doi:10.1017/S0003055408080234

The Qualities of Leadership: Direction, Communication, and Obfuscation

TORUN DEWAN *London School of Economics*
DAVID P. MYATT *University of Oxford*

What is leadership? What is good leadership? What is successful leadership? Answers emerge from our study of a formal model in which followers face a coordination problem: they wish to choose the best action while conforming as closely as possible to the actions of others. Although they would like to do the right thing and do it together, followers are unsure about the relative merits of their options. They learn about their environment and the likely moves of others by listening to leaders. These leaders bridge differences of opinion and become coordinating focal points. A leader's influence increases with her judgement (her sense of direction) and her ability to convey ideas (her clarity of communication). A leader with perfect clarity enjoys greater influence than one with a perfect sense of direction. When followers choose how much attention to pay to leaders, they listen only to the most coherent communicators. However, power-hungry leaders who need an audience sometimes obfuscate their messages, but less so when their followers place more emphasis on conformity than on doing the right thing.

Political scientists and commentators agree that leadership is central to the performance of organizations, and yet, fundamental questions remain open.[1] What does it mean to lead? What is good leadership? When is a leader successful? Which qualities contribute to good and successful leadership, and how do these qualities arise?

To answer these questions, we develop a formal theoretical model in which the actions of a mass of followers are shaped by the speeches made by leaders. Specifically, the followers are engaged in a coordination game: they each want to do the right thing, and do it together, but they lack full information about their environment. They form their opinions by listening to leaders, and then base their actions on what they hear. We characterize the influence of leaders with different skills both when their audiences are exogenous and when their followers can choose to whom to listen. A central message emerging from our analysis is that a leader's ability to communicate clearly to the masses is relatively more important than her ability to discover the best course of action for them. In an extension to our model, we also allow each leader to choose endogenously the clarity of her communication. Before fully describing our formal model and our results, in the remainder of this introductory section, we explain the motivation for this work.

Leadership can be important when political actors want to coordinate. As suggested by Calvert (1995), Myerson (2004), and Dewan and Myatt (2007), a leader can be focal: when a leader communicates, she helps to unify expectations about how a mass will act.[2] Leaders can also help people learn. As Levi (2006) argued recently, "leadership . . . provides the learning environment that enables individuals to transform or revise beliefs."

As an illustrative example, which we will use throughout this article as a vehicle for describing our general theory of leadership, consider the following stylized representation of a political party populated by a mass of activists. An activist advocates the policy he believes to be desirable. He may, however, not know which policy is best. An activist is also concerned with the cohesion of his party. A party is more successful when its members advocate similar policies, and less so when there is widespread discordance. Because of this, a party activist would like to advocate a policy that is in line with others; in the absence of common expectations, the "party line" may be hard to discern. In this situation, a party leader has influence via her communication. She might convey information to activists, perhaps via a direct speech to the party membership or via other media channels, and so aid them in their advocacy. This also has focal properties: her words could create a common viewpoint around which support can coalesce. This is important because an

Torun Dewan is Lecturer, London School of Economics & Political Science, London WC2A 2AE United Kingdom. E-mail: t.dewan@lse.ac.uk

David P. Myatt is Official Fellow, Nuffield College, University of Oxford, Oxford OX1 1NF United Kingdom. E-mail: david.myatt@economics.ox.ac.uk

We warmly thank Thomas Plümper, Ken Shepsle, Lee Sigelman, Chris Wallace, three anonymous referees, and seminar participants at Edinburgh, Essex, Oxford, Pisa, and Warwick, for helpful comments. We also thank Michael Chwe and the editorial team of the APSR for helping shape the final version.

[1] Of course, many authors have contributed toward theories of leadership. In the context of party leaders, we have learned how leaders manipulate the agenda (Riker 1996), serve as agents of their parties (Fiorina and Shepsle 1989), and choose policies that enhance their survival (Bueno de Mesquita et al. 2003). Economists have suggested that a leader may use "leading by example" to send a costly signal to others and hence encourage their efforts (Hermalin 1998). Of course, the work mentioned here is only a small sample; there is a large literature on leadership spanning several fields and disciplines. However, at an abstract level we believe that the questions posed here have not been fully answered.

[2] Thus, a leader might be interpreted as a coordinating "rational ritual" in the sense proposed by Chwe (2001). Here, we agree with Chwe (2001) that collective action problems extend beyond the free riding emphasized by Olson (1965) and include many situations in which coordination is required.

activist faces uncertainty not only about which is the best policy, but also about what others think is the best policy. Successful coordination depends upon accurate assessments of others' beliefs; leadership helps provide such assessments.

Within this framework, a good leader helps a mass of actors to achieve their goals: her communication fosters the understanding that is needed for them to choose the right actions, and to choose them together. However, a successful leader is one who has influence: her words impact upon the actions taken by the followers. The performance of a leader on both dimensions depends on her qualities. As Levi (2006) suggested, "[the] quality of government depends on the quality of institutions and constitutional design but also on the quality of leadership, and the accuracy of beliefs held by the population about the state of the world in which they live . . . ." But which qualities are relevant?

One such quality is a leader's *sense of direction*. A leader conveys her information about the best courses of action for the followers. The value of that information reflects the quality of her judgement. History provides us examples of those who appeared to know instinctively the best course to pursue. Of George Washington, for example, Ellis (2005) wrote: "his judgement on all the major political and military questions had invariably proved prescient . . . his genius was his judgement." Such a sense of direction might also reveal the action that is most compatible with the wider mass of political actors. For example, Carwardine (2003) argued that "to fathom the thinking of ordinary citizens and to reach out to them with uncommon assurance" was a central achievement of Abraham Lincoln. Of course, a sense of direction need not always be seen as the property of an individual: it might also arise from the combined wisdom of a cabinet of advisors.

A second relevant quality is a leader's *clarity of communication*. Good judgement is wasted unless a leader can effectively communicate her message: increased clarity enhances the informativeness of this message. Crucially, however, there is a second effect of increased clarity. When coordination is important, a follower not only wonders about the content of the message received from a leader, but also considers how others interpret it. A clear message is better able to act as a unifying focal point. Indeed, a speech that points everyone in the wrong direction, but is commonly interpreted, may sometimes be preferable to one that points in the right direction but lacks a common interpretation.

A clear communicator is a leader whose use of language leads to a common understanding of the message being communicated and the implications of that message. A poor communicator, by contrast, although not necessarily suffering from any speech defect, is unable to generate such a common understanding. Audience members may understand the words she utters, but each forms a different interpretation of their meaning; the errors are those of comprehension as well as diction. Arguably a gift for communication belonged to Andrew Jackson about whom Brand (2005) wrote ". . . his diction was clear and his purpose unmistakable. No one ever listened to a speech or a talk from Andrew Jackson who, when he was done, had the least doubt as to what he was driving at."

Communicative ability need not be solely due to innate oratorical flair, because messages might be transmitted indirectly via interlocutors. For example, a follower might hear a leader's views through a spokesperson, from political correspondents, or via other media sources. When a message is conveyed via multiple media, different followers hear different things and so clarity may be compromised. Different media regimes also provide variance in the clarity of communication. For example, in the United Kingdom, the shift to audio broadcasting of parliamentary debates in 1978, the introduction of televised debates in 1989, and similar changes to the coverage of party conference speeches allowed a wider audience to listen directly to the speeches made by leaders.

To assess the effect of these two key leadership qualities, we analyze a game in which followers want to choose the best action while conforming as closely as possible to the actions chosen by others. They listen and respond to leaders. In equilibrium, the relative influence of a leader and her followers' aggregate performance increase with her sense of direction and clarity of communication; good and successful leadership coincide.

An emphasis upon oratorical ability may seem quaint, belonging more to the world of Cicero than to the modern world of political communication. However, our results reveal that a leader's clarity of communication is relatively more important than her sense of direction: heuristically, a leader who can perfectly communicate an imperfect opinion has more influence than a leader who imperfectly communicates a perfect one. Driving this is the desire for unity: when a leader speaks clearly, followers rally around a commonly understood so-called "party line," even though it may differ from the ideal.

The importance of clarity suggests a further question: how might such clarity endogenously arise? The clarity of a leader's message is affected by whether followers listen to her: if they pay careful attention, then they understand what she has to say. However, paying attention to one leader entails being less attentive to another and so leads to a game in which followers endogenously decide to whom to listen. Indeed, the desire to coordinate means that a follower listens to those leaders who already attract the attention of others. This suggests that an elite subset of clear orators may attract attention and obtain influence by acting as focal points while others are ignored.

Of course, an ambitious leader desires power and influence. She may adapt her rhetorical strategy to attract attention to her views. This drives a wedge between good and successful leadership: a good leader helps followers take the right actions, whereas a successful leader enjoys decisive influence or attracts the biggest audience. The prominence of clear communicators in the elite who attract attention suggests that leaders will speak as clearly as possible. However, a near-perfect communicator delivers the essence of her message in a short period of time. A follower need not linger in her audience; having heard what he needs to hear, the follower moves on to listen to others.

This logic suggests a role for obfuscation: a leader might deliberately choose an opaque form of words, avoid speaking via transparent media, or communicate via interlocutors. Her aim is to hold on to the audience for longer while they digest her message, so dissuading followers from listening to competing leaders. Of course, her optimally chosen clarity will depend upon her own sense of direction and the qualities of competing leaders; for instance, a leader may speak more clearly when she has more to say.

The willingness of a leader to blur a message also depends critically on the importance of acting in accordance with others. When cohesion is important, followers emphasize the adoption of a commonly understood course of action: they pay more attention to clearer speakers, and reacting to this, leaders may communicate more clearly. Of course, clearer communication enhances the informativeness of a leader's message: a twist to our story is that followers who focus on conformity, thus ensuring that all activists are singing from the same hymn sheet, also develop a better understanding of the ideal course of action.

Our focus on rhetorical strategies connects our work to that of Riker (1996), while our emphasis on the (endogenous) clarity of leaders' communication relates to strategic ambiguity, whereby leaders are equivocal on policy in order to broaden their appeal (Shepsle 1970, 1972; Zeckhauser 1969). Equivocation and obfuscation are conceptually different. Although the former has received much attention in the political science literature, our theoretical emphasis on the latter is novel and provides insight into a common malady affecting our leaders: they are often not as clear in their communication as we would like them to be. We show that this distortion need not be caused by a lack of inherent ability: it can arise due to the competitive tension between power-hungry leaders.

COORDINATING PARTY ACTIVISTS

Our study of leadership builds upon a simple game in which players want to coordinate their actions in an uncertain environment. This basic game is strategically equivalent to the "beauty contest" scenario described by Morris and Shin (2002), although, as we explain in later sections, our information structure is richer. The terminology stems from Keynes (1936), who described popular newspaper competitions in which entrants chose the prettiest faces from a set of photographs. The winners were those whose choices were also the most popular. It was foolish for an entrant to follow solely his own opinion of beauty: a winning strategy would anticipate the choices of others because the rules promoted conformity. Keynes (1936, Chapter 12) explained:

> It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest ... we devote our intelligences to anticipating what average opinion expects the average opinion to be.

Keynes (1936) believed that this logic reflected elements of professional investment in the stock market. In recent years economic theorists have built formal models of the beauty contest parable (Angeletos and Pavan 2007; Morris and Shin 2002) and used them to analyze public announcements in monetary economies (Amato, Morris, and Shin 2002, Hellwig 2005), complementary investments (Angeletos and Pavan 2004), and the role of higher order beliefs in asset pricing (Allen, Morris, and Shin 2006).

As we will show, the beauty contest parable can be used to develop important insights into the focal role of leadership. We find it useful to describe our own version of the parable as a vehicle to convey our ideas. Rather than think of contestants choosing the prettiest faces, we can think of members of a political party advocating and campaigning for the best policies. The pressure for unity and conformity within political parties then provides an incentive to back a policy that is likely to be popular, which subsequently leads to the anticipation of the average opinion.

To tell this story more formally, we build a simultaneous-move game played by a unit mass of activist party members indexed by $t \in [0, 1]$.[3] An activist advocates a policy $a_t \in \mathbb{R}$. This might be interpreted as the position he supports at a party conference, or the policy he promotes during an election campaign. Drawing together the actions of all party members, the "party line" is the average policy advocated: $\bar{a} \equiv \int_0^1 a_t \, dt$.

A party activist pursues two objectives. Firstly, he would like to advocate the policy θ that best meets the party's needs. Secondly, he wishes to coordinate with others in his party. That is, a concern for party unity drives him to conform to the party line. We represent these twin concerns via a pair of quadratic loss functions:

$$u_t = \bar{u} - \underbrace{\pi(a_t - \theta)^2}_{\text{(i) concern for policy}} - \underbrace{(1-\pi)(a_t - \bar{a})^2}_{\text{(ii) desire for conformity}}.$$

π indexes an activist's relative concern for choosing the ideal policy compared to maintaining party unity.[4] To ensure that both concerns are present, we assume that $0 < \pi < 1$.

When activists share common knowledge of θ, then it is optimal for them to advocate the same ideal policy $a_t = \theta$. In fact, this is the unique Nash equilibrium of the game.[5] When this is so, there is no tension between

[3] The fundamental results of this paper continue to hold with a finite number of party members.

[4] The loss function $(a_t - \bar{a})^2$ captures the penalty of non-conformity. A different measure of mis-coordination is the aggregate distance $\int_0^1 (a_t - a_{t'})^2 \, dt'$ between an activist's action and the policies advocated by others. As we explain elsewhere Dewan and Myatt (2008) the use of this measure of party disunity does not change activists' behavior, but does introduce a role for the strategic use of rhetoric by benevolent leaders.

[5] For this, we require $\pi > 0$, When $\pi = 0$, activists care only about coordinating, and so there are infinitely many Nash equilibria: it is an equilibrium for all activists to back any arbitrary policy $a \neq \theta$.

the activists' twin objectives. It is tempting to conclude that there is no disagreement between party members.

This, however, is not the case. When θ is unknown, an activist is unsure of the best policy and must use any information at his disposal to form beliefs about θ. Activists do not share the same information because their prior experiences will be varied. Holding different views, they are divided over the best course of action. Of course, their disagreement is not fundamental: if only they were able to discover the true value of θ, then they would all reach agreement. Instead, disagreement stems from differences of opinion rather than of ideology. This source of division is particularly relevant to a political party: the members share common values, but nevertheless there is scope for debate about policy.

This discussion suggests that our theory is only applicable to a world in which all fundamental disagreement is absent. Once again, however, this is not the case: our results hold when party activists fundamentally differ in their policy preferences. In fact, the appropriate interpretation of θ is not that it is the single ideal policy for all members of the party, but rather it is the average ideal policy across the spectrum of party members.

OPTIMAL ADVOCACY

Remaining within our parable of a party's coordination problem, we turn our attention to the optimal advocacy decisions of the party's members. When θ is unknown, an activist is unsure of the best policy. He may also be unsure of the likely actions of others. Given this uncertainty, he maximizes his expected payoff $\mathrm{E}[u_t]$, where the expectation is taken with respect to his beliefs about the true underlying ideal policy θ and the party line $\bar{a}$. This maximization is equivalent to the minimization of $\pi\,\mathrm{E}[(a_t-\theta)^2]+(1-\pi)\,\mathrm{E}[(a_t-\bar{a})^2]$. The appropriate first-order condition yields the uniquely optimal advocacy choice

$$a_t = \pi\,\mathrm{E}[\theta] + (1-\pi)\,\mathrm{E}[\bar{a}],$$

which is a weighted average of the expected ideal policy, from the perspective of the activist, and his understanding of the average policy advocated by the party at large.

An activist's expectations are based on any information available to him: activist t observes n informative signals, which form a collection $\tilde{s}_t \in \mathbb{R}^n$ capturing all information relevant to his play. A signal might represent the activist's own research, the understanding gained from communication with other party members, or the influence of a leader. We postpone our description of the statistical properties of the signals until the next section. However, we do assume that the distribution of signals is symmetric across activists. What this means is that if activists t and t' observe the same signal realization, so that $\tilde{s}_t = \tilde{s}_{t'}$, then they share the same beliefs about the identity of the best policy and about the likely signals of other activists; put succinctly, activists are *ex ante* symmetric.

Whatever the informative signals represent, an activist's advocacy strategy is a mapping from signal realizations to policy choices; formally, $a_t = A_t(\tilde{s}_t) : \mathbb{R}^n \to \mathbb{R}$. A strategy profile might involve the use of different strategies by different players. However, once we seek (Bayesian Nash) equilbria, it is without loss of generality to restrict attention to symmetric strategy profiles, so that every player uses the same advocacy strategy $A(\cdot)$. This is because each individual is neglibly small and so, conditional on observing the same signal realization, two different activists see their world in the same way. Because any best reply is unique (and hence strict), this implies that they behave similarly.

An advocacy strategy yields a Bayesian Nash equilibrium when it specifies an optimal choice for an activist, given his beliefs, and when those beliefs are consistent with the partywide use of the strategy. Given activists use a strategy $A(\cdot)$, an activist's expectation of the party line is $\mathrm{E}[\bar{a}\,|\,\tilde{s}_t] = \mathrm{E}[A(\tilde{s}_{t'})\,|\,\tilde{s}_t]$ for $t' \neq t$. Similarly, his expectation of the ideal policy is $\mathrm{E}[\theta\,|\,\tilde{s}_t]$. Hence, the strategy $A(\cdot)$ forms an equilibrium if and only if

$$A(\tilde{s}_t) = \pi\,\mathrm{E}[\theta\,|\,\tilde{s}_t] + (1-\pi)\,\mathrm{E}[A(\tilde{s}_{t'})\,|\,\tilde{s}_t]. \qquad (\star)$$

Thus, an activist's strategy is a weighted average of his expectation of the ideal policy and his understanding of the average policy advocated by party members.[6] To obtain a solution to Equation ($\star$), we need to specify fully how signals help an activist form beliefs about the ideal policy θ and beliefs about the signals seen by other activists. To do this, we turn our attention to the details of the mechanism via which activists learn.

LEARNING FROM LEADERS

Leaders help a follower develop his beliefs about the world and about the likely moves of others. Once again, rather than explain how followers learn from leaders in an abstract setting, we describe our model within the context of the party activist story. Activists begin with no substantive knowledge of the ideal policy: they share a diffuse prior over θ.[7] They learn by listening to n party leaders indexed by $i \in \{1,\ldots,n\}$. The term "leader" can be viewed as a label for an informative signal; indeed, the sources of information available might extend beyond leaders and include the activists' own research and intraparty communication. Nevertheless, this personification is useful for exposition and is relevant when we subsequently introduce a role for strategic leaders.

Each leader forms an independent, unbiased, and private opinion of the ideal policy for the party.

[6] In our ongoing research (Dewan and Myatt 2008) we investigate the impact of individual policy biases on activists' and leaders' behaviors; we present a brief summary of one result here. Concretely, suppose that activist t has a (known to him) policy bias b_t so that his ideal policy is $\theta + b_t$ and so his payoff satisfies $u_t = \bar{u} - \pi(a_t - \theta - b_t)^2 - (1-\pi)(a_t - \bar{a})^2$. If the distribution of policy biases is independent of activists' information about θ then all of our results hold unchanged: in the analysis that follows we need only change an individuals optimal advocacy choice from $a_t = A(\tilde{s}_t)$ to $a_t = A(\tilde{s}_t) + \pi b_t$.

[7] It is straightforward to extend our analysis to a world in which activists share a common prior $\theta \sim N(\mu, \xi^2)$.

Formally, leader i observes an informative signal s_i satisfying

$$s_i \mid \theta \sim N\big(\theta, \kappa_i^2\big) \quad \text{and so} \quad \frac{1}{\kappa_i^2} = \textit{Sense of Direction},$$

where, conditional on θ, leaders' signals are statistically independent. These n underlying informative signals can be collected together to form the $n \times 1$ vector s of information available to the entire set of leaders. The variance term $\kappa_i^2 > 0$ captures an important skill: a leader's ability to discern the correct state of the world. When κ_i^2 is small, she is better able to assess policy (she is a skilled technocrat), and so the precision $1/\kappa_i^2$ indexes her sense of direction. A sharper sense of direction stems from the quality of the leader's judgement and also the quality of the information and advice that is available.

Our leaders address the mass of activists. It is perhaps easiest to think of each leader as speaking directly to her party's membership; for expositional ease, we will adopt this interpretation throughout most of our analysis. However, a "speech" is, more generally, a broad label for the channels of communication open to a leader.

A leader's speech conveys information about her opinion. We assume that her preferences over policy choices match those of party members, and so she describes the world as she sees it; strategic information transmission is ruled out. Alas, she is unable to communicate perfectly: each activist t observes the leader's signal plus noise. Formally,

$$\tilde{s}_{it} \mid s_i \sim N\big(s_i, \sigma_i^2\big) \quad \text{and so}$$

$$\frac{1}{\sigma_i^2} = \textit{Clarity of Communication}.$$

Conditional on s, the final signals received by activists (that is, how they interpret a speech) are statistically independent across activists and leaders. The variance $\sigma_i^2 > 0$ reflects a second important leadership skill: a leader's ability to communicate clearly. The precision $1/\sigma_i^2$ indexes her ability to express coherently her privately held opinions in a public forum. Because her clarity is imperfect, activists do not necessarily hear what the leader is trying to say and obtain different impressions of the leader's views.

The noise in speeches may seem comical. It is tempting to think of a leader suffering from some kind of speech impediment: she speaks to individual activists in turn, each time attempting to say the same thing, but stumbling over the words. This, however, is not our interpretation. Perhaps somewhat subtly, noisy speeches are consistent with words that are spoken perfectly but are interpreted in different ways: errors stem not from concerns about what is said but rather about what is meant. Thus, σ_i^2 will be large when a leader uses words to which the activists who follow her attach varying definitions.

Despite this interpretation, it might seem extreme to impose limits on the clarity of the communication process; presumably, a leader can choose unambiguous words and say precisely what she means, so that $\sigma_i^2 = 0$. We offer some justifications for setting $\sigma_i^2 > 0$.

Firstly, a leader's clarity is restricted whenever she is unable to speak directly and simultaneously to activists. Indeed, the ability to speak directly to an entire mass is a relatively recent phenomenon. Prior to radio and television, a leader would need to travel far and wide to reach an audience, her ability to communicate hindered by the physical demands of such a grueling schedule. As a substitute for direct communication, a leader might rely on interlocutors to provide accounts of her words. For example, in the United Kingdom, the Lobby serves as the gateway to the House of Commons. Ministerial briefings are delivered to lobby journalists. Although lobby members observe the same events, their account of proceedings will differ; imprecise reporting compromises clarity.

Secondly, even when a leader can communicate directly, comprehension may be limited. Given time, an articulate speaker may get her point across by removing, clause by clause, any vagueness. But such a legalistic approach can prove the antithesis of clarity—the language becomes long and tortuous, and the task of absorbing the message in its entirety moves beyond most listeners. The limits to the attention span of a listener may then impose a bound on the overall clarity of the message received. Such limits might be endogenously chosen by the listener: even if there is enough time for him to absorb a leader's words, doing so has an opportunity cost because it prevents him from paying careful attention to other leaders.

Thirdly, there may be endogenous limits to a leader's clarity. Even if a leader is able to speak succinctly using unambiguous words, she may choose not to do so. Indeed, in later sections, we argue that attention-seeking leaders will intentionally obfuscate. Rhetorical strategies that eschew clarity have been studied elsewhere, notably by Zeckhauser (1969) and by Shepsle (1972), who both considered the strategic advantage realized by a candidate who is vague over the policy she would implement if elected. Shepsle (1972), in particular, highlighted the "politician's advantage in speaking half-truths and in varying his appeals with variations in audience and political climate."[8]

Finally, whether a leader's views are less than perfectly clear is an empirical question. The evidence suggests that different people hold different opinions about the views of their leaders. For example, Alvarez and Franklin (1994) described survey respondents who were far less certain of their senator's political views than of their own. Moreover, in his earlier study of campaign effects in senatorial elections, Franklin (1991) noted that "candidate behavior has a substantial impact on the clarity of citizen perceptions."

[8] The consideration of equivocation relates our work to that of Aragones and Neeman (2000), who studied equilibrium levels of candidates' ambiguity. Our work contrasts with that of Mierowitz (2005). He considered politicians who learn about voters' preferences during primary elections. Candidates who refrain from committing to specific policies have greater flexibility upon receiving information from the primary.

Taken together, the variance parameters κ_i^2 and σ_i^2 are inversely related to the two qualities of leadership—sense of direction and clarity of communication—which are central to our article. Consider a politician who arguably embodied both qualities. As a backbencher from 1936 to 1940, Winston Churchill advocated preparation for war while Prime Ministers Baldwin and Chamberlain vacillated in the light of uncertain German military ambitions. History suggests that Churchill had a sharp sense of direction: he identified the threat and a military strategy to deal with it. A further Churchillian skill was communication; his speeches created a common understanding of the perils faced by the allies.

Rarely, however, does an individual embody both characteristics, and in such abundance, that she trumps all rivals. More usually, different leaders (or potential leaders) vary across these dimensions. A contemporary example involves the former British Prime Minister Tony Blair and his successor Gordon Brown. Although Brown is perceived as being among the most intellectually astute of his cohort (low $\kappa_{i'}^2$), he is sometimes regarded as a poor communicator (high σ_i^2). By contrast, although Blair's judgement was called into question (not least over his handling of the second Iraq war), he was widely perceived as one of the best communicators in the business. Blair's strength lay in the articulation of a coherent central message; it might be argued that he combined a lower σ_i^2 with a higher κ_i^2.

Before proceeding to analyze the response of activists to leaders, we pause to describe the relationship between our specification and that of Morris and Shin (2002). In the Morris–Shin world, agents learn via two information sources: an imperfect public signal that is commonly observed and interpreted, and private signals that are independently and identically distributed across agents. Interpreted in our framework, their public signal is a leader with perfect clarity of communication ($\sigma_i^2 = 0$) but an imperfect sense of direction ($\kappa_i^2 > 0$). Conversely, their private signal is a leader with a perfect sense of direction ($\kappa_i^2 = 0$) but imperfect clarity of communication ($\sigma_i^2 > 0$).[9] We analyze leaders (or, equivalently, information sources) with an arbitrary mix of these different attributes, and so our information structure is significantly richer than the Morris–Shin environment, because it allows for positive but imperfect correlation between the signals observed by activists.

FOLLOWING THE LEADERS

We now ask how activists react to the speeches they hear. When there is only one leader, an activist can do no better than to follow the advice given in her speech; this advice yields an unbiased estimate of the ideal policy and the party line, as long as others behave in the same way. However, when there is no clear leader apparent, different leaders must be assessed according to their competencies. This assessment is captured by an equilibrium policy advocacy strategy $A(\cdot)$ satisfying Equation ($\star$).

In principle, an equilibrium strategy could take a complicated functional form. Fortunately, however, we are able to focus our attention on a simple, robust, and easily interpreted class of strategies. Activists employ a linear strategy if

$$A(\tilde{s}_t) = \sum\nolimits_{i=1}^{n} w_i \tilde{s}_{it}$$

where w_i is a coefficient attached to the speech of the ith leader; this provides a convenient measure of this leader's influence on the actions of the mass.

Our focus on linear advocacy strategies stems from the use of the normal distribution in the specification of our model. Normality ensures that the conditional expectations of the ideal policy $\mathrm{E}[\theta \mid \tilde{s}_t]$, of other activists' signals $\mathrm{E}[\tilde{s}_{t'} \mid \tilde{s}_t]$, and of the leaders' underlying signals $\mathrm{E}[s \mid \tilde{s}_t]$ are all linear in $\tilde{s}_t$. If another activist uses a linear strategy, then the conditional expectation $\mathrm{E}[A(\tilde{s}_{t'}) \mid \tilde{s}_t]$ of his action is also linear in $\tilde{s}_t$. This implies that if all other activists use a linear strategy, then a best reply is to use a linear strategy. (That is, the class of symmetric and linear advocacy strategies is closed under best reply.) Pushing further, we obtain the following lemma (formal proofs are contained in the technical appendix.)

LEMMA 1. *There is a unique Bayesian Nash equilibrium involving the use of linear strategies, so that $A(\tilde{s}_t) = \sum_{i=1}^{n} w_i \tilde{s}_{it}$. This equilibrium satisfies $w_i > 0$ for all i and $\sum_{i=1}^{n} w_i = 1$.*

When the unique linear equilibrium is played, an activist advocates a weighted average of the policy recommendations that he hears. The weight w_i placed on a speech acts as an index of the orator's effectiveness; it can measure the success of a leader.

Of course, the possibility of nonlinear equilibria remains. Nevertheless, in the technical appendix, we explain how a mild further restriction on advocacy strategies rules out nonlinear equilibria. Furthermore, even if nonlinear equilibria exist (we conjecture that they do not), then further criteria suggest the selection of the unique linear equilibrium.[10]

Focusing on the partywide deployment of a linear advocacy strategy, we must find the weights placed on the leaders' speeches. One "brute force" approach (described in the appendix) would be to compute the conditional expectations $\mathrm{E}[\theta \mid \tilde{s}_t]$ and $\mathrm{E}[\tilde{s}_{t'} \mid \tilde{s}_t]$, and then proceed to solve Equation ($\star$). Here, however, we take a more subtle approach. We observe that the weights used by activists in equilibrium are precisely those that

[9] We assume that $\kappa_i^2 > 0$ and $\sigma_i^2 > 0$. However, Propositions 1 and 2 also hold when either $\kappa_i^2 = 0$ or $\sigma_i^2 = 0$.

[10] We have noted that we lean heavily upon the work of Morris and Shin (2002), and indeed, their model is obtained by setting $n = 2$, $\sigma_1^2 = 0$, and $\kappa_2^2 = 0$. They claimed that nonlinear equilibria do not exist. However, there is a small chink in the proof they used. We explain further in our technical appendix; the problem we highlight was also noted in a recent paper by Angeletos and Pavan (2007, 1112).

are desirable from the perspective of the entire party: they efficiently maximize party welfare.

LEMMA 2. *In the unique linear equilibrium, the weights placed on the speeches of the various leaders maximize the aggregate welfare of the party: the equilibrium is efficient.*

To verify this claim, consider a change in a_t. This change imposes externalities on others: an increase in a_t pushes up the party line $\bar{a}$, and so activist $t' \neq t$ enjoys a positive spillover if $a_{t'} > \bar{a}$ (his action is now closer to the party line) but suffers if $a_{t'} < \bar{a}$. However, looking across the entire party membership, these externalities sum to zero. This is because the party line $\bar{a}$ is the average policy advocated by the party, and so, in expectation, the policies advocated by individuals lie equally above and below $\bar{a}$. Because the various externalities from an activist's action cancel out, the total partywide effect of a marginal change in the policy he advocates is reflected in his own payoff. This means that he faces socially correct incentives (from the perspective of his party) at the margin.

We now calculate the weights that maximize welfare and so find the unique linear equilibrium. The party's welfare (the *ex ante* expected payoff of a party member) satisfies $\mathrm{E}[u_t] = \bar{u} - \pi\,\mathrm{E}[(a_t-\theta)^2] - (1-\pi)\,\mathrm{E}[(a_t-\bar{a})^2]$. Taking the first quadratic loss term, a_t is a weighted average of unbiased signals of θ, and so $\mathrm{E}[(a_t-\theta)^2] = \mathrm{var}[a_t \mid \theta]$. Turning to the second quadratic loss term, a_t is equal to $\bar{a}$ on average, and so $\mathrm{E}[(a_t-\bar{a})^2] = \mathrm{var}[a_t \mid s]$, where s is the vector of signals seen by the leaders. Assembling these elements,

$$\begin{aligned}\textit{Party Welfare} &\\ &= \bar{u} - \pi \underbrace{\sum_{i=1}^{n} w_i^2\left(\kappa_i^2+\sigma_i^2\right)}_{\text{(i) } \mathrm{var}[a_t \mid \theta]} - (1-\pi)\underbrace{\sum_{i=1}^{n} w_i^2\sigma_i^2}_{\text{(ii) } \mathrm{var}[a_t \mid s]}\\ &= \bar{u} - \sum_{i=1}^{n} w_i^2\left[\pi\kappa_i^2+\sigma_i^2\right].\end{aligned}$$

Notice that any noise in the information sources available to activists detracts from party welfare. Interestingly, the effective noise $\pi\kappa_i^2+\sigma_i^2$ corresponding to a leader's speech does not equally incorporate the variances κ_i^2 and σ_i^2. A lack of clarity in a leader's communication frustrates activists' coordination and lessens the information content of her speech. By contrast, a failing in her sense of direction, while affecting an activist's ability to infer the ideal policy, has no impact on coordination; it thus attracts a reduced coefficient of $\pi < 1$. These observations regarding the different effects of our leadership skills are reflected in the equilibrium weights that maximize party welfare.

Proposition 1. *The unique linear Bayesian Nash equilibrium advocacy strategy satisfies*

$$w_i = \frac{\hat{\psi}_i}{\sum_{j=1}^{n}\hat{\psi}_j}, \quad \textit{where} \quad \hat{\psi}_i = \frac{1}{\pi\kappa_i^2+\sigma_i^2}.$$

Party welfare is $\bar{u} - 1/[\sum_{i=1}^{n}\hat{\psi}_i]$. A leader's influence, indexed by $\hat{\psi}_i$, increases with both her sense of direction and her clarity of communication. The relative influence of relatively better communicators increases as activists' concern for party unity grows: if $(\sigma_i^2/\sigma_j^2) < (\kappa_i^2/\kappa_j^2)$, so that i enjoys a comparative strength in communication, then $\hat{\psi}_i/\hat{\psi}_j$ is decreasing in π.

The index $\hat{\psi}_i$ measures influence and therefore successful leadership; however, because welfare increases with $\sum_{i=1}^{n}\hat{\psi}_i$, it also measures good leadership. An influential leader (with a high value of $\hat{\psi}_i$) clearly communicates her sharp sense of direction. Unsurprisingly, a leader who excels on both dimensions enjoys more influence. Nevertheless, and because $w_i > 0$ for all i, even unskilled leaders enjoy some influence.

So which skill is more important? The presence of π in the index $\hat{\psi}_i$ suggests that a leader's ability to express her views clearly is more important than her ability to understand the political environment. Of course, this claim relies on an implicit assumption that it is appropriate to compare directly the variances κ_i^2 and σ_i^2. Even if this comparison is inappropriate, Proposition 1 reveals that the relative influence of relatively clear communicators grows as cohesion looms larger in the minds of activists. Furthermore, an inspection of $\hat{\psi}_i$ confirms that influence is entirely determined by absolute clarity (rather than relative clarity) as π vanishes to zero: if $\sigma_i^2 < \sigma_j^2$ then $w_i > w_j$ when π is sufficiently small.

To obtain further insight, recall once again that a leader helps activists learn about policy and coordinate. Her message about policy is muddled by two sources of noise: any errors of judgement on her part (the variance κ_i^2) plus any misunderstanding of what she says (the variance σ_i^2). Combining these sources of noise,

$$\tilde{s}_{it} \mid \theta \sim N\left(\theta, \kappa_i^2+\sigma_i^2\right) \quad \text{so that}$$

$$\psi_i \equiv \frac{1}{\kappa_i^2+\sigma_i^2} = \textit{Quality of Information}.$$

If activists cared only about discovering the best policy (so that $\pi = 1$), then the two components of a leader's skill set would be equally important. Each activist would choose an action $a_t = \mathrm{E}[\theta \mid \tilde{s}_t]$ without reference to others. Indeed, Bayesian updating yields

$$\mathrm{E}[\theta \mid \tilde{s}_t] = \sum_{i=1}^{n}\left(\frac{\psi_i}{\sum_{j=1}^{n}\psi_j}\right)\tilde{s}_{it},$$

so that the weight placed on each leader's speech is proportional to the quality of information ψ_i that the speech contains. However, when $\pi < 1$ activists care about coordination as well as policy, and so a leader's speech can act as a convenient focal point for them. For this to be true, it is useful if different activists tend to hear the same thing.

A measure of the commonality of messages received is the correlation between what is heard by different activists. To calculate this, note that the conditional covariance of two signals related to the same speech is $\operatorname{cov}[\tilde{s}_{it}, \tilde{s}_{it'} \mid \theta] = \kappa_i^2$. The correlation coefficient

$$\rho_i = \frac{\kappa_i^2}{\kappa_i^2 + \sigma_i^2} = \textit{Correlation of Messages}$$

depends on the relative strength of a leader's clarity of communication and sense of direction. When a leader becomes a perfect communicator ($\sigma_i^2 \to 0$), the correlation satisfies $\rho_i \to 1$ and everyone hears the same message; the speech becomes a public signal. However, when a leader becomes a perfect technocrat ($\kappa_i^2 \to 0$), the correlation satisfies $\rho_i \to 0$; the messages received become independent private signals of θ.

Correlation is important for expectations about what others hear. If $\rho_i = 1$ (a leader with perfect clarity), then $\mathrm{E}[\tilde{s}_{it'} \mid \tilde{s}_t] = \tilde{s}_{it}$: an activist knows that others hear what he hears. However, when $\rho_i < 1$, interpretations differ. In forming his beliefs, an activist recognizes that any information about the underlying ideal policy is useful in thinking about what the leader was trying to convey. Bayesian updating yields

$$\mathrm{E}[\tilde{s}_{it'} \mid \tilde{s}_t] = \rho_i \tilde{s}_{it} + (1 - \rho_i)\,\mathrm{E}[\theta \mid \tilde{s}_t].$$

A high correlation coefficient reinforces the party's response to a leader. If activists listen to that leader, then when ρ_i is large others will listen to that leader in order to anticipate the party line. In contrast, when ρ_i is small, they divert attention to others. Drawing these observations together, we can reformulate Proposition 1 in terms of ψ_i and ρ_i.

Proposition 2. *The unique linear Bayesian Nash equilibrium advocacy strategy satisfies*

$$\hat{\psi}_i = \frac{\psi_i}{(1 - \rho_i) + \pi\rho_i},$$

and hence a leader's influence increases with the quality of information she offers to activists and the correlation of the messages that they hear. Comparing two leaders i and j satisfying $\rho_i > \rho_j$ (so that leader i is a more coherent communicator) the influence of i relative to j grows as π falls. If leaders' labels are in strict order of decreasing correlation, so that $\rho_1 > \rho_2 > \cdots > \rho_n$, then there is some k such that w_i is locally decreasing in π for $i < k$ and locally increasing in π for $i > k$.

Because π is the weight placed on any deviation from the ideal policy, the remainder $1 - \pi$ is the desire for party unity. Proposition 2 reveals the determinants of good leadership:

$$\textit{Leadership} = \frac{\textit{Quality of Information}}{1 - [\textit{Correlation of Messages} \times \textit{Desire for Unity}]}.$$

Fixing the quality of information provided, coherent communication determines the effectiveness of leadership, and more so when there is a greater desire for party unity. In fact, it is useful to compare a perfect communicator ($\rho_i \approx 1$, so that $\hat{\psi}_i \approx \psi_i/\pi$) with a perfect technocrat ($\rho_j \approx 0$, so that $\hat{\psi}_j \approx \psi_j$). As π vanishes so that only party unity matters, the perfect communicator becomes far more influential than the perfect technocrat.

LISTENING TO LEADERS

To interpret our model, we might think of party members attending a large party conference where each listens to speeches made from the conference platform. An implicit assumption is that activists form a captive audience. Under this assumption, we concluded that the clearest communicators enjoy relatively more influence.

Of course, speeches convey information only if they are heard. Activists may abstain from listening to a particular speech or may not devote their full attention to it. The clarity of a leader's message depends on the willingness of her audience to listen; however, the decision to listen is endogenous. This is important when being informed is costly as, for example, when activists have limited attention spans and cannot listen to a leader indefinitely. Furthermore, there will be strategic interaction among activists when they decide to whom to listen. In common with Hellwig and Veldkamp (2008), we recognize that there may be coordination motives in information acquisition so that, using their language, if an activist wants to do what others do (coordination of policy advocacy), then he wants to know what others know (coordination of attention).

To analyze these effects, we extend our model. Activists are given a single unit of time (perhaps the duration of a party conference) to allocate to different leaders: activist t spends a proportion x_{it} of his time listening to what leader i has to say. We think of him as observing a sample of (noisy) observations of the leader's views. In this sense, the time spent listening represents the sample size. In the usual way, the sample variance declines with the sample size; equivalently, the precision of the aggregate signal is linearly increasing in x_{it}. This leads us to the specification

$$\tilde{s}_{it} \mid s_i \sim N\left(s_i, \frac{\sigma_i^2}{x_{it}}\right) \quad \text{and so} \quad \frac{x_{it}}{\sigma_i^2} = \textit{Clarity of Message},$$

so that the overall clarity of the message is the product of the leader's clarity of communication and the time spent deciphering what it is that she is trying to convey. A constraint $\sum_{i=1}^{n} x_{it} \leq 1$ captures the limited attention span of an activist: paying close attention to one leader carries an opportunity cost, because less attention is paid to others.

With this extension in hand, we analyze a game in which activists choose both to whom to listen and how to react to the speeches they hear. Specifically, activist t chooses $x_t \in \mathbb{R}_+^n$, satisfying the budget constraint on her time and then, given what she hears, chooses a policy to advocate. Payoffs are as before, and activists use a linear advocacy strategy.

As previously, while an activist imposes externalities on others via his effect on the party line $\bar{a}$, the positive and negative externalities cancel out (Lemma 2). Thus, to find the equilibrium, we can again maximize aggregate party welfare. Any strict equilibrium involves the symmetric choice of attention; hence, we can drop the subscript t so that each activist devotes a fraction of time x_i to leader i. Whereas leader i's clarity of communication is still indexed by $1/\sigma_i^2$, the clarity of the message received from her is now x_i/σ_i^2. Restricting attention to a linear advocacy strategy and exploiting Proposition 1,

Party Welfare

$$= \bar{u} - \frac{1}{\sum_{i=1}^{n} \hat{\psi}_i}, \quad \text{where} \quad \hat{\psi}_i = \frac{1}{\pi\kappa_i^2 + [\sigma_i^2/x_i]},$$

and so the equilibrium x maximizes $\sum_{i=1}^{n} \hat{\psi}_i$ subject to $\sum_{i=1}^{n} x_i \leq 1$. Because welfare is increasing in the attention paid to each leader, activists exhaust the time they have available. However, it may be that $x_i = 0$ for some i: activists may ignore some leaders. Evaluating which leaders receive attention can provide insights into the formation of a natural oligarchy of influential leaders; a necessary condition for a leader to have influence is that activists pay attention to her message. Before performing this evaluation, we label (without loss of generality) the leaders in order of decreasing clarity, so that $\sigma_1^2 \leq \cdots \leq \sigma_n^2$.[11]

Proposition 3. *When leaders' audiences are endogenous, there is a unique Bayesian Nash equilibrium involving the subsequent play of a linear advocacy strategy. Activists listen only to the clearest communicators: ordering leaders by decreasing clarity so that $\sigma_1^2 \leq \cdots \leq \sigma_n^2$, there is a unique $m \in \{1, \ldots, n\}$ such that $x_i > 0$ for $i \leq m$ and $x_i = 0$ for all $i > m$. For $i \leq m$,*

$$x_i = \frac{\sigma_i(K_m - \sigma_i)}{\pi\kappa_i^2}, \quad \textit{where} \quad K_m \equiv \frac{\pi + \sum_{j=1}^{m} \left[\sigma_j^2/\kappa_j^2\right]}{\sum_{j=1}^{m} \left[\sigma_j/\kappa_j^2\right]}.$$

The attention paid to a leader increases with her sense of direction, although not always with her clarity: for $i \leq m$, the attention x_i paid to a leader is locally increasing in her clarity when $\sigma_i > K_m/2$, but locally decreasing when $\sigma_i < K_m/2$. The elite size m minimizes K_m. It increases with activists' relative concern π for policy, but decreases with each leader's sense of direction.

When all leaders share the same communication skills ($\sigma_i = \sigma_j$ for all $i \neq j$), then all leaders enjoy an audience, and the attention paid to each is proportional to her sense of direction. However, when leaders differ in their coherence, richer results emerge: activists gravitate toward the clearest communicators. Correspondingly, once a leader's clarity falls below a threshold (i.e., when $\sigma_i > K_m$) activists will ignore her; such a leader can have no influence. Although intuitively one might think that a good sense of direction would demand attention, our result highlights the importance of getting the message across.[12]

Despite this finding, communicating too clearly can deflect attention toward others: when $\sigma_i < K_m/2$ for $i \leq m$ (the noise in a leader's speech is relatively low), an increase in her clarity reduces the attention paid to her. We return to this issue in due course, when we evalute a leader's incentive either to clarify or to obfuscate when speaking to her party.

Nevertheless, sufficient clarity is a prerequisite for successful leadership. One possibility emerging from Proposition 3 is that $m = 1$, so that activists listen only to the best communicator. Such a leader, should she exist, enjoys undivided attention and thus undiluted influence; she becomes a *de facto* dictator. But when will such a leader emerge?

Proposition 4. *Recall that we have (without loss of generality) ordered the leaders by decreasing clarity, so that $\sigma_1^2 \leq \cdots \leq \sigma_n^2$. The clearest communicator is a* de facto *dictator if and only if*

$$\sigma_2^2 \geq \sigma_1^2 \times \left[1 + \frac{\pi\kappa_1^2}{\sigma_1^2}\right]^2.$$

This fails when σ_1^2 is sufficiently small. Hence, for the clearest communicator to enjoy exclusive attention, she needs to communicate imperfectly. The clarity $\sigma_1^2 = \pi\kappa_1^2$ that best supports her dictatorship (by minimizing the right-hand side of the inequality) increases with her sense of direction and the desire for unity. If $\sigma_1^2 < \sigma_2^2$, then she enjoys exclusive attention if π is small enough.

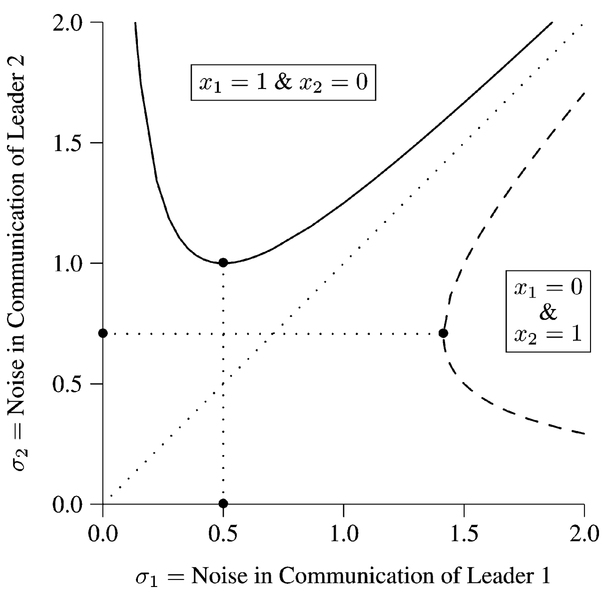
A *de facto* dictator must be the clearest communicator (Proposition 1). Her power and influence are maximized only when the attention of her followers is not diverted to others. For the dictator to enjoy this exclusive attention, however, the clarity of her clearest competitor must be sufficiently low; equivalently, σ_2^2 (and σ_i^2 for other leaders $i > 2$) must be large. Being the clearest communicator is not enough; $\sigma_1^2 < \sigma_2^2$ is sufficient for dictatorship in only two cases. The first case is when $\pi \to 0$, so that activists care only about party unity, and the clearest communicator is best able to describe a focal policy around which the party can rally. The second case is when $\kappa_1^2 \to 0$, so that the best communicator also enjoys an excellent sense of direction; she is a Churchillian leader who trumps all others.

A leader succeeds in monopolizing the agenda when the inequality in Proposition 4 is satisfied. This is easiest when σ_1^2 minimizes the right-hand side of the inequality; that is, when $\sigma_1^2 = \pi\kappa_1^2 > 0$. Figure 1 illustrates: with the parameter values shown, when Leader 1 chooses $\sigma_1^2 = 0.25$ (or $\sigma_1 = 0.5$ in the figure) and $\sigma_2^2 \geq 1$, then Leader 2 receives no attention and enjoys no influence. However, if Leader 1 speaks more clearly, then eventually Leader 2 attracts an audience. The lesson is clear: should an ambitious leader want to monopolize the

[11] If $\sigma_i^2 = \sigma_j^2$ for $i \neq j$, then this order is not unique, but Propositions 3 and 4 do hold as long as $\sigma_1^2 \leq \cdots \leq \sigma_n^2$.

[12] Proposition 3 differs from the central result of Hellwig and Veldkamp (2008) because they find multiple equilibria. The reason is that our activists can vary continuously the attention paid to a leader, whereas their model is equivalent to one in which a leader can be listened to for a fixed length of time, or not at all.

FIGURE 1. Paying Attention to a Single Leader



This figure uses the parameter choices: $\pi = \frac{1}{2}$, $\kappa_1^2 = \frac{1}{2}$, and $\kappa_2^2 = 1$. Leader 1 enjoys unreserved attention whenever σ_2 lies above the solid line; simililarly, leader 2 enjoys unreserved attention whenever σ_1 lies to the right of the broken line. The bullets indicate the values $\sigma_i^2 = \pi\kappa_i^2$ for $i \in \{1, 2\}$ that leader i would choose if she wanted to make it as difficult as possible for leader $j \neq i$ to receive any attention.

agenda, and so maintain complete influence, then she would need to avoid perfect clarity; better communication can sometimes divert attention toward others.[13] Moreover, the clarity $1/(\pi\kappa_1^2)$ that maximizes the range of *de facto* dictatorship increases with a leader's sense of direction; if she is to maintain exclusive attention, then a leader can get away with speaking more clearly only when she has more to say.

Propositions 3 and 4 suggest that a leader can sometimes attract attention by speaking less clearly; in fact, imperfect clarity is necessary if a leader is to maintain complete influence as a *de facto* dictator. To understand why this is so, recall that an activist gathers information to develop an understanding of his environment. Listening to leaders helps him to do this, but given time contraints he will not listen to a leader longer than he needs to. When a leader is a good communicator, an activist can discern her position in a short period of time; with time to spare, he moves on to gather more information.

It is also worthwhile noting that the focus of activists' attention depends upon their relative preference for policy versus party unity. From Proposition 3, m declines with the policy concern parameter π. Hence, as the desire for unity grows (so that π falls), the size of the elite shrinks; Proposition 4 reveals that a *de facto* dictator emerges when π is small enough. The intuition is natural: when activists care only about unity, then the information regarding policy provided by leaders is irrelevant, and all that matters is finding a clear focal point around which the membership can coalesce.

OBFUSCATION

So far, the characteristics of leaders have been exogenous. This seems reasonable for a leader's sense of direction, which, at the time of speaking to her followers, is likely to be beyond her control. However, a leader's clarity of communication may be more manipulable: the overall clarity of a message depends endogenously on both speaker and audience. For instance, a leader might be able to reduce her clarity; she may obfuscate.

Here, we extend our model by allowing for endogenous obfuscation. Of course, if a leader chooses her clarity, then we must consider her objectives. For instance, a benevolent party leader may want to maximize the welfare of her party's membership. Because activists make efficient decisions when choosing to whom to listen and whom to follow, the leader has no reason to distort her behavior. Indeed, she can best enhance her welfare by maximizing the information available, and she does so by speaking as clearly as possible.

Of course, the objectives of an ambitious leader may extend beyond benevolence, and therefore, there may be a variety of reasons for her to distort activists' decisions; obfuscation is one way to do this. For instance, a power-hungry leader may want to maximize her influence. We have already noted that such influence is maximized when a leader enjoys a position as a *de facto* dictator, and we have seen (Proposition 4) that this requires less than perfect clarity. The source of this requirement is the fact that a reduction in clarity can increase the attention devoted to a leader. Indeed, a leader must attract attention because it is a necessary component of exercising influence. To explore fully the desire for attention, we focus on this dimension of a leader's ambition and ask: what rhetorical strategy maximizes the attention paid to her? (Later in this article, we will return to consider other possible objectives and show that they too result in obfuscatory rhetoric.)

To answer this question, we study a simultaneous-move game in which each leader i chooses the variance σ_i^2 of the noise in her speech; her clarity of communication is the precision $1/\sigma_i^2$. We impose exogenous restrictions on the clarity of our n leaders: their choices must satisfy $\sigma_i^2 \geq \underline{\sigma}_i^2 > 0$. Under this specification, $1/\underline{\sigma}_i^2$ is an upper bound to a leader's clarity, which may represent her skill as an orator. Without loss of generality, we order leaders according to their communication skills so that $\underline{\sigma}_1^2 \leq \cdots \leq \underline{\sigma}_n^2$.

A leader's payoff is the attention x_i emerging endogenously from the choices of activists (Proposition 3). Of course, if her communication skills are poor (so that $\underline{\sigma}_i^2$ is large), then she may be

[13] Strictly speaking, there is no perfect clarity because $\sigma_i^2 > 0$ for all i. However, what we mean is that speaking with near-perfect clarity (allowing σ_1^2 to shrink toward zero) eventually deflects attention to other leaders.

ignored. Such a leader is indifferent between her feasible clarities, and so her behavior is not uniquely defined. To simplify our exposition, we assume (without loss of generality) that such a leader speaks as clearly as she can. Similarly, there may be some range of σ_i^2 a leader dictates (Proposition 4). This dictator must be the clearest communicator, and so we label her as $i = 1$. For such a leader, we assume (again without loss of generality) that she chooses her clarity in order to maximize the range of other leaders' clarities for which $x_1 = 1$, thereby defending her dictatorial position. Following Proposition 4, this means that she chooses $\sigma_1^2 = \pi\kappa_1^2$, as long as this is feasible, and otherwise will choose $\sigma_1^2 = \underline{\sigma}_1^2$. Combining these cases, if $x_1 = 1$, then we assume that $\sigma_1^2 = \max\{\pi\kappa_1^2, \underline{\sigma}_1^2\}$.

An attention-seeking leader must convey some information and cannot simply babble; if the noise in her speech is too large (from Proposition 3, when $\sigma_i > K_m$), she will be ignored. But this does not imply that she wants to speak with perfect clarity. When $\sigma_i < K_m/2$, the attention paid to the leader increases with noise added to her speech, and so she obfuscates.[14] In fact, we can obtain an upper bound to her optimally chosen clarity.

LEMMA 3. *If a leader's communication skills are strong, so that $\underline{\sigma}_i^2$ is sufficiently small, then she obfuscates by choosing $\sigma_i^2 > \underline{\sigma}_i^2$. Her optimally chosen clarity always satisfies $\sigma_i^2 \geq \tilde{\sigma}^2$, where*

$$\frac{1}{\tilde{\sigma}^2} \equiv \frac{1}{\pi}\sum_{i=1}^{n}\frac{1}{\kappa_i^2},$$

and so a sufficient condition for a leader to obfuscate is $\underline{\sigma}_i^2 < \tilde{\sigma}^2$.

Further insight emerges from an economic analogy. In our model, each activist is a consumer of costly information. He allocates time (rather than money) to n competing information products (leadership speeches). His purchase from leader i is the clarity x_i/σ_i^2 of her message. When a leader chooses her clarity to maximize the attention paid to her, she acts as would a revenue-maximizing oligopolist. Adding noise to her communication is equivalent to a price hike: it directly increases her revenue (in the form of the attention paid to him or her) for a given quantity (clarity of message); however, obfuscation prompts an activist to lower her demand for her product (speech) by substituting to others. Balancing the two effects of a change in clarity generates an intermediate solution.

Lemma 3 reveals that leaders will sometimes obfuscate. To move further, however, we find the unique equilibrium of our attention-seeking game.

Proposition 5. *There is a unique pure strategy Nash equilibrium of the attention-seeking game. Ordering leaders so that $\underline{\sigma}_1^2 \leq \cdots \leq \underline{\sigma}_n^2$, their clarities satisfy $\sigma_1^2 \leq \cdots \leq \sigma_n^2$. There is a unique $\tilde{m} \in \{0, 1, \ldots, n\}$ such that leaders $i > \tilde{m}$ speak with maximum clarity ($\sigma_i^2 = \underline{\sigma}_i^2$), whereas leaders $i \leq \tilde{m}$ obfuscate ($\sigma_i^2 > \underline{\sigma}_i^2$). Those who obfuscate choose the same clarity: if $i < j \leq \tilde{m}$, then $\sigma_i^2 = \sigma_j^2$.*

If $\underline{\sigma}_i^2 < \tilde{\sigma}^2$ for all i (so that all leaders have good communication skills), then $\tilde{m} = n$ (they all obfuscate), and the unique Nash equilibrium is symmetric, satisfying $\sigma_i^2 = \tilde{\sigma}^2$ for all i. Hence,

Clarity of Communication

$$= \frac{1}{\tilde{\sigma}^2} = \frac{1}{\pi}\sum_{i=1}^{n}\frac{1}{\kappa_i^2} = \frac{\textit{Aggregate Sense of Direction}}{1 - [\textit{Desire for Unity}]}.$$

An increase in any leader's sense of direction and activists' desire for unity: (1) (weakly) increases every leader's clarity of communication, (2) (weakly) reduces the number $\tilde{m}$ of leaders who obfuscate, and (3) (weakly) reduces the size m of the elite who attract an audience.

Linking Propositions 3 and 5, there are (potentially) three groups of leaders. Firstly, those leaders $i \leq \tilde{m} \leq m$ have excellent communication skills and yet do not exploit them; with oratorical flair in abundance, they nevertheless attract maximum attention by obfuscating. (In fact, their choices all satisfy $\sigma_i = K_m/2$.) Because these $\tilde{m}$ leaders choose the same clarity, the relative influence of each individual is determined by her sense of direction. By contrast, leaders in the second group $\tilde{m} < i \leq m$ have less developed oratorical skills and must strain to be clear to the best of their abilities in order to attract an audience. The exogenous limits to their oratorical skills mean that they do not obfuscate, but they are heard so long as $\underline{\sigma}_i < K_m$. A leader $i > m$ talks only to herself. To enjoy an audience, she must improve her oratorical skills: only when the noise in her speech is no greater than twice that of the obfuscating (and hence clearest) communicators will her views be heard.

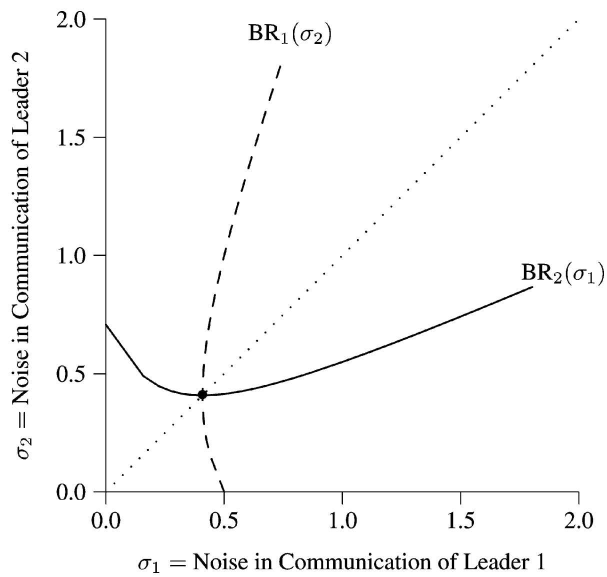
Of course, the three different classes of leader may collapse to a single group. For instance, when $\underline{\sigma}_n^2 < \tilde{\sigma}^2$, every leader will obfuscate. This will be so when there are few exogenous limits to clarity, perhaps a world in which a leader is able to deliver a commonly heard speech directly to the entire activist mass. Perhaps surprisingly, the unique equilibrium identified by Proposition 5 (illustrated in Figure 2) reveals that all leaders speak with the same clarity, even though they do not share a common sense of direction.

To understand why, consider the solution for x_i from Proposition 3. The attention (relative to others) paid to each leader increases proportionally with her sense of direction. However, the way in which it reacts to clarity is (approximately) the same for everyone via the term $\sigma_i(K_m - \sigma_i)$.[15] Returning to our analogy, local to the

[14] We expect an influence seeker to choose greater clarity than an attention seeker: fixing the attention paid to her, any increase in clarity increases the weight placed on a leader's speech and hence her influence.

[15] K_m depends on σ_i in a different way for each leader. However, what matters is the shape of $\sigma_i(K_m - \sigma_i)$ local to $K_m/2$. As the proof of Proposition 1 shows, $\partial K_m/\partial\sigma_i = 0$ when evaluated at $\sigma_i = K_m/2$.

FIGURE 2. Endogenous Leadership Clarity



This figure uses the parameter choices: $\pi = \frac{1}{2}$, $\kappa_1^2 = \frac{1}{2}$, and $\kappa_2^2 = 1$. It illustrates the two reaction (or best reply) functions for two attention-seeking leaders. The unique intersection corresponds to the unique Nash equilibrium; it lies on the 45-degree line, and hence, leaders speak with equal clarity in equilibrium.

equilibrium the demand curves (for clarity of message) faced by different leaders are the same shape; however, those with a better sense of direction benefit from proportionally higher demand for any given price. Because $1/\kappa_i^2$ simply scales the demand curve along a quantity axis, the revenue-maximizing price is independent of it. This leads naturally to a symmetric equilibrium (Figure 2).

As suggested previously, communication skill may be due to a combination of natural ability and the information technology available to all leaders. When such technology is poor, we would expect $\underline{\sigma}_i^2$ to be large for everyone. If $\underline{\sigma}_n^2 > \tilde{\sigma}^2$, then leaders speak with different clarities, and so attention (and influence) is biased away from those with poor communication skills. When information technology improves, $\underline{\sigma}_i^2$ may fall for everyone. If $\underline{\sigma}_n^2 < \tilde{\sigma}^2$, then we shift to an equilibrium in which everyone speaks with equal clarity, and so the sole determinant of attention and influence is a leader's sense of direction.

For example, when leaders are constrained to use primitive media (e.g., a series of private audiences), only the most oratorically skilled leaders, who can speak clearly and concisely, attract attention. As technology improves so that the same leadership speech can simultaneously be broadcast to all activists, then exceptional oratorical flair is no longer a prerequisite for success. Thus, while a leader in the mold of Theodore Roosevelt—combining physical stamina, oratorical flair, and the ability to deliver an effective stump speech to many different audiences—enjoys influence despite technology constraints, others, not so well endowed, can flourish only as technology advances. For example, it is well documented that Calvin Coolidge lacked the capacity to succeed on the stump. He owed much to broadcasting; in his own words [Cornwell (1957)],

> I am very fortunate that I came in with the radio. I can't make an engaging, rousing, or oratorical speech to a crowd ... but I have a good radio voice, and now I can get my messages across to them without acquainting them with my lack of oratorical ability.

Of course, if technology allows all leaders to communicate clearly, then all will obfuscate. This rhetorical behavior has (perhaps superficial) similarities with the "garbling" of messages in sender–receiver games analyzed by Crawford and Sobel (1982) and extended to political settings by Gilligan and Krehbiel (1987) and Li, Rosen, and Suen (2001), among others. In these "cheap talk" scenarios, an informed politician is restricted to sending garbled messages due to her commonly understood policy bias. Our leaders have no inherent policy bias, but the strategic incentive to obfuscate arises nevertheless.

Although a leader may obfuscate, her payoff does not depend on the policies advocated by activists; she has no direct incentive to misrepresent the truth as she sees it. Her only strategic move is to change the precision of the information she transmits. Thus, our leaders' speeches are signals in the statistical, rather than the game theoretic, sense; an activist need not anticipate any bias.[16] Our model thus lacks the strategic tensions that arise when leaders have policy preferences that differ from those of their followers.

Nevertheless, such biases could arise endogenously from attention-seeking behavior. For example, suppose that when θ falls within a specific range of values, information is more easily conveyed than when θ falls outside that range. This might be so when there is a partywide consensus about the correct policy. For example, a policy that forms the status quo might easily be described to an audience, whereas elucidating an alternative path requires a longer exposition. Alternatively, a particular policy (say left or right) may require more (or less) clarification. In such scenarios, there is an incentive for a leader to dissemble, pretending that she observed something she did not, though this need not exclude obfuscation as a rhetorical strategy. Whenever such biases are present, activists face a harder task in extracting information from a leader's speech.

Before concluding this analysis of endogenous clarity, we note that the incentive to obfuscate arises due to the nature of the "beauty contest" played by activists and the competitive tensions between attention-seeking leaders. Beyond our policy advocacy parable, these ingredients combine in other social situations where obfuscation is pervasive, and so our analysis

[16] In contrast, Hermalin (1998) described a theory of "leading by example" in which the costly action of a leader (sender) acts as a textbook signal of her private information to her followers (receivers).

provides more general insights. Consider the following variation: our activists are members of a religious mass who seek guidance from spiritual leaders. They wish that their everyday acts are in accordance with a divine plan unknown to them, but as members of a community must also act in concert with their fellows. Through their readings of sacred texts, meditation and other methods, religious leaders receive signals of these unknown truths and convey messages through speech. Our results suggest that when religious leaders compete for the attention span of the mass, they obfuscate, and so provide an interpretation for why lucid prose rarely emanates from the pulpit.

PARTY PERFORMANCE

When the exogenous limits to communication are not binding, Proposition 5 predicts that the (common) clarity of leaders' communication increases with every leader's sense of direction. An increase in a leader's sense of direction enhances demand for her information as activists divert attention away from others. This has a knock-on effect because it forces other leaders to compete harder, by increasing the clarity of their communication. This feeds back, in turn, to the original leader. One aspect of this effect seems natural: an exogenous increase in a leader's judgement increases the clarity of everyone, and so a leader speaks more clearly when he or she has more to say. More subtly, however, a leader speaks more clearly when others have more to say.[17] This suggests that an exogenous increase in leadership quality (an enhanced sense of direction) can promote a further endogenous increase (clearer communication) and hence better party performance.

Proposition 5 also predicts that leaders' rhetoric reacts to the relative preferences of the party membership. As activists emphasize party unity relative to choosing the right policy, so that π falls, leaders speak with increasing clarity. Coordination requires precise communication, and leaders respond by speaking more clearly.

We observe the full importance of this point when we consider the welfare and policy performance implications of attention-seeking leaders. Recall (Proposition 1) that party welfare increases with $\sum_{i=1}^{n} \hat{\psi}_i$, and so the index $\hat{\psi}_i$ measures both good leadership and successful leadership; it reacts positively to a leader's clarity as well as her sense of direction. Allowing the attention paid to leaders (and hence the overall clarity of their messages) to be determined endogenously, the situation becomes more complex. Clarity of communication remains critical in ensuring that a leader receives some attention. However, whereas increased clarity benefits activists and is thus a component of good leadership, an attempt to seek attention or to monopolize the agenda may induce a leader to reduce her clarity; a successful leader (as opposed to a good leader) may obfuscate. The vanity of attention seekers separates good and successful leadership.

Allowing our leaders to play a game in which they simultaneously choose their rhetorical strategies might be expected to complicate things further. In practice, it simplifies matters. Because (when $\underline{\sigma}_n^2 < \tilde{\sigma}^2$) all leaders choose the same clarity, the attention paid to each leader is proportional to her sense of direction. The leadership index reduces to

$$\hat{\psi}_i = \frac{1}{\pi\kappa_i^2 + \left[\sigma_i^2/x_i\right]} = \frac{1}{2\pi\kappa_i^2},$$

and so party welfare increases with the combined judgement of the leadership, but falls with activists' relative concern for policy. The effect of π is unsurprising, because the policy component of an activist's loss function reacts to two sources of noise rather than one; even fixing the behavior of all actors, our welfare measure will fall with an increase in π.

A more surprising insight is obtained by considering an objective performance index. A party exhibits good policy performance if the policies advocated by its members are close to the ideal policy. An appropriate measure here is the loss function $\int_0^1 (a_t - \theta)^2\, dt = \mathrm{E}[(a_t - \theta)^2]$. The inverse of this provides our measure of policy performance.

Proposition 6. *If $\underline{\sigma}_i^2 < \tilde{\sigma}^2$ for all i, so that all leaders choose the same clarity, then the equilibrium influence of a leader is proportional to her sense of direction. Furthermore,*

$$\textit{Policy Performance} \equiv \frac{1}{\mathrm{E}[(a_t - \theta)^2 \mid \theta]} = \frac{1}{1+\pi} \sum_{i=1}^{n} \frac{1}{\kappa_i^2},$$

which increases with the leaders' combined sense of direction but decreases with the policy concern parameter π; a greater desire for party unity improves the policy performance of the party.

Paradoxically, activists become better at advocating the best policy as they care less about doing so. Recognizing the endogenous quality of leadership provides the correct intuition. When activists desire unity, they seek out a common party line, and so leaders respond by speaking clearly. Clearer communication allows activists to develop a better understanding of their environment. Moving away from a desire to back good policies generates a need for coherent unifying leadership; this reduces the obfuscation of attention seekers and so improves policy performance.

Our analysis of policy performance also suggests a further possible objective for leaders and so an additional source of obfuscation. Consider a leader who cares only about policy performance and not about party unity. From her perspective, followers place too

[17] This does not imply a cross-sectional relationship between clarity of communication and sense of direction. For instance, if $\underline{\sigma}_i^2 \leq \tilde{\sigma}^2$ for all i, then all leaders communicate with equal clarity. However, when leaders speak with equal clarity, activists listen more carefully to those with a better sense of direction, and hence, leaders with good judgement are better able to get their messages across. So, looking across the set of leaders, a leader who has more to say is heard more clearly, even if she does not speak more clearly.

much emphasis on the speeches of relatively clear communicators and hence too little on those with a good sense of direction. If such a leader is clear but has poor judgement, then she may prefer to lose influence to others: obfuscation is one technique for doing so.[18] In fact, it is straightforward to construct reasonable parameter configurations in which this phenomenon arises. More generally, whenever a leader's concern for policy relative to coordination differs from that of her followers, incentives to obfuscate emerge.

CONCLUDING REMARKS

A leader can be influential when political actors want (1) to make informed policy choices and (2) to coordinate with each other. A leader's relative influence depends on her own qualities and those of other leaders: two such skills are her ability to communicate clearly (clarity of communication) and to judge the best policy (sense of direction).

We studied party activists who want to advocate the best policy and also to act in concert with their fellow party members. Adopting the weight activists place on a leader's speech as a measure of her influence, we found that, when leadership skills are exogenous, clarity of communication is the most important leadership attribute. When the attention paid to leaders is endogenous, a natural leadership elite of the clearest communicators emerges. Activists apportion their time among these elite communicators, but may pay more heed to those with (relatively) inferior communicative ability. Correspondingly, when attention-seeking leaders choose their clarity, some may obfuscate to retain their audience: these leaders choose the same levels of clarity; their clarity is increasing in the combined sense of direction of the leadership elite; and so in contrast to our earlier results, a leader's sense of direction becomes relatively more important.

We also asked whether leaders can be both good and successful. A good leader enhances party performance, aiding activists in pursuit of their twin goals to the best of her abilities. A successful leader commands attention. Only when skills are exogenous do good and successful leadership necessarily coincide. Otherwise, a leader may increase her success by obfuscating her message; activists receive less information and consequentially are less informed about their environment. Perhaps surprisingly, there is less obfuscation when activists place more emphasis on following the party line than pursuing the best policy. When a party emphasizes unity, it provides leaders with the necessary incentives so that good and successful leadership coincide and policy performance improves.

Few would deny that leaders play an important role in influencing people's actions, and indeed there is experimental evidence that they do (Güth et al. 2004; Humphreys, Masters, and Sandbu 2006). Yet, there has been no recent formal work that evaluates the influence of different leaders. This article has attempted to fill this gap. Building on suggestions made by Levi (2006, 10) that "leadership—both of government and within civil society—provides the agency that coordinates the efforts of others" we have explored a world with multiple potential leaders differentiated by their skill sets and extended our analysis to a world where leadership skills emerge endogenously. Our work provides, we hope, a small step in response to Levi's (2006, 11) claim that "still lacking is a model of the origins and means of ensuring good leadership."

TECHNICAL APPENDIX

Symmetry of Equilibria. In the text, we claimed that it is without loss of generality to restrict attention to symmetric strategy profiles. To see why, let us suppose that activists use different strategies. Activist t's expectation of the party line $\bar{a}$ is

$$\mathrm{E}[\bar{a}\,|\,\tilde{s}_t] = \mathrm{E}\left[\int_0^1 a_{t'}\,dt' \,\middle|\, \tilde{s}_t\right] = \int_0^1 \mathrm{E}[a_{t'}\,|\,\tilde{s}_t]dt'$$
$$= \int_0^1 \mathrm{E}[A_{t'}(\tilde{s}_{t'})\,|\,\tilde{s}_t]dt'.$$

This expectation depends upon the signal realization $\tilde{s}_t$ but not directly on the activist's player index t. This implies his set of best replies to strategies of others is independent of t. Because the loss function that he minimizes is strictly convex, his best reply is unique. Taking these observations together, we conclude that all activists will best reply with the same advocacy strategy $A(\cdot)$. Hence, $\mathrm{E}[\bar{a}\,|\,\tilde{s}_t] = \int_0^1 \mathrm{E}[A(\tilde{s}_{t'})\,|\,\tilde{s}_t]dt' = \mathrm{E}[A(\tilde{s}_{t'})\,|\,\tilde{s}_t]$. ■

Proof of Lemma 1. A linear strategy takes the form $A(\tilde{s}_t) = w \cdot \tilde{s}_t$ for some $n \times 1$ vector w, where "$\cdot$" is the usual vector product. Conditioning on the $n \times 1$ vector s of underlying signals observed by leaders and integrating across the unit mass of activists, the party line is $\bar{a} = w \cdot s$, and hence, $\mathrm{E}[\bar{a}\,|\,\tilde{s}_t] = w \cdot \mathrm{E}[s\,|\,\tilde{s}_t]$. The best reply $\mathrm{BR}[\tilde{s}_t\,|\,w]$ of activist t, given that he observes a signal vector $\tilde{s}_t$ and others play $A(\tilde{s}_{t'}) = w \cdot \tilde{s}_{t'}$, satisfies

$$\mathrm{BR}[\tilde{s}_t\,|\,w] = \pi\,\mathrm{E}[\theta\,|\,\tilde{s}_t] + (1-\pi)w \cdot \mathrm{E}[s\,|\,\tilde{s}_t].$$

Because an activist begins with a diffuse prior over θ (a nondiffuse prior is easily accommodated by incorporating an additional element into the signal vector $\tilde{s}_t$) and $\tilde{s}_t$ is a normally distributed signal of θ, the conditional expectation of θ satisfies $\mathrm{E}[\theta\,|\,\tilde{s}_t] = b \cdot \tilde{s}_t$ for some $n \times 1$ vector b, and similarly, $\mathrm{E}[s\,|\,\tilde{s}_t] = B\tilde{s}_t$ for some $n \times n$ inference matrix B. Hence,

$$\mathrm{BR}[\tilde{s}_t\,|\,w] = \pi b \cdot \tilde{s}_t + (1-\pi)w \cdot B\tilde{s}_t = \hat{w} \cdot \tilde{s}_t \quad \text{where}$$
$$\hat{w} = \pi b + (1-\pi)B'w.$$

This is a linear strategy, which verifies the first claim. A linear equilibrium corresponds to a vector w, satisfying $w = \pi b + (1-\pi)B'w$. The unique solution is $w = \pi[I - (1-\pi)B']^{-1}b$, where I is the $n \times n$ identity matrix, as long as $I - (1-\pi)B'$ has full rank. To find an explicit solution for w, we need only calculate b and B. Bayesian updating leads

[18] Indeed, it is precisely situations in which influence seekers do not want to obfuscate (so that influence is increasing in clarity) that can induce performance seekers to do so. Thus, it is when one force for obfuscation (the hunger for power) is absent that another arises.

to $b_i = \psi_i / \sum_{j=1}^n \psi_j$, where $\psi_i = 1/[\kappa_i^2 + \sigma_i^2]$, the quality of information term used in Proposition 2. Similarly,

$$B = \begin{bmatrix} \rho_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \rho_n \end{bmatrix} + \frac{1}{\sum_{j=1}^n \psi_j} \begin{bmatrix} (1-\rho_1)\psi_1 & \dots & (1-\rho_1)\psi_n \\ \vdots & \ddots & \vdots \\ (1-\rho_n)\psi_1 & \dots & (1-\rho_n)\psi_n \end{bmatrix},$$

where ρ_i is the correlation coefficient defined in the text. (Details of the Bayesian updating formulae are contained in a further supplementary appendix, available from the authors.) Applying this, we obtain $\mathrm{E}[s_i \mid \tilde{s}_t] = \rho_i \tilde{s}_{it} + (1-\rho_i)\,\mathrm{E}[\theta \mid \tilde{s}_t]$, the expression used prior to Proposition 2. $I - (1-\pi)B'$ has full rank, because it is a rank one update of a diagonal matrix and hence invertible. Thus, the solution for w, and hence the linear equilibrium, is unique. The coefficients must satisfy $\sum_{i=1}^n w_i = 1$ to ensure that a common shift in all signals results in the same shift in activists' actions. (Solving for w explicitly confirms this.) ■

Proof of Lemma 2. The effect of changing $\bar{a}$ is $\partial[\int_0^1 u_{t'}\,dt']/\partial\bar{a} = -2\int_0^1 (a_{t'} - \bar{a})\,dt' = 0$. ■

Proof of Proposition 1. From the proof of Lemma 1, the unique linear equilibrium satisfies $w = \pi[I - (1-\pi)B']^{-1}b$. Rather than calculate this directly (this approach is used in our supplementary appendix), we use Lemma 2: finding the (unique) equilibrium boils down to minimizing $\sum_{i=1}^n w_i^2[\pi\kappa_i^2 + \sigma_i^2]$ subject to $\sum_{i=1}^m w_i = 1$. Introducing the Lagrange multiplier λ, the first-order conditions take the form $2w_i[\pi\kappa_i^2 + \sigma_i^2] = \lambda$ for each i, or equivalently, $w_i = \lambda\hat{\psi}_i/2$; the joint solution yields the proposition's main claim. The welfare measure follows by substitution, and the comparative static claims follow by inspection. ■

Proof of Proposition 2. $\hat{\psi}_i$ follows from simple algebra, and the comparative static claims regarding ψ_i and ρ_i follow by inspection. Taking logarithms and differentiating,

$$\begin{aligned} \frac{\partial \log(\hat{\psi}_i/\hat{\psi}_j)}{\partial\pi} &= \left[\frac{-\rho_i\psi_i}{((1-\rho_i)+\pi\rho_i)^2} \times \frac{(1-\rho_i)+\pi\rho_i}{\psi_i}\right] \\ &\quad - \left[\frac{-\rho_j\psi_j}{((1-\rho_j)+\pi\rho_j)^2} \times \frac{(1-\rho_j)+\pi\rho_j}{\psi_j}\right] \\ &= \frac{\rho_j}{(1-\rho_j)+\pi\rho_j} - \frac{\rho_i}{(1-\rho_i)+\pi\rho_i} \\ &< 0 \quad \Leftrightarrow \quad \rho_i > \rho_j, \end{aligned}$$

which yields the final claim of the proposition. ■

Nonlinear Equilibria. In the text, we characterized the unique linear Bayesian Nash equilibrium of the beauty contest game, but noted that the possibility of nonlinear equilibria remained open. We also suggested that a mild restriction on advocacy strategies eliminates any nonlinear equilibria. Here, we expand upon our claims.

We use the following notation. (i) The subscript t indicates an expectation conditional on the information of activist t, so that $\mathrm{E}_t[\cdot] \equiv \mathrm{E}[\cdot \mid \tilde{s}_t]$. (ii) $\bar{\mathrm{E}}[\cdot] \equiv \int_0^1 \mathrm{E}_t[\cdot]\,dt$ is the average expectation across the mass of activists. (iii) For any positive integer k, we define $\bar{\mathrm{E}}^k[\cdot]$ inductively: $\bar{\mathrm{E}}^1[\cdot] \equiv \bar{\mathrm{E}}[\cdot]$ and $\bar{\mathrm{E}}^{k+1}[\cdot] \equiv \bar{\mathrm{E}}[\bar{\mathrm{E}}^k[\cdot]]$. (iv) $\mathrm{BR}[A(\cdot)] : \mathbb{R}^n \mapsto \mathbb{R}$ is the best reply of activist t, given that all others play $A(\cdot)$. (v) Finally, for any positive integer k, we define $\mathrm{BR}^k[A(\cdot)]$ inductively: $\mathrm{BR}^1[A(\cdot)] \equiv \mathrm{BR}[A(\cdot)]$ and $\mathrm{BR}^{k+1}[A(\cdot)] \equiv \mathrm{BR}[\mathrm{BR}^k[A(\cdot)]]$.

The extra condition that we impose prevents "exploding" higher-order expectations.

Definition. *Fix a policy advocacy strategy $A(\cdot)$ played by party members. Higher-order expectations of the party line are* nonexplosive *if* $\lim_{k\to\infty}[\alpha^k\,\mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]] = 0$ *for any* $\alpha \in (0, 1)$.

α^k vanishes exponentially as $k \to \infty$, and hence, this condition is easily satisfied. For instance, if we fix a linear strategy $w \cdot \tilde{s}_t$ and some $B > 0$ such that $|A(\tilde{s}_t) - w \cdot \tilde{s}_t| < B$ for all $\tilde{s}_t$, then higher-order expectations are nonexplosive.

Proposition. *Fix a policy advocacy strategy $A(\cdot)$ played by party members. If higher-order expectations of the party line are nonexplosive, then $\lim_{k\to\infty} \mathrm{BR}^k[A(\cdot)] = w \cdot \tilde{s}_t$, where $w \cdot \tilde{s}_t$, is the unique linear Bayesian Nash equilibrium described in Proposition 1. Hence, the only equilibrium with nonexplosive higher-order expectations is the unique linear equilibrium.*

The best reply to $A(\cdot)$ is $\mathrm{BR}[A(\cdot)] = \pi\,\mathrm{E}_t[\theta] + (1-\pi)\,\mathrm{E}_t[A(\cdot)]$ (where $\mathrm{E}_t[A(\cdot)] = \mathrm{E}[A(\tilde{s}_{t'}) \mid \tilde{s}_t]$ for $t' \neq t$). If all activists play $\mathrm{BR}[A(\cdot)]$, then the party line is $\bar{a} = \pi\bar{\mathrm{E}}[\theta] + (1-\pi)\bar{\mathrm{E}}[A(\cdot)]$. A new best reply $\mathrm{BR}[\mathrm{BR}[A(\cdot)]] = \mathrm{BR}^2[A(\cdot)]$ to the partywide play of $\mathrm{BR}[A(\cdot)]$ satisfies

$$\begin{aligned} &\mathrm{BR}^2[A(\cdot)] \\ &\quad = \pi\,\mathrm{E}_t[\theta] + (1-\pi)\,\mathrm{E}_t[\bar{a}] \\ &\quad = \pi\,\mathrm{E}_t[\theta] + (1-\pi)\,\mathrm{E}_t\left[\pi\bar{\mathrm{E}}[\theta] + (1-\pi)\bar{\mathrm{E}}[A(\cdot)]\right] \\ &\quad = \pi\left(\mathrm{E}_t[\theta] + (1-\pi)\,\mathrm{E}_t[\bar{\mathrm{E}}[\theta]]\right) + (1-\pi)^2\,\mathrm{E}_t\left[\bar{\mathrm{E}}[A(\cdot)]\right]. \end{aligned}$$

If all activists switch to play $\mathrm{BR}^2[A(\cdot)]$, then $\bar{a} = \pi(\bar{\mathrm{E}}[\theta] + (1-\pi)\bar{\mathrm{E}}^2[\theta]) + (1-\pi)^2\bar{\mathrm{E}}^2[A(\cdot)]$. Continuing inductively, for any positive integer k,

$$\begin{aligned} \mathrm{BR}^k[A(\cdot)] &= \pi\left[\sum_{j=0}^k (1-\pi)^j\,\mathrm{E}_t[\bar{\mathrm{E}}^j[\theta]]\right] \\ &\quad + (1-\pi)^{k+1}\,\mathrm{E}_t\left[\bar{\mathrm{E}}^k[A(\cdot)]\right]. \end{aligned}$$

We first consider $\sum_{j=0}^k (1-\pi)^j\,\mathrm{E}_t[\bar{\mathrm{E}}^j[\theta]]$. In the proof of Lemma 1, we noted that $\mathrm{E}_t[\theta] = b \cdot \tilde{s}_t$ for some $n \times 1$ vector b, and so $\bar{\mathrm{E}}[\theta] = b \cdot s$, where s is the vector of underlying signals observed by the leaders. Furthermore, $\mathrm{E}_t[s \mid \tilde{s}_t] = B\tilde{s}_t$ for the $n \times n$ inference matrix B described in Lemma 1. Combining these observations, $\mathrm{E}_t[\bar{\mathrm{E}}[\theta]] = \mathrm{E}_t[b \cdot s] = b \cdot \mathrm{E}_t[s] = b \cdot (B\tilde{s}_t)$ and so $\bar{E}^2[\theta] = b \cdot (Bs)$. Continuing inductively, $\mathrm{E}_t[\bar{\mathrm{E}}^j[\theta]] = b \cdot (B^j\tilde{s}_t)$. Hence,

$$\begin{aligned} &\lim_{k\to\infty}\left[\pi\sum_{j=0}^k (1-\pi)^j\,\mathrm{E}_t[\bar{\mathrm{E}}^j[\theta]]\right] \\ &\quad = b \cdot \left[\sum_{j=0}^\infty (1-\pi)^j B^j\right]\tilde{s}_t = b \cdot [I - (1-\pi)B]^{-1}\tilde{s}_t = w \cdot \tilde{s}_t, \end{aligned}$$

where $w \cdot \tilde{s}_t$ is the unique linear equilibrium strategy (Proposition 1). If the higher-order expectations generated by $A(\cdot)$ are nonexplosive, then the second component of $\mathrm{BR}^k[A(\cdot)]$ vanishes as $k \to \infty$, and so $\mathrm{BR}^k[A(\cdot)] \to w \cdot \tilde{s}_t$ for each $\tilde{s}_t$.

This argument provides one justification for our focus on linear equilibria: starting from a strategy $A(\cdot)$ with nonexplosive higher-order expectations, a tatonnement in which activists update via sequence of best replies, generating the sequence $\mathrm{BR}^k[A(\cdot)]$, converges to the unique linear equilibrium. Furthermore, because any equilibrium policy advocacy strategy satisfies $\mathrm{BR}^k[A(\cdot)] = A(\cdot)$ for all k, we can be assured that any nonlinear equilibria must generate explosive higher-order expectations of the party line.

The logic used here was employed by Morris and Shin (2002).They claimed to find a unique equilibrium, but did not prove that $(1-\pi)^{k+1}\,\mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]] \to 0$; as Angeletos and Pavan (2007) noted, their analysis was not quite watertight. Indeed, it is possible to find strategies that generate explosive higher-order expectations. Consider, for instance, a world in which $n = 1$, $\kappa_1^2 = 0$, and $\sigma_1^2 = \sigma^2 > 0$. Because there is only one leader, we drop the "i" subscript. If $A(\tilde{s}_t) = e^{\beta\tilde{s}_t}$ for some $\beta > 0$, then $\mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]] = e^{k\beta^2\sigma^2+\beta\tilde{s}_t}$, and so $(1-\pi)^{k+1}\,\mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]]$ diverges if $(1-\pi)e^{\beta^2\sigma^2} > 1$. ■

Before presenting the proof of Proposition 3, it is useful to derive the properties of K_m.

LEMMA 4. *Without loss of generality, order leaders so that $\sigma_1^2 \le \cdots \le \sigma_n^2$ and define*

$$K_i \equiv \frac{\pi + \sum_{i'=1}^{i}\left[\sigma_{i'}^2/\kappa_{i'}^2\right]}{\sum_{i'=1}^{i}\left[\sigma_{i'}/\kappa_{i'}^2\right]}.$$

(i) For all $j \le i$:, K_i is (strictly) increasing in κ_j^2 if and only if $\sigma_j(<) \le K_i$ and K_i is (strictly) increasing in σ_j^2 if and only if $\sigma_j(>) \ge K_i/2$. (ii) For all $i > 1$:, $\sigma_i(<) \le K_i$ if and only if $\sigma_i(<) \le K_{i-1}$. (iii) There exists a unique $m \in \{1,\ldots n\}$ such that $\sigma_i < K_i$ for all $i \le m$ and $\sigma_i \ge K_i$ for all $i > m$. (iv) $K_i > K_{i+1}$ for all $i < m$ and $K_i \le K_{i+1}$ for all $i \ge m$, and hence, $K_m \le K_i$ for all $i \in \{1,\ldots,n\}$. (v) m uniquely satisfies $\sigma_i < K_i \le \sigma_{i+1}$.

Proof. (i) K_i is (strictly) increasing in κ_j^2 if and only if it is (strictly) decreasing in $[1/\kappa_j^2]$:

$$\frac{\partial K_i}{\partial[1/\kappa_j^2]} = \frac{\sigma_j^2}{\sum_{i'=1}^{i}\left[\sigma_{i'}/\kappa_{i'}^2\right]} - \sigma_j\frac{\pi + \sum_{i'=1}^{i}\left[\sigma_{i'}^2/\kappa_{i'}^2\right]}{\left[\sum_{i'=1}^{i}\left[\sigma_{i'}/\kappa_{i'}^2\right]\right]^2}(<)$$
$$\le 0 \quad\Leftrightarrow\quad \sigma_j(<) \le K_i.$$

Next, we differeniate with respect to σ_j to obtain

$$\frac{\partial K_i}{\partial\sigma_j} = \frac{2\sigma_j}{\kappa_j^2\sum_{i'=1}^{i}\left[\sigma_{i'}/\kappa_{i'}^2\right]} - \frac{\pi + \sum_{i'=1}^{i}\left[\sigma_{i'}^2/\kappa_{i'}^2\right]}{\kappa_j^2\left[\sum_{i'=1}^{i}\left[\sigma_{i'}/\kappa_{i'}^2\right]\right]^2}$$
$$= \frac{2\sigma_j - K_i}{\kappa_j^2\sum_{i'=1}^{i}\left[\sigma_{i'}/\kappa_{i'}^2\right]}(>) \ge 0 \;\Leftrightarrow\; \sigma_j(>) \ge \frac{K_i}{2}.$$

(ii) A straightforward but tedious algebraic exercise confirms that $K_i = K_{i-1}$ if and only if $\sigma_i = K_{i-1}$. This implies that $K_i - \sigma_i = 0$ when evaluated at $\sigma_i = K_{i-1}$. Now,

$$\left.\frac{\partial K_i}{\partial\sigma_i}\right|_{\sigma_i=K_i} = \frac{\sigma_i/\kappa_i^2}{\sum_{i'=1}^{i}\left[\sigma_{i'}/\kappa_{i'}^2\right]} < 1.$$

This implies that $K_i - \sigma_i$ is strictly decreasing in σ_i when evaluated at $K_i - \sigma_i = 0$.
(iii) Define m to be largest member of $\{1,\ldots,n\}$ satisfying $\sigma_i < K_i$ for all $i \le m$. (Such an m can be found, because $\sigma_1 < K_1$ by inspection.) If $m = n$, the claim holds. If $m < n$, then $\sigma_{m+1} \ge K_{m+1}$. Because leaders are ordered by clarity, $\sigma_{m+2} \ge \sigma_{m+1} \ge K_{m+1}$ and hence, by using claim (ii), $\sigma_{m+2} \ge K_{m+2}$. Continuing inductively, $\sigma_j \ge K_j$ for all $j \ge m+1$.
(iv) Algebra confirms that $K_i < K_{i+1}$ if and only if $\sigma_{i+1} < K_i$. From claim (ii), this holds if and only if $\sigma_{i+1} < K_{i+1}$. In turn, and from claim (iii), this holds if and only if $i < m$.
(v) The first inequality holds if and only if $i \le m$ (from (iii)). The second inequality holds if and only if $\sigma_{i+1} \le K_{i+1}$ (from (ii)). This turn holds if and only if $i \ge m$ (from (iii)). ■

Proof of Proposition 3. In the light of Lemma 2, the equilibrium maximizes party welfare. Equivalently, it maximizes $\sum_{i=1}^n \hat{\psi}_i$ and so solves the problem

$$\max_{x\in\mathbb{R}_+^n}\sum_{i=1}^{n}\frac{1}{\pi\kappa_i^2 + \left[\sigma_i^2/x_i\right]}$$

subject to $x \in \mathbb{R}_+^n$ satisfying $\sum_{i=1}^n x_i = 1$. Differentiating, note that

$$\frac{\partial\hat{\psi}_i}{\partial x_i} = \frac{\sigma_i^2}{\left(\pi\kappa_i^2 + \left[\sigma_i^2/x_i\right]\right)^2 x_i^2} = \frac{\sigma_i^2}{\left(\pi x_i\kappa_i^2 + \sigma_i^2\right)^2},$$

which is positive and strictly decreasing in x_i. It follows that the objective function is strictly concave. The feasible set is compact and convex, and so there is a unique solution. The usual Kuhn–Tucker conditions are both necessary and sufficient. We use the Lagrange multiplier $\lambda > 0$ for the attention span constraint; the constraint binds because welfare is strictly increasing in attention and so $\lambda > 0$. For $x_i > 0$, the first-order condition is

$$\frac{\partial\hat{\psi}_i}{\partial x_i} = \frac{\sigma_i^2}{\left(\pi x_i\kappa_i^2 + \sigma_i^2\right)^2} = \lambda \quad\Rightarrow\quad \lambda < \frac{1}{\sigma_i^2}.$$

Hence, any leader who attracts an audience speaks with clarity exceeding λ. If $x_i = 0$, then

$$\left.\frac{\partial\hat{\psi}_i}{\partial x_i}\right|_{x_i=0} = \frac{1}{\sigma_i^2} \le \lambda,$$

and so the clarity of communication of an ignored leader falls below λ. Taken together, this means that the leaders who attract attention must be the best communicators: $x_i > 0$ if and only if $i \le m$ for some $m \in \{1,\ldots,n\}$. For this elite of m leaders,

$$\frac{\sigma_i^2}{\left(\pi x_i\kappa_i^2 + \sigma_i^2\right)^2} = \lambda \Leftrightarrow x_i = \frac{\sigma_i(1-\sigma_i\sqrt{\lambda})}{\pi\kappa_i^2\sqrt{\lambda}}$$
$$= \frac{\sigma_i(K_m - \sigma_i)}{\pi\kappa_i^2} \quad\text{where}\quad K_m \equiv \frac{1}{\sqrt{\lambda}}.$$

To find K_m (and the Lagrange multiplier $\lambda = 1/K_m^2$), we sum over the m-strong elite:

$$\sum_{i=1}^{m} x_i = \sum_{i=1}^{m}\frac{\sigma_i(K_m - \sigma_i)}{\pi\kappa_i^2} = 1 \quad\Leftrightarrow\quad \pi + \sum_{i=1}^{m}\frac{\sigma_i^2}{\kappa_i^2} = K_m\sum_{i=1}^{m}\frac{\sigma_i}{\kappa_i^2},$$

which solves for K_m. We need only find the unique m satisfying $\sigma_m < K_m \le \sigma_{m+1}$. Claims (iii)–(v) of Lemma 4 characterize the existence and properties of this m.

We turn to the comparative static claims, and for now assume that $K_m < \sigma_{m+1}$, so that a local change in K_m does not change the size of the elite. If $m = 1$, then $x_1 = 1$ is invariant to local parameter changes. For $m \ge 2$ and $i \le m$, differentiate x_i with respect to $[1/\kappa_i^2]$:

$$\frac{\partial x_i}{\partial[1/\kappa_i^2]} = \frac{\sigma_i(K_m - \sigma_i)}{\pi} + \frac{\sigma_i}{\pi\kappa_i^2}\frac{\partial K_m}{\partial[1/\kappa_i^2]} = \frac{\sigma_i(K_m - \sigma_i)}{\pi}$$
$$+ \frac{\sigma_i}{\pi\kappa_i^2}\left[\frac{\sigma_i^2}{\sum_{j=1}^{m}\left[\sigma_j/\kappa_j^2\right]} - \sigma_i\frac{\pi + \sum_{j=1}^{m}\left[\sigma_j^2/\kappa_j^2\right]}{\left[\sum_{j=1}^{m}\left[\sigma_j/\kappa_j^2\right]\right]^2}\right]$$
$$= \frac{\sigma_i(K_m - \sigma_i)}{\pi}\left[1 - \frac{\sigma_i/\kappa_i^2}{\sum_{j=1}^{m}\left[\sigma_j/\kappa_j^2\right]}\right] > 0.$$

Next, differentiate x_i with respect to σ_i:

$$\frac{\partial x_i}{\partial \sigma_i} = \frac{1}{\pi\kappa_i^2}\left[K_m - 2\sigma_i + \sigma_i \frac{\partial K_m}{\partial \sigma_i}\right] = \frac{K_m - 2\sigma_i}{\pi\kappa_i^2}\left[1 - \frac{\sigma_i/\kappa_i^2}{\sum_{j=1}^m \left[\sigma_j/\kappa_j^2\right]}\right].$$

This is (strictly) positive if and only if σ_i is (strictly) less than $K_m/2$. The comparative static claims also hold when $K_m = \sigma_{m+1}$: for any local parameter change that increases K_m, and so introduces leader $m+1$ into the elite, we replace K_m with K_{m+1} in the formulae derived previously, noting that $K_m = K_{m+1}$ when $K_m = \sigma_{m+1}$.

Finally, the size m of the elite is determined by the inequalities $\sigma_m < K_m \leq \sigma_{m+1}$. Changes in κ_i^2 for $i > m$ have no effect. A fall in σ_{m+1} (or in σ_i for any $i > m$, following an appropriate renaming of players) can expand the elite. An increase in π (by inspection) or in κ_i^2 for $i \leq m$ (from Lemma 4) increases K_m and so may expand the elite. ■

Proof of Proposition 4. For attention to be focused on the clearest communicator, we need

$$\sigma_2 \geq K_1 = \frac{\pi + \left[\sigma_1^2/\kappa_1^2\right]}{\sigma_1/\kappa_1^2} = \sigma_1\left[1 + \frac{\pi\kappa_1^2}{\sigma_1^2}\right].$$

Squaring yields the lower bound on σ_2^2 given in the proposition. Now,

$$\frac{\partial K_1^2}{\partial \sigma_1^2} = \left[1 + \frac{\pi\kappa_1^2}{\sigma_1^2}\right]^2 - 2\sigma_1^2\left[1 + \frac{\pi\kappa_1^2}{\sigma_1^2}\right]\frac{\pi\kappa_1^2}{\sigma_1^4} = 1 - \left[\frac{\pi\kappa_1^2}{\sigma_1^2}\right]^2.$$

This is increasing in σ_1^2, and hence, K_1^2 is convex in σ_1^2. Setting the derivative to zero yields $\sigma_1^2 = \pi\kappa_1^2$, as claimed. The remaining claims follow by inspection. ■

Proof of Lemma 3. A leader's optimally chosen clarity satisfies $\sigma_i = \max\{\underline{\sigma}_i, K_m/2\}$. To prove this claim, note that if $\sigma_i < (>) K_m/2$, then she would want to reduce (increase) her clarity. This is because either (i) $x_i \in (0, 1)$ is locally increasing (decreasing) in σ_i (from Proposition 3), or (ii) $x_i \in \{0, 1\}$ and the additional assumptions made for these cases apply. Hence, if $\underline{\sigma}_i \leq K_m/2$, then $\sigma_i = K_m/2$, and if $\underline{\sigma}_i > K_m/2$, then $\sigma_i = \underline{\sigma}_i$. We conclude that she obfuscates if $\underline{\sigma}_i < K_m/2$. To obtain a lower bound to $K_m/2$, let us choose a set of clarities for all leaders to minimize K_m. Following Proposition 3, this is achieved when $\sigma_j = K_m/2$ for all j. Algebraic manipulations confirm that this is so when $\sigma_j = \tilde{\sigma}$ for all j. Hence, leader i will never choose clarity $\sigma_i^2 < \tilde{\sigma}^2$, and so she obfuscates if $\underline{\sigma}_i^2 < \tilde{\sigma}^2$. ■

Proof of Proposition 5. Following the proof of Lemma 3, equilibrium clarity choices satisfy $\sigma_i = \max\{\underline{\sigma}_i, K_m/2\}$ and hence $\sigma_1^2 \leq \cdots \leq \sigma_n^2$. A leader obfuscates if and only if $\underline{\sigma}_i < K_m/2$, which holds if and only if $i \leq \tilde{m}$ for some $\tilde{m}$.

To prove the existence and uniqueness of the equilibrium, we construct the function $f(\sigma)$ in the following way. Fixing $\sigma \geq \underline{\sigma}_1$, let $\sigma_i = \max\{\sigma, \underline{\sigma}_i\}$ for each i, take the unique equilibrium from Proposition 3, and then set $f(\sigma) = K_m/2$. An equilibrium corresponds to either (i) $\sigma = \underline{\sigma}_1$ if $f(\sigma) < \underline{\sigma}_1$, or (ii) a fixed point $\sigma = f(\sigma)$ of this function.

To show that there is a unique such σ, we consider the properties of $f(\sigma)$. Firstly, it is positive and continuously differentiable in σ. Secondly, it crosses σ at most once and from above to below. To verify this second claim, note that

$$\left.\frac{\partial f(\sigma)}{\sigma}\right|_{\sigma = f(\sigma)} = \frac{1}{2}\sum_{i=1}^m \left.\frac{\partial K_m}{\partial \sigma_i}\right|_{\sigma_i = K_m/2} = 0,$$

and so $f(\sigma)$ is locally constant around a fixed point $\sigma = f(\sigma)$, which implies that $f(\sigma) - \sigma$ is decreasing. Hence, if $f(\sigma) \leq \underline{\sigma}_1$, then $f(\sigma) < \sigma$ for all $\sigma > \underline{\sigma}_1$, which implies that there is a unique equilibrium in which every leader speaks with maximum clarity. If $f(\sigma) > \underline{\sigma}_1$, then there is at most one fixed point $\sigma = f(\sigma)$ in the range $\sigma > \underline{\sigma}_1$. To show that there is such a fixed point, we need only show that $f(\sigma) < \sigma$ for some σ. Now, for all $\sigma \geq \underline{\sigma}_n$,

$$f(\sigma) = \frac{\pi + \sigma^2 \sum_{i=1}^n \left(1/\kappa_i^2\right)}{2\sigma \sum_{i=1}^n \left(1/\kappa_i^2\right)} < \sigma \quad \Leftrightarrow \quad \sigma^2 > \frac{1}{\sum_{i=1}^n \left(1/\kappa_i^2\right)}.$$

Hence, for $\sigma > f(\sigma)$ for σ large enough. The unique fixed point yields the unique equilibrium. Straightforward calculations reveal that $\tilde{\sigma}$ is this fixed point as long as $\tilde{\sigma} > \underline{\sigma}_n$.

We turn to the comparative static claims. K_m is (at least weakly) increasing in κ_i^2 for all i and in π. $f(\sigma)$ inherits the properties. An increase in $f(\sigma)$ raises the fixed point, hence reducing the clarity of all obfuscators. Equivalently, clarity is increasing in each leader's sense of direction and in activists' desire $(1-\pi)$ for party unity. The increase in clarity correponds to a reduction in the fixed point and hence in K_m. This reduces the number of leaders satisfying $\underline{\sigma}_i < K_m/2$ and $\underline{\sigma}_i < K_m$, and so, respectively, reduces $\tilde{m}$ and m. ■

Proof of Proposition 6. The inverse of policy performance satisfies

$$\mathrm{E}[(a_t - \theta)^2 \mid \theta] = \sum_{i=1}^n w_i^2\left(\kappa_i^2 + \frac{\sigma_i^2}{x_i}\right) = \frac{1+\pi}{\sum_{i=1}^n 1/\kappa_i^2},$$

where the second equality follows from substitution and simplification. This increases with π and decreases with the aggregate sense of direction of the leaders. ■

Construction of Figure 1. Uses the formula from Proposition 4. ■

Construction of Figure 2. We computed the reaction function for a leader with unconstrained clarity. Note that such a leader i chooses her clarity optimally when $\sigma_i = K_m/2$. Hence,

$$2\sigma_i = K_m = \frac{\pi + \sum_{j=1}^m \left[\sigma_j^2/\kappa_j^2\right]}{\sum_{j=1}^m \left[\sigma_j/\kappa_j^2\right]} = \frac{\pi\kappa_i^2 + A + \sigma_i^2}{B + \sigma_i}$$

$$\Leftrightarrow \sigma_i^2 + B\sigma_i - \left(A + \pi\kappa_i^2\right) = 0,$$

where $A \equiv \kappa_i^2 \sum_{j \neq i}[\sigma_j^2/\kappa_j^2]$ and $B \equiv \kappa_i^2 \sum_{j \neq i}[\sigma_j/\kappa_j^2]$. Solving for the positive root,

$$\sigma_i = -B + \sqrt{A + B^2 + \pi\kappa_i^2} = \sqrt{\pi\kappa_i^2 + \kappa_i^2 \sum_{j \neq i} \frac{\sigma_j^2}{\kappa_j^2} + \kappa_i^4 \left[\sum_{j \neq i} \frac{\sigma_j}{\kappa_j^2}\right]^2} - \kappa_i^2 \sum_{j \neq i} \frac{\sigma_j}{\kappa_j^2}.$$

For the case of two players, this reduces to

$$\sigma_i = \sqrt{\pi\kappa_i^2 + \sigma_j^2\left[\frac{\kappa_i^2}{\kappa_j^2} + \frac{\kappa_i^4}{\kappa_j^4}\right]} - \sigma_j \frac{\kappa_i^2}{\kappa_j^2},$$

which is the formula used in the construction of Figure 2. ■

REFERENCES

Allen, Franklin, Stephen Morris, and Hyun S. Shin. 2006. "Beauty Contests and Iterated Expectations in Asset Markets." *Review of Financial Studies* 19 (3): 719–52.

Alvarez, R. Michael, and Charles H. Franklin. 1994. "Uncertainty and Political Perceptions." *Journal of Politics* 56 (3): 671–88.

Amato, Jeffrey D., Stephen Morris, and Hyun S. Shin. 2002. "Communication and Monetary Policy." *Oxford Review of Economic Policy* 18 (4): 495–503.

Angeletos, George-Marios, and Alessandro Pavan. 2004. "Transparency of Information and Coordination in Economies with Investment Complementarities." *American Economic Review: AEA Papers and Proceedings* 94 (2): 91–8.

Angeletos, George-Marios, and Alessandro Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica* 75 (4): 1103–42.

Aragonés, Enriqueta, and Zvika Neeman. 2000. "Strategic Ambiguity in Electoral Competition." *Journal of Theoretical Politics* 12 (2): 183–204.

Brands, H. W. 2005. *Andrew Jackson: His Life and Times*. New York: Doubleday.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Calvert, Randall L. 1995. "The Rational Choice of Social Institutions: Cooperation, Coordination, and Communication." In *Modern Political Economy: Old Topics, New Directions*, ed. Jeffrey S. Banks and Eric A. Hanushek. Cambridge, UK: Cambridge University Press.

Carwardine, Richard J. 2003. *Lincoln*. Harlow: Pearson Longman.

Chwe, Michael Suk-Young. 2001. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton, NJ: Princeton University Press.

Cornwell, Elmer E., Jr. 1957. "Coolidge and Presidential Leadership." *Public Opinion Quarterly* 21 (2): 265–78.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–51.

Dewan, Torun, and David P. Myatt. 2007. "Leading the Party: Coordination, Direction and Communication." *American Political Science Review* 101 (4): 825–43.

Dewan, Torun, and David P. Myatt. 2008. "On the Rhetorical Strategies of Leaders: Speaking Clearly, Stepping Back, and Standing Down." *London School of Economics and University of Oxford*. Working Paper.

Ellis, Joseph J. 2005. *His Excellency: George Washington*. London: Faber.

Fiorina, Morris P., and Kenneth A. Shepsle. 1989. "Formal Theories of Leadership: Agents, Agenda-Setters, and Entrepreneurs." In *Leadership and Politics: New Perspectives in Political Science*, ed. Bryan D. Jones. Lawrence: University Press Kansas, 17–41.

Franklin, Charles H. 1991. "Eschewing Obfuscation? Campaigns and the Perception of US Senate Incumbents." *American Political Science Review* 85 (4): 1193–1214.

Gilligan, Thomas W., and Keith Krehbiel. 1987. "Collective Decisionmaking and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures." *Journal of Law, Economics, and Organization* 3 (2): 287–335.

Güth, Werner, M. Vittoria Levati, Matthias Sutter, and Eline van der Heijden. 2004. "Leadership and Cooperation in Public Goods Experiments." Jena, Germany: Max Planck Institute of Economics.

Hellwig, Christian. 2005. "Heterogeneous Information and the Welfare Effects of Public Information Disclosures." Working paper, UCLA.

Hellwig, Christian, and Laura Veldkamp. 2008. "Knowing What Others Know: Coordination Motives in Information Acquisition." *Review of Economic Studies*. Forthcoming.

Hermalin, Benjamin E. 1998. "Toward an Economic Theory of Leadership: Leading by Example." *American Economic Review* 88 (5): 1188–206.

Humphreys, Macartan, William Masters, and Martin E. Sandbu. 2006. "The Role of Leadership in Democratic Deliberations: Results from a Field Experiment in São Tomé and Principe." *World Politics*. Forthcoming.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.

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

Li, Hao, Sherwin Rosen, and Wing Suen. 2001. "Conflicts and Common Interests in Committees." *American Economic Review* 91 (5): 1478–97.

Mierowitz, Adam. 2005. "Informational Party Primaries and Strategic Ambiguity." *Journal of Theoretical Politics* 17 (1): 107–36.

Morris, Stephen, and Hyun S. Shin. 2002. "Social Value of Public Information." *American Economic Review* 92 (5): 1521–34.

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

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

Riker, William. 1996. *The Strategy of Rhetoric: Campaigning for the American Constitution*. New Haven, CT: Yale University Press.

Shepsle, Kenneth A. 1970. "A Note On Zeckhauser's 'Majority Rule with Lotteries on Alternatives': The Case of the Paradox of Voting." *Quarterly Journal of Economics* 84 (4): 705–10.

Shepsle, Kenneth A. 1972. "The Strategy of Ambiguity: Uncertainty and Electoral Competition." *American Political Science Review* 66 (2): 555–69.

Zeckhauser, R. 1969. "Majority Rule with Lotteries on Alternatives." *Quarterly Journal of Economics* 83 (2): 696–703.