Political Science and Political Economy Working Paper

Department of Government

London School of Economics

No. 1/2008

The Qualities of Leadership: Direction, Communication, and Obfuscation

Torun Dewan (LSE)

David P. Myatt (Oxford)

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

The Qualities of Leadership: Direction, Communication, and Obfuscation

Torun Dewan
Department of Government, London School of Economics
t.dewan@lse.ac.uk

David P. Myatt
Department of Economics, University of Oxford
david.myatt@economics.ox.ac.uk

Revised for the American Political Science Review. September 2007.

Abstract. Party activists wish to (i) advocate the best policy while (ii) unifying behind a party line. Activists learn about policy and the party line by listening to leaders. A leader's influence increases with her judgement on policy (sense of direction) and her ability to convey ideas (clarity of communication). A leader with perfect clarity enjoys greater influence than one with a perfect sense of direction. When activists choose how much attention to pay to leaders they listen only to the most coherent communicators. However, attention-seeking leaders sometimes obfuscate their messages, but less so when activists emphasize party unity.

Political scientists and commentators agree that good leadership is important, indeed fundamental, to the successful performance of organizations. 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, Smith, Siverson, and Morrow, 2003), but do not yet have answers to fundamental questions. What is leadership? When is leadership good and when is is it successful? Which qualities contribute to good and successful leadership, and how do these qualities arise?

Leadership can be important when political actors wish 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. Leaders can also help people to learn. As Levi (2006) argued recently, "leadership . . . provides the learning environment that enables individuals to transform or revise beliefs."

To illustrate, consider 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. He 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 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: a speech could yield a common viewpoint around which support can coalesce. This is important, since an 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 to provide such assessments.

Within this framework, a good leader helps activists to achieve their goals: her communication fosters the understanding that is needed for activists to advocate the right policy, and to advocate it together. On the other hand, a successful leader is one who has influence: her words impact upon the actions taken by activists. 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 policy for the party. The value of that information reflects the quality of her judgement. History provides us examples of those who, on the bigger issues of the day, appeared to know instinctively the best course to pursue. Of George Washington, for example, Ellis (2005) wrote "whatever minor missteps he had made along the way, 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: a leader may surround herself with a clique or cabinet of advisors; her sense of direction then arises from the combined judgements of such a group.

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, an activist 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 which points all activists in the wrong direction, but is commonly interpreted, may sometimes be preferable to one which 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 policy implications of that message. A poor communicator by contrast, though 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, since messages might be transmitted indirectly via interlocutors. For example, a follower might hear a leader's views through a party spokesperson, from political correspondents, or via other media sources. When a message is conveyed via multiple media, different activists 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 party activists wish (i) to advocate the best policy and yet (ii) to unify behind a common party line. Our activists listen and respond to leaders. In equilibrium, the relative influence of a leader and aggregate party performance increases 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 party unity: when a leader speaks clearly, activists rally around a commonly understood party line, even though it may differ from the ideal policy.

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 activists 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 activists endogenously decide to whom to listen. Indeed, the desire to coordinate means that an activist will choose to listen 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 for the party, while others are ignored.

Of course, an ambitious leader might adapt her rhetorical strategy to increase the attention paid to her. This drives a wedge between good and successful leadership: a good leader helps party members to advocate the right policies, whereas a successful leader enjoys the biggest audience. The prominence of clear communicators in the attention-seeking elite 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. An activist need not linger in her audience; keen to further his understanding of his environment, and having heard what he needs to hear, he quickly 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 her audience for longer while they digest her message, so preventing them from listening to others. Of course, her optimally chosen clarity will depend upon her own sense of direction and the qualities (both exogenous and endogenous) of competing leaders; for instance, she may speak more clearly when she has more to say.

The willingness of a leader to blur her message also depends critically on the importance of party unity. When unity is important, activists emphasize the adoption of a commonly understood party line. In pursuit of that goal they pay more attention to clearer speakers and, reacting to this incentive, leaders may communicate more clearly. Of course, clearer communication enhances the informativeness of a leader's message: a twist to our story is that parties which focus on party unity, thus ensuring that all activists are singing from the same hymn sheet, also develop a better understanding of policy.

Our focus on the rhetorical strategies of leaders relates our work to that of Riker (1996), whilst our emphasis on the (endogenous) clarity in leaders' communication has connections to the study of strategic ambiguity, whereby leaders are equivocal on policy in order to appeal to a broader section of voters (Zeckhauser, 1969; Shepsle, 1970, 1972b). Equivocation and obfuscation are conceptually different, though both are aspects of rhetorical strategy and may be observationally equivalent. Whilst 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 leaders, though, as we show, may be absent in some strategic scenarios.

COORDINATING PARTY ACTIVISTS

Our study of leadership builds upon a simple game in which party activists wish 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, since 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 have reached the third degree where 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 (Morris and Shin, 2002; Angeletos and Pavan, 2007) 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).

The beauty-contest parable can also offer new insights to political science. Rather than think of contestants choosing the prettiest faces, we can think of political actors 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 $t \in [0, 1]$. 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 follow the party line. We represent these twin concerns via a pair of quadratic loss functions:

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

Here π indexes an activist's relative concern for choosing the ideal policy compared to maintaining party unity. If $\pi = 1$ then an activist would be solely concerned with advocating the best policy; if $\pi = 0$, by contrast, then he would seek only to conform by minimizing the distance between himself and the party line.[1] We wish to consider situations in which both of these concerns are present and so we assume that $0 < \pi < 1$.

Optimal Advocacy

When activists share common knowledge of θ then it is optimal for them all to advocate the same ideal policy $a_t = \theta$. In fact, this is the unique Nash equilibrium of the game.[2] When this is so there is no tension between the activists' twin objectives.

However, 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 can be solved straightforwardly to yield 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 of the game. An informative signal might represent the activist's own private research, the understanding gained from communication with other party members, or the influence of a party leader. We postpone our description of the precise statistical properties of the signals until the next section of the paper. 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

[1]The loss function $(a_t - \bar{a})^2$ arguably represents a desire for conformity rather than party unity. A convenient measure of disunity might be the aggregate distance between an activist's advocated policy and the policies advocated by other party members, leading to the loss function $\int_0^1 (a_t - a_{t'})^2\,dt'$. As we explain in the technical appendix, the use of this "party disunity" loss function leads to very few changes in our results.

[2]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 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. Since 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 party-wide use of the strategy. Given activists use a strategy $A(\cdot)$, an activist's expectation of the party line is $\mathrm{E}[\bar{a} \mid \tilde{s}_t] = \mathrm{E}[A(\tilde{s}_{t'}) \mid \tilde{s}_t]$ for $t' \neq t$. Similarly, his expectation of the ideal policy is $\mathrm{E}[\theta \mid \tilde{s}_t]$. Hence the strategy $A(\cdot)$ forms an equilibrium if and only if

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

Thus an activist's strategy is a weighted average of the expected ideal policy and his understanding of the average policy advocated by the party at large.

To find an equilibrium we may solve Equation ($\star$) to find an equilibrium advocacy strategy $A(\cdot)$. However, we need to specify how signals help an activist to 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 can help an activist to develop his beliefs about policy and the likely actions of others. We assume that activists begin with no substantive knowledge of the true ideal policy: they share a diffuse prior over θ.[3] To acquire information 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 party leaders and include the activists' own research and intra-party communication. Nevertheless, this personification of the information sources is useful for exposition and is relevant when we subsequently introduce a role for strategic leaders who manipulate the clarity of the information that activists obtain.

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

Each leader forms an independent, unbiased, and private opinion of the ideal policy for the party. Formally, leader i observes an informative signal s_i satisfying

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

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 judge 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 to her.

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. Hence she has no incentive to misrepresent her views and describes the world as she sees it; strategic information transmission is ruled out. Alas, a leader is unable to communicate perfectly: each activist t observes the leader's private signal plus noise. Formally,

$$\tilde{s}_{it} \,|\, s_i \sim N(s_i, \sigma_i^2) \quad \text{and so} \quad \frac{1}{\sigma_i^2} = \textit{Clarity of Communication}.$$

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. Since 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 her 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 via a television or radio broadcast is a relatively recent phenomenon. Prior to these technological innovations, a leader would need to travel far and wide to reach a similar audience; her ability to communicate clearly 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 spoken words. For example, in the United Kingdom the Lobby serves as the gateway to the House of Commons. Ministerial briefings are delivered to an elite group of 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 all of a leader's words, doing so has an opportunity cost, since it prevents him from paying careful attention to the communication of 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 (1972b) who both considered the strategic advantage realized by a candidate who is vague over the policy she would implement if elected. Shepsle (1972b), in particular, highlighted the "politician's advantage in speaking half-truths and in varying his appeals with variations in audience and political climate." [4]

[4]Our focus on the optimal level of equivocation also relates our work to that of Aragonés and Neeman (2000) who, building on the literature cited here, focused on equilibrium levels of candidates' ambiguity. Our focus on strategic incentives to obfuscate and its effects on party welfare and information aggregation, relates our work to that of Mierowitz (2005). In contrast to our model, in which activists learn from leaders, he considered a world where politicians learn about voters' preferences during primary elections. Candidates who refrain from committing to specific policies have greater flexibility upon receiving information from the primary.

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." This helps to justify further theoretical study of clarity choice.

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 paper. 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. While Brown is perceived as amongst the most intellectually astute of his cohort (low κ_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 which is commonly observed and interpreted; and private signals which 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 ($\sigma_i^2 > 0$).[5] 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, since it allows for positive but imperfect correlation between the signals observed by activists.

[5]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 $\sigma_i^2 = 0$.

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, so 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 employs 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' information $\mathrm{E}[s \mid \tilde{s}_t]$ 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. (Stated somewhat more formally, the class of symmetric and linear advocacy strategies is closed under best reply.) Our formal statement of this observation, which is recorded in the following lemma, goes further (formal proofs are contained in the technical appendix.)

Lemma 1. *If all other activists $t' \neq t$ use a linear strategy $A(\tilde{s}_t) = \sum_{i=1}^{n} w_i \tilde{s}_{it}$ then the unique best reply for activist t is to use a linear strategy. There is a unique Bayesian Nash equilibrium involving the use of linear strategies. This satisfies $w_i > 0$ for all i and $\sum_{i=1}^{n} w_i = 1$.*

Unsurprisingly, the weight placed on a leader's speech is positive. Happily, the weights also sum to one, so that the policy advocated by an activist is 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; other things equal, it can measure the success of a leader.

Of course, the possibility of non-linear equilibria remains open. Nevertheless, in our technical appendix we explain how the imposition of a mild further restriction on the class of advocacy strategies we consider is sufficient to rule out such non-linear equilibria.

Furthermore, even if non-linear equilibria did exist (and we conjecture that they do not) then further criteria suggest the selection of the unique linear equilibrium.[6]

Focusing on the party-wide deployment of a linear advocacy strategy, we must find the weights placed on the leaders' speeches. One "brute force" approach to this task 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 ($\star$). Rather than doing this (although we explain the details of this method in the technical appendix) we take a rather more subtle approach. We observe that the weights use by activists in equilibrium are precisely those which are desirable from the perspective of the entire party; that is, 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; put simply, the equilibrium is efficient.*

To verify this claim, suppose that activist t contemplates a change in a_t. This change imposes externalities on others. For instance, 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 conveniently 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 individual party members lie symmetrically above and below $\bar{a}$. Since the various externalities from an activist's action cancel out, the total party-wide effect of a marginal change in the policy he advocates is completely reflected in his own expected payoff. This means that he faces socially correct incentives (from the perspective of his party) at the margin; in aggregate, party members choose their actions to maximize party welfare.

With Lemma 2 established, we now calculate the weights that maximize aggregate party welfare and hence find the linear equilibrium of our advocacy game. The party's welfare, which corresponds to the *ex ante* expected payoff of a randomly chosen activist, 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 party leaders. Putting these elements together,

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

[6]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 the unique linear equilibrium which they constructed is also the unique equilibrium; that is, they claimed that non-linear 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, p. 1112).

Notice that any noise in the information sources available to activists detracts from party welfare. Interestingly, the noise $\pi\kappa_i^2 + \sigma_i^2$ arising from a leader's speech does not equally weight the variances κ_i^2 and σ_i^2. A lack of clarity in a leader's communication frustrates activists' coordination as well as lessening the information content of her speech. By contrast, a failing in her sense of direction, whilst affecting an activist's ability to advocate the ideal policy, has no impact on the party membership's coordination; it thus attracts a reduced weight of π. These observations regarding the different effects of our leadership skills are reflected in the equilibrium weights which 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 better communicators increases as activists' concern for party unity grows: if $\sigma_i^2 < \sigma_j^2$ then ψ_i/ψ_j is decreasing in π.

The index $\hat{\psi}_i$ measures influence and therefore successful leadership; however, since 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 in both leadership skills (so that $\sigma_i^2 < \sigma_j^2$ and $\kappa_i^2 < \kappa_j^2$) enjoys more influence (so that $w_i > w_j$). Nevertheless, even unskilled leaders enjoy some influence: since $w_i > 0$ for all i, no leader dominates completely.

So which skill is more important? The presence of π in the index $\hat{\psi}_i$ suggests that a leader's ability to give clear expression to her views 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, the final claim of Proposition 1 reveals that relative influence of clear communicators grows as party cohesion looms larger in the minds of activists. Since $\sum_{i=1}^n w_i = 1$ this necessarily means that the absolute following of poorer communicators must fall as that of skilled orators grows. Furthermore, an inspection of $\hat{\psi}_i$ reveals that influence is entirely determined by clarity as π vanishes to zero.[7]

Corollary. *An increase in activists' concern for party unity shifts influence away from poor communicators and toward clear communicators: if leaders' labels are in order of decreasing clarity, so that $\sigma_1^2 < \sigma_2^2 < \cdots < \sigma_n^2$, then there is some k such that w_i is locally decreasing in π for $i < k$ and locally increasing in π for $i > k$. If π is sufficiently small then $w_1 > w_2 > \cdots > w_n > 0$.*

[7]For ease of exposition the corollary assumes that clarities strictly differ. There is little loss of generality; appropriately modified claims may be made whenever $\sigma_i^2 = \sigma_j^2$ for some $i \neq j$.

To obtain further insight recall once again that a leader helps activists to learn about policy and to 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,

$$\tilde{s}_{it} \,|\, \theta \sim N(\theta, \kappa_i^2 + \sigma_i^2) \quad \text{so that} \quad \psi_i \equiv \frac{1}{\kappa_i^2 + \sigma_i^2} = \textit{Quality of Information}.$$

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 $a_t = \mathrm{E}[\theta \,|\, \tilde{s}_t]$ without reference to others. Bayesian updating yields

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

so that the weight placed on each leader's speech is proportional to the quality of information ψ_i which 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.

An appropriate measure of the commonality of messages received is the correlation between what is heard by different activists. To calculate this correlation, note that the conditional covariance of two signals related to the same speech is $\mathrm{cov}[\tilde{s}_{it}, \tilde{s}_{it'} \,|\, \theta] = \kappa_i^2$. 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. On the other hand, 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 $\mathrm{E}[\tilde{s}_{it'} \,|\, \tilde{s}_t] = \tilde{s}_{it}$: 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

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

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.

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.

Since 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

We have studied a model in which activists received exogenous signals via the speeches of leaders. We might think of all party members attending a large party conference where each listens carefully 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 not only on the clarity of her communication but also on the willingness of her audience to listen; but the decision to listen is endogenous. This is important when being informed is costly as, for example, when activists have a limited attention span and are unable to listen to a leader indefinitely.

Given that activists choose to whom to listen, leaders may try and capture their attention. How much attention a leader receives depends upon her skills endowment; specifically the clarity of her communication and her sense of direction. Whilst the latter might be seen as an exogenous trait, the former is more manipulable; a leader may vary her clarity as and when the need arises. That is, the overall clarity of a message depends endogenously on both speaker and audience.

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, since 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,

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

and so the equilibrium x maximizes $\sum_{i=1}^{n} \hat{\psi}_i$ subject to $\sum_{i=1}^{n} x_i \leq 1$. Since welfare is increasing in the attention paid to each leader, activists will certainly 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 and which do not 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, and for simplicity of exposition, we order (without loss of generality) the leaders in order of decreasing clarity, so that $\sigma_1^2 \leq \cdots \leq \sigma_n^2$.[8]

Proposition 3. *When leaders' audiences are exogenous 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 $i \leq m$,*

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

Amongst the elite, the attention paid to a leader increases with her sense of direction though not always with the clarity of her communication: 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 size of the elite is the largest m satisfying $\sigma_m < K_m$. The elite's size increases with activists' concern π for policy versus party unity, but decreases with each leader's sense of direction.

When all n leaders share the same communication skills, so that $\sigma_i = \sigma_j$ for all $i \neq j$, then $m = n$, so that every leader enjoys an attentive audience. Furthermore, the attention paid to each leader is proportional to her sense of direction. However, when leaders differ in their coherence richer results emerge.

The leaders toward whom activists gravitate (so that $x_i > 0$) are the clearest communicators. Correspondingly, once a leader's clarity of communication falls below a threshold (that is, when $\sigma_i > K_m$ where m is the elite size) activists will ignore her; such a leader can have no influence. Whilst intuitively one might think that a good sense of direction would demand attention, our result highlights the importance of getting the message across.

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 of communication remains a pre-requisite for successful leadership. One possibility emerging from Proposition 3 is that $m = 1$, so that activists pay attention only to the leader with the best communication skills, ignoring the speeches made by others. 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?

[8] For leaders with the same clarity of communication (so that, for instance, $\sigma_i^2 = \sigma_j^2$ for $i \neq j$) this order is not uniquely defined. Nevertheless, this does not matter: Propositions 3 and 4 hold so long as $\sigma_1^2 \leq \cdots \leq \sigma_n^2$.

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*

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

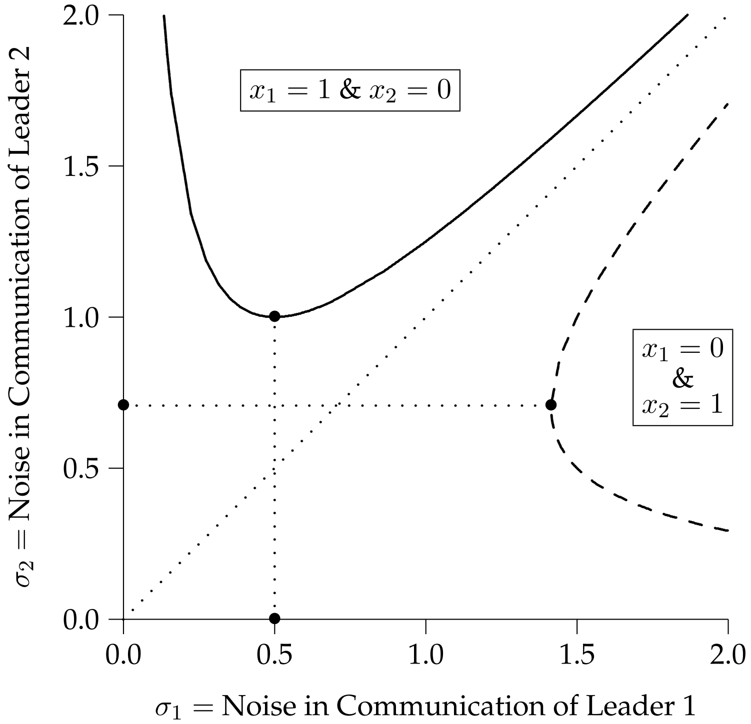
The right-hand side of this inequality is convex in σ_1^2, is minimized by $\sigma_1^2 = \pi\kappa_1^2$, and explodes as $\sigma_1^2 \to 0$. Hence, for the clearest communicator to enjoy exclusive attention as a de facto *dictator she needs to communicate imperfectly. The clarity which best supports her dictatorship (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 so long as π is sufficiently small.*

A *de facto* dictator must be the clearest communicator (Proposition 3). For her to enjoy 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 provide a focal policy around which the party membership 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: if a leader wishes to monopolize the agenda, and therefore maintain complete influence, then she needs to avoid perfect clarity; better communication can sometimes divert attention toward others.[9] 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

[9] Strictly speaking, there is no perfect clarity since $\sigma_i^2 > 0$ 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.



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; similiarly, 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.

FIGURE 1. Paying Attention To A Single Leader

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, the size m of the elite to whom activists listen declines with the policy-concern parameter π. Hence, as the desire for party unity grows (so that π falls) the size of the leadership elite shrinks. Indeed, an inspection of Proposition 4 reveals that the best communicator will become a *de facto* dictator 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 to find a clear focal point around which the membership can coalesce.

ATTRACTING ATTENTION

We have examined a world in which information transmission depends endogenously on a leader's audience. We now allow a leader's clarity to be chosen endogenously. We assume that leaders crave attention: a leader selects the clarity which maximizes the attention paid to her. Simply, leaders like a large and attentive audience.

Formally, 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 of communication which represents 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 a leader's communication skills are poor ($\underline{\sigma}_i^2$ is large) then she may be unable to attract an audience, so that $x_i = 0$ for all $\sigma_i^2 \geq \underline{\sigma}_i^2$. In this case, we assume (without loss of generality, and solely for ease of exposition) that if she is unable to attract listeners then she chooses to speak as clearly as she can.[10]

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 wishes to speak with perfect clarity. When $\sigma_i < K_m/2$, the attention paid to her increases with noise added to her speech, and so she obfuscates. 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 her) for a given quantity (clarity of message); on the other hand, obfuscation prompts an activist to lower his demand for her product (speech) by substituting to others. Balancing the two effects of a change in clarity generates an intermediate solution.

[10]This is equivalent to specifying lexicographic preferences: she wishes to maximize the attention paid to her but, subject to that, speaks as clearly as possible. This ensures that a leader's best reply is unique.

Lemma 3 reveals that leaders will sometimes obfuscate, and offers some predictions via the bound $\sigma_i^2 \geq \tilde{\sigma}^2$ on a leader's clarity. To move further, however, we seek an equilibrium of the attention-seeking game played by leaders.

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

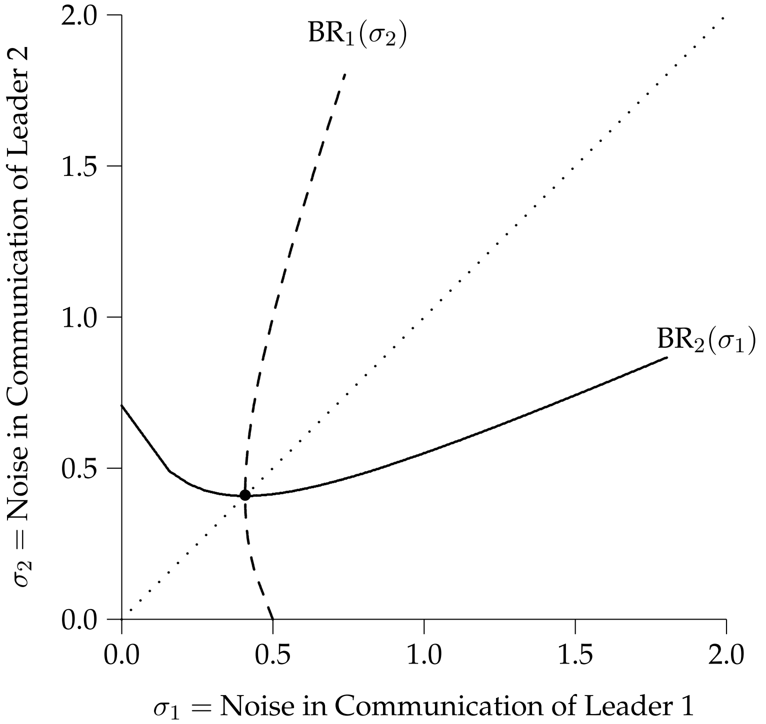
If $\underline{\sigma}_i^2 < \tilde{\sigma}^2$ for all i (so that all leaders have good communication skills) then $\tilde{m} = n$ (they all obfuscate) and the unique Nash equilibrium is symmetric, satisfying $\sigma_i^2 = \tilde{\sigma}^2$ for all i. Hence,

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

Each leader's clarity of communication increases with every leader's sense of direction and activists' desire for party unity. Finally, the number $\tilde{m}$ of leaders who obfuscate and the number m of leaders who attract an audience both decrease with activists' desire for party unity.

Linking Propositions 3 and 5, there are (potentially, at least) 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, the clarity choices of these leaders solves $\sigma_i = K_m/2$, where K_m is the expression in Proposition 3.) Since members of this group of $\tilde{m}$ 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 attain an audience. The exogenous limits to their oratorical skills mean that they do not obfuscate, but they maintain an audience as long as their clarity satisfies $\sigma_i \leq \sigma_i = K$. When $i > m$, then a leader talks only to herself. To attain an audience she must raise her clarity. When the noise in her speech is no greater than twice that of the clearest communicators, her views will 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 choose to obfuscate. This will be so when there are few exogenous limits to the clarity of a leader's message; 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 Nash equilibrium identified by Proposition 5 (illustrated in Figure 2) reveals that in this situation all leaders speak with the same clarity, even though they do not share a common sense of direction.



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.

FIGURE 2. Endogenous Leadership Clarity

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 attention reacts to clarity is (approximately) the same for each leader via the term $\sigma_i(K_m - \sigma_i)$.[11] Returning to our analogy, local to the 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. Since $1/\kappa_i^2$ 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).

Following our earlier discussion we might breakdown $\underline{\sigma}_i^2$ into a term, common to all leaders, that reflects reliance of clarity on available sources of information technology, and another representing idiosyncratic noise due to variable communicative skills. When leaders are constrained to use a primitive source of communication, (we might suppose

[11]The shape is only approximately the same because K_m depends on σ_i in a different way for each leader. However, what really matters is the shape of $\sigma_i(K_m - \sigma_i)$ local to $K_m/2$. As the proof of Proposition 3 shows, $\partial K_m/\partial\sigma_i = 0$ when evaluated at $\sigma_i = K_m/2$.

that they are only able to speak to activists via a series of private audiences, so that $\underline{\sigma}_i^2$ is large) only the most audible leaders, whose speeches are easy to comprehend, 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 pre-requisite for successful leadership. Thus, whilst a leader in the mould 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 Coolidge lacked the physical capacities to succeed on the stump. His elevation to the presidency owed much to his use of radio broadcasting. In his own words, quoted in Cornwell Jr (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."

Even when technology allows leaders to communicate clearly, some will obfuscate. This rhetorical behavior of leaders 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), Li, Rosen, and Suen (2001), and Persico (2004), amongst others. In these "cheap-talk" scenarios an informed politician can never credibly reveal what she knows and is restricted to sending garbled messages due to her commonly understood policy bias. Our model lacks these strategic tensions; leaders have no inherent policy bias and can credibly reveal their signal of the true state of the world if they so wish. The strategic incentive to obfuscate arises nevertheless.

This incentive arises due to the nature of the "beauty contest" played by activists and the competitive tensions between leaders. Of course these key ingredients combine in other social situations where obfuscation is pervasive, and so our analysis provides more general insights. As one example, consider a variation on our party activists scenario: 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 temporal 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 provides an interpretation for why lucid prose is rarely heard from the pulpit.

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 since 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: a leader speaks more clearly when she has more to say. More subtly, however, a leader speaks more clearly when others have more to say. 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 policy closest to θ, that is as 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 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 and, therefore, enjoys influence. 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. Since (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

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

By inspection, the key determinant of a leader's success becomes her sense of direction.

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,*

$$\text{Party Welfare} = \bar{u} - \frac{2\pi}{\sum_{i=1}^{n}[1/\kappa_i^2]},$$

which increases with the leaders' combined sense of direction and the desire for party unity. Furthermore the variance $\mathrm{E}[(a_t - \theta)^2 \mid \theta]$ of activists' actions around the ideal policy grows with π, and so a greater desire for party unity improves the policy performance of the party.

Party welfare falls as π (activists' relative concern for policy) grows. This need not be surprising, since the policy component of an activist's loss function reacts to two sources of noise rather than one. More surprising, however, is the fact that the variance of activists' actions around the ideal policy falls as π shrinks. Recall that π indexes the extent to which activists care about choosing the ideal policy. Thus, paradoxically, activists become better at advocating the best policy as they care less about doing so.

Recognizing the endogenous quality of leadership corrects the intuition. When activists desire unity, leaders respond by speaking clearly. In so doing they generate common expectations about the party line. When such a common understanding emerges, the party acts more cohesively. But this is not the only effect since, when leaders speak more clearly, activists are able also to understand better their political 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.

Concluding Remarks

Few would deny that leadership plays an important role in influencing people's actions, and indeed there is experimental evidence that they do (Humphreys, Masters, and Sandbu, 2006; Güth, Levati, Sutter, and van der Heijden, 2004). Yet there has been no recent formal work which evaluates the influence of different leaders. This paper has attempted to fill this gap. Building on suggestions made by Levi (2006,p. 10) that "leadership—both of government and within civil society—provides the agency that coordinates the efforts of others," as well as earlier work on the 'focal' role of leadership actions, we explored a world with multiple potential leaders differentiated by their skill sets. We analyzed their ability to make policy judgements (sense of direction) and their ability to express their views (clarity of communication). To explore this issue we developed a model, loosely based the Keynes 'beauty contest" scenario, where a mass of party activists wish to (i) advocate the best policy and (ii) unify behind a common party line. To learn about their environment and to form expectations about the likely actions taken by fellow partisans, they listen to the speeches made by party leaders. We asked how much weight individuals place on the expressed views of their leaders, and according to the skill they possess.

In our model leadership speeches act are signals allowing players to coordinate, and they also influence their followers' beliefs. Leaders' speeches are signals in the statistical sense rather than the game-theoretic sense; the only strategic move of a leader is to change the precision of information that she transmits. When she does so her pay-off does not depend on the underlying state of the world, and so an activist has no need to filter the information he receives. Moreover, our model lacks the strategic tensions that arise would arise when leaders have inherent policy biases. When such biases are present, activists face a harder task in extracting information from a leader's speech.

Nevertheless, and despite these limitations, our model does provide a first-shot at assessing how a leader's influence is affected by her skill set. Treating a leader's skills as exogenous, we found that clarity of communication is relatively more important than sense of direction. Extending our model to allow activists to choose how much time to allocate to different leaders, so that clarity is endogenous, we found that a necessary condition for a leader to receive any attention is that she belongs to the party's clearest communicators. Moreover, if she is to monopolize the agenda then she must be the most coherent communicator. Analyzing how activists apportion their time amongst the best communicators (so that no single leader dominates proceedings) we found that they may pay more attention to those with (relatively) inferior communicative ability. Correspondingly, we found when that, when attention-seeking leaders choose their clarity, they obfuscate in order to retain their audience. In equilibrium all members of the leadership elite choose the same levels of clarity. Thus, and in contrast to our earlier result, a leader's influence is solely determined by her sense of direction.

We have also tackled the thorny issue of whether leaders can be both good and successful in a model which brings together the information aggregation properties of leadership and the political ambitions of leaders. In our framework, a good leader helps activists understand their environment and coordinate with each other, whilst a successful leader maximizes the attention paid to her. When leaders' skills are exogenous, good leadership and successful leadership are indistinguishable. But when activists choose the amount of attention they give to each leader (and leaders react by choosing the clarity of their communication) a leader's desire for attention, a key component of her political ambitions, drives a wedge between good and successful leadership: A leader increases her success by obfuscating her message; activists receive less information; and consequentially activists are less informed about their environment.

These welfare effects, due to the attention-seeking concerns of leaders, depend critically upon the weighting activists placed on their twin goals. Perhaps surprisingly, there is less obfuscation when activists place more emphasis on following the party line rather than pursuing the best policy. Our message is that when a party emphasizes unity, it provides leaders with the necessary incentives so that good and successful leadership coincide.

This key result is related to our assumption that leaders are attention-seekers. One justification for this focus is that the attention given to a leader is a component of her overall influence; a leader who cares about her influence would do well to maximize the attention paid to her views. Indeed in some cases (as for example when a leader receives undivided attention) attention-seeking and influence-maximizing coincide. Moreover, in most cases the behaviors associated with these motives are similar.[12] Whilst our results are robust to different motivational assumptions they also incorporate straightforward extensions of the model. For example, our results hold when we consider activists who have access to information sources other than leadership. In sum we hope that our model provides a small step in response to Levi's (2006, p. 11) claim that "still lacking is a model of the origins and means of ensuring good leadership."

[12]An influence-seeking leader would ideally wish to monopolize the agenda. Proposition 4 reveals that she must be the best communicator; but if her clarity is too great then activists also follow others. Hence, some obfuscation is necessary if a leader is to dictate. However, when clarity is chosen simultaneously by all leaders then we would expect to see greater clarity than in a pure attention-seeking game. To see this, begin with the equilibrium described in Proposition 5. A marginal increase in clarity has no effect on the attention paid to a leader. However, it does increase the weight placed on her speech and so increases her influence. Thus we expect to see more obfuscation when leaders seek attention rather than influence.

Unity and Conformity. If Footnote 1 we noted that the loss function $(a_t - \bar{a})^2$ might reflect a desire to conform rather than a preference for unity. A better measure of disunity might be $\int_0^1 (a_t - a_{t'})^2 \, dt'$. However, it is straightforward to confirm that

$$\int_0^1 (a_t - a_{t'})^2 \, dt' = (a_t - \bar{a})^2 + \int_0^1 (a_{t'} - \bar{a})^2 \, dt'.$$

Hence if we used $\int_0^1 (a_t - a_{t'})^2 \, dt'$ to measure disunity an activist's payoff would become

$$\tilde{u}_t \equiv \bar{u} - \pi(a_t - \theta)^2 - (1 - \pi) \int_0^1 (a_t - a_{t'})^2 \, dt' = u_t - (1 - \pi) \int_0^1 (a_{t'} - \bar{a})^2 \, dt'.$$

The final term is independent of a_t and hence is irrelevant to the decision-making of activist t, and so the behavior stemming from this revised specification would be observationally equivalent to that arising from the specification used in the main text.

One element of our analysis is, however, affected by the use of the revised "party disunity" loss function. When $\tilde{u}_t$ is an activist's payoff, party welfare becomes

$$\mathrm{E}[\tilde{u}_t] = \mathrm{E}[u_t] - (1 - \pi) \operatorname{var}[a_{t'} \mid \theta] = \mathrm{E}[u_t] - (1 - \pi) \sum_{i=1}^{n} w_i^2 \sigma_i^2.$$

The equilibrium advocacy strategy maximizes $\mathrm{E}[u_t]$ rather than $\mathrm{E}[\tilde{u}_t]$. What this means is that, from a welfare perspective, activists place too little weight on relatively clear leaders. (Of course, if we specified policy performance as our welfare measure, captured by the expected loss $\mathrm{E}[(a_t - \theta)^2]$, then we would reach the opposite conclusion.) □

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

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

Notice that 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. Since the loss function which 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,

$$\mathrm{E}[\bar{a} \mid \tilde{s}_t] = \int_0^1 \mathrm{E}[A(\tilde{s}_{t'}) \mid \tilde{s}_t] dt' = \mathrm{E}[A(\tilde{s}_{t'}) \mid \tilde{s}_t]. \qquad \square$$

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 $\mathrm{E}[\bar{a} \mid \tilde{s}_t] = w \cdot \mathrm{E}[s \mid \tilde{s}_t]$. The best reply $\mathrm{BR}[\tilde{s}_t \mid w]$ 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

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

Since an activist begins with a diffuse prior over θ (a non-diffuse 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 $\mathrm{E}[\theta \mid \tilde{s}_t] = b \cdot \tilde{s}_t$ for some $n \times 1$ vector b, and similarly $\mathrm{E}[s \mid \tilde{s}_t] = B\tilde{s}_t$ for some $n \times n$ inference matrix B. Hence

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

This is a linear strategy, which verifies the first claim of the lemma. 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, so 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 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,

$$B = \begin{bmatrix} \rho_1 & 0 & \dots & 0 \\ 0 & \rho_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \rho_n \end{bmatrix} + \frac{1}{\sum_{j=1}^{n} \psi_j} \begin{bmatrix} (1-\rho_1)\psi_1 & (1-\rho_1)\psi_2 & \dots & (1-\rho_1)\psi_n \\ (1-\rho_2)\psi_1 & (1-\rho_2)\psi_2 & \dots & (1-\rho_2)\psi_n \\ \vdots & \vdots & \ddots & \vdots \\ (1-\rho_n)\psi_1 & (1-\rho_n)\psi_2 & \dots & (1-\rho_n)\psi_n \end{bmatrix},$$

where ρ_i is the correlation coefficient defined in the text. (Full details of the Bayesian updating formulae are contained in a further not-for-publication 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 in the text prior to Proposition 2. It is straightforward to confirm that $I - (1-\pi)B'$ has full rank (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 the a common shift in all signals results in the same shift in activists' actions. (Solving for w explicitly confirms this.) □

Proof of Lemma 2. Local changes in the party line $\bar{a}$ have no net effect on the party:

$$\frac{\partial}{\partial \bar{a}} \int_0^1 u_{t'} \, dt' = -2 \int_0^1 (a_{t'} - \bar{a}) \, dt' = 0.$$ □

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 the solution directly (this approach is contained, for completeness, in our not-for-publication supplementary appendix) we instead build upon 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. The expression for $\hat{\psi}_i$ follows from simple algebra, and the comparative static claims regarding ψ_i and ρ_i follow by inspection. Taking logarithms and differentiating with respect to π we obtain

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

which yields the final claim of the proposition. □

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

Our exposition uses 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) $\bar{E}[\cdot] \equiv \int_0^1 \mathrm{E}_t[\cdot]\,dt$ 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 $\bar{\mathrm{E}}^{k+1}[\cdot] = \bar{\mathrm{E}}[\bar{\mathrm{E}}^k[\cdot]]$. (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 which 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* non-explosive *if* $\lim_{k\to\infty}\left[\alpha^k\,\mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]]\right] = 0$ *for any* $\alpha \in (0,1)$.

Clearly, α^k vanishes to zero exponentially as $k \to \infty$, and hence this condition is easily satisfied for many advocacy strategies. For instance, it is automatically satisfied if $A(\cdot)$ is bounded. A weaker sufficient condition is that $A(\cdot)$ is bounded by linear strategies: if we fix a linear strategy $w \cdot \tilde{s}_t$ and there exists 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 non-explosive.

Proposition. *Fix a policy advocacy strategy $A(\cdot)$ played by party members. If higher-order expectations of the party line are non-explosive 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 non-explosive 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 satisfies $\bar{a} = \pi\bar{\mathrm{E}}[\theta] + (1-\pi)\bar{\mathrm{E}}[A(\cdot)]$. A new best reply $\mathrm{BR}[\mathrm{BR}[A(\cdot)]] = \mathrm{BR}^2[A(\cdot)]$ to the party-wide play of $\mathrm{BR}[A(\cdot)]$ satisfies

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

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,

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

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

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

where $w \cdot \tilde{s}_t$ is the unique linear (Proposition 1). It remains to consider the second component of $\mathrm{BR}^k[A(\cdot)]$. If the higher-order expectations generated by $A(\cdot)$ are non-explosive then this second component 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 any non-explosive policy strategy $A(\cdot)$ a tatonnement in which party activists update their behavior via sequence of best replies, generating the sequence $\mathrm{BR}^k[A(\cdot)]$, converges to the unique linear equilibrium. Furthermore, since any equilibrium policy advocacy strategy satisfies $\mathrm{BR}^k[A(\cdot)] = A(\cdot)$ for all k, we can be assured that any non-linear 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. However, they implicitly assumed that $(1-\pi)^{k+1}\, \mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]] \to 0$. As

Angeletos and Pavan (2007) noted, this means that their analysis was not quite watertight. Indeed, it is possible to find strategies which generate explosive higher-order expectations of the party line. Consider, for instance, a world in which $n = 1$, $\kappa_1^2 = 0$, and $\sigma_1^2 = \sigma^2 > 0$. (The model in the text specifies $\kappa_i^2 > 0$ but can be modified appropriately.) Since there is only one leader, we drop the "i" subscript. Now consider the advocacy strategy $A(\tilde{s}_t) = e^{\beta \tilde{s}_t}$ for some $\beta > 0$. Straightforward calculations confirm that $\mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]] = e^{k\beta^2\sigma^2 + \beta\tilde{s}_t}$, and so $(1-\pi)^{k+1}\,\mathrm{E}_t[\bar{\mathrm{E}}^k[A(\cdot)]]$ diverges if $(1-\pi)e^{\beta^2\sigma^2} > 1$. □

Proof of Proposition 3. A strict Bayesian Nash equilibrium entails the play of a unique strict best reply by each player-type, and this best reply is independent of the player's label, since each player is negligible. Hence any strict equilbrium must be symmetric. Any externalities arising from a player's behavior feed via the party line $\bar{a}$. As argued in the text, positive and negative externalities cancel yielding a zero net effect at the margin. Hence, the equilibrium maximizes party welfare. Now,

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

and so the equilibrium attention levels must solve

$$\max_{x \in \mathbb{R}_+^n} \sum_{i=1}^{n} \frac{1}{\pi\kappa_i^2 + [\sigma_i^2/x_i]} \quad \text{subject to} \quad \sum_{i=1}^{n} x_i \leq 1.$$

The constraint function is convex. The objective is strictly concave, since its ith component is concave in x_i. To see this, differentiate $\hat{\psi}_i$ with respect to x_i, to obtain

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

which is positive and strictly decreasing in x_i. Since the objective function is strictly concave, there is a unique solution and hence a unique equilibrium. The usual Kuhn-Tucker conditions are both necessary and sufficient. Introducing the Lagrange multiplier $\lambda > 0$ for the attention-span constraint (the constraint binds since welfare is strictly increasing in attention and so the multiplier is strictly positive) 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 attention must speak with clarity exceeding λ. It follows that if a leader is ignored, (that is $x_i = 0$), then it must be the case that

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

and so the clarity of communication of such a 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 $i \leq m$ for some $m \in \{1, \dots, n\}$. For this elite of m leaders,

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

To find the value of K (and hence the value $\lambda = 1/K^2$ of the Lagrange multiplier) we sum over the m-strong elite and set $\sum_{i=1}^{m} x_i = 1$. Hence

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

which solves for K given in the proposition. We need only check that the sense of direction of any leader $i > m$ falls below λ, or equivalently $\sigma_i \geq K$. If so, and so long as $\sigma_i < K$ for all $i \leq m$, then we have a solution. If $\sigma_{m+1} < K$ then we expand the elite to size $m+1$. We continue doing this until we find the equilibrium elite size m. Finally, we turn to the comparative-static claims. Differentiate x_i with respect to σ_i to obtain:

$$\frac{\partial x_i}{\partial \sigma_i} = \frac{1}{\pi\kappa_i^2}\left[K - 2\sigma_i + \sigma_i \frac{\partial K}{\partial \sigma_i}\right].$$

Taking K from the proposition and differentiating with respect to σ_i we obtain

$$\frac{\partial K}{\partial \sigma_i} = \frac{(2\sigma_i/\kappa_i^2)\left(\sum_{j=1}^{m}[\sigma_j/\kappa_j^2]\right) - (1/\kappa_i^2)\left(\pi + \sum_{j=1}^{m}[\sigma_j^2/\kappa_j^2]\right)}{\left(\sum_{j=1}^{m}[\sigma_j/\kappa_j^2]\right)^2} = \frac{2\sigma_i - K}{\kappa_i^2 \sum_{j=1}^{m}[\sigma_j/\kappa_j^2]},$$

which upon substitution back into $\partial x_i/\partial \sigma_i$ yields:

$$\frac{\partial x_i}{\partial \sigma_i} = \frac{K - 2\sigma_i}{\pi\kappa_i^2}\left[1 - \frac{\sigma_i}{\kappa_i^2 \sum_{j=1}^{m}[\sigma_j/\kappa_j^2]}\right] > 0 \quad \Leftrightarrow \quad \sigma_i < \frac{K}{2},$$

since the bracketed term is strictly positive. For the final claim, notice that (other things equal) K is increasing in π. An increase in π makes it harder to satisfy $\sigma_i > K$, and so m must increase with π. □

Proof of Proposition 4. For attention to be focused on a elite of size m, the clarity of $i > m$ must satisfy $\sigma_i^2 \geq K^2$ where K is from Proposition 3. For the special case of $m = 1$,

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

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

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

This is increasing in σ_1^2 and hence K^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 Proposition 5. x_i is increasing in σ_i for $\sigma_i < K/2$ and decreasing for $\sigma_i > K/2$ and so is maximized if and only if $\sigma_i = K/2$. To find the effect of a parameter on a best reply we find its effect on K: if a parameter increases K then leader i responds by raising σ_i until $\sigma_i = K/2$ once more, reducing her clarity. We observe that K increases with π, and so clarity increases with the desire for party unity. Differentiating K with respect to $1/\kappa_i^2$,

$$\frac{\partial K}{\partial[1/\kappa_i^2]} = \frac{\sigma_i^2\left(\sum_{j=1}^m[\sigma_j/\kappa_j^2]\right) - \sigma_i\left(\pi + \sum_{j=1}^m[\sigma_j^2/\kappa_j^2]\right)}{\left(\sum_{j=1}^m[\sigma_j/\kappa_j^2]\right)^2} = \frac{\sigma_i^2 - \sigma_i K}{\sum_{j=1}^m[\sigma_j/\kappa_j^2]},$$

which is positive if and only if $\sigma_i > K$. Of course, if $\sigma_i > K$ then $x_i = 0$ and leader i is not entered in the formula for K. Thus, if leader i attracts any attention, so that $\sigma_i < K$, then K is decreasing in her sense of direction. To find the equilibrium of the attention-seeking game we set $\sigma_i = \sigma = K/2$ for all i. Substituting $K = 2\sigma$ into the definition of K,

$$2\sigma = \frac{\pi + \sigma^2\sum_{j=1}^m[1/\kappa_j^2]}{\sigma\sum_{j=1}^m[1/\kappa_j^2]} \quad\Leftrightarrow\quad 2\sigma^2 = \frac{\pi}{\sum_{j=1}^m[1/\kappa_j^2]} + \sigma^2,$$

yielding the solution from the proposition. The final claim follows by inspection. □

Proof of Proposition 6. Party welfare is $\bar{u} - 1/\sum_{i=1}^n \hat{\psi}_i$. Substituting in the expression for $\hat{\psi}_i$ given in the text (obtained by simple algebra) yields the welfare stated in the proposition. The comparative-static claims follow straightforwardly. Regarding policy performance,

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

where the 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. To construct this figure we computed the reaction function for a leader i. Note that a leader i chooses her clarity of optimally when $\sigma_i = K/2$. Hence

$$2\sigma_i = K = \frac{\pi + \sum_{j=1}^m[\sigma_j^2/\kappa_j^2]}{\sum_{j=1}^m[\sigma_j/\kappa_j^2]} = \frac{\pi\kappa_i^2 + A + \sigma_i^2}{B + \sigma_i} \quad\Leftrightarrow\quad \sigma_i^2 + B\sigma_i - (A + \pi\kappa_i^2) = 0,$$

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 these 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, F., S. MORRIS, AND H. S. SHIN (2006): "Beauty Contests and Iterated Expectations in Asset Markets," *Review of Financial Studies*, 19(3), 719–52.

ALVAREZ, R. M., AND C. H. FRANKLIN (1994): "Uncertainty and Political Perceptions," *Journal of Politics*, 56(3), 671–688.

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

ANGELETOS, G.-M., C. HELLWIG, AND A. PAVAN (2006): "Dynamic Global Games of Regime Change: Learning, Multiplicity, and Timing of Attacks," *Econometrica*, forthcoming.

ANGELETOS, G.-M., AND A. PAVAN (2004): "Transparency of Information and Coordination in Economies with Investment Complementarities," *American Economic Review: AEA Papers and Proceedings*, 94(2), 91–8.

——— (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103–42.

ARAGONÉS, E., AND Z. NEEMAN (2000): "Strategic Ambiguity in Electoral Competition," *Journal of Theoretical Politics*, 12(2), 183–204.

BARTLETT, M. S. (1951): "An Inverse Matrix Adjustment Arising in Dicriminant Analysis," *Annals of Mathematical Statistics*, 22(1), 107–111.

BRAND, H. W. (2005): *Andrew Jackson: His Life and Times*. Doubleday.

BUENO DE MESQUITA, B., J. D. MORROW, R. M. SIVERSON, AND A. SMITH (2002): "Political Institutions, Policy Choice, and the Survival of Leaders," *British Journal of Political Science*, 32(4), 559–90.

BUENO DE MESQUITA, B., A. SMITH, R. M. SIVERSON, AND J. D. MORROW (2003): *The Logic of Political Survival*. MIT Press, Cambridge, MA.

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

CARWARDINE, R. J. (2003): *Lincoln*. Pearson Longman, Harlow.

CORNWELL JR, E. E. (1957): "Coleridge and Presidential Leadership," *Public Opinion Quarterly*, 21(2), 265–78.

CRAWFORD, V. P., AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–51.

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

EDMOND, C. (2005): "Information Manipulation, Coordination, and Regime Change," *Working Paper*.

ELLIS, J. J. (2005): *His Excellency: George Washington*. Faber, London.

FIORINA, M. P., AND K. A. SHEPSLE (1989): "Formal Theories of Leadership: Agents, Agenda-Setters, and Entrepreneurs," in *Leadership and Politics: New Perspectives in Political Science*, ed. by B. D. Jones, pp. 17–41. University Press Kansas, Lawrence.

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

GILLIGAN, T. W., AND K. KREHBIEL (1987): "Collective Decisionmaking and Standing Committees: An Informational Rationale for Restrictive Ammendment Procedures," *Journal of Law, Economics, and Organization*, 3(2), 287–335.

GÜTH, W., M. V. LEVATI, M. SUTTER, AND E. VAN DER HEIJDEN (2004): "Leadership and Cooperation in Public Goods Experiments," *Max Planck Institute of Economics*.

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

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

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

KROSZNER, R. S., AND T. STRATMANN (2005): "Corporate Campaign Contributions, Repeat Giving and the Rewards to Legislator Reputation," *Journal of Law and Economics*, 48(2), 41–71.

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

LI, H., S. ROSEN, AND W. SUEN (2001): "Conflicts and Common Interests in Committees," *American Economic Review*, 91(3), 1478–97.

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

MORRIS, S., AND H. S. SHIN (2002): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–34.

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

PAGE, B. I. (1996): "The Theory of Political Ambiguity," *American Political Science Review*, 70(3), 742–52.

PERSICO, N. (2004): "Committee Design with Endogenous Information," *Review of Economic Studies*, 71, 165–91.

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

SHEPSLE, K. 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.

——— (1972a): "Parties, Voters and the Risk Environment: A Mathematical Treatment of Electoral Competition under Uncertainty," in *Probability Models of Collective Decisison-Making*, ed. by R. G. Niemi, and H. F. Weisberg, pp. 273–97. Charles E. Merrill, Columbus, Ohio.

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

SIGELMAN, L., AND E. H. BUELL, JR. (2004): "Avoidance or Engagement? Issue Convergence in US Presidential Campaigns, 1960–2000," *American Journal of Political Science*, 48(4), 650–61.

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

Bayesian Inference from Leaders' Speeches

Suppose that an activist begins with the prior $\theta \sim N(\mu, \xi^2)$. (The model in the text corresponds to the limiting case of $\xi^2 \to \infty$, so that the prior is improperly diffuse. We can also incorporate a prior by introducing an extra "leader" satisfying $\kappa_0^2 = \xi^2$ and $\sigma_0^2 = 0$.) Conditional on θ, the n leaders observe independent signals satisfying $s_i \,|\, \theta \sim N(\theta, \kappa_i^2)$. Unconditionally, the signals are joint normally distributed, satisfying $\mathrm{var}[s_i] = \xi^2 + \kappa_i^2$ and $\mathrm{cov}[s_i, s_j] = \xi^2$ for $i \neq j$. That is,

$$s \sim N(\mu e, \Omega) \quad \text{where} \quad \Omega = \begin{bmatrix} \xi^2 + \kappa_1^2 & \xi^2 & \dots & \xi^2 \\ \xi^2 & \xi^2 + \kappa_2^2 & \dots & \xi^2 \\ \vdots & \vdots & \ddots & \vdots \\ \xi^2 & \xi^2 & \dots & \xi^2 + \kappa_n^2 \end{bmatrix}.$$

and where e is the $n \times 1$ unit vector, so that μe is the $n \times 1$ vector with each entry equal to μ. A party activist $t \in [0, 1]$ observes a vector of n signals-of-signals. That is,

$$\tilde{s}_t \,|\, s \sim N(s, \Sigma) \quad \text{where} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n^2 \end{bmatrix}.$$

Here we focus our attention on an activist's beliefs about the leaders' underlying signals. Given the normal specification, Bayesian updating leads to $s \,|\, \tilde{s}_t \sim N(\mathrm{E}[s \,|\, \tilde{s}_t], \mathrm{var}[s \,|\, \tilde{s}_t])$ where

$$\mathrm{E}[s \,|\, \tilde{s}_t] = [\Omega^{-1} + \Sigma^{-1}]^{-1}[\Omega^{-1}(\mu e) + \Sigma^{-1}\tilde{s}_t] \quad \text{and} \quad \mathrm{var}[s \,|\, \tilde{s}_t] = [\Omega^{-1} + \Sigma^{-1}]^{-1}.$$

To evaluate the various components, we begin by observing that

$$\Sigma^{-1} = \begin{bmatrix} \gamma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \gamma_n \end{bmatrix} \quad \text{where} \quad \gamma_i \equiv \frac{1}{\sigma_i^2}.$$

Next we evaluate Ω^{-1} by applying the Sherman-Morrison formula for inverting rank-one updates to a non-singular matrix (Bartlett, 1951). To see this formula in action, consider an $n \times n$ matrix M with full rank and two $n \times 1$ vectors x and y. The $n \times n$ dyadic product xy' has unit rank, and $M + xy'$ is a rank-one update to the non-singular matrix M. Assuming that M is symmetric for simplicity (as it is for our application) the matrix-inversion formula is:

$$[M + xy']^{-1} = M^{-1} - \frac{M^{-1}xy'M^{-1}}{1 + x'M^{-1}y}.$$

To apply the formula, note that $\Omega = K + (\xi e)(\xi e)'$ where e is the $n \times 1$ unit vector, as before, and

$$K = \begin{bmatrix} \kappa_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \kappa_n^2 \end{bmatrix}.$$

Hence to apply the formula we set $M = K$ (which is symmetric) and $x = y = \xi e$. This yields:

$$\Omega^{-1} = K^{-1} - \frac{\xi^2 (K^{-1}e)(K^{-1}e)'}{1 + \xi^2 e' K^{-1} e}.$$

The diagonal matrix K is straightforward to invert, yielding

$$K^{-1} = \begin{bmatrix} \delta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \delta_n \end{bmatrix} \quad \text{where} \quad \delta_i \equiv \frac{1}{\kappa_i^2}.$$

Let us write δ for the $n \times 1$ vector with ith element δ_i and $|\delta| \equiv \sum_{i=1}^n \delta_i$. Using this notation, $K^{-1}e = \delta$ and $e'K^{-1}e = |\delta|$, and so

$$\Omega^{-1} = K^{-1} - \frac{\xi^2 \delta \delta'}{1 + \xi^2 |\delta|}.$$

Next we combine our expressions for Ω^{-1} and Σ^{-1} in order to evaluate

$$[\Omega^{-1} + \Sigma^{-1}]^{-1} = \left[\Sigma^{-1} + K^{-1} - \frac{\xi^2 \delta \delta'}{1 + \xi^2 |\delta|} \right]^{-1}.$$

The matrix $\Omega^{-1} + \Sigma^{-1}$ is, once again, a rank-one update of a non-singular matrix, taking the Sherman-Morrison form with $M = \Sigma^{-1} + K^{-1}$ and $x = -y = \xi\delta/\sqrt{1 + \xi^2|\delta|}$. Hence,

$$[\Omega^{-1} + \Sigma^{-1}]^{-1} = [\Sigma^{-1} + K^{-1}]^{-1} + \frac{\xi^2 ([\Sigma^{-1} + K^{-1}]^{-1}\delta)([\Sigma^{-1} + K^{-1}]^{-1}\delta)'}{1 + \xi^2|\delta| - \xi^2 \delta' [\Sigma^{-1} + K^{-1}]^{-1} \delta}.$$

The matrix inversion straightforwardly yields

$$[\Sigma^{-1} + K^{-1}]^{-1} = \begin{bmatrix} 1/(\gamma_1 + \delta_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1/(\gamma_n + \delta_n) \end{bmatrix}.$$

Next, it is convenient to write $\hat{\delta} \equiv [\Sigma^{-1} + K^{-1}]^{-1}\delta$ for the $n \times 1$ vector with ith element $\hat{\delta}_i \equiv \delta_i/(\delta_i + \gamma_i)$. Using this notation, we obtain

$$[\Omega^{-1} + \Sigma^{-1}]^{-1} = [\Sigma^{-1} + K^{-1}]^{-1} + \frac{\xi^2 \hat{\delta} \hat{\delta}'}{1 + \xi^2 (|\delta| - \delta' \hat{\delta})}.$$

Evaluating the denominator of the fraction, notice that

$$|\delta| - \delta'\hat{\delta} = \sum_{i=1}^n \delta_i - \sum_{i=1}^n \frac{\delta_i^2}{\delta_i + \gamma_i} = \sum_{i=1}^n \frac{\delta_i \gamma_i}{\delta_i + \gamma_i} = \sum_{i=1}^n \frac{1}{\kappa_i^2 + \sigma_i^2} = |\psi|,$$

where ψ is the $n \times 1$ vector with ith element ψ_i, and where ψ_i is defined in the paper as quality of information drawn from the ith leader; $|\psi|$ is simply shorthand for $\sum_{i=1}^n \psi_i$. Hence,

$$[\Omega^{-1} + \Sigma^{-1}]^{-1} = [\Sigma^{-1} + K^{-1}]^{-1} + \frac{\xi^2 \hat{\delta} \hat{\delta}'}{1 + \xi^2 |\psi|}.$$

This is the covariance matrix of an activist's posterior beliefs about s. Allowing the prior to disappear, so that $\xi^2 \to \infty$, this covariance matrix becomes

$$\lim_{\xi^2\to\infty} \mathrm{var}[s \mid \tilde{s}_t] = [\Sigma^{-1} + K^{-1}]^{-1} + \frac{\hat{\delta}\hat{\delta}'}{|\psi|}.$$

Of more direct interest to us is the posterior mean. Recall that this satisfies

$$\mathrm{E}[s \mid \tilde{s}_t] = [\Omega^{-1} + \Sigma^{-1}]^{-1}[\Omega^{-1}(\mu e) + \Sigma^{-1}\tilde{s}_t].$$

Let us evaluate the term involving the prior mean μ. Allowing the prior variance to explode,

$$\Omega^{-1}e = \left[K^{-1} - \frac{\xi^2\delta\delta'}{1+\xi^2|\delta|}\right]e = \left(1 - \frac{\xi^2|\delta|}{1+\xi^2|\delta|}\right)\delta \longrightarrow 0 \quad \text{as} \quad \xi^2 \to \infty.$$

Hence, as the prior vanishes, we obtain

$$\lim_{\xi^2\to\infty} \mathrm{E}[s \mid \tilde{s}_t] = B\tilde{s}_t \quad \text{where} \quad B = \left[[\Sigma^{-1} + K^{-1}]^{-1} + \frac{\hat{\delta}\hat{\delta}'}{|\psi|}\right]\Sigma^{-1}.$$

Plugging in the various elements, we obtain the explicit expression

$$B = \begin{bmatrix} \frac{\gamma_1}{\gamma_1+\delta_1} & 0 & \dots & 0 \\ 0 & \frac{\gamma_2}{\gamma_2+\delta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\gamma_n}{\gamma_n+\delta_n} \end{bmatrix} + \frac{1}{|\psi|}\begin{bmatrix} \frac{\gamma_1\delta_1^2}{(\gamma_1+\delta_1)^2} & \frac{\gamma_2\delta_1\delta_2}{(\gamma_1+\delta_1)(\gamma_2+\delta_2)} & \dots & \frac{\gamma_n\delta_1\delta_n}{(\gamma_1+\delta_1)(\gamma_n+\delta_n)} \\ \frac{\gamma_1\delta_1\delta_2}{(\gamma_1+\delta_1)(\gamma_2+\delta_2)} & \frac{\gamma_2\delta_2^2}{(\gamma_2+\delta_2)^2} & \dots & \frac{\gamma_n\delta_2\delta_n}{(\gamma_2+\delta_2)(\gamma_n+\delta_n)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\gamma_1\delta_n\delta_1}{(\gamma_n+\delta_n)(\gamma_1+\delta_1)} & \frac{\gamma_2\delta_n\delta_2}{(\gamma_n+\delta_n)(\gamma_2+\delta_2)} & \dots & \frac{\gamma_n\delta_n^2}{(\gamma_n+\delta_n)^2} \end{bmatrix}.$$

We can simplify this inference matrix by noting that

$$\frac{\gamma_i\delta_i}{\gamma_i+\delta_i} = \frac{1}{\kappa_i^2+\sigma_i^2} = \psi_i \quad \text{and} \quad \frac{\gamma_i}{\gamma_i+\delta_i} = \frac{\kappa_i^2}{\kappa_i^2+\sigma_i^2} = \rho_i$$

where ρ_i is the correlation coefficient defined in the paper. This means that

$$B = \begin{bmatrix} \rho_1 & 0 & \dots & 0 \\ 0 & \rho_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \rho_n \end{bmatrix} + \frac{1}{|\psi|}\begin{bmatrix} (1-\rho_1)\psi_1 & (1-\rho_1)\psi_2 & \dots & (1-\rho_1)\psi_n \\ (1-\rho_2)\psi_1 & (1-\rho_2)\psi_2 & \dots & (1-\rho_2)\psi_n \\ \vdots & \vdots & \ddots & \vdots \\ (1-\rho_n)\psi_1 & (1-\rho_n)\psi_2 & \dots & (1-\rho_n)\psi_n \end{bmatrix}.$$

Plugging everything in, this means that $\lim_{\xi^2\to\infty} \mathrm{E}[s_i \mid \tilde{s}_t] = \rho_i\tilde{s}_{it} + (1-\rho_i)\,\mathrm{E}[\theta \mid \tilde{s}_t]$.

Alternative Proof of Proposition 1

The unique linear equilibrium satisfies $w = \pi[I - (1-\pi)B']^{-1}b$. Now,

$$I-(1-\pi)B' = \begin{bmatrix} 1-(1-\pi)\rho_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1-(1-\pi)\rho_n \end{bmatrix} - \frac{(1-\pi)}{|\psi|}\begin{bmatrix} (1-\rho_1)\psi_1 & \dots & (1-\rho_n)\psi_1 \\ \vdots & \ddots & \vdots \\ (1-\rho_1)\psi_n & \dots & (1-\rho_n)\psi_n \end{bmatrix}.$$

Recalling that $\hat{\delta}_i = \delta_i/(\delta_i + \gamma_i) = 1 - \rho_i$, the final matrix is the dyadic product of ψ and $\hat{\delta}$. Also,

$$1 - (1-\pi)\rho_i = 1 - \frac{(1-\pi)\kappa_i^2}{\kappa_i^2 + \sigma_i^2} = \frac{\pi\kappa_i^2 + \sigma_i^2}{\kappa_i^2 + \sigma_i^2} = \frac{\psi_i}{\hat{\psi}_i}$$

where $\hat{\psi}_i$ is the influence index described in the text. Hence,

$$I - (1-\pi)B' = \tilde{\Psi} - \frac{(1-\pi)\psi\hat{\delta}'}{|\psi|} \quad \text{where} \quad \tilde{\Psi} = \begin{bmatrix} \psi_1/\hat{\psi}_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \psi_n/\hat{\psi}_n \end{bmatrix}.$$

Writing $|\hat{\psi}| \equiv \sum_{i=1}^n \hat{\psi}_i$ the solution for w given in the text is $w = \hat{\psi}/|\hat{\psi}|$. Now,

$$\begin{aligned}[I - (1-\pi)B']w = \frac{1}{|\hat{\psi}|}\left[\tilde{\Psi} - \frac{(1-\pi)\psi\hat{\delta}'}{|\psi|}\right]\hat{\psi} &= \left[\frac{1}{|\hat{\psi}|} - \frac{(1-\pi)\hat{\delta}'\hat{\psi}}{|\psi||\hat{\psi}|}\right]\psi \\ &= \left[\frac{|\psi| - (1-\pi)\hat{\delta}'\hat{\psi}}{|\hat{\psi}|}\right]\frac{\psi}{|\psi|} = \pi b.\end{aligned}$$

The second equality is obtained from observing that $\tilde{\Psi}\hat{\psi} = \psi$. The final equality is obtained by observing that $b = \psi/|\psi|$ and

$$|\psi| - (1-\pi)\hat{\delta}'\hat{\psi} = \sum_{i=1}^n \Big(\psi_i - (1-\pi)\hat{\psi}_i(1-\rho_i)\Big) = \sum_{i=1}^n \psi_i\left(1 - \frac{(1-\pi)(1-\rho_i)}{1-(1-\pi)\rho_i}\right) = \pi|\hat{\psi}|.$$

Hence $[I - (1-\pi)B']w = \pi b$, which yields a "brute force" proof of Proposition 1.