On the Rhetorical Strategies of Leaders: Speaking Clearly, Standing Back, and Stepping Down

Torun Dewan (LSE)

David P. Myatt (Oxford)

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

On the Rhetorical Strategies of Leaders:
Speaking Clearly, Standing Back, and Stepping Down

Torun Dewan

Department of Government and STICERD, London School of Economics and Political Science

`t.dewan@lse.ac.uk`

David P. Myatt

Nuffield College and Department of Economics, University of Oxford

`david.myatt@economics.ox.ac.uk`

April 17, 2009.[1]

Abstract. Followers wish to coordinate their actions in an uncertain environment. A follower would like his action to be close to some ideal (but unknown) target; to reflect his own idiosyncratic preferences; and to be close to the actions of others. He learns about his world by listening to leaders. Followers fail to internalize the full benefits of coordination and so place insufficient emphasis on the focal views of relatively clear leaders. A leader sometimes stands back, by restricting what she says, and so creates space for others to be heard; in particular, a benevolent leader with outstanding judgement gives way to a clearer communicator in an attempt to encourage unity amongst her followers. Sometimes a leader receives no attention from followers, and sometimes she steps down (says nothing); hence a leadership elite emerges from the endogenous choices of leaders and followers.

In an uncertain world, leaders can help followers to make more informed decisions and to coordinate; indeed, Schelling (1960, p. 91) suggested that "the coordination game lies behind the stability of institutions and traditions and perhaps the phenomenon of leadership itself" and recently others have emphasized the role of leaders as coordinating focal points (Calvert, 1995; Myerson, 2004; Dewan and Myatt, 2007, 2008; Bolton, Brunnermeier, and Veldkamp, 2008b). A leader with good judgement can provide useful information to resolve uncertainties, and a leader who communicates clearly can provide a common message around which followers' actions can coalesce.

Within this setting a variety of leadership institutions might emerge, including a single focal leader or an oligarchic elite.[2] In this context, what factors determine the form of a leadership institution? Do the leaders who endogenously emerge succeed in helping their followers to achieve their objectives? And what rhetorical strategies might leaders employ to pursue their own objectives, whether benevolent or otherwise?

---

[2]A leader could personify any focal information source and so, under this broader definition, leadership institutions can include published media, organized conferences, or even informal social discussion.

To answer these questions we analyze a variant of the formal model of Dewan and Myatt (2008) in which followers play a coordination game. A follower would like his action to be close to some (unknown) common target; to reflect his own idiosyncratic preferences; and to be close to the actions of others. He learns about the target and about others' likely behavior by listening to leaders. The set of leaders is broad enough to encompass information sources which extend beyond the leadership provided by individuals, and so followers may incorporate their prior beliefs, talk to others, read newspapers, and engage in private research. We find it helpful to think of a follower as the member of a party whose action corresponds to the advocacy of a policy. In this context the common target is the party's ideal policy; the idiosyncratic preference is the party member's personal policy bias; and the desire to coordinate stems from a concern for party unity.

A leader's skills are her ability to ascertain the right thing to do (judgement) and her ability to convey ideas (clarity). A follower is influenced by a weighted average of what he hears. He places relatively greater weight on the words of relatively clearer communicators; they are better placed to coordinate him with others. Furthermore, when followers choose to whom to listen then they focus attention on the clearest communicators.

Nevertheless, followers' decisions are inefficient: their emphasis on relatively clear leaders is socially insufficient. Clear orators promote unity by offering a focal point for actions, but followers fail to incorporate fully the impact of unity; a follower would improve (on average) the lot of others by moving toward the perceived center.

Insufficient unity suggests a role for a leader as a strategic actor. We allow leaders to place varying emphases on the twin objectives of reaching an ideal target and coordinating on a common goal. For a party leader, for example, the former objective corresponds to her concern for policy whilst the latter objective reflects her desire for party unity. A unity-seeking leader, for example, is someone who would like followers to place greater emphasis on the coordination motive; a benevolent leader satisfies this definition. In contrast, a policy-seeking leader would like to see greater emphasis on the group target.

We allow each leader to alter her clarity (she either speaks clearly or obfuscates) or the length of time for which her views are heard (she either stands forth or stands back). By standing forth and speaking clearly a leader (often, but not always) attracts influence; however, she may be better served by pushing influence toward other leaders. Both obfuscation and standing back serve to do this; the difference is that standing back (limiting what she says) frees up time for followers to listen to others. Amongst other results, we find that a unity-seeking leader stands back if and only if she is a poorer communicator, and a policy-seeking leader stands back if and only if she is a clearer communicator. In some circumstances a leader may cede all of her influence by not speaking at all, or by

speaking incoherently. If she does this then she steps down from the leadership platform, giving way to other leaders.

Rhetorical strategies provide a mechanism by which some leaders come to the fore whilst others play a lesser role. Thus our model allows us to analyze the size and character of the leadership elite that emerges endogenously. For example, in a two-leader scenario, leadership could be provided by one leader or by both, and this distinction corresponds to the institutional forms of dictatorship and oligarchy respectively. When followers entertain the views of both leaders, but one of them steps down, then a would-be oligarchy becomes a dictatorship. Conversely, if followers coalesce around the views of a single leader then if that de facto dictator stands back then she allows space for other opinions; a follower-generated dictatorship reverts to an oligarchy. We show that, when followers pay attention to both leaders, a dictatorship can emerge when leaders wish to promote unity. A special case is when leaders are benevolent, since such leaders recognize that followers place insufficient weight on the coordination motive. Conversely, a dictator magnanimously yields to others when she cares little for the coordination motive.

A discussion of related literature is postponed to our concluding section, and so we highlight only a selection of contributions here. The central tension between doing the right thing and doing it together relates to the "beauty contest" scenario described by Keynes (1936, Chapter 12) and developed formally by Morris and Shin (2002, 2005), Angeletos and Pavan (2004, 2007, 2008), Hellwig (2005), and Calvó-Armengol and de Martí Beltran (2007, 2009), amongst others. As we noted in our opening remarks, we view leaders as helping to inform followers' actions and as focal points for their coordination (Schelling, 1960; Calvert, 1995; Myerson, 2004; Dewan and Myatt, 2007, 2008; Dickson, 2008).

Turning to leaders' skills, the leadership acts used to communicate credibly a leader's information were considered by Hermalin (1998, 2007), whilst a leader's judgement (her ability to spot the right thing to do) is important for recent theories offered by Bolton, Brunnermeier, and Veldkamp (2008b) and by Majumdar and Mukand (2008). The analysis of rhetorical strategies relates our work to that of Hafer and Landa (2007), who examined situations where agents have latent dispositions to some types of arguments. Obfuscation is related to strategic ambiguity (Shepsle, 1972; Page, 1976; McKelvey, 1980), an idea which has been developed more recently (Meirowitz, 2005; Blume and Board, 2009) and also occurs in the contemporary economics literature (Ferreira and Rezende, 2007). Finally our idea of "standing back" resonates with the analyses of voting models which suggest that abstention can be beneficial (Feddersen and Pesendorfer, 1997).

A Model of Policy Performance and Party Unity

Our study of leadership takes place within the context of a game in which players (that is, the followers) wish to coordinate their actions in an uncertain environment. A follower would like his action to be close to some (unknown) common target; to reflect his own idiosyncratic preferences; and to be close to the actions of others.

There are many social situations of this type. For example, freedom fighters are most effective when they coordinate an attack by using complementary tactics. They fail if they do not discover the identity of the best target, or when differences of opinion or of expectations frustrate coordination. As a second example, the religious members of a congregation may seek revelation of a fundamental truth unknown to them, and wish to live in accordance with their (possibly different) perceptions of that truth. They are drawn together by a desire for harmony but driven apart by personal prejudices and proclivities. In a more mundane setting, committee members may hold different opinions about the right solution to a prevailing problem and yet desire a coordinated response.

It is helpful to use a specific example to explain the mechanisms of our (intentionally) abstract model, and so we follow Dewan and Myatt (2008) by using a political party as a vehicle for exploring our ideas. Party activists wish to promote and support a policy which is best suited to their party's environment, but are also concerned with party unity. Problems may arise when the best policy is unknown and when individuals entertain idiosyncratic policy biases. Our aim is to assess the role played by leaders who bridge differences of opinion and unite the party despite any policy disagreements. Benevolent leaders seek to enhance the aggregate welfare of activists; alternatively, leaders may differ from their followers in the relative emphasis which they place on policy performance and party unity. We will characterize the rhetorical strategies—speaking clearly, standing back, and stepping down—which a leader may seek to employ.

To express our ideas formally we analyze a simultaneous-move game played by a mass of activist party members indexed by m.[3] Activist m must advocate a policy $a_m \in \mathbb{R}$. The policy a_m may be interpreted as the position he supports while attending a party conference, or the policy he promotes on the doorstep during an election campaign. Similarly, the real line $\mathbb{R}$ can be interpreted as a familiar left-right policy spectrum.

An activist would like to advocate the best policy. We write θ for the ideal (but initially unknown) policy for the party; this is the right thing to do when individual considerations are put aside. An activist's personal policy bias is b_m. Combining these elements,

---

[3]Throughout our paper we assume that the player set is a unit mass, so that $m \in [0, 1]$. However, our model can be easily modified to a world with a finite collection of M activists.

his payoff declines as his chosen action deviates away from $\theta + b_m$ via a quadratic-loss function $(a_m - \theta - b_m)^2$. We assume that the policy bias varies across the party membership: it has zero mean (without loss of generality) and its variance β^2 captures the extent of fundamental disagreement amongst the party membership. Finally, when (in the next section) we specify the activist's sources of information about θ we assume that those sources are independent of his personal policy bias.

A second motivation for an activist is party unity: he wishes to coordinate with his fellow activists. A simple way of expressing the desire for unity is to imagine that he aims to minimize the difference between his own action and those of others. We capture this desire via the quadratic loss $\mathrm{E}[(a_m - a_{m'})^2]$ where the expectation is taken by looking across other party members indexed by m'.[4]

Summarizing, an activist balances his concerns for policy and for party unity. He would like do the right thing (subject to his own policy bias, of course) but do it together with others. We write the weights he places on his two concerns as π and $1 - \pi$ respectively. Bringing the ingredients together yields a payoff:

$$u_m = \bar{u} - \underbrace{\pi(a_m - \theta - b_m)^2}_{\text{(i) concern for policy}} - \underbrace{(1 - \pi)\,\mathrm{E}[(a_m - a_{m'})^2]}_{\text{(ii) desire for unity}} \tag{1}$$

where the expectation in the desire-for-unity term is taken by looking across the party-wide set of other activists. An activist maximizes his expectation of u_m conditional on any information available to him and on the conjectured behavior of others.

Since we have yet to describe activists' information sources our game is not fully defined; we specify fully those sources in the next section. Nevertheless, we can already explore the relationship between activists' privately optimal actions and those that maximize the aggregate welfare of the party's membership. Noting that any expectations are taken with respect to the beliefs of activist m, and are therefore conditional on any information available to him, the expected loss from his policy concern is $\mathrm{E}[(a_m - \theta - b_m)^2]$. Turning to his concern for party unity, his loss can be decomposed so that

$$\underbrace{\mathrm{E}[(a_m - a_{m'})^2]}_{\text{disunity}} = \underbrace{\mathrm{E}[(a_m - \bar{a})^2]}_{\text{non-conformity}} + \underbrace{\mathrm{E}[(a_{m'} - \bar{a})^2]}_{\text{others' non-conformity}} \quad \text{where} \quad \bar{a} \equiv \mathrm{E}[a_{m'}]. \tag{2}$$

$\bar{a}$ is the party line: the average policy advocated across the party's membership. Notice that the first component of disunity corresponds to an activist's non-conformity with the party line. The second element corresponds to other activists' non-conformity and so is seen as exogenous by activist m. Thus an individual acts to minimize a weighted average

---

[4]More formally, the loss from disunity experienced by activist m is $\mathrm{E}[(a_m - a_{m'})^2] \equiv \int_0^1 (a_m - a_{m'})^2\, dm'$. In a world with a finite collection of M activists, disunity can be defined as $\sum_{m' \neq m} (a_m - a'_m)^2 / (M - 1)$.

of his expected deviation from his ideal policy and his expected non-conformity. That is, a_m is chosen to minimize $\pi \, \mathrm{E}[(a_m - \theta - b_m)^2] + (1-\pi) \, \mathrm{E}[(a_m - \bar{a})^2]$. Solving the optimization problem straightforwardly generates a simple lemma.

Lemma 1. *An activist's individually optimal action is $a_m = \pi(\mathrm{E}[\theta] + b_m) + (1 - \pi) \, \mathrm{E}[\bar{a}]$ where the expectations are conditional on his information. This is a weighted average of his perception of the ideal policy and of the party line. His action places insufficient weight on party unity: party welfare would be enhanced if he put more emphasis on following the party line.*

From a party-welfare perspective, activists are insufficiently concerned with following the party line. Examining (2), notice that an activist benefits from the conformity of others; he would be better off if his fellow party members all moved inward. Alas, activists do not internalize this spillover effect: extreme views prevail even though all would benefit from a move to the center. This suggests an investigation into institutional remedies that might help activists resolve their collective-action problem. Leadership can provide one such resolution, and we explore its role in the following section.

LEADERSHIP IN AN UNCERTAIN WORLD

An activist is uncertain about which policy is best for the party and which policy is likely to be advocated by others; he knows his own policy bias b_m but begins with no substantive knowledge of θ.[5] Prior to acting, he seeks to understand his environment. He learns by observing a collection of informative signals. These signals may stem from various sources: an activist entertains his own prior belief, engages in private correspondence and discussion with others (perhaps at a party conference), or observes a more public signal of the best policy.

Here we think of the informative signals as n speeches made by n party leaders, and the ingredients of these speeches as reflecting different leaders' skill sets. The leaders' speeches capture all external information relevant to an activist's play. A so-called speech can also be seen as a label for any information source. The properties of an informative signal (for instance, the knowledge of a newspaper's correspondents and the eloquence of its writers) can replicate the qualities of a conceptual leader.

What then are the ingredients of a speech? We consider two crucial components of a leader's skill set: the quality of her judgement and the clarity of her message. Considering the first component, a leader is distinguished from activists in that she has substantive

---

[5]Formally, we assume that activists share a diffuse prior over θ although it is straightforward to extend our analysis to a world in which activists share a common prior belief $\theta \sim N(\mu, \xi^2)$.

information about the underlying ideal policy θ. Specifically, prior to making a speech, leader $i \in \{1, \ldots, n\}$ forms an opinion s_i about the best policy:

$$s_i = \theta + \eta_i \quad \text{where} \quad \eta_i \sim N(0, \kappa_i^2) \quad \text{so that} \quad \frac{1}{\kappa_i^2} \equiv \textit{Quality of Judgement.} \qquad (3)$$

Conditional on θ we assume (with little loss of generality) that the leaders' opinions are independent; hence the the collection of noise terms (η_i for each i) are uncorrelated.[6] Some opinions are more accurate reflections of θ than others: when κ_i^2 is small a leader is better able to assess policy. Inverting this variance, the precision $1/\kappa_i^2$ is a leader's quality of judgement. Good judgement may stem from innate ability, or it may arise from the quality of a leader's private information sources.

Upon forming her opinion a leader addresses the activists. We assume that she has no policy bias and conveys truthfully her opinion. However, the clarity of her message may be compromised: whilst some leaders speak clearly and coherently, others are less audible or adopt clumsy rhetoric. Furthermore, the clarity of her message may also be influenced by both the length of time for which she speaks and the willingness of activists to listen to her. For a particular leader i, these factors are reflected in two parameters σ_i^2 and x_{im} which we discuss shortly. Formally, when a leader stands to deliver a speech s_i, an activist m forms an imperfect interpretation $\tilde{s}_{im}$ where

$$\tilde{s}_{im} = s_i + \varepsilon_{im} \quad \text{where} \quad \varepsilon_{im} \sim N\left(0, \frac{\sigma_i^2}{x_{im}}\right) \quad \text{so that} \quad \frac{x_{im}}{\sigma_i^2} \equiv \textit{Clarity of Message.} \qquad (4)$$

We assume that the various noise terms ε_{im} are uncorrelated, and so each activist receives a conditionally independent signal of a leader's opinion.[7]

The noise in the communication between a leader and follower is determined by the variance σ_i^2/x_{im}. The denominator x_{im} represents the length of time which activist m devotes to listening to what leader i has to say. It is equivalent to a sample size and so it is natural to assume that the precision x_{im}/σ_i^2 increases linearly with it. The numerator σ_i^2 is determined by the oratorical flair of the leader or by the nature of the medium through which a message is transmitted. Under this specification, and so long as $\sigma_i^2 > 0$, different activists hear different things. Variations in opinion emerge amongst activists that would be absent if a leader communicated with perfect clarity (so that $\sigma_i^2 = 0$). Furthermore, these are disagreements over the content of a leader's message and so are differences of opinion, rather than the fundamental differences which arise from activists' personal policy biases.

A speech by a clear communicator can help to unite a party: activists not only understand the message, but also know that others listening to the same speech perceive its content

---

[6]Unconditionally, of course, the opinions are correlated since they are all opinions about θ.

[7]This is without loss of generality: any common shock in the ε_{im} noise terms can be absorbed into η_i.

similarly. By contrast, if a leader lacks communicative ability then her speech may be divisive: activists are unclear what message is being conveyed, form different opinions about which policy the leader recommends they advocate, and so activists fail to develop a common understanding of the merits of different policies.

For now we treat a leader's clarity of communication σ_i^2 and each activist's attention x_{im} as exogenous, and we also assume that $x_{im} = x_i$ for all m. Later in the paper, these parameters will become endogenous: activists will choose to whom to listen, and leaders will choose both the length of their speeches (hence placing an upper bound on x_i) and will be free to obfuscate (hence increasing σ_i^2) if they so wish.

<div align="center">OPTIMAL POLICY ADVOCACY</div>

In the context of a Bayesian Nash equilibrium (in which activists react optimally to beliefs which are consistent with the anticipated play of others) it is without loss of generality to restrict attention to symmetric equilibria in which all activists use the same advocacy strategy. Formally this means that activist m selects an action $a_m = A(\tilde{s}_m, b_m)$ where $A(\cdot, \cdot) : \mathbb{R}^{n+1} \mapsto \mathbb{R}$ is a mapping from the speeches heard by an activist and his own policy bias into the policy space. Two activists who share the same policy bias and interpret the speeches they hear in the same way will take the same action.

In order to compute his optimal move, an activist forms beliefs about the underlying ideal policy and about the policies advocated by others. Combining these elements with Lemma 1, the strategy $A(\cdot, \cdot)$ forms an equilibrium if and only if

$$A(\tilde{s}_m, b_m) = \pi \left(\mathrm{E}[\theta \,|\, \tilde{s}_m] + b_m \right) + (1 - \pi) \mathrm{E}[A(\tilde{s}_{m'}, b'_m) \,|\, \tilde{s}_m]. \tag{5}$$

Biases are orthogonal to other elements of an activist's world, and so b_m tells an activist nothing about either θ or the likely actions of others. Using (5), this implies that an equilibrium strategy takes the form $A(\tilde{s}_m, b_m) = \hat{A}(\tilde{s}_m) + \pi b_m$, where $\hat{A}(\tilde{s}_m) : \mathbb{R}^n \mapsto \mathbb{R}$ is the action taken by a party member with no personal agenda.

A natural class of strategies to examine is those of the form $\hat{A}(\tilde{s}_m) = \sum_{i=1}^n w_i \tilde{s}_{im}$ where $\sum_{i=1}^n w_i = 1$. This means that (absent the bias term) the policy advocated by an activist party member is a weighted average of the speeches he hears. Such a strategy is easily interpreted: each coefficient is a convenient measure of a leader's influence. Furthermore, linear strategies emerge naturally because the conditional expectations (or regressions) $\mathrm{E}[\theta \,|\, \tilde{s}_m]$ and $\mathrm{E}[\tilde{s}_{m'} \,|\, \tilde{s}_m]$ are linear, a conclusion which follows from our normality assumptions. This means that if other activists use a linear strategy, then a best reply is also linear; this ensures the existence of a (unique) equilibrium in which linear strategies are used and which the equality $\sum_{i=1}^n w_i = 1$ is satisfied. Finally, a linear equilibrium is unique within

a broader class of advocacy strategies, and recent work in economics suggests that the equilibria of games similar to ours always involve linearity.[8]

An important element of our first main result, as well as those that follow, is the relative clarity of a leader. The message of leader i is clearer than leader j if $(x_i/\sigma_i^2) > (x_j/\sigma_j^2)$, and she has better judgement if $(1/\kappa_i^2) > (1/\kappa_j^2)$. Of additional interest to us, however, is a leader's comparative advantage in clarity; that is, clarity relative to judgement.

Definition. *Leader i is comparatively clearer than leader j if and only if $x_i\kappa_i^2/\sigma_i^2 > x_j\kappa_j^2/\sigma_j^2$.*

Comparative clarity is linked to the (conditional) correlation of the messages heard by different activists. Conditional on θ, the correlation coefficient ρ_i between $\tilde{s}_{im}$ and $\tilde{s}_{im'}$ is $\rho_i = \kappa_i^2/(\kappa_i^2 + (\sigma_i^2/x_i))$; hence leader i is comparatively clearer than j if and only if $\rho_i > \rho_j$.

With our notation, definitions, and model apparatus in place we are ready to describe the unique linear Bayesian Nash equilibrium and so characterize the influence of leaders.

Proposition 1. *There is a unique equilibrium involving the use of linear strategies:*

$$A(\tilde{s}_m, b_m) = \pi b_m + \sum_{i=1}^{n} w_i^* \tilde{s}_{im} \quad where \quad w_i^* \propto \frac{1}{\pi \kappa_i^2 + (\sigma_i^2/x_i)}. \tag{6}$$

A leader's influence increase with the quality of her judgement, with the clarity of her communication, and with the attention she receives; it is independent of the heterogeneity β^2 of activists' biases. The relative influence of comparatively clearer communicators increases as the coordination motive grows: w_i^*/w_j^* is decreasing in π if and only if leader i is comparatively clearer than j.

While an activist responds to her personal policy bias, the influences of the various leaders do not depend on the heterogeneity (measured by β^2) of idiosyncratic opinions. This is because the leaders' speeches reveal nothing about the private elements of preferences; the party's ideal policy and any signals of it are orthogonal to the private preference components. This implies that fundamental disagreements do not always influence the role of leaders nor (as we shall see) any strategic decisions taken by them.

Unsurprisingly, a leader with better skills enjoys more influence. However, there is an emphasis (which increases with the concern for unity) on better communicators. This is welcome since (Lemma 1) there unity is under-supplied: a clear message unites extreme

---

[8]Consider, for instance, strategies which do not diverge away from linearity: an advocacy strategy $\hat{A}(\cdot)$ where for some set of weights $A(\tilde{s_m}) - \sum_{i=1}^{n} w_i \tilde{s}_{im}$ remains bounded for any $\tilde{s}_m$. Within this class, the only equilibrium is linear: the argument of Dewan and Myatt (2008) extends straightforwardly. In other recent work, Hellwig and Veldkamp (2009) noted that the arguments of Angeletos and Pavan (2008) may be deployed to verify uniqueness in games of this kind, while Calvó-Armengol, de Martí Beltran, and Prat (2009) demonstrated uniqueness when players' signals are drawn from a truncated normal distribution. When the player set is finite and a quadratic-payoff game admits an exact potential function (as it does here) then Theorem 4 of Radner (1962) can be used to establish uniqueness.

opinions by serving as a focal point around which the party base can unite. Neverthe-less, an activist's failure to incorporate the external benefits of a move toward the center ensures that the equilibrium weights are inefficient.

Proposition 2. *The most efficient linear strategy is for activists to use weights $w_i^\dagger$ satisfying*

$$w_i^\dagger \propto \frac{1}{\pi^\dagger \kappa_i^2 + (\sigma_i^2/x_i)} \quad where \quad \pi^\dagger = \frac{\pi}{2 - \pi}. \tag{7}$$

This satisfies $\pi^\dagger < \pi$, confirming that activists place too little emphasis on party unity. In equi-librium activists place too little weight on the speeches of comparatively clear communicators. The welfare-maximizing weights are independent of the heterogeneity β^2 of activists' biases.

We conclude this section by contrasting our model with that of Dewan and Myatt (2008). They used a similar information structure but specified followers' preferences very dif-ferently. Firstly, they imposed no idiosyncratic policy biases; this is equivalent to setting $\beta^2 = 0$. Secondly, their activists care about conformity (via $\mathrm{E}[(a_m - \bar{a})^2]$) rather than party unity (via $\mathrm{E}[(a_m - a_{m'})^2]$). Despite these important differences, the equilibrium weights placed on leaders' speeches are unchanged. There are two reasons for this. Firstly, β^2 has no effect because (as we have observed) leaders' speeches are orthogonal to personal biases.[9] Secondly, and following the decomposition of the party disunity term in (2), an activist's decision only changes the first non-conformity component. Nevertheless, the second component (the non-conformity of others) remains a component of party welfare. This is important because here the inefficiency of the equilibrium creates a role for strate-gic actions by benevolent leaders who may seek to enhance party unity.

<div align="center">STRATEGIC LEADERS</div>

Whilst leaders inform their followers actions, this does not fully resolve the implicit collective-action problem. In our policy-advocacy world, activists do not internalize the full benefits of unity (Lemma 1) and so weight leaders' views inefficiently (Proposition 2). What might make leadership more effective in achieving activists' goals? To provide one answer, we allow leaders' skills to be endogenous and leaders to take measures—rhetorical strategies—to alleviate the problems that they perceive.

Before analyzing rhetorical strategies, we specify leaders' motives. Each leader shares similar objectives to the activists: she would like them to advocate policies which best suit the party's needs and she prizes party unity. We allow leaders to weight these objectives differently from activists, although the case where they weight them similarly (benevolent

---

[9]An interaction between idiosyncratic biases and leadership may arise if the distribution of policy biases is uncertain and a leader is able to sample the views of the party base before delivering her speech.

leaders who care about welfare) is a prominent case of interest for us. More formally, $\tilde{\pi}_i$ is the weight placed by leader i on her concern for policy, so that $1 - \tilde{\pi}_i$ is the weight placed on unity. The payoff of leader i may then be specified as

$$v_i = \bar{v} - \tilde{\pi}_i \, \mathrm{E}[(a_m - \theta - b_m)^2] - (1 - \tilde{\pi}_i) \, \mathrm{E}[(a_m - a_{m'})^2], \qquad (8)$$

where the indices m and m' represented generic activists, and expectations are taken across the party's membership.[10] A benevolent leader (who cares about welfare) is easily obtained by setting $\tilde{\pi}_i = \pi$. Later in the paper we allow leaders' objectives to differ, so that $\tilde{\pi}_i$ varies with i. For now, however, we simplify our exposition by supposing that all n leaders share the same objective, so that $\tilde{\pi}_i = \tilde{\pi}$ and $v_i = v$ for all i.

Examining a leader's payoff we can break down the disunity term $\mathrm{E}[(a_m - a_{m'})^2]$ into two components, just as we did in (2). From the perspective of an activist only the first non-conformity component matters. However, for a leader both components matter. Indeed, it is straightforward to confirm that a leader's expected payoff satisfies

$$v = \bar{v} - \tilde{\pi} \, \mathrm{E}[(a_m - \theta - b_m)^2] - 2(1 - \tilde{\pi}) \, \mathrm{E}[(a_m - \bar{a})^2], \qquad (9)$$

which (by inspection) places greater relative emphasis on the non-conformity loss function. Incorporating the behavior of activists, it is then straightforward to calculate a leader's expected payoff. This takes a similar form to an activist's expected payoff, with the proviso that the influence of the errors-of-judgement variance is downplayed:

$$v \propto \text{constant} - \sum_{i=1}^{n} w_i^2 \left[\pi^{\ddagger} \kappa_i^2 + \frac{\sigma_i^2}{x_i} \right] \quad \text{where} \quad \pi^{\ddagger} \equiv \frac{\tilde{\pi}}{2 - \tilde{\pi}}. \qquad (10)$$

The importance of the errors-of-judgement variance terms κ_j^2 is determined by $\pi^{\ddagger}$. This contrasts with the coefficient π which is used in the equilibrium strategy. Put another way, whereas π drives followers' decisions, $\pi^{\ddagger}$ is the parameter that a leader would like them to use. Only when $\pi^{\ddagger} = \pi$ is a leader content with their behavior.

Definition. *A leader is a* unifying leader *if $\pi^{\ddagger} < \pi$ and she is a* policy seeking leader *if $\pi^{\ddagger} > \pi$.*

Following logic similar to that of Proposition 2, leaders would like followers to use weights

$$w_i^{\ddagger} \propto \frac{1}{\pi^{\ddagger} \kappa_i^2 + (\sigma_i^2 / x_i)}. \qquad (11)$$

The difference between a leader's desired weights and the equilibrium weights placed depends on a comparison of $\pi^{\ddagger}$ and π, and is recorded in the next simple result.

---

[10]Somewhat more formally, $\mathrm{E}[(a_m - \theta - b_m)^2] \equiv \int_0^1 (a_m - \theta - b_m)^2 \, dm$ and similarly for $\mathrm{E}[(a_m - a_{m'})^2]$.

Proposition 3. *A unifying leader would like followers to place more weight on comparatively clearer communicators (if $w_i^\ddagger - w_i^* > 0 > w_j^\ddagger - w_j^*$ then i must be comparatively clearer than j). A policy-seeker would like followers to place less weight on comparatively clearer communicators.*

A special case is benevolent leadership. According to our terminology, a benevolent leader is automatically a unifying leader and so (as noted in Proposition 2) she would like activists to respond more strongly to comparatively clearer leaders. Thus even when the interests of leaders and their followers coincide, they disagree about how activists' actions should react to their information.

Whenever $\pi^\ddagger \neq \pi$ leaders would like to change the behavior of their followers. This raises the possibility that a leader may deploy a rhetorical strategy in the pursuit of this aim. One option open to a leader is to reduce the clarity of her speech. This is obfuscation, and corresponds to an increase in σ_i^2. Similarly, she may limit the length of time for which she speaks. In doing so she (at least partially) stands back from the pulpit, a move which reduces x_i. Both of these strategies succeed in harming the clarity of her message, although as we shall see they can also raise the clarity of other leaders' messages. The direct effect of an increase in σ_i^2/x_i is to harm the expected payoff of a leader. However, an indirect effect is to influence the equilibrium weights placed on the various leaders' speeches, and this strategic effect can succeed in raising a leader's payoff. We proceed, therefore, by studying optimal rhetorical strategies.

SPEAKING CLEARLY

Other things equal, one might think that a leader who gives clear expression to her views will be more successful than one who does not. Yet, a common perception remains that politicians do not always transparently express their views. Moreover, our literature review (which concludes our paper) shows that both theory and evidence support the view that leaders strategically choose how clearly to convey ideas.

A novel logic to explain obfuscatory rhetoric was given by Dewan and Myatt (2008). In their model followers choose to whom to listen (we also allow for endogenous attention later in this paper) and egocentric leaders seek attention. An attention-seeking leader sometimes obfuscates: if she speaks less clearly then her audience must listen for longer in order to digest her message. A problem with their theory is that it relies on a rather extreme leadership motive. Here we explore whether obfuscation is part of a leader's rhetorical armor when her objectives are directly related to followers' actions.

We begin by supposing that all leaders speak as clearly as they can, and ask whether leader i would like to reduce her clarity by increasing σ_i^2 or, equivalently, by reducing x_i.

The direct effect of this is harmful; however, the indirect strategic effect is to erode the influence of leader i (a reduction in w_i^*) and enhance that of other leaders (proportional increases in w_j^* for $j \neq i$). A necessary condition for obfuscation is that the strategic effect is positive, which in turn implies that we must find a leader who wishes to lose influence. Proposition 3 has already given us some insights into this: for instance, unifying leaders would prefer to see influence shift toward comparatively clearer communicators. To make this insight more precise it is useful to define rather more formally an appropriate measure of a leader's comparative clarity.

Definition. *Fixing the activists' policy-concern parameter π, a leader's comparative clarity is*

$$\gamma_i \equiv \frac{\rho_i}{1 - (1 - \pi)\rho_i} \quad where \quad \rho_i = \frac{\kappa_i^2}{\kappa_i^2 + (\sigma_i^2/x_i)}. \tag{12}$$

According to this definition, a leader's comparative clarity γ_i is an increasing function of the correlation coefficient ρ_i, which in turn depends on the clarity of a leader's message relative to the quality of her judgement. The measure γ_i varies from $\gamma_i = 0$ (when $\kappa_i^2 = 0$) up to $\gamma_i = 1/\pi$ (when $\sigma_i^2 = 0$). Defining comparative clarity in this particular way is useful because it allows us to state a simple condition which determines when it is optimal for a leader to reduce her clarity by increasing locally σ_i^2.

Lemma 2. *Leader i finds it optimal to obfuscate locally, so that $\partial v/\partial \sigma_i^2 > 0$, if and only if*

$$(\pi - \pi^\ddagger)(\bar{\gamma} - \gamma_i) > \frac{1}{2} \quad where \quad \bar{\gamma} \equiv \sum_{i=1}^{n} w_i^* \gamma_i. \tag{13}$$

A leader's payoff v is quasi-concave in σ_i^2 if $\pi^\ddagger > \pi$, and quasi-convex in σ_i^2 if $\pi^\ddagger < \pi$.

The right-hand side of the "obfuscation inequality" in (13) reflects the direct cost of obfuscation, whereas the left-hand side captures the (sometimes beneficial) strategic effect.

The strategic effect is helpful whenever $\pi - \pi^\ddagger$ and $\bar{\gamma} - \gamma_i$ share the same sign.[11] There needs to be enough distance between the desired policy-unity balance of leaders and followers ($|\pi - \pi^\ddagger|$ needs to be large) for obfuscation to be optimal; if not, then the direct effect dominates. Notice also that there needs to be a sufficiently large gap between the comparatively clarity of a leader and the leadership-wide average. The size of the gap $\gamma_i - \bar{\gamma}$ is, in turn, bounded above by $1/\pi$, and so the obfuscation inequality can hold only for certain values of π. With Lemma 2 in hand, we can evaluate the incentive to obfuscate for leaders with different skills, and also assess the extent of that obfuscation.

---

[11]There is a strategic benefit to obfuscation when a comparatively poor communicator ($\gamma_i < \bar{\gamma}$) wishes to promote unity ($\pi^\ddagger < \pi$) or when a comparatively good speaker ($\gamma_i > \bar{\gamma}$) wishes to see better policy ($\pi^\ddagger > \pi$).
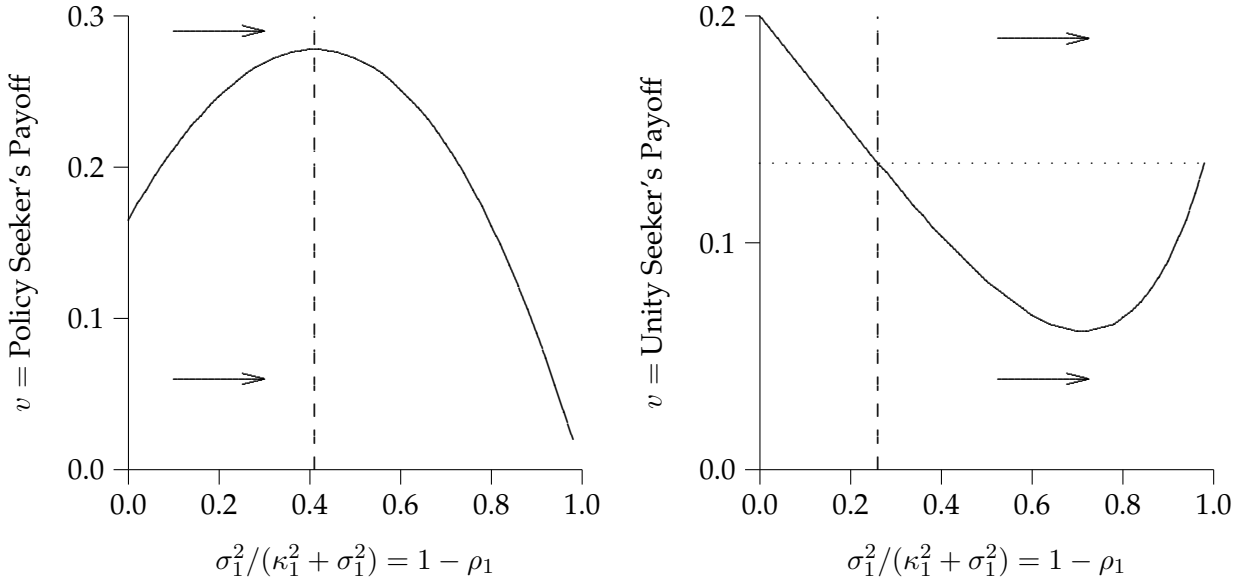
Notes: These figures illustrate the use of obfuscatory rhetoric by plotting the payoff v of a leader $i = 1$ against a measure of her obfuscation.

Obfuscation is an increase in σ_1^2, which is shown relative to $\kappa_1^2 + \sigma_1^2$. Notice:

$$\frac{\kappa_1^2}{\kappa_1^2 + \sigma_1^2} = \frac{1/(\kappa_1^2 + \sigma_1^2)}{1/\kappa_1^2} = \frac{\text{Informativeness of the Leader's Speech}}{\text{Quality of the Leader's Judgement}}.$$

This is a measure of the information received by followers relative to the information available to the leader. This means that

$$\frac{\sigma_1^2}{\kappa_1^2 + \sigma_1^2} = 1 - \frac{\text{Informativeness of the Leader's Speech}}{\text{Quality of the Leader's Judgement}}$$

is the proportion of information lost due to errors in communication. This proportion increases with the extent of a leader's obfuscation.

Both figures use these parameter choices: $n = 2$ leaders; $\beta^2 = 0$ so that fundamental differences are absent; $x_1 = x_2 = 1$ without loss of generality.

The left-hand figure shows the payoff of a policy seeker. It uses these parameters: $\pi = 0.1$ so that activists care mainly about unity; $\pi^{\ddagger} = 1$ so that leaders care only about policy; $\bar{v} = 1$; $\sigma_2^2 = 1$ and $\kappa_2^2 = 0$, so that leader 2 has perfect judgement but imperfect clarity; and $\kappa_1^2 = 1$ so that the two leaders are equally informative so long as leader 1 speaks with perfect clarity. Complete clarity ($\sigma_1^2 = 0$) is not optimal for leader 1: if her natural clarity is to the left of the broken line then she faces an incentive to obfuscate.

The right-hand figure shows the payoff of a unifying leader. It uses these parameters: $\pi = 1$ so that activists care only about policy; $\pi^{\ddagger} \approx 0.05$ so that leaders care mainly about unity; $\bar{v} = 0.25$; $\sigma_2^2 = 0$ and $\kappa_2^2 = 1$, so that leader 2 has perfect clarify but imperfect judgement; and $\kappa_1^2 = 1$ so that the two leaders are equally informative so long as leader 1 speaks with perfect clarity. By inspection, either maximum clarity or maximum obfuscation can be optimal for leader 1: if her natural clarity falls to the right of the broken line then it is optimal for obfuscate maximally.

FIGURE 1. Obfuscation by Policy-Seeking and Unifying Leaders

Proposition 4. *(i) A policy-seeking leader obfuscates only if she is a comparatively clear communicator. A necessary condition for obfuscation is that followers' policy concern is sufficiently weak: if $\pi > \frac{2}{3}$ then maximum clarity is optimal. Although partial obfuscation can be desirable, it is never optimal for a policy seeker to obfuscate completely by choosing $\sigma_i^2 \to \infty$.*

(ii) A unifying leader obfuscates only if she is a comparatively poor communicator. If she obfuscates then she does so completely by choosing $\sigma_i^2 \to \infty$ and so babbles incoherently. A necessary condition for obfuscation is that $\pi^{\ddagger}$ is sufficiently small relative to π.

(iii) For a benevolent leader the gap $\pi - \pi^{\ddagger}(= \pi - \pi^{\dagger})$ is never large enough for the obfuscation inequality (from (13) of Lemma 2) to hold, and so she never obfuscates.

A policy seeker ($\pi^{\ddagger} > \pi$) obfuscates if she thinks that activists are insufficiently concerned with policy. When this is so, she wishes to enhance the influence of leaders with comparatively good judgement. A problem emerges: wisdom is no guarantee of eloquence. A technocrat with an excellent grasp of policy but with poor communication skills may lose out to a clearer communicator whose words have less substance. When a policy-seeking leader is comparatively clear ($\gamma_i > \bar{\gamma}$) obfuscation can help to redress the perceived imbalance. Nevertheless, the extent of her obfuscation is limited and so she always contributes something to the debate. The intuition is straightforward: she obfuscates only if she enjoys too much influence as a comparatively clear communicator; but if she obfuscates enough then she is no longer comparatively clear and so the strategic incentive falls away. The left-hand panel of Figure 1 illustrates this effect.

By contrast a unifying leader ($\pi^{\ddagger} < \pi$) obfuscates only if her clarity is relatively poor ($\gamma_i < \bar{\gamma}$). If the latter effect is large enough, then the strategic incentive ensures that the obfuscation inequality (13) holds. However, if the strategic incentive dominates then it must also do so as the communication error increases endogenously. In lowering her absolute clarity (raising σ_i^2) a leader also lowers her relative clarity (a fall in γ_i) and so ensures that the obfuscation inequality is more easily satisfied. Following this self-reinforcement logic, a unifying leader who obfuscates does so completely by maximizing the noise in her communication; this contrasts with her policy-seeking counterpart. The right-hand panel of Figure 1 illustrates the "all or nothing" rhetorical strategy.

Proposition 4 provides a further insight: the size of the leadership elite is influenced by rhetorical strategies. If a unity-seeking leader obfuscates then her all-or-nothing strategy effectively means that she removes herself from the elite. She does so because she fears that a diversity of opinion may frustrate coordination. In obfuscating all the way (babbling incoherently, or avoiding all reference to policy issues) she resigns any de facto authority. As illustrated by the right-hand panel of Figure 1, this can generate a de facto

dictatorship. However, this does not happen when leaders are policy seekers: they would like activists to weight the opinions of a wider group more evenly, and so they obfuscate only so that activists form a balanced view of the merits of different policies.

Here we have demonstrated (Figure 1) conditions under which obfuscation may arise. We note, however, that the incentives of a benevolent leader are too closely aligned with those of activists for obfuscation ever to be desirable. This suggests that such a leader does not engage in rhetorical manipulation to achieve her goals. However, as we shall see, this is not the case once we allow for endogenous information acquisition.

<div align="center">ENDOGENOUS ATTENTION</div>

So far activists have chosen their actions based on the information available to them but have had no control over the nature of that information. Here we extend our model to include endogenous information acquisition: activists choose to whom to listen.

Recall that the noise in communication is determined by the variance σ_i^2/x_{im} where x_{im} is the length of time which activist m devotes to listening to leader i. Time is limited, and so we impose the constraint $\sum_{i=1}^n x_{im} \leq t$. Thus, a follower can improve his understanding of one leader only at the expense of clouding his understanding of another.

An activist's choice of attention is a move in a game played with other activists. The coordination motive ensures that he would like to do what others do, and so naturally he would like to know what others know.[12] Thus we study a simultaneous-move game in which activist m chooses both the attention paid to each leader and the weight placed on the leader's speech in his policy-advocacy decision, subject to the budget constraint on his time. He does this to maximize his expected payoff prior to the realization of the signals obtained from listening to the leaders' speeches, and once those signals are received he implements the action $a_m = \pi b_m + \sum_{i=1}^n w_{im}\tilde{s}_{im}$.[13] It is straightforward to confirm that this action is optimal ex post given that the weights w_{im} are chosen optimally ex ante.

Happily, our listening-then-following game has a unique symmetric equilibrium. Fixing the equilibrium attention x_i^* paid to each of the leaders, the equilibrium weights are those described in Proposition 1, with the proviso that $w_i^* = 0$ if and only if $x_i^* = 0$; a leader has influence if and only if she receives attention. Of more interest here are the determinants

---

[12]This feature of information acquisition was elegantly emphasized by Hellwig and Veldkamp (2009).

[13]Somewhat more formally, we consider a simultaneous move game in which each player's choice is a vector in $\mathbb{R}^{2n}$. This vector breaks down into a first component $x_m \in \mathbb{R}_+^n$ satisfying $\sum_{i=1}^n x_{im} \leq t$, and a second component $w_m \in \mathbb{R}_+^n$ satisfying $\sum_{i=1}^n w_{im} = 1$. Note that a player's expected payoff is not defined when $x_{im} = 0$ and $w_{im} > 0$. For such cases, payoffs are obtained by taking the limit as $x_{im} \to 0$.

of that attention. Taking the weights placed on the various leaders' speeches as given, the equilibrium attention levels minimize

$$\sum_{i=1}^{n} (w_i^*)^2 \left[\pi \kappa_i^2 + \frac{\sigma_i^2}{x_i} \right]. \tag{14}$$

It is easy to confirm that x_i^* is proportional to $w_i^* \sigma_i$, and so the attention devoted to a leader increases with the weight placed on her speech: if a leader is influential, then it is important for her followers to listen carefully to her. Fixing leaders' influences, the attention paid to a leader falls with her clarity: if a leader is a poor communicator then her followers may compensate by paying greater heed to her words.[14] This does not automatically imply that a leader attracts attention by obfuscating, since obfuscation also reduces a leader's influence and so indirectly limits the attention she attracts.

Since x_i^* is proportional to w_i^* any leader who has influence attracts attention; similarly, any leader who attracts attention also enjoys influence. Does this imply that all members of the leadership elite enjoy both influence and attention? Not necessarily. A loss of attention can spark a loss of influence, which drives away a leader's audience and so causes a further loss of influence; such a downward spiral can converge to a situation in which a leader is ignored. Thus we ask: which leaders receive attention, and which, if any, are ignored? How does the amount of attention a leader receives depend on her natural abilities? Our next formal result answers these questions.

Proposition 5. *There is a unique n^* such that activists pay attention only to an elite comprising the n^* clearest leaders, and ignore the others. Within the elite, there is a constant such that*

$$x_i^* = \frac{\sigma_i(\text{constant} - \sigma_i)}{\pi \kappa_i^*}. \tag{15}$$

The size n^* of the leadership elite falls as the concern for unity grows. Amongst the elite, the absolutely clearer communicators are also comparatively clearer. The attention paid to each leader and the size of the elite grow with the time available to activists.

This proposition is closely related to a result of Dewan and Myatt (2008). However, there are important differences. Firstly, the objectives of our activists differ substantially; they prize unity rather than conformity and they entertain personal policy biases. Secondly, here we allow for a general time constraint t and offer the new result that, perhaps unsurprisingly, activists listen to more leaders when they have more time available to do so. Thirdly, just below we will extend our analysis to a world in which activists are able to choose endogenously the total time spent listening to leaders.

---

[14]This last effect was used by Dewan and Myatt (2008) to explain obfuscation by attention-seeking leaders: increasing σ_i directly increases the attention paid to a leader. However, in assessing obfuscation as an attention-grabbing technique it is also important to note that w_i^* is decreasing in σ_i^2.

Interestingly, whereas all leaders have influence when their views are heard this is no longer the case when audiences are endogenous. A follower-generated dictatorship can emerge in which activists give prominence to the views of one leader, whilst ignoring those of another; simple calculations confirm that this happens if $\sigma_j \geq \sigma_i + (\pi\kappa_i^2/\sigma_i)$ for some i and all $j \neq i$, and in this case $x_i^* = 1$. So long as the clearest leader is a strictly better communicator than everyone else, this inequality is satisfied if π is sufficiently small: when activists wish to focus almost entirely on the need for unity then a de facto dictatorship arises, and potentially damaging differences of opinion are minimized.

A further observation is that clarity, rather than judgement, is critical in determining a leader's membership of the elite. However, dictatorship cannot emerge when a leader is a perfect communicator. This is because, as in Dewan and Myatt (2008) a very clear leader is understood quickly; this frees up time for her audience to listen to others.

Proposition 5 imposes a budget constraint upon activists: they must divide a fixed period of time between different leaders. This is responsible for the exclusion of some leaders from positions of influence; if instead t grows unboundedly large then eventually all leaders will be heard. This leads us to ask: what if t is chosen endogenously?

To answer this question, we briefly consider a world in which each activist m can choose the total time spent listening and in doing so incurs a cost of $C(\sum_{i=1}^n x_{im})$, where $C(\cdot)$ satisfies familiar decreasing-returns properties.[15] It is straightforward to show that an activist's marginal benefit from increasing the time spent listening to leaders is decreasing, and so there is unique symmetric choice of total attention.

Proposition 6. *With a cost-of-attention function $C(\cdot)$, the total time spent listening to leaders and the number of leaders who attract attention fall as the activists' concern for unity rises. From the perspective of a benevolent social planner, activists spend too little time listening to leaders.*

The inefficiency described in the final claim arises because an activist fails to recognize that others benefit from his conformity. Listening for longer helps conformity, and so exerts a (positive) externality on others which is not incorporated into private decisions.

<div align="center">STANDING BACK</div>

We have characterized the equilibrium behavior of activists when they choose to whom to listen. We now ask whether their attention choices are efficient, and relatedly ask whether goal-oriented leaders would wish to distort those choices.

---

[15]We assume that $C(\cdot)$ is strictly increasing, convex, and continuously differentiable, and $C'(0) = 0$.

Fixing the total time spent listening to leaders and the weights placed on their speeches, the attention levels are efficiently chosen. However, the weights placed on speeches are inefficient; Proposition 2 confirms that activists place too little weight on comparatively clear speeches. A benevolent planner would like influence to shift toward clearer communicators, and one way of achieving this is to distort attention toward them; after all, increased attention generates increased influence. This feature of the equilibrium (too little attention is paid to better communicators once the inefficiency of the equilibrium influence of leaders is taken into account) has important and interesting implications. It remains possible that a dictatorship, whilst desirable from a welfare perspective, is not attainable when activists alone determine the attention given to different leaders.

To be more specific, consider a two-leader world ($n = 2$) in which the first leader is the clearest communicator and without loss of generality we set $t = 1$. If the second leader receives any attention then a benevolent social planner would prefer the activists to devote attention $x_1^\dagger > x_1^*$ to the first leader, and consequently $x_2^\dagger < x_2^*$ to the second leader; this scenario is considered in Figure 2. Notice that for the parameter configuration illustrated it is socially optimal to set $x_1^\dagger = 1$; that is, the dictatorship of the first leader maximizes expected welfare. Nevertheless, activists listen to both leaders.

Invoking the idea of a social planner who would enforce efficient attention levels if she could, we consider political institutions that could enhance welfare. Suppose that discussion time were allocated to leaders by a central authority. We might think of this authority as represented by a moderator of the leadership debate. This institutional innovation ensures that leaders can only speak for as long as the moderator grants them the floor. The discussion so far generates the following proposition.

Proposition 7. *In a two-leader world a benevolent moderator of debate allocates more time, relative to the equilibrium attention levels chosen by activists, to the clearer communicator. Sometimes a benevolent leader prefers to stop the poorer communicator from speaking at all.*

Can welfare enhancing effects arise endogenously, for example, via the adoption of rhetorical strategies by leaders? To explore this possibility we continue with a world in which activists allocate a single unit of time between two leaders; setting $t = 1$ is without loss of generality for what follows. The two leaders can, however, choose to limit the amount of time that they air their views. Specifically, each leader i speaks for a length of time $\bar{x}_i$, forcing activists to choose an attention level below this upper bound. Of course, if in equilibrium $\bar{x}_i > x_i^*$ for each leader then the leaders' moves have no effect. On the other hand, if $\bar{x}_i < x_i^*$ then leader i stands back from the leadership debate; by doing so, she diverts attention to the other leader.
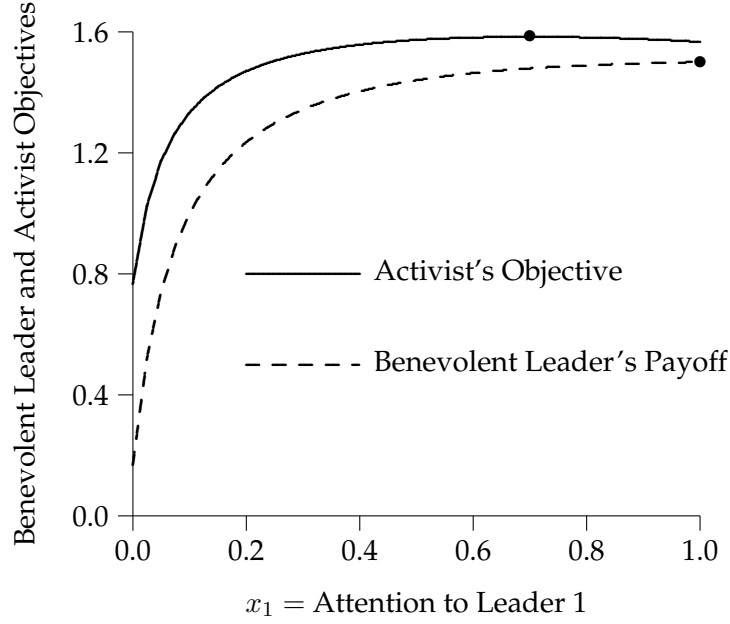
Notes: This figure illustrates the contrast between the objectives of an activist and a benevolent social planner. It uses the following parameter configuration: $\pi = 0.333$, $\kappa_1^2 = 1$, $\kappa_2^2 = 1$, $\sigma_1^2 = 0.1$, and $\sigma_2^2 = 0.9$. Hence Leader 1 is a relatively clear communicator. In equilibrium activists devote a fraction $x_1^* \approx 0.7$ of their time to listening to the first leader, and the remainder to the second leader. A benevolent planner would prefer them to focus exclusively on listening to the first leader.

FIGURE 2. Equilibrium vs. Efficient Attention

For now the leaders share the same preferences; they agree on the ideal division of attention between their speeches. If they wish activists to devote more time to the first leader (for instance) then this can be achieved if the second leader stands back from the debate. Our formal result reveals that (generically) one leader always wishes to stand back.

Proposition 8. *In a two-leader world, if both leaders receive attention from activists then one of them wishes to stand back. Leader i stands back if and only if $(\pi - \pi^\ddagger)(\sigma_i - \sigma_j) > 0$. Thus a unifying leader stands back if and only if she is the poorer communicator; a policy seeker stands back if and only if she is the clearer communicator. The only relevant leadership characteristic in determining whether a leader stands back is her clarity of communication.*

When the objectives of activists and leaders diverge then there is always a leader who desires less attention than activists bestow upon her. This contrasts with our analysis of obfuscation, when the strategic effect was often insufficient to offset the (harmful) direct effect. The reason that standing back is always optimal for someone is that, starting from the equilibrium attention levels, the direct effect of the first step back from the lectern is

negligible. By standing back a leader harms the ability of her followers to learn from her, but they are free to use the time saved to listen to someone else.[16]

When $\pi^\ddagger > \pi$ a leader feels that activists pay too much attention to the clearer communicator and so she would like them to give greater consideration to the views of the leader who is less eloquent. So, when she is the better communicator she receives more attention than she desires. By constraining the length of time that she holds the floor, a policy-seeking leader can divert attention to others by standing back. Similar logic (related to that given in our analysis of obfuscation) applies to the $\pi^\ddagger < \pi$ case.

Somewhat surprisingly, the only characteristic relevant to the decision to stand back is the (absolute) clarity of a leader's communication; a leader's quality of judgement does not enter into the statements of Proposition 8. The reason is that what really matters is a leader's comparative clarity, and once attention is endogenously chosen (following Proposition 5) the leader with the better absolute clarity also has the better comparative clarity; any difference in the leaders' qualities of judgements is already accounted for in the activists' equilibrium choice of attention.

We conclude this section by recalling that a benevolent leader is automatically a unifying leader; she recognizes (from a welfare perspective) that activists pay too little attention to the clearer leader who can act as a unifying focal point. Thus, a simple corollary to Proposition 8 reinforces the logic of Proposition 7.

Corollary to Proposition 8: *A benevolent leader stands back in favor of a clearer communicator.*

STEPPING DOWN

We have seen (via Proposition 5) that a follower-generated dictatorship can emerge in which activists listen to one leader whilst ignoring others. In other cases, activists may wish to listen to more than one leader but (as Figure 2 illustrates) a dictatorship can emerge when a (sometimes benevolent) leader stands back in order to deflect attention away from herself. Thus the size as well as character of the leadership elite may be determined by the rhetorical strategies that leaders deploy.

This prompts us to examine how the size of the leadership oligarchy is determined by the rhetorical strategies of leaders under a wide range of leadership motives. In essence we are investigating the most fundamental choice a leader faces: a leader can have influence only if she steps up and makes her views known; alternatively she may step down

---

[16]Formally, moving x_1 away from the equilibrium level x_1^* has only a second-order direct effect on activists; however, it has a first-order effect in helping to rebalance the influence of the leaders.

and thereby deflect all attention to others. Within the context of the two-leader world, a decision to step down is a special case of standing back.

When can the rhetorical strategy of standing back change the size of a leadership elite? Consider again two leaders who wish to encourage unity. The clearer leader always enjoys attention in equilibrium, and (following Proposition 8) only the poorer communicator ever stands back. In this case, standing back can result in move from oligarchy to a leader-prompted dictatorship. Next consider a world in which the two leaders wish to promote policy. In this world, it is the clearer communicator who stands back; doing so creates room for the poorer communicator, which can serve to expand the size of the leadership elite. Thus whether the rhetorical strategies of leaders result in an expansion or contraction of the elite turns on whether those leaders are unifiers or policy seekers.

Proposition 9. *Consider a world with two leaders, and without loss of generality label them so that $\sigma_1^2 \leq \sigma_2^2$. Suppose that the leaders wish to promote unity, so that $\pi^{\ddagger} < \pi$. If*

$$\left[1 + \frac{\pi \kappa_1^2}{\sigma_1^2}\right] \left[1 + \frac{(2\pi^{\ddagger} - \pi)\kappa_1^2}{\sigma_1^2}\right] < \frac{\sigma_2^2}{\sigma_1^2} < \left[1 + \frac{\pi \kappa_1^2}{\sigma_1^2}\right]^2 \tag{16}$$

then activists wish to pay attention to both leaders, but nevertheless the second leader steps down in favor of the first. The stepping down condition (the left inequality) is most easily satisfied when the second leader is a relatively poor communicator and the first leader has good judgement.

Suppose instead that leaders wish to promote policy, so that $\pi^{\ddagger} > \pi$. If

$$\left[1 + \frac{\pi \kappa_1^2}{\sigma_1^2}\right] \left[1 + \frac{(2\pi^{\ddagger} - \pi)\kappa_1^2}{\sigma_1^2}\right] > \frac{\sigma_2^2}{\sigma_1^2} > \left[1 + \frac{\pi \kappa_1^2}{\sigma_1^2}\right]^2 \tag{17}$$

then activists prefer to devote all of their attention to first leader. Nevertheless, this first leader stands back in order to create room for the views of the second leader.

Proposition 9 brings into full focus how the size and character of the leadership elite is shaped by rhetorical strategies. When leaders emphasize unity, then an oligarchy becomes a dictatorship. This can occur because a leader steps down. However, recalling an earlier result (Proposition 4) it can also occur because a leader with an extreme preference for unity, and who is herself a relatively poor communicator, obfuscates completely, thereby make herself irrelevant to the leadership debate. Conversely, when leaders are policy seekers and activists veer toward a relatively clear would-be dictator, leaders stand back so that the voices of a broader oligarchy are heard. Under a wider oligarchic arrangement the relative influence of leaders, and the attention they receive, is determined by the response of followers to leaders' innate characteristics as well as the rhetorical strategies adopted; policy-seeking leaders stand back and may obfuscate (although not completely) to ensure that a wide range of opinions is heard.
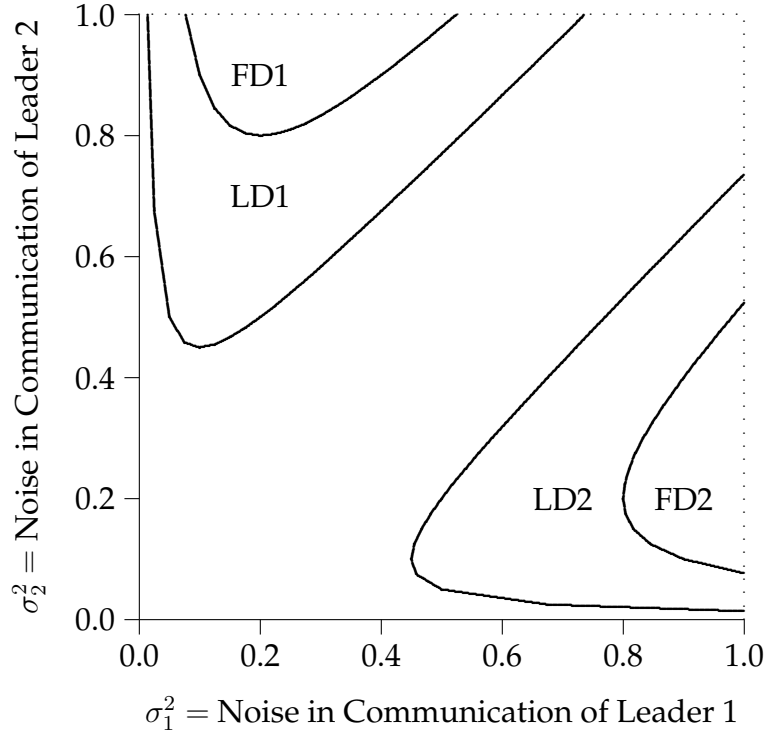
FIGURE 3. Endogenous Leadership Regimes

Notes: The figure illustrates five regions with different leadership elites. Two leaders have equal qualities of judgement, so that $\kappa_1^2 = \kappa_2^2 = 1$. Other parameter choices are $\pi = 0.2$ and $\pi^{\ddagger} = 0.135$, so that the two leaders are unifiers. In regions FD1 and FD2 there is follower-generated dictatorship: only one leader receives attention in these regions. In regions LD1 and LD2 there is a leader-generated dictatorship: activists would like to listen both leaders, but the poorer communicator steps down. In the remaining region, both leaders attract attention and both speak, although the poorer communicator limits the length of her speech by standing back.

A DIVIDED OLIGARCHY

By standing back or stepping down, leaders can determine the allocation of influence amongst the elite. So far we have assumed that there is no disagreement amongst the elite's members. Here we consider briefly a two-leader world in which leaders disagree amongst themselves. Specifically, we suppose that the first leader is a better communicator than the second (so that $\sigma_1^2 < \sigma_2^2$). Unlike previous sections the leaders do differ in the relative weight they wish activists to place on policy versus conformity; formally, $\pi_1^{\ddagger} \neq \pi_2^{\ddagger}$.

One possibility is that both leaders are policy seekers or both leaders are unifiers. Following our earlier logic, only one leader wishes to stand back; if the inequality contained in Proposition 8 is satisfied for leader i then it cannot be satisfied for leader $j \neq i$. Since only one leader stands back, the other stands forth and so activists remain able to devote all of their time to listening to leaders. Put simply, Proposition 8 holds.

A second situation is when the leaders preferences oppose, but where their relative skills suit their preferences. This is a situation where the clearer leader is a unifier (and so does not wish to stand back) while the poorer communicator is a policy seeker (and so also faces no incentive to stand back). Since neither leader can achieve her objectives by limiting what she says, we obtain an equilibrium in which neither leader stands back and so activists devote their unconstrained equilibrium attention levels x_1^* and x_2^* to the pair.

The final situation is one in which leaders preferences are not well aligned with their skills: the clearer leader is a policy seeker, while the poorer communicator seeks greater unity. Both leaders wish to lose influence, and so both wish to stand back, so long as the space created by doing so is used by another orator.

Proposition 10. *Consider a world with two leaders labeled so that the first leader is a clearer communicator. If both leaders are policy seekers or both are unifiers then Proposition 8 holds. If the first leader is a unifier while the second is a policy seeker then neither stands back. However, if the first leader is a policy seeker while the second is a unifier then both leaders face an incentive to stand back. There can be multiple equilibria in which one leader stands back.*

The final claim raises the possibility of multiple equilibria whenever $\pi_1^{\ddagger} > \pi > \pi_2^{\ddagger}$. The reason for this is that both leaders wish to stand back. However, standing back differs from obfuscation precisely because it frees time for followers to listen to others; in contrast, obfuscation reduces the clarity of a message while creating no extra room for anyone else. If both leaders stand back then followers are unable to use the time freed by one leader to listen to another. This, of course, makes standing back equivalent to obfuscation, and so it becomes less attractive. So, if one leader stands back then the other leader is (often, but not always) unwilling to do so. (Specifically, this arises when a leader is willing to stand back but not willing to obfuscate.) It is straightforward to see that there can be a range of equilibria in which the two leaders divide the debating time between themselves, and each stand back a little further if only the other would speak up in her place.

RELATED LITERATURE AND CONCLUDING REMARKS

There is a large social-scientific literature which studies the broad notion of leadership; textbooks such as Northouse (2007) and Yukl (2008) offer surveys. Perhaps surprisingly, however, there are relatively few formal theoretical models. Indeed, Levi (2006, p. 11) observed that there is "a large literature on leadership" which "demonstrates that leadership clearly can make a difference" and yet "still lacking is a model of the origins and means of ensuring good leadership." Nevertheless, there has been recent growth in the

formal theoretical literature examining leadership and communication in different social situations. Our paper forms part of that growth.[17]

Within economics, recent theoretical contributions were surveyed by Bolton, Brunnermeier, and Veldkamp (2008a). Some of those focus on acts of leadership. Notably, Hermalin (1998, 2007) considered a situation in which a leader enjoys private knowledge about the nature of a task contemplated by her followers. The leader benefits from her followers' efforts and so wishes them to believe that the task is worthwhile. To communicate credibly she must "lead by example" by sending a costly signal. Other economists have considered leadership styles in a management context (Rotemberg and Saloner, 1993, 2000; Van den Steen, 1993). A manager's style is determined by her preferences over different possible decisions over which she has residual control in an incomplete-contracting environment. A top-level principal then chooses the style of manager in order to best resolve a moral-hazard problem. Within political science, researchers have considered a variety of leadership objectives, including a desire to survive in office (Bueno de Mesquita, Morrow, Siverson, and Smith, 2002) or to manipulate the agenda to achieve desired outcomes (Riker, 1996). Often a leader faces conflicting objectives; for instance, Canes-Wrone, Herron, and Shotts (2001) analyzed a situation in which a leader faces a trade-off between doing the right thing and taking an action that is politically convenient.

Our notion of leadership accords with the textbook description (Northouse, 2007, p. 3) of "a process whereby an individual influences a group of individuals to achieve a common goal." Our followers form a group who would like to the right thing and do it together, but are incompletely informed. We view leadership as a mechanism that allows them to make informed choices and to coordinate. The idea that leaders can help explain when and how their followers can successfully pursue a common goal is central to a recent paper by Majumdar and Mukand (2008), and the coordinating role of leaders has been considered in recent papers by Calvert (1995), Myerson (2004), Dewan and Myatt (2007, 2008), Dickson (2008), and Bolton, Brunnermeier, and Veldkamp (2008b).

Our modeling technique uses the "beauty contest" games which were introduced to the recent literature by Morris and Shin (2002), but which are related to the classic team-decision problems of Radner (1962) and Marshak and Radner (1972). Variants of such games have been used to study public announcements in monetary economies (Amato, Morris, and Shin, 2002; Hellwig, 2005), complementary investments (Angeletos and Pavan, 2004), asset pricing (Allen, Morris, and Shin, 2006), and the networked communication of information (Calvó-Armengol and de Martí Beltran, 2007, 2009). In using this

---

[17]There have also been recent empirical developments, including work by Wantchekon (2003, 2007), Humphreys, Masters, and Sandbu (2006), and Goodall, Kahn, and Oswald (2008).

framework to examine leadership we relate to both Dewan and Myatt (2008) and subsequent work by (Bolton, Brunnermeier, and Veldkamp, 2008a). A distinguishing feature of our model is that the clarity of a message is endogenous: followers choose to whom to listen, and leaders choose how much to say and how clearly to speak. Endogenous information acquisition in beauty contests was studied by Hellwig and Veldkamp (2009) who noted that complementarity between players' actions can lead to complementarity of decisions in an associated information-acquisition game; endogenous networked communication in organizations was studied in a very nice paper by Calvó-Armengol, de Martí Beltran, and Prat (2009) which also relates to work by Chwe (1999, 2000); and Myatt and Wallace (2009) offered a result which is closely related to our Proposition 6.

In our model and its antecedents, leadership facilitates information transmission and provides 'focal' coordination. But (as always) there are trade-offs: when offering an opinion a leader provides valuable information, but also adds to the diversity of views. Her rhetorical strategy responds to this trade-off. By adapting her rhetoric a leader affects the influence of her views on mass opinion. For example, we find that when a leader places more emphasis on unity (relative to party activists) she may restrict communication, either by obfuscating or by standing back from the leadership podium, and thereby deflect influence to other leaders better able to serve as unifying focal points.

Of course, the insight that obfuscation may emerge as a rhetorical strategy is not new: researchers have considered leaders who obfuscate in order to attract attention (Dewan and Myatt, 2008) or to convince followers of the veracity of their message (Hafer and Landa, 2008). Here, obfuscation emerges under a wider range of political objectives—it emerges whenever a leader wishes her followers to place very different emphases on the twin concerns of policy and unity. More generally, rhetorical strategies respond to the relative skills of the leadership elite. For example, whilst a benevolent leader never obfuscates, she may stand back or even step down, although only when she is a relatively poor communicator. By contrast when a leader emphasizes policy over unity, she will never stand down but may obfuscate (partially) or stand back only if she is a relatively clear communicator. The relationship between rhetorical strategies and institutional rules of debate connects our work to that of Hafer and Landa (2007, 2008); when leaders are benevolent then the length of time a leaders' views are heard corresponds to the amount of floor time that a (benevolent) moderator of the debate would allocate to her.

A major final contribution of the paper is to show how the size and character of the leadership elite emerges endogenously and in response to leaders' rhetorical strategies. A leader might obfuscate or stand back so far that she effectively refrains from any meaningful discourse. By stepping down from the podium she pulls her hat out of the ring and

so constrains her followers' choices. Such rhetorical maneuvers are, of course, akin to the heresthetical technique (Riker, 1996, p. 9) of "composing the alternatives among which political actors must choose—in such a way that even those who do not wish to do so are compelled by the structure of the situation to support the heresthetician's purpose." Our main results characterize such situations; in particular, when leaders emphasize unity, then a single, focal, and unifying leader emerges because others step out of the limelight.

OMITTED PROOFS

Proof of Lemma 1. $\pi \, \mathrm{E}[(a_m - \theta - b_m)^2] + (1 - \pi) \, \mathrm{E}[(a_m - \bar{a})^2]$ is strictly convex in a_m and so the first-order condition $a_m = \pi(\mathrm{E}[\theta] + b_m) + (1 - \pi) \, \mathrm{E}[\bar{a}]$ is necessary and sufficient for a solution. The final statement follows from the argument in the text. Formally, with a finite set of activists:

$$\frac{\partial \sum_{m' \neq m} u_{m'}}{\partial m} = -2 \left(a_m - \frac{\sum_{m' \neq m} a_{m'}}{M - 1} \right) > 0 \quad \Leftrightarrow \quad a_m < \frac{\sum_{m' \neq m} a_{m'}}{M - 1},$$

and so activist m exerts a net positive externality on others by moving toward the center. Also with a finite collection of activists, the first-order condition is

$$a_m = \pi(\mathrm{E}[\theta] + b_m) + (1 - \pi) \, \mathrm{E}\left[\frac{\sum_{m' \neq m} a_{m'}}{M - 1} \right]$$

$$= \pi_M(\mathrm{E}[\theta] + b_m) + (1 - \pi_M) \, \mathrm{E}\left[\frac{\sum_{m'} a_{m'}}{M} \right] \quad \text{where} \quad \pi_M \equiv \frac{(M - 1)\pi}{M - \pi},$$

which yields $a_m = \pi(\mathrm{E}[\theta] + b_m) + (1 - \pi) \, \mathrm{E}[\bar{a}]$ in the limit as $M \to \infty$. $\qquad\square$

Proof of Proposition 1. The argument given in the text implies that an equilibrium strategy takes the form $A(\tilde{s}, b_m) = \hat{A}(\tilde{s}_m) + \pi b_m$ where $\hat{A}(\tilde{s}_m)$ is the action taken by an unbiased activist. Noting that $\mathrm{E}[A(\tilde{s}_{m'}, b'_m) \,|\, \tilde{s}_m] = \mathrm{E}[\hat{A}(\tilde{s}_{m'}) \,|\, \tilde{s}_m]$, and using (5), $\hat{A}(\tilde{s}_m)$ must satisfy

$$\hat{A}(\tilde{s}_m) = \pi \, \mathrm{E}[\theta \,|\, \tilde{s}_m] + (1 - \pi) \, \mathrm{E}[\hat{A}(\tilde{s}_{m'}) \,|\, \tilde{s}_m].$$

This is condition $(\star)$ from Dewan and Myatt (2008, p. 354). Applying their Proposition 1 generates:

$$w_i^* = \frac{\psi_i^*}{\sum_{j=1}^{n} \psi_j^*} \quad \text{where} \quad \psi_i^* \equiv \frac{1}{\pi \kappa_i^2 + (\sigma_i^2 / x_i)}.$$

The first three comparative-static claims hold by inspection. For the final such claim,

$$\frac{w_i^*}{w_j^*} = \frac{\psi_i^*}{\psi_j^*} = \frac{\pi \kappa_j^2 + (\sigma_j^2 / x_j)}{\pi \kappa_i^2 + (\sigma_i^2 / x_i)} = \frac{\kappa_j^2 + (\sigma_j^2 / x_j)}{\kappa_i^2 + (\sigma_i^2 / x_i)} \times \frac{(1 - \rho_j) + \pi \rho_j}{(1 - \rho_i) + \pi \rho_i},$$

which by inspection is decreasing in π whenever $\rho_i > \rho_j$. $\qquad\square$

Derivation of Equation (10). Set $a_m = \pi b_m + \sum_{i=1}^{n} w_i \tilde{s}_{im}$ and evaluate (9) to obtain

$$v = \bar{v} - \tilde{\pi}(1 - \pi)\beta^2 - (2 - \tilde{\pi}) \sum_{i=1}^{n} w_i^2 \left[\pi^{\ddagger} \kappa_i^2 + \frac{\sigma_i^2}{x_i} \right] \quad \text{where} \quad \pi^{\ddagger} \equiv \frac{\tilde{\pi}}{2 - \tilde{\pi}}. \qquad\square$$

Proof of Propositions 2 and 3. A leader's desired weights maximize the expected payoff v from (10). This is equivalent to minimizing $\sum_{i=1}^{n} w_i^2 \left[\pi^\ddagger \kappa_i^2 + (\sigma_i^2/x_i) \right]$ subject to $\sum_{i=1}^{n} w_i = 1$. The solution satisfies (11). Recycling the argument used to prove the final claim of Proposition 1, $w_i^\ddagger/w_j^\ddagger$ is decreasing in $\pi^\ddagger$ if and only if $\rho_i > \rho_j$. Thus if an increase in a leader's concern for unity ($\pi^\ddagger \downarrow$) raises $w_i^\ddagger$ then it must raise the influence of those who are comparatively clearer than i. This yields Proposition 3. The claims of Proposition 2 follow from (11) and as corollaries of Proposition 3. $\quad\square$

Proof of Lemma 2. Note that $\partial \psi_i^*/\partial[\sigma_i^2/x_i] = -(\psi_i^*)^2$. Using this fact, for $j \neq i$:

$$\frac{\partial w_j^*}{\partial[\sigma_i^2/x_i]} = \psi_i^* w_i^* w_j^* \quad \text{and} \quad \frac{\partial w_i^*}{\partial[\sigma_i^2/x_i]} = -\psi_i^* w_i^*(1 - w_i^*).$$

Differentiate v from (10) with respect to $[\sigma_i^2/x_i]$ obtain

$$\frac{\partial v}{\partial[\sigma_i^2/x_i]} = -\underbrace{(2 - \tilde{\pi})(w_i^*)^2}_{\text{direct effect}} - \underbrace{(2 - \tilde{\pi}) \sum_{j=1}^{n} 2 w_j^* \left(\pi^\ddagger \kappa_j^2 + \frac{\sigma_j^2}{x_j} \right) \frac{\partial w_j^*}{\partial[\sigma_i^2/x_i]}}_{\text{strategic effect}}.$$

Next substitute in the terms derived previously to yield

$$\frac{\partial v/\partial[\sigma_i^2/x_i]}{2(2-\tilde{\pi})(w_i^*)^2} = -\frac{1}{2} + (1 - w_i^*)\psi_i^* \left(\pi^\ddagger \kappa_i^2 + \frac{\sigma_i^2}{x_i} \right) - \sum_{j \neq i} w_j^* \psi_j^* \left(\pi^\ddagger \kappa_j^2 + \frac{\sigma_j^2}{x_j} \right)$$

$$= -\frac{1}{2} - \sum_{j \neq i} w_j^* \left[\frac{\pi^\ddagger \kappa_j^2 + (\sigma_j^2/x_j)}{\pi \kappa_j^2 + (\sigma_j^2/x_j)} - \frac{\pi^\ddagger \kappa_i^2 + (\sigma_i^2/x_i)}{\pi \kappa_i^2 + (\sigma_i^2/x_i)} \right]$$

$$= -\frac{1}{2} - (\pi^\ddagger - \pi) \sum_{j=1}^{n} w_j^* \left[\frac{\kappa_j^2}{\pi \kappa_j^2 + (\sigma_j^2/x_j)} - \frac{\kappa_i^2}{\pi \kappa_i^2 + (\sigma_i^2/x_i)} \right] = -\frac{1}{2} + (\pi - \pi^\ddagger)(\bar{\gamma} - \gamma_i).$$

The sign of the derivative of v is determined by the final expression. The expression can also be used to examine the second derivative of v when evaluated at a stationary point of v. That is,

$$\left. \frac{\partial^2 v}{\partial[\sigma_i^2/x_i]^2} \right|_{\partial v/\partial[\sigma_i^2/x_i]=0} > 0 \quad \Leftrightarrow \quad (\pi - \pi^\ddagger) \frac{\partial[(\bar{\gamma} - \gamma_i)]}{\partial[\sigma_i^2/x_i]} > 0.$$

Differentiating $(\bar{\gamma} - \gamma_i)$ we obtain

$$\frac{\partial[(\bar{\gamma} - \gamma_i)]}{\partial[\sigma_i^2/x_i]} = \frac{\partial}{\partial[\sigma_i^2/x_i]} \sum_{j \neq i} w_j^* \left[\frac{\kappa_j^2}{\pi \kappa_j^2 + (\sigma_j^2/x_j)} - \frac{\kappa_i^2}{\pi \kappa_i^2 + (\sigma_i^2/x_i)} \right]$$

$$= \sum_{j \neq i} \frac{\partial w_j^*}{\partial[\sigma_i^2/x_i]} \left[\frac{\kappa_j^2}{\pi \kappa_j^2 + (\sigma_j^2/x_j)} - \frac{\kappa_i^2}{\pi \kappa_i^2 + (\sigma_i^2/x_i)} \right] + \frac{(1 - w_i^*)\kappa_i^2}{[\pi \kappa_i^2 + (\sigma_i^2/x_i)]^2}$$

$$= \sum_{j \neq i} \psi_i^* w_i^* w_j^* \left[\frac{\kappa_j^2}{\pi \kappa_j^2 + (\sigma_j^2/x_j)} - \frac{\kappa_i^2}{\pi \kappa_i^2 + (\sigma_i^2/x_i)} \right] + \frac{(1 - w_i^*)\kappa_i^2}{[\pi \kappa_i^2 + (\sigma_i^2/x_i)]^2}$$

$$= \sum_{j \neq i} w_j^* \left[\frac{\psi_i^* w_i^* \kappa_j^2}{\pi \kappa_j^2 + (\sigma_j^2/x_j)} + \frac{(1 - w_i^*)\kappa_i^2}{[\pi \kappa_i^2 + (\sigma_i^2/x_i)]^2} \right] > 0.$$

Hence the sign of the second derivative of v is determined by $\pi - \pi^{\ddagger}$:

$$\left.\frac{\partial^2 v}{\partial[\sigma_i^2/x_i]^2}\right|_{\partial v/\partial[\sigma_i^2/x_i]=0} > 0 \quad \Leftrightarrow \quad \pi > \pi^{\ddagger}.$$

Hence v is quasi-convex in $[\sigma_i^2/x_i]$ if $\pi^{\ddagger} < \pi$ (leaders are unifiers) and quasi-concave if $\pi > \pi^{\ddagger}$ (leaders are policy seekers). This confirms the final statement of the lemma. $\qquad\square$

Proof of Proposition 4. The first claim of (i) follows from Lemma 2. Turning to the second claim of (i), beginning from maximum clarity $(\pi^{\ddagger} - \pi)(\gamma_i - \bar{\gamma}) > 1/2$ must be satisfied. The term $(\gamma_i - \bar{\gamma})$ is bounded above by $1/\pi$, and $(\pi^{\ddagger} - \pi)$ is bounded above by $1 - \pi$. Hence a necessary condition for obfuscation by a policy-seeker is $(1 - \pi)/\pi > 1/2$, or $\pi > 2/3$ as claimed. For the third and final claim of (i), if σ_i^2 is sufficiently large then $\gamma_i < \bar{\gamma}$ and so the inequality of (13) must fail.

The first claim of (ii) follows from Lemma 2; the second claim follows from the quasi-convexity of the leader's payoff when $\pi^{\ddagger} < \pi$. If a unifying leader obfuscates completely then the inequality of (13) must hold at some point; in particular it must hold when $\gamma_i = 0$. So a necessary condition for obfuscation by a unifier leader is that $(\pi - \pi^{\ddagger})/\pi > 1/2$, so that $\pi^{\ddagger}$ needs to be sufficiently small. For a benevolent leader $\pi^{\ddagger} = \pi/(2 - \pi)$, and this necessary inequality is not satisfied. $\qquad\square$

Proof of Proposition 5. An activist chooses to whom to listen and then picks a policy. Restricting attention to linear equilibria, it is equivalent to examine a situation in which he simultaneously chooses to whom to listen and also the weights he will place on the speeches he hears, and does so to maximize his ex ante expected payoff. Hence, we study a simultaneous-move game in which activist m chooses $x_m \in \mathbb{R}_+^n$ and $w_m \in \mathbb{R}^n$ subject to the constraint $\sum_{i=1}^n x_{im} \leq t$. Let us consider a strategy profile in which an activist of interest chooses (x_m, w_m) whereas all others choose the same strategy (x^*, w^*). Our aim is to characterize the optimal choice of (x_m, w_m). We simplify exposition, without loss of generality, by supposing that policy biases are absent.

Note that $a_m = \sum_{i=1}^n w_{im} \tilde{s}_{im}$ and $\bar{a} = \sum_{i=1}^n w_i^* s_i$. Thus activist m minimizes

$$\pi \, \mathrm{E}\left[\left(\sum_{i=1}^n w_{im}(\tilde{s}_{im} - \theta)\right)^2\right] + (1 - \pi) \, \mathrm{E}\left[\left(\left[\sum_{i=1}^n w_{im} \tilde{s}_{im}\right] - \left[\sum_{i=1}^n w_i^* s_i\right]\right)^2\right].$$

The first term is unboundedly large if $\sum_{i=1}^n w_{im} \neq 1$ and so we impose the constraints $\sum_{i=1}^n w_{im} = 1$ and $\sum_{i=1}^n w_i^* = 1$. Given that this is so, it is straightforward to confirm that the first term satisfies

$$\mathrm{E}\left[\left(\sum_{i=1}^n w_{im}(\tilde{s}_{im} - \theta)\right)^2\right] = \sum_{i=1}^n w_{im}^2 \left(\kappa_i^2 + \frac{\sigma_i^2}{x_{im}}\right).$$

Similar calculations reveal that the second term satisfies

$$\mathrm{E}\left[\left(\left[\sum_{i=1}^n w_{im} \tilde{s}_{im}\right] - \left[\sum_{i=1}^n w_i^* s_i\right]\right)^2\right] = \sum_{i=1}^n \left[\frac{w_{im}^2 \sigma_i^2}{x_{im}} + (w_{im} - w_i^*)^2 \kappa_i^2\right].$$

Bringing these elements together, an activist's ex ante expected loss is

$$\sum_{i=1}^{n} \left[w_{im}^2 \left(\pi \kappa_i^2 + \frac{\sigma_i^2}{x_{im}} \right) + (1 - \pi)(w_{im} - w_i^*)^2 \kappa_i^2 \right],$$

and he minimizes this loss subject to the relevant constraints. This is a nicely behaved minimization problem with a unique solution. This is the same solution for every activist, and so in equilibrium terms of the form $(w_{im} - w_i^*)^2$ disappear. Fixing the attention levels, the relevant weights (influences of leaders) can be obtained from Proposition 1. Turning to the equilibrium attention levels, the unique solution can be characterized using Lagrange methods. Introducing the multiplier λ on the constraint $\sum_{i=1}^{n} x_{im} \leq t$, the first-order conditions with respect to positive attention levels takes the form $w_{im}^2 \sigma_i^2 = \lambda x_{im}^2$. Clearly, a leader receives attention in equilibrium if and only if she has influence. Let us label leaders to that the first n^* leaders have influence. Evaluated at equilibrium, we substitute in the solutions for w_i^* (from Proposition 1) to obtain

$$x_i^* \sqrt{\lambda} = \frac{\sigma_i \psi_i^*}{\sum_{j=1}^{n^*} \psi_j^*} \quad \text{where} \quad \psi_i^* = \frac{1}{\pi \kappa_i^2 + (\sigma_i^2 / x_i^*)}.$$

Following re-arrangement,

$$x_i^* = \frac{\sigma_i (K_{n^*} - \sigma_i)}{\pi \kappa_i^2} \quad \text{where} \quad K_{n^*} \equiv \frac{1}{\lambda \sum_{i=1}^{n^*} \psi_i^*}.$$

This is only valid if $\sigma_i < K$, so that $x_i \geq 0$ is satisfied. Moreover, it is easily checked that if $\sigma_j < K$ for some $j > n^*$ then it is optimal for an activist to introduce both positive attention and influence to the jth leader. Hence the set of n^* leaders who receive attention comprises the n^* clearest leaders. Since this is so, without loss of generality we label leaders so that $\sigma_1^2 \geq \ldots \sigma_n^2$.

It remains, however, to determine n^*. To do this we must first solve for K_{n^*}. To do this, substitute the solutions for x_i^* into the binding constraint $\sum_{i=1}^{n^*} x_i = t$ to yield

$$t = -\sum_{j=1}^{n^*} \frac{\sigma_i^2}{\pi \kappa_i^2} + K_{n^*} \sum_{j=1}^{n^*} \frac{\sigma_i}{\pi \kappa_i^2} \quad \Leftrightarrow \quad K_{n^*} = \frac{\pi t + \sum_{j=1}^{n^*} (\sigma_i^2 / \kappa_i^2)}{\sum_{j=1}^{n^*} (\sigma_i / \kappa_i^2)}.$$

To find n^* we need only find integer n^* satisfying $\sigma_{n^*} \leq K_{n^*} < \sigma_{n^*+1}$. (If $n^* = n$ then we need only the first inequality.) An argument of Dewan and Myatt (2008) extends here to show that there is a unique such n^*, which generates the unique equilibrium. Finally, the Lagrange multiplier satisfies:

$$\lambda = \frac{\pi}{\sum_{i=1}^{n^*} (K_{n^*} - \sigma_i) / \kappa_i^2}.$$

The comparative-static claims in the proposition follow by inspection. For instance, an increase in t increases K_j for all $j \in \{1, \ldots, n\}$ and so must increase both n^* and x_i^* for each i. Notice also that the Lagrange multiplier λ is decreasing in K and so decreasing in t. Since the Lagrange multiplier is the marginal benefit of a relaxation of the budget constraint on an activist's time, this means that there are decreasing returns to the overall length of an activist's attention span. $\square$

Proof of Proposition 6. The proof of Proposition 5 demonstrated that there are decreasing returns to an activist's attention span. Hence there is unique choice of t^* satisfying $\lambda = C'(t^*)$. The claim regarding the activists' concern for unity is proven by demonstrating that λ is increasing in π for each t. This can be readily confirmed by differentiating the expression for λ:

$$\frac{\partial \lambda}{\partial \pi} = \frac{1}{\sum_{i=1}^{n^*}(K_{n^*} - \sigma_i)/\kappa_i^2} - \frac{\pi}{[\sum_{i=1}^{n^*}(K_{n^*} - \sigma_i)/\kappa_i^2]^2}\sum_{j=1}^{n^*}\frac{1}{\kappa_i^2}\frac{\partial K_{n^*}}{\partial \pi} > 0 \quad \Leftrightarrow$$

$$\sum_{i=1}^{n^*}\frac{K_{n^*} - \sigma_i}{\kappa_i^2}\frac{\pi t\sum_{i=1}^{n^*}(1/\kappa_i^2)}{\sum_{i=1}^{n^*}(\sigma_j/\kappa_i^2)} \quad \Leftrightarrow \quad \frac{\sum_{i=1}^{n^*}(1/\kappa_i^2)}{\sum_{i=1}^{n^*}(\sigma_i/\kappa_i^2)} > \frac{\sum_{i=1}^{n^*}(\sigma_i/\kappa_i^2)}{\sum_{i=1}^{n^*}(\sigma_i^2/\kappa_i^2)}$$

$$\Leftrightarrow \quad \sum_{i \neq j}\frac{\sigma_i^2 + \sigma_j^2 - 2\sigma_i\sigma_j}{\kappa_i^2\kappa_j^2} > 0.$$

This final inequality holds by inspection: each numerator term is equal to $(\sigma_i - \sigma_j)^2 > 0$. The final claim of the proposition follows from the argument given in the text. □

Proof of Propositions 7 and 8. Using (10), a leader seeks to minimize $\sum_{i=1}^n(w_i^*)^2\left[\pi^\ddagger\kappa_i^2 + (\sigma_i^2/x_i)\right]$, where w_i^* are the equilibrium weights. Beginning from the equilibrium attention levels x_i^*, there is no first-order direct effect of changing the pattern of attention. Thus the only first-order effect is to change indirectly the equilibrium weights. Specializing to a two-leader world, a goal-oriented leader will wish to lower the attention paid to her if and only if she wishes to lower the weight placed on her speech; this is so if and only if $(\pi - \pi^\dagger)(\gamma_i - \gamma_j) < 0$; that is, when the better communicator is a policy seeker or when the poorer communicator is a unifier. To obtain the claims of Proposition 8 and the first claim of Proposition 7 we need only show that $\gamma_i > \gamma_j$ if and only if $\sigma_i < \sigma_j$. This follows from substituting the equilibrium attention levels into the formula for comparative clarity. The remaining claim of Proposition 7 is a corollary of Proposition 9. □

Proof of Proposition 9. Using an expression from the proof of Lemma 2:

$$\frac{\partial v}{\partial[\sigma_i^2/x_i]} \propto (w_i^*)^2\left[-\frac{1}{2} + (\pi - \pi^\ddagger)(\bar{\gamma} - \gamma_i)\right] \quad \Rightarrow$$

$$\frac{\partial v}{\partial x_i} \propto \left(\frac{w_i^*\sigma_i}{x_i}\right)^2\left[\frac{1}{2} - (\pi - \pi^\ddagger)(\bar{\gamma} - \gamma_i)\right] \propto \frac{\partial v}{\partial x_i} \propto \left(\frac{\sigma_i}{\pi\kappa_i^2 x_i + \sigma_i^2}\right)^2\left[\frac{1}{2} - (\pi - \pi^\ddagger)(\bar{\gamma} - \gamma_i)\right].$$

Evaluation these expressions at $x_i = 0$ (so that $\gamma_i = 0$) and at $x_j = t = 1$ (so that $\gamma_j = \bar{\gamma}$):

$$\left.\frac{\partial v}{\partial x_i}\right|_{x_i=0} \propto \frac{1}{\sigma_i^2}\left[\frac{1}{2} - (\pi - \pi^\ddagger)\bar{\gamma}\right] \quad \text{and} \quad \left.\frac{\partial v}{\partial x_j}\right|_{x_j=1} \propto \left(\frac{\sigma_j}{\pi\kappa_j^2 + \sigma_j^2}\right)^2\frac{1}{2}.$$

Consider a situation in which $n = 2$, $x_i = 0$, and $x_j = 1$. Then:

$$\left[\frac{\partial v}{\partial x_i} - \frac{\partial v}{\partial x_j}\right]_{x_i=0,x_j=1} > 0 \quad \Leftrightarrow \quad 1 - 2(\pi - \pi^\ddagger)\bar{\gamma} > \frac{\sigma_i^2\sigma_j^2}{(\pi\kappa_j^2 + \sigma_j^2)^2}.$$

So, if this inequality holds then a goal-oriented leader would like attention to move away from the de facto dictator j and toward leader i. To fully evaluate this expression, notice that when leader

j receives all of the attention then $\bar{\gamma} = \kappa_j^2/(\pi\kappa_j^2 + \sigma_j^2)$. Hence the inequality becomes

$$1 - \frac{2(\pi - \pi^\ddagger)\kappa_j^2}{\pi\kappa_j^2 + \sigma_j^2} > \frac{\sigma_i^2\sigma_j^2}{(\pi\kappa_j^2 + \sigma_j^2)^2} \quad \Leftrightarrow \quad \left[1 + \frac{(2\pi^\ddagger - \pi)\kappa_j^2}{\sigma_j^2}\right]\left[1 + \frac{\pi\kappa_j^2}{\sigma_j^2}\right] > \frac{\sigma_i^2}{\sigma_j^2}.$$

When this inequality fails then it is preferable for leader j to retain all of the attention; that is, if her desired emphasis on policy is $\pi^\ddagger$ and the inequality holds then leader i prefers to step down. These calculations yields the first and third displayed inequalities in the statement of the proposition. The remaining inequalities concern the activists' desire to listen to a leader, and can be readily obtained by setting $\pi^\ddagger = \pi$ in the expressions derived here. □

Proof of Proposition 10. Follows from the argument given in the main text. □

References

ALLEN, F., S. MORRIS, AND H. S. SHIN (2006): "Beauty Contests and Iterated Expectations in Asset Markets," *Review of Financial Studies*, 19(3), 719–752.

AMATO, J. D., S. MORRIS, AND H. S. SHIN (2002): "Communication and Monetary Policy," *Oxford Review of Economic Policy*, 18(4), 495–503.

ANGELETOS, G.-M., AND A. PAVAN (2004): "Transparency of Information and Coordination in Economies with Investment Complementarities," *American Economic Review: AEA Papers and Proceedings*, 94(2), 91–98.

——— (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103–1142.

——— (2008): "Policy with Dispersed Information," *Journal of the European Economic Association*, 7(1), 11–60.

BLUME, A., AND O. BOARD (2009): "Intentional Vagueness," unpublished paper, University of Pittsburgh.

BOLTON, P., M. BRUNNERMEIER, AND L. VELDKAMP (2008a): "Economists' Perspectives on Leadership," unpublished paper, New York University.

——— (2008b): "Leadership, Coordination and Mission-Driven Management," NBER Working Paper 14339.

BUENO DE MESQUITA, B., J. D. MORROW, R. M. SIVERSON, AND A. SMITH (2002): "Political Institutions, Policy Choice, and the Survival of Leaders," *British Journal of Political Science*, 32(4), 559–590.

CALVERT, R. L. (1995): "The Rational Choice of Social Institutions: Cooperation, Coordination, and Communication," in *Modern Political Economy: Old Topics, New Directions*, ed. by J. S. Banks, and E. A. Hanushek. Cambridge University Press, Cambridge, UK.

CALVÓ-ARMENGOL, A., AND J. DE MARTÍ BELTRAN (2007): "Communication Networks: Knowledge and Decision," *American Economic Review: AEA Papers and Proceedings*, 97(2), 86–91.

——— (2009): "Information Gathering in Organizations: Equilibrium, Welfare, and Optimal Network Structure," *Journal of the European Economic Association*, 7(1), 116–161.

CALVÓ-ARMENGOL, A., J. DE MARTÍ BELTRAN, AND A. PRAT (2009): "Endogenous Communication in Endogenous Organizations," unpublished paper, Universitat Pompeu Fabra.

CANES-WRONE, B., M. C. HERRON, AND K. W. SHOTTS (2001): "Leadership and Pandering: A Theory of Executive Policymaking," *American Journal of Political Science*, 45(3), 532–550.

CHWE, M. S.-Y. (1999): "Structure and Strategy in Collective Action," *American Journal of Sociology*, 105(1), 128–156.

——— (2000): "Communication and Coordination in Social Networks," *Review of Economic Studies*, 67(1), 1–16.

DEWAN, T., AND D. P. MYATT (2007): "Leading the Party: Coordination, Direction amd Communication," *American Political Science Review*, 101(4), 825–843.

——— (2008): "The Qualities of Leadership: Communication, Direction and Obfuscation," *American Political Science Review*, 102(3), 351–368.

DICKSON, E. S. (2008): "Leadership, Followership, and Beliefs About the World: Theory and Experiment," unpublished paper, New York University.

FEDDERSEN, T. J., AND W. PESENDORFER (1997): "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, 65(5), 1029–1058.

FERREIRA, D., AND M. REZENDE (2007): "Corporate Strategy and Information Disclosure," *RAND Journal of Economics*, 38(1), 164–184.

GOODALL, A. H., L. M. KAHN, AND A. J. OSWALD (2008): "Why Do Leaders Matter? The Role of Expert Knowledge," IZA Discussion Paper 3583.

HAFER, C., AND D. LANDA (2007): "Deliberation as Self-Discovery and Institutions for Political Speech," *Journal of Theoretical Politics*, 19(3), 329–360.

——— (2008): "Majoritarian Debate," unpublished paper, New York University.

HELLWIG, C. (2005): "Heterogeneous Information and the Welfare Effects of Public Information Disclosures," unpublished paper, UCLA.

HELLWIG, C., AND L. VELDKAMP (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition," *Review of Economic Studies*, 76(1), 223–251.

HERMALIN, B. E. (1998): "Toward an Economic Theory of Leadership: Leading by Example," *American Economic Review*, 88(5), 1188–1206.

——— (2007): "Leading for the Long Term," *Journal of Economic Behavior and Organization*, 62(1), 1–19.

HUMPHREYS, M., W. MASTERS, AND M. E. SANDBU (2006): "The Role of Leadership in Democratic Deliberations: Results from a Field Experiment in São Tomé and Príncipe," *World Politics*, 58(4), 583–622.

KEYNES, J. M. (1936): *The General Theory of Employment, Interest and Money*. Macmillan, London.

LEVI, M. (2006): "Why We Need a New Theory of Government," *Perspectives on Politics*, 4(1), 5–19.

MAJUMDAR, S., AND S. W. MUKAND (2008): "The Leader as Catalyst: On Leadership and the Mechanics of Institutional Change," Queen's University Economics Department Working Paper 1128.

MARSHAK, J., AND R. RADNER (1972): *The Economic Theory of Teams*. Yale University Press, New Haven.

MCKELVEY, R. (1980): "Ambiguity in Spatial Models of Policy Formation," *Public Choice*, 35(4), 385–402.

MEIROWITZ, A. (2005): "Informational Party Primaries and Strategic Ambiguity," *Journal of Theoretical Politics*, 17(1), 107–136.

MORRIS, S., AND H. S. SHIN (2002): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–1134.

——— (2005): "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, No. 2(2005), 1–43.

MYATT, D. P., AND C. WALLACE (2009): "Endogenous Information Acquisition in Coordination Games with Multiple Signals," unpublished paper, University of Oxford.

MYERSON, R. B. (2004): "Justice, Institutions, and Multiple Equilibria," *Chicago Journal of International Law*, 5(1), 91–107.

NORTHOUSE, P. G. (2007): *Leadership: Theory and Practice*. Sage, fourth edn.

PAGE, B. I. (1976): "The Theory of Political Ambiguity," *American Political Science Review*, 70(3), 742–752.

RADNER, R. (1962): "Team Decision Problems," *Annals of Mathematical Statistics*, 33(3), 857–881.

RIKER, W. (1996): *The Strategy of Rhetoric: Campaigning for the American Constitution*. Yale University Press, New Haven.

ROTEMBERG, J. J., AND G. SALONER (1993): "Leadership Style and Incentives," *Management Science*, 39(11), 1299–1318.

——— (2000): "Visionaries, Managers, and Strategic Direction," *RAND Journal of Economics*, 31(4), 693–716.

SCHELLING, T. C. (1960): *The Strategy of Conflict*. Harvard University Press, Cambridge, MA.

SHEPSLE, K. A. (1972): "The Strategy of Ambiguity: Uncertainty and Electoral Competition," *American Political Science Review*, 66(2), 555–569.

VAN DEN STEEN, E. (1993): "Organizational Beliefs and Managerial Vision," *Journal of Law, Economics, and Organization*, 21(1), 256–283.

WANTCHEKON, L. (2003): "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin," *World Politics*, 55(3), 399–422.

——— (2007): "Expert Information, Public Deliberation, and Electoral Support for "Good" Governance: Experimental Evidence from Benin," unpublished paper, New York University.

YUKL, G. A. (2008): *Leadership in Organizations*. Pearson Education, sixth edn.