On the rhetorical strategies of leaders: Speaking clearly, standing back, and stepping down

Torun Dewan
London School of Economics and Political Science, London, UK

David P Myatt
London Business School, London, UK

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.

1. Introduction

In an uncertain world, leaders can help followers to take more informed decisions and to coordinate; indeed, Schelling (1960: 91) suggested that 'the coordination game lies behind the stability of institutions and traditions and perhaps the phenomenon of leadership itself'. Recently others have emphasized the role of leaders as coordinating focal points (Bolton et al., 2008b; Calvert, 1995; Dewan and Myatt, 2007, 2008; Myerson,

Corresponding author:
Torun Dewan, London School of Economics and Political Science, Houghton St, London WC2A 2AE, London, UK.
Email: t.dewan@lse.ac.uk

2004). 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. Previous analysis has, however, taken the size and character of the leadership elite as given, and hence as non-responsive to the strategies that leaders deploy. Here we ask the following: 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 how do the critical features of the leadership elite respond to the rhetorical strategies that leaders employ?

We explore how leaders' rhetorical strategies shape the alternatives available to followers. In this sense, our work relates to Riker's concept of heresthetic, a term he used (Riker, 1996: 9) to describe 'the art of setting up situations—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'. We assess two critical elements of a leader's rhetorical strategy: the first is her choice of clarity – she might speak as clearly as her natural abilities allow, or alternatively, she may obfuscate; the second is the length of time for which she speaks. By endogenously manipulating these components of her rhetorical strategy, a leader shapes the choice set of followers, affects her own influence and, indirectly, that of other leaders. The joint set of rhetorical strategies thus give shape to the leadership elite that forms.

To explore the role that rhetoric plays in shaping alternatives, we develop an extension to 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. 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.

We build our analysis in several steps. Firstly, and to provide a role for strategic leadership, we show that in the absence of leadership unity is undersupplied. We then characterize the influence of leaders on their followers' actions and show that a follower places relatively greater weight on the words of relatively clearer communicators and, when possible, he listens to them for longer; they are better placed to coordinate him with others. Nevertheless, followers' decisions are inefficient; they fail to incorporate fully the impact of unity. A follower would improve (on average) the lot of others by moving toward the perceived center, and would do this by shifting toward clearer (and hence focal) leaders.

These preliminary results allow us to explore how leaders can help followers to achieve their objectives, or perhaps guide them towards actions that leaders find desirable. Insufficient unity opens a window of opportunity for strategic leadership. In particular it suggests a role for a benevolent leader, although we also consider other objective functions for leaders (for example, some leaders might prefer that followers place even more weight on the policy target). We allow each leader to pursue her objectives by altering

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).

A natural finding is that 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. We show that 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 show that a relatively poor communicator stands back if she prefers to see a greater emphasis on the unity motive, whereas a clear communicator gives way only when she feels that followers place too much emphasis on the coordination motive. We also demonstrate circumstances in which 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.

Building on these insights, we directly assess the relationship between rhetorical strategies and the properties of the endogenous leadership elite that forms. We do so with the help of a simple two-leader scenario in which leadership could be provided by one leader (de facto dictatorship) or by both (oligarchy). 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 listen only to a single leader then the de facto dictator allows space for other opinions by standing back: a dictatorship reverts to an oligarchy. We show conditions under which the rhetorical strategies of leaders induce such regime switches. In one example, a benevolent leader, with poor communication skills, and recognizing that followers place insufficient weight on coordination, steps down to promote unity. In another, a leader with good communication skills recognizes that followers place insufficient weight on reaching the target policy, and stands back to allow other leaders with diverse views to be heard. In sum, by adapting their strategies leaders directly influence the choice sets of followers.

A discussion of related literature is postponed to our concluding section, and so we highlight only a few contributions here. Our focus on the role of leadership in addressing the undersupply of party unity clearly relates our work to that of Cox and McCubbins (1993). In that account, however, unity is undersupplied due to the private objectives of party actors. In our model we find, somewhat surprisingly, that the personal preferences of followers play no role in outcomes. Moreover, the mechanism for resolving party tensions is different: in Cox and McCubbins (1993), leaders wield sticks and offer carrots to enhance efficiency; in our model leaders supply information (and manipulate that information via their rhetorical strategies).

The tension between doing the right thing and doing it together, central to our account, relates to the 'beauty contest' scenario described by Keynes (1936: Chapter 12) and developed 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 (Calvert, 1995; Dewan and Myatt, 2007, 2008; Dickson, 2008; Myerson, 2004; Schelling, 1960).

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 et al. (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 arguments. Obfuscation is related to strategic ambiguity (McKelvey, 1980; Page, 1976; Shepsle, 1972), 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 'stepping down' resonates with the analyses of voting models which suggest that abstention can be beneficial (Feddersen and Pesendorfer, 1997).

2. A model of policy performance and party unity

Our study uses a game in which players (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 disagree about the right solution to a 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 have other objectives. We will characterize the rhetorical strategies – speaking clearly, standing back, and stepping down – which a leader may seek to employ.

Formally, we analyze a simultaneous-move game played by a mass of activist party members indexed by $m \in [0, 1]$.[1] Activist m must advocate a policy $a_m \in \mathbb{R}$ which may be interpreted as the position he supports at a party conference, or the policy he promotes on the doorstep during an election campaign; hence $\mathbb{R}$ 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 and b_m for the personal bias of activist m. An activist's payoff declines as his action deviates from $\theta + b_m$ via a quadratic-loss function $(a_m - \theta - b_m)^2$. 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. When (in the next section) we specify an activist's sources of information about θ we assume that those sources are independent of his bias. A second motivation for an activist is party unity: he wishes to coordinate

with others. We capture this via the quadratic loss $E[(a_m - a_{m'})^2]$ where the expectation is taken by looking across other party members indexed by m'.[2]

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) but do it 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:

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

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 information sources in the next section. Nevertheless, we can already explore some properties of activists' actions. 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 $E[(a_m - \theta - b_m)^2]$. His loss from party disunity is

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

Here $\bar{a}$ is the party line: the average policy advocated across the party's membership. The first component of disunity is an activist's non-conformity with the party line. The second is others' non-conformity and so is seen as exogenous by activist m. Thus an individual acts to minimize a weighted average of his expected deviation from his ideal policy and his expected non-conformity; he minimizes $\pi E[(a_m - \theta - b_m)^2] + (1 - \pi)E[(a_m - \bar{a})^2]$. This generates a simple lemma.

Lemma 1. *An activist's individually optimal action is* $a_m = \pi(E[\theta] + b_m) + (1 - \pi)E[\bar{a}]$ *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.*

Notice that each activist benefits from the conformity of others; he would be better off if his colleagues all moved inward. Alas, activists do not internalize this spillover. This suggests an investigation into institutional remedies (including leadership) that might help activists resolve their collective-action problem.

3. 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.[3] Prior to acting, he seeks to understand his environment. He learns by observing a collection of informative signals. These 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. We think of the signals as n speeches made by n party leaders. The leaders' speeches capture all external information relevant to an activist's play. A 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 components of a leader's skill set: her judgement and her clarity. Considering the first component, a leader is distinguished from activists in that she has substantive 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:

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

Conditional on θ we assume (with little loss of generality) that the leaders' opinions are independent; hence the collection of noise terms (η_i for each i) are uncorrelated. (Unconditionally, of course, the opinions are correlated since they are all opinions about θ.) Some opinions are more accurate reflections of θ than others: when κ_i^2 is small a leader is better able to assess policy, and so 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, activist m forms an imperfect interpretation $\tilde{s}_{im}$ where

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

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; any common shock is absorbed into η_i.[4]

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 leader i. It is equivalent to a sample size and so it is natural 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 understand that others listening to the same speech perceive its content 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 σ_i^2 and each activist's attention x_{im} as exogenous, and we assume that $x_{im} = x_i$ for all m. Later in the paper, these parameters will become endogenous: activists will choose to whom they listen, and leaders will choose both the length of their speeches (so placing an upper bound on x_i) and will be free to obfuscate (so increasing σ_i^2) if they wish.

4. Optimal policy advocacy

We look for a Bayesian Nash equilibrium. It is without loss of generality to restrict attention to equilibria in which all activists use the same advocacy strategy, so that activist m selects an action $a_m = A(\tilde{s}_m, b_m)$ based only the speeches he hears and his own policy bias. Building upon Lemma 1, $A(\cdot, \cdot)$ is an equilibrium strategy if and only if

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

Biases are orthogonal to everything else, and so b_m reveals nothing about θ or the actions of others. Using (5), this means 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 someone with no personal agenda. It is also natural to consider linear strategies, $\hat{A}(\tilde{s}_m) = \sum_{i=1}^n w_i \tilde{s}_{im}$ and where $\sum_{i=1}^n w_i = 1$. This means that (absent the bias term) an activist advocates a weighted average of the speeches he hears, where each coefficient measures the corresponding leader's influence. Linear strategies are natural because our normality assumptions ensure the linearity of the conditional expectations $\mathrm{E}[\theta \mid \tilde{s}_m]$ and $\mathrm{E}[\tilde{s}_{m'} \mid \tilde{s}_m]$. This means that if other activists use a linear strategy, then a best reply is also linear. This guarantees the existence of a (unique) linear equilibrium; this equilibrium satisfies $\sum_{i=1}^n w_i = 1$.[5]

An important element of our first main result 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 is a leader's comparative advantage in clarity.

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 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 $\rho_i > \rho_j$.

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 increases with her judgement, her clarity, and with the attention she receives; it is independent of the extent β^2 of fundamental disagreement. 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.

Notice that each leader's influence does 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 better leader enjoys more influence. However, there is an emphasis (which increases with the concern for unity) on better communicators. This is welcome since (Lemma 1) unity is undersupplied: a clear message unites extreme opinions by serving as a focal point around which the party can unite. Nevertheless, 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 equilibrium 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). In their model idiosyncratic biases are absent, and activists care about conformity (via $E[(a_m - \bar{a})^2]$) rather than unity (via $E[(a_m - a_{m'})^2]$). Despite these differences, the equilibrium weights placed on leaders' speeches are unchanged. There are two reasons for this. Firstly, β^2 has no effect because leaders' speeches are orthogonal to personal biases.[6] Secondly, and following the decomposition 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 here because the inefficiency of the equilibrium creates a role for strategic actions by benevolent leaders who may seek to enhance party unity.

5. Strategic leaders

Whilst leaders inform their followers' actions, this does not fully resolve the implicit collective-action problem. 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 leadership) is a prominent case of interest. 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

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

where the indices m and m' represented generic activists, and expectations are taken across the party's membership.[7] A benevolent leader 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 exposition by supposing that 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). In contrast to the activists' case, for a leader both components matter. Indeed,

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

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:

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

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

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

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.

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.

6. Speaking clearly

One might think that a leader who expresses herself clearly will be successful. Yet, a common perception remains that politicians do not always communicate transparently. 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.[8] Here we explore whether obfuscation is part of a leader's rhetorical armor when her objectives are related to those of her followers.

We ask whether leader i would like to reduce her clarity by increasing σ_i^2 or, equivalently, by reducing x_i. The direct effect is harmful; however, the indirect strategic effect is to push influence away from leader i and toward others. A necessary condition for obfuscation is that the strategic effect is positive, which necessitates 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*

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

Hence a leader's comparative clarity γ_i is an increasing function of the correlation coefficient ρ_i. The measure γ_i varies from $\gamma_i = 0$ (when $\kappa_i^2 = 0$) up to $\gamma_i = 1/\pi$ (when $\sigma_i^2 = 0$).

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.[9] There needs to be enough distance between the desires 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 comparative 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.

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 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 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 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.

Figure 1 illustrates 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 that

$$\frac{\kappa_1^2}{\kappa_1^2 + \sigma_1^2} = \frac{1/(\kappa_1^2 + \sigma_1^2)}{1/\kappa_1^2} = \frac{\text{Informativeness of the Leader's Speech}}{\text{Quality of the Leader's Judgement}}.$$
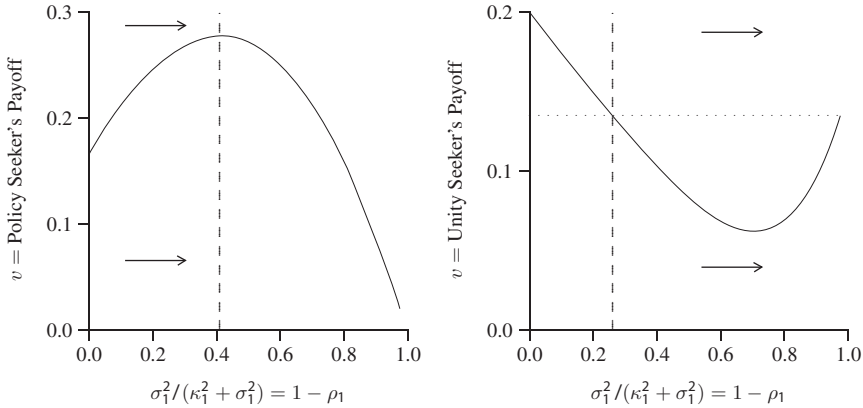
Figure 1. Obfuscation by policy-seeking and unifying leaders.

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.

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.[10] 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 for a benevolent leader obfuscation is undesirable. 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.

7. Endogenous attention

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 they 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 he would like to know what others know.[11] 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 $a_m = \pi b_m + \sum_{i=1}^{n} w_{im} \tilde{s}_{im}$.[12] 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 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: followers wish to listen to influential leaders. 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.[13] 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. This does not imply that all leaders enjoy influence and attention. 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 attention a leader receives depend on her abilities?

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*

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

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 (unsurprisingly) activists listen to more leaders when they have more time available to do so. Thirdly, below we extend our analysis to a world in which activists are able to choose the total time spent listening to leaders.

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 another; 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 for a cost of $C(\sum_{i=1}^{n} x_{im})$, where $C(\cdot)$ satisfies familiar decreasing-returns properties.[14] 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.

8. Standing back

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

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 the 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 clearer. 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.
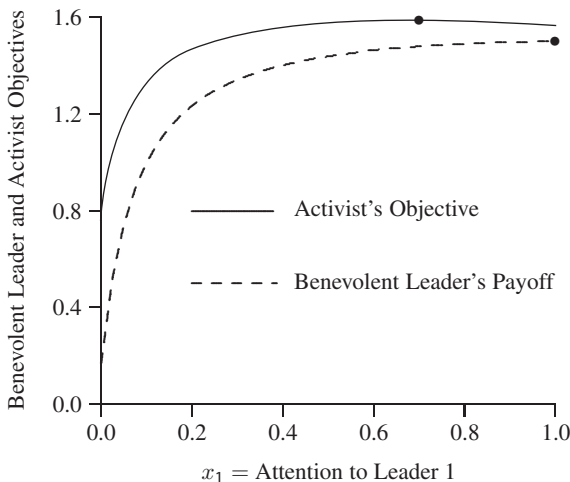

Figure 2. Equilibrium vs. efficient attention.

Invoking the idea of a social planner who wishes to enforce efficient attention levels, 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 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.

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 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.[15] Figure 2 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.

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 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 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 judgement is already accounted for in activists' equilibrium attention choices.

We conclude by recalling that a benevolent leader is automatically a unifying leader; from a welfare perspective, activists pay too little attention to the clearer leader who can act as a focal point.

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

9. Stepping down

We have seen (via Proposition 5) that a follower-generated dictatorship can emerge in which activists listen only to one leader. 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. 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 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.

Consider again two leaders who wish to encourage unity. The clearer leader always enjoys attention in equilibrium, and (Proposition 8) only the poorer communicator ever stands back. In this case, standing back can result in a 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 they 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

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

then activists prefer to devote all of their attention to the 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 leadership elite is shaped by rhetorical strategies; the various cases (determined by the inequalities in the proposition) are illustrated in Figure 3, which shows 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 to 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. When leaders emphasize unity, then an oligarchy becomes a dictatorship. This can occur because a leader steps down. However (recalling 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 making herself irrelevant to the debate. Conversely, when leaders are policy seekers and activists veer toward a relatively clear would-be dictator, leaders stand back so that the
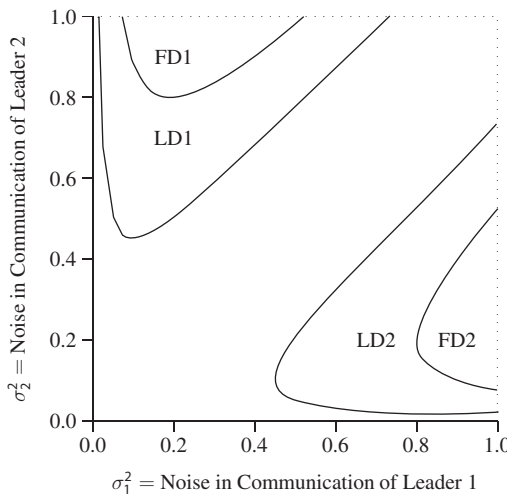

Figure 3. Endogenous leadership regimes.

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.

10. A divided oligarchy

By standing back or stepping down, leaders can change 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. Specifically, we suppose that the first leader is a better communicator (so that $\sigma_1^2 < \sigma_2^2$). Unlike previous sections the leaders 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 are unifiers. Following earlier logic, only one wishes to stand back; if the inequality contained in Proposition 8 is satisfied for leader i then it cannot be satisfied for $j \neq i$. 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; 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 aligned with their skills: the clearer leader is a policy seeker, while the poorer communicator seeks unity. Both leaders wish to lose influence, and so both wish to stand back, so long as the space created 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.

11. Related literature and concluding remarks

The broad notion of leadership is, of course, widely studied; textbooks such as Northouse (2007) and Yukl (2008) offer surveys. Perhaps surprisingly, however, there are relatively few formal models. 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.[16]

Within economics, recent theories were surveyed by Bolton et al. (2008a). Some of those focus on acts of leadership. Notably, Hermalin (1998, 2007) considered a leader with private knowledge about the nature of a task that she would like her followers to perform. To communicate credibly she 'leads 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 selects the style of manager in order to resolve a moral-hazard problem. Within political science, researchers have considered a variety of leadership objectives, including survival in office (Bueno de Mesquita et al., 2002) and agenda manipulation (Riker, 1996). Often a leader faces conflicting aims; for instance, Canes-Wrone et al. (2001) analyzed a situation in which a leader balances doing the right thing against taking an action that is politically convenient.

Our notion of leadership accords with the textbook description (Northouse, 2007: 3) of 'a process whereby an individual influences a group of individuals to achieve a common goal'. Our followers would like to do the right thing and do it together, but are incompletely informed; leadership 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 recent work by Majumdar and Mukand (2008), and the coordinating role has been considered by Calvert (1995), Myerson (2004), Dewan and Myatt (2007, 2008), Dickson (2008), and Bolton et al. (2008b).

Our modeling technique uses the 'beauty contest' games popularized by Morris and Shin (2002).[17] In using this framework to examine leadership we relate to Dewan and Myatt (2008) and subsequent work by Bolton et al., (2008a). A distinguishing feature of our model is that a message's clarity is endogenous: followers choose to whom they listen, and leaders choose how to speak. Endogenous information acquisition in beauty contests was studied by Hellwig and Veldkamp (2009) who noted that the complementarity of actions can lead to the complementarity of information-acquisition decisions; endogenous networked communication in organizations was studied in a very nice paper by Calvó-Armengol et al. (2009) which also relates to Chwe (1999, 2000), and Myatt and Wallace (2012) 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 a leader's opinion 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 those better able to serve as focal unifiers.

Of course, the insight that obfuscation may emerge is not new: a leader may obfuscate to attract attention (Dewan and Myatt, 2008) or to convince followers of the veracity of their message (Hafer and Landa, 2008). Here, obfuscation arises under a wider range of objectives. More generally, rhetorical strategies respond to the relative skills of the elite. The relationship between such 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 leader's views are heard corresponds to the 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 elite emerges. A leader might obfuscate or stand back so far that she effectively refrains from any 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 techniques (Riker, 1996) discussed in our introductory remarks. 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.

Appendix A. 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 m exerts a net positive externality by moving toward the center. Also, the first-order condition is

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

where $\pi_M \equiv (M - 1)\pi/(M - \pi)$ 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. We have argued that an equilibrium strategy satisfies $A(\tilde{s}, b_m) = \hat{A}(\tilde{s}_m) + \pi b_m$. From $\mathrm{E}[A(\tilde{s}_{m'}, b'_m) \,|\, \tilde{s}_m] = \mathrm{E}[\hat{A}(\tilde{s}_{m'}) \,|\, \tilde{s}_m]$, and using (5), $\hat{A}(\tilde{s}_m) = \pi \, \mathrm{E}[\theta \,|\, \tilde{s}_m] + (1 - \pi) \, \mathrm{E}[\hat{A}(\tilde{s}_{m'}) \,|\, \tilde{s}_m]$. This is condition ($\star$) from Dewan and Myatt (2008: 354). From their Proposition 1,

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

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

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

Proof of Propositions 2 and 3. A leader's desired weights maximize 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 proof of 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 increased concern for unity ($\pi^{\ddagger} \downarrow$) raises $w_i^{\ddagger}$ then it raises the influence of those who are comparatively clearer. This yields Proposition 3. The claims of Proposition 2 follow from (11) and as corollaries of Proposition 3. $\square$

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 $[\sigma_i^2/x_i]$ so as to obtain

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

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 $\pi - \pi^{\ddagger}$:

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

Hence v is quasi-convex in $[\sigma_i^2/x_i]$ if $\pi^{\ddagger} < \pi$ and quasi-concave if $\pi > \pi^{\ddagger}$. $\quad\square$

Proof of Proposition 4. The first claim of (i) follows from Lemma 2. For the second claim, beginning from maximum clarity, $(\pi^{\ddagger} - \pi)(\gamma_i - \bar{\gamma}) > 1/2$ must be satisfied. $(\gamma_i - \bar{\gamma})$ is bounded above by $1/\pi$, and $(\pi^{\ddagger} - \pi)$ is bounded 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 final claim of (i), if σ_i^2 is sufficiently large then $\gamma_i < \bar{\gamma}$ and so (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 (13) must hold at some point, in particular when $\gamma_i = 0$. So a necessary condition for such obfuscation is $(\pi - \pi^{\ddagger})/\pi > 1/2$, so that $\pi^{\ddagger}$ needs to be sufficiently small. For a benevolent leader $\pi^{\ddagger} = \pi/(2 - \pi)$, and this inequality fails. $\quad\square$

Proof of Proposition 5. Restricting to linear equilibria, an activist's two-step listen-then-follow strategy is equivalent to choosing simultaneously his attention levels and the weights he place on the speeches he hears, with the aim of maximizing his ex ante expected payoff. So, 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 $\sum_{i=1}^n x_{im} \leq t$. Let us consider a strategy profile in which activist m chooses (x_m, w_m) whereas all others choose (x^*, w^*). Our aim is to characterize the optimal choice of (x_m, w_m). We simplify exposition, without loss of generality, by omitting policy biases.

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 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 take 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 so 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 an 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 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 multiplier λ is decreasing in K and so decreasing in t. Since λ 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 length of an activist's attention span. □

Proof of Proposition 6. The proof of Proposition 5 demonstrated that there are decreasing returns to t. Hence there is a unique 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. Differentiating,

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

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

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

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 $\sum_{i=1}^{n}(w_i^*)^2[\pi^\ddagger \kappa_i^2 + (\sigma_i^2/x_i)]$, where w_i^* are the equilibrium weights. Beginning from the equilibrium 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 leader wishes 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 of these expressions at $x_i = 0$ (so that $\gamma_i = 0$) and at $x_j = t = 1$ (so that $\gamma_j = \bar{\gamma}$) gives

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

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

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

If this holds then a goal-oriented leader would like to shift attention from the de facto dictator j toward leader i. 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

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

If this fails then leader j prefers 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. The proof follows from the argument given in the main text. □

Acknowledgements

Notes

1. We assume that the player set is a unit mass, but our model can be easily modified to use only a finite set of followers.
2. More formally, the loss from disunity experienced by activist m is $E[(a_m - a_{m'})^2] \equiv \int_0^1 (a_m - a_{m'})^2 \, dm'$. In a world with a finite collection of M activists, disunity can be defined as $\sum_{m' \neq m} (a_m - a_m')^2/(M-1)$.

3. Formally, activists share a diffuse prior over θ. Our model can be extended with a common prior $\theta \sim N(\mu, \xi^2)$.

4. The absence of a policy bias for the leader might imply that there is no tension between the set of leaders. In fact, tensions can arise due to different leaders' perceptions as to the best policy. In earlier work (Dewan and Myatt, 2008) we allowed leaders to have private motivations which consisted of their desire to maximize the attention given to their speech.

5. A linear equilibrium is unique within a broader class of advocacy strategies. Consider strategies which do not diverge away from linearity: a 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) applies. 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 et al. (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.

6. 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.

7. 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 $\mathrm{E}[(a_m - a_{m'})^2]$.

8. A novel logic to explain obfuscatory rhetoric was given by Dewan and Myatt (2008). In their model followers choose to whom they listened (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.

9. 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 ($\pi^{\ddagger} > \pi$).

10. Our use of this term follows from the fact that, in this case, a single leader has influence over the actions taken by activists. Here, a dictatorship emerges due to the actions taken voluntarily by the leadership elite.

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

12. 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 $x_{im} \rightarrow 0$.

13. 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.

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

15. 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.

16. There is also recent empirical work (Goodall et al., 2008; Humphreys et al., 2006; Wantchekon, 2003, 2007).

17. These games are related to classic team-decision problems (Marshak and Radner, 1972; Radner, 1962). They have been used to study public announcements in monetary economies (Amato et al., 2002; Hellwig, 2005), complementary investments (Angeletos and Pavan, 2004), asset pricing (Allen et al., 2006), and networked communication (Calvó-Armengol and de Martí Beltran, 2007, 2009).

References

Allen F, Morris S and Shin HS (2006) Beauty contests and iterated expectations in asset markets. *Review of Financial Studies* 19: 719–752.

Amato JD, Morris S and Shin HS (2002) Communication and monetary policy. *Oxford Review of Economic Policy* 18: 495–503.

Angeletos G-M and Pavan A (2004) Transparency of information and coordination in economies with investment complementarities. *American Economic Review: AEA Papers and Proceedings* 94: 91–98.

Angeletos G-M and Pavan A (2007) Efficient use of information and social value of information. *Econometrica* 75: 1103–1142.

Angeletos G-M and Pavan A (2008) Policy with dispersed information. *Journal of the European Economic Association* 7: 11–60.

Blume A and Board O (2009) Intentional vagueness. Unpublished paper, University of Pittsburgh.

Bolton P, Brunnermeier M and Veldkamp L (2008a) Economists' perspectives on leadership. Unpublished paper, New York University.

Bolton P, Brunnermeier M and Veldkamp L (2008b) Leadership, coordination and mission-driven management. NBER Working Paper 14339.

Bueno de Mesquita B, Morrow JD, Siverson RM and Smith A (2002) Political institutions, policy choice, and the survival of leaders. *British Journal of Political Science* 32: 559–590.

Calvert RL (1995) The rational choice of social institutions: cooperation, coordination, and communication. In: Banks JS and Hanushek EA (eds) *Modern Political Economy: Old Topics, New Directions*. Cambridge, UK: Cambridge University Press.

Calvó-Armengol A and de Martí Beltran J (2007) Communication networks: knowledge and decision. *American Economic Review: AEA Papers and Proceedings* 97: 86–91.

Calvó-Armengol A and de Martí Beltran J (2009) Information gathering in organizations: equilibrium, welfare, and optimal network structure. *Journal of the European Economic Association* 7: 116–161.

Calvó-Armengol A, de Martí Beltran J and Prat A (2009) Endogenous communication in complex organizations. Unpublished paper, Universitat Pompeu Fabra.

Canes-Wrone B, Herron MC and Shotts KW (2001) Leadership and pandering: a theory of executive policymaking. *American Journal of Political Science* 45: 532–550.

Chwe MS-Y (1999) Structure and strategy in collective action. *American Journal of Sociology* 105: 128–156.

Chwe MS-Y (2000) Communication and coordination in social networks. *Review of Economic Studies* 67: 1–16.

Cox GW and McCubbins MD (1993) *Legislative Leviathan: Party Government in the House.* Berkeley, CA: University of California Press.

Dewan T and Myatt DP (2007) Leading the party: coordination, direction amd communication. *American Political Science Review* 101: 825–843.

Dewan T and Myatt DP (2008) The qualities of leadership: communication, direction and obfuscation. *American Political Science Review* 102: 351–368.

Dickson ES (2008) Leadership, followership, and beliefs about the world: theory and experiment. Unpublished paper, New York University.

Feddersen TJ and Pesendorfer W (1997) Voting behavior and information aggregation in elections with private information. *Econometrica* 65: 1029–1058.

Ferreira D and Rezende M (2007) Corporate strategy and information disclosure. *RAND Journal of Economics* 38: 164–184.

Goodall AH, Kahn LM and Oswald AJ (2008) Why do leaders matter? The role of expert knowledge. IZA Discussion Paper 3583.

Hafer C and Landa D (2007) Deliberation as self-discovery and institutions for political speech. *Journal of Theoretical Politics* 19: 329–360.

Hafer C and Landa D (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 Veldkamp L (2009) Knowing what others know: coordination motives in information acquisition. *Review of Economic Studies* 76: 223–251.

Hermalin BE (1998) Toward an economic theory of leadership: leading by example. *American Economic Review* 88: 1188–1206.

Hermalin BE (2007) Leading for the long term. *Journal of Economic Behavior and Organization* 62: 1–19.

Humphreys M, Masters W and Sandbu ME (2006) The role of leadership in democratic deliberations: results from a field experiment in São Tomé and Príncipe. *World Politics* 58: 583–622.

Keynes JM (1936) *The General Theory of Employment, Interest and Money.* London: Macmillan.

Majumdar S and Mukand SW (2008) The leader as catalyst: on leadership and the mechanics of institutional change. Queen's University Economics Department Working Paper 1128.

Marshak J and Radner R (1972) *The Economic Theory of Teams.* New Haven: Yale University Press.

McKelvey R (1980) Ambiguity in spatial models of policy formation. *Public Choice* 35: 385–402.

Meirowitz A (2005) Informational party primaries and strategic ambiguity. *Journal of Theoretical Politics* 17: 107–136.

Morris S and Shin HS (2002) Social value of public information. *American Economic Review* 92: 1521–1134.

Morris S and Shin HS (2005) Central bank transparency and the signal value of prices. Brookings Papers on Economic Activity No. 2: 1–43.

Myatt DP and Wallace C (2012) Endogenous information acquisition in coordination games with multiple signals. *Review of Economic Studies* 79: 340–374.

Myerson RB (2004) Justice, institutions, and multiple equilibria. *Chicago Journal of International Law* 5: 91–107.

Northouse PG (2007) *Leadership: Theory and Practice*, 4th ed. London: Sage Publications.

Page BI (1976) The theory of political ambiguity. *American Political Science Review* 70: 742–752.

Radner R (1962) Team decision problems. *Annals of Mathematical Statistics* 33: 857–881.

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

Rotemberg JJ and Saloner G (1993) Leadership style and incentives. *Management Science* 39: 1299–1318.

Rotemberg JJ and Saloner G (2000) Visionaries, managers, and strategic direction. *RAND Journal of Economics* 31: 693–716.

Schelling TC (1960) *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Shepsle KA (1972) The strategy of ambiguity: uncertainty and electoral competition. *American Political Science Review* 66: 555–569.

Van den Steen E (1993) Organizational beliefs and managerial vision. *Journal of Law, Economics, and Organization* 21: 256–283.

Wantchekon L (2003) Clientelism and voting behavior: evidence from a field experiment in Benin. *World Politics* 55: 399–422.

Wantchekon L (2007) Expert information, public deliberation, and electoral support for 'good' governance: experimental evidence from Benin. Unpublished paper, New York University.

Yukl GA (2008) *Leadership in Organizations*, 6th ed. Pearson Education.