On Leaders' Judgement and Trustworthy Associates

Torun Dewan[1] **Francesco Squintani**[2]

January, 2015.

Preliminary and Incomplete.

Abstract. Members of a group value informed decisions and hold ideological preferences. A leader is granted authority to take a decision on their behalf. Good leadership depends on characteristics of moderation and judgement. The latter emerges endogenously via the advice that is communicated by members to the leader. Trustworthy advice requires ideological proximity to the leader. An implication is that the group may chose a relatively extreme leader with a large number of trustworthy allies. Paradoxically, this may happen even when it is in the group's collective interest to choose a more moderate one. We develop our analysis further in the context of two-party competition where each party chooses a leader who implements her preferred policy if elected. Then a party may choose an extreme leader who defeats a moderate one chosen by the opposing party. Our results highlight the importance of party cohesion and the relations between a leader and her party. We show that these can be more important to securing electoral victory than proximity of a leader's position to the median voter.

E la prima coniettura che si fa del cervello d'uno signore, é vedere li uomini che lui ha d'intorno. *[Transl.: The first opinion which one forms of a prince, and of her understanding, is by observing the men she has around her.]* Niccolo' Machiavelli, Il Principe, Ch. 22.

[1]London School of Economics and Political Science. email: t.dewan@lse.ac.uk
[2]Warwick University. email: F.Squintani@Warwick.ac.uk

1. INTRODUCTION

What makes a good leader? A simple answer imputes good leadership to individual leader's qualities, such as wisdom, judgement, or the ability to communicate with others, and presumes that this qualities are partly innate, and partly the outcome of a leader's experience and training.[3] A famous point raised by Niccolo' Machiavelli in Chapter 22 of "The Prince", however, leads us to investigate more deeply how good leadership may arise. Machiavelli argued that the first opinion that one forms of a prince, and of his understanding, is by observing the men he has around him. Further, he describes the greatness of Pandolfo Petrucci, prince of Siena and his valent minister Antonio da Venafro. He highlights that Pandolfo's ability as a ruler depended upon the information and good judgement provided by Antonio. The lesson we take from this suggestion is that what makes a good leader are not only her individual qualities. A leader's good judgement is also a consequence of the valuable advice of her network of *trustworthy associates*, i.e., of the individuals that can be relied upon to truthfully reveal valuable information to the leader.[4] Hence, good leadership may emerge because of determinants external to the leader, through her interaction with other actors in the governance process, and because of these actors' individual characteristics.

To explore this novel notion of leadership we develop a simple formal model. A group of politicians collectively selects a leader who is thereby granted authority to take a decision on their behalf. The politicians payoffs depend on an uncertain state of the world about which each is independently, privately and imperfectly informed. Prior to the decision being taken, politicians can communicate their information to each other. They value informed decisions and so, all other things equal, they would prefer the leader, whoever she may be, to be as informed as possible. However, they also have ideological proclivities that influence their choice of leader and that can prohibit truthful revelation of information obtained from private sources.

[3]Recent work has highlighted that democratically elected leaders are more likely to have higher academic credentials than un-elected ones Besley and Reynal-Querol (2011), and have analysed the party allocation of educated politicians to competitive races Galasso and Nannicini (2011). Further, since Stokes (1963), studies on electoral competition have long recognized the importance of candidates' valence and competence to secure electoral victory (McCurley and Mondak, 1995; Ansolabehere, Snyder, and Stewart, 2001; Burden, 2004; Stone and Simas, 2010).

[4]The Machiavellian suggestion allows for several interpretations. One is that the authority of an instrinsically valent leader is not threatened even if surrounding herself with highly capable, potentially hostile, associates. For example, Kearns Goodwin (2005) relates the political genius of Abraham Lincoln to her ability in forming a cabinet consisting of erstwhile rivals to her Presidency. Relatedly, an intrinsically valent leader may great benefit from her ability to appoint and manage a team that will guide her effectively.

In this setting a leader's judgement emerges endogenously via intra-group communication. Specifically, it depends upon the size of the group of truthful advisers a politician could rely upon were she selected as leader. Our first result is intuitive. We show that a leader can rely on truthful advice only from like-minded associates, whose ideology does not differ too much from her own. A larger group of such associates translates into more informed decisions, here. This simple and intuitive result establishes our take on the Machiavellian lesson: A leader's wisdom and judgement are determined by the men she has around her.

Building on this insight, we then ask what are the characteristics of a good leader? He is defined as the one that the group should choose when maximizing their joint welfare. Because ideology influences her choice, a moderate leader is perhaps desirable. Nevertheless, a good leader relies on her judgement that is determined by the number of allies in her circle. The choice of a good leader thus suggests a potential tradeoff between moderation, on the one hand, and judgement on the other. Further, this tradeoff is also present when considering the majority preferred leader, instead of the optimal one. Although preferences in our model are not (in general) single-peaked, the majority choice is determined by the median politician. This follows from the fact that the group's preferences satisfy a single-crossing condition. The median also trades-off competence and moderation when dictating the leader's election, albeit the weights she places on these two features differ from those that maximize welfare.

Most importantly, we can relate this tradeoff between ideological moderation and good judgment 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 relates to "global" properties of the ideology distribution, while judgement is related to "local" ones. We note that in standard models of collective choice that build on Blacks theorem, only changes to the global properties of the ideological distribution–specifically changes to the identity of the median–matter to outcomes.

Next we ask: under which conditions is the trade-off we described relevant to the collective choice? We consider a benchmark case where politicians views are evenly spread apart, or there are more politicians with moderate views than with extreme ones. There, "local" properties of the ideological distribution are irrelevant. As a moderate politician can then be no less informed than any other, she should take the decision on the group's behalf and indeed will be chosen to do so.

Beyond this benchmark case, we do not need to look far for a different scenario to emerge. Allowing for some clustering of politicians, a relatively extreme leader can yield a higher welfare to the group than the most moderate politician, and can be the leader preferred by the majority. This is so because she has a larger network of trustworthy, ideologically close associates and so can reach more informed decisions.

Further, in this case, the majority rule can lead to an inefficient choice: the elected leader need not be the optimal one. This much may not appear to be novel in the literature (although, here, the elected leader need not maximize the aggregate group welfare even if the ideology distribution is symmetric around the median, which therefore coincides with the mean). However, and most importantly, deeper insights emerge once we recognise our model as one of implicit delegation. As in other models of collective choice, the median player is decisive. But she need not take the leadership upon herself. Instead she can, though need not, delegate to another politician. When doing so she considers only her own preferences. So there may be welfare consequences: she may assume the leadership when it is in the groups interest that she delegates to another with better judgement; or, conversely, and more intriguingly, she may delegate when it is in the groups interest that she makes decisions herself. When she does so, then the elected leader has extreme views relative to the optimal leader. The surprising implication is that from a welfare perspective, the outcomes of *the election place too much emphasis* on the leader's judgment.

Our model reaches further surprising conclusions. The importance of local properties of the ideological distribution implies that a change in ideology of players other than the median can be consequential. This contrasts with standard models of collective choice. We explore this possibility via straightforward comparative statics exercises: fixing the ideology of other group members we ask what will happen as a player moves to the left or to the right of the spectrum. Performing these exercises reveals that a rightward shift in the ideology of one group member can have an opposite effect on leadership choice, making it more likely that a leftist leader is chosen, and vice-versa.

In sum, our simple model of leadership choice building upon the insight that a leader's judgement arises endogenously via her relations with others, yields the following possibilities with respect to its primitives: (i) changes in the ideological preferences of group members other than the median can impact the groups decision; (ii) a group may choose relatively extreme leaders, even when it is not in their interest to do so; and (iii) an ideological shift of group members can influence the groups choice of leader in the opposite direction. When combined these insights reveal that the outcomes of the collective choice of leader can change dramatically, when including information aggregation and advice in the study of leadership.

We discuss the implications of our findings for a wide range of applications, though focus attention on one in particular: the leader's selection within each one of two parties engaged in electoral competition. A conjecture is that many of our surprising findings would disappear in the presence of party competition. The pressure of competition may lead parties to chose moderate leaders, whose ideologies are close to the median voter's, as in the classic spatial models.

To assess this lead, we build a model in which each party chooses a leader via an internal election (involving politicians, members, and/or registered voters) before a general election that is contested by the two party leaders. The winner of the general election implement her preferred policies, as in the citizen candidate model of Osborne and Slivisky (1996) and Besley and Coate (1997).

In the benchmark case without communication between politicians, parties select as leaders their candidates whose ideologies are closest to the median voter's. Allowing for strategic communication completely changes the picture, and parties may opt for less moderate candidates. In fact competition between groups via an electoral contest between leaders may even strengthen our findings. There exist circumstances (with evenly spread apart political candidates' views) in which a party would choose a moderate leader in the absence of electoral competition, while choosing a relatively extreme when in competition with an opposing party.

Furthermore, our model yields several new insights to the study of party competition.

First, in contrast to existing models, we show that a political candidate may turn a winning situation into a losing one by *moderating her policy position*. Such a move can alter interparty relations in a way that reduces her leadership potential. Specifically, a leader may become more isolated when moderating her position, and so less able to aggregate valuable information and deliver informed policies.

Second, and in contrast to standard spatial models, any change in the ideological composition of a party can have an impact on its electoral fortunes. For example, the movement of some moderates toward the extremes or the moderation of views amongst extremists can enhance electoral prospects. Arguably, it was not the choice of Tony Blair as leader of the party that was decisive in transforming the Labour Party's electoral fortunes, it was the fact that leftists in the party moderated their position and became his trustworthy allies. By the same token, if some politicians become less moderate in their views, this can harm the prospects of a party even when they have a moderate leader who appeals to the electorate.

Third, our focus on local properties of the ideological distribution highlights the importance of party cohesion. A party that is relatively cohesive can win election even though the views of its leaders are far from the median. Although there is a large literature on cohesion with respect to legislative behaviour, there is no model, to our knowledge, that draws the connection between ideological cohesion and the electoral fortunes of parties.

Finally, and as in the case of a single group, a change in the ideological preference of a politician can change outcomes in the opposite direction. While such non-monotonic comparative statics are surprising, as they do not arise in standard spatial models, they nevertheless can help explain several instances of leadership selection that we discuss.

2. Related Literature

While we shall comment on and discuss our contributions throughout our essay, here we precede our analysis by briefly pointing out some of the main related literature. We contribute to a small but growing formal literature that develops different notions of leadership. Hermalin (1998) develops the notion of leading by example whereby a leader provides a costly signal that aligns followers' incentives with her own. Dewan and Myatt (2008) develop the notion of focal leadership that draws on earlier work by Schelling (1960) and Calvert (1995). Within the context of a beauty contest model, leaders are exogenous information sources that help party activists to advocate the best policies and coordinate their actions. Our notion of leadership is one where a leader's judgement depends on trustworthy allies who act as information sources: she must listen to others in order to reach more informed decisions.

We build a model in which a leaders characteristics are derived from first principles, and draw a distinction between a leaders judgement and her moderation. Relatedly Dewan and Myatt (2008, 2012) contrast a leader's judgement with her ability to communicate clearly; Bolton, Brunnermeier and Velkamp (2010) highlight the role of a leader's overconfidence with respect to a signal that informs a mission statement that she wants wants followers to coordinate on; Egorov and Sonin (2010) focus on the tradeoff between competence and loyalty to the leader.

In our model, a leaders judgement stems from the fact that others will truthfully convey policy relevant information to her. Information is then aggregated via internal party debates. As is customary in the formal-theoretic literature, we identify such debate with verbal (cheap talk) communication between privately-informed participants seeking to arrive at a collective choice. Landa and Meirowitz (2009) point out that political debate is meaningful, in the sense that it changes the preferences of politicians over outcomes, only when policy relevant information is not common knowledge. Consequently we develop our insights within the context of information aggregation via political debate, using the model by Galeotti, Ghiglino, and Squintani (2009) of multi-player communication between imperfectly informed players.

That model has numerous applications in the political science literature. Patty and Penn (2013) study information transmission in small networks of decision makers; Patty (2013) determines the optimal exclusion and inclusion policies to maximize information sharing among cabinet members; Gailmard and Patty (2009) study transparency and optimal delegation by a principal to informed agents; Dewan, Galeotti, Ghiglino, and Squintani (2011) investigate the optimal assignment of decision-making power in the executive of a parliamentary democracy; Penn (2014) studies the formation of stable aggregation of different units within an association; Dewan and Squintani (2014) the formation of party factions.

Unlike these papers, our current paper focuses on leadership, and derives a large set of novel and distictive findings. By studying the majority election of a group of politician's leader, we identifies instances in which the elected leader is not the median politician, and demonstrate that this may occur even if it were in the best interest of the group that the median politician (who acts as a dictator in the leader's election) took leadership upon herself. Further, we formulate and study a novel model of two party competition in which the voters anticipate multi-player communication within parties.

Our model provides the first derivation from first principles of electoral candidate's valence, in the form of good judgement. Valence is usually defined as 'all candidate's characteristics that benefit all voters regardless of their ideologies.' In this sense, one is led to believe that a leader's valence is independent of her ideology. However, in our microfoundation, a candidate's valence turns out to be related and partly determined by the ideological distribution of politicians in her own group. Among the formal theoretical models that have been developed to study the implication of valence on candidates' policies and electoral outcomes, one can see

Ansolabehere and Snyder (2000); Groseclose (2001); Aragones and Palfrey (2002); Callander and Wilkie (2007); Aragones and Palfrey (2002); Bernhardt, Camara, and Squintani (2011).

3. Model

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

Our model follows the multi-player communication set up by Galeotti, Ghiglino, and Squintani (2013). 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 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 $b_1 \leq b_2 \leq ... \leq b_n$. 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's 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 is thus a function m_i that depends on both s_i and j.

Communication between politicians allows information to be transferred: adopting the standard terminology, and up to relabelling of messages, we say that each communication strategy from i to j may be either *truthful*, so that $m_{ij}(s_i) = s_i$ for $s_i \in \{0, 1\}$; or, alternatively, it may be *"babbling"*, and in this case $m_{ij}(s_i)$ does not depend on s_i. We interpret the politicians who adopt the truthful strategy with respect to j as the trustworthy associates of that leader.[5]

After communication takes place, the leader chooses y so as to implement her preferred policy. We denote a decision strategy by leader j as $y_j : \{0,1\}^N \to \mathbb{R}$. Given the received messages $\hat{\mathbf{m}}_{-j}$, by sequential rationality, j chooses $\hat{y}_j$ to maximize her expected utility. So given the quadratic loss specification of players' payoffs, she chooses:

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

For a given a leader j, an equilibrium consists of the strategy pair $(\mathbf{m}, \mathbf{y})$ and a set of beliefs that are consistent with equilibrium play. Our equilibrium concept is pure-strategy Perfect Bayesian Equilibrium. Fixing the leader, there may be multiple equilibria $(\mathbf{m}, \mathbf{y})$. For example, the strategy profile where all players "babble" is always an equilibrium. The multiplicity of equilibrium predictions implies that the ranking of leader selection is not well defined: For the same leader j, different equilibria yield different player payoffs. This makes leadership selection difficult as it depends upon which equilibrium will be played. To deal with this issue we assume that for a given the leader j, politicians coordinate on the equilibria $(\mathbf{m}, \mathbf{y})$ that provides the highest expected payoffs to all politicians.[6] The selection of these equilibria is standard in games of communication. And it allows us to focus attention on leadership selection.

We consider two forms of leader selection.

The first one addresses a normative question: which leader would maximize politicians' welfare if chosen? Following the utilitarian principle, we define welfare as the sum of players' expected payoffs. Formally, define the *optimal leader* as player j who induces equilibria $(\mathbf{m}, \mathbf{y})$ with the largest sum of expected payoffs:

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

We denote $W^*(j)$ as the maximal payoffs associated with selection of the optimal leader.

[5]Individuals adopting the babbling strategy with a leader j can be intepreted as "yes-men", who always give the same advice to the leader, regardless of their information.

[6]Indeed, it can be easily shown that for any given leader j, each politicians' ranking among the possible equilibria $(\mathbf{m}, \mathbf{y})$ is the same (see Galeotti, Ghiglino, and Squintani (2013), Theorem 2.)

The second determines which player will be elected by majority rule. Player i's payoff when j is chosen as leader solves

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

Once again, this is associated with the equilibrium $(\mathbf{m}, \mathbf{y})$ that provides the highest expected payoff among the equilibria induced by j: we denote it as $U_i^*(j)$. 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 perfectly informed, a politician becomes better informed the more other members truthfully reveal their signals to her. 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 trustworthy associates is related to key primitives of our model, namely the ordering of ideological biases within the group.

4.1. **Trustworthy Advisers.** The equilibrium communication structure given any chosen leader j is easily derived, following Corollary 1 by Galeotti, Ghiglino, and Squintani (2013). We let $d_j(\mathbf{m})$ be the number of politicians willing to truthfully advise j were she to lead group. These politicians form the group of trustworthy associates of j. We prove (in the Appendix) that the profile $\mathbf{m}$ is an equilibrium if and only if, whenever i is truthful to j,

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

An important consequence of the equilibrium condition (2) is that truthful communication from politician i to leader j becomes less likely with an increase in the difference between their ideological positions.[7]

[7] A perhaps more surprising effect is that the possibility for i to communicate truthfully with j decreases with the information held by j in equilibrium. To see why communication from i to j is less likely to be truthful when j is well informed in equilibrium, suppose that $b_i > b_j$, so that i's ideology is to the right of j's bliss point. Suppose j is well informed and that politician i deviates from the truthful communication strategy –she reports $\hat{m}_{ij} = 1$ when $s_i = 0$–then she will induce a small shift of j's action to the right. Such a small shift in j's action is always beneficial in expectation to i, as it brings j's action closer to i's (expected) bliss point. Hence, politician i will not be able to truthfully communicate the signal $s_i = 0$. By contrast, when j has a small number of players

We can use this result to derive how informed politician j would be in the event where she becomes leader.

First we note that the term $d_j(\mathbf{m})$ is a function of the communication strategies 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. It proves useful then to define d_j^* as the maximal $d_j(\mathbf{m})$. That is, the most information that j could obtain under the assumption that any politician who could communicate truthfully will in fact do so. 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*.

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 the 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 $\{3, 4, 5, 6, 7\}$, 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 j's network of trustworthy associates d_j^*.

Proposition 1. *For any politician j, the maximal equilibrium number of truthful associates d_j^* is the unique $d \in \mathbb{N}$ which solves the equation*

$$N_j\left(\frac{1}{2(d+3)}\right) = d. \quad (3)$$

This result 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

communicating with her, then i's report $\hat{m}_{ij} = 1$ moves j's action to the right significantly, possibly beyond i's bliss point. In this case, biasing rightwards j's action may result in a loss for politician i and so she would prefer to report truthfully- that is, she will not deviate from the truthful communication strategy.

politicians closest to j have bias distant less than $1/12$ from b_j, i.e. they have a bias in the interval $(-1/12, 1/12)$. Then we can conclude that these three politicians would provide truthful advice to j, if she were selected as the 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 ideological neighborhood of a leader j to which a politician need 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.

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 of them 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, in our model.

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 distance between b_j and the average ideology $\sum_{i \in N} b_i/n$. 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:

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

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 $n[6(d_j^*+3)]^{-1}$. Evidently, a more moderate leader, whose bias is closer to average ideology $\sum_{i\in N} b_i/n$, makes the aggregate ideological loss $\sum_{i\in N}(b_i - b_j)^2$ smaller.

Further, the residual variance $[6(d_j^* + 3)]^{-1}$ is inversely related to the size of the leader's maximal informant set d_j^* and hence to her judgement.[8] Thus, optimal leader selection takes into account each politicians' 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 $\mathrm{b} = \{b_1, ..., b_n\}$ in the group. In fact, every element of b, even extreme ones, matters for the determination of the average ideology $\sum_{i\in N} b_i/n$. 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 (2). The leader's understanding is thus defined by those close to her, or adopting Machiavelli's text, by *"the men he has around him"*. This analysis of the role played by the local ideological distribution is novel in the large contemporary and formal literature on collective choice; though it echoes the insights of Machiavelli made in his masterpiece of 500 years ago.

We summarize our findings as follows.

Proposition 2. *A good leader j maximizes $W^*(j)$. Hence, optimal leadership requires ideological moderation: the leader j's policy should reflect the diversity of views in the group. Optimal leadership also requires judgement. This stems from the information that the leader j obtains from the politicians she has around her: the close-minded associates defined in proposition 1.*

5.2. **Electing the Leader.** We now determine which politician is elected as leader by simple majority decision in the group of politicians. The implications of majority rule are not straightforward in this context, because the politicians' utilities are not single-peaked with

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

respect to the leader's identity, and so Black's theorem cannot be applied. In fact, each player i's utility as a function of the leader's identity j is:

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

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^*. Player's preferences need not be single peaked because a politician who is ideologically distant may in fact be better informed, and so have better judgement, than one who is ideologically similar. However, we can establish the weaker result that utility functions are single-crossing.

Lemma 1. *The utility functions $U_i(j)$ are single crossing in i and j: if $i < j$ and $i' < i''$ then $U_j(i') > U_j(i'') \Rightarrow U_i(i') > U_i(i'')$; and if $i > j$ and $i'' > i'$ then $U_j(i'') > U_j(i') \Rightarrow U_i(i'') > U_i(i')$.*

As a consequence of this result we can appeal to a result by Gans and Smart (1996) to show that the player with median ideology will be act as a "dictator"in the election. The unique Condorcet winner of the election game is the politician j who maximizes the expected payoff of the median player.

Proposition 3. *The group I elects as leader the player j who maximizes the utility $U_m^*(j)$ of the median politician $m = (n+1)/2$. The collective choice considers the ideological proximity of any player j to m, as well as j's judgement that is determined by her number of close-minded associates.*

Having established the outcome of majority election, we can compare it with the optimal leader selection by inspecting expressions (4) and (5), the latter for $i = m$. As for leader selection, there is a trade-off between moderation and judgement, here: the Condorcet winner j should keep both the ideological loss $(b_m - b_j)^2$ and the residual variance $\frac{1}{6(d_j^*+3)}$ as low as possible. Just like optimal leadership, the majority choice trades off the desire for a moderate leader with the desire for a leader with good judgement, i.e., with a large group of close-minded associates.

Nevertheless, a leader chosen by the majority trades off moderation and judgment optimally with respect to the median's politician preferences, whereas the optimally selected leader trades off moderation and judgment optimally for the whole group. As the weights placed on these two features of good leadership are different, the majority choice may differ from the optimal leader, as we detail in the next section. Surprisingly, we will identify instances in which the median politician's utility $U^*(\cdot)$ place less weight on moderation (and more on judgement), than the group's welfare $W^*(\cdot)$, so that the majority choice may be inefficient because it places *too much weight* on the leader's judgement. This surprising finding highlights the predictive importance of judgement in the positive study of leadership. And the finding is even more surprising when reinterpreting the decisive median's vote in the election, as a model of delegation of authority by the median to (possibly) another politician. Surprisingly, the median has a strong incentive to delegate, here, as it may relinquish control to another politician, even when it is the group's interest that she exercise power herself!

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 $m = (n+1)/2$, while the optimal one would be the one whose bias is the closest to the average bias $\sum_{i=1}^{n} b_i/n$. 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 $\mathbf{b}$), the most moderate politician is the optimal leader and the majority-preferred one. Evidently, this is the case, for example, when there are only 3 politicians in the group. Then, the middle one is the most moderate, and it cannot be the case that she is less informed than either of the extreme ones: If she is willing to communicate truthfully with their neighbors, then at least one of them is willing to be truthful to her. [9]

Moving beyond the three-player case, we illustrate sufficient conditions for optimal selection to deliver the most moderate politician as leader. Doing so, we consider the situation in which

[9]This reasoning can be pushed one step further. The most extreme politicians 1 and n can never be chosen as leaders, as they do not have better judgement than their more moderate neighbors. Simply put: if a player $i > 1$ is willing to communicate truthfully to 1 in equilibrium, then also 1 is willing to communicate truthfully to i, and i can count on left neighbors $k < i$ who may be willing to communicate to her, whereas 1 does not have any left neighbors available.

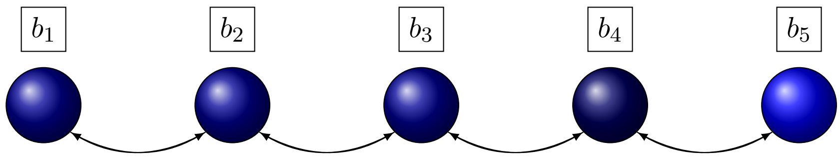

FIGURE 1. Leadership Choice with Equidistant Bias: this figure depicts the case where $b_i - b_{i-1} = \beta$ for all politicians $i = 2, ..5 and$1/12¡$\beta \leq 1/10$. An arrow linking two politicians illustrates that truthful communication is sustained between them.

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 politician $m = (n+1)/2$ is the most moderate, and has at least as many trustworthy advisers as any other politician. So, 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 the groups behalf. Similarly, she would be the unique choice of the majority, were an election held.

We formalize this insight in the following proposition, in which we prove an even stronger result. We show that the median politician m is always elected by the majority as leader, and should always optimally lead the group of politicians, also when the ideology distribution is symmetric around m and 'single peaked' at m, in the sense that politicians are more ideologically clustered as they get closer to the median politician m. Formally, we define the ideology distribution b as 'single peaked' and symmetric at m, when for any $i = 1, ..m - 1$, $b_{i+1} - b_i$ weakly increases in i, and $b_{i+1} - b_i = b_{n-i+1} - b_{n-i}$. Evidently, the case in which politicians' ideologies are evenly distributed on the line, so that 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 in the definition of b single peaked and symmetric at m.

Proposition 4. *When politicians' ideologies* b *are single peaked and symmetric at* m*, the median politician* m *is also maximally competent. 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 b_2, b_3, and b_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.

6.2. **The Case with 5 Politicians.** Proposition 4 relates the core characteristics of leadership to reveal that, with equidistant biases, the most moderate politician is also (weakly) the

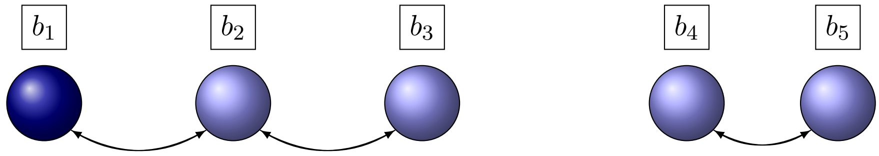

FIGURE 2. Leadership Choice with Non-equidistant Bias: illustrates a case where $b_1 = -(\alpha + \beta)$, $b_2 = -\beta$, $b_3 = 0$, $b_4 = \gamma$ and $b_5 = \beta + \gamma$. An arrow linking two politicians illustrates that truthful communication is sustained between them. Here $\alpha \leq 1/10$, $\beta \leq 1/10$ and $\delta \leq 1/10 < \gamma$. So each player is informed by her immediate neighbours.

better informed and so has better judgement. 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 not evenly distributed. We explore this possibility by studying in depth the case of 5 politicians, where, without loss of generality, we denote the ideologies as $b_1 = -(\alpha + \beta)$, $b_2 = -\beta$, $b_3 = 0$, $b_4 = \gamma$, $b_5 = \gamma + \delta$. This case is rich enough to identify properties of the ideology distribution that lead to the novel and interesting results we introduced earlier. At the same time, it is sufficiently simple to be manageably exposited.

When calculating the optimal and majority-elected leader, we first note that it can never be optimal that players 1 and 5 lead the group, or that they will be elected by the group as leaders (see footnote 6.1). Players 2 and 4 can be chosen as leaders if and only if they have better judgement than player 3. Also, interchanging α with δ and β with γ players 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 equation (3), we calculate (in the appendix) all cases for which the condition $d_2^* > d_3^*$ holds, and determine the restriction each one of them imposes on the parameters α, β,γ and δ. Here, we illustrate our findings in the case in which $d_2^* = 2$ and $d_3^* = 1$. This holds when $\alpha \leq 1/10$, $\beta \leq 1/10$ but $\gamma > 1/10$. 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 again the arrows indicate the trustworthy associates of each player in the event she were chosen as leader. As we can see, the ideologies are not distributed at even distances, and this may induce a tradeoff between judgement and moderation.

The choice between the more moderate politician 3 and politician 2, who has a better judgment, is resolved in favor of the former or the latter depending on whether the views of politician 2 are too extreme, or sufficiently moderate. Formally, the views of 2 are more extreme the larger is β. Hence, we can relate the choice between politician 2 and 3 as leader, according to the size of β relative to the other primitives of the model as demonstrated by the following result.

Lemma 2. *Consider the case of* 5 *politicians, with the above restrictions:* $W^*(2) \geq W^*(4)$, $U_3^*(2) \geq U_3^*(4)$, $\alpha \leq 1/10$, $\beta \leq 1/10$ *and* $\gamma > 1/10$.

- *If* $\beta < \frac{\sqrt{30}}{60}$, *then the Condorcet winner is politician* 2, *else, the most moderate politician* 3 *wins the majority choice.*
- *Letting* $\phi = \delta - \alpha + 2\gamma > 1/10$, *if* $\beta < \tau(\phi) \equiv \frac{\sqrt{6}}{12}\sqrt{24\phi^2 + 1} - \phi$, *then the optimal leader is* 2, *and else it is* 3.
- *There is a unique* $\bar{\phi} > 1/10$ *such that* $\tau(\phi) > \frac{\sqrt{30}}{60}$ *for all* $\phi < \bar{\phi}$ *whereas* $\tau(\phi) < \frac{\sqrt{30}}{60}$ *for all* $\phi > \bar{\phi}$.

The result defines a welfare threshold, $\tau(\phi)$. The group is better off when player 2 takes the decision if and only if β, the ideological distance between players 2 and 3, is below $\tau(\phi)$. This threshold, in turn, depends upon the values of α, δ, and γ. Intuitively, it is optimal that player 2 leads the group when her better judgement, combined with the benefits to those the left of the spectrum (α) are not outweighed by the ideological loss incurred by those to the right ($\gamma + \delta$).

Lemma 2 also defines a majority threshold. This takes a much simpler form as it depends only on the views of the median player. She may obtain a more informed outcome when 2 takes the decision and this yields a constant addition to her utility. But this comes at an ideological cost β. Thus the group will choose 2 as leader if and only if β is below a threshold given by the constant $\frac{\sqrt{30}}{60}$.

The final part of lemma 2 reveals that in equilibrium the welfare threshold $\tau(\phi)$ can either be larger or smaller than the majority threshold $\frac{\sqrt{30}}{60}$, depending on how large ϕ is. In the former case, we can distinguish three possibilities, on the basis of β. For small β, i.e., $\beta < \frac{\sqrt{30}}{60}$, the non-moderate politician 2 is both the Condorcet winner and the optimal leader; for large

β, specifically, $\beta > \tau(\phi)$, the most moderate politician 3 is both the Condorcet winner and the optimal leader; in the intermediate case in which $\frac{\sqrt{30}}{60} < \beta < \tau(\phi)$, the optimal leader is the most moderate politician 2, whereas the majority elects the non-moderate politician 3. We have identified an instance in which majority voting leads to an inefficient leader's choice. The finding is somewhat expected: The median politician 3, acting as a dictator in the majority vote, retains leadership for herself, although it would be optimal for the group if she delegated to the better informed politician 2.

More interesting and unexpected is the case in which $\tau(\phi) < \frac{\sqrt{30}}{60}$. Again, for small β, the non-moderate politician 2 is both the Condorcet winner and the optimal leader, and for large β the most moderate politician 3 is both the Condorcet winner and the optimal leader. But now, the inefficiency that arises in the intermediate case in which $\tau(\phi) < \beta < \frac{\sqrt{30}}{60}$ is surprising: the optimal leader is the most moderate politician 3, whereas the majority elects the non-moderate politician 2. In other terms, the median 'majority dictator', politician 3, delegates leadership to the better informed politician 2, although it would be optimal for the group if she retained it for herself.

The logic behind this result is simple. The median politician may trade off moderation and competence in a way that differs from the optimal choices made by a social planner. Starting from her ideal point, the median politician 3, who by definition is moderate relative to the entire spectrum of opinion in the group, may sacrifice policy for a more informed outcome. In our 5 player example the median politician, 3 will indeed do so when $b_2 - b_3$ is sufficiently small and $d_2 > d_3$. A change to the left in identity of the decision-maker then benefits the median; and of course the leftist members of the group, 1 and 2. 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 the rightwing politicians 4 and 5, than it is beneficial to the leftist politicians 1 and 2. Then it may be the case that a social planner would force the median politician to take the decision, if she could.

In addition to these surprising findings on the contrast between optimal leadership and majority selection of leaders, the analysis of the 5 politician case allows us to uncover also some surprising comparative statics results: A player changing her ideology from right (left) to left (right) can induce a shift in leadership in the opposite direction.[10]

To illustrate, consider a benchmark case with evenly distributed ideologies players, in which politicians 2, 3, and 4 can all count on the truthful advice of their ideological neighbors, so that $1/10 < \alpha = \beta = \gamma = \delta \leq 1/8$. Then following Proposition 4, 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 away from the bias of the median player 3 so that they are no longer truthful to one another (i.e., suppose that γ increases so as to become larger than $1/8$). Then 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 him, despite not being the most moderate politician. Indeed, by Lemma 2, we know that this is the case when $\beta < \frac{\sqrt{30}}{60}$. 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., $\alpha = \beta = \gamma = \delta > 1/8$. 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 α and β decrease, so that they both become smaller than $1/8$). Because player 4 is still not truthful to 3, the leadership switches from the median player 3 to the centre-right player 2, again, when $\beta < \frac{\sqrt{30}}{60}$. Here, the ideological movement of players towards more moderate positions may lead to the capture of the control of the politicians' group.

In both cases, somewhat unexpectedly, changes in the ideology distributions by which all politicians weakly move their ideology to the right lead to a shift in the group decision to the left.

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

Proposition 5. *When ideology is not evenly distributed on the line, a politician other than the most moderate one can be the optimal leader and the majority choice.*

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

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

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

6.3. **Discussion.** An important insight from proposition 5 emerges when we recognize our model as one of strategic delegation 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 whereas Chari, Jones, and Marimon (1997) suggest that the opposite occurs in voting contexts.[11] 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 groups interest that she execute the decision herself.

Viewing our model as one of implicit delegation, then proposition 5 reveals a failure of the famous ally principle. This principle 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 principles and agents as a 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 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.

[11]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.

A large body of literature has explored the process of nominations and appointments to majoritarian institutions. For example Shepsle and Rohde and Krehbiel 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-dependant institutions, as well as multi-player interactions with each of these institutions, the possible set of strategies to consider are 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. 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. As noted by Clarke, "the former case is essentially an application of the median voter theorem to the supreme court," as it rests on, "the assumption that an opinion must gain the assent of four justices, the median justice and four justices on one side or another." 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 then the median may achieve majority support. In her review of the field Clarke suggests that relaxing the complete information assumption in standard models may yield new insights and 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 Blacks 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.

The question we ask is whether the central insights of our analysis of group choice survive when the leader is subject to competition.

7. Electoral Competition and Endogenous Valence

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 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 with the party leader, although this 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 Slivinsky (1996) and Besley and Coate (1997), the winner of the election implements her ideal policy. Here, though, she does so only after consultation with informed leading politicians in her own party. 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 leading politicians in party A consist of the set of politicians $N_A = \{1, ..., q\}$ and those in B consist of politicians $N = \{q+1, ..., n\}$. 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 only assume 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 (to simplify the exposition, we consider only ideology and party profiles such that there is only one such politician in each party).[12] The eventual winner $j \in \{a, b\}$ of the general

[12] These assumptions are very weak and would be satisfied in a number of micro-founded models of leader selection within our framework. For example, it may be that only the leading politicians in the party have real influence on party leadership. Another possibility is to say that each 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.

election then implements the final policy $\hat{y} \in \mathbb{R}$. Following the citizen-candidate paradigm, candidates cannot credibly commit to their electoral promises and so will implement their preferred policy if elected. But unlike in basic citizen-candidate models, the winner of the election makes his decision only after consultation with her party leading politicians, as we detail below.

There are 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:

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

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 politician i has some private information on θ. After the general election takes place, each politician 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 the elected politician has truthful associates only within her own party, and so restrict attention to equilibria in which the politicians from the opposite party do not reveal any information to her.

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

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

Because each voter i evaluates a candidate j's judgement favorably, regardless of her ideology, we take a candidate judgment to be her *valence*. Such a candidate j's valence is here endogenously determined by j's network of trustworthy party fellows. Hence, ours is a model of electoral competition with endogenous valence.

As a consequence of Lemma 1, voter 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 in the general election.

7.2. **Results.** A natural benchmark for comparison is an otherwise identical model in which players are not allowed to 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.*

An immediate consequence of this result is the following. The winner of the general election is the politician with ideology closest to zero, also in our model of elections with endogenous valence, when politicians are sufficiently ideologically distant from each other. In this case, they do not truthfully communicate in equilibrium.

Departing from this simple case, it is immediate to see that, in our model of electoral competition, party leaders need not be moderate, because of the same logic as in the previous section. Again, politicians with a large network of truthful informants may be the best prospects for the median voter even if they have relatively extreme ideologies. Further, our model of electoral competition allows us to uncover additional insights, beyong the findings of the previous section.

The first one reveals 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. This finding stands in sharp contrast with Proposition 4, in the previous section. The reason

FIGURE 3. Party Competition with Equidistant Bias: The figure illustrates a case where politicians $1, 2, 3$ belong to party A and $3, 4, 5$ to party B. Within each party each politician is informed by her immediate neighbour. Then for some parameter values provided in the text, party A chooses candidate 2 and Party B chooses candidate 5 as leader with each equally likely to win the general election.

is that, when politicians are partitioned in competing parties, the most moderate politician within the electorate as a whole, may well be at an extreme end of the ideological spectrum within her party. This constrains the pool of trustworthy associates within her party, and may hamper her capability of making informed decisions if elected in office.

Proposition 6. *Even if the ideologies* b *of the potential candidates* $N_A \cup N_B$ *are evenly distributed, so that there exists* β *and that* $b_{i+1} - b_i = \beta$ *for 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 belong to party A and the others 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 elected in the primaries 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 $1/20 < \beta < \sqrt{15}/60$. 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.

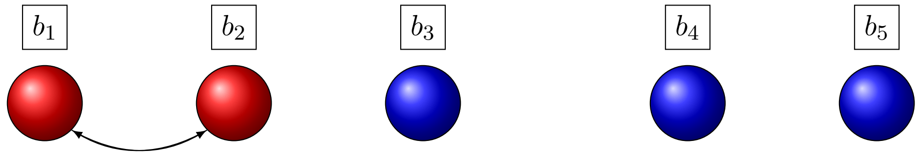

FIGURE 4. Party Competition with a majoritarian party: illustrates a case where politicians 1 and 2 are in party A and $3, 4, 5$ are in party B;$b_1 = -(\alpha + \beta)$, $b_2 = -\beta$, $b_3 = \gamma$, $b_4 = \gamma + \delta$ and $b_5 = \gamma + 2\delta$ and $\gamma < \beta$. The figure depicts within party communication where arrows linking two politicians illustrates that truthful communication is sustained between them. Here $\alpha \leq 1/8$ and $\delta > 1/8$.

The second novel insight of our analysis of electoral competition with endogenous valence is a surprising finding which highligths the value of a party's ideological cohesiveness when engaging in political competition. Interpreting a party's cohesiveness as the ideological distance among its leading members, we find that a more cohesive party can defeat a larger, less cohesive, party in a general election, even in the case that the larger party can draw information from a larger set of informed leading party members, and even if the large party's leader's views are close to the median voter's. This happens because the leader of the more cohesive party can count on more trustworthy associates than her opponent in the general election, and hence the median voter anticipates that she will have a better judgment if elected.

Finally, we find that the outcome of the election may depend on the whole ideological distribution in often a very subtle way. For example, it may happen that a party leader becomes ideologically more moderate, and loses the general election as a result. (Of course, also the opposite can happen: a politician may lose the election by becoming more extremist). Intuitively, this occurs because, by becoming more moderate, the party leader makes the ideological distance with the other leading members of the party too large, and loses the capability of gathering truthful advice from them. This result is, to the best of our knowledge, both novel and unsupported by any variant of the standard spatial model to be found in the literature.

Proposition 7. *A large party may lose the election to a smaller, more cohesive party, even if it can draw information from a larger number of leading members, and even if its leader is the general election candidate whose views are closest to the median voter's.*

The outcome of the election may depend on the whole ideological distribution of leading politicians a subtle way: For example, a party leader may moderate her views closer to the median voter's, and loses the general election as a result.

The result can be demonstrated by means of the following example, illustrated in figure 4. Suppose that there are 5 politicians, with ideologies $b_1 = -(\alpha+\beta)$, $b_2 = -\beta$, $b_3 = \gamma$, $b_4 = \gamma+\delta$ and $b_5 = \gamma + 2\delta$ with $\gamma < \beta$. Politicians 1 and 2 belong to party A and $3, 4, 5$ belong to party B. In this example, party B is not only larger in the sense that there is a larger set of leading politicians from whom the selected leader can draw information, but also in the sense that it can express a leader, player 3, whose views are the closest to the median voter in the general election. In fact, the mid-point between the ideological views of the most moderate politicians in parties A and B, that is defined by $M = (-\beta+\gamma)/2$, is to the left of zero, the ideal point of the median voter. 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 be 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 $\alpha \leq 1/8$ whereas $\delta > 1/8$, and is depicted in Figure 4. It is then not difficult to find conditions under which the median voter prefers to elect politician 2 from party A rather than politician 3 from party B, so that, in equilibrium, B will lose the election despite being the largest party. Specifically, we find in the appendix that this is the case if and only if $(\beta-\gamma)(\beta+\gamma) < 1/72$: this condition is satified when the mid-point $M = (-\beta+\gamma)/2 < 0$ is not too far from zero, the median voter's ideal point.

To see that a leader can lose the election by moving closer to the median voter, consider politician 2 the leader of party A. Suppose that we start from a situation in which α is close to $1/8$, candidate 2 is barely within range of 1 for be to be truthful to her. If politician 2 moves ideologically closer to the median voter, and 1's position remains fixed, then the condition $\alpha \leq 1/8$ will not be satisfied anymore, candidate 2 will lose the informational advantage over 3 and she will lose the election.[13]

[13]The claim that a candidate can lose the election by moving closer to the median voter can also be proved by making γ larger and reducing δ so that b_4 and b_5 remain constant and b_3 moves closer to b_4 thereby making it possible for the leader 3 to gather politician 4's truthful advice.

8. Conclusion and Discussion

This paper has proposed a novel take on leadership, relating a leader's judgemen to the truthful advice of her trustworthy associates, and, in turn, relating advice to ideological proximity to the leader. As a result, we have formulated a theory of how leaders emerge and of what makes a good leader starting from first principle. By setting the analysis within a group of politician that may or may not be in electoral competition with another party, we have uncovered a number of novel results, which can explain documented facts that cannot be reconciled with previous theories.

This is the first paper providing a theory of leader's characteristics from first principles. Its value is therefore also in stimulating further research on this important topic. One interesting question, which we leave for future research is how our results would change if different politicians had different access to information. It is possible that ideologically close politicians are likely to gather information largely from the same sources. As a result, the advice of associates who are too ideologically close may be less valuable than the truthful advice of 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 it would survive in any model of communication since Crawford and Sobel, 1982). As a result, all our possibility results would extend to this enriched model. Further, it would be interesting to determine the optimal composition of advisers to the leader. We conjecture that the optimal set of advisors 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 theirrecommendations would not be credible, it is possible that politicians who are too close ideologically would be also excluded not to crowd out more valuable less ideologically close advisors (see the discussion in footnote 7.)

We shall end our essay by returning to more general points that stem from the Machiavellian lesson: a leader's wisdom and judgement are determined by the men she has around her. Our take on this lesson is that a leader's judgement emerges endogenously through interaction with trustworthy associates who provide valuable information to a leader. Set in the context of party competition, the lesson takes on new resonance and delivers new insights that can be summarised as follows. The trusting relations that a leader enjoys with her party are important to party success. Such relations depend, in turn, upon the ideological distance between politicians in a party: the "local" properties of the ideological distribution. These relations are fragile as changes in the ideological position of a leader, or those of her trustworthy allies, can have electoral implications. When a leader moves to the center-ground, or her allies move to the extremes, a potential winning situation can turn into a losing one.

One way to conceptualise this insight is via the notion of party cohesion. As proposition 7 above shows, ideologically cohesive parties can win elections even against parties whose leaders are ideologically closer to the median voter and so, intuitively, cohesion can be an electoral asset. There are, of course, many reasons why this might be the case: a cohesive party may be better able to coordinate legislative activity, and/or provide commitment for its leader. But in parliamentary systems where it is generally high, legislative cohesion is unlikely to provide a source of electoral advantage. Therefore, and with respect to such cases at least, our take on Machiavelli's insight provides a more compelling mechanism linking cohesion with electoral success. Put simply, cohesion sustains intra-party deliberations that, in turn, inform a leader's judgement.

Party cohesion can, then, be related to a leader's judgement. We can go further, in fact, in saying that the cohesion of a party is more important than the moderation of its leader. As we have shown, a moderate leader of one party can be defeated by an extreme leader of a smaller, yet more cohesive, party. It follows, as a corollary to proposition 7, that the moderation of its leader is neither a necessary nor sufficient condition for party success. In order to assess the chances of a party at the elections, it is essential to consider the relationship between the party leader's ideology and the ideology of the other leading members. A cohesive party stands a higher chance of victory than even a larger party.

This important result is novel in the formal literature on party competition, and it provides a new interpretation of several instances of party transformation and success. A remarkable success story has been the transformation of the British Labour party during the 1990's (though arguably the process was initiated much earlier), culminating in the election for three terms of prime minister Blair. Tony Blair's moderation has been much commented upon as pivotal to New Labours success. Our result suggests, in fact, that it may have been the cohesion of New Labour that made possible its remarkable success.

We have defined cohesion in terms of the trustworthy relations that bind party politicians to their leader. 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 of his memoirs) "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 extremes 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 aspect of New Labour's success and Blair's leadership of the party. More generally, our model provides a formal framework that links such inter-personal partisan ties to a leaders judgement and her success.

References

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.

BERNHARDT, D., O. CAMARA, 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).

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

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," *Games and Economic Behavior*, 60, 262–286.

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

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

COUGHLIN, P. (2001): "In defense of unanimous jury verdicts: mistrials, communication and strategic voting," *American Political Science Review*, 94, 975–79.

DEWAN, T., A. GALEOTTI, C. GHIGLINO, AND F. SQUINTANI (2011): "Information Aggregation and Optimal Executive Structure," *Working paper*.

DEWAN, T., AND D. P. MYATT (2008): "The Qualities of Leadership: Communication, Direction and Obfuscation," *American Political Science Review*, 102(3), 351–368.

GAILMARD, S., AND J. W. PATTY (2009): "Separation of Powers, Information, and Bureaucratic Structure," .

GALASSO, V., AND T. NANNICINI (2011): "Competing on Good Politicians," *American Political Science Review*, 105(1).

GALEOTTI, A., C. GHIGLINO, AND F. SQUINTANI (2009): "Information Transmission in Networks," *Working paper*.

GALEOTTI, A., C. GHIGLINO, AND F. SQUINTANI (2013): "Information Transmission Networks," *Journal of Economic Theory*, Forthcoming.

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.

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

LANDA, D., AND A. MEIROWITZ (2009): "Game Theory, Information, and Deliberative Democracy," *American Journal of Political Science*, 53(2), 427–444.

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.

MORGAN, J., AND P. STOCKEN (2008): “Information Aggregation in Polls,” *American Economic Review*, 98(3), 864–896.

PATTY, J. W. (2013): “A Theory of Cabinet-Making: The Politics of Inclusion, Exclusion, and Information,” .

PATTY, J. W., AND E. M. PENN (2013): “Sequential Decisionmaking and Information Aggregation in Small Networks,” .

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

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.

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 a politician i holds n bits of information, i.e. she holds the private signal s_i and $n-1$ politicians truthfully reveal their signal to her. The probability that l out of such n signals equal one, conditional on θ is

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

Hence, politician i's posterior is

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

the expected value is

$$E\left(\theta|l,n\right)=\frac{l+1}{n+2},$$

and the variance is

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

■

Derivation of Expression 2. Consider any player j, and let $C_j(\mathbf{m})$ be the set of players truthfully communicating with j in equilibrium. The equilibrium information of j is thus $d_j(\mathbf{m})=\#C_j(\mathbf{m})+1$, the cardinality of $C_j(\mathbf{m})$ plus j's signal. Consider any player $i\in C_j(\mathbf{m})$. Let s_R be the vector containing s_j and the (truthful) messages of all players in $C_j(\mathbf{m})$ except i. Let also $y^{j}_{s_R,s}$ be the action that j would take if she has information s_R and believed in the signal s sent from player i, analogously, $y^{j}_{s_R,1-s}$ is the action that j would take if she has information s_R and believed in the signal $1-s$ sent from player i. Agent i reports truthfully signal s to the leader j if and only if

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

Using the identity $a^2-b^2=(a-b)(a+b)$ and simplifying, we obtain:

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

Next, observing that

$$y^j_{s_R,s} = b_j + E\left[\theta|s_R, s\right],$$

we obtain

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

Denote

$$\Delta\left(s_R,s\right) = E\left[\theta|s_R,s\right] - E\left[\theta|s_R,1-s\right].$$

Observing that:

$$f(\theta,s_R|s) = f(\theta|s_R,s)\Pr(s_R|s),$$

and simplifying, we get:

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

Furthermore,

$$\int_0^1 \theta f(\theta|s_R,s)d\theta = E[\theta|s_R,s],$$

and

$$\int_0^1 f(\theta|s_R,s)E[\theta|s_R,s]d\theta = E[\theta|s_R,s],$$

because $E[\theta|s_R,s]$ does not depend on θ. Therefore, we obtain:

$$\begin{aligned} & -\sum_{s_R\in\{0,1\}^{d_j(\mathbf{m})}} \left[\Delta\left(s_R,s\right)\left(\frac{E\left[\theta|s_R,s\right]+E\left[\theta|s_R,1-s\right]}{2} + b_j - b_i - E\left[\theta|s_R,s\right]\right)\right] P(s_R|s) \\ = & -\sum_{s_R\in\{0,1\}^{d_j(\mathbf{m})}} \left[\Delta\left(s_R,s\right)\left(-\frac{E\left[\theta|s_R,s\right]-E\left[\theta|s_R,1-s\right]}{2} + b_j - b_i\right)\right] P(s_R|s) \geq 0. \end{aligned}$$

Now, note that:

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

where l is the number of digits equal to one in s_R. Hence, we obtain that agent i is willing to communicate to agent j the signal $s=0$ if and only if:

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

or

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

and note that this condition is redundant if $b_j - b_i > 0$. On the other hand, she is willing to communicate to agent j the signal $s=1$ if and only if:

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

or

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

and note that this condition is redundant if $b_j - b_i < 0$. Collecting the two conditions yields:

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

i.e., expression (2). ■

Proof of Proposition ??. Because $N_j(\cdot)$ is an increasing step function, and $1/[2(d+2)]$ strictly decreases in d, whereas the identity function is strictly increasing in d, there is a unique solution to equation (3). From equilibrium condition 4, we see that maximization of $W(\mathbf{m},\mathbf{y})$ is equivalent to maximization of the equilibrium information $d_j(\mathbf{m})$. Inspection of the equilibrium condition 2 shows that the maximal information of the leader j can be calculated according to equation (3). ■

Derivation of equilibrium welfare, expression 4. Assume $(\mathbf{m}, \mathbf{y})$ is an equilibrium. The ex-ante expected utility of each player i is:

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

where Ω_j denotes the equilibrium information of the leader j. Hence

$$\begin{aligned} Eu_i(\mathbf{m}, \mathbf{y}) &= -E\left[(b_j - b_i)^2 + \left(E\left[\theta|\,\Omega_j\right] - \theta\right)^2 - 2(b_j - b_i)\left(E\left[\theta|\,\Omega_j\right] - \theta\right); \mathbf{m}\right] \\ &= -\left[(b_j - b_i)^2 + E\left[\left(E\left[\theta|\,\Omega_j\right] - \theta\right)^2; \mathbf{m}\right] \right. \\ &\quad \left. -2(b_j - b_i)\left(E[E\left[\theta|\,\Omega_j\right]; \mathbf{m}] - E[\theta; \mathbf{m}]\right)\right], \end{aligned}$$

by the law of iterated expectations, $E[E\left[\theta|\,\Omega_j\right]; \mathbf{m}] = E[\theta; \mathbf{m}]$, and by definition $E\left[\left(E\left[\theta|\,\Omega_j\right] - \theta\right)^2; \mathbf{m}\right] = \sigma_j^2(\mathbf{m})$.

Further, note that the equilibrium information Ω_j of the leader may be represented as any vector in $\{0,1\}^{d_j(\mathbf{m})+1}$. Letting l be the number of digits equal to one in any such vector, we obtain

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

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

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

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

we obtain:

$$\begin{aligned} E\left[\left(E\left[\theta|\,\Omega_j\right] - \theta\right)^2; \mathbf{m}\right] &= \frac{1}{d_j(\mathbf{m}) + 2}\left[\sum_{l=0}^{d_j(\mathbf{m})+1} V\left(\theta|l, d_j(\mathbf{m}) + 1\right)\right] \\ &= \sum_{l=0}^{d_j(\mathbf{m})+1} \frac{(l+1)\,(d_j(\mathbf{m}) - l + 2)}{(d_j(\mathbf{m}) + 2)\,(d_j(\mathbf{m}) + 3)^2\,\left(d_{a(k)}(\mathbf{m}) + 4\right)} \\ &= \frac{1}{6(d_j(\mathbf{m}) + 3)}. \end{aligned}$$

■

Proof of Proposition 4. 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 who belong to $N_j(b)$ and have biases b_i to the left of b_j is the same as the number of politicians i who belong to $N_j(b)$ and have biases b_i to the right of b_j. 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. Note now that the equation (3) can be written as:

$$\phi(j,d) = N_j\left(\frac{1}{2(d+3)}\right) - d = 0,$$

and that $\phi(j,d)$ decreases in d because $N_j(b)$ weakly increases in b and $\frac{1}{2(d+3)}$ decreases in d. Hence, the integer d which solves $\phi(j,d) = 0$ is maximal for the index(es) j which 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 competence, and should always be selected as group leader. ■

Analysis of the 5 Player Case in Section 6, Proof of Lemma ?? and of Proposition 5. We calculate all the parameter regions in which $d_3^* > d_2^*$. $d_3^* = 0$ if $\beta > 1/8$ and $\gamma > 1/8$; so that $d_2^* \leq 1$ as 3 will never be truthful to 2. Specifically, $d_2^* = 1$ if $\alpha \leq 1/8$. $d_3^* = 1$ if $\beta \leq 1/8$ and $\gamma > 1/10$; so that $d_2^* \leq 2$ as 4 will never be truthful to 2. Specificall, $d_2^* = 2$ if $\alpha \leq 1/10$ and $\beta \leq 1/10$. $d_3^* = 1$ if $\beta > 1/10$ and $\gamma \leq 1/8$; so that $d_2^* \leq 1$ as 3 will never be truth ful to 2. $d_3^* = 2$ if $\beta \leq 1/10$ and $\gamma \leq 1/10$ but $\alpha + \beta > 1/12$ and $\gamma + \delta > 1/12$; so that $d_2^* \leq 3$ as 5 will never be truthful to 2. Specifically, $d_2^* = 3$ if $\beta + \gamma \leq 1/12$ and $\alpha \leq 1/12$. $d_3^* = 3$ if $\alpha + \beta \leq 1/12$ and $\gamma \leq 1/12$ but $\gamma + \delta > 1/14$; so that $d_2^* \leq 3$ as 5 will never be truthful to 2. $d_3^* = 3$ if $\alpha + \beta > 1/14$

but $\beta \leq 1/12$ and $\gamma + \delta \leq 1/12$; so that $d_2^* \leq 4$. Specifically, $d_2^* = 4$ if $\beta + \gamma + \delta \leq 1/16$ and $\alpha \leq 1/16$.

Using expression (4), we can calculate the aggregate expected payoffs for selecting either politician 2 or 3 as the leader:

$$\begin{aligned} W^*(2) &= -\alpha^2 - \beta^2 - (\beta+\gamma)^2 - (\beta+\gamma+\delta)^2 - 5\frac{1}{6(2+3)}, \\ W^*(3) &= -(\alpha+\beta)^2 - \beta^2 - \gamma^2 - (\gamma+\delta)^2 - 5\frac{1}{6(1+3)}. \end{aligned}$$

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

$$W^*(2) - W^*(3) = \frac{1 - 24\beta^2 - 48\beta\phi}{24} > 0 \text{ or } \beta < \tau(\phi) \equiv \frac{\sqrt{6}}{12}\sqrt{24\phi^2 + 1} - \phi$$

where $\phi = \delta - \alpha + 2\gamma > 1/10$. It is easy to verify that the threshold $\tau(\phi)$ is strictly decreasing in ϕ, with $\tau(1/10) \approx 0.1273$, that $\tau(\phi)$ is strictly positive for any ϕ and equals zero only in the limit as ϕ approaches infinity.

In sum, we conclude that, whenever β is sufficiently small — i.e., smaller than $1/10$ and than $\tau(\phi)$, $\alpha \leq 1/10$ and $\gamma > 1/10$, 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 is more competent, as it can count on two ideologically close trustworthy associates, whereas 3 has only one; and it is not too much more extremist than 3, as β is small.

It is interesting to compare this situation with the equidistant case in which $b_{i+1} - b_i$ is constant for all $i = 1, ..., 4$. The simplest way to morph the equidistant case into the case in which 2 should be selected as leader is to start from the equidistant bias situation in which $b_{i+1} - b_i = \beta \leq \tau(2\beta)$, i.e., $\beta \leq \sqrt{30}/60 \approx 0.0913$ and that the centre-right politician 4 extremizes away from the median, so as to increase $b_4 - b_3 = \gamma$ beyond $1/10$. Paradoxically, by doing so, she will make the final optimal decision move towards the opposite ideological direction, as the centre-left politician 2 will become more competent than the median politician 3.

Another way to morph the equidistant bias case into the situation in which 2 is the optimal leader is as follows. Suppose that, initially $b_{i+1} - b_i = \gamma > 1/10$. Suppose that the leftist politicians 1 and 2 moderate their views, so that $b_3 - b_2$ becomes smaller than $\tau(\phi)$ and α becomes smaller than $1/10$. As a result, they manage to move the optimal group policy towards their views, by making the centre-left politician 2 the leader, in lieu of the median politician 3. Putting together these two ideology morphisms, we uncover the value of moderation in this example. Moving closer to median may turn policy in the politicians' ideological direction, whereas moving far from the median may turn policy in the opposite ideological direction.

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 (**??**):

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

the median politician 3 will grant leadership to player 2 whenever

$$U_3^*(2) - U_3^*(3) = \frac{1-120\beta^2}{120} > 0 \text{ or } \beta < \frac{\sqrt{30}}{60}.$$

Hence, we obtain that, whenever β is smaller than $\sqrt{30}/60 \approx 0.0913$, $\alpha \leq 1/10$ and $\gamma > 1/8$, the median politician 3 will prefer to delegate leadership to the centre-left politician 2, instead of retaining it for herself. In light of Proposition 3, we then conclude that politician 2 is the Condorcet winner of the election game, when $\beta \leq \sqrt{30}/60$, $\alpha \leq 1/10$ and $\gamma > 1/8$. Again, this is because 2 is more competent, as it 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. And again, this situation can be morphed from the equidistant bias case by assuming that the centre-right politician 4 extremizes away from the median, so as to increase $b_4 - b_3 = \gamma$ beyond $1/10$. Paradoxically, by doing so, she will hurt herself: The centre-left politician 2 becomes more competent than the median politician 3; and defeats 3 in the election game. As a result, the group's implemented policy $\hat{y}$ moves to the left.

Having concluded that the parameter β, the ideological difference between 2 and 3 is crucial in determining who should be selected, or will be elected as the leader, it is interesting to compare election and selection of the leader. Because $\tau(\phi)$ is strictly decreasing in ϕ, $\tau(1/10) > 1/10$ and $\tau(\phi) \to 0$ as $\phi \to \infty$, it is immediate to see that there is a unique threshold $\bar{\phi} > 1/10$ such that $\tau(\phi) > \sqrt{30}/60$ for all $\phi < \bar{\phi}$ and $\tau(\phi) < \sqrt{30}/60$ for all $\phi > \bar{\phi}$.

This result implies that, whenever $\phi < \bar{\phi}$, there is an interval of the parameter β 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. The result is intuitive: when ϕ is small so that δ and 2γ are too large relative to α, the ideological loss borne by the right-wing players 4 and 5 for switching leadership from the median politician 3 to the centre-left politician 2 is not too large relative to the gain by extreme-left politician 1. This makes selecting 2 as the leader more likely optimal in the aggregate sense. As the median politician 3 does not care about the other players payoffs when deciding whether to delegate to 2 or not, she may wind up suboptimally retaining leadership for herself.

However, a surprising result occurs when $\phi > \bar{\phi}$, so that δ and 2γ are sufficiently large relative to α. For values of β larger than $\tau(\phi)$ but smaller than $\sqrt{30}/60$, the Condorcet winner is the centre-left politician 2 despite the fact that optimal leader is the median politician 3. The intuition is analogous to the case $\phi < \bar{\phi}$, although this time, when player 3 disregards the other players' payoffs, she downplays the prerogatives of players 4 and 5, instead of the ones of player 1. But the result is nevertheless striking. In the election game, the median politician 3 single-handedly delegates leadership to the 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 ??, Proof of Proposition 6. 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 $b_4 = \beta/2$. The leftist biased politician belong to party A and the rightwing politicians to party B. The values of m_A and m_B are immaterial. Unless politicians 2 and 5 can count of more trustworthy advisers than 3 and 4, the latter will be elected in the primaries, and tie the general election, in equilibrium. Because of symmetry of $\mathbf{b}$, let me now just focus on the challenge between 2 and 3. Because 3 can rely on 2, if 3 communicates to 2 in equilibrium, it follows that 2's only fighting chance to be more competent than 3 is that $d_2^* = 2$ and $d_3^* = 1$, which requires that $\beta \leq 1/10$ and that $2\beta > 1/10$.

The median voter has bias zero, and decides the general election. Her utility for electing candidate j is:

$$U_0(j) = -b_j^2 - \frac{1}{6(d_j^* + 3)}.$$

Because of symmetry of $\mathbf{b}$, if $U_0(2) > U_0(3)$, then there cannot be an equilibrium in which party A elects 3 as its candidate in the general election; if they did, in fact, party B would elect 5 as their candidate and win the election. In fact, when $U_0(2) > U_0(3)$, the unique equilibrium of the game has candidates 2 and 5 win the primaries and 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 $\beta < \sqrt{15}/60$, we conclude that when $1/20 < \beta < \sqrt{15}/60 \approx 0.0645$, the winners of the general election are not the most moderate politicians 3 and 4, despite the fact that politicians' ideologies are evenly distributed in the ideological spectrum. ■

Analysis of the 5 Player Example in Section ?? and Proof of Proposition 7 Suppose that there are 5 politicians, with ideologies $b_1 = -\alpha$, $b_2 = -\beta$, $b_3 = \gamma$, $b_4 = \gamma + \delta$ and $b_5 = \gamma + 2\delta$. Again, leftist politicians belong to party A and rightwing ones belong to party B. We assume that $\gamma < \beta$, so that party B is not only majoritarian in the sense that it has more politicians who can run in the general election, but also in the sense that it can express a candidate, player 3, with views closer to the median voter in the general election. In this sense, we say that party B is ideologically majoritarian in the partition of voters in the general election. One way to conceptualize this idea is noting that the mid-point $M = (-\beta + \gamma)/2$ between the ideologies of the marginal politicians in parties A and B is to the left of the median voter. Hence, if there were no possibility of communicating private information to the elected policy-holder, party B would always win the election by selecting the most moderate politician, player 3.

However, the least moderate politician 2 has a fighting chance as gaining the vote of the median voter in the general election if she is more competent than politician 3. Evidently, this may only occur if $d_2^* = 1 > d_3^* = 0$ as there is only another informed politician in party A. This situation requires that $\alpha - \beta \leq 1/8$, whereas $\delta > 1/8$, so that party A is more ideologically cohesive, and can express candidates who are more competent than the candidates available to party B, in the sense that A's candidate can trust the advice of his party companion, whereas B's candidate need to decide on their on. It is then not difficult to find conditions under which the median voter prefers to elect politician 2 than politician 3. It is enough that

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

where $\Delta = \beta^2 - \gamma^2$. Hence, even if politician 3 were very close to the median voter, she may still lose the general election because her party, B, is less ideologically cohesive than her opponent's party, A. This happens despite the fact that A is ideologically minotarian, as long the ideological handicap is not too large. Formally, it is required that $\Delta = \beta^2 - \gamma^2 = (\beta -$

$\gamma)\,(\beta+\gamma) < 1/72$ and this condition can be easily related to the mid-point $M = (-\beta+\gamma)/2 < 0$ not being too far from zero, the median voter's ideal point.

To prove the claim that candidate 2 can lose the election by moving closer to the median voter, suppose that we start from a situation in which $\alpha - \beta$ is close to $1/8$, candidate 2 is barely within range of 1 for be to be truthful to her. If politician 2 moves ideologically closer to the median voter (i.e., β decreases), then the condition $\alpha - \beta \leq 1/8$ will not be satisfied anymore, candidate 2 will lose the informational advantage over 3 and she will lose the election.

To conclude, note that the claim that a candidate can lose the election by moving closer to the median voter can also be proved by making γ larger and reducing δ so that b_4 and b_5 remain constant and b_3 moves closer to b_4 thereby making it possible to gather its truthful advice. ■