Leadership with Trustworthy Associates

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

April 2017.

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

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

> E la prima coniettura che si fa del cervello d'uno signore, é vedere li uomini che lui ha d'intorno.
>
> Niccolo' Machiavelli, Il Principe, Ch. 22.

1. Introduction

Who should rule? Which individual characteristics are required for good and successful leadership? These questions are central to political writing and thought. They are addressed in Plato's Republic and, perhaps most famously, in Machiavelli's masterpiece *Il Principe*. A central contention of Machiavelli is that good governance stems from the characteristics of a ruler and that these determine political success. This view is shared by contemporary political scientists. Since Stokes (1963), studies have recognized the importance to electoral success of a candidates' valence– a term used to describe competence, talent, good judgement, or honesty, that are generally desirable qualities of a leader.[3]

These leadership characteristics might be seen as innate. According to the social historian Thomas Carlyle leaders (or heroes in his terminology) were made rather than born. Alternatively such characteristics might be the product of education and training. Plato, for example, believed the education of future leaders to be a core function of the state. There is, however, a different view. Core leadership characteristics may stem from the relations a leader enjoys with others in the governance process and the advise that she obtains from them. We find this perspective in Aristotle who, in Politics III.16, 1287 27-35, argues that advice from friends is central to a leader's judgement:

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

A related theme emerges in Machiavelli's masterpiece, *Il Principe*. In chapter 22, he writes on knowledge acquisition and making use of trusted advisors. There, he famously argues that:

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

One interpretation, according to the idiom "a person is known for the company she seeks," suggests that we infer a leader's quality from the type of person she associates with. Another, perhaps more intriguing interpretation, is that a leader's qualities arise *because* of those she

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

associates with.[4]Machiavelli entertains this second interpretation, attributing the greatness of Pandolfo Petrucci, prince of Siena, to the relationship she enjoyed with her valent minister Antonio da Venafro. Indeed, Machiavelli highlights that Pandolfo's ability as a ruler depended upon the information and good judgement provided by Antonio. [5]

In this paper we develop a novel theory of leadership that relates a leader's judgement to the links she enjoys through association with others in the governance process. These relations can benefit a leader and enhance group welfare. Generally, one might think of intra-group linkages as facilitating financial transactions, the flow of information to a leader, and/or as helping build trust, all of which can benefit the group. Building on the insights of Aristotle, Machiavelli and others, we analyze the relations a leader forms with other group members whose advise she may benefit from. Such advise can help her form better judgement and so take more informed decisions. But, for this to be so, a leader must be able to trust the advise obtained. That is, the advise must be truthful.

To explore this notion of leadership we analyze a group that collectively chooses a leader who is granted authority to take a decision on their behalf. Players' payoffs depend on an uncertain state of the world about which each is independently, privately, and imperfectly informed. Each, however, has different preferred outcomes that reflect their idiosyncratic ideological preferences. After the leader is chosen, but before the decision is taken, group members may advise the leader, whoever she or she may be. Such advise takes the form of cheap-talk communication.

First, we study the endogenous formation of a leader's network of trustworthy associates: those the leader can rely upon for truthful advise. The results are intuitive. We show that a leader can rely on truthful advice only from those whose ideological preference is similar to her own. Taking the next step we show that a leader's judgement depends upon the number of trustworthy associates that she has. A larger group of such associates translates into more informed decisions. This intuitive result establishes our take on the Machiavellian lesson: A leader's wisdom and judgement are determined by those she has around her. And it resonates with Aristotle's claim: a good leader has many friends, who are ideologically similar to her, and whose advise she benefits from.[6] Thus we find that the support of friends, that is of course

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

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

[6]There are likely many reasons why politicians are more effective when assisting a leader with whom they share similar views. Our model of advice may be taken as a partial microfoundation for this stylized fact.

necessary in order to defeat opponents in the quest for leadership, also provides a source for better and more effective leadership.

In light of these results, we then ask: what are the characteristics of a good leader? She is defined as the one that the group should choose when maximizing their joint welfare. In line with the classic texts we find that moderation (or temperance) is desirable. Nevertheless, a good leader relies on her judgement that is determined by the number of allies in her circle. Depending on the distribution of ideological views in the group, a moderate leader may be isolated in that that she cannot rely on anyone's advice. Thus a tradeoff arises between moderation, on the one hand, and judgement on the other.

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

What then are the characteristics of the chosen leader? Although individual preferences in our model are not single-peaked, so we cannot use a straightforward application of Black's Theorem, majority choice is determined by the median politician. This follows from the fact that the group preference satisfies a single-crossing condition. Since the majority choice is determined by the median player, one might expect a moderate leader to be chosen. On the other hand, a leaders network is also an important consideration. Indeed we find, in line with common intuition and the practical wisdom of Aristotle, that a leader may be chosen because she has many friends. Here, these friends act as a leader's trustworthy associates (or "mouths and ears" in Aristotle's terminology). In line with Aristotle's view, a large network of friends translates into better judgement. A direct prediction stems from this. The group leader may in fact be be relatively extreme. Indeed, this is so when ideological opinions are concentrated at the extremes.

Next, we compare the group's choice of leader under majority rule with the optimal choice. In doing so we address an issue raised by Levi and Ahlquist (2011) in their survey of the literature. They note that "leadership does not always improve aggregate welfare and we need to know more about the conditions under which it does and it does not." On this point, we show that the chosen leader maximizes aggregate welfare when the ideology of group members is clustered around the median. However, when ideologies are more dispersed then leadership fails to improve aggregate welfare. More precisely, a social planner would have that the group choose a different leader.

Perhaps it is unsurprising that the group may choose the wrong leader from a welfare perspective. It is of course well known that majority rule can lead to suboptimal outcomes. Deeper insights emerge once we recognize our model as one of implicit (strategic) delegation, as first analyzed by Schelling (1960). The incentive of the median politician to delegate arises when another politician has more trustworthy associates and so will take a more informed decision. When choosing whether to delegate, however, the median considers only her own preferences. Moreover, the weight she places on leadership attributes (moderation and judgement) may differ from the optimal ones. A surprising consequence is that she may delegate to a relatively extreme leader when a more moderate one (such as herself) would better serve the group interest. The upshot of this result is a reversal of the famous "ally principle,' which states that delegation should take place to an ally who is as close as possible to the principal.[7] A surprising implication, from a welfare perspective, is that the group choice places *too much* emphasis on a leader's judgment, and too little on her moderation.

A useful exercise is to consider how the group's choice of leader changes when the preferences of its members change. This might occur due to transition in the group's membership or from exogenous shocks to members' preferences. In models of collective choice that build on Black's celebrated theorem only the identity (and hence opinion) of the median (player or committee member) matters for decisions made. Consequently, any change to the distribution of views within a group that leaves the identity of the median unchanged has no effect on policy outcomes. Although Black's theorem holds with respect to leadership choice, our comparative static analysis produces sharply different predictions to those of that canonical model. We find that changes in the ideological views of group members can affect leadership and hence policy choice *even though* they do not alter the identity of the "global" median player. This result stems from the fact that such ideological change affects the size of a leader's network of trustworthy associates. If some players become more extreme (moderate) in their views then a moderate (extreme) leader may loose important allies and can no longer benefit from their advise. This affects her ability to exercise good judgement and hence her prospects of being chosen.

These new theoretical results shed light on empirical questions arising from applications of the spatial model in political science. Consider, for example, work analyzing appointments made by the President to the Supreme Court that are approved by the Senate (Krehbiel, 2007; Rohde and Shepsle, 2007). These models of complete information, based on Blacks' theorem, assess the sequentially rational behaviour of senate members who anticipate the policy impact of such appointments. Specifically, a rational senate member considers whether a proposed appointee changes the identity of the median court member. She is the critical player, since, "an opinion must gain the assent of four justices, the median justice and four justices on one side or another" (Rohde and Shepsle, 2007). This reasoning does not sit well with common intuition that the viewpoints of all players are relevant to decision-making. Moreover, empirical evidence on Supreme Court appointments is in line with this intuition. The specific ideology of the President's nominee matters, in fact, and in ways that confound spatial reasoning. The evidence shows that extreme justices are less likely to have their nomination

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

confirmed. Clark (2012) reviews this literature and notes that the facts are difficult to reconcile with existing theoretical models. He argues that an explanation requires relaxing the complete information assumption that underpins those models. Our analysis suggests that his conjecture is correct. Indeed our model of decision-making in small groups with incomplete information supports the common intuition that the viewpoints of all group members are relevant to decisions that are made. Moreover, and as a consequence, were an external body (such as a Senate) to consider strategic appointments to a group (such as a court), then the set of nominees rejected would be larger than that predicted by a model in which only the median matters.

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

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

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

2. Our Contribution to the Related Literature

While we shall comment on and discuss our contributions throughout, here we precede our analysis by briefly pointing out some of the main related literature and themes. In developing our theory of leadership where a leader is connected via a network, 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). [8] Canes-Wrone, Herron, and Shotts (2001) draw a distinction between "leadership"–the act of implementing a policy that a leader believes to be correct– and "pandering" to a majority. Relatedly, Canes-Wrone (2006) develops a notion of "transformative leadership": in the context of an agenda-setting model, a leader (the President) strategically chooses whether to bring an issue to to the public's attention anticipating that (the pivotal) member of Congress will move toward the public's position. In our She must listen to others in order to reach more informed decisions.

We study verbal (cheap talk) communication between privately-informed participants who provide advise to a leader anticipating that such advise may affect her decisions. Our insights are developed within the context of multi-player communication between imperfectly informed players as studied by Galeotti, Ghiglino, and Squintani (2013). There are numerous applications of multi-player communication in the political science literature: Patty and Penn (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 (2016) analyze the formation of party factions. Our contribution is in developing the multi player communication model to deliver a large set of distinctive findings on leadership and extending these in the context of party competition in which voters anticipate (multi-player) communication within parties.

Several papers share our focus on characteristics that make a leader desirable.Dewan and Myatt (2008); Warren (2012) contrast a leader's judgement with her ability to communicate

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

clearly. Bolton, Brunnermeier and Velkamp (2010) highlight the role of a leader's "overconfidence". Egorov and Sonin (2010) focus on the tradeoff between competence and loyalty to the leader. Besley and Reynal-Querol (2011) show that democratically elected leaders are more likely to have higher academic credentials than unelected ones. Relatedly, Galasso and Nannicini (2011) view talented leaders as a scarce resource and analyze party allocation of competent politicians, proxied by their education level, across electoral districts. We draw a distinction between a leader's judgement and her moderation. A key contribution here is in studying leader characteristics that are derived from first principles.

As mentioned in our introductory notes, our model can be seen as one of implicit strategic delegation initiated by Schelling (1960) to which recent more contributions include Harstad (2010) and Chari, Jones, and Marimon (1997). The question we ask is when and why a political leader would confer decision making authority to specialized or better informed bureaucrats (see Huber and Shipan 2002, for a review). According to perceived wisdom, a political principal would prefer delegating to a bureaucrat with views that are the most similar to her own. The logic behind this so called ally principle has recently been challenged. Bendor and Meirowitz (2004) identify a trade-off between a bureaucrats information and ideological proximity as a reason for the its failure.[9] Our work advances this insight in noting that while politicians may delegate to bureaucrats with a mandate limited to policy implementation, they may also delegate the act of decision-making to other politicians due to the fact that they are better informed. The failure of the ally principle that we identify– delegation to an ideologically distant politician even in situations where this is detrimental to the group as a whole–is based solely on observables (the ideologies of political actors) that initiated its formulation.

Finally our model relates to a large literature on candidate valence defined as candidate's characteristics that benefit all voters regardless of their ideology. Many formal theoretical models have analyzed the implication of valence on candidate policies and electoral outcomes (Ansolabehere and Snyder, 2000; Groseclose, 2001; Aragones and Palfrey, 2002; Callander and Wilkie, 2007; Aragones and Palfrey, 2002; Bernhardt, Camara, and Squintani, 2011). To our knowledge we provide the first derivation from first principles of electoral candidate's valence, in the form of good judgement. In the standard definition of valence, it is independent of ideology. Here, in our microfoundation, a leadership candidate's valence is related to, and partly determined by, the ideological distribution of politicians in her group.

[9]The analysis of Bendor and Meirowitz (2004) explores a large number of features that may shape the form and extent of delegation, including policy uncertainty, risk aversion by the political leader, costly information acquisition by agents, ex-ante control, ex-post auditing, ex-post control, multiple political principals By and large, the analysis shows that the ally principle and other assertions from the informal literature on delegation may or may not hold, or hold only with important qualifications, depending on detailed game form or preference assumptions that represent different institutions or instances of delegation. (See, also Epstein and Halloran, 1999, Gilligan and Krehbiel, 1989, 1990, and Huber and McCarthy, 2004).

3. Model

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

Our players are a group of politicians $N = \{1, ..., n\}$ who are faced with a decision $\hat{y} \in \mathbb{R}$. One amongst them– a leader–makes the decision on the group's behalf. The utility of each politician i depends on how well $\hat{y}$ matches an unknown state of the world θ. Politicians are ideologically differentiated and so the utility of i depends also on her ideological bias b_i. Bringing these elements together in a familiar quadratic loss form, we suppose that, were she to know θ, politician i's payoff $u_i(\hat{y}, \theta)$ would be a function of 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 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.[10]

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:

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

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

For a given 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" (their message is not informative of their signal) is always an equilibrium. Because of equilibrium multiplicity, the ranking of leaders and the leadership selection depend upon the choice of equilibrium: for the same leader j, different equilibria yield different player payoffs. To avoid ambiguities, we assume that for a given leader j, politicians coordinate on the equilibria $(\mathbf{m}, \mathbf{y})$ that provides the highest expected payoffs to all politicians.[11] The selection of these equilibria is standard in games of communication and allows us to focus attention on leadership selection.[12]

We consider two forms of leader selection.

The first one addresses our normative question: which leader would maximize politicians' welfare if chosen? Following the utilitarian principle, 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.

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,

[11] 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).

[12] Extensions could consider equilibria in which a politician threatens to babble (thus not transmit information to the elected leader) as a way to force her own election as leader.

namely the ordering of ideological biases within the group. Therefore we can relate a leader's judgement to the same ordering.

4.1. **A Leader's Judgement.** The equilibrium 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 the 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.[13] We use this result to derive how informed politician j would be in the event where she becomes leader.

First we note that the term $d_j(\mathbf{m})$ 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 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

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

[13] 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 be unable to communicate truthfully a signal $s_i = 0$. By contrast, when j has a small number of players communicating with her, then i's report $\hat{m}_{ij} = 1$ moves j's action significantly to the right, and 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.

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*

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

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

5. Selecting the Leader

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

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:

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

Expression (4) decomposes the welfare function into two elements: the aggregate ideological loss $\sum_{i\in N}(b_i - b_j)^2$ associated with the decision taken by j, and the aggregate residual variance of her decision $n[6(d_j^*+3)]^{-1}$. 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.[14] 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 $\mathbf{b} = \{b_1, ..., b_n\}$ in the group. In fact, every element of $\mathbf{b}$, even extreme ones, matters for the determination of the average ideology $\sum_{i\in N} b_i/n$. 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 she has around her"*. This analysis of the role played by the local ideological distribution is, to our knowledge, novel in the large contemporary and formal literature on collective choice; though it echoes the insights of Machiavelli made in 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: leader j's policy should reflect the diversity of views in the group. Optimal leadership also requires judgement. This stems from the information that 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 a simple majority decision taken within the group. Each player i's utility as a function of the leader's identity j is:

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

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

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

leadership choice under majority rule is far from straightforward, we can, nevertheless, make progress by establishing 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 appeal to that by Gans and Smart (1996) to show that the player with median ideology will determine the outcome of the election. The unique Condorcet winner of the election game is the politician j who maximizes the expected payoff of the median player.

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

6. What Makes a Good Leader?

Our analysis relates the characteristics that define good leadership -moderation and judgement- to the communication structure that emerges in the equilibrium of our model. The importance of the former is well known. Indeed it is easy to see that if there were no informative signals (or just no communication) in this game, then the chosen leader would be the median politician $m = (n+1)/2$, 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 $\mathbf{b}$), is the most moderate politician the optimal leader and the majority-preferred one. Evidently, this is the case when there are only 3 politicians in the group. For then the median is the most moderate politician and she cannot be less informed than either of the others: If she is willing to communicate truthfully with her neighbors, then at least one of them is willing to be truthful to her.[15]

Moving beyond the three-player case, we illustrate sufficient conditions such that the optimal leader is also the most moderate politician. Doing so, we consider the situation in which politicians are distributed at even distances with respect to their ideology on the line. Because n is odd, and by symmetry of the ideological distribution, the median $m = (n+1)/2$ is the most moderate and has at least as many trustworthy advisers as any other politician. Then there is no tradeoff between a leader's moderation and her judgement. As the median politician is as informed as anyone else she should take the decision on behalf of the group and, indeed, she would be the unique choice of the majority.

We formalize this insight in the following proposition that proves a stronger result. We show that the median politician m is elected by the majority as leader and is also the optimal leader when the ideology distribution is symmetric around m and 'single peaked' at m, in the sense that politicians are more ideologically clustered as they get closer to m. Formally, we define the ideology distribution $\mathbf{b}$ as 'single peaked' and symmetric at m when for any $i = 1, ..m-1$, $b_{i+1} - b_i$ weakly increases in i, and $b_{i+1} - b_i = b_{n-i+1} - b_{n-i}$.[16]

Proposition 4. *When politicians' ideologies* $\mathbf{b}$ *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.

Proposition 4 relates the core characteristics of leadership to reveal that, with single-peaked and symmetric biases (as in the case of equidistant biases depicted in Figure 1.) the most moderate politician is also (weakly) the better informed and so has better judgement. The

[15]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.

[16]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 the definition of $\mathbf{b}$ as single peaked and symmetric at m.

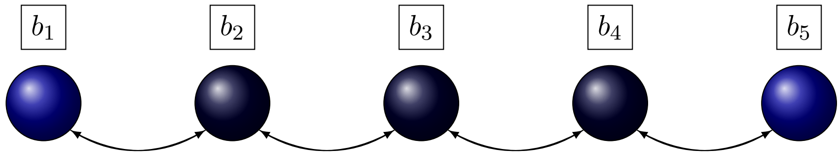

FIGURE 1. Leadership Choice with Equidistant Bias: this figure depicts the case where $b_i - b_{i-1} = \beta$ for all politicians $i = 2, ..5$ and $1/12 < \beta \leq 1/10$. An arrow linking two politicians illustrates that truthful communication is sustained between them.

significance of our result lies in the fact that, when ideologies in the group are evenly distributed, there is no tradeoff between moderation and the capacity to gather reliable advice. This underlines that a necessary condition for a politician who is not the most moderate to be the optimal leader or the elected one is that ideologies are not evenly distributed, but instead are clustered in a nonhomogenous way.[17] We explore this tradeoff next, focussing on the case of 5 politicians for ease of exposition.

6.2. **The Case with 5 Politicians.** To explore the tradeoff between moderation and judgement we study in-depth the case of 5 politicians, where, without loss of generality, we denote the ideologies as $b_1 = -(\alpha + \beta)$, $b_2 = -\beta$, $b_3 = 0$, $b_4 = \gamma$, $b_5 = \gamma + \delta$. Our case is suitably rich (allowing us to identify properties of the ideology distribution that provide novel and interesting findings) yet simple (so that we can do so in a clean and clear manner).

First we note that it can never be optimal that 1 and 5 lead the group and neither will they be elected by the group as leaders. Players 2 and 4 can be chosen as leaders if and only if they have better judgement than player 3. Also, interchanging α with δ and β with γ 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 γ is such that (in contrast to Figure 1) player 4 can no longer communicate truthfully with 3. As the player ideologies are not distributed at even distances there may

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

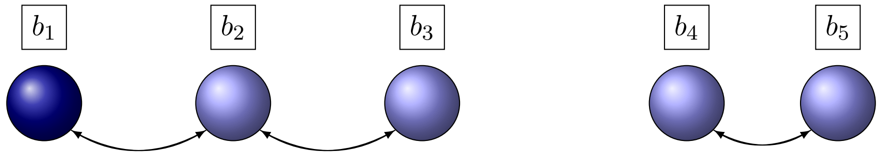

FIGURE 2. Leadership Choice with Non-equidistant Bias: illustrates a case where $b_1 = -(\alpha + \beta)$, $b_2 = -\beta$, $b_3 = 0$, $b_4 = \gamma$ and $b_5 = \beta + \gamma$. An arrow linking two politicians illustrates that truthful communication is sustained between them. Here $\alpha \leq 1/10,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

be a tradeoff between judgement and moderation and, moreover, surprising consequences of ideological shifts (such as that of player 4 in Figure 2 relative to Figure 1) which we discuss below.

6.3. **The Tradeoff between Judgement and Moderation.** Whether the choice between the more moderate politician 3 and politician 2, who has better judgment, is resolved in favor of either depends upon β. This determines how extreme is 2 relative to 3. In fact, we can relate the choice between politician 2 and 3 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 $\tau(\phi)$ which, in turn, depends upon the values of α, δ, and γ. Intuitively, it is optimal that 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 simpler form as it depends only on the median player. she may obtain a more informed outcome when 2 takes the decision and this yields a constant addition to her utility. This comes at an ideological cost β. Thus the group chooses 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 $\frac{\sqrt{30}}{60} < \beta < \tau(\phi)$, the optimal leader is the most informed politician 2, whereas the majority elects politician 3. This identifies an instance in which majority voting leads to a moderate but inefficient choice of leader.

A perhaps more interesting and unexpected case arises when $\tau(\phi) < \frac{\sqrt{30}}{60}$. Again, for small β, the non-moderate politician 2 is both the Condorcet winner and the optimal leader, and for large β it is politician 3. Now inefficiency arises in the intermediate case in which $\tau(\phi) < \beta < \frac{\sqrt{30}}{60}$. Although the optimal leader is the most moderate politician 3, the majority elects instead a relatively extreme leader in politician 2.

The logic behind this result is simple. The median politician may trade off moderation and competence in a way that differs from the optimal choices made by a social planner. Starting from her ideal point, she may sacrifice a policy more in line with her bias for a more informed outcome. In our 5 player example, she will indeed do so when $b_2 - b_3$ is sufficiently small and $d_2 > d_3$. Choosing a leftist leader then benefits the median and of course leftist members of the group. But it harms the right-wing members 4 and 5, who bear costs $(b_4 - b_2)^2$ and $(b_5 - b_2)^2$ respectively. Because the ideological loss function $(b_i - b_j)^2$ is convex in the ideological distance $|b_i - b_j|$, the leadership move from 3 to 2 is more harmful to rightwing politicians than it is beneficial to the leftist ones. Then it may be the case that a social planner would force the median politician to take the decision, if only she could.

6.4. **Comparative Statics.** Further analysis of the five-player case uncovers some interesting comparative statics results: A player changing her ideology from right (left) to left (right) can induce a shift in leadership in the opposite direction.[18]

To illustrate, consider a benchmark case with players ideologies evenly distributed apart and where politicians 2, 3, and 4 can all count on the truthful advice of their ideological neighbors, so that $1/10 < \alpha = \beta = \gamma = \delta \leq 1/8$. 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 rightward and so away from that of the median player and that, as a consequence, they are no longer truthful to one another (i.e., suppose that γ increases so as to become larger than $1/8$). Now politician 3 has lost a (previously) trustworthy associate. It is now possible that the centre-left politician 2 is the Condorcet winner of the election game– 3 will delegate authority to her, despite not being the most moderate politician. Indeed, by Lemma 2, we know that this is the case when $\beta < \frac{\sqrt{30}}{60}$. Hence, the ideological movement of a player towards a more extreme position may induce a leadership change in the opposite direction.

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

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, leadership switches from the median player 3 to the centre-right player 2, again, when $\beta < \frac{\sqrt{30}}{60}$. Here, the rightward ideological movement of leftist players moves leadership choice (and hence policy outcomes) in the opposite direction. Thus moderation allows them to capture control of leadership of the group.

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

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.*

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) control of the group, and turn the group policy closer to (away from) their views.

6.5. **Discussion.** The findings of our positive study of leadership highlight the predictive importance of judgement. The fact that the majority choice may place too much weight on this characteristic is perhaps surprising, the more so when interpreting the decisive median vote in the election as a decision to delegate authority to leaders with specific characteristics. This notion goes back to Schelling (1960) who in his seminal book *The Strategy of Conflict*, discussed the use of delegates with particular characteristics as a way to credibly commit a negotiating party to a position. He suggested that agents in bargaining situations may transfer power to stubborn negotiators.[19] 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.

As noted in our earlier discussion of related literature, when viewing our model as one of implicit delegation, 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.

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

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.

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 set of possible strategies to consider is large. These models are tractable, however, due to the assumption that within the Senate and the Court the pivotal player (politician, judge) is the one with the median preference. Appointments can then be considered with respect to whether or not they change the identity of that player, and, hence, these models go by the description of "move the median" games. 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 Clark (2012) , "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. As already noted in our introductory remarks, in her review of the field, Clarke suggests that relaxing the complete information assumption in standard models may yield new insights. Indeed our analysis would appear to confirm that this is in fact the case.

We postpone a more extensive application of our ideas to these cases to future research. Here, instead, we focus attention on an immediate and we believe first order extension of our model. As already noted, our analysis of group choice of leader provides insights that differ from those provided by a straightforward application of Black's Theorem. A noted application of 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.

7. A Model With Electoral Competition

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

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

7.1. **Model.** Suppose that there are two parties, A and B. The 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 assume only that each party selects as leader the strongest possible candidate, defined as the politician within the party who defeats the largest possible number of candidates from the other party in the general election. To simplify the exposition, we consider only ideology and party profiles such that there is only one such politician in each party.[20] The eventual winner $j \in \{a, b\}$ of the general 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. Unlike in basic citizen-candidate models, the winner of the election makes her decision only after consultation with her party leading politicians.

[20] 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.

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 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 evaluates a candidate j on the basis of both j's ideological proximity $(b_i - b_j)^2$ and 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 can think of judgment as *valence*. Here, it is endogenously determined by j's network of trustworthy party fellows.

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

7.2. **Policy Divergence.** A natural benchmark for comparison is an otherwise identical model in which players 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.

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.

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.*

Whenever politicians are sufficiently ideologically distant from each other that they can never communicate truthfully (even to their closest ideological neighbour) then the winner of the general election is the candidate with bias closest to zero. Beyond this simple case, it is immediate that, in our model of electoral competition, party leaders need not be moderate. The analysis follows our earlier logic: politicians with a large network of truthful informants may be preferred by the median voter even if they have relatively extreme ideologies. In fact we can reveal new insights. First we show that the winner of the general election need not be the most moderate politician (i.e., the politician whose ideology is closest to the median voter), even in circumstances in which the politicians' ideologies are evenly distributed in the ideological spectrum. Thus our finding stands in sharp contrast with Proposition 4. Why so? When politicians are partitioned in competing parties the most moderate politician (with respect to the electorate as a whole) is at the extreme end of the ideological spectrum within her own party. This constrains the pool of trustworthy associates she can rely upon and so hampers her ability to take informed decisions if elected to office.

Proposition 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 (lighter shading) belong to party A and the others (darker shading) to party B.

Following our earlier analysis, unless politicians 2 and 5 can count on more trustworthy advisers than 3 and 4, in equilibrium, the latter will be 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.

This result provides a new take on the documented divergence among candidates in two party elections. Even if two parties compete for power, they will not necessarily the most moderate candidates, so that convergence to the median will not take place. This finding complements earlier explanation for policy divergence that abstract from party competition (e.g., the citizen candidate models by **?** or **?**, or the model by **?** and **?**).

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

Finally, we find that the outcome of the election may depend on the whole ideological distribution and often in a very subtle way. For example, suppose that a party leader becomes more moderate then she may lose the general election as a result of her new found moderation. (Of course, the opposite can happen: a politician may lose the election by becoming more extreme). Intuitively, this occurs because, by becoming more moderate, the party leader increases the ideological distance between herself and others in the party. And , as a consequence, she loses the benefit of the truthful advise they would otherwise have provided her with. This result is, to the best of our knowledge, both novel and unsupported by any variant of the standard spatial model found in the literature.

Proposition 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 though its leader is the candidate in the general election whose views are closest to the median voters.*

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), and lose the general election as a result.

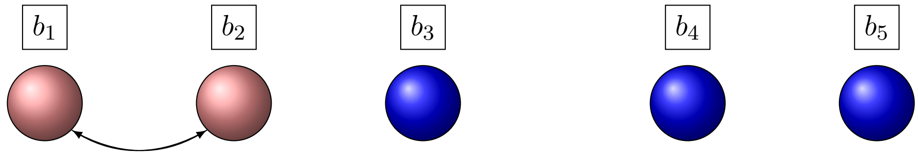

FIGURE 4. Party Competition with a majoritarian party: illustrates a case where politicians 1 and 2 are in party A and $3, 4, 5$ are in party B;$b_1 = -(\alpha + \beta)$, $b_2 = -\beta$, $b_3 = \gamma$, $b_4 = \gamma + \delta$ and $b_5 = \gamma + 2\delta$ and $\gamma < \beta$. The figure depicts within 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 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, there are two senses in which party B is advantaged: there is a larger set of leading politicians from whom the leader could draw upon for information; and in player 3, it has a potential leader whose views are closest to those of the median voter. 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 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 politician 2 from party A to any politician from party B, so that, in equilibrium, B will lose the election despite its advantage. Specifically, we find in the appendix that this is the case if and only if $(\beta - \gamma)(\beta + \gamma) < 1/72$: this condition is satisfied 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 as 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 and so 1 is a trustworthy associate of 2. If politician 2 moves ideologically closer to the median voter, and 1's position remains fixed, then the condition $\alpha \leq 1/8$ will no longer be satisfied. Thus candidate 2's judgement is no longer better than that of politician 3 and so, were they to contest the election, 2 would lose.[21]

[21]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.

7.4. **Discussion.** Our results link the notion of cohesiveness to a party's electoral success and suggests that anything that increases (decreases) cohesion will have positive (negative) electoral consequences[22] These theoretical results are consistent with a common understanding that 'divided parties don't win' (Worcester, Mori Publications 2002) and empirical findings. For example in the UK in three elections previous to Labours election victory in 1997, each of which provided the Conservatives with a majority, Labour were the party perceived to be the most divided. The British General Election Study of 1983 reveals that 24.2% of the electorate perceived the Conservatives to be divided whereas 87.9% perceived Labour to be so. In 1987 the corresponding figures were 42.1% for the Conservatives and 66.9% for Labour, and in 1992, 27.1% for the Conservatives and 62.8% for Labour. In 1997, however, more of the electorate (40.6%) perceived the Conservatives to be divided than perceived Labour to be so (18.9%) (Source, British General Election Studies, Heath, Jowell, Curtice et al 1983, 1987, 1992, 1997).

Our results on the importance of the cohesiveness of parties relate to debates on party size, and how this affects their political viability and effectiveness, that dates back at least to Michels (1911). Our model highlights that the size of party may be less important than the ideological disparity of views within it. Indeed our results suggest that the formation of a leadership clique or oligarchy that closed down the possibility of fruitful internal debate would be damaging.[23]

Our results on the relevance of party cohesiveness and of the choice of a leader that is not estranged from the party are especially relevant to understand the events that led to the victory of Labour victory in the 1997 UK elections. That Labour victory is often related to its leader Tony Blair's moderation and his eschewing of the left-wing policies of predecessors. Our result suggests a different narrative, namely that it was the ideological cohesion of "New Labour" rather than the moderation of its leader that was important. Indeed, a corollary to proposition 7 is that the moderation of its leader is neither a necessary nor sufficient condition for party success. Instead, and in order to assess the chances of electoral victory, one should consider the relationship between the party leader and those of his associates. In his memoirs Blair talks about the team of politicians who advised him and on whom he could rely, amongst them Gordon Brown, John Reid, and David Blunkett. Referring to the latter he states (page 34-35) "his loyalty and commitment to New Labour, I never doubted." Yet whereas Blair himself had always been a moderate and so natural moderniser, Blunkett had moved from the left toward the center. He had been a leader of Sheffield Council, one of Britain's most left leaning councils during the 1980's. His personal ideological change was noted in an article by the Economist in 2001, which described him in the following terms: "a municipal socialist when Thatcherism was rampant, he came to understand the limitations of the old left. This made him a genuine Blairite." Our analysis suggests that the ideological odyssey of Blunkett

[22]Party cohesiveness is studied in McGann (2002) and McGann, Koetzle, and Grofman (2002).

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

(amongst others) that allowed him to become a trustworthy associate of Blair, and that might be seen as a small episode in Labour history, should be viewed as a central component of its electoral success.[24]

8. Conclusion and Discussion

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

We conclude by discussing some useful extensions of our ideas. An interesting question is how our results would change if different politicians had different access to information. It is possible, for example, that ideologically close politicians gather information largely from the same sources. As a result, the advice of associates who are too ideologically close may be less valuable than that provided by more distant ones. Of course, our main result, that associates who are too ideologically distant are not truthful to the leader in equilibrium, would survive in a model in which information is possibly correlated among politicians. As a result, all of our possibility results would extend to this enriched environment. Further, it may prove interesting to determine the optimal composition of advisers to the leader. This subject is the argument of our current research. 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 their recommendations would not be credible, and it is possible that politicians who are too close would be excluded so as not to crowd out more valuable less ideologically close advisors (see discussion in footnote 13.)[25]

[24]Our lessons resonate with British politics today. At the time of writing the Conservative Party, under its new leader Theresa May, enjoys a large lead over its Labour rivals. Both parties are divided, and have recently been involved in bruising leadership contests. Yet within the Conservative Party, even bitter rivals such as Michael Gove and Borris Johnson have, allegedly, agreed "to keep lines of communication open." By contrast, lines of communication between the Labour leader Jeremy Corbyn and her closest allies in the Parliamentary Labour Party are frayed. Chris Bryant, one of several MPs who have called upon Mr Corbyn to quit states: Its a bit of a problem if Jeremy wont even see the seven people in his shadow cabinet who he appointed this week." In the same passage, Bryant goes on to speak of the leader having developed a "bunker mentality"(*Corbyn Has Bunker Mentality Say Challengers*, FT.com, July 3rd, 2016.)

[25]A distinct reason for why leaders may choose advisers with diverse policy preferences is that they may decide to pitch advocates with opposite views to argue about the pros and cons of policy choices with uncertain consequences. Policy makers would then rely on verifiable information disclosed by the advocates, and discard unverifiable political advice. By selecting advocates with opposite views, the best incentives are provided for the advocates to expend effort in investigating the matter and acquiring information. Instead, our paper focuses on

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

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

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

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

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

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

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

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

——— (1997): “The Economics of Split-Ticket Voting in Representative Democracies,” *The American Economic Review*, 87(5), pp. 957–976.

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

DEWAN, T., A. GALEOTTI, C. GHIGLINO, AND F. SQUINTANI (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.

communication of non-verifiable information: the possibility that politicians lie makes transmission reliable only from ideologically close associates.

DEWAN, T., AND F. SQUINTANI (2016): "In Defence of Factions," *American Journal of Political Science*, 60(4).

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 (2013): "Information Transmission Networks," *Journal of Economic Theory*, 148(5), 136–159.

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

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

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

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

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

KREHBIEL, K. (2007): "Supreme Court pointments as a Move the Median Game," *American Journal of Political Science*, 51(2).

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

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

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

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

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

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

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

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

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

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

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

WARREN, P. L. (2012): "Allies and Adversaries: Appointees and Policymaking Under Separation of Powers," *Journal of Law, Economics, Organization*, 28(3), pp. 407–446.

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

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

Hence, politician i's posterior is

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

the expected value is

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

and the variance is

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

■

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

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

Denote

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

Observing that:

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

and simplifying, we get:

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

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:

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

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

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

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

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

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:

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

■

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 2 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 7, 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 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 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 7 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 her 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

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

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 $\Delta = \beta^2 - \gamma^2 = (\beta - \gamma)\left(\beta + \gamma\right) < 1/72$ 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. ■