Torun Dewan and Francesco Squintani

Leadership with trustworthy associates

Article (Accepted version)
(Refereed)

Original citation:
Dewan, Torun and Squintani, Francesco Leadership with trustworthy associates. American Political Science Review. ISSN 0003-0554 (In Press)
DOI:

Leadership with Trustworthy Associates[1]

Torun Dewan[2] **Francesco Squintani**[3]

October 2017

Abstract. Group members value informed decisions and hold ideological preferences. A leader takes a decision on their behalf. Good leadership depends on characteristics of moderation and judgement. The latter emerges (endogenously) via advice communicated by "trustworthy associates". Trustworthy advice requires ideological proximity to the leader. A group may choose a relatively extreme leader with a large number of such associates. Paradoxically, this can happen though it is in the group's collective interest to choose a moderate leader. To assess whether these insights persist when political groups compete, we embed our analysis in a model of elections. Each of two parties chooses a leader who implements her preferred policy if elected. We find that a party may choose an extreme leader who defeats a moderate candidate chosen by the opposing party. Our results highlight the importance of party cohesion and the relations between a leader and her party. These can be more important to electoral success than proximity of a leader's position to the median voter.

[1]We thank for their comments audiences at several seminars, including Harvard University, Stanford University, University of Chicago, Berkeley University, Princeton University, University of Rochester, and Washington University in St. Louis, and at several conferences.
[2]London School of Economics and Political Science. email: t.dewan@lse.ac.uk
[3]Warwick University. email: F.Squintani@Warwick.ac.uk

E la prima coniettura che si fa del cervello d'uno signore, é vedere li uomini che lui ha d'intorno.

Niccolo' Machiavelli, Il Principe, Ch. 22.

1. Introduction

Who should rule? Which individual characteristics are required for good and successful leadership? These questions are central to political writing and thought. They are addressed in Plato's Republic and, perhaps most famously, in Machiavelli's masterpiece *Il Principe*. A central contention of Machiavelli is that good governance stems from the characteristics of a ruler and that these determine political success. This view is shared by contemporary political scientists. Since Stokes (1963), studies have recognized the importance to electoral success of a candidates' valence– a term used to describe competence, talent, good judgement, that are generally desirable qualities of a leader.[4]

The personal attributes that contribute toward good leadership might be innate, as in the "great man" theory espoused by the social historian Thomas Carlyle. Alternatively they might be the product of education and training. Plato, for example, believed the education of future leaders to be a core function of the state. There is, however, a different view. Leadership characteristics may stem from the relations a leader enjoys with others in the governance process and the advice that she obtains from them. We find this perspective in Aristotle who, in Politics III.16, 1287 27-35, argues that advice from friends is central to a leader's judgement:

> "It would perhaps be accounted strange if someone, when judging with one pair of eyes and one pair of ears, and acting with one pair of feet and hands, could see better than many people with many pairs, since, as things stand, monarchs provide themselves with many eyes, ears, hands and feet. For they appoint as co-rulers those who are friends to themselves and to their rule. If they are not his friends, they will not do as the monarch chooses. But suppose they are friends to him and to his rule – well, a friend is someone similar and equal, so if he thinks they should rule, he must think that those who are equal and similar to her should rule like him."

A related theme emerges in Machiavelli's *Il Principe*. Our opening quote comes from chapter 22, where he writes on knowledge acquisition and making use of trusted advisers, and is translated as:

> "The first opinion that one forms of a prince, and of his understanding, is by observing the men he has around him."

[4]See McCurley and Mondak (1995); Ansolabehere, Snyder, and Stewart (2001); Burden (2004); Stone and Simas (2010), amongst others.

One interpretation, according to the idiom "a person is known for the company she seeks," suggests that we infer a leader's quality from the type of person she associates with. Another, perhaps more intriguing interpretation, is that a leader's qualities arise *because* of those she associates with.[5] Machiavelli entertains this second interpretation, attributing the greatness of Pandolfo Petrucci, Prince of Siena, to the relationship he enjoyed with his valent minister Antonio da Venafro. Indeed, Machiavelli highlights that Pandolfo's ability as a ruler depended upon the information and good judgement provided by Antonio.[6]

In this paper we develop a novel theory of leadership that relates a leader's judgement to the relations she forms with other group members whose advice she may benefit from. Such advice can help her form better judgement and so take more informed decisions. But, for this to be so, a leader must be able to trust the advice obtained. That is, the advice must be truthful. To explore this notion of leadership we analyze a group that collectively chooses a leader who is granted authority to take a decision on their behalf. Players' payoffs depend on an uncertain state of the world about which each is independently, privately, and imperfectly informed. Each, however, has different preferred outcomes that reflect their idiosyncratic ideological preferences. After the leader is chosen, but before the decision is taken, group members may advise the chosen leader, whoever she may be. Such advice takes the form of cheap-talk communication.

First, we study the endogenous formation of a leader's network of trustworthy associates: those the leader can rely upon for truthful advice. The results are intuitive. We show that a leader can rely on truthful advice only from those whose ideological preferences are similar to her own. Taking the next step we show that a leader's judgement depends upon the number of trustworthy associates that she has. A larger group of such associates translates into more informed decisions. This intuitive result establishes our take on the Machiavellian lesson: A leader's wisdom and judgement are determined by those she has around her. And it resonates with Aristotle's claim: a good leader has many friends, who are ideologically similar to her, and whose advice she benefits from.

In light of these results, we then ask: what are the characteristics of a good leader? She is defined as the one that the group should choose when maximizing their joint welfare. In line with the classic texts we find that moderation (or temperance) is desirable. Nevertheless, a good leader relies on her judgement that is determined by the number of allies in her circle.

[5]A different interpretation is that an instrinsically good leader is not threatened even when surrounded by highly capable, if potentially hostile, associates. For example, Kearns Goodwin (2005) relates the political genius of Abraham Lincoln to his ability in forming a cabinet consisting of erstwhile rivals to his Presidency.

[6]These two prominent examples develop the theme that reliable advice from friends, allies, and associates leads to better judgement and successful leadership. This view was in fact quite general in the Middle Ages. Recent analysis of a collection of the "mirror for princes" texts (a class of texts offering advice on governance, developed in both Christian Europe and the Islamic world in the Middle Ages, of which Machiavelli's work is the most famous) uses state of the art text-as-data measurement techniques developed by political scientists (Blaydes, Grimmer, and McQueen, 2013). This textual analysis reveals a prominent theme referring to the characteristics of exemplary rulers, such as their moderation (or temperance) and judgement. Within this theme, a main subtopic highlights the importance of a leader's relations with others.

Depending on the distribution of ideological views in the group, a moderate leader may be isolated in that she cannot rely on anyone's advice. Thus a tradeoff arises between moderation, on the one hand, and judgement on the other.

Our model delivers a simple mathematical equation that describes this potential tradeoff and the optimal choice of leader. This equation reveals that the tradeoff is related to different properties of the distribution of views in the group. A leader's moderation is understood with respect to the entire spectrum of views. Her judgement, by contrast, depends upon the concentration of viewpoints similar to her own. Put otherwise, moderation refers to "global" properties of the ideology distribution, while judgement is related to "local" ones.

What then are the characteristics of the chosen leader? Since the majority choice is determined by the median player, one might expect a moderate leader to be chosen. On the other hand, a leader's network is also an important consideration. Indeed we find, in line with common intuition and the practical wisdom of Aristotle, that a leader may be chosen because she has many friends. Here, a large network of friends translates into better judgement. A direct prediction stems from this. The group leader may in fact be relatively extreme. Indeed, this is so when ideological opinions are concentrated at the extremes.

Next, we compare the group's choice of leader under majority rule with the optimal choice and so address an issue raised by Ahlquist and Levi (2011) who note that "leadership does not always improve aggregate welfare and we need to know more about the conditions under which it does and it does not." We show that the chosen leader maximizes aggregate welfare when the ideology of group members is evenly distributed or clustered around the median. However, when ideologies are clustered away from the median, then she fails to achieve maximal aggregate welfare. Deeper insights emerge once we recognize our model as one of implicit (strategic) delegation, as first analyzed by Schelling (1960). The incentive of the median politician to delegate arises when another politician has more trustworthy associates and so will take a more informed decision. When choosing whether to delegate, however, the median considers only her own preferences. A surprising consequence is that she may delegate to a relatively extreme leader when a more moderate one (such as herself) would better serve the group interest. The upshot of this result is a reversal of the famous "ally principle," which states that delegation should take place to an ally who is as close as possible to the principal.[7] A surprising implication, from a welfare perspective, is that the group choice places *too much* emphasis on a leader's judgement, and too little on her moderation.

A useful exercise is to consider how the group's choice of leader changes when the preferences of its members change. This might occur due to transition in the group's membership or from exogenous shocks to members' preferences. In models of collective choice that build on Black's celebrated theorem only the identity (and hence opinion) of the median (player or committee member) matters for decisions made. Consequently, any change to the distribution of views

[7]See Bendor, Glazer, and Hammond (2001) for a review of this literature.

within a group that leaves the identity of the median unchanged has no effect on policy outcomes. Our comparative static analysis produces sharply different predictions. Changes in the ideological views of group members can affect leadership and hence policy choice *even though* they do not alter the identity of the "global" median player. This result stems from the fact that such ideological changes affect a leader's network of trustworthy associates. If some players become more extreme (moderate) in their views then a moderate (extreme) leader may lose important trustworthy allies. This affects her ability to exercise good judgement and hence her prospects of being chosen.

These new theoretical results shed light on empirical questions arising from applications of the spatial model in political science, such as for example work analyzing appointments made by the President to the Supreme Court that are approved by the Senate (Krehbiel, 2007; Rohde and Shepsle, 2007). In these models of complete information, based on Blacks' theorem, a Senate member's vote is based only on whether a proposed appointee changes the identity of the median court member. This reasoning does not sit well with common intuition that the viewpoints of all players are relevant to decision-making. Moreover, empirical evidence shows that extreme justices are less likely to have their nomination confirmed. Clark (2012) reviews this literature and notes that the facts are difficult to reconcile with existing theoretical models. He argues that an explanation requires relaxing the complete information assumption that underpins those models. Our analysis of decision-making in small groups with incomplete information supports the common intuition that the viewpoints of all group members are relevant, and suggests that Clark's conjecture is correct.

The main body of our paper explores the idea that a leader's judgement depends on her close associates and so, in turn, on the local distribution of preferences in the group. Next we check whether the empirical consequences of that assertion are robust when considering electoral competition between groups. To explore this, we study internal leadership contests (involving politicians, members, and/or registered voters) in two parties whose leaders then contest a general election after which the winner implements her preferred policies. A conjecture is that our surprising findings will disappear with competition that (as illustrated in the classic spatial model of Downs) provides incentives for parties to moderate their position. While that conjecture is correct in the absence of (strategic) communication within parties, it no longer holds true when a leader's judgement depends upon advice obtained from others. Parties may choose relatively extreme leaders even when more moderate candidates are available and, moreover, doing so can enhance their chances of electoral victory. In fact, and surprisingly, our results are stronger as a consequence of competition. That is, there exist circumstances in which the most moderate available political candidate would be elected as leader in the absence of electoral competition (i.e., if all politicians belong to a single group), whereas two-party competition would cause the election of relatively extreme leader.

A surprising comparative static prediction of our model involves the ideological direction of leadership change. A rightward shift in the ideology of a party politician can have an opposite effect on leadership choice, making it more likely that a leftist leader is chosen, and vice-versa. This non-monotonicity has further unexpected implications when considering party competition. We find that a party can turn a winning (losing) situation into a losing (winning) one when moderates (extremists) become more moderate (extreme). Moreover we illustrate how a political leader can turn a potential winning situation into a losing one by *moderating* her policy position. In so doing she reduces her leadership potential, becomes isolated and less well-placed to benefit from advice of others, and is unable to deliver informed policies.

Bringing these insights together reveals the importance of a party's cohesion on its electoral success. Our results suggest that the electoral success of relatively moderate leaders is not due to their moderation *per se*, but to the fact that their parties are cohesive. Correspondingly, we argue that the success of moderate leaders (e.g., Tony Blair and Bill Clinton) can be related to the fact that key figures in their party had moderated their own opinions. Indeed excerpts from Blair's autobiography suggest that his judgement during his first term in office depended upon the advise that was provided by trustworthy allies (such as David Blunkett) who themselves had moved from the hard to the centre left of the party.

2. Our Contribution to the Related Literature

While we shall comment on and discuss our contributions throughout, here we precede our analysis by briefly pointing out some of the main related literature and themes. We contribute to a small but growing formal literature that develops different notions of leadership. For example, Hermalin (1998) develops the notion of leading by example whereby a leader provides a costly signal that aligns followers' incentives with her own. Canes-Wrone, Herron, and Shotts (2001) draw a distinction between "leadership"–the act of implementing a policy that a leader believes to be correct– and "pandering" to a majority. Relatedly, Canes-Wrone (2006) develops a notion of "transformative leadership": in the context of an agenda-setting model, a leader (the President) strategically chooses whether to bring an issue to the public's attention anticipating that (the pivotal) member of Congress will move toward the public's position. Dewan and Myatt (2007) develop the notion of focal leadership that draws on earlier work by Schelling (1960) and Calvert (1995). Here a leader is connected via a network to trustworthy associates who influence her judgement through truthful communication of privately held information.[8] Leadership emerges via majority decision taken by group members who anticipate the formation of a leader's network. While the phenomena we describe–leadership, judgement, and trustworthiness–are macro level processes (and subject to different interpretations) here they emerge endogenously from a model built on sparse assumptions.

We study verbal (cheap talk) communication between privately-informed participants who provide advice to a leader anticipating that such advice may affect her decisions. Our insights are developed within the context of multi-player communication between imperfectly

[8]We thus develop the notion of "relational leadership" found in the social psychology literature, see Uhl-Bien (2006).

informed players as studied by Galeotti, Ghiglino, and Squintani (2013). There are numerous applications of multi-player communication in the political science literature: Patty and Penn (2014) study information transmission in small networks of decision makers; Patty (2013) determines the optimal exclusion and inclusion policies to maximize information sharing in meetings; Dewan, Galeotti, Ghiglino, and Squintani (2015) investigate the optimal assignment of decision-making power in the executive of a parliamentary democracy; Penn (2016) studies the formation of stable aggregation of different units within an association; Dewan and Squintani (2016) analyze the formation of party factions. Our contribution is in developing the multi player communication model to deliver a large set of distinctive findings on leadership and extending these in the context of party competition in which voters anticipate (multi-player) communication within parties.

Other models of leadership relate to individual characteristics such as honesty, courage and/or generosity as in the "great man theory" mentioned in our introductory remarks and so share our focus on characteristics that make a leader desirable. Dewan and Myatt (2007, 2008) contrast a leader's judgement with her ability to communicate clearly.[9] Bolton, Brunnermeier and Velkamp (2010) highlight the role of a leader's "overconfidence". Egorov and Sonin (2010) focus on the tradeoff between competence and loyalty to the leader. Besley and Reynal-Querol (2011) show that democratically elected leaders are more likely to have higher academic credentials than unelected ones. Relatedly, Galasso and Nannicini (2011) view talented leaders as a scarce resource and analyze party allocation of competent politicians, proxied by their education level, across electoral districts. We draw a distinction between a leader's judgement and her moderation. A key contribution here is in studying leader characteristics that are derived from first principles.

As mentioned in our introductory notes, our model can be seen as one of implicit strategic delegation initiated by Schelling (1960) to which recent more contributions include Harstad (2010) and Chari, Jones, and Marimon (1997). The question we ask is when and why a political leader would confer decision making authority to specialized or better informed bureaucrats (see Huber and Shipan 2002, for a review). According to perceived wisdom, a political principal would prefer delegating to a bureaucrat with views that are the most similar to her own. The logic behind this so called ally principle has recently been challenged. Bendor and Meirowitz (2004) identify a trade-off between a bureaucrat's information and ideological proximity as a reason for the its failure. Our work advances this insight in noting that while politicians may delegate to bureaucrats with a mandate limited to policy implementation, they may also delegate the act of decision-making to other politicians due to the fact that they are better informed.

Finally our model relates to a large literature on candidate valence defined as candidate's characteristics that benefit all voters regardless of their ideology. Many formal theoretical

[9]Within the context of a Keynsian beauty contest model, leaders are exogenous information sources that help party activists to advocate the best policies and coordinate their actions.

models have analyzed the implication of valence on candidate policies and electoral outcomes (Ansolabehere and Snyder, 2000; Groseclose, 2001; Aragones and Palfrey, 2002; Callander and Wilkie, 2007; Aragones and Palfrey, 2002; Bernhardt, Câmara, and Squintani, 2011). We provide a derivation from first principles of electoral candidate's valence, in the form of good judgement. In the standard definition of valence, it is independent of ideology. Here, in our microfoundation, a leadership candidate's valence is related to, and partly determined by, the ideological distribution of politicians in her group.

3. MODEL

This section sets out our basic model of leadership in a group of politicians who value informed decisions, and hold ideological preferences. The distinctive feature of our model is that a leader gathers advice from politicians before making her decision.

Our players are a group of politicians $N = \{1, ..., n\}$ who are faced with a decision $\hat{y} \in \mathbb{R}$. One amongst them—a leader—makes the decision on the group's behalf. The utility of each politician i depends on how well $\hat{y}$ matches an unknown state of the world θ. Politicians are ideologically differentiated and so the utility of i depends also on her ideological bias b_i. Bringing these elements together in a familiar quadratic loss form, we suppose that, were she to know θ, politician i's payoff $u_i(\hat{y}, \theta)$ would be a function of $\hat{y}$ according to:

$$u_i(\hat{y}, \theta) = -(\hat{y} - \theta - b_i)^2 .$$

With this specification each politician i's ideal policy is $\theta + b_i$: she would like the policy implemented to be related to the state while accounting for her idiosyncratic bias. We assume without loss of generality, that $b_1 \leq b_2 \leq ... \leq b_n$, and use the notation $\beta_i = b_{i+1} - b_i$, for all $i = 1, ..., n-1$. The vector of ideologies $\mathbf{b} = \{b_1, ..., b_n\}$ is common knowledge. The unknown state θ is uniformly distributed on $[0, 1]$.

Each politician i has some private information on θ. Specifically, conditional on θ, i holds a signal s_i, which takes the value one with probability θ and zero with probability $1 - \theta$. Politicians can communicate these signals to the leader before the decision is taken. A player's willingness to provide truthful advice may depend on who among them is selected as the leader. For example, a player i may be unwilling to truthfully reveal a signal $s_i = 1$ if her ideology b_i is to the left of the group leader's ideology. Supposing that player j is selected as the leader, we say that each politician i may send a message $\hat{m}_{ij} \in \{0, 1\}$ to her. A pure communication strategy of player i to j is thus a function m_{ij} that depends on s_i. Given leader j, let $\mathbf{m}_{-j}$ be the profile of communication strategies m_{ij} of players $i \neq j$. After communication takes place, the leader chooses $\hat{y}$ so as to implement her preferred policy. We denote a decision strategy by leader j as $y_j : \{0, 1\}^n \to \mathbb{R}$.

For a given leader j, an equilibrium consists of the strategy pair $(\mathbf{m}_{-j}, y_j)$ and a set of beliefs that are consistent with equilibrium play. Our equilibrium concept is pure-strategy Perfect

Bayesian Equilibrium. Up to relabelling of messages, each equilibrium pure communication strategy m_{ij} from a player i to a leader j may be either *truthful,* in that i reveals her signal to j, so that $m_{ij}(s_i) = s_i$ for $s_i \in \{0, 1\}$, or *"babbling"* (that is, uninformative of s_i), and in this case $m_{ij}(s_i)$ does not depend on s_i.[10] We interpret the politicians who adopt the truthful strategy with respect to j as the trustworthy associates of that leader.

Fixing a leader j, there may be multiple equilibria $(\mathbf{m}_{-j}, y_j)$. For example, the strategy profile where all players "babble" is always an equilibrium. Because of equilibrium multiplicity, the ranking of leaders and the leadership selection depend upon the choice of equilibrium: for the same leader j, different equilibria yield different player payoffs. To avoid ambiguities, we assume that for a given leader j, politicians coordinate on the equilibria $(\mathbf{m}_{-j}, y_j)$ that provide the highest expected payoffs to all politicians.[11] The selection of these equilibria is standard in games of communication and allows us to focus attention on leadership selection.

We consider two forms of leader selection.

The first one addresses our normative question: which leader would maximize politicians' welfare if chosen? Following the utilitarian principle, the welfare $W(\mathbf{m}_{-j}, y_j)$ associated with an equilibrium $(\mathbf{m}_{-j}, y_j)$ is the sum of players' expected payoffs:

$$W(\mathbf{m}_{-j}, y_j) = -\sum_{i \in N} E[(y_j - \theta - b_i)^2].$$

We denote $W^*(j)$ as the maximal equilibrium welfare when j is chosen as leader, and define the *optimal leaders* as the players j who maximize $W^*(j)$.

The second determines which player will be elected by majority rule. When j is the leader, each player i's payoff in an equilibrium $(\mathbf{m}_{-j}, y_j)$ is

$$U_i(\mathbf{m}_{-j}, y_j) = -E[(y_j - \theta - b_i)^2].$$

We denote as $U_i^*(j)$ the payoff associated with the equilibrium that maximizes each player's payoff among the equilibria induced by j. The Condorcet winner is the player j who defeats any other player k in a direct vote among alternatives j and k. As this winner need not be well defined when n is even, (then, the majority vote may result in a tie), we restrict attention to groups with an odd number of politicians.

4. A Leader's Trustworthy Associates

In our model a leader is informed via communication from members of the group. This takes the form of costless, or so-called "cheap talk", messages. As no one member of the group is

[10] For brevity, we abstract from the analysis of mixed strategy equilibria, which is cumbersome. In the three player case, Galeotti, Ghiglino, and Squintani (2013) demonstrates a mixed-strategy equilibrium in which one player communicates truthfully to the decision-maker, and the other one mixes between truthful communication and babbling. For some bias parameters, this equilibrium is more informative than any pure-strategy equilibrium.

[11] It can be easily shown that for any given leader j, each politician i's ranking among the possible equilibria $(\mathbf{m}_{-j}, y_j)$ is the same (see Galeotti, Ghiglino, and Squintani (2013), Theorem 2).

perfectly informed, a politician becomes better informed the more other members truthfully reveal their signals to her in equilibrium. Such politicians form her circle of *trustworthy associates*. We first define and characterize this concept before calculating its size for an arbitrary leader j. We show that the circle of equilibrium trustworthy associates is related to key primitives of our model, namely the ordering of ideological biases within the group. Therefore we can relate a leader's judgement to the same ordering.

4.1. **A Leader's Judgement.** The equilibrium strategies $(\mathbf{m}_{-j}, y_j)$ given any chosen leader j are easily derived from the analysis leading to Corollary 1 in Galeotti, Ghiglino, and Squintani (2013). Given the received messages $\hat{\mathbf{m}}_{-j}$ and her signal s_j, by sequential rationality, the leader j chooses $\hat{y}$ to maximize her expected utility. Because of the quadratic loss specification of players' payoffs, she chooses:

$$y_j(s_j, \hat{\mathbf{m}}_{-j}) = b_j + E[\theta | s_j, \hat{\mathbf{m}}_{-j}]. \tag{1}$$

Let $d_j(\mathbf{m}_{-j})$ be the number of politicians willing to truthfully advise j were she to lead the group. These politicians form the group of trustworthy associates of j. We prove (in the Appendix) that the profile $\mathbf{m}_{-j}$ is an equilibrium if and only if, whenever i is truthful to j,

$$|b_i - b_j| \leq \frac{1}{2\left[d_j(\mathbf{m}_{-j}) + 3\right]}. \tag{2}$$

An important consequence of the equilibrium condition (2) is that truthful communication from politician i to leader j is possible only if the ideological positions of i and j are sufficiently close. We use this result to derive how informed politician j would be in the event where she becomes leader.

First we note that the term $d_j(\mathbf{m}_{-j})$ is a function of the equilibrium communication strategies $\mathbf{m}_{-j}$ deployed by group members. In particular, whenever i can be truthful to j in equilibrium, then there is another equilibrium in which i "babbles" when communicating with j: since she babbles j will ignore her, and given this response there exists no profitable deviation for i. Fixing j's leadership, the equilibria $(\mathbf{m}_{-j}, y_j)$ that provides the highest expected payoffs to all politicians are such that j's information $d_j(\mathbf{m}_{-j})$ is maximal. That is, we consider the most information that j could obtain when any politician who could communicate truthfully in equilibrium will in fact do so.[12] This allows us to define d_j^* as the maximal size of the group of politician who form j's trustworthy associates. Straightforwardly, we can relate the maximal size of this group to a leader's equilibrium *judgement*.

[12]This equilibrium selection can also be motivated by the concept of focal leadership as in Calvert (1995): a leader provides a focal mechanism allowing followers to coordinate on truthful communication, with the consequence that information is aggregated optimally. Extensions could consider equilibria in which a politician threatens to babble (thus not transmit information to the elected leader) as a way to try and force her own election as leader. In the extreme case all players could commit to babble, communication would break down in equilibrium, and the most moderate politician would be always elected as leader. More interestingly, early voters could have an advantage in sequential voting because of a forward induction argument: by voting for herself, an early voter would "signal" to the others that she plans to babble if not elected.

Next we derive this leadership characteristic from first principles. In particular we can define it as a consequence of j's ideological position relative to that of other politicians in her party. To do so we define the function $N_j : \mathbb{R} \to \mathbb{N}$ as the ideological "neighbourhood" function of politician j. For any real number b, the quantity $N_j(b)$ is the number of politicians whose ideology is within distance b of her own, i.e., the number of politicians whose ideology is in j's ideological neighbourhood of size b. To formally define the function N_j, we exploit the fact that politicians are ordered according to their bias, so that

$$N_j(b) = \max\{i \in N : |b_i - b_j| \leq b\} - \min\{i \in N : |b_i - b_j| \leq b\},$$

for any real number b. For example, if the group of players who are truthful to leader $j = 5$ is $\{3, 4, 6, 7\}$, then $N_j(b) = 7 - 3 = 4$. We use the function $N_j(\cdot)$ combined with the equilibrium condition (2) to calculate the maximal size of any politician j's network of trustworthy associates

$$d_j^* = \max\left\{d \in \mathbb{N} : N_j\left(\frac{1}{2(d+3)}\right) \leq d\right\}. \tag{3}$$

This provides a simple rule to calculate d_j^* by counting the number of politicians other than j that are ideologically close to her. For example, suppose that $b_j = 0$ and the three politicians closest to j have bias distance less than $1/12$ from b_j, i.e. they have a bias in the interval $(-1/12, 1/12)$. These politicians would provide truthful advice to j were she to be selected as leader. For j to have one more trustworthy associate it must be that no member of that circle has a bias further from $b_j = 0$ than $1/14$. Interestingly, the size of the ideological neighborhood of leader j (to which a politician needs to belong to be trustworthy to j) decreases in the number of associates truthful to j. For example, a politician i with bias b_i of distance within $1/10$ and $1/8$ to b_j will be truthful to j if and only if j has no other trustworthy associate.[13]

5. Selecting the Leader

Having defined the size of a leader's network of trustworthy associates, we now turn to the question of leadership selection. We define the optimal leader as one who maximizes group welfare. In the absence of a mechanism that ensures the first best choice, it is natural to ask which leader would be chosen by the group when each casts a vote with the outcome determined by majority rule. Using the result of the previous section we show that the characteristics of optimal and majority-preferred leadership can be derived from first principles.

[13]Costly information acquisition does not change our main insights. Although it is possible that the set of associates of leader j is different and so leadership choice may differ in this scenario, nevertheless, with costly acquisition, our first main result– that a leader can rely on truthful advice only from those whose ideological preferences are similar to her own–continues to hold. If a player anticipates that he will be unable to convey any information to the leader, then he will not acquire costly information in the first place. Conversely those who do acquire information are those able to convey this information: they are ideologically close to the leader. As a consequence of this finding we can then use the same technique to derive the size of a leader's network of trustworthy associates and the tradeoffs identified in our key results to follow shall continue to apply.

5.1. **The Optimal Leader.** We first show that optimal leader selection involves trading off a politician's ideological moderation and her judgement. To formalize this insight, we denote politician j's *moderation* as $-\left|b_j - \sum_{i\in N} b_i/n\right|$, the contrary of the distance between b_j and the average ideology $\bar{b} \equiv \sum_{i\in N} b_i/n$. We have defined d_j^* as the maximal size of a leader's network of trustworthy associates. It is but a small step to relate this number to her *judgement,* the second critical and endogenous leadership characteristic. When combining the information obtained from others with her own view, a leader forms an independent judgement of the best course of action. Thus a leader's judgement is (strictly) increasing in the number of informative signals she obtains from her trustworthy associates.

In fact, and armed with these definitions, we can prove that the equilibrium ex-ante sum of players' payoffs $W^*(j)$ can be rewritten as:

$$W^*(j) = -\sum_{i\in N}(b_i - b_j)^2 - n\frac{1}{6(d_j^* + 3)}. \tag{4}$$

Expression (4) decomposes the welfare function into two elements: the aggregate ideological loss $\sum_{i\in N}(b_i - b_j)^2$ associated with the decision taken by j, and the aggregate residual variance of her decision $n\frac{1}{6(d_j^*+3)}$. Evidently, a more moderate leader, whose bias is closer to average ideology $\bar{b}$, makes the aggregate ideological loss $\sum_{i\in N}(b_i - b_j)^2$ smaller. Further, the residual variance $\frac{1}{6(d_j^*+3)}$ is inversely related to the size of the leader's maximal informant set d_j^* and hence to her judgement.[14] The optimal leader j maximizes $W^*(j)$. Thus, optimal leader selection takes into account each politician's moderation and her endogenous judgement that are related to the core primitives of our model, namely the ideologies of members of the group.

Leader j's moderation can be understood spatially as the relative position of j's bias b_j with respect to the whole ideology distribution $\mathbf{b} = \{b_1, ..., b_n\}$ in the group. In fact, every element of $\mathbf{b}$, even extreme ones, matters for the determination of the average ideology $\bar{b}$. In this sense, moderation is a "global" property of j's ideology b_j with respect to the distribution $\mathbf{b} = \{b_1, ..., b_n\}$.

On the other hand, judgement is a "local" property of j's ideology b_j within $\mathbf{b} = \{b_1, ..., b_n\}$: it depends only on how many other politicians are ideologically close to j, in the sense defined by equation (3). The leader's understanding is thus defined by those close to the leader, or adopting Machiavelli's text, by *"the men he has around him".* This analysis of the role played by the local ideological distribution is, to our knowledge, novel in the large contemporary and formal literature on collective choice; though it echoes the insights of Machiavelli made in *Il Principe*, 500 years ago.

We summarize our findings as follows.

[14] Mathematically, the residual variance $\frac{1}{6(d_j^*+3)}$ corresponds to the inverse of the precision of the leader's decision.

Proposition 1. *The optimal leader j is determined by ideological moderation, the proximity of b_j to the average group ideology $\bar{b}$, and by good judgement, her number d_j^* of close-minded associates.*

5.2. **Electing the Leader.** We now determine which politician is elected as leader by a simple majority decision taken within the group. Each player i's utility as a function of the leader's identity j is:

$$U_i^*(j) = -(b_i - b_j)^2 - \frac{1}{6(d_j^* + 3)}. \tag{5}$$

As in equation (4), the first term on the right hand side illustrates the ideological loss $-(b_i - b_j)^2$ suffered by each member of the group i when j is chosen as leader. The second term illustrates player i's preference for an informed leader j, as it increases in the judgement d_j^*. We note that player utilities are not single-peaked with respect to a leader's identity: a politician who is ideologically distant may in fact be better informed, and so have better judgement, than one who is ideologically similar. While Black's theorem does not apply in this setting, so leadership choice under majority rule is far from straightforward, we can, nevertheless, make progress by showing that utility functions are single-crossing.

Lemma 1. *The utility functions $U_i^*(j)$ are single crossing in i and j: if $i < i'$ and $j < j'$, then $U_{i'}^*(j) > U_{i'}^*(j')$ implies $U_i^*(j) > U_i^*(j')$; and if $i > i'$ and $j > j'$, then $U_{i'}^*(j) > U_{i'}^*(j')$ implies $U_i^*(j) > U_i^*(j')$.*

As a consequence of this result, we appeal to Theorem 2 by Gans and Smart (1996) to show that the player with median ideology will determine the outcome of the election. The unique Condorcet winner of the election game is the politician j who maximizes the expected payoff of the median player.

Proposition 2. *The group elects as leader the player j who maximizes the utility $U_m^*(j)$ of the median politician $m \equiv (n+1)/2$. The collective choice considers the ideological proximity of any player j to m, as well as j's judgement d_j^* that is determined by her number of close-minded associates.*

Having established the outcome of a majority election, we can compare it with the optimal leader selection by inspecting expressions (4) and (5), the latter for $i = m$. As in the earlier case there is a trade-off between moderation and judgement: the Condorcet winner j keeps both the ideological loss $(b_m - b_j)^2$ and the residual variance $\frac{1}{6(d_j^*+3)}$ as low as possible. Just as with optimal leadership, the majority choice involves a trade off between the desire for a moderate leader and that for a leader with good judgement which, in turn, stems from having a large group of close-minded associates. Beyond this similarity there is a critical difference and it is this: whereas a majority preferred leader makes this trade off decision by considering only her own payoff, by contrast, an optimally selected leader would consider the preferences of the entire group. Straightforwardly, and as the weights placed on these two features of good leadership are different in our key expressions, the majority choice of leader may not be optimal. As we shall see, the implications are surprising in that we identify instances in which the median politician's utility U_m^* places less weight on moderation (and more on judgement) than the group's welfare W^*. Thus majority choice may be inefficient because it places *too much weight* on the leader's judgement.

6. What Makes a Good Leader?

Our analysis relates the characteristics that define good leadership—moderation and judgement—to the communication structure that emerges in the equilibrium of our model. The importance of the former is well known. Indeed it is easy to see that if there were no informative signals (or just no communication) in this game, then the chosen leader would be the median politician $m,$ while the optimal one would be the one whose bias is the closest to the average bias $\bar{b}$. On the other hand, the role played by judgement, that in turn is related to a leader's trustworthy associates, is novel and central to the results that follow.

6.1. **Moderate Leadership.** A natural question to ask our model is thus under which conditions on the primitive parameters (the ideology distribution $\mathbf{b}$), is the most moderate politician the optimal leader and the majority-preferred one. Evidently, this is the case when there are only 3 politicians in the group. For then the median is the most moderate politician and she cannot be less informed than either of the others: If she is willing to communicate truthfully with her neighbors, then at least one of them is willing to be truthful to her.[15]

Moving beyond the three-player case, we illustrate sufficient conditions such that the optimal leader is also the most moderate politician. Doing so, we consider the situation in which politicians are distributed at even distances with respect to their ideology on the line. Because n is odd, and by symmetry of the ideological distribution, the median m is the most moderate and has at least as many trustworthy advisers as any other politician. Then there is no tradeoff between a leader's moderation and her judgement. As the median politician is as informed as anyone else she should take the decision on behalf of the group and, indeed, she would be the unique choice of the majority.

We formalize this insight in the following proposition that proves a stronger result. We show that the median politician m is elected by the majority as leader and is also the optimal leader when the ideology distribution is symmetric around m and "single peaked" at m, in the sense that politicians are weakly more ideologically clustered as they get closer to m. Formally, we define the ideology distribution $\mathbf{b}$ as single peaked and symmetric at m when for every $i = 1, ..., m-1$, $\beta_i = b_{i+1} - b_i$ weakly increases in i and $\beta_i = \beta_{n-i}$. Evidently, the case in which

[15]This reasoning can be pushed one step further. The most extreme politicians 1 and n can never be chosen as leaders, as they cannot have better judgement than their neighbors, who are also more moderate. If player 2 is truthful to 1 in the most informative equilibrium, then also 1 is truthful to 2, whereas every other player $i > 2$ is ideologically closer to 2 than to 1, so that if i is truthful to 1, then i is also truthful to 2 in the most informative equilibrium (and not necessarily viceversa).

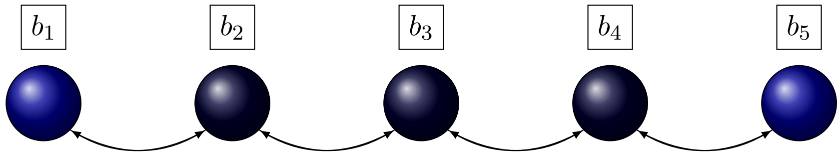

FIGURE 1. Leadership Choice with Equidistant Bias: this figure depicts the case where $b_{i+1} - b_i = \beta$ for all politicians $i = 1, ..., 4$ and $1/12 < \beta \leq 1/10$. An arrow linking two politicians illustrates that truthful communication is sustained between them.

politicians' ideologies are evenly distributed on the line (i.e., there is a constant $\beta > 0$ such that $b_{i+1} - b_i = \beta$ for all $i = 1, ..., n-1$) is a limit case covered by the definition of $\mathbf{b}$ as single peaked and symmetric at m.

Proposition 3. *When politicians' ideologies* $\mathbf{b}$ *are single peaked and symmetric at* m, *the median politician* m *has also the best judgement. She is the optimal leader, and will always be elected as leader by majority.*

The result is depicted in Figure 1 for $n = 5$ and ordered left-right biases b_1 to b_5. In the figure, for each of three politicians $j = 2, 3, 4$, their maximal amount of equilibrium information is $d_j^* = 2$. Then the optimal leader, and the one who is indeed chosen by the group, is player 3.

Proposition 3 relates the core characteristics of leadership to reveal that, with biases single-peaked and symmetric at m (as in the case of equidistant biases depicted in Figure 1) the most moderate politician is also (weakly) better informed and so has better judgement. The significance of our result lies in the fact that, when ideologies in the group are single-peaked and symmetric at m, there is no tradeoff between moderation and the capacity to gather reliable advice. This underlines that a necessary condition for a politician who is not the most moderate to be the optimal leader or the elected one is that ideologies are clustered away from the median.[16] We explore this tradeoff next, focussing on the case of 5 politicians for ease of exposition.

6.2. **The Case with 5 Politicians.** To explore the tradeoff between moderation and judgement we study in-depth the case of 5 politicians, parametrized by the bias differences $\beta_i = b_{i+1} - b_i$ for players $i = 1, ..., 4$. Our case is suitably rich (allowing us to identify properties of the ideology distribution that provide novel and interesting findings) yet simple (so that we can do so in a clean and clear manner).

First we note that it can never be optimal that 1 and 5 lead the group and neither will they be elected by the group as leaders. Players 2 and 4 can be chosen as leaders if and only if

[16]One example of such ideological clusters has been documented in Argentina. Politicians and policy experts come from two separate "schools." One is the traditional Peronist or leftist "Intelligentsia," mainly composed of social scientists and administrators that are entrenched in the Argentinian tradition. The second school are the "Chicago/Minnesota boys," economists trained in "fresh water" US PhD programs. Similar ideological clusters appear in countries such as France.

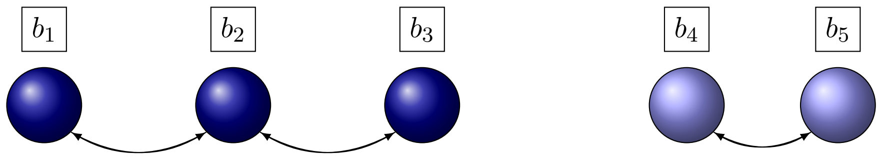

FIGURE 2. Leadership Choice with Non-equidistant Bias: in this figure, $\beta_1 \leq 1/10$, $\beta_2 \leq 1/10$ and $\beta_4 \leq 1/10 < \beta_3$. All immediate neighbor pairs truthfully communicate, with the exception of 3 and 4.

they have better judgement than player 3. Also, interchanging β_1 with β_4 and β_2 with β_3 then players 2 and 4 are symmetric to each other. Hence, it is with no loss of generality that we restrict attention to parameter values for which $W^*(2) \geq W^*(4)$, so that welfare is strictly greater when 2 is the leader rather than 4, and for which $U_3^*(2) \geq U_3^*(4)$, so that the only possible Condorcet winners are 2 and 3.

Player 2 has better judgement than 3 when she can count on more trustworthy associates, that is when $d_2^* > d_3^*$. Using condition (3), we calculate (in the Appendix) all cases for which the condition $d_2^* > d_3^*$ holds, and determine the restrictions each one of them imposes on the parameters β_i, $i = 1, ..., 4$. Here, we illustrate our findings in the case in which $d_2^* = 2$ and $d_3^* = 1$. This holds when $\beta_1 \leq 1/10$, $\beta_2 \leq 1/10$ but $\beta_3 > 1/10$. In this case, players 1 and 3 are sufficiently close to 2 to be trustworthy, whereas the median politician 3 can trust only 2, but not 4. So politician 2 has better judgement, while 3 is more moderate. This scenario is illustrated in Figure 2 where $\beta_3 = b_4 - b_3$ is such that (in contrast to Figure 1) player 4 can no longer communicate truthfully with 3. As the player ideologies are not distributed at even distances there may be a tradeoff between judgement and moderation and, moreover, surprising consequences of ideological shifts (such as that of player 4 in Figure 2 relative to Figure 1) which we discuss below.

6.3. **The Tradeoff between Judgement and Moderation.** Whether the choice between the more moderate politician 3 and politician 2, who has better judgement, is resolved in favor of either depends upon $\beta_2 = b_3 - b_2$. This determines how extreme is 2 relative to 3. This pins down the majority choice: if 2 will make a more informed decision than 3, then 3 would be happy for her to do so so long as the ideological distance between them is not too large. This allows us to define a majority threshold with respect to β_2. In fact, we can relate the optimal choice and the majority choice between politician 2 and 3 according to the size of β_2 relative to the other primitives of the model, as demonstrated by the Lemma below that allows us to define a welfare threshold. Let $\delta = b_5 - b_3 + b_4 - b_3 - (b_2 - b_1)$ be the sum of the ideological distance from 3 of the players to the right of 3 (i.e., 4 and 5) net of the distance from 2 of player 1, the only player to the left of 2.

Lemma 2. *Consider the case of* 5 *politicians, with the above restrictions:* $W^*(2) \geq W^*(4)$, $U_3^*(2) \geq U_3^*(4)$, $\beta_1 \leq 1/10$, $\beta_2 \leq 1/10$, $\beta_3 > 1/10$ *(and hence* $\delta > 1/10$*).*

- *If* $\beta_2 < \frac{1}{2\sqrt{30}} \approx 0.0913$, *then the Condorcet winner is politician* 2, *else, the most moderate politician* 3 *wins the majority choice.*
- *If* $\beta_2 < \tau(\delta) \equiv \sqrt{\delta^2 + 1/24} - \delta$, *then the optimal leader is* 2, *and else it is* 3.
- *There is a unique* $\bar{\delta} > 1/10$ *such that* $\tau(\delta) > \frac{1}{2\sqrt{30}}$ *for all* $\delta < \bar{\delta}$ *and* $\tau(\delta) < \frac{1}{2\sqrt{30}}$ *for all* $\delta > \bar{\delta}$.

The result defines a welfare threshold, $\tau(\delta)$. The group is better off when player 2 takes the decision if and only if β_2, the ideological distance between players 2 and 3, is below $\tau(\delta)$. Intuitively, it is optimal that 2 leads the group when her better judgement, combined with the benefits to those the left of the spectrum ($b_2 - b_1$) are not outweighed by the ideological loss incurred by those to the right ($b_4 - b_3 + b_5 - b_3$). We prove in the Appendix that τ is strictly decreasing in δ, with $\tau(1/10) > 1/10$ and $\tau(\delta) \to 0$ for $\delta \to \infty$.

Lemma 2 also defines the majority threshold $\frac{1}{2\sqrt{30}}$. This takes a simpler form as it depends only on the median player. She may obtain a more informed outcome when 2 takes the decision and this yields a constant addition to her utility. This comes at an ideological cost β_2. Thus the group chooses 2 as leader if and only if β_2 is below a threshold given by the constant $\frac{1}{2\sqrt{30}}$.

The final part of lemma 2 reveals that in equilibrium the welfare threshold $\tau(\delta)$ can either be larger or smaller than the majority threshold $\frac{1}{2\sqrt{30}}$ depending on how large δ is. In the former case, we can distinguish three possibilities on the basis of β_2: for small β_2, i.e., $\beta_2 < \frac{1}{2\sqrt{30}}$, the non-moderate politician 2 is both the Condorcet winner and the optimal leader; for large β_2, specifically, $\beta_2 > \tau(\delta)$, the most moderate politician 3 is both the Condorcet winner and the optimal leader; in the intermediate case $\frac{1}{2\sqrt{30}} < \beta_2 < \tau(\delta)$, the optimal leader is the most informed politician 2, whereas the majority elects politician 3. This identifies an instance in which majority voting leads to a moderate but inefficient choice of leader.

A perhaps more interesting and unexpected case arises when $\tau(\delta) < \frac{1}{2\sqrt{30}}$. Again, for small β_2, the non-moderate politician 2 is both the Condorcet winner and the optimal leader, and for large β_2 it is politician 3. Now inefficiency arises in the intermediate case in which $\tau(\delta) < \beta_2 < \frac{1}{2\sqrt{30}}$. Although the optimal leader is the most moderate politician 3, the majority elects instead a relatively extreme leader in politician 2.

The logic behind this result is simple. The median politician may trade off moderation and judgement in a way that differs from the optimal choices made by a social planner. Starting from her ideal point, she may sacrifice a policy more in line with her bias for a more informed outcome. In our 5 player example, she will indeed do so when β_2 is sufficiently small and $d_2^* > d_3^*$. Choosing a leftist leader then benefits the median and of course leftist members of the group. But it harms the right-wing members 4 and 5, who bear costs $(b_4 - b_2)^2$ and $(b_5 - b_2)^2$ respectively. Because the ideological loss function $(b_i - b_j)^2$ is convex in the ideological distance $|b_i - b_j|$, the leadership move from 3 to 2 is more harmful to rightwing politicians than it is beneficial to the leftist ones. Then it may be the case that a social planner would force the median politician to take the decision, if only she could.[17]

6.4. **Comparative Statics.** Further analysis of the five-player case uncovers some interesting comparative statics results. Our comparative static exercise is as follows. We first consider leadership choice with a given set of primitives (players ideological biases) and then consider what happens to leadership choice when a player changes her ideology from right (left) to left (right). We show that this can in fact induce a shift in leadership in the opposite direction.[18]

To illustrate, consider a benchmark case with players ideologies evenly distributed apart and where politicians 2, 3, and 4 can all count on the truthful advice of their ideological neighbors, so that β_i is constant in $i = 1, ..., 4$, and $1/12 < \beta_i \leq 1/10$. Then following Proposition 3, politician 3 is (strictly) most moderate and has (weakly) better judgement among the five; hence she is elected as leader and this choice is also optimal for the group. Suppose now that the ideology of the centre-right player 4 moves rightward and so away from that of the median player 3 and that, as a consequence, they are no longer truthful to one another (i.e., suppose that β_3 increases so as to become larger than $1/10$). Now politician 3 has lost a (previously) trustworthy associate. It is now possible that the centre-left politician 2 is the Condorcet winner of the election game—3 will delegate authority to her, despite not being the most moderate politician. Indeed, by Lemma 2, we know that this is the case when $\beta_2 < \frac{1}{2\sqrt{30}}$. Hence, the ideological movement of a player towards a more extreme position may induce a leadership change in the opposite direction.

Conversely, suppose that the 5 politicians are such that, in the benchmark case with evenly distributed ideologies, there is no truthful communication across players, i.e., β_i is constant across $i < n$ and $\beta_i > 1/8$. Suppose now that the leftist politicians 1 and 2 become more moderate, so that now politician 2 can count on the truthful advice of players 1 and 3 (formally, suppose that β_1 and β_2 decrease, so that they both become smaller than $1/8$). Because player 4 is still not truthful to 3, leadership switches from the median player 3 to the centre-right player 2, again, when $\beta_2 < \frac{1}{2\sqrt{30}}$. Here, the rightward ideological movement of leftist players moves leadership choice (and hence policy outcomes) in the opposite direction. Thus moderation allows them to capture control of leadership of the group.

In both cases, and unexpectedly, a change in the ideological distribution by which politicians (weakly) move right leads to a shift in the group decision to the left.

[17] In a situation of multiparty competition with advisors the analysis is not straightforward, as Lemma 1 does not apply. Nevertheless the tradeoff between moderation and judgement may be present. Dewan and Squintani(2016) analysed a world with multiple groups that emerged endogenously and whose influence increased in proportion to their size. The application of that model is to internal party factions. They show the benefits of (possibly) extreme factions that aggregate information and that induce a tradeoff with moderation. A direct extension of their model is to government formation between different parties.

[18] Such non-monotonocities are of course ruled out in the optimal selection of the leader in the absence of communication; and, following on from comments above, neither can they occur in the absence of communication when the leader is elected under the Condorcet procedure.

We summarize our findings for this section in the following result.

Proposition 4. *When the distribution of ideologies is not single peaked and symmetric, a politician other than the most moderate one can be the optimal leader and the majority choice.*

Relative to the group of politicians, the median player weighs judgement more than moderation: she may lead when it is in the group's interest that another with better judgement is chosen; and she may not be chosen when it is optimal that she leads.[19]

If politicians become more moderate (extremist), they capture (lose) control of the group, and turn the group policy closer to (away from) their views.

6.5. **Discussion.** The findings of our positive study of leadership highlight the predictive importance of judgement. The fact that the majority choice may place too much weight on this characteristic is perhaps surprising, the more so when interpreting the decisive median vote in the election as a decision to delegate authority to leaders with specific characteristics. This notion goes back to Schelling (1960) who in his seminal book *The Strategy of Conflict*, discussed the use of delegates with particular characteristics as a way to credibly commit a negotiating party to a position. He suggested that agents in bargaining situations may transfer power to stubborn negotiators.[20] Seen in this context, we note that our model is one where the median player can choose either to take the decision herself or delegate to another politician. She chooses the latter option when another member of the group has more information and so better judgement. The surprising, and we believe novel, finding is that the median may delegate to another when it is in the group's interest that she executes the decision herself.

As noted in our earlier discussion of related literature, when viewing our model as one of implicit delegation, Proposition 4 reveals a failure of the "ally principle", that states that the principal will always delegate to an agent who is ideologically closest to her. Indeed it has been noted that when viewing the set of possible principals and agents as heterogenous groups rather than as unitary actors, and when agents are imperfectly informed, then the ally principle may not hold. Our model combines these elements—multiple players with different

[19]One interesting question is whether these kinds of inefficiencies could be eliminated allowing for collective leadership, instead of considering only individual leaders. Within this paper's model, collective leadership can be represented as vectors α of "shares of leadership" such that $\alpha_i \geq 0$ for all players $i = 1, ..., n$, and $\sum_{i=1}^{n} \alpha_i = 1$, identifying the support of α as the set of leaders in the group. We show in Appendix that, while the possibility of collective leadership weakly improves utilitarian welfare over individual leaders, it does not eliminate the possibility of the inefficiencies identified here. The analysis for 5 players of Lemma 2 is extended to show that, also when allowing for shared leadership, there exist bias profiles b such that the optimal shared leadership vector α gives all authority to the median politician 3, but majority voting selects the vector α' that gives all authority to player 2.

[20]By contrast Chari, Jones, and Marimon (1997) suggest that the opposite occurs in voting contexts. Harstad (2010) draws a distinction between the political power of extreme politicians and the bargaining power of more moderate ones, and analyzes the trade off between them.

preferences—and reveals conditions on the primitives of our model under which the ally principle holds and those where it does not and, moreover, provides a framework within which to understand the welfare consequences of the failure of the ally principle.

Beyond this normative perspective, our analysis has consequences for the empirical analysis of a number of institutional settings operating under majoritarian principles, where, following Black, the decisive player is the median. We mention two possible applications of our ideas.

A large body of literature has explored the process of nominations and appointments to majoritarian institutions. For example, Krehbiel (2007) and Rohde and Shepsle (2007) have analysed the process by which nominations are made to the Supreme Court by an ideologically disposed President and majority approved (or not) by a Senate, in which senators anticipate the consequences of such an appointment on court decisions that are likewise made under majority rule. As the situation involves multiple inter-dependent institutions, as well as multi-player interactions with each of these institutions, the set of possible strategies to consider is large. These models are tractable, however, due to the assumption that within the Senate and the Court the pivotal player (politician, judge) is the one with the median preference. Appointments can then be considered with respect to whether or not they change the identity of that player, and, hence, these models go by the description of "move the median" games. Specifically, a rational Senate member considers whether a proposed appointee changes the identity of the median court member. She is the critical player, since, "an opinion must gain the assent of five justices, the median justice and four justices on one side or another" (Rohde and Shepsle, 2007). Our analysis suggests, by contrast, that it is not just the identity of the median that is important in determining a groups choice under majority rule. This implies that the results of the "move the median" game may be different when considering preferences that depend on private information.

A second and related research topic is the writing of the Supreme Court decision. The exact procedure is elaborate, but again things simplify if one assumes that the opinion is either directly written by the median justice (referred to as the median justice model) or must be approved by her as part of a bargaining process. A straightforward extension of our five player group, depicted in Figure 2, to a nine member Court would yield different insights. Specifically, our analysis suggests that the opinion of a justice other than the median may achieve majority support. As already noted in our introductory remarks, in his review of the field, Clark suggests that relaxing the complete information assumption in standard models may yield new insights. Indeed our analysis would appear to confirm that this is in fact the case.

We postpone a more extensive application of our ideas to these cases to future research. Here, instead, we focus attention on an immediate and we believe first order extension of our model. As already noted, our analysis of group choice of leader provides insights that differ from those provided by a straightforward application of Black's Theorem. A noted application of Black's ideas is via the workhorse spatial model of party competition. We study a version of that model with two parties who each choose a leader who then competes in a general election.

7. A Model With Electoral Competition

The analysis in the previous section reveals that a relatively extreme leader may be chosen if she has good judgement. Also, it highlighted a peculiar non-monotonic comparative static result: a rightward ideological shift by a politician can induce a leftist choice of leader, and vice-versa. Next we explore whether these surprising effects survive political competition. Will political groups such as parties choose relatively extreme leaders when their candidates face an electoral test in the form of a general election?

In order to explore this, we analyze a model of two party competition that incorporates different democratic selection methods. We consider a world where each party first chooses an electoral candidate (who we identify as the party leader, although this is not needed for our arguments) via an internal election involving politicians, members, and/or registered votes. Party leaders then compete in a general election. As in the now standard citizen candidate model of Osborne and Slivinski (1996) and Besley and Coate (1997), the winner of the election implements her ideal policy. The difference (with the standard model) is that she does so only after consultation with other informed politicians in her own party. We assume that politicians and the electorate as a whole value informed decisions made by elected office holders, but are ideologically differentiated and anticipate final outcomes when casting their votes.

7.1. **Model.** Suppose that there are two parties, A and B. The top politicians in party A consist of the set of players N_A and those in B consist of politicians N_B. At the beginning of the game, parties chooses leaders $\{a, b\}$. To make our results general, we do not commit to a specific leader choice model. We assume only that each party selects as leader the strongest possible candidate, defined as the politician within the party who defeats the largest possible number of candidates from the other party in the general election.[21] To simplify the exposition, we consider only ideology and party profiles such that there is only one such politician in each party. The eventual winner $j \in \{a, b\}$ of the general election then implements the final policy $\hat{y} \in \mathbb{R}$. Candidates cannot commit to electoral promises, and so the winner implements her preferred policy after consultation with her party's top politicians.

There is a continuum of citizens, which includes the finite set of politicians $N_A \cup N_B$. The preferences of each citizen k, including politicians, are expressed by:

$$u_k(\hat{y}, \theta) = -(\hat{y} - \theta - b_k)^2,$$

[21]These assumptions would be satisfied in a number of micro-founded models of leader selection within our framework. For example, suppose that each top politician in either party can participate in a primary, held under plurality rule and at a small cost $c > 0$ to herself, to become the leader of the party. These primaries yield leaders who obtain (small) ego rents, $r > c$, and only citizens registered with the party can vote in the primaries.

where b_k is the ideological bias of citizen k relative to the median voter in the general election, who we assume to have bias equal to zero, without loss of generality. As before, the utility of k depends on how well $\hat{y}$ matches an unknown state of the world θ together with her ideological bias b_k. We single out politicians who belong to the set $N_A \cup N_B$, denote them with indexes i, and maintain the assumption that b_i is increasing in i and therefore that all politicians in A are to the left of all politicians in B.

The remainder of our model is as before. Each top politician i has some private information on θ. After the general election takes place, each i observes a signal $s_i \in \{0,1\}$ of θ such that $\Pr(s_i = 1|\theta) = \theta$. And before the elected policy-maker j chooses $\hat{y}$, each politician i can communicate by sending a message $\hat{m}_{ij} \in \{0,1\}$ to j. We assume that there are social conventions stifling communication across parties, and that the elected politician has truthful associates only within her own party. We thus consider equilibria in which the politicians from the opposite party do not reveal any information to her.

As in the previous section, each voter k evaluates a candidate j on the basis of both j's ideological proximity $(b_k - b_j)^2$ and her judgement, identified by the number of j's trustworthy party fellows d_j^*, according to the, by now usual, decomposition:

$$U_k^*(j) = -(b_k - b_j)^2 - \frac{1}{6(d_j^* + 3)}.$$

Because each voter k evaluates a candidate j's judgement favorably, regardless of her ideology, we can think of judgement as *valence*. Here, it is endogenously determined by j's network of trustworthy party fellows.

As a consequence of Lemma 1, preferences satisfy the single crossing condition with respect to the choice of leader in the general election. Moreover, the play of weakly undominated strategies in (the subgame that represents) the election implies that each voter chooses her preferred candidate $j \in \{a, b\}$. As a consequence, candidate a will be elected with certainty if and only if $U_0(a) > U_0(b)$, where we take 0 to be the index of the median voter.

7.2. **Policy Divergence.** A natural benchmark for comparison is an otherwise identical model in which players do not communicate to the elected leader before she chooses the policy $\hat{y}$. As no communication can take place, so no information about θ can be aggregated, only the vector of ideologies $\mathbf{b}$ are relevant to votes cast in either primary or general election. As these are common knowledge, the game then boils down to a simple one of perfect information. It is then straightforward to prove that $U_0(a) > U_0(b)$ only when the policy bias of leader a is closer to 0 than that of b, so that the most moderate candidate in each party is chosen as leader.

Fact 1. *Suppose that politicians cannot communicate to the politician j who wins the general election. Then the winner of the general election is the player whose ideology is closest to that of the median voter in the electorate.*

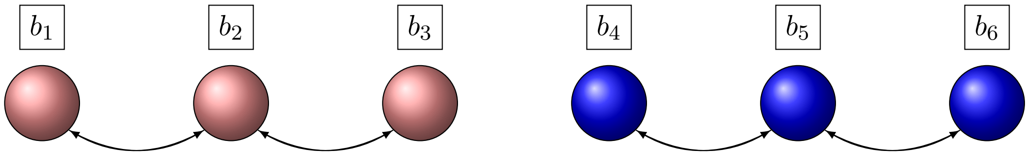

FIGURE 3. Party Competition with Equidistant Bias: The figure illustrates a case where politicians $1, 2, 3$ belong to party A and $4, 5, 6$ to party B. Within each party each politician is informed by her immediate neighbour.

Whenever politicians are sufficiently ideologically distant from each other that they can never communicate truthfully (even to their closest ideological neighbour), then the winner of the general election is the candidate with bias closest to zero. Beyond this simple case, it is immediate that, in our model of electoral competition, party leaders need not be moderate. The analysis follows our earlier logic: politicians with a large network of truthful informants may be preferred by the median voter even if they have relatively extreme ideologies. In fact we can reveal new insights. First we show that the winner of the general election need not be the most moderate politician (i.e., the politician whose ideology is closest to the median voter), even in circumstances in which the politicians' ideologies are evenly distributed in the ideological spectrum. Thus our finding stands in sharp contrast with Proposition 3. Why so? When politicians are partitioned in competing parties, the most moderate politician (with respect to the electorate as a whole) is at the extreme end of the ideological spectrum within her own party. This constrains the pool of trustworthy associates she can rely upon and so hampers her ability to take informed decisions if elected to office.

Proposition 5. *Even if the ideologies* b *of the potential candidates* $N_A \cup N_B$ *are evenly distributed, so that* $b_{i+1} - b_i = \beta$ *for some* β *and all* $i = 1, ..., n-1$, *it need not be the case that the winner of election is one of the most moderate politicians.*

This insight is demonstrated by the 6-player example depicted in figure 3. There are 6 politicians, with ideologies such that $b_{i+1} - b_i = \beta$ for all $i = 1, ..., 5$, arranged symmetrically around the median ideology zero, so that $b_3 = -\beta/2$ and $b_4 = \beta/2$. The leftist politicians $1, 2$, and 3 (lighter shading) belong to party A and the others (darker shading) to party B.

Following our earlier analysis, unless politicians 2 and 5 can count on more trustworthy advisers than 3 and 4, in equilibrium, the latter will be selected by their parties and tie the general election. Because of the symmetry of b we can focus attention on the challenge between 2 and 3 for leadership of party A. Politician 2 has better judgement when $d_2^* = 2$ and $d_3^* = 1$, which requires that $\beta \leq 1/10$ and that $2\beta > 1/10$.

It is then relatively straightforward to identify a condition on β such that the median voter in the general election would prefer that candidate 2 is chosen by party A, that is $U_0(2) > U_0(3)$. Specifically, we show in the Appendix that this is the case when $1/20 < \beta < \frac{1}{4\sqrt{15}} \approx 0.0645$. By symmetry, and since the median has zero bias, it must also be that $U_0(5) > U_0(4)$. Hence, in the unique equilibrium, party A chooses politician 2 as leader, and party B chooses politician 5. In sum, the chosen leaders are not the most moderate candidates 3 and 4.

This result provides a new take on the documented divergence among candidates in two party elections. Even if two parties compete for power, they will not necessarily be the most moderate candidates, so that convergence to the median will not take place.[22]

7.3. **Moderation and Party Cohesiveness.** The second novel insight highlights the value of a party's ideological cohesiveness. Interpreting a party's cohesiveness as the ideological distance among its top politicians, we find that a more cohesive party can defeat a larger, less cohesive, one in a general election. This can occur even though the larger party can draw information from a larger set of informed politicians. And it can occur even though the leader of the larger party has ideological views that are closer to those of the median voter. Why? The leader of the more cohesive party can count on more trustworthy associates than her opponent. The median voter anticipates that, as a consequence, she will have better judgement.

Finally, we find that the outcome of the election may depend on the whole ideological distribution and often in a very subtle way. For example, suppose that (repeating the comparative static exercise from earlier) the ideology of the leader of a party becomes more moderate. Should she be elected then she will implement her preferred policy which, in turn depends upon the advise she obtains. Surprisingly she may in fact lose the general election as a result of her new found moderation. (Of course, the opposite can happen: a politician may lose the election if her views become more extreme). Intuitively, this occurs because, as the leader's views become more moderate, the ideological distance with others in the party increases. As a consequence, the leader may lose the benefit of truthful advice from erstwhile allies. This result is, to the best of our knowledge, both novel and unsupported by any variant of the standard spatial model found in the literature.

Proposition 6. *A large party may lose the election to a smaller, more cohesive party, even if it can draw information from a larger number of top party politicians and though its leader is the candidate in the general election whose views are closest to the median voter's.*

The outcome of the election may depend on the whole ideological distribution of the top party politicians in subtle ways: For example, a party leader may lose the general election only because her views become more moderate (closer to the median voter).

The result can be demonstrated by means of the following example, illustrated in figure 4. Suppose that there are 5 politicians, with $b_2 < 0 < b_3$ and $|b_3| < |b_2|$. Politicians 1 and 2 belong to party A and $3, 4, 5$ belong to party B. In this example, there are two senses in which party

[22]The possibility that a less moderate candidate may win an election because of her higher valence was earlier conjectured by Stokes (1963). In the literature, valence is defined residually (as all candidate characteristics valued by voters regardless of their ideologies). Our model of elections and advice provides a microfoundation of valence that refers to a leader's judgement.

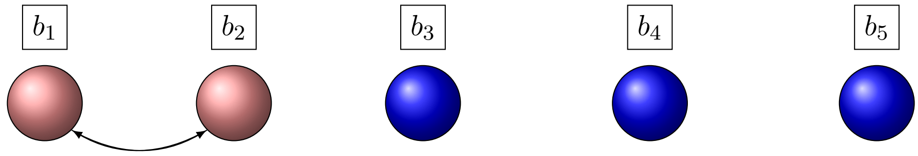

FIGURE 4. Party Competition with a Majoritarian Party: illustrates a case where politicians 1 and 2 are in party A and $3, 4, 5$ are in party B; $b_2 < 0 < b_3$, $b_3 < |b_2|$, $\beta_1 \leq 1/8$, $\beta_3 > 1/8$ and $\beta_4 > 1/8$. Politicians truthfully communicate within party A but not within party B.

B is advantaged: there is a larger set of top politicians from whom the leader could draw upon for information; and in player 3, it has a potential leader whose views are closest to those of the median voter, whose ideal point is equal to zero. This implies that, were there no possibility of communicating private information, party B would always win the election by selecting politician 3 as its leader. Party B is not only larger, it is also "ideologically majoritarian" in that, in contrast to party A, it is able to put forward candidates whose ideological perspective appeals to a majority of the electorate.

Given the advantage of B in fielding more moderate candidates, following Lemma 1, a politician from party A can only be elected due to her better judgement. For this to be the case it must be that $d_2^* = 1 > d_3^* = 0$, as there is only one other informed politician in party A. This situation requires that $\beta_1 \leq 1/8$, $\beta_3 > 1/8$ and $\beta_4 > 1/8$, and is depicted in Figure 4. It is then not difficult to find conditions under which the median voter prefers politician 2 from party A to any politician from party B, so that, in equilibrium, B will lose the election despite its advantage. Specifically, we find in the Appendix that this is the case if and only if $b_2^2 - b_3^2 < 1/72$: this condition is satisfied when $|b_2|$ is not too much farther than $|b_3|$ from zero, the median voter's ideal point.

To see that a leader can lose the election if her ideology moves closer to the median voter, consider politician 2 as leader of party A. Suppose that we start from a situation in which β_1 is smaller than but close to $1/8$. Politician 1 is a trustworthy associate of 2, but if the ideal platform of politician 2 moves closer to the median voter's, and 1's views remains fixed, then the condition $\beta_1 \leq 1/8$ will no longer be satisfied. Thus candidate 2's judgement is no longer better than that of politician 3 and so, were they to contest the election, 2 would lose.

7.4. **Discussion.** Our results link the notion of cohesiveness to a party's electoral success and suggests that anything that increases (decreases) cohesion will have positive (negative) electoral consequences.[23] These theoretical results are consistent with a common understanding that divided parties do not win elections, and with empirical findings. For example, in the three UK elections previous to Labour's election victory in 1997, each of which provided the Conservatives with a majority, Labour were the party perceived to be the most divided. The

[23]Party cohesiveness is studied in McGann (2002) and McGann and Grofman (2002).

British General Election Study of 1983 reveals that 24.2% of the electorate perceived the Conservatives to be divided whereas 87.9% perceived Labour to be so. In 1987 the corresponding figures were 42.1% for the Conservatives and 66.9% for Labour, and in 1992, 27.1% for the Conservatives and 62.8% for Labour. In 1997, however, more of the electorate (40.6%) perceived the Conservatives to be divided than perceived Labour to be so (18.9%). See Heath, Jowell, and Curtice (1983, 1987, 1992, 1997).

Our results on the importance of the cohesiveness of parties relate to debates on party size, and how this affects their political viability and effectiveness, that dates back at least to Michels (1911). Our model highlights that the size of party may be less important than the ideological disparity of views within it. Indeed our results suggest that the formation of a leadership clique or oligarchy that closed down the possibility of fruitful internal debate would be damaging.[24]

Our results on the relevance of party cohesiveness and of the choice of a leader that is not estranged from the party are especially relevant to understand the events that led to the victory of Labour in the 1997 UK elections. That Labour victory is often related to its leader Tony Blair's moderation and his eschewing of the left-wing policies of predecessors. Our result suggests a different narrative, namely that it was the ideological cohesion of "New Labour" rather than the moderation of its leader that was important. Indeed, a corollary to Proposition 6 is that the moderation of its leader is neither a necessary nor sufficient condition for party success. Instead, and in order to assess the chances of electoral victory, one should consider the relationship between the party leader and those of his associates. In his memoirs Blair talks about the team of politicians who advised him and on whom he could rely, amongst them Gordon Brown, John Reid, and David Blunkett. Referring to the latter he states (page 34-35) "his loyalty and commitment to New Labour, I never doubted." Yet whereas Blair himself had always been a moderate and so natural moderniser, Blunkett had moved from the left toward the center. He had been a leader of Sheffield Council, one of Britain's most left leaning councils during the 1980's. His personal ideological change was noted in an article by the Economist in 2001, which described him in the following terms: "a municipal socialist when Thatcherism was rampant, he came to understand the limitations of the old left. This made him a genuine Blairite." Our analysis suggests that the ideological odyssey of Blunkett (amongst others) that allowed him to become a trustworthy associate of Blair, and that might be seen as a small episode in Labour history, should be viewed as a central component of its electoral success.

8. Conclusion and Discussion

Our paper develops a theory of leadership that focuses upon a leader's relations with others in the governance process and, in particular, the importance of her network of close friends and

[24]Either there can no truthful communication between the leader and those outside his clique, and so it would not matter, or (in the case where truthful communication were possible) it would be damaging.

allies. These are people the leader can rely upon to truthfully reveal any private information they hold. A large network of such allies translates into better judgement and better policy. In studying the endogenous formation of such networks we have been able to formulate a theory of leadership choice in which a leader's core characteristic, her judgement or ability to ascertain the best course of action, stems from first principles. Analysis of our model uncovers a set of results that can plausibly explain documented facts, such as the election of extreme leaders and the impact of ideological change within a collective body (such as a party or committee), that are not easily reconciled with previous theories.

We conclude by discussing some useful extensions of our ideas. An interesting question is how our results would change if different politicians had different access to information. It is possible, for example, that ideologically close politicians gather information largely from the same sources. As a result, the advice of associates who are too ideologically close may be less valuable than that provided by more distant ones. Of course, our main result, that associates who are too ideologically distant are not truthful to the leader in equilibrium, would survive in a model in which information is possibly correlated among politicians. As a result, all of our possibility results would extend to this enriched environment. Further, it may prove interesting to determine the optimal composition of advisers to the leader. This is the subject of our current research. We conjecture that the optimal set of advisers would only include politicians whose ideological distance from the leader is neither too large nor too small: ideologically distant advisers would not be consulted as their recommendations would not be credible, and it is possible that politicians who are too close would be excluded so as not to crowd out more valuable less ideologically close advisers.[25]

References

AHLQUIST, J., AND M. LEVI (2011): "Leadership: What it means, what it does, and what we want to know about it," *Annual Review of Political Science*, 14, 1–24.

ANSOLABEHERE, S., AND J. SNYDER (2000): "Valence politics and equilibrium in spatial election models," *Public Choice*, 103, 327–336.

ANSOLABEHERE, S., J. SNYDER, AND C. STEWART (2001): "Candidate Positioning in U.S. House Elections," *American Journal of Political Science*, 45, 136–159.

ARAGONES, E., AND T. PALFREY (2002): "An Economic Model of Representative Democracy," *Journal of Economic Theory*, 103, 131–161.

BENDOR, J., A. GLAZER, AND T. HAMMOND (2001): "Theories of Delegation," *Annual Review of Political Science*, 4, 235–269.

[25]A distinct reason why leaders may choose advisers with diverse policy preferences is that they can decide to pitch advocates with opposite views to argue about the pros and cons of policy choices with uncertain consequences. Policy makers would then rely only on the verifiable information disclosed by advocates, as they cannot trust any unverifiable information conveyed. Selecting advocates with opposite views provides incentives to expend effort in investigating and acquiring information. Instead, our paper focuses on communication of non-verifiable information: the possibility that politicians lie makes transmission reliable only from ideologically close associates.

BENDOR, J., AND A. MEIROWITZ (2004): “Spatial Models of Delegation,” *American Political Science Review*, 98, 293–310.

BERNHARDT, D., O. CÂMARA, AND F. SQUINTANI (2011): “Competence and Ideology,” *Review of Economic Studies*, 78, 487–522.

BESLEY, T., AND M. REYNAL-QUEROL (2011): “Do Democracies Select More Educated Leaders?,” *American Political Science Review*, 105(3), 552–566.

BESLEY, T. J., AND S. COATE (1997): “An Economic Model of Representative Democracy,” *Quarterly Journal of Economics*, 112(1), 85–114.

BLAYDES, L., J. GRIMMER, AND A. MCQUEEN (2013): “Mirrors for Prices and Sultans: Advice on the Art of Governance in the Medieval Christian and Islamic Worlds,” *Working Paper*.

BURDEN, B. C. (2004): “Candidate Positioning in US Congressional Elections,” *British Journal of Political Science*, 34, 211–227.

CALLANDER, S., AND S. WILKIE (2007): “Lies, Damned Lies, and Political Campaigns,” *Games and Economic Behavior*, 60, 262–286.

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

CANES-WRONE, B. (2006): *Who Leads Whom? Presidents, Policy, and the Public*. Chicago University Press, Chicago.

CANES-WRONE, B., M. C. HERRON, AND K. W. SHOTTS (2001): “Leadership and pandering: a theory of executive policymaking,” *American Journal of Political Science*, 45(3), 332–350.

CHARI, V. V., L. E. JONES, AND R. MARIMON (1997): “The Economics of Split-Ticket Voting in Representative Democracies,” *American Economic Review*, 87(5), pp. 957–976.

CLARK, T. (2012): “Bargaining and Opinion Writing in the US Supreme Court,” in *New Directions in Judicial Politics*, ed. by K. McGuire. Routledge.

DEWAN, T., A. GALEOTTI, C. GHIGLINO, AND F. SQUINTANI (2015): “Information Aggregation and Optimal Executive Structure,” *American Journal of Political Science*, 2, 475–494.

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

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

DEWAN, T., AND F. SQUINTANI (2016): “In Defence of Factions,” *American Journal of Political Science*, 60(4), 860–881.

GALASSO, V., AND T. NANNICINI (2011): “Competing on Good Politicians,” *American Political Science Review*, 105(1), 79–99.

GALEOTTI, A., C. GHIGLINO, AND F. SQUINTANI (2013): “Information Transmission Networks,” *Journal of Economic Theory*, 148(5), 136–159.

GANS, J., AND M. SMART (1996): “Majority Voting with Single Crossing Preferences,” *Journal of Public Economics*, 59, 219–237.

GROSECLOSE, T. (2001): “A Model of Candidate Location when One Candidate Has a Valence Advantage,” *American Journal of Political Science*, 45, 862–886.

HARSTAD, B. (2010): “Strategic Delegation and Voting Rules,” *Journal of Public Economics*, 94(12), 102–113.

HEATH, A., R. JOWELL, AND J. CURTICE (1983): “1983 British Election Study Cross-Section,” *University of Essex*.

——— (1987): “1987 British Election Study Cross-Section,” *University of Essex*.

——— (1992): “1992 British Election Study Cross-Section,” *University of Essex*.

——— (1997): “1997 British Election Study Cross-Section,” *University of Essex*.

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

KEARNS GOODWIN, D. (2005): *Team of Rivals: The Political Genius of Abraham Lincoln*. Simon and Schuster.

KREHBIEL, K. (2007): “Supreme Court Appointments as a Move the Median Game,” *American Journal of Political Science*, 51(2), 231–240.

MCCURLEY, C., AND J. J. MONDAK (1995): “The Influence of Incumbents’ Competence and Integrity in U.S. House Elections,” *American Journal of Political Science*, 39, 864–885.

MCGANN, A. J. (2002): “Theories of Delegation,” *Journal of Theoretical Politics*, 14, 37–70.

MCGANN, A. J., W. K., AND B. GROFMAN (2002): “How an ideologically concentrated minority can trump a dispersed majority: Nonmedian voter results for plurality, run-off, and sequential elimination elections,” *American Journal of Political Science*, 46, 134–147.

OSBORNE, M. J., AND A. SLIVINSKI (1996): “A Model of Political Competition with Citizen-Candidates,” *Quarterly Journal of Economics*, 111, 65–96.

PATTY, J. W. (2013): “A Theory of Cabinet-Making: The Politics of Inclusion, Exclusion, and Information,” *Washington University at St. Louis*.

PATTY, J. W., AND E. M. PENN (2014): “Sequential Decision Making and Information Aggregation in Small Networks,” *Political Science Research and Methods*, 2, 1–29.

PENN, E. M. (2016): “Engagement, Disengagement, or Exit: A Theory of Equilibrium Associations,” *American Journal of Political Science*, 60, 322–336.

ROHDE, D. W., AND K. SHEPSLE (2007): “Advising and Consenting in the 60 Vote Senate: Strategic Appoinments to the Supreme Court,” *Journal of Politics*, 69, 664–677.

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

STOKES, D. (1963): “Spatial Models of Party Competition,” *American Political Science Review*, 57, 368–372.

STONE, W., AND E. SIMAS (2010): “Candidate Valence and Ideological Positions in U.S. House Elections,” *American Journal of Political Science*, 54, 371–288.

UHL-BIEN, M. (2006): “Relational Leadership Theory: Exploring the social processes of leadership and organizing,” *The Leadership Quarterly*, 17, 654–676.

ONLINE APPENDIX NOT SUBMITTED FOR PUBLICATION

Equilibrium beliefs. In our model a politicians' equilibrium updating is based on the standard Beta-binomial model. Suppose that the leader j holds $d+1$ bits of information, i.e. she holds the private signal s_j and d politicians truthfully reveal their signals to her. The probability that l out of such $d+1$ signals equal one, conditional on θ is

$$f\left(l|\theta, d+1\right) = \frac{(d+1)!}{l!\,(d+1-l)!}\theta^l \left(1-\theta\right)^{(d+1-l)}.$$

Hence, politician j's posterior on θ is

$$f\left(\theta|l, d+1\right) = \frac{(d+2)!}{l!\,(d+1-l)!}\theta^l \left(1-\theta\right)^{(d+1-l)},$$

the expected value of θ is

$$E\left(\theta|l, d+1\right) = \frac{l+1}{d+3},$$

and the variance is

$$V(\theta|l, d+1) = \frac{(l+1)\,(d+2-l)}{(d+3)^2\,(d+4)}.$$

■

Derivation of Expression (2). Fix a leader j, consider a communication strategy profile $\mathbf{m}_{-j}$ and suppose that it is an equilibrium together with the strategy y_j in expression (1). Let $C_j(\mathbf{m}_{-j})$ be the set of players truthfully communicating with the leader j in the equilibrium. The equilibrium information of j is thus $d_j(\mathbf{m}_{-j})+1 = |C_j(\mathbf{m}_{-j})|+1$, the cardinality of $C_j(\mathbf{m}_{-j})$ plus j's signal s_j. Consider any player $i \in C_j(\mathbf{m}_{-j})$. Let $\mathbf{s}_{-i}(\mathbf{m}_{-j})$ be any vector containing s_j and the (truthful) messages of all players in $C_j(\mathbf{m}_{-j})$ except i. Let also $y_j(s_i, \mathbf{s}_{-i}(\mathbf{m}_{-\mathbf{j}}))$ be the action that j takes if she has information $\mathbf{s}_{-i}(\mathbf{m}_{-j})$ and believes in the signal s_i sent from player i, analogously, $y_j(1-s_i, \mathbf{s}_{-i}(\mathbf{m}_{-j}))$ is the action that j takes if she has information $\mathbf{s}_{-i}$ and believes in the signal $1-s_i$ sent from player i. Simplifying notation, player i does not deviate from reporting truthfully signal s_i to the leader j if and only if

$$-\int_0^1 \sum_{\mathbf{s}_{-i}\in\{0,1\}^{d_j}} \left[(y_j(s_i, \mathbf{s}_{-i}) - \theta - b_i)^2 - (y_j(1-s_i, \mathbf{s}_{-i}) - \theta - b_i)^2\right] f(\theta, \mathbf{s}_{-i}|s_i)d\theta \geq 0.$$

Simplifying, we obtain:

$$-\int_0^1 \sum_{\mathbf{s}_{-i}\in\{0,1\}^{d_j}} (y_j(s_i, \mathbf{s}_{-i}) - y_j(1-s_i, \mathbf{s}_{-i})) \left[\frac{y_j(s_i, \mathbf{s}_{-i}) + y_j(1-s_i, \mathbf{s}_{-i})}{2} - (\theta + b_i)\right] f(\theta, \mathbf{s}_{-i}|s_i)d\theta \geq 0.$$

Next, observing that

$$y_j(s_i, \mathbf{s}_{-i}) = b_j + E\left[\theta|s_i, \mathbf{s}_{-i}\right],$$

we obtain

$$-\int_0^1 \sum_{\mathbf{s}_{-i}\in\{0,1\}^{d_j}} \left(E\left[\theta|s_i,\mathbf{s}_{-i}\right]-E\left[\theta|1-s_i,\mathbf{s}_{-i}\right]\right)\cdot$$

$$\cdot \quad \left[b_j+\frac{E\left[\theta|s_i,\mathbf{s}_{-i}\right]+E\left[\theta|1-s_i,\mathbf{s}_{-i}\right]}{2}-\theta-b_i)\right] f(\theta,\mathbf{s}_{-i}|s_i)d\theta \geq 0.$$

Denoting

$$\Delta\left(s_i,\mathbf{s}_{-i}\right)=E\left[\theta|s_i,\mathbf{s}_{-i}\right]-E\left[\theta|1-s_i,\mathbf{s}_{-i}\right],$$

observing that:

$$f(\theta,\mathbf{s}_{-i}|s_i)=f(\theta|\mathbf{s}_{-i},s_i)\Pr(\mathbf{s}_{-i}|s_i),$$

and simplifying, we get:

$$-\sum_{\mathbf{s}_{-i}\in\{0,1\}^{d_j}} \int_0^1 \Delta\left(s_i,\mathbf{s}_{-i}\right)\left(\frac{E\left[\theta|s_i,\mathbf{s}_{-i}\right]+E\left[\theta|1-s_i,\mathbf{s}_{-i}\right]}{2}+b_j-b_i-\theta\right) f(\theta|\mathbf{s}_{-i},s_i)\Pr(\mathbf{s}_{-i}|s_i)\geq 0.$$

Furthermore, using

$$\int_0^1 \theta f(\theta|\mathbf{s}_{-i},s_i)d\theta = E\left[\theta|s_i,\mathbf{s}_{-i}\right],$$

we obtain:

$$-\sum_{\mathbf{s}_{-i}\in\{0,1\}^{d_j}} \int_0^1 \Delta\left(s_i,\mathbf{s}_{-i}\right)\left(\frac{E\left[\theta|s_i,\mathbf{s}_{-i}\right]+E\left[\theta|1-s_i,\mathbf{s}_{-i}\right]}{2}+b_j-b_i-E\left[\theta|s_i,\mathbf{s}_{-i}\right]\right) f(\theta|\mathbf{s}_{-i},s_i)\Pr(\mathbf{s}_{-i}|s_i)$$

$$=-\sum_{\mathbf{s}_{-i}\in\{0,1\}^{d_j}} \int_0^1 \Delta\left(s_i,\mathbf{s}_{-i}\right)\left(-\frac{\Delta\left(s_i,\mathbf{s}_{-i}\right)}{2}+b_j-b_i\right) f(\theta|\mathbf{s}_{-i},s_i)\Pr(\mathbf{s}_{-i}|s_i)\geq 0.$$

Now, note that, for any number $l=0,...,d_j$ of digits equal to one in $\mathbf{s}_{-i}(\mathbf{m}_{-j})$,

$$\begin{aligned}
\Delta\left(s_i,\mathbf{s}_{-i}\right) &= E\left[\theta|s_i,\mathbf{s}_{-i}(\mathbf{m}_{-j})\right]-E\left[\theta|1-s_i,\mathbf{s}_{-i}(\mathbf{m})\right]\\
&= E\left[\theta|l+s_i,d_j(\mathbf{m}_{-j})+1\right]-E\left[\theta|l+1-s_i,d_j(\mathbf{m}_{-j})+1\right]\\
&= \left(l+1+s_i\right)/\left(d_j(\mathbf{m}_{-j})+3\right)-\left(l+2-s_i\right)/\left(d_j(\mathbf{m}_{-j})+3\right)\\
&= \begin{cases} -1/\left(d_j(\mathbf{m}_{-j})+3\right) \text{ if } s_i=0\\ 1/\left(d_j(\mathbf{m}_{-j})+3\right) \text{ if } s_i=1.\end{cases}
\end{aligned}$$

We obtain that player i communicates truthfully the signal $s_i=0$ to player j if and only if:

$$-\left(\frac{-1}{d_j(\mathbf{m}_{-j})+3}\right)\left(-\frac{-1}{2(d_j(\mathbf{m}_{-j})+3)}+b_j-b_i\right)\geq 0,$$

or

$$b_i-b_j\leq\frac{1}{2\left(d_j(\mathbf{m})+3\right)},$$

and note that this condition is redundant if $b_i-b_j<0$.

Likewise, i communicates truthfully the signal $s_i=1$ to player j if and only if:

$$-\left(\frac{1}{d_j(\mathbf{m}_{-j})+3}\right)\left(-\frac{1}{2(d_j(\mathbf{m}_{-j})+3)}+b_j-b_i\right)\geq 0,$$

or

$$b_i - b_j \geq -\frac{1}{2\left(d_j(\mathbf{m}_{-j})+3\right)},$$

and note that this condition is redundant if $b_i - b_j > 0$.

Collecting the two conditions yields expression (2). ■

Derivation of equilibrium welfare, expression (4). We consider any equilibrium $(\mathbf{m}_{-j}, y_j)$. The ex-ante expected utility of each player i is:

$$\begin{aligned} Eu_i(\mathbf{m}_{-j}, y_j) &= -E\left[(y_j(s_j, \hat{\mathbf{m}}_{-j}) - \theta - b_i)^2\right] \\ &= -E\left[(b_j + E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right] - \theta - b_i)^2\right]. \end{aligned}$$

Hence

$$\begin{aligned} Eu_i(\mathbf{m}_{-j}, y_j) &= -E\left[(b_j - b_i)^2 + \left(E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right] - \theta\right)^2 - 2(b_j - b_i)\left(E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right] - \theta\right)\right] \\ &= -\left[(b_j - b_i)^2 + E\left[\left(E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right] - \theta\right)^2\right] - 2(b_j - b_i)\left(E[E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right]] - E[\theta]\right)\right], \end{aligned}$$

by the law of iterated expectations, $E[E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right]] = E[\theta]$, so we obtain

$$Eu_i(\mathbf{m}_{-j}, y_j) = -(b_j - b_i)^2 - E\left[\left(E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right] - \theta\right)^2\right].$$

Letting l be the number of digits equal to one in the $(d_j(\mathbf{m_{-j}})+1)$-digit leader's information vector $(s_j, \hat{\mathbf{m}}_{-j})$,

$$\begin{aligned} E\left[\left(E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right] - \theta\right)^2\right] &= \int_0^1 \sum_{l=0}^{d_j(\mathbf{m_{-j}})+1} \left(E\left[\theta | l, d_j(\mathbf{m}_{-j})+1\right] - \theta\right)^2 f(l | d_j(\mathbf{m}_{-j})+1, \theta) d\theta \\ &= \int_0^1 \sum_{l=0}^{d_j(\mathbf{m_{-j}})+1} \left(E\left[\theta | l, d_j(\mathbf{m}_{-j})+1\right] - \theta\right)^2 \frac{f\left(\theta | l, d_j(\mathbf{m}_{-j})+1\right)}{d_j(\mathbf{m}_{-j})+2} d\theta, \end{aligned}$$

where the second equality follows from $f(l|d_j(\mathbf{m}_{-j})+1, \theta) = f(\theta | l, d_j(\mathbf{m}_{-j})+1)/(d_j(\mathbf{m}_{-j})+2)$.

Because the variance of a beta distribution of parameters l and $d+1$, is

$$V(\theta | l, d+1) = \frac{(l+1)(d+2-l)}{(d+3)^2(d+4)},$$

we obtain:

$$\begin{aligned} E\left[\left(E\left[\theta | s_j, \hat{\mathbf{m}}_{-j}\right] - \theta\right)^2\right] &= \frac{1}{d_j(\mathbf{m}_{-j})+2}\left[\sum_{l=0}^{d_j(\mathbf{m}_{-j})+1} V\left(\theta | l, d_j(\mathbf{m_{-j}})+1\right)\right] \\ &= \sum_{l=0}^{d_j(\mathbf{m}_{-j})+1} \frac{(l+1)\left(d_j(\mathbf{m}_{-j})+2-l\right)}{\left(d_j(\mathbf{m}_{-j})+2\right)\left(d_j(\mathbf{m_{-j}})+3\right)^2\left(d_j(\mathbf{m}_{-j})+4\right)} \\ &= \frac{1}{6(d_j(\mathbf{m}_{-j})+3)}. \end{aligned}$$

■

Proof of Lemma 1. We note that

$$\begin{aligned} U_i^*(j) &= -(b_i-b_j)^2 - \left[6\left(d_j^*+3\right)\right]^{-1} = -(b_i-b_{i'}+b_{i'}-b_j)^2 - \left[6\left(d_j^*+3\right)\right]^{-1} \\ &= -(b_i-b_{i'})^2 - (b_{i'}-b_j)^2 - 2(b_i-b_{i'})(b_{i'}-b_j) - \left[6\left(d_j^*+3\right)\right]^{-1} \\ &= -(b_i-b_{i'})((b_i-b_{i'})+2(b_{i'}-b_j)) + U_{i'}^*(j) \\ &= -(b_i-b_{i'})(b_i+b_{i'}-2b_j) + U_{i'}^*(j) \end{aligned}$$

and

$$U_i^*(j') = -(b_i-b_{i'})\left(b_i+b_{i'}-2b_{j'}\right) + U_{i'}^*(j').$$

If $i<i'$, $j<j'$ and $U_{i'}^*(j) > U_{i'}^*(j)$, then $U_i^*(j) > U_i^*(j')$ is implied by

$$-(b_i-b_{i'})(b_i+b_{i'}-2b_j) \geq -(b_i-b_{i'})\left(b_i+b_{i'}-2b_{j'}\right)$$

or, because $i<i'$, by

$$b_i+b_{i'}-2b_j \geq b_i+b_{i'}-2b_{j'}$$

which is implied by $j<j'$. ■

Proof of Proposition 3. Suppose that there is a constant $\beta>0$ such that $b_{i+1}-b_i=\beta$ for all $i=1,...,n-1$. Then, for any real number $b>0$, the size of ideological neighborhood $N_j(b)$ is constant in j for all players j such that the number of politicians $i<j$ who belong to $N_j(b)$ is the same as the number of politicians $i>j$ who belong to $N_j(b)$. Formally, letting $\bar{\imath}_j(b)=\max\{i\in N : |b_i-b_j|\leq b\}$ and $\underline{i}_j(b)=\min\{i\in N : |b_i-b_j|\leq b\}$, we have that $N_j(b)=2\lfloor b/\beta\rfloor+1$, for any j such that $\bar{\imath}_j(b)-j=j-\underline{i}_j(b)$, where the notation $\lfloor b/\beta\rfloor$ denotes the largest integer smaller than b/β.

The remaining players j are constrained by the boundaries of the ideology spectrum b_1 and b_n in the size of their ideological neighborhood $N_j(b)$, so that it is either the case that $\bar{\imath}_j=n$, in which case $N_j(b)=\lfloor b/\beta\rfloor+1+\bar{\imath}_j(b)-j$, or that $\underline{i}_j=1$, in which case $N_j(b)=\lfloor b/\beta\rfloor+1+j-\underline{i}_j(b)$; and in both cases $N_j(b)<2\lfloor b/\beta\rfloor+1$.

Because $m=(n+1)/2$, by construction $N_m(b)=2\lfloor b/\beta\rfloor+1$ for all values of b, and hence $N_m(b)\geq N_j(b)$ for all other politician j and values of b. We note that $N_j(b)$ weakly increases in b and $\frac{1}{2(d+3)}$ decreases in d, and hence d_j^* is maximal for the index(es) j that maximize the function $N_j(\cdot)$. That is to say, when there is a constant $\beta>0$ such that $b_{i+1}-b_i=\beta$ for all $i=1,...,n-1$, the median politician m weakly dominates all other politicians in terms of judgement, and should always be selected as group leader. ■

Analysis of the 5 Player Case in Section 6, Proof of Lemma 2 and of Proposition 4. We calculate all the parameter regions in which $d_2^*>d_3^*$. We first note that $d_3^*=0$ if $\beta_2>1/8$ and $\beta_3>1/8$; so that $d_2^*\leq 1$ as 3 will never be truthful to 2, and $d_2^*=1$ if $\beta_1\leq 1/8$. We then see that $d_3^*=1$ if $\beta_2\leq 1/8$ and $\beta_3>1/10$; so that $d_2^*\leq 2$ as 4 will never be truthful to 2, and $d_2^*=2$ if $\beta_1\leq 1/10$ and $\beta_2\leq 1/10$. Also, we see that $d_3^*=1$ if $\beta_2>1/10$ and $\beta_3\leq 1/8$; so that $d_2^*\leq 1$ as

3 will never be truthful to 2. Then, we note that $d_3^* = 2$ if $\beta_2 \leq 1/10$, $\beta_3 \leq 1/10$, $\beta_1 + \beta > 1/12$ and $\beta_3 + \beta_4 > 1/12$; so that $d_2^* \leq 3$ as 5 will never be truthful to 2, and $d_2^* = 3$ if $\beta_2 + \beta_3 \leq 1/12$ and $\beta_1 \leq 1/12$. Further, we note that $d_3^* = 3$ if $\beta_1 + \beta_2 \leq 1/12$, $\beta_3 \leq 1/12$ and $\beta_3 + \beta_4 > 1/14$; so that $d_2^* \leq 3$ as 5 will never be truthful to 2. Finally we see that $d_3^* = 3$ if $\beta_1 + \beta_2 > 1/14$, $\beta_2 \leq 1/12$ and $\beta_3 + \beta_4 \leq 1/12$; so that $d_2^* \leq 4$, and $d_2^* = 4$ if $\beta_2 + \beta_3 + \beta_4 \leq 1/16$ and $\beta_1 \leq 1/16$.

We consider the case in which $W^*(2) > W^*(4), U_3^*(2) > U_3^*(4)$, $\beta_1 \leq 1/10, \beta_2 \leq 1/10, \beta_3 > 1/10$ and hence $\delta = \beta_4 - \beta_1 + 2\beta_3 > 1/10$, $d_2^* = 2, d_1^* = 1$. Using expression (4), we can calculate the aggregate expected payoffs for selecting either politician 2 or 3 as the leader:

$$\begin{aligned} W^*(2) &= -\beta_1^2 - \beta_2^2 - (\beta_2 + \beta_3)^2 - (\beta_2 + \beta_3 + \beta_4)^2 - 5\frac{1}{6(2+3)}, \\ W^*(3) &= -(\beta_1 + \beta_2)^2 - \beta_2^2 - \beta_3^2 - (\beta_3 + \beta_4)^2 - 5\frac{1}{6(1+3)}. \end{aligned}$$

The centre-left politician 3 is optimally selected as the leader whenever

$$W^*(2) - W^*(3) = -2\delta\beta_{23} - \beta_2^2 + \frac{1}{24} > 0 \text{ or } \beta_2 < \tau(\delta) \equiv \sqrt{\delta^2 + 1/24} - \delta$$

It is easy to verify that the threshold $\tau(\delta)$ is strictly decreasing in δ, with $\tau(1/10) \approx 0.1273 > 1/10$, that $\tau(\delta)$ is strictly positive for any δ and equals zero only in the limit as δ approaches infinity.

In sum, we conclude that, whenever β_2 is sufficiently small — i.e., smaller than $1/10$ and than $\tau(\delta)$, $\beta_1 \leq 1/10$ and $\beta_3 > 1/10$, then the centre-left politician 2 should be optimally selected as the leader in lieu of the most moderate candidate, politician 3. This is because 2 has better judgement, as it can count on two trustworthy associates, whereas 3 has only one; and 2 is not too much more extremist than 3, as β_2 is small.

Turning to studying the election of the leader by majority vote, we first calculate player 3's payoffs for selecting politician 2 or 3 as the leader, using expression (5):

$$U_3^*(2) = -\beta_2^2 - \frac{1}{6(1+3)} \text{ and } U_3^*(3) = -\frac{1}{6(3)},$$

the median politician 3 will delegate leadership to player 2 whenever

$$U_3^*(2) - U_3^*(3) = \frac{1 - 120\beta_{23}^2}{120} > 0 \text{ or } \beta_2 < \frac{1}{2\sqrt{30}} \approx 0.0913.$$

In light of Proposition 2, we obtain that, whenever β_2 is smaller than $\frac{1}{2\sqrt{30}}$, $\beta_1 \leq 1/10$ and $\beta_3 > 1/10$, the politician 2 is the Condorcet winner of the election game. Again, this is because 2 can count on two ideologically close trustworthy associates, whereas 3 has only one, and because 2 does not hold views too different from the ones of 3.

It is interesting to compare this situation with the equidistant case in which $b_{i+1} - b_i$ is constant for all $i = 1, ..., 4$ and smaller than $\frac{1}{2\sqrt{30}}$. Suppose that the centre-right politician 4 extremizes her ideology b_4 away from the median b_3, so as to increase β_3 beyond $1/10$. Paradoxically, by doing so, she will make the elected leader's ideology move in the opposite direction, as the centre-left politician 2 will gain better judgement than the median politician 3, and win the election. Equivalently, suppose that, initially $b_{i+1} - b_i = \beta > 1/10$ for all $i = 1, ..., 4$. If the leftist politicians 1 and 2 moderate their views, so that β_2 becomes smaller than $\frac{1}{2\sqrt{30}}$ and β_1 becomes smaller than $1/10$, then they move the elected leader's decision towards their views, by making the centre-left politician 2 the leader, in lieu of the median politician 3.

We now compare election and selection of the leader. Because $\tau(\delta)$ is strictly decreasing in δ, $\tau(1/10) > 1/10$ and $\tau(\delta) \to 0$ as $\delta \to \infty$, it is immediate to see that there is a unique threshold $\bar{\delta} > 1/10$ such that $\tau(\delta) > \frac{1}{2\sqrt{30}}$ for all $\delta < \bar{\delta}$ and $\tau(\delta) < \frac{1}{2\sqrt{30}}$ for all $\delta > \bar{\delta}$. This implies that, whenever $\delta < \bar{\delta}$, there exists an interval $\left(1/[2\sqrt{30}], 1/10\right)$ of the parameter β_2 such that the centre-left politician 2 should be optimally selected as leader but the median politician 3 is the Condorcet winner of the election game. A surprising result occurs when $\delta > \bar{\delta}$, so that $b_5 - b_3$ and $b_4 - b_3$ are sufficiently large relative to $b_2 - b_1$. For values of β_2 larger than $\tau(\delta)$ but smaller than $\frac{1}{2\sqrt{30}}$, the Condorcet winner is the centre-left politician 2 despite the fact that optimal leader is the median politician 3. In the election game, the median politician 3 delegates leadership to a less moderate politician, 2, despite the fact that it would be optimal for the group if she retained leadership for herself. ■

Analysis of the 6 Player Example in Section 7, Proof of Proposition 5. Suppose that there are 6 politicians, with ideologies such that $b_{i+1} - b_i = \beta$ for all $i = 1, ..., 5$, arranged symmetrically around the median ideology zero, so that $b_3 = -\beta/2$ and $b_4 = \beta/2$. Politicians 1, 2, 3 belong to party A, and politicians $4, 5, 6$ to party B. Unless politicians 2 and 5 can count of more trustworthy advisers than 3 and 4, the latter will be selected by their parties and tie the general election, in equilibrium. Because of symmetry of $\mathbf{b}$, let us now just focus on the selection of party A candidates. Candidate 1 will never be selected, so we consider 2 and 3. Because 3 can rely on 2, if 3 communicates to 2 in equilibrium, it follows that the only case in which 2 has better judgement than 3 is when $d_2^* = 2$ and $d_3^* = 1$, which requires that $\beta \leq 1/10$ and that $2\beta > 1/10$.

Because of symmetry of $\mathbf{b}$, if $U_0(2) > U_0(3)$, then there cannot be an equilibrium in which party A selects 3 as its candidate in the general election; if they did, in fact, party B would select 5 as candidate and win the election. When $U_0(2) > U_0(3)$, the unique equilibrium of the game has candidates 2 and 5 tie the general election. Simplifying this condition, we obtain:

$$U_0(2) - U_0(3) = -(\beta + \beta/2)^2 - \frac{1}{6(2+3)} - \left[-(\beta/2)^2 - \frac{1}{6(1+3)}\right] = \frac{1}{120}\left(1 - 240\beta^2\right) > 0.$$

Because the last inequality holds if and only if $\beta < \frac{1}{4\sqrt{15}}$, we conclude that when $1/20 < \beta < \frac{1}{4\sqrt{15}}$, the winners of the general election are not the most moderate politicians 3 and 4, despite the fact that the politicians' ideologies are evenly distributed on the line. ■

Analysis of the 5 Player Example in Section 7 and Proof of Proposition 6 Suppose that there are 5 politicians, with $b_2 < 0 < b_3$. Politicians $1, 2$ belong to party A and $3, 4, 5$

belong to party B, and we assume that $b_3 < -b_2$. Party B has more informed politicians, and it can also select a candidate, player 3, whose views are closer to the median voter. If there were no communication to the winner of the general election, party B would always win by selecting politician 3. However, politician 2 wins the general election if she has better judgement than player 3. As there is only another informed politician in party A, this may only happen if $d_2^* = 1 > d_3^* = 0$, and this requires $\beta_1 \leq 1/8$, $\beta_3 > 1/8$ and $\beta_4 > 1/8$. Party A is more ideologically cohesive, and can express a candidate, 2, with a larger network of trustworthy associates than any candidate available to party B. The median voter turns out to prefer to elect politician 2 than politician 3 whenever

$$U_0(2) - U_0(3) = -b_2^2 - \frac{1}{6(1+3)} - \left[-b_3^2 - \frac{1}{6(3)}\right] = \frac{1}{72} - (b_2^2 - b_3^2) > 0,$$

i.e., $b_2^2 - b_3^2 < 1/72$.

To prove the claim that candidate 2 can lose the election by moving closer to the median voter, suppose that we start from an ideology profile $\mathbf{b}$ such that β_1 is smaller than but close to $1/8$. If politician 2 moves ideologically closer to the median voter (i.e., $-b_2$ decreases), then the condition $\beta_1 \leq 1/8$ will not be satisfied anymore, candidate 2 will lose the truthful advice of party fellow 1, in turn losing the informational advantage over 3, and the general election. ■

Shared leadership. Consider a group of politicians $i = 1, ..., n$. Suppose that, instead of electing a single leader j, it is possible to select a vector α of shares of leadership α_j for $j = 1, ..., n$ such that $\alpha_j \geq 0$ for all j, and $\sum_{j=1}^n \alpha_j = 1$. For every vector α, its support $L_\alpha \equiv \{j : \alpha_j > 0\}$ denotes the associated set of leaders. The communication by each player i to the leaders L_α may be *private* (hence, the message $\hat{m}_{ij} \in \{0, 1\}$ sent by i to j may differ across $j \in L_\alpha$), or *public* (and then $\hat{m}_{ij}$ must be the same for all $j \in L_\alpha$).

A vector of authority shares α determines the mixture over outcomes:

$$y(\mathbf{s}, \mathbf{m}; \alpha) = \sum_{j=1}^{n} \alpha_j \left[b_j + E\left[\theta | s_j, \mathbf{m}_{-j}\right]\right],$$

given the signals $\mathbf{s} = (s_j)_{j=1}^n$ and the equilibrium communication strategies $\mathbf{m} = (\mathbf{m}_{-j})_{j=1}^n$. And this yields each player i expected utility:

$$U_i(\mathbf{s}, \mathbf{m}; \alpha) = -\sum_{j=1}^{n} \alpha_j (b_i - b_j)^2 - \sum_{j=1}^{n} \alpha_j \frac{1}{6\left(d_j^*(\mathbf{m}_{-j}) + 3\right)}.$$

In terms of optimal choice, the possibility of choosing α optimally improves utilitarian welfare over single leadership weakly by definition, in our model. It is easy to find examples where it improves utilitarian welfare strictly—see Example 1 in Dewan, Galeotti, Ghiglino, and Squintani (2015), for instance.

Let us consider now the majority choice among share of leadership vectors α. The space of vectors α can be linearly ordered according to the mixture over biases $\bar{b}(\alpha) = \sum_{j=1}^n \alpha_j b_j$. It is immediate to then extend the proof of Lemma 1 to this environment. As a consequence, the set of Condorcet winning share of leadership vectors α coincides with the set of vectors α that maximize the expected payoff of the median player m.

The same kinds of inefficiency described in Lemma 2 and Proposition 4 extends to this richer environment. As we now demonstrate, there are examples, parametrized by the bias vector b, in which the optimal share of leadership vector α differs from the majority choice.

We consider the 5-player case studied in Section 6.3, and so assume $\beta_1 \leq 1/10$, $\beta_2 \leq 1/10$, $\beta_3 > 1/10$, and hence $\delta > 1/10$. Suppose that $\tau(\delta) < \frac{1}{2\sqrt{30}}$, and that $\tau(\delta) < \beta_2 < \frac{1}{2\sqrt{30}}$. As shown in Lemma 2, the optimal leader is 3, but 3 delegates to 2 who is better informed, because $d_2^* = 2$ and $d_3^* = 1$. Allowing for shared leadership, it would be possible to get 4 to communicate truthfully to 3 only if including 5 in the set of leaders L_α, and considering public communication. With private communication, 4 would not be truthful to 3 in equilibrium, as it would wish to distort the decision y_3 regardless of the message $\hat{m}_{45}$ she sends to player 5. Using Lemma 1 of Dewan, Galeotti, Ghiglino, and Squintani (2015), there is an equilibrium in which players 2 and 4 are truthful to 3 and 5 if and only if:

$$|b_4 - (\gamma_3 b_3 + \gamma_5 b_5)| \leq \gamma_3 \frac{1}{2(d_3+2)} + \gamma_5 \frac{1}{2(d_5+2)} \tag{6}$$

$$|b_2 - (\gamma_3 b_3 + \gamma_5 b_5)| \leq \gamma_3 \frac{1}{2(d_3+2)} + \gamma_5 \frac{1}{2(d_5+2)} \tag{7}$$

where $\gamma_3 = \frac{\alpha_3/2(d_3+2)}{\alpha_3/2(d_3+2)+\alpha_5/2(d_5+2)}$ and $\gamma_5 = \frac{\alpha_5/2(d_5+2)}{\alpha_3/2(d_3+2)+\alpha_5/2(d_5+2)}$, and $\alpha_3 + \alpha_5 = 1$.

Here, because $b_5 - b_3 > b_4 - b_3 > 1/10$, player 3 and 5 cannot be truthful to each other, hence $d_3 = 2$ and $d_5 = 2$. Conditions (6) and (7) become:

$$|b_4 - (\alpha_3 b_3 + (1-\alpha_3)b_5)| = \alpha_3\beta_3 - (1-\alpha_3)\beta_4 \leq \frac{1}{10}$$

$$|b_2 - (\alpha_3 b_3 + (1-\alpha_3)b_5)| = \alpha_3\beta_2 + (1-\alpha_3)(\beta_2+\beta_3+\beta_4) \leq \frac{1}{10}.$$

Condition (6) is satisfied tightly for $\alpha_3 = \frac{\beta_4+1/10}{\beta_3+\beta_4}$, plugging this into condition (7), we obtain:

$$\frac{\beta_4+1/10}{\beta_3+\beta_4}(\beta_2 - 1/10) + \left(1 - \frac{\beta_4+1/10}{\beta_3+\beta_4}\right)(\beta_2+\beta_3+\beta_4-1/10) = \beta_2+\beta_3-1/5 \leq 0.$$

that is violated for $\beta_3 > 1/5 - \beta_2$, i.e., $\beta_3 > 1/10$, because $\beta_2 \leq 1/10$. We conclude that, for $0 < \beta_2 \leq 1/10$, $0 < \beta_4 \leq 1/10$ and $\beta_3 > 1/10$, it is not possible to get 2 and 4 to communicate truthfully to 3 in equilibrium with any shared leadership vector α. In other terms, $d_3^* \leq 1$ in equilibrium.

Suppose further that $\tau(\delta) < \frac{1}{2\sqrt{30}}$, noting that $\tau(\delta) = \sqrt{\delta^2 + 1/24} - \delta$ decreases in $\delta = \beta_4 + 2\beta_3 - \beta_1$, so that the condition $\tau(\delta) < \frac{1}{2\sqrt{30}}$ is satisfied for $\delta > \tau^{-1}\left(\frac{1}{2\sqrt{30}}\right) = 1/\sqrt{30} \approx 0.182\,57$, and does impose any upper bound on β_3. Consider any β_2 such that $\tau(\delta) < \beta_2 < \frac{1}{2\sqrt{30}}$, and note that $0 < \tau(\delta) < \beta_2 < \frac{1}{2\sqrt{30}} < 1/10$. The proof of Lemma 2 implies that, because $d_2^* = 2$ and $d_3^* = 1$, the optimal leader is 3, but 3 prefers to delegates to 2 who is better informed, and hence 2 is elected by majority voting. The same kind of inefficiency described in Lemma 2 and Proposition 4 extends to the environment that includes the possibility of shared leadership.

We conclude by noting that a different way to define shared leadership would be to fix a system α of sharing rules α_L for all possible sets of leaders $L \subseteq \{1, ..., n\}$, and restrict the optimal and majority choice only to the set of leaders L given the system α. For example, α could be an "egalitarian system" such that $\alpha_{jL} = 1/|L|$ for all sets of leaders L, and all $j \in L$. Regardless of the selected/elected set of leaders L, each leader $j \in L$ has equal share of power. Alternatively, the system α could include forms of seniority among politicians.

It is obvious that fixing the system α and selecting L optimally is a weak improvement upon optimal individual leadership, and that it is weakly dominated by optimal selection of a vector of shares α. Further, the extended example above demonstrates that the kinds of inefficiency described in Lemma 2 occur also in this environment. There are examples, parametrized by the bias vector b and leadership sharing system α, in which the optimal choice of L given α differs from the majority choice. ■