the voter receives a noisy signal of the state, then under mild conditions, versions of [Proposition 1](#), [Proposition 2](#), and [Proposition 3](#) continue to hold.

5.5. Pre-election private information about the state

We have assumed that candidates have no private information about the policy-relevant state prior to the election. The [Supplementary Appendix](#) relaxes this assumption. We identify there an informative cheap-talk equilibrium when the extent of private information candidates have about the policy-relevant state is small relative to that about their own congruence. In that equilibrium, campaign statements are informative not only about candidates' congruence (and actions if elected), but also the policy-relevant state. Specifically, a non-congruent candidate only reveals that he is non-congruent when his private information sufficiently favors high states. As the voter's belief about the state (and the elected PM's congruence if he has not revealed himself as non-congruent) then depends on *both* candidates' announcements, so does the elected PM's behavior, despite the PM fully learning the state after the election. We also discuss in the [Supplementary Appendix](#) why, when the strength of reputation con-

cerns is large, communication about congruence remains central to the welfare benefits of informative cheap talk even when candidates have some private information about the state.

6. 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? Furthermore, how could cheap-talk campaigns affect candidates' post-election behavior?

This paper has developed a simple rationale for why costless and non-binding electoral communication can be informative and also influence policymaking. 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 behavior by officeholders. Plainly, there are a number reasons outside our model that electoral accountability is desirable. A novel lesson from our analysis is that cheap talk in elections can mitigate the distortions induced by accountability.

We close by mentioning some additional issues.

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 $\psi \geq 0$ 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 $\psi > 0$, 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 ψ . 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 ψ 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 $\psi > 0$, a semi-separating equilibrium is compatible with $\sigma > 1/2$, i.e., 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 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](#) and [Subsection 5.2](#). Then, even though the voter's welfare from electing a new policymaker may be non-monotonic in the probability of his congruence, the voter's welfare from re-electing an incumbent is monotonic in that probability. More generally, though, what if the voter's welfare from re-electing an incumbent is also non-monotonic in the probability of congruence, e.g., because there are no term limits? This is an interesting avenue for future research.

Broader implications. 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.³⁰ While we have focussed in this paper on the implications for information revelation in elections, we believe it would also be fruitful to study the phenomenon in other contexts in which reputational distortions are important, such as judiciaries, media, and organizations. For example, [Shapiro \(2016\)](#) argues that media reports would be more informative if journalists' partisan leanings were known; our results suggest that it may be possible for journalists to (partially) reveal such information themselves.

³⁰ [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 is entirely intuitive, the known devil can be better than the unknown angel.

Appendix: Proofs

Proof of Proposition 1. The discussion preceding the proposition explained why Equation 6 characterizes (interior) equilibria.

Step 1: We first establish that Equation 6 has a unique solution s_0^* . Since

$$\frac{1 - F(s_0^* - b)}{1 - F(s_0^*)} \geq 1 \geq \frac{F(s_0^* - b)}{F(s_0^*)}, \quad (9)$$

the right-hand side (RHS) of Equation 6 is non-negative for all s_0^* . The left-hand side (LHS) is non-negative if and only if $s_0^* \geq (\bar{a} + \underline{a})/2$. Hence, any solution has $s_0^* \geq (\bar{a} + \underline{a})/2$; we restrict attention in the remainder of the proof to this domain. Existence of a solution follows from continuity, as the RHS of Equation 6 is bounded in s_0^* while the LHS tends to ∞ as $s_0^* \rightarrow \infty$. For uniqueness, it is sufficient to show that the RHS of Equation 6 is non-increasing, because the LHS is strictly increasing.

Differentiating the RHS of Equation 6 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

$$\begin{aligned} RHS' &= \frac{k}{2(\bar{a} - \underline{a})} V' \left(\frac{1}{1 + \alpha G(s)} \right) [-1(1 + G(s))^{-2} \alpha G'(s)] - V' \left(\frac{1}{1 + \alpha H(s)} \right) [-1(1 + H(s))^{-2} \alpha H'(s)] \\ &= \frac{k\alpha}{2(\bar{a} - \underline{a})} \left[V' \left(\frac{1}{1 + \alpha H(s)} \right) \frac{H'(s)}{(1 + \alpha H(s))^2} - V' \left(\frac{1}{1 + \alpha G(s)} \right) \frac{G'(s)}{(1 + \alpha G(s))^2} \right], \end{aligned} \quad (10)$$

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 (10) 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 (9), simplifies to

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

The above inequality holds for all $s \geq (\bar{a} + \underline{a})/2$ because f is log-convex on that domain (part 2 of Assumption 1) and hence has a non-increasing hazard rate on that domain (An, 1998, Remark 5(i)).³¹

³¹ An (1998, Remark 5(i)) establishes that a cumulative distribution \tilde{F} with support $[x, \infty)$, where $x \in \mathbb{R}$, and log-convex density \tilde{f} has a non-increasing hazard rate on $[x, \infty)$. Let $x := (\bar{a} + \underline{a})/2$ and $\tilde{f}(s) := f(s)/(1 - F(s))$ for $s \geq x$. Then, on the domain $[x, \infty)$, f log-convex is equivalent to \tilde{f} log-convex (using the fact that a non-negative function $l(\cdot)$ is log-convex if and only if $l(\lambda s + (1 - \lambda)t) \leq [l(s)]^\lambda [l(t)]^{1 - \lambda}$ for all s, t and $\lambda \in [0, 1]$), which

Step 2: Let the unique solution to [Equation 6](#) be denoted $s_0^*(p, k)$. Since both sides of [Equation 6](#) are continuously differentiable in all arguments, the implicit function theorem (which can be invoked because the derivative of the LHS with respect to s_0^* is 1 while that of the RHS is non-positive, by the argument in Step 1) ensures that $s_0^*(p, k)$ is continuously differentiable in p and k .

Step 3: We now prove parts 1 and 2 of [Proposition 1](#).

For part 1, note that when $k > 0$, our assumption that $V(\cdot)$ is strictly increasing ensures that the RHS of [Equation 6](#) is strictly positive for any $p \in (0, 1)$. Therefore, $s_0^*(p, k) > (\bar{a} + \underline{a})/2$ for any $p \in (0, 1)$ and $k > 0$. However, when $p \in \{0, 1\}$ the RHS is equal to 0, and hence $s_0^*(0, k) = s_0^*(1, k) = (\bar{a} + \underline{a})/2$.

For part 2, fix an arbitrary $p \in (0, 1)$. First note that $s_0^*(p, k)$ is strictly increasing in k because the RHS of [Equation 6](#) is non-increasing in s_0^* (by Step 1) and strictly increasing in k , given $s_0^* \geq (\bar{a} + \underline{a})/2$. That $s_0^*(p, 0) = (\bar{a} + \underline{a})/2$ follows from the fact that the RHS of [Equation 6](#) is 0 when $k = 0$. That $s_0^*(p, k) \rightarrow \infty$ as $k \rightarrow \infty$ follows from the fact that, for any s_0^* , the RHS tends to ∞ as $k \rightarrow \infty$. \square

Proof of [Lemma 1](#). Recalling the definition

$$U(\tau) \equiv - \int_{\underline{s}}^{\tau} (\underline{a} - s)^2 f(s) ds - \int_{\tau}^{\infty} (\bar{a} - s)^2 f(s) ds,$$

we compute

$$U'(\tau) = (\bar{a} - \underline{a})(\bar{a} + \underline{a} - 2\tau) f(\tau). \quad (11)$$

Partially differentiating [Equation 7](#) and suppressing the arguments of $s_0^*(\cdot)$,

$$\begin{aligned} \mathcal{U}_k(p, k) &= [pU'(s_0^*) + (1-p)U'(s_0^* - b)] \frac{\partial s_0^*}{\partial k} \\ &\propto pU'(s_0^*) + (1-p)U'(s_0^* - b) \\ &= (\bar{a} - \underline{a})[(\bar{a} + \underline{a} - 2s_0^*)pf(s_0^*) + (\bar{a} + \underline{a} - 2s_0^* + 2b)(1-p)f(s_0^* - b)] \\ &\propto \left(\frac{\bar{a} + \underline{a}}{2} - s_0^* \right) + \frac{(1-p)bf(s_0^* - b)}{pf(s_0^*) + (1-p)f(s_0^* - b)}, \end{aligned} \quad (12)$$

where the first proportionality uses $\frac{\partial s_0^*}{\partial k} > 0$ ([Proposition 1](#)), the equality uses [Equation 11](#), and the second proportionality obtains from a division by $2(\bar{a} - \underline{a})(pf(s_0^*) + (1-p)f(s_0^* - b)) > 0$.

Fix any $p \in (0, 1)$. Expression (12) is strictly positive as $k \rightarrow 0$ because $s_0^* \rightarrow \frac{\bar{a} + \underline{a}}{2}$ as $k \rightarrow 0$ ([Proposition 1](#)) and the last fraction in (12) is strictly positive and bounded away from zero as $s_0^* \rightarrow \frac{\bar{a} + \underline{a}}{2}$. Analogously, (12) is strictly negative for large k because $s_0^* \rightarrow \infty$ as $k \rightarrow \infty$ and the last fraction is always less than one. Therefore, it suffices to show that expression (12) has a unique zero, i.e., that

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

implies $\tilde{f}/(1 - \tilde{F})$ non-increasing, which is equivalent to $f/(1 - F)$ non-increasing.

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 $\left[\frac{\bar{a} + \underline{a}}{2} - b, \infty\right)$ (part 2 of [Assumption 1](#)), $s_0^* \geq (\bar{a} + \underline{a})/2$, and $b > 0$. \square

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

Part 1: Partially differentiating [Equation 7](#) with respect to p yields

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

where the second equality uses [Equation 11](#).

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

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

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 $(\bar{a} + \underline{a})/2$.

That $\mathcal{U}(\cdot, k)$ is uniquely maximized at $p = 1$ follows from [Proposition 1](#) establishing that $s_0^*(1, k) = (\bar{a} + \underline{a})/2 = s_{FB}$, 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 $s_0^*(p, 0) = (\bar{a} + \underline{a})/2$,

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

Since $\lim_{k \rightarrow \infty} s_0^*(p, k) = \infty$ ([Proposition 1](#)),

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

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

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

or, equivalently, if and only if

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

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

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

which is precisely what was assumed in part 3 of [Assumption 1](#).

Therefore, $\mathcal{U}(p, 0) > \mathcal{U}(0, 0) > \lim_{k \rightarrow \infty} \mathcal{U}(p, k)$, 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 $k > \hat{k}(p)$. Hence, $\mathcal{U}_k(p, \hat{k}(p)) < 0$, and $\hat{k}(\cdot)$ is continuous by the implicit function theorem.

To see that $\hat{k}(p) \rightarrow \infty$ as $p \rightarrow 0$ or as $p \rightarrow 1$, 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 $(\bar{a} + \underline{a})/2$, (ii) for any k , $s_0^*(p, k) > (\bar{a} + \underline{a})/2$ for any $p \in (0, 1)$ but $s_0^*(p, k) \rightarrow (\bar{a} + \underline{a})/2$ as $p \rightarrow 0$ or as $p \rightarrow 1$, and (iii) $\mathcal{U}(p, k)$ is given by expression [Equation 7](#) whereas $\mathcal{U}(0, 0) = U((\bar{a} + \underline{a})/2 - b)$, 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

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

Note that because of how we set v_{θ} ([fn. 12](#)),

$$\begin{aligned} v_{\theta} &= \int_{\underline{s}}^{\frac{\bar{a}+\underline{a}}{2}-\theta} (\underline{a} - s - \theta)^2 f(s) ds + \int_{\frac{\bar{a}+\underline{a}}{2}-\theta}^{\infty} (\bar{a} - s - \theta)^2 f(s) ds \\ &= -\tilde{U}(s_{\theta}, \theta), \end{aligned} \tag{13}$$

where $s_{\theta} = (\bar{a} + \underline{a})/2 - \theta$ is the threshold type θ would use in the absence of reputation concern.

For the rest of the proof, fix any $p \in (0, 1)$ and $k > 0$. 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 8](#), 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

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

where the first equality uses the definition of $W(\cdot)$ and $\tilde{U}(\cdot)$, the first inequality uses $s_\theta^*(\cdot) > s_\theta$ (and s_θ is the unique maximizer of $\tilde{U}(\cdot, \theta)$), 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

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

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

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

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 $\bar{p}(p, k) > \underline{p}(p, k)$, 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: $s_b^*(p, k) > (\bar{a} + \underline{a})/2 = s_0$. 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,

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

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) $(\bar{a} - s - b)^2 < (\underline{a} - s - b)^2$ if $s > s_0 - b$ and (ii) $(\bar{a} - s - b)^2 - (\underline{a} - s - b)^2 < (\bar{a} - s)^2 - (\underline{a} - s)^2$ for any s . \square

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 $m_A \neq m_B$ the voter must elect the candidate who announced j , and if $m_A = m_B$ 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.

Finally, note that there cannot be an equilibrium with $\mu^0 > \mu^b = 0$ because that would induce $p^0 = 1 > p^b$ and hence $\mathcal{U}(p^0, k) > \mathcal{U}(p^b, k)$. It follows that there does not exist a separating equilibrium and any semi-separating equilibrium has $1 = \mu^0 > \mu^b > 0$. \square

Proof of Lemma 4. First note that $k^* = \min\{\hat{k}(p) : p \in (0, 1)\}$, where for any $p \in (0, 1)$, $\hat{k}(p)$ was defined in part 2 of Proposition 2 as the unique positive solution to $\mathcal{U}(p, \hat{k}(p)) = \mathcal{U}(0, 0)$. It follows from the properties of $\hat{k}(\cdot)$ established in Proposition 2 that $k^* \in (0, \infty)$. That $p^*(k) = 0$ if and only if

$k < k^*$ then follows from the definition of $p^*(\cdot)$, that $k^* = \min\{\hat{k}(p) : p \in (0, 1)\}$, and $\mathcal{U}_p(0, k) > 0$ for all k ([Proposition 2](#)).

Next, 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, since $k = \hat{k}(p^*(k))$ for $k \geq k^*$ and $\hat{k}(\cdot)$ is continuous and unbounded ([Proposition 2](#)), it follows that $p^*(k) \rightarrow 1$ as $k \rightarrow \infty$. \square

Proof of Proposition 3. We show that a semi-separating equilibrium exists if and only if $p \in (0, p^*(k))$; note that this condition implies $k \geq k^*$. By [Lemma 3](#), any semi-separating equilibrium has $1 = \mu^0 > \mu^b > 0$ and voter believes $p^0 > p > p^b = 0$ 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 $p \in (0, p^*(k))$, 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 $w^0 := p + (1-p)\mu^b \in (0, 1)$ 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 $W(b, 0, k) = 0$ ([Lemma 2](#)), and 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,

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

As the LHS of [Equation 15](#) is strictly increasing in σ while the RHS is strictly decreasing in it, there is at most one value of σ that solves [Equation 15](#). When $\sigma = 0$, the RHS of [Equation 15](#) is strictly larger than the LHS because of [Assumption 2](#), $w^0 < 1$, and $W(b, p^0, k) < k$ ([Lemma 2](#)). When $\sigma = 1/2$, the LHS is strictly larger than the RHS because $W(b, p^0, k) > 0$ 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 $p^*(k) \rightarrow 1$ as $k \rightarrow \infty$ ([Lemma 4](#)). \square

Proof of Corollary 1. As explained before the corollary, the result follows from [Proposition 2](#). \square

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

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

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 $\hat{k}(p) \rightarrow \infty$ as $p \rightarrow 1$ ([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 $k > k_1$, $\frac{\partial \mathcal{U}(p, k)}{\partial k} < 0$. \square

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 $s_b^*(p, k) > (\bar{a} + \underline{a})/2$ and define

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

(This is well-defined by [Proposition 1](#).) Since $U(\cdot)$ is strictly decreasing above $(\bar{a} + \underline{a})/2$, 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

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

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

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

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

References

- ACEMOGLU, D., G. EGOROV, AND K. SONIN (2013): "A Political Theory of Populism," *Quarterly Journal of Economics*, 128, 771–805.
- AGRANOV, M. (2016): "Flip-Flopping, Primary Visibility and Selection of Candidates," *American Economic Journal: Microeconomics*, 8, 61–85.
- 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.
- ASH, E., M. MORELLI, AND R. VAN WEELDEN (2017): "Elections and Divisiveness: Theory and Evidence," *Journal of Politics*, 79, 1268–1285.
- 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.
- BIDWELL, K., K. CASEY, AND R. GLENNERSTER (2016): "Debates: Voting and Expenditure Responses to Political Communication," Unpublished.
- 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.
- CARRILLO, J. D. AND M. CASTANHEIRA (2008): "Information and Strategic Political Polarization," *Economic Journal*, 118, 845–874.
- 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.
- DAVIS, J. H. (2015): "Mount McKinley Will Again Be Called Denali," *New York Times*, August 30.
- 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.
- FERNANDEZ-VASQUEZ, P. (2014): "Signaling Policy Positions in Election Campaigns," Unpublished.
- FOX, J. AND M. STEPHENSON (2015): "The Welfare Effects of Minority-Protective Judicial Review," *Journal of Theoretical Politics*, 27, 499–521.
- FUDENBERG, D. AND J. TIROLE (1991): *Game Theory*, Cambridge, MA: MIT Press.
- GRILLO, E. (2016): "The Hidden Cost of Raising Voters' Expectations: Reference Dependence and Politicians' Credibility," *Journal of Economic Behavior and Organization*, 130, 126–143.
- GROSSER, J. AND T. R. PALFREY (2014): "Candidate Entry and Political Polarization: An Antimedean 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.
- HOFFMAN, M. AND E. LYONS (2017): "A Time to Make Laws and a Time to Fundraise? On the Relation between Salaries and Time Use for State Politicians," Unpublished.

- 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 (2015): "Information Revelation and Pandering in Elections," Unpublished.
- KARTIK, N. AND R. VAN WEELDEN (2017): "Reputation Effects and Incumbency (Dis)Advantage," 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.
- MCGREGOR, J. (2010): "Why Mike Bloomberg is a Real Leader," *Washington Post*, August 15.
- MOEN, E. R. AND C. RIIS (2010): "Policy Reversal," *American Economic Review*, 100, 1261–1268.
- MORELLI, M. AND R. VAN WEELDEN (2013): "Ideology and Information in Policymaking," *Journal of Theoretical Politics*, 25, 412–439.
- MORRIS, S. (2001): "Political Correctness," *Journal of Political Economy*, 109, 231–265.
- OBAMA, B. (2014): "Remarks by the President on the Economy," *White House Office of The Press Secretary*, July 10.
- OSBORNE, M. J. AND A. SLIVINSKI (1996): "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics*, 111, 65–96.
- PANOVA, E. (2017): "Partially Revealing Campaign Promises," *Journal of Public Economic Theory*, 19, 312–330, forthcoming.
- 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.
- ROGOWSKI, J. AND P. TUCKER (2016): "Moderate, Extreme, or Both? How Voters Respond to Ideologically Unpredictable Candidates," Unpublished.

- SCHARFSTEIN, D. S. AND J. C. STEIN (1990): "Herd Behavior and Investment," *American Economic Review*, 80, 465–79.
- SCHNAKENBERG, K. (2016): "Directional Cheap Talk in Electoral Campaigns," *Journal of Politics*, 78, 527–541.
- SHAPIRO, J. (2016): "Special Interests and the Media: Theory and an Application to Climate Change," *Journal of Public Economics*, 144, 91–108.
- 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.
- SULKIN, T. (2009): "Campaign Appeals and Legislative Action," *Journal of Politics*, 71, 1093–1108.
- TOMZ, M. AND R. P. VAN HOUWELING (2009): "The Electoral Implications of Candidate Ambiguity," *American Political Science Review*, 103, 83–98.
- WITTMAN, D. (1983): "Candidate Motivation: A Synthesis of Alternatives," *American Political Science Review*, 77, 142–157.