

# Political Science and Political Economy Working Paper

Department of Government

London School of Economics

No. 7/2006

## Leading the Party: Coordination, Direction, and Communication

Torun Dewan (LSE)

David P. Myatt (Oxford)



THE LONDON SCHOOL  
OF ECONOMICS AND  
POLITICAL SCIENCE ■

# Leading the Party: Coordination, Direction, and Communication

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

*Discussion Paper. September 14, 2006.*<sup>1</sup>

**Abstract.** Party activists face a coordination problem: a critical mass (a *barrier to coordination*) must advocate a single policy alternative if the party is to succeed. The *need for direction* is the degree to which the merits of the alternatives respond to the underlying mood of the party. An individual's ability to assess the mood is his *sense of direction*. These factors combine to form an index of both the desirability and the feasibility of leadership: we call this index *Michels' Ratio*. A sovereign party conference gives way to leadership by an individual or oligarchy if and only if Michels' Ratio is sufficiently high. Leadership enhances the clarity of intra-party communication, but weakens the response of policy choices to the party's mood.

In her recent Presidential Address to the American Political Science Association, Margaret Levi (2006) called for a new theory of government. A central ingredient in her proposed recipe was leadership; she argued (p. 13) that "human agency, through leadership, belief reformation, preference formation, and widespread constituent support, provides the yeast, the missing ingredient of a dynamic theory of effective government." Few would deny that leadership is an essential feature of political organization, and yet (Levi, 2006, p. 11) "still lacking is a model of the origins and means of ensuring good leadership."

We take a small step in response to Levi's call. Building on her suggestion that (p. 10) "leadership—both of government and within civil society—provides the agency that coordinates the efforts of others" we develop a formal model in which the direction provided by leadership helps to coordinate the actions of a mass. We ask: is such direction best provided by one, a few, or the many? These institutional forms correspond to *de facto* dictatorship, oligarchy, and pure democracy. While an answer to this first question reveals the relative desirability of these institutional forms, we must also consider their feasibility: when will members of a democratic body voluntarily follow the lead taken by either a single individual or an elite subset of their membership?

---

<sup>1</sup>We thank participants at the August 2006 APSA meetings for their helpful comments and queries.

The “Iron Law of Oligarchy” emerging from Michels’ (1915) study of labor movements and political parties offers unequivocal answers to these two questions. Michels realized that success might require the effective coordination of an activist base, arguing that the

“importance and influence of the working class are directly proportional to its numerical strength. But for the representation of that numerical strength organization and coordination are indispensable.” (Michels, 1915, p. 19)

This view stemmed from Michels’ study of the German Socialist Party.<sup>2</sup> Whilst nominally adhering to the principle of conference sovereignty, key decisions were taken within the party’s *fraktion* meetings; conference rallied behind its leadership. Michels’ argued that an activist mass would voluntarily renounce its democratic rights and defer to an oligarchy.<sup>3</sup>

To provide our own answers, we model the coordination problem faced by an activist mass; this is a stylized representation of a party’s membership. Specifically, each party member must advocate one of two policy alternatives *A* and *B*. A policy succeeds if and only if a critical fraction of the party (we call this a *barrier to coordination*) supports it; an uncoordinated party splits and fails. Party members would like the best policy to succeed. A stylized interpretation of the mid-1990s reform of the British Labour party helps to fix ideas. Policy *A* would be the adoption of Tony Blair’s “New Labour” program, while the alternative policy *B* would be the retention of “old” Labour ideals. Whatever the merits of these alternatives, a unified party could challenge for power; in contrast, a split may well have relegated the party to another term in the wilderness.

Given their common objective, one mode of behavior seems obvious: all activists should advocate the best policy. This is difficult when the identity of the “best policy” is uncertain. This will be so when the merits of policies depend upon the political situation in which the party operates, since party members’ assessment of this situation may differ. Here leadership may help since, as Levi (2006) suggested, “leadership ... provides the learning environment that enables individuals to transform or revise beliefs.” However, other institutions, such as a party conference, might also provide a learning environment. Institutions differ in how they allow actors to achieve a common understanding. The need for such an understanding is most pressing when the merits of policies react strongly to the underlying situation; that is, when there is a strong *need for direction*.

<sup>2</sup>The Socialist Party was a forerunner of the German Social Democratic Party. It helped to provide political organization to the largest mass labor movement in 19th century Europe. For an application of Michels’ work to British political parties see McKenzie (1963), and for a critical review see Hands (1971).

<sup>3</sup>Moving further, Michels claimed that elites develop to pursue goals that betray the original movement. Modern interpretations have highlighted the similarities between this dynamic and the process of agency drift. This being so, some have suggested that *ex ante* control mechanisms can counteract the tendency of leadership to drift (Lupia and McCubbins, 2000). Koelble (1996), for example, in using the arguments from the literature on Congress, suggested that “party activists prefer strict monitoring of the actions of the representatives, mechanisms for screening and selection which force the agent to report and act as delegate rather than as a trustee for the party.” We do not address the agency issue, focusing instead on the conditions under which the sovereignty of party conference will give way to a leadership elite.

It is instructive to begin with an environment in which leadership is absent. We imagine activists gathering to discuss policy. At such a conference, an activist attends meetings, listens to opinions, engages in conversations, and eavesdrops on others: he receives a private signal of the party's mood, and the signal's precision is his *sense of direction*. Differences in signal realizations (different activists hear different things) can generate disagreement over the desirability of policies. So, whilst a conference helps to provide direction, the private nature of opinion formation can prevent party members from forming a common view of the party's goals; there is a lack of clarity in the party's communication.

Formally, the conference leads to a game played by party members. An activist bases his action ( $A$  or  $B$ ) on his assessment (a private signal) of the party's mood. Contemplating his decision, the activist will recognize that his advocacy counts only when he is pivotal to a policy's success. Conditional on his information, he will consider the relative likelihood of his pivotality for policy  $A$  versus policy  $B$ . As his assessment moves in favor of  $A$ , this relative likelihood grows. His optimal strategy will be a threshold rule: advocate  $A$  rather than  $B$  if and only if the perceived mood of the party is sufficiently favorable to  $A$ . We identify a unique equilibrium in which all party members use the same threshold rule.

An advantage of the conference is that the actions of the mass respond to the need for direction. When policy  $A$  is far more desirable than  $B$ , the aggregate mood of the party will reflect this. Activists' assessment of this mood will point toward  $A$ , and so (given the threshold rule) policy  $A$  will succeed. A disadvantage is that the party may fail to coordinate. If the policies are evenly matched, so that there is a mood of indifference throughout the party, then the private signals of different activists may well point in different directions. When activists advocate different policies the party splits.

Given that a split is undesirable, can party members coordinate on the same policy irrespective of its merits? For instance, they might all back policy  $B$  unless their private signals reveal that policy  $A$  is extremely desirable. Such a bias toward  $B$  might be self-reinforcing; an activist will reason that should he be pivotal to the success of a policy, then he is more likely to be pivotal for  $B$ . We refer to this effect as positive feedback.

This is not, however, the only effect we need to consider. When activists are biased toward  $B$ , the support for policy  $A$  achieves the critical mass required for success only if its true merits are exceptional. Contemplating this, an individual activist realizes that if he is pivotal to the success of  $A$ , then policy  $A$  must have a critical mass of support, and so must be extremely desirable. Thus, conditional on his pivotality (the situation in which advocacy matters) an individual activist has a strong preference for  $A$ . This negative-feedback effect is analogous to the swing voter's curse of Feddersen and Pesendorfer (1996): an *activist's curse* prompts a switch away from policies to which others are biased.

The activist's curse also moves the party away from safe policies. Suppose that  $A$  represents a dramatic shift from the status quo, so that a greater fraction of the party needs to

coordinate if radical reform is to succeed; it is more ambitious, since it must vault a high barrier to coordination. Recycling the logic above, an activist can be pivotal for  $A$  only when its merits are very strong, since only then will party-wide support for  $A$  achieve a critical mass. Once again, conditional on his pivotality, an individual will have a preference for  $A$ . The implication is that activists will pursue lofty goals with extra zeal.

Bringing these observations together, a conference is desirable because it aggregates the opinions of party members and so successful policies respond to the need for direction. Nevertheless, it brings the risk of mis-coordination, and the activist's curse biases party members toward ambitious goals. In some circumstances, then, dissatisfaction might lead to what Michels (1915, pp. 35–41) described as a “need for leadership felt by the mass.” If a leader could dictate the policy choice, then mis-coordination might be avoided. The penalty would be that the leader might not reflect the true mood of the party. To assess the need for leadership, then, we pose the following question: if a leading activist could choose between (i) choosing the policy herself, and (ii) allowing the policy to emerge from a party conference, then what would she do?

The answer incorporates three factors. Firstly, a policy's success requires a critical mass of support. The sum of the critical-mass measures for the alternatives indexes the barriers to coordination. Secondly, the strength of the relationship between the policies' merits and the true underlying party mood indexes the need for direction. Finally, the ability to discover the best policy relies upon the accuracy of an activist's assessment of the party mood. This accuracy is an individual's sense of direction. Our analysis reveals that it is natural to combine these three elements into a single measure  $R$ . We call it *Michels' Ratio*:

$$R = \frac{\text{Barriers to Coordination} \times \text{Sense of Direction}}{\text{Need for Direction}}.$$

We find that a leading party member would prefer to dictate policy if and only if Michels' Ratio is sufficiently large. Direct choice by the leader is desirable when barriers to coordination are high because the risk of mis-coordination is large. She is better able to select the best policy when her sense of direction is particularly sharp. However, when the need for direction is particularly pressing (so that the policies' merits depend strongly on the party's mood) then conference becomes more desirable, precisely because its outcome reflects the party's mood. Drawing these elements together, we show that the criterion for the mass desirability of leadership is  $R > 1$ . When this holds all party members are dissatisfied with the conference: each would prefer to dictate policy in a leadership role.

When  $R < 1$  the desire to dictate is felt only by some. The activists who are content with the outcome of a conference and the risk of mis-coordination that this entails are “moderates” in the sense that their signals fail to identify clearly the best policy: they are unsure of the right thing to do. By contrast, those with signals that point decisively toward one of the policies retain a desire to lead. They are “extremists” not because of their preferences, but because of their confidence in assessing the party's mood.

Understanding the roles played by barriers to coordination, the need for direction, and party members' sense of the underlying party mood helps us to determine the desirability of leadership. But when is leadership feasible? To address this question, suppose that a leader stands up and makes a speech which is heard by all of the party, and so commonly communicates her views to the party mass. Activists can now draw upon two sources of information when forming their assessment of the party's mood: firstly, a private signal emerging from conference; and secondly, a public signal from the leader's speech.

Suppose that the speech favors policy  $A$ . This will lead an activist to reappraise positively his opinion of the party mood with regard to  $A$ , and so increases the likelihood that if he is pivotal then he is pivotal for  $A$ . Contemplating this, he realizes that since others hear the same speech, they too will begin to move toward  $A$  and away from  $B$ . This reinforces his bias toward  $A$ : a bandwagon begins to roll. However, a second force—the activist's curse—slows this bandwagon. When the positive feedback sparked by the speech overcomes the negative feedback from the activist's curse, the bandwagon rolls on until all support for policy  $B$  evaporates.

Positive feedback is related to the information content of the leader's speech which is determined, in turn, by her sense of direction. Furthermore, the relative likelihood of pivotal events responds strongly to activists' assessments of the party mood when these events are (heuristically) far apart. This is so when the barriers to coordination are high. On the other hand, the activist's curse is large when the merits of the policies depend heavily on swings in the party's mood; that is, when the need for direction is pressing. Bringing these factors together we find that positive-feedback dominates negative-feedback if and only if  $R > 1$ . In summary, the leader's speech fully coordinates the party mass when Michels' Ratio is sufficiently large; this is precisely the criterion that generates the need for leadership felt by the mass. Put simply, leadership that fully coordinates the party is feasible if and only if it is desired by the mass.

If  $R < 1$  then while the party is not susceptible to leadership by an individual, it may defer to an oligarchy. Suppose that an elite of  $k$  party members successfully share their assessments of the party's mood, and hence reach a common view. If a representative is able to communicate perfectly the elite's view, then party members will hear a speech of greater precision; the speech conveys a better sense of direction. By pooling information, the elite enhances the positive-feedback effect, and full coordination is feasible. Our analysis reveals that an elite of  $k$  party members is able to do this if and only if  $k \times R > 1$ : the minimum feasible size of a *de facto* oligarchy is inversely related to Michels' Ratio.

Following a brief literature review, we describe the coordination problem faced by party members. We then characterize behavior in a leaderless world, where decisions are based on information arising from a stylized party conference. Our study of the role of leadership is in three steps: firstly, we assess the need for leadership; secondly, we investigate the feasibility of successful leadership; and finally we consider leadership by an elite.

## RELATED LITERATURE

Before proceeding to construct the model of activism and leadership that underpins our arguments, we first review a small selection of the literature upon which we build.

Party activists play a *global game* (Carlsson and van Damme, 1993); following Morris and Shin (2003), it is a game “of incomplete information whose type space is determined by the players each observing a noisy signal of the underlying state.” Here, the “underlying state” is the party’s mood and the “noisy signal” includes information gleaned from other party members (a private signal) or from a leader’s speech (a public signal). Economists have used global games to model many phenomena, including currency crises (Morris and Shin, 1998), bank runs (Goldstein and Pauzner, 2005), and debt pricing (Morris and Shin, 2004). However, the approach has seen little use in political science.

We also build upon Myatt’s (2006) analysis of strategic voting in plurality-rule three-candidate elections. He explored the properties of a stable voting equilibrium with multi-candidate support, contrasting with the “Duvergerian” equilibria in previous studies (Cox, 1994, 1997; Palfrey, 1989; Myerson and Weber, 1993). Whereas in those earlier models voters commonly understand the electoral situation, Myatt’s voters base their decisions on private signals. Since signals may point in different directions all candidates receive votes in equilibrium, and so supporters of two candidates who share a dislike of a third candidate fail to fully coordinate their actions. This feature is present in our model, since differences in private signal realizations generate a risk of mis-coordination. The key difference between the two papers is this: in Myatt’s model voters know their own preferences, but preferences differ across the electorate; in ours, party members share the same preferences but must use signals to identify the policies’ merits.

The emergence of a leader corresponds to the introduction of a commonly observed public signal of the party’s mood. For this reason, our paper relates directly to the work of Morris and Shin (2002) and subsequent debates (Angeletos and Pavan, 2004; Svensson, 2006; Morris, Shin, and Tong, 2006). They considered the welfare effects of enhanced dissemination of public information in the context of a coordination game.

Finally, some insights stem from the literature on information aggregation. A negative-feedback effect was central to the work of Feddersen (1992) and Feddersen and Pesendorfer (1997, 1998). For example, Feddersen and Pesendorfer (1998) showed that the unanimity of jury verdicts can increase the likelihood of a conviction irrespective of the true state of the world: when a juror is pivotal in securing a conviction under unanimity, he will put less emphasis on his own private information than on what must be true given that others are making informed decisions. A similar effect exists here: a high barrier to coordination for one of the policies leads activists to move toward it.

## COORDINATING ACTIVISM

The setting for our analysis is a simultaneous-move binary-action common-interest game in which party activists must decide which of two policy platforms to advocate. Since the players share common payoffs, the game lacks the tension between private and social interests that is central to the view of collective action popularized by Olson (1965). Nevertheless, the actors face a collective-action problem in a different sense: a critical mass must coordinate on one of the policies if that platform is to be successfully adopted by their party. Failure to coordinate results in a party split, and subsequent electoral failure.

**The Game.** Formally,  $n$  party activists simultaneously decide to support either platform  $A$  or platform  $B$ . We imagine  $n$  to be large, so that the player set is either the entire party membership or a large and representative subset of them. We write  $x$  for the total number of activists who advocate platform  $A$ , so that  $n - x$  back  $B$ . The actions combine to yield one of three outcomes: (i) policy  $A$  succeeds; (ii) policy  $B$  succeeds; or (iii) the party is split, and fails to move forward. For two fractions  $p_A$  and  $p_B$  satisfying  $1 > p_A > p_B > 0$ ,

$$\text{Outcome} = \begin{cases} \text{Platform } A & \text{if } \frac{x}{n} > p_A, \\ \text{Failure} & \text{if } p_A \geq \frac{x}{n} \geq p_B, \text{ and} \\ \text{Platform } B & \text{if } p_B > \frac{x}{n}. \end{cases}$$

$p_A$  represents the coordination required for the success of  $A$ , and  $1 - p_B$  is the coordination required for  $B$ . We think of situations in which a 50 : 50 party split would lead to failure; this corresponds to the restriction  $p_A > \frac{1}{2} > p_B$ .<sup>4</sup> For later use, we define indices  $\pi_A$  and  $\pi_B$  which measure the severity of the coordination problem. Writing  $\Phi(\cdot)$  for the cumulative distribution function of the standard normal distribution, the *barrier to coordination* for platform  $A$  is a parameter  $\pi_A > 0$  satisfying  $p_A = \Phi(\pi_A)$ . Similarly, the barrier to coordination  $\pi_B > 0$  for policy  $B$  satisfies  $p_B = 1 - \Phi(\pi_B)$ . Platform  $A$  is a more challenging policy option than platform  $B$  whenever  $\pi_A - \pi_B > 0$ , and so  $\pi_A - \pi_B$  indexes the relative height of the barriers to coordination. Their aggregate height is the sum  $\pi_A + \pi_B$ . These expressions,  $\pi_A - \pi_B$  and  $\pi_A + \pi_B$ , feature prominently in our formal results.

Our stylized description of mid-1990s British Labour Party reform illustrates our specification: Platform  $A$  is the “New Labour” agenda, and  $B$  represents “old” Labour; a split means another term in the wilderness. Given that  $B$  is some kind of status quo, we might expect  $\pi_A - \pi_B > 0$ ; contentious aspects of the New Labour agenda require greater unity.<sup>5</sup>

<sup>4</sup>This restriction, which helps to simplify the exposition, is imposed throughout. However, the formal results continue to hold when either  $1 > p_A > p_B > \frac{1}{2}$  or  $\frac{1}{2} > p_A > p_B > 0$ .

<sup>5</sup>The case  $p_A > 1 - p_B$  is illustrated in Figure 1(b). If frustration with old Labour failings means that greater consensus is needed to retain existing policies, then the appropriate specification might be  $p_A < 1 - p_B$ .



**Payoffs.** The party activists, who share common interests, care only about the outcome. A failure to coordinate is undesirable, and so generates a zero payoff to everyone, whereas successful coordination yields a strictly positive payoff.<sup>6</sup> For some  $u_A > 0$  and  $u_B > 0$ ,

$$\text{Common Payoff} = \begin{cases} u_A & \text{if } \frac{x}{n} > p_A, \\ 0 & \text{if } p_A \geq \frac{x}{n} \geq p_B, \text{ and} \\ u_B & \text{if } p_B > \frac{x}{n}. \end{cases}$$

Since they share common interests, all players would like to coordinate on the best policy. Nevertheless, when information is complete there are many Nash equilibria. For instance, there is an equilibrium in which everyone backs  $A$  (so that  $x = n$ ) and an equilibrium in which everyone backs  $B$  (so that  $x = 0$ ). There are many other pure-strategy equilibria: any pattern of support where no individual is pivotal is an equilibrium.<sup>7</sup>

These equilibria include play where the party is split, and others where it succeeds. Despite this multiplicity, one equilibrium seems focal: if  $u_A > u_B$ , then surely everyone should coordinate on policy  $A$ ? Alas, this can only be obvious when it is common knowledge that  $u_A > u_B$ . If not, then an activist might worry that others think (or even suspect that others worry that he thinks) that  $u_B < u_A$ . In short, an analysis based on a complete-information game relies heavily on a common understanding of the policies' merits.

**A Need For Policy Direction.** To move further, we build a richer game in which the desirability of the policies is uncertain. In the context of a pure-strategy equilibrium players may entertain the possibility of pivotal events in which they may play a part.

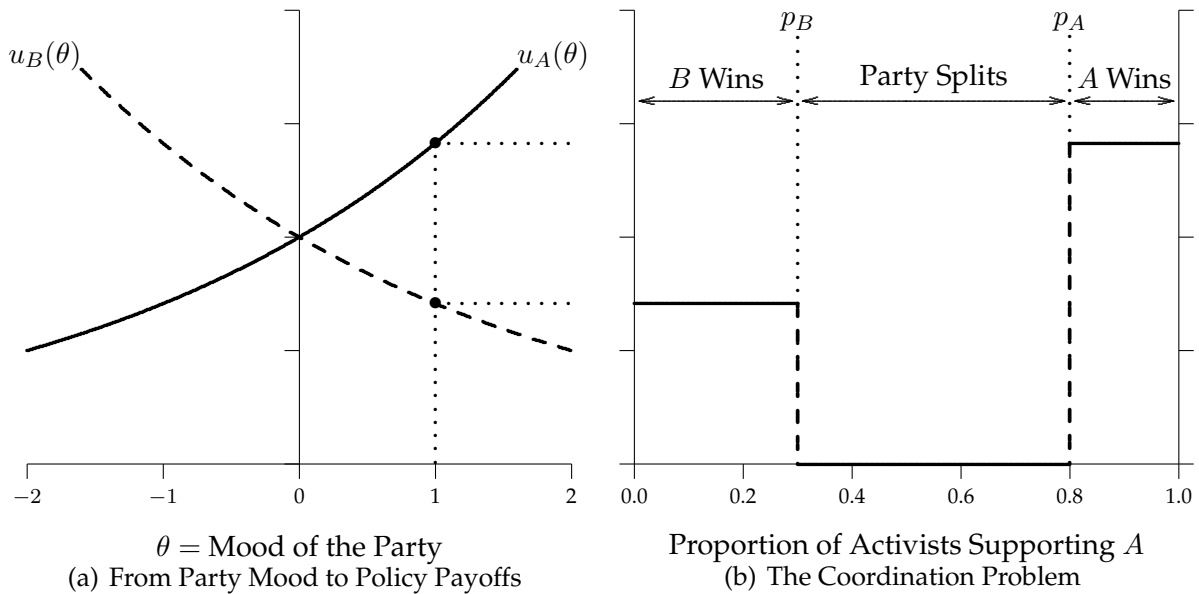
Formally, the payoffs  $u_A(\theta)$  and  $u_B(\theta)$  depend upon an underlying (and *ex ante* uncertain) real-valued state variable  $\theta$ . We assume that  $u_A(\theta)$  is increasing, and  $u_B(\theta)$  is decreasing, so that an increase in  $\theta$  favors the desirability of  $A$  relative to  $B$ . The state variable  $\theta$  represents the underlying political situation; it could depend on socioeconomic variables, the preferences of an electorate, or the evolving ideology of the party membership at large. We refer to it as the *mood of the party*. The congruence of policies with the party mood will then determine their desirability. When policy preferences respond strongly to mood, then there is a pressing *need for direction*, since the merits of a platform depend

<sup>6</sup>Each activist must choose to back one of the policies; indifferent abstention is disallowed. However, given that mis-coordination is possible and everyone strictly dislikes failure, abstention from participation is a (weakly) dominated strategy. This contrasts with the work of Feddersen and Pesendorfer (1996). They studied a world in which either  $A$  or  $B$  always wins, and found a strict incentive for abstention.

<sup>7</sup>The two pure-strategy profiles which are not equilibria correspond to values of  $x$  satisfying

$$\frac{x+1}{n} > p_A \geq \frac{x}{n} \quad \text{or} \quad \frac{x}{n} \geq p_B > \frac{x-1}{n}.$$

For instance, in the first situation a player who chooses  $B$  could switch to  $A$ , avoiding a failure and hence yielding a payoff gain of  $u_A$ . Such pivotal situations are exceptional, since other values of  $x$  (where no individual can affect the outcome) yield a pure-strategy Nash equilibrium. We could, of course, characterize mixed-strategy equilibria but this would only expand the embarrassment of riches.



(a) From Party Mood to Policy Payoffs

(b) The Coordination Problem

Panel (a) illustrates the relationship between the underlying political situation and the payoffs from successful coordination. For the particular value of  $\theta = 1$  highlighted, we see that  $\frac{u_A(\theta)}{u_B(\theta)} = 2$ , and so (if only they knew the party's true mood) activists would wish to coordinate on platform A. Panel (b) illustrates the barriers to coordination faced by activists. In this case,  $p_A > 1 - p_B$ , and hence policy platform A might be described as a "loftier" goal. When the proportion of activists supporting A falls into the central region, the party splits and fails.

FIGURE 1. Coordinating Activism

heavily on the value taken by  $\theta$ . We use a parameter  $\lambda > 0$  to index this need for direction, and we adopt the following specification:

$$u_A(\theta) = \exp\left[\frac{\lambda\theta}{2}\right] \quad \text{and} \quad u_B(\theta) = \exp\left[-\frac{\lambda\theta}{2}\right].$$

Figure 1(a) illustrates. When  $\theta$  is zero an activist is indifferent between the two policies. As the mood swings to the right then the payoff from platform A relative to platform B grows. It is convenient to express the payoffs in terms of log relative preference:

$$\log\left[\frac{u_A}{u_B}\right] = \lambda \times \theta \quad \text{where} \quad \begin{cases} \lambda = \text{need for direction, and} \\ \theta = \text{mood of the party.} \end{cases}$$

Of course, the log relative preference for A over B could be a general increasing function of the party mood; our specification is a linear approximation to such a relationship.

The party mood  $\theta$  is unknown, and so activists must use any informative signals at their disposal to form beliefs about it. If choices depend on such signals, then they will also depend on the realization of  $\theta$ . Hence an activist must contemplate the party mood for two reasons: firstly, he must assess the merits of the policies; and, secondly, he must consider the likelihood of pivotal events in which his advocacy makes a difference.

**Optimal Activism.** Whereas an analysis of our game requires us to specify the signals via which activists assess the party mood, we pause to discuss optimal behavior. An activist’s choice is relevant only when he is *pivotal*. For instance, he is pivotal for policy  $A$  ( $\mathcal{P}_A$  is shorthand for this event) when support for  $A$  is one step short of  $p_A$ . Similarly,  $\mathcal{P}_B$  is the situation in which the activist is pivotal for policy  $B$ . If choices respond to signal realizations, then these two pivotal events will have positive probability.

If an activist is pivotal for  $A$  then his support for  $A$  yields a payoff of  $u_A(\theta)$ . Of course,  $\theta$  is uncertain, and his expected payoff is  $E[u_A(\theta) | \mathcal{P}_A]$ . Notice that the merits of policy  $A$  are conditional on the pivotal event  $\mathcal{P}_A$ . This event happens with probability  $\Pr[\mathcal{P}_A]$ , and so the net payoff from backing  $A$  is  $\Pr[\mathcal{P}_A] \times E[u_A(\theta) | \mathcal{P}_A]$ . Similarly, the net payoff from backing  $B$  is  $\Pr[\mathcal{P}_B] \times E[u_B(\theta) | \mathcal{P}_B]$ . It is strictly optimal to support  $A$  if and only if

$$\Pr[\mathcal{P}_A] E[u_A(\theta) | \mathcal{P}_A] > \Pr[\mathcal{P}_B] E[u_B(\theta) | \mathcal{P}_B],$$

noting that the probabilities and expectations are also conditional on any other information available to an activist. This criterion may be conveniently re-written as

$$\text{Choose } A \iff \underbrace{\log \frac{\Pr[\mathcal{P}_A]}{\Pr[\mathcal{P}_B]}}_{\text{(i) relative likelihood}} + \underbrace{\log \frac{E[u_A(\theta) | \mathcal{P}_A]}{E[u_B(\theta) | \mathcal{P}_B]}}_{\text{(ii) relative conditional preference}} > 0. \quad (\star)$$

Hence the activist optimally balances (i) the relative likelihood of his pivotality for  $A$  versus  $B$ ; and (ii) his relative preference for the two policies, where this preference is evaluated conditional on his pivotality in each of the two cases.

## THE PARTY CONFERENCE

We first study behavior in a leaderless world. We find a unique equilibrium in which activists respond to privately observed signals of the party’s underlying mood.

**Intra-Party Communication.** A “party conference” is a metaphor for intra-party discussion of the underlying political situation. While discussion might occur in a variety of settings, we think of a large conference meeting. An activist gains (via meetings, conversations, and eavesdropping) a feel for the mood of the party; he receives a *private signal* of this mood. If the conference is representative of the wider party membership, then signals will be correct on average. Thus, if party members could pool their information then they would discover the true party mood and hence the correct policy platform.<sup>8</sup>

Unfortunately, signals differ: an activist who attends fringe meetings may get a different sense of the party mood than an activist who spends his time in the conference bar. This is likely to be important since, as an activist wanders around conference, he will ask himself whether others share his sense of the direction in which the party is going.

<sup>8</sup>Our results extend immediately to a setting in which the aggregate mood of the party conference imperfectly reflects the wider political situation and so fails to capture fully the merits of the policy alternatives.

**A Sense of Direction.** Activists begin with no knowledge of the party mood; they have uniform (improper) priors over  $\theta$ . Activist  $i \in \{1, 2, \dots\}$  then sees a private signal, where conditional on the party mood signals are identically and independently distributed:<sup>9</sup>

$$m_i | \theta \sim N\left(\theta, \frac{1}{\psi}\right) \quad \text{where} \quad \begin{cases} \theta = \text{mood of the party, and} \\ \psi = \text{sense of direction.} \end{cases}$$

Hence a signal is equal to the true party mood plus normal noise.  $\psi$  is the inverse of the variance of this noise, and so indexes the signal's precision: it is an activist's *sense of direction*. Conditional on his signal, his updated beliefs satisfy  $\theta | m_i \sim N(m_i, [1/\psi])$ .

**Thresholds.** We consider a natural class of strategies in which choices react positively to signals: an activist operates a *threshold rule* if he backs  $A$  rather than  $B$  if and only if his assessment of the party mood exceeds some threshold  $m$ . Equivalently

$$\text{Choose } A \iff m_i > m,$$

and so he chooses  $B$  if  $m_i < m$ .<sup>10</sup> Allowing the threshold to take values  $m \in \{-\infty, \infty\}$ , the class of threshold strategies includes those where an activist ignores his private signal and always advocates  $A$  (corresponding to  $m = -\infty$ ) or always advocates  $B$  (for  $m = \infty$ ). However, for any real-valued threshold, actions respond to signals in a non-trivial way.

**Definition.** When  $m$  takes on a finite real value, a *threshold rule* is signal responsive.

Conditional on the party mood, an activist backs  $A$  with probability  $p \equiv \Pr[m_i > m | \theta]$ . Writing  $\Phi(\cdot)$  for the cumulative distribution function of the standard normal, this probability satisfies  $p = \Phi[\sqrt{\psi}(\theta - m)]$ . Clearly, support for  $A$  increases as  $\theta$  grows, so that the party's mood swings to the right. It is also decreasing in the threshold used by activists: a reduction in  $m$  yields a move toward platform  $A$  and away from platform  $B$ .

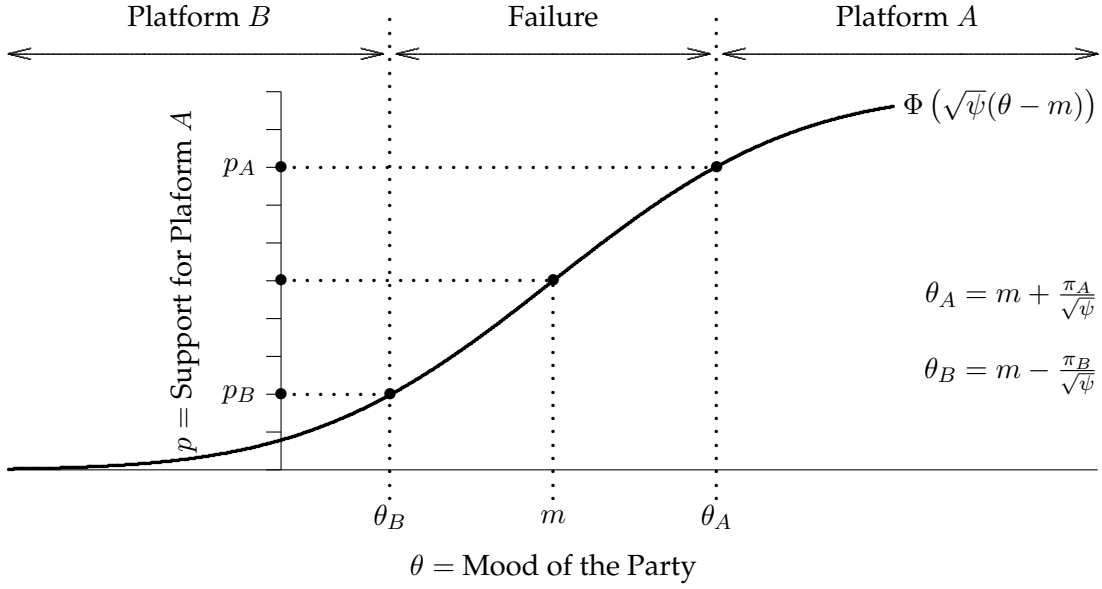
Whereas we call  $p$  the "support" for platform  $A$ , the actual fraction of activists who advocate  $A$  is  $\frac{x}{n}$ . However, if attendance at the conference is large, so that  $n \rightarrow \infty$ , then the Law of Large Numbers ensures that this fraction will almost always be close to  $p$ .

**Party Success and Failure.** Given that the party membership is large, policy  $A$  will succeed if and only if  $p > p_A$ . Equivalently, it succeeds if and only if the true party mood exceeds  $\theta_A$ , where  $\theta_A$  is the critical value that satisfies  $p_A = \Phi[\sqrt{\psi}(\theta_A - m)]$ ; similarly,  $B$  succeeds if and only if  $\theta < \theta_B$  where  $p_B = \Phi[\sqrt{\psi}(\theta_B - m)]$ . Summarizing,

$$\text{Outcome} = \begin{cases} \text{Platform } A & \text{if } \theta > \theta_A, \\ \text{Failure} & \text{if } \theta_A \geq \theta \geq \theta_B, \\ \text{Platform } B & \text{if } \theta_B > \theta, \end{cases} \quad \text{where} \quad \begin{cases} \theta_A = m + \frac{\pi_A}{\sqrt{\psi}}, & \text{and} \\ \theta_B = m - \frac{\pi_B}{\sqrt{\psi}}, \end{cases}$$

<sup>9</sup>Our results extend easily to environments in which the activists' signals are conditionally correlated.

<sup>10</sup>All results hold for any tie-break rule adopted for the (measure zero) case where  $m_i = m$ .



This figure illustrates the relationship between the party mood  $\theta$  and the support  $p$  for policy  $A$  (the support for  $B$  is  $1 - p$ ) when activists use the threshold rule:

$$\text{Choose } A \iff m_i > m.$$

When  $\theta = m$ , the party is 50 : 50 split between the platforms. For  $A$  to succeed, the party mood needs to swing right to  $\theta > \theta_A$ , and for  $B$  to succeed the mood must swing left to  $\theta < \theta_B$ . The mood is indecisive when  $\theta_A < \theta < \theta_B$ . Observe that

$$\theta_A - \theta_B = \frac{\pi_A + \pi_B}{\sqrt{\psi}},$$

and so the “zone of mis-coordination” for an indecisive party mood increases with the barriers to coordination but decreases with activists’ sense of direction.

FIGURE 2. Critical Moods of the Party

where we have used our indices of the barriers to coordination  $\pi_A \equiv \Phi^{-1}(\pi_A)$  and  $\pi_B \equiv \Phi^{-1}(1 - p_B)$ . The two critical party moods,  $\theta_A$  and  $\theta_B$ , partition the range of party moods into three segments according to the outcome. Figure 2 provides an illustration.

We make two observations. Firstly, when activists use a threshold  $m$  they react to their assessments and so policy responds to the party mood. Secondly, such strategies bring the risk of mis-coordination: when  $\theta_A > \theta > \theta_B$  the party’s mood (relative to the threshold  $m$ ) is one of indifference and the party fails. The size of this “zone of mis-coordination” is determined by  $(\pi_A + \pi_B)/\sqrt{\psi}$ , and so is increasing with the combined height of the barriers to coordination but is decreasing in activists’ sense of direction.

We now turn our gaze away from the conference at large and back to an individual’s decision. He will use his signal to assess the party mood and hence (i) the relative likelihood of his pivotality for  $A$  versus  $B$ , and (ii) his relative preference for  $A$  versus  $B$ . We study these two factors in turn, before characterizing behavior in equilibrium.

**Relative Likelihood.** When the party conference is large, and others adopt a threshold  $m$ , the support for  $p$  will be determined by  $\theta$ . An activist will realize that he can be pivotal for  $A$  only when  $\theta \approx \theta_A$ , and pivotal for  $B$  only when  $\theta \approx \theta_B$ . Thus, conditional on his private signal  $m_i$ , activist  $i$  will compute the relative likelihood of  $\theta_A$  versus  $\theta_B$ . Given that his posterior beliefs are normally distributed around  $m_i$  with precision  $\psi$ , the log relative likelihood takes a particularly simple form, as the following lemma confirms. (The proofs to this lemma and our other formal results are contained in the technical appendix.)

**Lemma 1.** *Fixing a threshold  $m$  used by other members of the party, and conditional on the private signal  $m_i$  of activist  $i$ , the log relative likelihood of being pivotal for  $A$  versus  $B$  satisfies*

$$\log \frac{\Pr[\mathcal{P}_A]}{\Pr[\mathcal{P}_B]} \longrightarrow (\pi_A + \pi_B) \sqrt{\psi} \times (m_i - m) \quad \text{as } n \rightarrow \infty.$$

*This is increasing in the activist's private assessment of the party mood, and decreasing in the threshold used by others. The size of the relative-likelihood effect increases with the barriers to coordination and an activist's sense of direction, but is unaffected by the need for direction.*

When assessing the likelihood of pivotal events, the activist judges the party mood relative to the threshold used by others. As he perceives a swing to the right (that is, a rise in  $\theta$ ) it is more likely that if he is pivotal then he will be pivotal for  $A$ . This response depends not only on his sense of direction, but also on the height of the barriers to coordination. As illustrated in Figure 2, as  $\pi_A$  and  $\pi_B$  grow the zone of mis-coordination widens. This implies that the two critical moods,  $\theta_A$  and  $\theta_B$ , move further apart and so are easier to distinguish. This being so, an activist can more readily ascertain which of the two critical events (whether he is pivotal for  $A$  or for  $B$ ) is more likely.<sup>11</sup>

Lemma 1 also reveals a positive-feedback effect. As others bias toward  $A$  (a fall in  $m$ ) the log likelihood ratio grows, which pushes activist  $i$  toward advocacy of  $A$ . This suggests a positive bandwagon: if an activist believes that others will bias toward a policy, then he will be tempted to follow them. This is not the whole story, however, since we must also consider the activist's relative preference (conditional on his pivotality) for  $A$  versus  $B$ .

**Relative Preference.** When contemplating the policies' merits, an activist considers his strategic environment. He recognizes that the payoff from policy  $A$  only matters when his action is critical to its success. Thus, he cares about the expected payoff from  $A$  when he is pivotal: this is  $E[u_A(\theta) | \mathcal{P}_A]$ . When the party is large he is pivotal only when  $\theta \approx \theta_A$ , and hence  $E[u_A(\theta) | \mathcal{P}_A] \longrightarrow u(\theta_A)$  as  $n \rightarrow \infty$ . Considering the behavior of others, the activist can work out what is true of the world (that is, the underlying party mood) when a pivotal event occurs. This means that the precise realization of his private signal does not matter. Similar considerations for the payoff from policy  $B$  lead to the following lemma.

<sup>11</sup>The need for direction does not enter Lemma 1. This is because the activist considers only the likelihood of  $\mathcal{P}_A$  and  $\mathcal{P}_B$ ; the need for direction comes into play only when an activist considers the policies' merits.

**Lemma 2.** Fixing a threshold  $m$ , the log relative conditional preference for  $A$  over  $B$  satisfies

$$\log \frac{\mathbb{E}[u_A(\theta) | \mathcal{P}_A]}{\mathbb{E}[u_B(\theta) | \mathcal{P}_B]} \longrightarrow \lambda \left[ m + \frac{\pi_A - \pi_B}{2\sqrt{\psi}} \right] \quad \text{as } n \rightarrow \infty.$$

This relative preference for  $A$  increases with the bias of others toward  $B$  and with the relative height of the barriers to coordination. The size of the second effect decreases with an activist's sense of direction, and the size of both effects increases with the overall need for direction.

Lemma 2 reveals a negative-feedback effect: as others lean toward  $A$  (a fall in the threshold  $m$ ) then the relative preference for  $A$  falls, which (*cf.* Lemma 1) pushes an activist toward  $B$ . Similarly, when  $A$  is easy to achieve, in the sense of having a lower barrier to coordination, an activist swings against it: he is relatively keen to pursue lofty goals. Borrowing the “swing voter’s curse” terminology from Feddersen and Pesendorfer (1996), we bring these effects under the umbrella of the *activist’s curse*.

This curse bites whenever the odds are stacked against a policy. If  $B$  is hard to achieve, or if others are biased toward  $A$ , then, on reflection,  $B$  can only win when the party mood leans strongly to the left and hence  $\theta$  is low. This implies that if  $B$  is on the verge of success then, since other factors work against  $B$ , the mood of the party must heavily favor it; the policy must be extremely desirable. Notice that (much like the winner’s curse of auction theory) the activist contemplates what must be true about the world given the situation in which he finds himself, namely that he is pivotal. The implications of such considerations depend upon the strength of the connection between the policies’ payoffs and the underlying state of the world; this is simply the need for direction  $\lambda$ .

**Equilibrium.** Lemmas 1–2 characterize the effect of a threshold strategy deployed by the party membership on the incentives of an individual activist. We now seek an equilibrium of our party-conference game within the class of threshold strategies.

**Definition.** A threshold equilibrium is a threshold rule “advocate platform  $A \Leftrightarrow m_i > m^*$ ” such that when all other party members use it: (i) an activist with a signal  $m_i > m^*$  finds it optimal to back policy  $A$  when the party size is sufficiently large; and (ii) an activist with a signal  $m_i < m^*$  finds it optimal to back policy  $B$  when the party size is sufficiently large.

Two threshold equilibria are the fully coordinated threshold rules with  $m^* = \infty$  and  $m^* = -\infty$ . For instance, when party members use the threshold  $m^* = -\infty$  they all ignore the realization of their private signals and back policy  $A$ . Given that they do this, the pivotal events  $\mathcal{P}_A$  and  $\mathcal{P}_B$  never happen, and so actions are (trivially) optimal.

In contrast, when an equilibrium threshold  $m^*$  takes a real value, decisions respond to signal realizations. For such a signal-responsive equilibrium the outcome is uncertain and so the events  $\mathcal{P}_A$  and  $\mathcal{P}_B$  occur with positive probability.

A threshold  $m^*$  forms a signal-responsive equilibrium when, in a large party, an activist with a private signal  $m_i = m^*$  is indifferent between the alternatives. Assembling the relative-likelihood and relative-preference effects,

$$m_i = m^* \implies \lim_{n \rightarrow \infty} \left[ \log \frac{\Pr[\mathcal{P}_A]}{\Pr[\mathcal{P}_B]} + \log \frac{\mathbb{E}[u_A(\theta) | \mathcal{P}_A]}{\mathbb{E}[u_B(\theta) | \mathcal{P}_B]} \right] = 0.$$

Inspecting Lemma 1, when  $m_i = m^*$  the relative-likelihood effect disappears. Given that a private signal is equal to the threshold used by others,  $\mathcal{P}_A$  and  $\mathcal{P}_B$  are equally likely in a large party: formally,  $\log[\Pr[\mathcal{P}_A]/\Pr[\mathcal{P}_B]] \rightarrow 0$  as  $n \rightarrow \infty$ . This means that the equilibrium threshold  $m^*$  is tied down by the relative preference term from Lemma 2, which must equal zero. This term is increasing in the difference between the barriers to coordination  $\pi_A$  and  $\pi_B$ , so that there is a preference bias toward policy  $A$  whenever  $A$  is more difficult to achieve (that is  $\pi_A > \pi_B$ ). The equilibrium threshold  $m^*$  must, therefore, offset this effect. Setting the relative preference to zero, and solving for  $m^*$  we obtain the next result.

**Proposition 1.** *There is a unique signal-responsive threshold equilibrium, with threshold*

$$m^* = \frac{\pi_B - \pi_A}{2\sqrt{\psi}}.$$

*There is a bias toward policy  $A$  if and only if its barrier to coordination is higher. The size of the bias  $|m^*|$  decreases with the sense of direction  $\psi$ , and is invariant to the need for coordination  $\lambda$ .*

One might think that an exogenous increase in  $\pi_A$  makes a success for policy  $A$  less likely. However, there is an endogenous reaction: in equilibrium activists are biased in their advocacy toward a policy that is more difficult to achieve; they are enthusiastic in their pursuit of lofty goals. Using our expressions for  $\theta_A$  and  $\theta_B$ , and substituting for  $m^*$ ,

$$\theta_A = \frac{\pi_A + \pi_B}{2\sqrt{\psi}} \quad \text{and} \quad \theta_B = -\frac{\pi_A + \pi_B}{2\sqrt{\psi}},$$

so that the zone of mis-coordination is a symmetric interval around zero. If  $\pi_A > \pi_B$ , for instance, then the endogenous equilibrium bias toward  $A$  precisely offsets the exogenous bias toward  $B$ . Turning attention back to Figure 2, the political outcome ( $A$ ,  $B$ , or failure) depends only upon the location of the zone of mis-coordination which, in turn, depends only on the aggregate height of the barriers to coordination. If we were to raise the barrier faced by one policy, while lowering the other, then this zone would not move. The important point to be made here is that a key feature of the institutional backdrop, namely that some policies require greater consensus, does not influence policy.

Two further features are worthy of note. Firstly, a signal-responsive threshold ensures that policy choices react to the party mood and so favors the use of a conference as a central institutional mechanism for aggregating opinions. Secondly, there is a risk of mis-coordination: the party fails when  $\theta_A > \theta > \theta_B$ . Note, however, that the relative importance of these effects depends on the need for direction and yet, despite this,  $\lambda$  plays no role in the equilibrium threshold strategy.



**Feedback and Stability.** Before assessing the performance of the conference, we highlight the feedback effects that are central to the incentives of activists, and justify our focus on the signal-responsive equilibrium described in Proposition 1.

An activist's advocacy will depend upon the threshold rule used by others. As  $m$  increases, both his relative likelihood of influencing the party's chosen policy and his relative preference for the policies are affected. Inspecting Lemma 1, notice that as  $m$  rises (so that others bias toward  $B$ ) the relative likelihood falls at rate  $(\pi_A + \pi_B) \times \sqrt{\psi}$ , so that he too is led toward policy  $B$ . This is *positive feedback*. On the other hand, Lemma 2 reveals that, as others bias toward  $A$ , his relative preference for  $A$  versus  $B$  rises at a rate  $\lambda$ ; this *negative feedback* stems from the activist's curse. Combining these two effects, we see that the positive-feedback effect exceeds the negative-feedback effect if and only if

$$(\pi_A + \pi_B) \times \sqrt{\psi} > \lambda.$$

When this holds, the net effect of a bias amongst other party members is to push an activist in the same direction. It proves useful to write this criterion in a modified form.

**Definition.** We define Michels' Ratio  $R$  to be a single index combining the aggregate barriers to coordination  $\pi_A + \pi_B$ , the need for direction  $\lambda$ , and an activist's sense of direction  $\psi$ . It satisfies

$$R \equiv \frac{(\pi_A + \pi_B) \times \sqrt{\psi}}{\lambda},$$

so that positive feedback exceeds negative feedback if and only Michels' Ratio satisfies  $R > 1$ .

Michels' Ratio features prominently in the results that follow; the index  $R$  is fundamental to both the feasibility and desirability of leadership and oligarchy.

When  $R > 1$  it seems that a "bandwagon" might form behind a policy. To explore this hypothesis, consider a situation where  $\pi_A = \pi_B$ , so that  $m^* = 0$ . Now suppose that party members switch to use a lower threshold strategy  $m < 0$ , so that they are biased toward  $A$ . When  $R > 1$ , this increases the incentive for an activist to back  $A$ ; in particular, an activist who observes a message  $m_i = 0$  will certainly advocate policy  $A$ .

Whilst this may suggest that a bandwagon begins to roll in favor of  $A$ , this is not quite so. To see why, consider an activist who observes a message  $m_i = m$ . For such a party member, the relative likelihood term disappears and thus the only factor he will consider is his relative preference term. Of course, the activist's curse is integral to this term and ensures that he now faces a strict incentive to back  $B$ . The attempt to induce a bandwagon effect in favor of policy  $A$  fails, since an activist with a signal equal to this new threshold will strictly wish to back  $B$ . This discussion leads to a formal definition of stability.

**Definition.** A threshold equilibrium  $m^*$  is stable if: (i) for larger thresholds  $m > m^*$  an activist with a private signal  $m_i = m$  would strictly prefer to back policy  $A$ ; and (ii) for smaller thresholds  $m < m^*$  an activist with a private signal  $m_i = m$  would strictly prefer to back policy  $B$ .

Given this definition and the discussion above, heuristically at least, the signal-responsive equilibrium threshold  $m^*$  described in Proposition 1 is stable. In contrast, the equilibria in which activists pay no attention to the realizations of their private signals are not.

**Proposition 2.** *In a conference environment, the signal-responsive equilibrium (Proposition 1) is stable. The fully coordinated equilibria in which activists ignore their signals are unstable.*

This provides one justification (there are others) for a focus on signal-responsive equilibria in a conference environment.<sup>12</sup> The important observation here is that the relative likelihood term depends not just on the threshold  $m$  used by others, but on the difference  $m_i - m$  between that threshold and the message received by an activist. Put more crudely, the positive-feedback effect is always eliminated when we consider an activist whose signal is equal to the threshold used by others and thus, in equilibrium, negative feedback dominates. Interestingly, the positive-feedback term returns to play a role once we introduce the possibility (later in the paper) of a publicly observed leader’s speech.

#### THE NEED FOR LEADERSHIP

Michels (1915, pp. 35–41) suggested that the transition from democracy to oligarchy is driven by a “need for leadership felt by the mass.” Here we analyze this need by assessing the desire of an individual to abandon conference and dictate the party’s policy.

**Guiding the Conference.** We begin with this thought experiment: we ask activist  $l$  (we call her a leader) to choose the threshold used by others. She has three options: (a)  $m = -\infty$ , so that everyone backs  $A$ ; (b)  $m = \infty$ , so that everyone backs  $B$ ; or (c) some interior threshold. Options (a) or (b) reflect a desire to lead dictatorially. Option (c) exploits the information-aggregation properties of a conference but risks a coordination failure; this choice reflects a desire for leadership to shape policy, but not to dictate it.

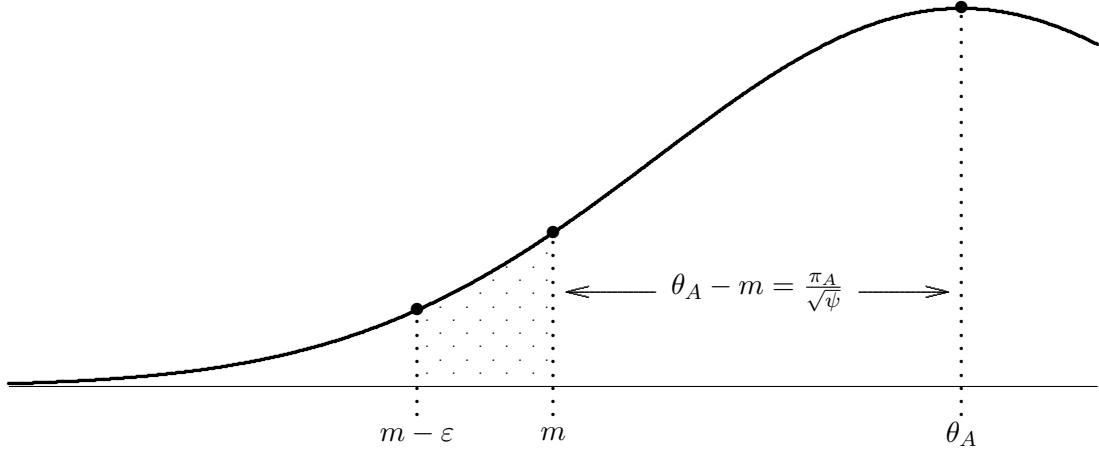
Conducting this thought experiment, we write  $U(m, m_l)$  for the expected payoff enjoyed by leading activist  $l$  when she sees a signal  $m_l$ , when others use a threshold  $m$ , and when the party is large, so that  $U(m, m_l)$  is evaluated as  $n \rightarrow \infty$ . Hence,

$$U(m, m_l) = \Pr[\theta > \theta_A] \times \mathbb{E}[u_A(\theta) \mid \theta > \theta_A] + \Pr[\theta < \theta_B] \times \mathbb{E}[u_B(\theta) \mid \theta < \theta_B],$$

where the probabilities and expectations are conditional on  $m_l$ . Consider a small change in the threshold. A fall in  $m$  shifts the zone of mis-coordination to the left, and so moves the party away from  $B$  and toward  $A$ . Formally,

$$\frac{\partial U}{\partial m} < 0 \quad \Leftrightarrow \quad \underbrace{\log \frac{\Pr[\mathcal{P}_A]}{\Pr[\mathcal{P}_B]}}_{\text{(i) relative likelihood}} + \underbrace{\log \frac{\mathbb{E}[u_A(\theta) \mid \mathcal{P}_A]}{\mathbb{E}[u_B(\theta) \mid \mathcal{P}_B]}}_{\text{(ii) relative conditional preference}} + \underbrace{\frac{\pi_B^2 - \pi_A^2}{2}}_{\text{(iii) feasibility effect}} > 0, \quad (\dagger)$$

<sup>12</sup>For instance, one justification is that the responsive equilibrium is optimal *ex ante*. Prior to the revelation of their private signals activists would all wish to commit to the play of the signal-responsive equilibrium.



This figure illustrates the effect of a reduction in the threshold  $m$  when the party mood hits the critical value  $\theta_A$ . The solid line represents the distribution of private signals across the party membership. A reduction in the threshold from  $m$  to  $m - \varepsilon$  increases the mass of activists who back  $A$ . This increase is represented by the shaded area. As  $\theta_A$  and  $m$  move further apart, the effect of the  $\varepsilon$  fall in the threshold is weakened. The distance between  $m$  and  $\theta_A$  is  $\pi_A/\sqrt{\psi}$ , and so a reduction in the threshold is less effective when the barrier to coordination for policy  $A$  is high.

FIGURE 3. Explaining the Feasibility Effect

where (i) and (ii) are evaluated at their limiting values as  $n \rightarrow \infty$ . These terms are identical to those that influence advocacy decisions. This is because the fall in  $m$  has an effect only when  $\theta \approx \theta_A$  or  $\theta \approx \theta_B$ ; that is, when activists are pivotal. An increase in the relative likelihood of  $\mathcal{P}_A$  versus  $\mathcal{P}_B$  leads to a desire to reduce  $m$  thus increasing the likelihood of a success for  $A$ . The relative preference for  $A$  is greater when there is a bias toward  $B$ : in selecting a threshold to be used by others, an activist's curse remains.

The third term on the right-hand side of (†) is absent from the advocacy criterion of (★). It measures the difference in the heights of the barriers to coordination. Other activists (if left to their own devices) favor the ambitious policy objective. Other things equal, activist  $l$  wishes to offset this bias; if  $\pi_B > \pi_A$  so that the barrier to coordination on policy  $A$  is lower, the leader would prefer to lower the threshold  $m$  and bias conference toward  $A$ .

To see why, note that the number of activists who switch sides following a change in  $m$  depends upon the party mood. Those who switch have signals close to  $m$ . If  $|m - \theta|$  is large then there are relatively few of these "marginal" activists since they are far from the median signal of  $\theta$ . Figure 3 illustrates this effect. From the definitions of  $\theta_A$  and  $\theta_B$ :

$$|m - \theta_A| = \frac{\pi_A}{\sqrt{\psi}} \quad \text{and} \quad |m - \theta_B| = \frac{\pi_B}{\sqrt{\psi}}.$$

Suppose that  $\pi_A < \pi_B$ , so that the barrier to coordination on policy  $A$  is lower. When  $\theta = \theta_A$ , there are relatively many marginal activists, whereas when  $\theta = \theta_B$  there are relatively few. Since more activists switch, following a change in  $m$ , when  $\theta = \theta_A$  compared to when

$\theta = \theta_B$ , the pivotal event  $\mathcal{P}_A$  becomes more important than the event  $\mathcal{P}_B$ . This provides an enhanced incentive to push the threshold  $m$  down and bias party members toward  $A$ .

Summarizing, if we delegate the choice of a threshold to a leading activist, she will (other things equal) bias the choice of threshold toward the policy that is “more feasible” in the sense that it faces a lower barrier to coordination. Her intervention acts as a counterweight to the endogenous bias toward an ambitious goal. However, given the presence of the term (i) in the criterion (†), she will also wish to move the threshold in the direction of her own private signal. For instance, if  $m_l > 0$  then she may wish to depress the threshold. If so, then how low will she go?

To answer, suppose that the criterion (†) holds. A reduction in  $m$  feeds back into terms (i) and (ii): the relative likelihood of  $\mathcal{P}_A$  versus  $\mathcal{P}_B$  increases due to positive feedback, whilst the relative conditional preferences falls due to the negative feedback from the activist’s curse. (The feasibility effect is unaffected.) When  $R > 1$ , the positive effect on (i) is greater than the negative effect on (ii); since positive feedback exceeds negative feedback the criterion remains positive and thus the leader faces a further incentive to lower  $m$ . Iterating this logic, she wishes to lower  $m$  all the way to  $m = -\infty$ . Similarly, if she wishes to raise  $m$  then she wishes to raise it all the way to  $m = +\infty$ . When  $R > 1$ ,  $U(m, m_l)$  is a quasi-convex or “U-shaped” function of  $m$ , and so the leader would prefer to choose one of the two extreme values for the threshold. In essence, she would prefer to dictate the policy of the party and sees no role for a democratic party conference.

If Michels’ Ratio satisfies  $R < 1$ , so that negative feedback exceeds positive feedback, then a reduction in  $m$  will lower the criterion of (†). In fact,  $U(m, m_l)$  is a quasi-concave or “hill-shaped” function of  $m$ . Thus, when  $R < 1$ , there is a unique interior threshold  $m^\dagger$  that maximizes the leader’s payoff. In essence, she would prefer a situation where, although she chooses the threshold, conference remains sovereign in that it determines the overall policy outcome. We summarize our thought experiment in the next proposition.

**Proposition 3.** *If  $R > 1$  then a leader prefers to dictate the party’s policy rather than allow a conference threshold strategy to operate. If  $R < 1$ , then she prefers others to use a threshold:*

$$m^\dagger = m^* - \frac{R}{1 - R} m_l$$

where  $m^*$  is the equilibrium threshold from Proposition 1. This new threshold  $m^\dagger$  is decreasing in her own private signal, and so, relative to the equilibrium threshold, the leader prefers others to operate a threshold that is biased toward her own private opinion of the party mood.

Proposition 3 suggests that Michels’ Ratio indexes the desirability of leadership. When  $R > 1$  a leader wishes to push the threshold to an extreme: a desire to dictate party policy. When  $R < 1$  then, whilst not wishing to dictate, she wishes to guide conference toward a particular threshold. The extent of the desire to guide is itself indexed by  $R$ : as  $R$  increases a leader acts with less restraint in moving the threshold toward an extreme.

**The Desire to Lead.** The index  $R$  brings together three factors which drive a desire to lead. Michels' Ratio increases with the aggregate height of the barriers to coordination and with an activist's sense of direction, but decreases with the need for direction.

The desire for leadership arises, most directly, when there are high barriers to coordination, since the zone of mis-coordination (Figure 2) is wide: ensuring that activists work together is more important than ensuring that the correct policy is chosen.

A willingness to take up the reins of leadership is affected by an activist's sense of direction. When  $\psi$  is large, she has confidence in her signal and in her ability to lead the party in the right direction. If  $\psi$  is low, it is more prudent to exploit the information-aggregation properties of the advocacy game, although the desire to guide conference remains.

When ultimate power resides with conference, policy is more likely to be in tune with the party mood. This is desirable when there is a greater need for policy direction, so that  $\lambda$  is large. When this is so, leaving control with the conference reduces the risk that a leader will choose to implement an inferior policy. But, when  $\lambda$  is small, the cost of leading the party toward an inferior outcome is small also and thus she retains a desire to dictate.

**Extremists and Moderates.** In our thought experiment we asked what threshold a leading activist would choose if she could ensure that others would use it. Of course, other party members will use the equilibrium threshold  $m^*$  rather than a leader's preferred threshold  $m^\dagger$ ; these thresholds coincide if and only if the leader's signal is neutral. When it is not, the leader may prefer to dictate even though  $R < 1$ . That is,  $R > 1$  is sufficient to generate a desire to lead, but not necessary.

When  $R < 1$ , some are content with conference, but others, given the opportunity, would like to lead. A leader with a signal  $m_l = 0$  lacks a clear signal of the party mood, and so is content with the information-aggregation properties of a conference; formally  $m^\dagger = m^*$  in this case. She is a *moderate* with an equivocal assessment of the party's mood. In contrast, a leading activist with a very high or very low private signal is an *extremist* who is confident in her assessment and so feels that she is well suited to dictating party policy. This suggests a taxonomy where only the moderates favor the party conference.

**Proposition 4.** *If  $R < 1$  then a leading activist is content with conference if and only if  $|m_l| < \tilde{m}$ , for some  $\tilde{m} > 0$ . Otherwise, if  $|m_l| > \tilde{m}$ , she would prefer to dictate by choosing the policy herself.  $\tilde{m}$  rises with the need for direction, but falls with the aggregate height of the barriers to coordination and the sense of direction. Changing these parameters,  $\tilde{m} \rightarrow 0$  as  $R \rightarrow 1$ .*

It may seem obvious that an extremist would prefer to dictate. This issue is a little more subtle, since an extremist can also be confident that a party conference will select the best policy. After all, she believes that the party mood leans heavily in one direction. Hence others will also receive strong signals and will almost certainly coordinate on the policy of her choice. This argument suggests that an extremist will be almost indifferent between

the two options. Nevertheless, and as the proof of Proposition 4 shows, she is relatively more confident in her own assessment than in the performance of conference.

Our definition of an extremist differs from common usage: an extremist is a leading activist who obtains a strong signal and so has a strong sense of the correct thing to do. Convinced of her opinions she is loathe to provide conference with a mandate and prefers to dictate. By contrast, a moderate's weak signal means that she is happy to allow others to take their cue from a sovereign democratic institution such as the party conference. Our taxonomy thus suggests a conflict within a party organization concerning the best mechanism for achieving policy coordination. This conflict coincides with differing private interpretations of the collective interest, and emerges despite the lack of inherent divergence of preferences in our model specification.

### FOLLOWING THE LEADER

We have analyzed an individual's desire to take a leading role. Following Michels' (1915, p. 36) claim that "the renouncement of the exercise of democratic rights is voluntary" we now assess the willingness of others to follow that lead. An activist attempts to lead by speaking to conference. When will the membership defer to her message?

**A Leader Speaks.** A party member  $l$  is chosen at random and makes a stand as leader. She gives a perfectly communicated speech  $s = m_l$  describing her views, so that others commonly observe her assessment of the party's mood. Hence party members observe a public signal  $s$  of  $\theta$ , where  $s | \theta \sim N(\theta, [1/\psi])$ . Activist  $i$ 's updated beliefs satisfy

$$\theta | (s, m_i) \sim N\left(\frac{s + m_i}{2}, \frac{1}{2\psi}\right).$$

The expectation of the party mood is an average of a private and public signal, and has twice its former precision. Importantly, the same public signal is used by all activists, and so party members begin to share the same perspective on the party's mood. This shared perspective is, of course, based on the leader's speech.

**Relative Likelihood and Relative Preference.** Once again, activist  $i$  supports  $A$  if and only if  $m_i > m$  for some threshold  $m$ . When he considers the likelihood of the pivotal events  $\mathcal{P}_A$  and  $\mathcal{P}_B$  his beliefs will now be influenced by the public signal.

**Lemma 3.** *Fixing a threshold  $m$  used by others, conditional on the private signal  $m_i$  of activist  $i$  and the speech  $s$  of the leader, the log relative likelihood of being pivotal for  $A$  versus  $B$  satisfies*

$$\log \frac{\Pr[\mathcal{P}_A]}{\Pr[\mathcal{P}_B]} \longrightarrow \frac{\pi_B^2 - \pi_A^2}{2} + 2(\pi_A + \pi_B)\sqrt{\psi} \times \left(\frac{m_i + s}{2} - m\right) \quad \text{as } n \rightarrow \infty.$$

*This is increasing in the activist's private assessment of the party mood and the leader's public speech expressing her views of the party mood. It is decreasing in the threshold used by others.*

In the absence of a leader's intervention a marginal activist (with a signal realization of  $m_i = m$ ) is unconcerned by the relative likelihood of the pivotal events; following Lemma 1, the log relative likelihood term is zero. Inspecting Lemma 3 this is no longer so when a public signal, in the form of a perfectly communicated leadership speech, is available: the relative likelihood of the pivotal events is influenced by the speech  $s$ .

This observation can open the gate to a bandwagon in favor of one of the policies. For instance, a bandwagon may begin to roll in favor of policy  $A$  when party members use a threshold  $m < 0$ . This generates positive feedback via the relative-likelihood effect, and negative feedback via the relative-preference effect of the activist's curse. In the absence of a leader's speech, the positive-feedback effect disappears. However, once a leader speaks, positive feedback can accelerate the bandwagon. Of course, the negative feedback stemming from the activist's curse (Lemma 2 continues to hold) can slow the bandwagon; we will consider this issue when we assess the stability of signal-responsive equilibria.

**A Threshold Equilibrium with Leadership.** A leader would like the party to use her preferred threshold. For  $R < 1$ , this is  $m^\dagger$  from Proposition 3. We now ask whether the leader's speech (a public signal to the party's membership) helps in the pursuit of her aims. We write  $m^*$  for the equilibrium threshold in the absence of a speech and  $m^\diamond$  for a new equilibrium threshold following the commonly observed public signal.

**Proposition 5.** *Following the observation of the leader's speech  $s$ , there is a unique signal-responsive threshold equilibrium in which party members use the threshold*

$$m^\diamond = m^* - \left[ \frac{R}{1-R} \right] s.$$

*If Michels' Ratio satisfies  $R < 1$ , then this signal-responsive equilibrium is stable. Hence if a leading activist  $l$  is able to make a public speech  $s = m_l$  then the threshold used by others in a stable equilibrium will be precisely her preferred threshold  $m^\dagger$  emerging from Proposition 3.*

Hence a leadership speech, which is perfectly communicated to the conference floor, results in the adoption of a common threshold which is precisely that chosen by the leader. When the leader announces her signal (she makes the speech  $s = m_l$ ) to the party membership at large, then other activists know what she knows. Given that they have all information at the leader's disposal, and they share her preferences, they act in the way that she would like them to act. The leader's original frustration with the conference-based threshold equilibrium, reflected in the difference  $m^* - m^\dagger$ , stemmed from differences of opinion that, in turn, were caused by different signal realizations.

**The Feasibility of Leadership.** We recall (Proposition 3) that Michels' Ratio provides an index of the desirability of leadership. It also (Proposition 5) indexes the willingness of a conference to adapt its behavior to a leader's speech. As  $R$  increases conference responds,

and more so when the leader's message is strong (so that  $|s|$  is large). Since conference moves its threshold in response to her speech, there is a direct effect of the leader's intervention on policy: Michel's Ratio indexes the feasibility of a leader's influence.

Combined with our previous results, an interesting and yet subtle insight emerges. We note that, when  $R < 1$ , a leader would always wish to guide conference toward the use of her own preferred threshold  $m^\dagger$ . When her speech communicates her assessment perfectly, it is persuasive: it results in the adoption of her desired threshold. An activist would lead if she could—and she can, since successful leadership is feasible.

Of course, if perfect communication is elusive, then the situation is more complex. Suppose that the leader is unable to explain her views, but is allowed (if she wishes) to intervene and choose the party's policy. It is in this situation (at least for  $R < 1$ ) that only an extremist with a strong signal of the correct thing to do would choose to dictate; following Proposition 4, a moderate prefers the sovereignty of conference. No such difference arises when the leader can explain herself: she would always wish to do so and hence ensure the use of her preferred threshold rule. This extremist-moderate divide arises only when a clear public speech is impossible. This suggests to us the importance of rhetoric in the classical sense: a democratic assembly is constrained by the limits of communication.

**Feedback and Stability.** We noted that our concept of threshold rules admit fully coordinated equilibria in which activists ignore their signals. In a conference environment, our focus on signal-responsive equilibria was justified by the stability of the threshold  $m^*$ .

This justification applies to the new threshold  $m^\diamond$  so long as  $R < 1$ . To see this consider a fall in the threshold to  $m < m^\diamond$ , which moves the party from  $B$  to  $A$ . This increases the relative likelihood term: a bias toward  $A$  is self-reinforcing, putting further upward pressure on  $m$ . However, it also lowers the relative conditional preference for  $A$  versus  $B$ : since  $B$  is now relatively harder to achieve, activists bias toward it. When  $R < 1$  the negative-feedback dominates the positive-feedback effect, and thus a marginal increase in the threshold fails to stimulate full coordination on  $A$ .<sup>13</sup> Of course, if  $R > 1$  then this argument fails, shifting focus away from signal-responsive equilibria.

**Proposition 6.** *If Michels' Ratio satisfies  $R > 1$  then the signal-responsive equilibrium with threshold  $m^\diamond$  is unstable, but the fully coordinated equilibria in which activists ignore their signals (one equilibrium in which everyone backs  $A$ , and the other where they all back  $B$ ) are locally stable.*

Here "locally stable" means this: for  $m < m^\diamond$ , an activist with a signal  $m_i = m$  strictly prefers to back  $A$ , and for  $m > m^\diamond$  an activist  $m_i = m$  strictly prefers to back  $B$ . This means, for instance, that once a threshold falls below the signal-responsive threshold  $m^\diamond$ , and  $R > 1$ , there is pressure to drive the party toward full coordination on policy  $A$ .

<sup>13</sup>This comparison of positive and negative feedback did not work in a conference environment, since the relative-likelihood term disappeared when we considered a marginal activist with a signal  $m_i = m$ .



**Focal Leadership and Perfect Coordination.** Using stability as an equilibrium criterion, when  $R > 1$  Proposition 6 tells us that we must look toward the two fully coordinated equilibria in which activists ignore their private assessment of the party mood. But which of these two equilibria, if any, will the mass adopt?

Following Proposition 3 a leader is delighted to provide a focal resolution to this coordination problem since she wishes to dictate. For instance, a clear and unambiguous public announcement that “everyone should back  $A$ ” is an obvious focal point. Of course, we are saying nothing new here; many examples of such focal points were suggested in a range of complete-information coordination games described in the classic work of Schelling (1960). However, we can offer a further justification. Consider the following speech:

“I stand before you as leader, and say to you that my assessment of the party mood is  $s = m_l$ . Based on this, I would like you all to advocate policy  $A$  if and only if your own assessment of the party mood is higher than  $m^\dagger$ . Having heard my views, but prior to discussing the merits of policy amongst yourselves, you would all unanimously wish to commit to following my recommendation. Now that you have formed your own private opinions, you have no reason to deviate from my recommendation. Hence you should follow my advice. Given that you do, I have no reason to mislead you over my own assessment of the party mood.”

The leader is providing a valid assessment of the party mood based on her signal. She conveys the reasons for her recommendation, and it is this feature of her appeal which makes it focal. Were she an outsider, who could not make such an assessment, her speech would lack focal properties, despite the fact that it is commonly understood. The leader is simply asking others to do what they would wish to commit themselves to *ex ante*, and noting that they have no reason to deviate *ex post*. Finally, the leader notes that she has no reason to misrepresent her views. In all, her request for others to follow her is compelling.

Note that the leader need not necessarily make an explicit policy recommendation (although she can) since activists are able to calculate the threshold that they mutually prefer *ex ante* based on their common observation of the leader’s public signal. Our argument is independent of the value taken by  $R$ , and leads to the following proposition.

**Proposition 7.** *Suppose that an activist stands as leader, and makes a perfectly communicated speech describing her views and describing a policy choice. If she does so, then party members ignore their private information and precisely adopt her recommended policy if and only if  $R > 1$ .*

Moving beyond the leader’s rhetoric, when  $R > 1$  there are further justifications for unification behind the leader’s preferred policy. Consider a world in which, prior to the leader’s speech, activists employ the (stable) threshold  $m^*$  from Proposition 1. The leader then speaks, with a speech  $s > 0$  in favor of policy  $A$ . This speech will cause an individual

party activist to reappraise positively the relative likelihood of  $\mathcal{P}_A$  versus  $\mathcal{P}_B$ . Given that others use the threshold  $m^*$ , he now finds it optimal to use a lower threshold  $m < m^*$ .

Of course, an activist might then anticipate that other party members will follow the same thought process. If he does, then he now expects them to use a lower threshold than before. Since  $R > 1$ , positive feedback exceeds negative feedback, and hence he will find it optimal to push down his own threshold still further. This heuristic “fictitious play” exercise continues until full coordination on policy  $A$  results.

## OLIGARCHY

We have explored situations in which an activist mass hands control to a *de facto* dictator. We now consider their response to an organized group (a clique) within their ranks. If activists ignore their private signals and follow the lead provided by this clique then the democratic rule by conference is replaced by the rule of an oligarchy.

**The Formation of a Clique.** Suppose that a clique of  $k$  party activists join together and share their views, and that  $k$  is small enough to allow a mutual understanding to form. We can imagine a meeting where each member of the clique is able to speak clearly to her colleagues. They reach a consensus and share common beliefs about the party mood.

Whereas the beliefs of an individual have precision  $\psi$ , those of the clique (an average of their signals) have precision  $k\psi$ , since they are based on  $k$  conditionally independent signals of the party mood.<sup>14</sup> Hence  $k$  indexes the sense of direction of the clique relative to that of an individual. This scenario is equivalent to one in which a leader is better able to assess policies than other party members. By studying the formation of a clique, we merely provide a micro-foundation for such a sharpened sense of direction.

**Guiding the Conference.** Previously, we conducted a thought experiment in which we asked a leading activist to choose the threshold used by others. She could dictate party policy, or defer back to conference with her own choice of threshold. From Proposition 3, she prefers to dictate if and only if  $R > 1$ ; a leader’s confidence in her own assessment of the party mood is sufficient for her to abandon conference.

We now conduct this experiment when the  $k$ -strong clique of activists have formed a common view. Since the precision of their shared beliefs is  $k\psi$  rather than  $\psi$ , they have greater confidence in their ability to do the right thing. The same logic presented in our analysis of the leader’s thought experiment applies here. However, the key criterion determining the desire for dictatorial leadership becomes  $kR > 1$  rather than  $R > 1$ .

<sup>14</sup>Alternatively, we might specify conditional correlation of the clique’s signals. For instance, if the clique forms from an impromptu meeting in the conference bar, then their signals might be based on similar information sources, and so might be correlated conditional on the party mood  $\theta$ . Whereas the average of their signals continues to provide a sufficient statistic for beliefs about  $\theta$ , its effective precision is lower. Thus, when  $k$  activists share conditionally correlated signals, the effective size of the clique is lower than  $k$ .

**Proposition 8.** *If  $kR > 1$  then a  $k$ -strong clique prefers to dictate policy rather than allow a conference threshold strategy to operate. If  $kR < 1$ , then they prefer others to use a threshold:*

$$m^\ddagger = m^* - \frac{kR}{1 - kR} \bar{m}$$

where  $m^*$  is the equilibrium threshold from Proposition 1 and where  $\bar{m}$  is the average private signal amongst the clique.  $m^\ddagger$  is decreasing in the clique's shared assessment of the party mood; they prefer others to operate a threshold that is biased toward their shared opinion.

Michels' Ratio continues to drive the desire to lead and the comparative-static properties of the clique's preferred threshold precisely match those for the case of a single leader. In fact, Proposition 3 is a special case of Proposition 8 for  $k = 1$ .

We can also extend the extremist-moderate taxonomy of Proposition 4 to a  $k$ -strong clique. Recall that this classification arose when we asked a leader to choose between dictating policy herself and retaining the equilibrium threshold  $m^*$  from the party conference.

**Proposition 9.** *If  $kR < 1$  then a  $k$ -strong clique is content with conference if and only if  $|\bar{m}| < \tilde{m}$ , for some  $\tilde{m} > 0$ . Otherwise, if  $|\bar{m}| > \tilde{m}$ , they would prefer to dictate by choosing the policy themselves.  $\tilde{m}$  rises with the need for direction, but falls with the aggregate height of the barriers to coordination and the sense of direction. Changing these parameters,  $\tilde{m} \rightarrow 0$  as  $kR \rightarrow 1$ .*

**From Elite to Oligarchy.** We have assessed the desire of a clique to take a leading role, and now turn to consider the feasibility of group-based leadership. Suppose that the  $k$ -strong clique becomes an elite: a group of activists who are able to communicate clearly their views to the party membership. Equivalently, the elite is able to put forward a single representative who can perfectly express their views  $s = \bar{m}$  by a speech to conference. Following this speech, activist  $i$ 's updated beliefs will satisfy

$$\theta | (s, m_i) \sim N \left( \frac{ks + m_i}{k + 1}, \frac{1}{(k + 1)\psi} \right).$$

He pays more attention to the views of the elite than he does to his own signal. His perspective is shared by others, enhancing the positive-feedback effect. Unsurprisingly, positive feedback exceeds negative feedback if and only if  $kR > 1$ , and this criterion is central to our final proposition which extends Propositions 5–7 to the case of an elite.

**Proposition 10.** *Following a speech by a  $k$ -strong elite, there is a unique signal-responsive equilibrium in which party members use a threshold  $m^\ddagger$ . If  $kR < 1$ , then this equilibrium is stable. If  $kR > 1$ , then it is unstable, but the fully coordinated equilibria are (locally) stable. Activists following the advice of the elite will always play a stable equilibrium. They ignore their private information (they defer to a de facto oligarchy) if and only if  $kR > 1$ . Hence  $1/R$  is the minimum size of a successful Michelsian oligarchy. This size increases with the need for direction  $\lambda$ , but decreases with the height  $\pi_A + \pi_B$  of the barriers to coordination and the sense of direction  $\psi$ .*

The first element confirms the existence of a stable signal-responsive equilibrium when  $kR < 1$ . The equilibrium threshold  $m^\ddagger$  is that preferred by the elite. The wider party membership is happy to follow the elite's advice, since this is how they would play if they could commit *ex ante*. The elite shapes policy, but conference remains sovereign.<sup>15</sup>

The second element concerns the case  $kR > 1$  when the signal-responsive equilibrium (which would involve a bias against the public signal) is no longer stable. Moreover, the elite would ideally like to see the full coordination of the party on the policy indicated by their shared view. Such a fully coordinated equilibrium is (locally) stable.

The third element describes the emergence of an oligarchy. As it was under the leadership of an individual, the advice of the leading elite is compelling. Prior to the realization of their private signals, but after listening to the elite's communication, party members would unanimously wish to follow the elite's advice. Thus, even in the absence of an explicit recommendation, the mass would like to do the elite's bidding. When  $kR > 1$ , they perfectly coordinate on a single policy and the elite becomes a *de facto* oligarchy.

The final element of Proposition 10 reveals that the desirability and feasibility of leadership by both individuals and elites are intrinsically linked by the index  $R$ .

**The Size of Oligarchy.** The inverse of Michels' Ratio provides a lower bound to the size of Michelsian oligarchy. The precision of the elite's aggregate signal of the party mood is increasing in  $k$ . An activist mass gives way to the elite only when this precision is sufficiently high. Adopting a more general interpretation of the parameter  $k$ , an elite becomes an oligarchy so long as its sense direction is sufficiently sharp.

While  $k > 1/R$  turns an elite into an oligarchy, further increases in  $k$  will enhance the quality of the elite's policy recommendations. Given that this is the case, will the elite grow without bound and incorporate the entire party membership? As it does so, it will begin to reflect the aggregation properties of the general conference.

The problem which now arises is one of communication. As we noted above, in the absence of perfect communication a leader will prefer to dictate than to defer back to conference. An absence of perfect communication may also frustrate the efficacy of a larger elite. When the elite's membership is large it may find it difficult to aggregate successfully its views and communicate a coherent and easily understood message to a wider audience. This issue may place an upper bound on the feasible size of a Michelsian oligarchy.

---

<sup>15</sup>We have a mind a situation where  $k > 1$ , so that an elite has a better sense of direction than an individual. However, our arguments continue to apply when  $k < 1$ . Letting  $k \rightarrow 0$ , the criterion  $kR < 1$  will be satisfied and the (now uninformative) elite's preferred threshold converges to  $m^*$ . Of course, this is the threshold from the signal-responsive equilibrium described in Proposition 1. This provides further justification for our focus on the unique signal-responsive equilibrium in a conference environment, since the case  $k \rightarrow 0$  corresponds to removal of any public information and hence a return to a diffuse public prior belief over  $\theta$ .

## TOWARD A THEORY OF LEADERSHIP

Leadership was central to Levi's (2006, p. 5) "desire to understand what makes for good governments and how to build them." In our view a leader's commonly interpreted assessment of the party mood can provide a focal point for the coordination of an activist mass. Our analysis thus adds to the literature on the role of institutions in solving coordination problems (Calvert, 1995; Weingast, 1997; Myerson, 2004). We note, however, a subtle difference between our approach and these precedents. In a world of incomplete information, leadership is important not just because it provides a common understanding of how the game will be played, but also because it provides payoff-relevant information concerning the merits of competing policies. Of course, leadership is not the only means of providing such information: we have compared leadership to a more democratic form of information aggregation, namely a stylized party conference.

In assessing these different institutional forms we were inspired by Michels (1915) who famously argued that a system of direct democracy would give way to decision-making by an oligarchic elite. His conceptualization of two mutually incompatible types of internal governance provides food for further formal analysis which we have provided here. We have provided micro-foundations for his claim that a "need for leadership" exists in mass psychology:

"Though it grumbles occasionally, the majority is really delighted to find persons who will take the trouble to look after its affairs. In the mass, and even in the organized mass of the labor parties, there is an immense need for direction and guidance." (Michels, 1915, p. 38)

In our world, the need for direction, barriers to coordination, and an activist's sense of direction combine to give a single measure (Michels' Ratio) of leadership. It indexes not only the willingness of activists to modify their behavior in light of a leader's speech, but also their willingness to abandon the conference forum altogether and follow a leader's prescription; it is a complete index of the feasibility of leadership. Moreover, the desirability of leadership is determined by the same combination of variables.

Echoing Michels' claim, the need for leadership is felt when barriers to coordination are high (so that the coordination problem is severe) and when a leader's sense of direction is sharp (so that she knows what to do). In contrast, our "need for direction" works in favor of a party conference: while leadership enhances the clarity of intra-party communication and so avoids the penalty of mis-coordination, it lessens the response of policy to the party mood. Our analysis thus contributes to an understanding of the trade-off between the responsiveness of policy outcomes and concentration of power in the form of dictatorship or oligarchy, central to the formal analysis of social-choice mechanisms.

Our results are based on a conceptualization of leadership different from that in previous formal approaches to the subject, which cast leaders in the role of agents under the control of a legislative body (Fiorina and Shepsle, 1989). In those models, a leader possesses skills necessary to the achievement of collective goals; the gap in expertise between leaders and followers underlies a common agency problem. An interesting feature of our analysis is that neither the desirability or the feasibility of leadership depend upon a given skill set which the leader may possess. Nevertheless, the framework developed here for analyzing the role of leadership could be extended to address this issue. For example, in our model the establishment of an oligarchy allows individual members of the oligarchy to pool their information; this, in turn, allows an oligarchy to convey more precise information. Of course, individuals differ also in their ability to evaluate information and convey messages. An extension of our model would allow for an exploration of this and other individual traits, important to the role of leadership in sustaining coordination.

Our analysis of the basic coordination problem faced by activists is devoid of factional conflict: activists share common interests and values but differ in their informed opinions of the path the party should take. Our approach captures a key element of intra-party division; that which pertains not to core values, but how best to achieve goals related to those values; uncertainty over how to achieve goals underpins division, as reflected in our analysis. Even in our common value game, a degree of factionalism may, nevertheless, emerge: those with neutral signals are more likely to place their trust in the sovereignty of conference, whilst those with extreme signals (pointing strongly in favor of a policy option) are willing to abandon conference as a central democratic institution.

The absence of conflict of interest in our model implies that a leader's communication always provides meaningful information; a leader's rhetoric is not "cheap talk". Our results suggest, however, that the clarity of a leader's communication is also important. Conference, acting as a central democratic institution, can (in aggregate) correctly assess the merits of policy, but as a mechanism for communication its performance is poor. A leader, by contrast, can convey only her private assessment of the party mood, but is able to do so with clarity. Moreover, the ability of a leader to convey clearly her message is relevant to institutional choice. For example, when a leader has an extreme signal then, faced with a choice, she would always wish to guide conference toward the use of her preferred threshold rather than to dictate; her willingness and ability to do so depends on her ability to communicate perfectly. Our focus on communication thus contributes to a broader understanding of different forms of governance, such as democracy and oligarchy, which until now have been studied formally only under the guise of commitment problems with regard to economic redistribution (Acemoglu and Robinson, 2000, 2001).

Finally, and perhaps most importantly, our results suggest a formal analysis of the role of rhetoric in effective leadership. Our next step will be to pursue this line of enquiry.

## TECHNICAL APPENDIX

Here we develop a formal model which encompasses the three scenarios (conference, leadership, and oligarchy) considered in the text, and provide proofs to Lemmas 1–3 and Propositions 1–10.

**Beliefs.** Activist  $i$  updates a diffuse prior over  $\theta$  following signals  $m_i | \theta \sim N(\theta, 1/\psi)$  and  $s | \theta \sim N(\theta, 1/[k\psi])$ . ( $k = 0$  is a conference,  $k = 1$  is a leader, and  $k > 1$  is a  $k$ -strong elite.) Conditional on  $\theta$ , signals are independent. Updating to form a posterior  $G(\theta | s, m_i)$  with density  $g(\theta | s, m_i)$ ,

$$\theta | (m_i, s) \sim N\left(\frac{ks + m_i}{k+1}, \frac{1}{(k+1)\psi}\right) \Rightarrow G(\theta | s, m_i) = \Phi\left(\sqrt{(k+1)\psi}\left[\theta - \frac{ks + m_i}{k+1}\right]\right). \quad (1)$$

$\Phi(\cdot)$  is the distribution function of the standard normal. Now suppose that other activists employ a threshold strategy. Conditional on  $\theta$ , an activist backs  $A$  with probability  $p$  where  $p = \Phi(\sqrt{\psi}(\theta - m))$ . Writing  $F(p | s, m_i)$  and  $f(p | s, m_i)$  for the distribution and density of beliefs about  $p$ ,

$$f(p | s, m_i) = \frac{1}{dp/d\theta} \times g(\theta | s, m_i) = \frac{g(\theta | s, m_i)}{\sqrt{\psi} \times \phi(\Phi^{-1}(p))} \quad \text{where} \quad \theta = m + \frac{\Phi^{-1}(p)}{\sqrt{\psi}}, \quad (2)$$

and where  $\phi(\cdot)$  is the density of the standard normal.

**Pivotal Probabilities.** Fixing activist  $i$ , and in a slight abuse of notation, write  $x \in \{0, 1, \dots, n-1\}$  for the number of others who advocate  $A$ . Conditional on  $\theta$ , party members back  $A$  with probability  $p$ , and hence  $x$  is a draw from the binomial with parameters  $p$  and  $n-1$ . However,  $p$  is uncertain and activist  $i$  must take expectations to form  $\Pr[x | s, m_i] = \int_0^1 \binom{n-1}{x} p^x (1-p)^{n-1-x} f(p | s, m_i) dp$ . Activist  $i$  is pivotal for the success of  $A$  if and only if  $x \leq p_A n < x+1$ . We write  $x_n^A$  for the unique value of  $x$  that satisfies these inequalities, and note that  $[x_n^A/n] \rightarrow p_A$  as  $n \rightarrow \infty$ . Clearly,

$$\Pr[\mathcal{P}_A | s, m_i] = \int_0^1 \binom{n-1}{x_n^A} p^{x_n^A} (1-p)^{n-1-x_n^A} f(p | s, m_i) dp.$$

This probability vanishes as  $n \rightarrow \infty$ . However, applying Proposition 1 of Chamberlain and Rothschild (1981),  $n \times \Pr[\mathcal{P}_A | s, m_i] \rightarrow f(p_A | s, m_i)$  and  $n \times \Pr[\mathcal{P}_B | s, m_i] \rightarrow f(p_B | s, m_i)$  as  $n \rightarrow \infty$ . This (in essence) is the Law of Large Numbers: when  $n$  is large, the proportion supporting policy converges in probability to  $p$ . Hence, pivotal probabilities are determined by beliefs about  $p$  via the density  $f(p | s, m_i)$ . While the probability of being pivotal vanishes with  $1/n$ , an activist cares not about the *absolute* probability of being pivotal, but rather the *relative* likelihood of  $\mathcal{P}_A$  and  $\mathcal{P}_B$ :

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[ \log \frac{\Pr[\mathcal{P}_A | s, m_i]}{\Pr[\mathcal{P}_B | s, m_i]} \right] &= \log \frac{f(p_A | s, m_i)}{f(p_B | s, m_i)} = \log \frac{\phi(\Phi^{-1}(p_B))}{\phi(\Phi^{-1}(p_A))} + \log \frac{g(\theta_A | s, m_i)}{g(\theta_B | s, m_i)} \\ &= \log \frac{\phi(\pi_B)}{\phi(\pi_A)} + \log \frac{g(\theta_A | s, m_i)}{g(\theta_B | s, m_i)} = \frac{\pi_A^2 - \pi_B^2}{2} + \log \frac{g(\theta_A | s, m_i)}{g(\theta_B | s, m_i)}. \end{aligned} \quad (3)$$

The first equality follows from Chamberlain-Rothschild. The second follows from Equation (2). The third follows from substitution of  $\pi_A$  and  $\pi_B$ , and the symmetry of the normal which ensures that  $\phi(\pi_B) = \phi(-\pi_B)$  and  $-\Phi^{-1}(1-p_B) = \Phi^{-1}(p_B)$ . The fourth follows from substitution into the standard normal density  $\phi(z) \propto \exp(-z^2/2)$ . The notation  $\theta_A$  and  $\theta_B$  is from the main text, where

$$\theta_A \equiv m + \frac{\Phi^{-1}(p_A)}{\sqrt{\psi}} = m + \frac{\pi_A}{\sqrt{\psi}}, \quad \text{and} \quad \theta_B \equiv m + \frac{\Phi^{-1}(p_B)}{\sqrt{\psi}} = m - \frac{\pi_B}{\sqrt{\psi}},$$

where the last equality again follows from the symmetry of the standard normal. Taking the posterior beliefs of activist  $i$  from Equation (1) and evaluating at  $\theta_A$ , we obtain

$$\begin{aligned} g(\theta_A | s, m_i) &= \sqrt{(k+1)\psi} \phi \left( \sqrt{(k+1)\psi} \left[ \theta_A - \frac{ks + m_i}{k+1} \right] \right) \\ &\propto \exp \left( -\frac{(k+1)\psi}{2} \left[ \theta_A - \frac{ks + m_i}{k+1} \right]^2 \right) = \exp \left( -\frac{(k+1)\psi}{2} \left[ m + \frac{\pi_A}{\sqrt{\psi}} - \frac{ks + m_i}{k+1} \right]^2 \right) \\ &= \exp \left( -\frac{k+1}{2} \left[ \psi \left( m - \frac{ks + m_i}{k+1} \right)^2 + 2\sqrt{\psi}\pi_A \left( m - \frac{ks + m_i}{k+1} \right) + \pi_A^2 \right] \right), \end{aligned} \quad (4)$$

where for the second step we have applied the formula for the density of the normal, and omitted the multiplicative constant that will be shared with the density  $g(\theta_B | s, m_i)$ . The final two equalities follow from substitution of  $\theta_A$  and algebraic manipulation. Similarly,

$$g(\theta_B | s, m_i) \propto \exp \left( -\frac{k+1}{2} \left[ \psi \left( m - \frac{ks + m_i}{k+1} \right)^2 - 2\sqrt{\psi}\pi_B \left( m - \frac{ks + m_i}{k+1} \right) + \pi_B^2 \right] \right). \quad (5)$$

Combining the expressions from Equations (4) and (5), we obtain

$$\log \frac{g(\theta_A | s, m_i)}{g(\theta_B | s, m_i)} = (k+1) \left[ \frac{\pi_B^2 - \pi_A^2}{2} + \sqrt{\psi}(\pi_A + \pi_B) \left( \frac{ks + m_i}{k+1} - m \right) \right]. \quad (6)$$

Substituting Equation (6) into Equation (3) we obtain

$$\lim_{n \rightarrow \infty} \left[ \log \frac{\Pr[\mathcal{P}_A | s, m_i]}{\Pr[\mathcal{P}_B | s, m_i]} \right] = \frac{k[\pi_B^2 - \pi_A^2]}{2} + (k+1)\sqrt{\psi}(\pi_A + \pi_B) \left( \frac{ks + m_i}{k+1} - m \right). \quad (7)$$

**Conditional Preference.** Here we study payoffs conditional on the events  $\mathcal{P}_A$  and  $\mathcal{P}_B$ . Our treatment is informal; a formal analysis is available from us upon request. In a large party, the proportion of activists supporting policy  $A$  converges in probability to  $p$ . Hence  $\mathcal{P}_A$  occurs if and only if  $p \approx p_A$ , or equivalently  $\theta \approx \theta_A$ . So, as  $n \rightarrow \infty$ ,  $\mathbb{E}[u_A(\theta) | \mathcal{P}_A] \rightarrow u_A(\theta_A) = \exp(\lambda\theta_A/2)$ . Hence,

$$\lim_{n \rightarrow \infty} \log \frac{\mathbb{E}[u_A(\theta) | \mathcal{P}_A]}{\mathbb{E}[u_B(\theta) | \mathcal{P}_B]} = \log \frac{u_A(\theta_A)}{u_B(\theta_B)} = \frac{\lambda(\theta_A + \theta_B)}{2} = \lambda \left[ m + \frac{\pi_A - \pi_B}{2\sqrt{\psi}} \right]. \quad (8)$$

*Proof of Lemmas 1–3.* Apply Equations (7) and (8).  $\square$

**Optimal Advocacy.** We now consider the decision of an activist given that the party is large.

$$\begin{aligned} &\lim_{n \rightarrow \infty} \left[ \log \frac{\Pr[\mathcal{P}_A]}{\Pr[\mathcal{P}_B]} + \log \frac{\mathbb{E}[u_A(\theta) | \mathcal{P}_A]}{\mathbb{E}[u_B(\theta) | \mathcal{P}_B]} \right] \\ &= \underbrace{\frac{k[\pi_B^2 - \pi_A^2]}{2} + (k+1)\sqrt{\psi}(\pi_A + \pi_B) \left( \frac{ks + m_i}{k+1} - m \right)}_{\text{from Equation (7)}} + \underbrace{\lambda \left[ m + \frac{\pi_A - \pi_B}{2\sqrt{\psi}} \right]}_{\text{from Equation (8)}} \\ &= \lambda \left[ \frac{k(\pi_B - \pi_A)(\pi_A + \pi_B)\sqrt{\psi}}{2\lambda\sqrt{\psi}} + (k+1) \frac{\sqrt{\psi}(\pi_A + \pi_B)}{\lambda} \left( \frac{ks + m_i}{k+1} - m \right) + m + \frac{\pi_A - \pi_B}{2\sqrt{\psi}} \right] \\ &= \lambda R \left[ m_i + ks + \left( k - \frac{1}{R} \right) m^* - \left( k + 1 - \frac{1}{R} \right) m \right]. \end{aligned} \quad (9)$$



The second equality follows from re-arrangement. The third equality follows from the substitution of Michels' Ratio  $R$  and  $m^*$  from Proposition 1 and further manipulation. Observe that this expression is increasing in  $m_i$ , and hence any optimal strategy (in a large party) is a threshold rule. Furthermore, when an activist's signal is equal to the threshold used by others:

$$m_i = m \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \left[ \log \frac{\Pr[\mathcal{P}_A]}{\Pr[\mathcal{P}_B]} + \log \frac{\mathbb{E}[u_A(\theta) | \mathcal{P}_A]}{\mathbb{E}[u_B(\theta) | \mathcal{P}_B]} \right] = \lambda R \left[ ks + \left( k - \frac{1}{R} \right) (m^* - m) \right], \quad (10)$$

which is increasing in  $m$  if and only if  $kR < 1$ . For a threshold equilibrium, this needs to be zero:

$$\lambda R \left[ ks + \left( k - \frac{1}{R} \right) (m^* - m) \right] = 0 \quad \Leftrightarrow \quad m = m^* - \left[ \frac{Rk}{1 - Rk} \right] s. \quad (11)$$

When a public signal is absent ( $k = 0$ ) then the equilibrium threshold is  $m^*$  from the statement of Proposition 1. If a public signal is in favor of policy  $A$  (that is, when  $s > 0$ ) then the equilibrium threshold is pushed down so long as  $Rk < 1$ . The extent of this effect is increasing in  $R$  and  $k$ .

*Proof of Propositions 1 and 2.* Apply Equations (10) and (11) with  $k = 0$ . □

**Choosing a Threshold.**  $U(m, m_l) = \int_{-\infty}^{\theta_B} u_B(\theta)g(\theta | m_l) d\theta + \int_{\theta_A}^{\infty} u_A(\theta)g(\theta | m_l) d\theta$ . Recalling that both  $\theta_A$  and  $\theta_B$  are linearly increasing in  $m$ , differentiate this payoff to obtain

$$\begin{aligned} \frac{\partial U}{\partial m} = u_B(\theta_B)g(\theta_B) - u_A(\theta_A)g(\theta_A) < 0 & \quad \Leftrightarrow \quad \log \frac{u_A(\theta_A)}{u_B(\theta_B)} + \log \frac{g(\theta_A | m_l)}{g(\theta_B | m_l)} > 0 \\ & \quad \Leftrightarrow \quad \lim_{n \rightarrow \infty} \left[ \log \frac{\Pr[\mathcal{P}_A | m_l]}{\Pr[\mathcal{P}_B | m_l]} + \log \frac{\mathbb{E}[u_A(\theta) | \mathcal{P}_A]}{\mathbb{E}[u_B(\theta) | \mathcal{P}_B]} \right] + \frac{\pi_B^2 - \pi_A^2}{2} > 0, \end{aligned}$$

where the final equivalence is from (3). The criterion is (†) from the main text. The sum of the first two terms (it is the same as (9) setting  $m_i = m_l$  and  $k = 0$ ) is strictly decreasing in  $m$  if  $R > 1$  and strictly increasing if  $R < 1$ . Thus (for the generic case of  $R \neq 1$ ) there is a unique  $m^\dagger$  satisfying  $\partial U / \partial m = 0$ . This must be a local minimum (and hence global minimum, since there is only one stationary point) if  $R > 1$ . Hence, when  $R > 1$ ,  $U(m, m_l)$  is maximized by choosing either  $m \rightarrow \infty$  or  $m \rightarrow -\infty$ . If  $R < 1$  then  $m^\dagger$  yields a global maximum. Explicitly,

$$\begin{aligned} \frac{\partial U}{\partial m} = 0 & \quad \Leftrightarrow \quad \underbrace{(\pi_A + \pi_B)\sqrt{\psi}(m_l - m^\dagger)}_{\text{from Equation (7) with } k=0} + \lambda \underbrace{\left[ m^\dagger + \frac{\pi_A - \pi_B}{2\sqrt{\psi}} \right]}_{\text{from Equation (8)}} + \frac{\pi_B^2 - \pi_A^2}{2} = 0 \\ & \quad \Leftrightarrow \quad m^\dagger = \frac{\pi_B - \pi_A}{2\sqrt{\psi}} - \frac{(\pi_A + \pi_B)\sqrt{\psi}m_l}{\lambda - \sqrt{\psi}(\pi_A + \pi_B)} = m^* - \left[ \frac{R}{1 - R} \right] m_l, \quad (12) \end{aligned}$$

where the solution for  $m^\dagger$  follows from simple algebraic manipulation.

*Proof of Proposition 3.* Apply Equation (12). □

**Extremists and Moderates.** If  $z \sim N(\mu, \sigma^2)$  then for real-valued constants  $b$  and  $H > L$ ,

$$\int_L^H \exp(bz) d\Phi \left( \frac{z - \mu}{\sigma} \right) = \exp \left( b\mu + \frac{b^2\sigma^2}{2} \right) \times \left[ \Phi \left( \frac{H - \mu - b\sigma^2}{\sigma} \right) - \Phi \left( \frac{L - \mu - b\sigma^2}{\sigma} \right) \right]. \quad (13)$$

We can use (13) to calculate a leader's expected payoff. Suppose that she believes  $\theta \sim N(\mu, \sigma^2)$ . Write  $U_A = \mathbb{E}[u_A(\theta)]$  and  $U_B[u_B(\theta)]$  for her payoffs when she dictates the adoption of policies  $A$

and  $B$  respectively. For  $U_A$  we set  $b = \lambda/2$ , and for  $U_B$  we set  $b = -\lambda/2$ . Hence

$$U_A = \exp\left(\frac{\lambda\mu}{2} + \frac{\lambda^2\sigma^2}{8}\right) \quad \text{and} \quad U_B = \exp\left(-\frac{\lambda\mu}{2} + \frac{\lambda^2\sigma^2}{8}\right) = \exp(-\lambda\mu) \times U_A. \quad (14)$$

$U_A > U_B$  if and only if  $\mu > 0$ : a leader implements  $A$  if and only if she expects the party mood to favor it. Next, consider her payoff when others use a threshold  $m$ . Policy  $A$  wins if  $\theta > \theta_A$ , policy  $B$  wins if  $\theta < \theta_B$ . So, writing  $\mathcal{I}[\cdot]$  for the indicator function and applying (13),

$$\begin{aligned} U_C &= \mathbb{E}\left[\exp\left(-\frac{\lambda\theta}{2}\right) \times \mathcal{I}[\theta < \theta_B]\right] + \mathbb{E}\left[\exp\left(\frac{\lambda\theta}{2}\right) \times \mathcal{I}[\theta > \theta_A]\right] \\ &= U_B \times \Phi\left(\frac{\theta_B - \mu}{\sigma} + \frac{\lambda\sigma}{2}\right) + U_A \times \left[1 - \Phi\left(\frac{\theta_A - \mu}{\sigma} - \frac{\lambda\sigma}{2}\right)\right] \\ &= U_A \times \left[\exp(-\lambda\mu) \times \Phi\left(\frac{\theta_B - \mu}{\sigma} + \frac{\lambda\sigma}{2}\right) + \Phi\left(\frac{\mu - \theta_A}{\sigma} + \frac{\lambda\sigma}{2}\right)\right], \end{aligned}$$

where the final equality stems from  $U_B = \exp(-\lambda\mu) \times U_A$  and from the symmetry of the normal. Now, suppose that the leader observes a signal  $m_l$  with precision  $k\psi$  so that  $\theta | m_l \sim N(m_l, [1/k\psi])$ . Without loss of generality, we set  $m_l > 0$ , so that  $U_A > U_B$ . Now, setting  $\mu = m_l$  and  $\sigma^2 = 1/k\psi$ ,

$$\frac{U_C}{U_A} = \exp(-\lambda m_l) \times \Phi\left(\sqrt{k\psi}(\theta_B - m_l) + \frac{\lambda}{2\sqrt{k\psi}}\right) + \Phi\left(\sqrt{k\psi}(m_l - \theta_A) + \frac{\lambda}{2\sqrt{k\psi}}\right). \quad (15)$$

If others adopt a threshold  $m^*$  then  $\theta_A = (\pi_A + \pi_B)/2\sqrt{\psi}$  and  $\theta_B = -(\pi_A + \pi_B)/2\sqrt{\psi}$ . Hence

$$\begin{aligned} \frac{U_C}{U_A} &= \exp(-\lambda m_l) \times \Phi\left(X - \sqrt{k\psi}m_l\right) + \Phi\left(X + \sqrt{k\psi}m_l\right) \\ \text{where } X &\equiv \frac{\lambda(1 - kR)}{2\sqrt{k\psi}} = \frac{\lambda}{2\sqrt{k\psi}} - \frac{\sqrt{k}(\pi_A + \pi_B)}{2}, \quad (16) \end{aligned}$$

following algebraic manipulation. We use this derivation in the next proof.

*Proof of Proposition 4.* Consider a leader with a signal  $m_l > 0$ . (The case  $m_l < 0$  is symmetric.) If she were to dictate then she would implement policy  $A$  and enjoy a payoff of  $U_A$ . By deferring to the equilibrium threshold of conference she enjoys a payoff  $U_C$ . She strictly prefers conference if

$$\begin{aligned} \frac{U_C}{U_A} > 1 &\Leftrightarrow \exp(-\lambda m_l) \times \Phi\left(X - \sqrt{k\psi}m_l\right) > 1 - \Phi\left(X + \sqrt{k\psi}m_l\right) \\ &\Leftrightarrow Y(m_l) \equiv \log\left[\frac{\Phi\left(X - \sqrt{k\psi}m_l\right)}{1 - \Phi\left(X + \sqrt{k\psi}m_l\right)}\right] - \lambda m_l > 0, \end{aligned}$$

where  $k = 1$  for a single leader. At  $m_l = 0$  (a neutral signal) this criterion becomes

$$\frac{U_C}{U_A} > 1 \Leftrightarrow \log\left[\frac{\Phi(X)}{1 - \Phi(X)}\right] > 0 \Leftrightarrow 2\Phi(X) > 1 \Leftrightarrow X > 0.$$

This last inequality holds if and only if  $kR < 1$ . Since  $Y(m_l)$  is continuous in  $m_l$ , there will be some region of signals close to zero for which the leader strictly prefers to follow conference. In fact, there exists a unique critical private signal  $\tilde{m}$  such that (for positive signals)  $Y(m_l) > 0$  if and only if  $m_l < \tilde{m}$ . The proof follows from the claim that  $Y(m_l)$  is strictly decreasing in  $m_l$  with a

derivative that is bounded away from zero. To prove this claim, first write

$$Y(m_l) = \log \left[ \frac{1 - \Phi(\sqrt{k\psi}m_l - X)}{1 - \Phi(\sqrt{k\psi}m_l + X)} \right] - \lambda m_l.$$

Next differentiate to obtain

$$\begin{aligned} Y'(m_l) &= \sqrt{k\psi} \left[ \frac{\phi(\sqrt{k\psi}m_l + X)}{1 - \Phi(\sqrt{k\psi}m_l + X)} - \frac{\phi(\sqrt{k\psi}m_l - X)}{1 - \Phi(\sqrt{k\psi}m_l - X)} \right] - \lambda \\ &= \sqrt{l\psi} \left[ h(\sqrt{k\psi}m_l + X) - h(\sqrt{k\psi}m_l - X) \right] - \lambda, \end{aligned}$$

where  $h(z) \equiv \phi(z)/[1 - \Phi(z)]$  is the hazard rate of the standard normal. Applying the mean-value theorem, there is some  $z$  satisfying  $\sqrt{k\psi}m_l + X > z > \sqrt{k\psi}m_l - X$  such that

$$Y'(m_l) = 2\sqrt{k\psi}Xh'(z) - \lambda = \lambda(1 - kR)h'(z) - \lambda,$$

where the second equality follows from the definition of  $X$  in (16). Now, the hazard  $h(z)$  of the standard normal is an increasing and convex function of its argument, and is asymptotically linear, so that  $h(z) - z \rightarrow 0$  as  $z \rightarrow \infty$ , and hence  $h'(z) \leq 1$  for all  $z$ . Hence, for  $kR < 1$ ,

$$Y'(m_l) \leq \lambda(1 - kR) - \lambda = -\lambda kR < 0.$$

Thus  $Y(m_l)$  is strictly decreasing in  $m_l$ , and the derivative is bounded away from zero. Hence  $Y(m_l) < 0$  for  $m_l$  sufficiently large, and there is a unique  $\tilde{m}$  such that  $Y(\tilde{m}) = 0$ .  $\square$

*Proof of Propositions 5 and 6.* Apply Equations (10) and (11) with  $k = 1$ .  $\square$

*Proof of Propostion 7.* From the argument given in the main text.  $\square$

*Proof of Proposition 8.* With an average signal of  $\tilde{m}$ , a  $k$ -strong clique's beliefs about the  $\theta$  are, modifying (1) appropriately, captured by the density  $g(\theta | \tilde{m}) = \sqrt{k\psi}\phi(\sqrt{k\psi}(\theta - \tilde{m}))$  where  $\phi(\cdot)$  is the density of the standard normal. Following derivations analogous to those leading up to (6),

$$\log \frac{g(\theta_A | \tilde{m})}{g(\theta_B | \tilde{m})} = k \left[ \frac{\pi_B^2 - \pi_A^2}{2} + \sqrt{\psi}(\pi_A + \pi_B)(\tilde{m} - m) \right]. \quad (17)$$

The clique's expected payoff  $U(m, \tilde{m})$  is locally decreasing in  $m$  (so that they favor a shift toward policy  $A$ ) if and only if  $g(\theta_A | \tilde{m})u_A(\theta_A) > g(\theta_B | \tilde{m})u_B(\theta_B)$ , or, upon substitution,

$$\log \frac{g(\theta_A | \tilde{m})}{g(\theta_B | \tilde{m})} + \log \frac{u_A(\theta_A)}{u_B(\theta_B)} = \underbrace{k \left[ \frac{\pi_B^2 - \pi_A^2}{2} + \sqrt{\psi}(\pi_A + \pi_B)(\tilde{m} - m) \right]}_{\text{from Equation (17)}} + \underbrace{\lambda \left[ m + \frac{\pi_A - \pi_B}{2\sqrt{\psi}} \right]}_{\text{from Equation (8)}} > 0. \quad (18)$$

The left-hand side of this inequality is strictly decreasing in  $m$  (implying that  $U(m, \tilde{m})$  is quasi-convex in  $m$ ) if and only if  $kR > 1$ . So, if  $kR > 1$ , the clique prefers to dictate policy. For  $kR < 1$ , setting the expression in (18) to zero and solving for  $m$  yields  $m^\ddagger$ .  $\square$

*Proof of Propostion 9.* For general  $k$ , the proof of Proposition 4 applies.  $\square$

*Proof of Propostion 10.* The first and second claims follow from an application of Equations (10) and (11) for general  $k$ , giving a solution for the equilibrium threshold equal to  $m^\ddagger$ . The remaining claims follow by inspection or from the arguments given in the main text.  $\square$

## REFERENCES

- ACEMOGLU, D., AND J. A. ROBINSON (2000): "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *Quarterly Journal of Economics*, 115(4), 1167–99.
- (2001): "A Theory of Political Transitions," *American Economic Review*, 91(4), 938–63.
- 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.
- CALVERT, R. L. (1995): "The Rational Choice of Social Institutions: Cooperation, Coordination, and Communication," in *Modern Political Economy*, ed. by J. S. Banks, and E. A. Hanushek. Cambridge University Press, Cambridge, UK.
- CARLSSON, H., AND E. VAN DAMME (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61(5), 989–1018.
- CHAMBERLAIN, G., AND M. ROTHCHILD (1981): "A Note on the Probability of Casting a Decisive Vote," *Journal of Economic Theory*, 25(1), 152–62.
- COX, G. W. (1994): "Strategic Voting Equilibria under the Single Nontransferable Vote," *American Political Science Review*, 88(3), 608–21.
- (1997): *Making Votes Count*. Cambridge University Press, Cambridge.
- FEDDERSEN, T. J. (1992): "A Voting Model Implying Duverger's Law and Positive Turnout," *American Journal of Political Science*, 36(4), 938–62.
- FEDDERSEN, T. J., AND W. PESENDORFER (1996): "The Swing Voter's Curse," *American Economic Review*, 86(3), 408–24.
- (1997): "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, 65(5), 1029–58.
- (1998): "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review*, 92(1), 23–35.
- FIORINA, M. P., AND K. A. SHEPSLE (1989): "Formal Theories of Leadership: Agents, Agenda-Setters, and Entrepreneurs," in *Leadership and Politics*, ed. by B. D. Jones, pp. 17–41. University Press Kansas, Lawrence.
- GOLDSTEIN, I., AND A. PAUZNER (2005): "Demand-Deposit Contracts and the Probability of Bank Runs," *Journal of Finance*, 60(3), 1293–327.
- HANDS, G. (1971): "Robert Michels and the Study of Political Parties," *British Journal of Political Science*, 1(2), 155–72.
- KOELBLE, T. A. (1996): "Economic Theories of Organization and the Politics of Institutional Design in Political Parties," *Party Politics*, 2(2), 251–63.
- LEVI, M. (2006): "Why We Need a New Theory of Government," *Perspectives on Politics*, 4(1), 5–19.

- LUPIA, A., AND M. D. MCCUBBINS (2000): "Representation or Abdication? How Citizens Use Information to Help Delegation Succeed," *European Journal of Political Research*, 37(3), 291–307.
- MCKENZIE, R. T. (1963): *British Political Parties: the Distribution of Power within the Conservative and Labour Parties (1955)*, Heinemann, London.
- MICHELS, R. (1915): *Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracy*. Batoche Books, Kitchener, Ontario, 2001 edition translated by Eden and Cedar Paul and available from <http://onlinebooks.library.upenn.edu/>.
- MORRIS, S., AND H. S. SHIN (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review*, 88(3), 587–97.
- (2002): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–34.
- (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics, the Eighth World Congress*, ed. by M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge University Press, Cambridge, UK.
- (2004): "Coordination Risk and the Price of Debt," *European Economic Review*, 48(1), 133–53.
- MORRIS, S., H. S. SHIN, AND H. TONG (2006): "Social Value of Public Information: Morris and Shin (2002) is Actually Pro-Transparency, not Con: Reply," *American Economic Review*, 96(1), 453–55.
- MYATT, D. P. (2006): "On the Theory of Strategic Voting," *Review of Economic Studies* (forthcoming).
- MYERSON, R. B. (2004): "Justice, Institutions, and Multiple Equilibria," *Chicago Journal of International Law*, 5(1), 91–107.
- MYERSON, R. B., AND R. J. WEBER (1993): "A Theory of Voting Equilibria," *American Political Science Review*, 87(1), 102–114.
- OLSON, M. (1965): *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press, Cambridge, MA.
- PALFREY, T. R. (1989): "A Mathematical Proof of Duverger's Law," in *Models of Strategic Choice in Politics*, ed. by P. C. Ordeshook, pp. 69–92. University of Michigan Press, Ann Arbor.
- SCHELLING, T. C. (1960): *The Strategy of Conflict*. Harvard University Press, Cambridge, MA.
- SVENSSON, L. E. O. (2006): "Social Value of Public Information: Comment: Morris and Shin (2002) is Actually Pro-Transparency, not Con," *American Economic Review*, 96(1), 448–52.
- WEINGAST, B. R. (1997): "The Political Foundations of Democracy and the Rule of Law," *American Political Science Review*, 91(2), 245–63.