On Leadership: Coordination and Endogenous Information Acquisition under Internal Disunity

A thesis in accordance with the requirements for the degree of

Bachelor of Science in Economics

Author: S.N.Geesing
Student Number: 342570
Supervisor: Dr. Dana Sisak

Department of Economics
Erasmus School of Economics
Erasmus University Rotterdam

Abstract

We propose a quadratic pay-off Keynesian "beauty-contest" model to analyze leadership and activism within political parties. We allow for endogenous information acquisition as well as heterogeneity with respect to the optimal policy positions that leaders advocate. Party leaders send signals to party activists given a certain inherent noise and may choose to bias the signal. Party activists choose how to divide their attention and which policy position to advocate based on the signals they receive. A special feature of our model is that we consider affine instead of linear strategies for party activists. The model is analyzed for perfectly informed and fully naive activists as well as an intermediate case. Furthermore we allow for costless and costly signal biasing by leaders. We find that for perfectly informed and fully naive activists the information acquisition/weighting process is independent from the biases added by leaders. Attention is paid only to the best communicators, regardless of bias. With perfect information biases are perfectly filtered out, and thus leaders are indifferent between any bias. With full naivety, biases are fully absorbed and the unique equilibrium is a faction dictatorship, independent of the extremity of the dictators' position. With costly biasing the extreme equilibria dissapear and only leaders that share an optimal policy position with the activists consitute an equilibrium. In the case in which naive and informed activists are mixed the biasing process is no longer independent of the information acquisition/weighting process. The unique equilibrium remains one of dictatorship.

Contents

Americans are apt to be unduly interested in discovering what average opinion believes average opinion to be.

– *John Maynard Keynes, General Theory of Employment, Interest, and Money*

Chapter 1

Introduction

In many scenarios of social-scientific interest, decision-makers seek actions that are both matched to some unknown underlying feature of the world and matched to the actions taken by others. As such, agents must balance their own directional preferences on the one hand whilst maintaining coordination on the other hand. A prominent example is found within the field of industrial organization, where a firm has an incentive to set a price close to that of its competitors whilst still reacting optimally to a set of other market conditions. This setting can be applied to a large variety of other settings, such as various organizational settings and, as in this thesis, political parties.

Analyzing such situations is far from trivial. Keynes noted this already in his commentary on the famous beauty contest, a game organized by a popular newspaper at the time. Here contestants were asked to choose the six prettiest faces from a hundred photographs. Those who picked the most popular faces are then eligible for a prize. As Keynes (1936) noted:

"It is not a case of choosing those [faces] that, to the best of one's judgment, are really the prettiest, nor even those that average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees."

Indeed, a serious analysis of the incentives involved in such situations was practically impossible until the development of more advanced game-theoretic tools in the latter half of the 20th century. As a result, the beauty contest model saw a revival within the literature through the late twentieth and early twenty-first century (see the following chapter for an overview). A particularly interesting aspect of such coordination games is the role of information. In order to make an optimal decision, decision-makers require information on the state of the world as well as the actions taken by others. Since the actions taken by others are in turn based on the information they receive, understanding which information sources are used by the decision makers is crucial within this framework. The insight behind this notion is simple: if one is to match one's action to someone else's action, one needs to be aware of the information that is available to the other party and on which information that action is based.

Realizing that this simple insight is crucial, a natural question is to consider which characteristics or features of an information source determine whether or not it is taken into account within the decision process. We are thus concerned with a setting in which specific attention is paid to specific sources and others are ignored. In other words, information is acquired endogenously. In particular, we can model this situation as one in which decision-makers choose to pay attention to a certain signal and then receive that signal with a certain noise that is based on the attention attributed to it. Certain features of the decision-makers and the signals can then be varied to see how the decision process is affected under different conditions.

An interesting interpretation of this type of modelling is the following: instead of considering the signal-senders as just sources of information, consider them to be leaders of the group of decision makers. Within this interpretation, the leaders inform the group about the best action for the group to take, and decision-makers base their decision on the information they receive from such a leader, whilst simultaneously maintaining the coordination motive. Leaders may vary in several dimensions. First of all, some leaders may be more qualitified to make an accurate judgement of the best action, i.e. be a better technocrat. Secondly, leaders may differ in their ability to communicate with the decision-makers. Both these aspects introduce a certain noise in the signal. Decision-makers then choose to pay attention or to ignore certain leaders and choose an optimal action based on the signals they receive.

The interpretation above has been the setting in several recent studies, we refer to the next chapter for an overview of these. In particular, the model has been used to model relations within a political party. Here party activists wish to advocate the optimal policy based on certain outside conditions. For example, they may want to maximize the votes they will receive in an upcoming election by choosing the best economic policy given the current state of the economy. At the same time, it is important that these activists advocate similar policies. Different people from the same party telling a different story is, after all, a surefire way to lose an election. In order to achieve these goals, they listen to the party leaders. Party leaders vary in their ability to communicate and in the ability to determine the correct policy. As such, activists may want to pay less/more attention to certain leaders than others, or ignore some leaders altogether. It is this setting that is considered in this thesis.

Whilst previous studies have allowed leaders to vary with respect to these characteristics, leaders have typically been assumed to be identical in one respect: all leaders send signals about the same policy position. This can imply one of two things: first of all, it can imply that the state of the world can be measured purely objectively. That is, the policy setting can be evaluated independently of any policy preferences. In such a case leaders would all send the same unbiased signal. Secondly however, we see that even in the absence of complete objectivity such a case could occur. This would happen if there was complete unity amongst the party leaders, meaning that there is complete agreement about the (possibly subjective) optimal policy decision. Clearly, these two conditions rarely hold in reality. Often political values are not objectively measurable and are subjective to some degree at least. This particularly applies to issues related to social justice and economic fairness. At the same time, it is highly unlikely that all leaders within a party agree completely on the policy to be pursued, especially in electoral systems that are characterized by few parties. A good example of this is the United States, in which a wide range of views can be found amongst both Democrats and Republicans. In systems with more parties (such as those commonly found in mainland Europe) such extreme disparities within parties are less visible, but not uncommon. A good example is found within economically right-wing parties, which are often split between conservative and liberal factions.

This observation motivates this thesis, in which we will focus on internal disunity. More precisely, leaders receive information about the state of the world, and can then choose whether or not to bias the information. This bias will be determined based on an individual policy preference. At the same time however, it is also in the interest of the party leaders to maintain unity within the party. As such, we introduce individual utility functions for leaders which are decreasing in the variance of the actions taken by the party activists. Party leaders thus have an incentive to balance their own policy preferences against the unity of the party. In this setting the role of information is clearly vital. In particular, the extent to which party activists are aware of the biases of individual leaders is likely to greatly affect the outcome of the analysis. As such we will consider several scenarios with respect to the information available, and derive equilibrium conditions for each of these settings. In particular, we will consider perfectly informed and completely naive activists, as well as mixed case in which we have both informed and naive activists. Analyzing this model allows us to draw interesting conclusions about how leaders and activists deal with disunity within the leadership and under what conditions leaders with extreme positions may be able to gain influence. The interaction of individual bias and individual clarity in particular will be interesting to investigate. Clearly, the conclusions of this analysis generalize readily to other organizational settings and many of the various examples given above and in the following chapter. As such, the proposed extension promises to provide an interesting addition to the existing economic literature on (political) leadership.

The structure of this thesis will be as follows: first an overview of the existing literature will be provided in the next chapter. After this, Chapter 3 will consider the model in detail, and provide an outline of the basic informational structure maintained throughout this thesis. Chapter 4 will comprise the bulk of this paper and will consider solutions and equilibrium conditions for the model within various settings with different informational assumptions. We will also present some simple comparative statics and graphical analyses. Finally, we provide a conclusion and a discussion of our results.

Chapter 2

Literature Review

The significance of coordination as a motive within social and organizational settings has not escaped economists over the last decades. Keynes applied this setting to the modelling of prices in equity markets (Keynes, 1936). His description was based on a contest which was popular in US newspapers around the time and it is here that the commonly used 'beauty contest' terminology has its origin. Over the last decades, a broad literature has developed surrounding varying applications of this setting. An especially interesting literature has developed surrounding the "shared knowledge" notion of corporate culture and the importance herein of "doing the right thing together" (Cremer, 1990, 1993). Typically a team-theoretic framework was applied here. A common finding was the "trade-off faced by a firm between accumulating diversified knowledge about the environment and providing common ground for decisions" (Marschak & Radner, 1972). This trade-off is cleary analogous to the coordination motive as previously described.

More recently interest has been renewed following a landmark paper by Morris and Shin (2002). The main features of the setting have since typically been captured in a class of quadratic pay-off beauty contest games. Since then these games have been applied to a range of fields. An example of this is investment games (Angeletos & Pavan, 2004). Applications were also found within the field of industrial organization, and to monopolistic competition in particular (Hellwig, 2005). Furthermore we have seen applications to financial markets (Allen, Morris, & Shin, 2006) and a wide range of other economic problems (Angeletos & Pavan, 2007). This type of modelling even managed to pierce the barrier between micro and macroeconomics and has been applied to study monetary policy, see for example Amato, Morris & Shin (2002). A particularly creative application is political leadership as in Dewan & Myatt (2008). This last paper in particular motivates much of this thesis.

In the above papers (with the exception of Dewan and Myatt) information is acquired exogenously. That is, signals, be they public or private, are received as they are sent. Several more recent papers have allowed for the acquisition of information to be endogenous. A notable example is Hellwig and Veldkamp (2009). They particularly focus on the existence of equilibria under both private and public information. Their main conclusion is that allowing for public information opens up for multiple equilibria, since public signals inform players both about the underlying state of the world and the likely actions of other players. Private information does not have this second characteristic, thus removing a key ingredient for multiple equilibria. It is of note that in this paper, players choose whether or not to pay to receive a signal. Alternatively, in Dewan and Myatt (2008) they divide their time continuously between sampling different information sources.

Another approach that deserves separate attention comes from the rational inattention literature (see e.g. Sims, (Sims, 1998, 2003, 2005, 2006)). Here players face information-processing constraints. This implies that there is a limited amount of information that can be communicated and absorbed. As such, agents face a constant trade-off: information absorbed from one source prohibits the agent from absorbing information from another. Of particular interest here is the balance between public and private information and the impact of aggregate and idiosyncratic shocks that feed through these sources (Mackowiak and Wiederholt, 2009). The constraints here are set by the so-called "Shannon capacity". This concept hails from information theory and is related to the minimum bandwidth required for succesful communication, assuming appropriately coded messages (MacKay, 2003, Cover and Thomas, 2006).

These three approaches have been unified in Myatt and Wallace (2011) in which it is shown that these three cases simply correspond to different cost-structure specifications. They find that that the nature of equilibrium within endogenous information acquisition in coordination games turns upon the nature of the cost function which players face. It has to be noted that this paper is in many ways a companion to the Dewan and Myatt paper, and some of its conclusions are very similar. For example, both papers indicate that the relative noise of signal crucially determines whether that signal is paid attention to. Furthermore they both find that the clarity of the signal (that is, the noise introduced in sending the signal from sender to receiver) is more important in determining whether or not a signal is paid attention to than the noise with which the sender receives information about the state of the world (speaking in the terminology of Dewan and Myatt, the communication skill of the leader is more important than whether or not the leader is a good technocrat).

By interpreting signal senders as being advisors rather than leaders, we break into the wider of field of decision-maker/advisor or DM-A models. Here decision-makers have access to an advisor and have to decide on the implementation of certain decision based on the advise of these advisors. It is interesting to note that within this literature it is common to focus on the preferences of the advisor relative to the decision maker rather than issues of clarity and attention. An exception to this is Persico (2004). Interesting findings from this field are many, but we will highlight several of the more important ones here. A crucial one is the "ally principle". It implies that it is typically more efficient to have a decision maker and advisor that share preferences (Swank, Swank & Visser, 2008). Whilst this finding is highly intuitive, others are not. An example of this is the finding that messages sent by advisors that have very different preferences from the decision maker can contain more information than those sent by like-minded advisors (Dominguez-Martinez, Swank & Visser, 2008). The logic here is that only under a very limited number of circumstances an opposed advisor will be in agreement with the decision maker, whilst like-mined advisors will approve a decision in a wide range of cases.

Chapter 3

The Model

3.1 Timing and Strategies

In this thesis we examine the quadratic-loss function "beauty-contest" coordination game from Myatt and Wallace (2011), interpreted as modelling the relationship between party leaders and party activists as presented by Dewan and Myatt (2008). We adapt this game by introducing a similar quadratic-loss function for party leaders. Furthermore we allow for leaders with individual policy preferences. Leaders are able to freely choose a 'bias" in their signals according to their preferences. Formally, a two-stage sequential game is played by n players, referred to as leaders, and a unit mass of players indexed by $\ell \in [0,1]$, referred to as party activists. The timing of the game is as follows:

1. A leader selects an individual policy position (or bias) $b_i \in \Re, i \in [1,\ldots,n]$. Note that leaders can freely select their bias from the reals; that is, we allow for infinitely extreme biases and policy preferences. This is in order to be consistent with the formulation of the signal noise terms which are (as we shall see) normally distributed and can therefore, in theory, take on infinitely large values as well. As we will see in the next chapter, for most equilibria limiting the bias to some interval $[\underline{M}, \bar{M}]$ does not change the results, although cases at the bounds of the interval would need to be considered separately. Also note that the bias b_i can be different from the policy preference $\bar{b}_i$, which is given exogenously. Furthermore, any leader i can perfectly observe both the bias b_j and the policy preference $\bar{b}_j$ of the other $n-1$ leaders. This step can be interpreted as the leaders preparing themselves to address (that is signal) the activists.

2. Next, the party activists determine an information acquistion policy $z_\ell \in \Re_n^+$. The policy indicates the attention paid to the n leaders, with the ith element of z_ℓ indicating the attention paid to the ith leader. Note that throughout section 2 of this thesis different assumptions will be made regarding the knowledge of the players with respect to the leaders' biases and policy preferences athis stage. In particular, we will consider perfectly and imperfectly informed party activists, as well as completely naive party activists.

3. After the information acquistion policy is determined, the activists observe a vector of n signals $x_\ell \in \Re^n$. These signals inform the activists about the unobserved state variable θ. The precision of these signals depends on the choice of z_ℓ and they are furthermore biased depending on the selected biases b_i.

4. In the final stage, the activists take an action a_ℓ, which is contingent on the signal. This action is real-valued such that $a_\ell \in \Re$.

A party leader's strategy is a singleton $\{b_i\}$ whereas an activists strategy is a pair $\{z_\ell, A_\ell(\cdot)\}$. Here the function $A_\ell(\cdot) : \Re^n \mapsto \Re$ signifies how the n observed signal realizations are mapped to the action a_ℓ. Therefore we have that $A_\ell(x_\ell) = a_\ell$ with $a_\ell \in \Re$ and $x_\ell \in \Re^n$. As such, an equilibrium is described by $\{\{b_i\}, \{z_\ell, A_\ell(\cdot)\}\}$ for $\forall i \in [1,\ldots,n]$ and $\forall \ell \in [0,1]$. Intuitively, party leaders choose which policy they wish to communicate to the activists. The activists determine which leaders to pay attention to based on the clarity of their signals and the position the leader advocates. Lastly, activists determine which position to advocate themselves based on the information they have received from the party leaders.

3.2 Pay-Offs

The key feature of any beauty-contest game is that players have an incentive to balance their own interests with the unity of actions within the group. In the case of our political party, party activists wish to advocate a policy as close as possible to the optimal policy as well maintain unity within the party. This is captured in a quadratic loss function, which is derived directly from Myatt and Wallace (2011). An activist's pay-off depends on the proximity of his/her action a_ℓ to the underlying state variable θ. At the same time, the pay-off depends on the proximity of this action to the average action, which is given by $\bar{a} \equiv \int_0^1 a_\ell d\ell$. Furthermore, information acquisition z_ℓ is costly. This yields the following utility function:

$$u_\ell = \bar{u} - (1-\gamma)(a_\ell - \theta)^2 - \gamma(a_\ell - \bar{a})^2 - C(z_\ell) \tag{3.1}$$

Here γ gives the desire for unity. In line with both Dewan and Myatt (2008) and Myatt and Wallace (2011) we restrict $\gamma \in (-1, 1)$. Thus we also allow $\gamma < 0$, that is we allow activists to have a dislike to unity. The cost function $C(z_\ell)$ is assumed to be increasing, convex, and differentiable. Party leaders have a similar quadratic loss function, be it with one major difference: since leaders do not directly take an action we need to construct a utility function that measures the overall direction and unity of the party. As such, the leaders' pay-offs depend on the proximity of the *average* party action from the leaders' individual policy preferences. Furthermore, the pay-offs depend on the *variance* in the party actions. Formally, this yields the following utility function:

$$v_i = \bar{v} - (1-\gamma)(\bar{a} - \bar{b}_i - \theta)^2 - \gamma \int_0^1 (a_\ell - \bar{a})^2 d\ell \tag{3.2}$$

A few things need to be noted here. First of all, it is assumed that the concern for unity γ is identical for party leaders and party activists. That is, γ is identical in equations 3.1 and 3.2. One can interpret γ as indicating how important having a highly unified party is within the given political system. As such, γ is exogenously determined by the institutional characteristics of the political system such as its voting rules, the existence and relative positions and strength of other political parties, and voter behaviour. Naturally, one can imagine a scenario in which γ differs between leaders and activists or even between individual leaders and activists. This is however outside the scope of interest of this thesis. Secondly it is to be noted that the individual policy preference $\bar{b}_i$ should be interpreted as the deviation from the overall party policy position θ. Therefore the policy position of an individual party leader equals $\theta + \bar{b}_i, \bar{b}_i \in \Re$.

A final note concerns the last term in the party leader's utility function 3.2, which represents party unity. Dewan and Myatt (2008) discuss that the specification used in the party activists utility function 3.1 represents a need for conformance rather than a call for unity and suggest the following specification instead:

$$\int_0^1 (a_\ell - a'_\ell)^2 d\ell'$$

They then continue to show that this specification is behaviourally equivalent to the specification used in the activists utility function 3.1 since the following holds:

$$\int_0^1 (a_\ell - a'_\ell)^2 d\ell' = (a_\ell - \bar{a})^2 + \int_0^1 (a'_\ell - \bar{a})^2 d\ell'$$

Since the second term on the RHS is independent of ℓ we have that the behaviour of activist ℓ is fully determined by the first term on the RHS. In the case of the party leader however, there is no reason to assume that the last term in 3.2 is independent of b_i. In fact, in general we would expect $a_\ell, \forall \ell \in [0, 1]$ to be dependent on b_i, and therefore $\bar{a}$ is also dependent on b_i. As such, the last term in 3.2 captures overall party unity independent of the specific subscript ℓ yet dependent on b_i. Of course, the criticism that this term captures a need to conform rather than a call for unity still applies. It can however be argued that in many social settings (and also in the political setting that is examined in this paper) a need for conformance is at least equally valid as a description of real-world incentives.

3.3 Signals & Information

Before the equilibria and their characteristics can be discussed, the signal and information structure needs to be determined. An advantage of the particular specification of the bias is that the variance-covariance structure as found in Myatt and Wallace remains intact. This follows directly from the basic properties of (co)variances. In particular, for any random variables x and y and any scalars a and c we have that:

$$\sigma(x, y) = \sigma(x + a, y) = \sigma(x, y + c) = \sigma(x + a, y + c)$$

Since we have generally defined the bias b_i to be a non-random scalar the original covariance/variance structure thus remains intact. We assume that activists start with no information on the underlying state of nature (thus sharing an improper prior over θ). The activists receive n signals from the n party leaders, the ith signal satisfying the following:

$$x_{i\ell} = \theta + b_i + \eta_i + \epsilon_{i\ell}, \quad \eta_i \sim N(0, \kappa_i^2), \quad \epsilon_{i\ell} \sim N(0, \frac{\xi_i^2}{z_{i\ell}}) \tag{3.3}$$

The noise terms η_i and $\epsilon_{i\ell}$ are independently distributed. Equation 3.3 has a straightforward interpretation. First of all, there is the "sender noise" term. The sender imperfectly observes the state of the world and thus adds noise. This is captured in the terms $\theta + \eta_i$, which represents the actual state of the world and the sender noise. The accuracy of the signal can thus be indexed by the precision, which is the reciprocal of the noise variance $\frac{1}{\kappa_i^2}$. In turn, the sender or leader can freely choose a bias to include in the signal captured in the b_i term. This signal is consequently sent to the receivers, who are the party activists. On the receiver end there is also noise, which is captured by the $\epsilon_{i\ell}$ term. The noise is composed of the clarity of the noise, which is indexed by $\frac{1}{\xi_i^2}$ and the attention paid by activist ℓ to the ith leader, which is given by $z_{i\ell}$. Overall clarity is the composition of these two factors (which counteract each other) and can be indexed by the precision $\frac{z_{i\ell}}{\xi_i^2}$. Higher values of ξ_i indicate less natural clarity with respect to the ith leader's message. Thus total clarity increases linearly with $z_{i\ell}$, so more attention paid to a signal by an activist increases the specific clarity of that signal towards the specific activist. Note that $z_{i\ell} = 0$ implies that no attention is paid by the activist to the ith information source. In that case, signal noise $\frac{\xi_i^2}{z_{i\ell}} \to \infty$ as $z_{i\ell} \to 0$ so the signal becomes complete noise.

The covariance structure is also identical to the Myatt and Wallace paper. That is, conditional on θ, the information sources are independent but activists' observations are correlated. Between two activists ℓ and ℓ' we have that the covariance is determined by the sender noise introduced by leader i. Thus we have:

$$\text{cov}[x_{i\ell}, x_{i\ell'}] = \kappa_i^2$$

Interestingly, this shows that as long as a signal has imperfect accuracy (κ_i^2 >0) signals move together. Formally, this implies the following specification:

$$x_{i\ell}|\theta \sim N(\theta + b_i, \sigma_{i\ell}^2), \quad \mathrm{cov}[x_{i\ell}, x_{i\ell'}|\theta] = \rho_{i\ell\ell'}\sigma_{i\ell}\sigma_{i\ell'}, \quad \forall \ell \neq \ell', \quad \forall i \tag{3.4}$$

From the basic definitions of variances and correlations we thus get the following:

$$\sigma_{i\ell}^2 = \kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}, \quad \rho_{i\ell\ell'} = \kappa_i^2 \left[\left(\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}} \right) \left(\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell'}} \right) \right]^{-\frac{1}{2}} \tag{3.5}$$

From this specification, it is easily seen that the model allows for both completely public and completely private information and the continuum in between. We can think of completely private information as a signal that is uncorrelated with all other actvists' signals whilst completely public information is correlated to all other activists' signals with correlation coefficient equal to 1. As we can see, the model allows for a wide variety of signal characteristics, with some being partially private and/or partially public. The degree of correlation is determined by the individual accuracy of the signal given by κ_i^2. Since this error term is identical for all activists, the degree of covariance between signals is linearly related with this term. If $\kappa_i^2 = 0$ the signals are purely private. This is because any variance in the signals is only determined by $\epsilon_{i\ell}$ which genearlly is unique for each activist and based on the individual information acquisition policy. With respect to the nature of information, this model thus generalizes many of the previous studies into similar models (see the literature review for an overview). With the information structure, utility functions and strategy space described and constructed, we can now move on to finding equilibria.

Chapter 4

Results

4.1 Strategy Class

Players' strategies are given by the singleton $\{b_i\}$ and the pair $\{z_\ell, A_\ell(\cdot)\}$. Here $A_\ell(x_\ell)$ specifies activists ℓ's response to the signal realization x_ℓ. In theory, $A_\ell(x_\ell)$ could be any function over x_ℓ. The literature however typically restricts $A_\ell(x_\ell)$ to be a linear strategy. A strategy is linear if there exists a vector of weights $w_\ell \in \Re^n$ such that $A_\ell(x_\ell) = \sum_i^n w_{i\ell} x_{i\ell} = w'_\ell x_\ell$. It is easy to show that any best response to a linear strategy is also a linear strategy. Furthermore, Dewand and Myatt (2008) show that if one restricts the class of non-linear strategies to those bounded by linear strategies (that is, if for a linear function $\bar{A}(x_\ell)$ it holds that $|A(x_\ell) - \bar{A}(x_\ell)|$ remains bounded for all x_ℓ) all equilibrium strategies must be linear. As such, there are many convincing arguments to impose linearity on $A(x_\ell)$. In this thesis however, we will extend our strategy space to allow for *affine* equilibria.

Definition *A strategy $A_\ell(x_\ell)$ is affine in x_ℓ if there are weights $w_\ell \in \Re^n$ and a scalar $c_\ell \in \Re$ such that $A_\ell(x_\ell) = \sum_i^n w_{i\ell} x_{i\ell} + c_\ell = w'_\ell x_\ell + c_\ell$.*

Clearly, the linear strategy is a specific case of this affine strategy for which $c_\ell = 0$. Before we discuss some of the properties of affine strategies, let us discuss the motivation for extending our class of strategies to allow for such affine transformations. In particular, consider the following stylized example: Assume there is one leader ($n = i = 1$) who can communicate with perfect clarity and accuracy (that is $\kappa_i^2 = \xi_i^2 = 0$). Furthermore, party activists can perfecly observe the applied bias b_i. Thus the signal x_ℓ received by activist ℓ is now non-random and thus perfectly accurate and all party activists observe $x_\ell = \theta + b_i$. Since no individual errors are introduced any equilibrium must clearly be perfectly symmetrical. As such, the equilibrium strategy is the one that minimizes the first term of the quadratic loss function:

$$A(x_\ell)^* = \arg\min_{A(x_\ell)} \ \mathbb{E}\left[(A(x_\ell) - \theta)^2\right]$$

The intuitively obvious solution is to ensure that $A(x_\ell) = \theta$ such that the quadratic term is minimized at 0. But this involves an affine transformation over x_ℓ. In particular consider $w_{i\ell} = 1$ and $c_\ell = -b_i$. We then have:

$$A(x_\ell) = w_{i\ell} x_{i\ell} + c_\ell = (\theta + b_i) - b_i = \theta$$

In more general terms, affine transformations allow party activists to *de-bias* the leaders' messages based on their best estimation of the individual biases b_i. Party activists will want to compensate for the bias in the signal to stay as close as possible to the optimal policy. This is captured in the following lemma:

Lemma 1 *For any given signal accuracy and clarity, the first quadratic loss term is minimized only for unbiased strategies. That is,* $\mathbb{E}\left[(A(x_\ell)-\theta)^2\right]$ *is minimized for any* κ_i^2, ξ_i^2 *and* z_ℓ *if and only if* $\mathbb{E}\left[A(x_\ell)\right] = \theta$.

Proof To prove this proposition we prove the following inequality for any random variable X and scalar a:

$$\mathbb{E}\left[(X-a)^2\right] \geq var(X) = \mathbb{E}\left[(X-\mu)^2\right],\ \mu = \mathbb{E}\left[X\right]$$

We now write:

$$\mathbb{E}\left[(X-a)^2\right] - \mathbb{E}\left[(X-\mu)^2\right] = \mathbb{E}\left[X^2\right] - 2a\mathbb{E}\left[X\right] + a^2 - \mathbb{E}\left[X^2\right] + 2\mu\mathbb{E}\left[X\right] - \mu^2 = a^2 - 2a\mu + \mu^2 = \mathbb{E}\left[(a-\mu)^2\right]$$

Clearly we have that $\mathbb{E}\left[(a-\mu)^2\right] \geq 0$ with equality holding only for $a = \mu = \mathbb{E}\left[X\right]$. So for any given variance, the term is only minimized if a is equal to the expectation of X. Now set $a = \theta$ and $X = A(x_\ell)$. Clearly then, for our quadratic loss term to be minimized we require $\mathbb{E}\left[A(x_\ell)\right] = \theta$. This proves the lemma. ■

The above lemma provides a more formal motivation for the use of affine strategies since they offer a natural solution to unbiasing. In fact, the above lemma provides us with a simple way to derive the scalar component c_ℓ for any affine strategy. Consider the linear component of the affine strategy, which is given by $\sum_i^n w_{i\ell}x_{i\ell}$. Now we can define the scalar component to be defined as follows:

$$c_\ell = \theta - E\left[\sum_i^n w_{i\ell}x_{i\ell}\right] \tag{4.1}$$

Lemma 1 now guarantees that, for a given variance and symmetric strategies, the first quadratic loss term is minimized. As we shall see later on, the above definition of the scalar component reduces finding the equilibrium in the perfect information scenario to a triviality.

Similarly to the linear case, another motivation for sticking to affine strategies is that the best responses to affine strategies are themselves affine strategies. To see this, differentiate the quadratic objective function w.r.t. the strategy played by activist ℓ:

$$\frac{\partial u_\ell}{\partial A_\ell(x_\ell)} = -2(1-\gamma)(A_\ell(x_\ell)-\theta) - 2\gamma(A_\ell(x_\ell)-\bar{a}) = 0 \Rightarrow A_\ell(x_\ell) = (1-\gamma)\mathbb{E}\left[\theta|x_\ell\right] + \gamma\mathbb{E}\left[\bar{a}|x_\ell\right]$$

Now we use the fact that $\mathbb{E}\left[\bar{a}|x_\ell\right] = \mathbb{E}\left[\int_0^1 a_{\ell'}d\ell'|x_\ell\right]$. Since the expectation is independent of the label ℓ' we get that $\mathbb{E}\left[\bar{a}|x_\ell\right] = \mathbb{E}\left[a_{\ell'}|x_\ell\right]$. Of course, we also have that $a_{\ell'} = A(x_{\ell'})$. Plugging this in yields the following:

$$A_\ell(x_\ell) = (1-\gamma)\mathbb{E}\left[\theta|x_\ell\right] + \gamma\mathbb{E}\left[A(x_{\ell'})|x_\ell\right]$$

Now, the first expecation term is linear in x_ℓ by the linear properties of expectations of normally distributed variables. The second expectation term is affine in x_ℓ by assumption. Any linear combination of affine and linear combinations is itself affine, so the best response strategy is also affine. This thus shows that the best response strategy to affine strategies is itself affine. Before moving on to finding the equlibrium in the perfect information case, one final property of affine strategies needs to be determined, which is now presented as a lemma.

Lemma 2 *For any affine strategy* $A_\ell(x_\ell) = \sum_i^n w_{i\ell}x_{i\ell} + c_\ell$ *it holds that* $\sum_i^n w_{i\ell} = 1$. *That is, the weights attached to the individual signals received from party leaders by party activists sum to one.*

Proof[1]. Consider an affine equilibrium strategy profile $A(x_\ell) = w'x_\ell + c_\ell$. We have shown above that the following holds for any equilibrium (best-response) strategy:

$$A_\ell(x_\ell) = (1-\gamma)\mathbb{E}\left[\theta|x_\ell\right] + \gamma\mathbb{E}\left[A(x_{\ell'})|x_\ell\right]$$

By the properties of the normal distribution, we have that $\mathbb{E}\left[\theta|x_\ell\right] = a'x_\ell$ where the elements of a sum to one. We also have that $\mathbb{E}\left[A(x_{\ell'})|x_\ell\right] = w'\mathbb{E}\left[x_{\ell'}|x_\ell\right] + c_{\ell'}$. Once again by normality we have that $\mathbb{E}\left[x_{\ell'}|x_\ell\right] = Bx_\ell$ where B is an $n \times n$ inference matrix. This matrix also has the properties that all rows sum to one. Combining all this we thus have:

$$A_\ell(x_\ell) = w'_\ell x_\ell + c_\ell = (1-\gamma)a'x_\ell + \gamma\left(w'Bx_\ell + c_{\ell'}\right) = \left((1-\gamma)a' + \gamma w'B\right)x_\ell + \gamma c_{\ell'}$$

Seeing as all elements of a sum to one, and all rows of B sum to one, this equality can only hold if all elements of w also sum to one. So we have that $\sum_i^n w_{i\ell} = 1$. Furthermore, this also holds for all best replies to this affine strategy so we can assume it holds for all ℓ. This proves the lemma. $\blacksquare$.

We have now established some basic properties of the affine equilibrium and provided several motivations for its use in our setting. As we will see in the next two sections, these properties go a long way towards establishing the equilibria in the simplest settings: complete information and complete naivety.

4.2 Equilibrium: Complete Information

We now consider our model in a setting of perfect information. This implies that party activists can perfectly observe the individual biases b_i for all n party leaders. To derive the equilibrium strategies under this condition, we will proceed by backward induction. That is, we solve for $A_\ell(x_\ell)$ for a given information acquisition policy $z\ell$ and given biases b_i. As discussed in the previous section, we limit ourselves to affine strategies of the form $A_\ell(x_\ell) = w'_\ell x_\ell + c_\ell$. Afterwards, we solve for z_ℓ. Finally, we solve for the individual biases introduced by the leaders. For ease of notation we now introduce the bias vector $b \in \Re^n$, for which the ith entry b_i is the bias introduced by leader i.We now start by rewriting the activists' objective function (equation 3.1). First we note that the following holds for any affine strategy:

$$a_\ell - \theta = \sum_i^n w_{i\ell}(\eta_i + \epsilon_{i\ell} + b_i) + c_\ell$$

By squaring and taking expectations we get the following:

$$\mathbb{E}(a_\ell - \theta)^2 = \sum_i^n w_{i\ell}^2\left(\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right) + \left(\sum_i^n w_{i\ell}b_i + c_\ell\right)^2 \tag{4.2}$$

Furthermore, $\bar{a} = \theta + \sum_i^n w_i\left(\eta_i + b_i\right) + \bar{c}$. Here $\bar{c} = \int_0^1 c_{\ell'}d\ell'$. This yields that $a_\ell - \bar{a} = \sum_i^n w_{i\ell}\epsilon_{i\ell} + \sum_i^n (w_{i\ell} - w_i)(\eta_i + b_i) + (c_\ell - \bar{c})$. Squaring and taking expectations yields:

$$\mathbb{E}(a_\ell - \bar{a})^2 = \sum_i^n w_{i\ell}^2\frac{\xi_i^2}{z_{i\ell}} + \sum_i^n (w_{i\ell} - w_i)^2(\kappa_i^2 + b_i^2) + 2\sum_i^n b_i(w_{i\ell} - w_i)(c_\ell - \bar{c}) + (c_\ell - \bar{c})^2 \tag{4.3}$$

Now combining equations 4.2 and 4.3 we get the following expression:

[1]Note that this proof is nearly identical to the one provided for the linear case in Myatt and Wallace (2011)

$$\mathbb{E}(u_\ell) = \bar{u} - \underbrace{\sum_i^n w_{i\ell}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right] + (1-\gamma)\left(\sum_i^n w_{i\ell} b_i + c_\ell\right)^2}_{L_1(z_\ell, w_\ell, c_\ell)} +$$

$$\underbrace{\gamma\left[\sum_i^n (w_{i\ell} - w_i)^2(\kappa_i^2 + b_i^2) + 2\sum_i^n b_i(w_{i\ell} - w_i)(c_\ell - \bar{c}) + (c_\ell - \bar{c})^2\right]}_{L_2(w_\ell, w, c_\ell, c)} - C(z_\ell) \quad (4.4)$$

Note that for any symmetric equilibrium, the second term (labeled L_2) vanishes as in a symmetric[2] equilibrium $w_i = w_{i\ell}$ and $c = \bar{c}$. More generally, w_ℓ and c_ℓ have no first order effect local to w and $\bar{c}$. As such, we can ignore the L_2 term when looking for an equilibrium. Therefore, maximizing the objective function is equivalent to minimizing the sum of L_1 and the information acquisition costs $C(z_\ell)$. This allows us to define the activists' equilibrium strategies for a given bias b in this situation of perfect information.

Lemma 3 *An activists strategy $\{z_\ell, w_\ell, c_\ell\}$ forms a symmetric equilibrium under perfect information if and only if it solves:*

$$\min \sum_i^n w_{i\ell}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right] + (1-\gamma)\left(\sum_i^n w_{i\ell} b_i + c_\ell\right)^2 + C(z_\ell) \text{ subject to } \sum_i^n w_{i\ell} = 1 \quad (4.5)$$

Proof The expression above follows directly from the discussion above, and the constraint follows from Lemma 2. Given that $C(z_\ell)$ is convex (which it is by assumption) the loss function is convex in all its arguments and the unique (global) solution to this optimization problem is thus obtained from the relevant first order conditions. ∎

With this established, the solution follows readily by deriving the relevant first-order conditions and solving the system of equations. This yields our first equilibrium which is characterized by the following proposition:

Proposition 1 *With perfect information there exist infinitely many equilibria. In each equilibrium, the activists' strategies $\{z_\ell, w_\ell, c_\ell\}$ are symmetric and unique and given by the following:*

$$w_{i\ell}^* = \frac{\psi_i}{\sum_j^n \psi_j}, \; c_\ell^* = -\sum_i^n w_{i\ell}^* b_i^*, \; z_{i\ell}^* = \frac{\xi_i w_{i\ell}^*}{\sqrt{C'(z)}} \text{ with } \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}}$$

Leaders are indifferent between all $b_i \in \Re$ and therefore have infinitely many best responses. The equilibrium strategy space is therefore the entirety of $\Re$. Alternatively, each bias vector $b \in \Re^n$ constitutes an equilbrium bias vector.

Proof The minimization problem is solved through the use of the method of Lagrange multipliers. This yields the following Lagrangian:

$$\mathcal{L}(w_\ell, z_\ell, c_\ell, \lambda) = \sum_i^n w_{i\ell}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right] + (1-\gamma)\left(\sum_i^n w_{i\ell} b_i + c_\ell\right)^2 + C(z_\ell) + \lambda\left(\sum_i^n w_{i\ell} - 1\right) \quad (4.6)$$

[2] Symmetric with respect to the activists' strategies.

Taking derivatives yields the following first-order conditions:

$$\frac{\partial \mathcal{L}\left(w_\ell, z_\ell, c_\ell\right)}{\partial w_{i\ell}} = 2w_{i\ell}\left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right] + 2(1-\gamma)\left(w_{i\ell}b_i + c_\ell\right) + \lambda = 0 \tag{4.7}$$

$$\frac{\partial \mathcal{L}\left(w_\ell, z_\ell, c_\ell\right)}{\partial z_{i\ell}} = -\frac{\xi_i^2 wi_i^2}{z_{i\ell}^2} + \frac{\partial C(z_\ell)}{\partial z_{i\ell}} = 0 \tag{4.8}$$

$$\frac{\partial \mathcal{L}\left(w_\ell, z_\ell, c_\ell\right)}{\partial c_\ell} = 2(1-\gamma)\left(\sum_i^n w_{i\ell}b_i + c_\ell\right) = 0 \tag{4.9}$$

$$\frac{\partial \mathcal{L}\left(w_\ell, z_\ell, c_\ell\right)}{\partial \lambda} = \sum_i^n w_{i\ell} - 1 = 0 \tag{4.10}$$

Condition 4.9 can be rearanged to yield the following solution for the scalar component of the affine strategy c_ℓ, yielding the first component of the proposition:

$$c_\ell^* = -\sum_i^n w_{i\ell}b_i \tag{4.11}$$

Note that this also follows from Lemma 1, since the above expression is exactly the debiasing factor that was discussed there. Substituting this specification into equation 4.7 and rearranging yields:

$$2w_{i\ell}\left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right] + \lambda = 0$$

Since this equality must hold for all $w_{i\ell}$ we have that the following holds for all i, j and some constant K:

$$w_{i\ell}\left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right] = w_{j\ell}\left[(1-\gamma)\kappa_j^2 + \frac{\xi_j^2}{z_{j\ell}}\right] = K$$

Now define ψ_i as follows:

$$\frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}} \equiv \psi_i$$

Then we have $w_{i\ell} = K\psi_i$ for all i. Using equation 4.11 this yields that $K\sum_j^n \psi_j = 1$ or alternatively that $K = \frac{1}{\sum_j^n \psi_j}$. Plugging this in for K above yields our final solution for the weights:

$$w_{i\ell}^* = \frac{\psi_i}{\sum_j^n \psi_j} \text{ with } \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}} \tag{4.12}$$

To find the equilibrium information acquisition policy, we solve equation 4.8 for $z_{i\ell}$, writing $C'(z) = \frac{\partial C(z_\ell)}{\partial z_{i\ell}}$:

$$\frac{\xi_i^2 w_{i\ell}^2}{z_{i\ell}^2} = C'(z) \Rightarrow z_{i\ell}^2 = \frac{\xi_i^2 w_{i\ell}^2}{C'(z)} \Rightarrow z_{i\ell} = \frac{\xi_i w_{i\ell}}{\sqrt{C'(z)}}$$

Thus our final solution for the information acquisition policy is as follows:

$$z_{i\ell}^* = \frac{\xi_i w_{i\ell}^*}{\sqrt{C'(z)}} \tag{4.13}$$

Finally, we turn to the leaders. Generally, we would have to rewrite the leaders' objective function and maximize wrt. the individual bias b_i. In this case however, we can solve for the leaders' equilibrium strategy by making a simple observation. We begin by writing out he activists strategies:

$$A_\ell(x_\ell) = \sum_i^n w_{i\ell}^* x_{i\ell} + c_\ell = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell} + b_i) + c_\ell$$

Now we plug in that in an equilibrium we have that $c_\ell^* = -\sum_i^n w_{i\ell} b_i$:

$$A_\ell(x_\ell) = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell} + b_i) - \sum_i^n w_{i\ell} b_i = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell})$$

As we can thus see, each activists' individual policy choice under the equilibrium strategy is independent of the individual bias b_i. As such we have that a_ℓ and $\bar{a}$ are both also independent of b_i. In turn, this implies that the leader's objective function is completely independent of b_i under the equilibrium strategy. As a result of that, leaders are indifferent between any $b_i \in \Re$. Therefore the equilibrium bias vector is all of $\Re^n$. That is, there exist infinitely many best responses to the activists' equilibrium strategies and therefore also infinitely many equilibria. This concludes the proof. ■

It is clear from the proposition that with perfect information, activists react exactly the same as they would in the model with no bias. The only difference is that the signal is de-biased first. As such, with perfect information the existance of a bias leaves the information acquisition policy unchanged. This result follows from the fact that with perfect information, information acquisition is a separate process altogether, since the influence of the bias is perfectly removed by the debiasing process. Activists can perfectly see through the leaders' rethoric and perfectly able to distill the information regarding the state of the world that they wish to uncover. As such, any difference between the optimal policy and the actual policy advocated by party activists is the result of the noise introduced by the leaders, whilst the bias is filtered out completely. The bias introduced by leaders therefore has no impact on the expected policy advocated. The leaders, realizing that party advocates are able to perfectly see through the veil of rhetoric, are as a result indifferent with respect to the bias they introduce. Another result of this proposition is that all comparative statics results as derived by Myatt and Wallace (2012) extend directly to this setting. In particular, we have that clarity of communication takes priority over accuracy with respect to information acquisition.

4.3 Equilibrium: Complete Naivety

In the previous section it was assumed that all party activists perfectly observe the bias introduced by the party leaders. In that situation we logically concluded that the activists react by compensating for the bias, making it irrelevant. As a result of this, party leaders are indifferent towards the bias they introduced. In this section, we consider the other extreme. Here, party activists are completely naive, implying that they are unaware of the fact that party leaders introduce a bias. We can therefore think of these activists as being ignorant of the fact that leaders can be unbiased. As we shall see, the result of this is that they 'blindly' follow their leaders, without taking into consideration their individual biases.

Formally, we can think of this setting by considering the signals that activists *believe* they receive and contrast these to the signals the activists *actually* receive. Activists believe they receive unbiased signals, which is composed as follows:

$$x_{i\ell} = \theta + \eta_i + \epsilon_{i\ell}, \quad \eta_i \sim N(0, \kappa_i^2), \quad \epsilon_{i\ell} \sim N(0, \frac{\xi_i^2}{z_{i\ell}}) \tag{4.14}$$

In reality, activists receive biased signals, which are given as before:

$$x_{i\ell} = \theta + b_i + \eta_i + \epsilon_{i\ell}, \quad \eta_i \sim N(0, \kappa_i^2), \quad \epsilon_{i\ell} \sim N(0, \frac{\xi_i^2}{z_{i\ell}}) \tag{4.15}$$

Simply put, naive activists always observe $b_i = 0$ whilst generally $b_i \neq 0$. A further assumption we make is that leaders are aware of the fact that the activists are naive, and will as such base their biases on this. With this in mind, it is trivial to derive the activists' equilibrium strategies, since they follow directly from Proposition 1 by setting $b_i = 0, \forall i$. This yields that their information acquisition policy z_ℓ and their signal weighting w_ℓ remain the same, with the only difference being that $c_\ell = 0$. This can also be easily seen by considering that activists' equlibrium strategies now satisfy the following:

$$\min \sum_i^n w_{i\ell}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right] + C(z_\ell) \textit{ subject to} \sum_i^n w_{i\ell} = 1$$

Now note that this optimization problem yields identical first-order conditions on z_ℓ and w_ℓ. The big difference between the setting of complete information and the current setting is that the discrepancy between the perceived signals and the actual received signals implies that party leaders can now affect the activists' advocated policy. To see this, note that activists perceive they are advocating the following under the equilibrium strategies:

$$A_\ell(x_\ell) = \sum_i^n w_{i\ell}^* x_{i\ell} + c_\ell = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell}) + c_\ell = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell})$$

Whilst in reality, a bias is present in the signal. As a result, they are in fact advocating the following instead:

$$A_\ell(x_\ell) = \sum_i^n w_{i\ell}^* x_{i\ell} + c_\ell = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell} + b_i) + c_\ell = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell} + b_i)$$

So the policy advocated by activist ℓ is no longer independent of b_i, and as such party leaders can influence the party line through biasing. Based on this specification and the leader's objective function we can formulate the following lemma:

Lemma 4 *A party leader's strategy* $\{b_i\}$ *forms an equilibrium if and only if it solves:*

$$\min \mathbb{E}\left[\left(\bar{a} - \theta - \bar{b}_i\right)^2\right] \textit{ given } A_\ell(x_\ell) = a_\ell = \sum_i^n w_{i\ell}(\theta + \eta_i + \epsilon_{i\ell} + b_i) \tag{4.16}$$

Proof Party leaders maximize their expected utility given the party activists' equilibrium strategies. Thus leaders maximize the following:

$$\mathbb{E}\left[v_i\right] = \bar{v} - (1-\gamma)\mathbb{E}\left[(\bar{a} - \bar{b}_i - \theta)^2\right] - \mathbb{E}\left[\int_0^1 (a_\ell - \bar{a})^2 d\ell\right]$$

Let us evaluate the last term. Given that the term within the integral is independent of the label ℓ (as noted in the previous section) we can rewrite the above as follows:

$$\mathbb{E}\left[\int_0^1 (a_\ell - \bar{a})^2 d\ell\right] = \int_0^1 \mathbb{E}\left[(a_\ell - \bar{a})^2\right] d\ell$$

Furthermore we have that $a_\ell - \bar{a} = \sum_i^n w_{i\ell}\epsilon_{i\ell} + \sum_i^n (w_{i\ell} - w_i)(\eta_i + b_i) + (c_\ell - \bar{c})$. In our equilibrium we have that $w_{i\ell} = w_i$ and that $c_\ell = \bar{c} = 0$. Thus we find that $a_\ell - \bar{a} = \sum_i^n w_{i\ell}\epsilon_{i\ell}$. Squaring and taking expectations yields:

$$\mathbb{E}\left[(a_\ell - \bar{a})^2\right] = \sum_i^n w_{i\ell}^2 \frac{\xi_i^2}{z_{i\ell}}$$

Plugging this into our expression for the expected quadratic loss term yields:

$$\mathbb{E}\left[\int_0^1 (a_\ell - \bar{a})^2 d\ell\right] = \int_0^1 \left(\sum_i^n w_{i\ell}^2 \frac{\xi_i^2}{z_{i\ell}}\right) d\ell \tag{4.17}$$

Now we know from Proposition 1 that both w_i and z_i are independent of b_i. Therefore the entire expression 4.17 is independent of b_i and thus has no first-order effect. Therefore it can be ignored when looking for an equilibrium. As such, minimizing the first quadratic loss term is equivalent to maximizing the objective function. This concludes the proof. ■

Given our symmetric equilibrium we have that $\mathbb{E}[\bar{a}] = \mathbb{E}[a_\ell]$ with any deviation in the realization resulting purely from the noise terms ϵ_i and η_i. In particular, we have that $\mathbb{E}[\bar{a}] = \theta + \sum_i^n w_i b_i$. Note however that there may be leaders who receive no attention, that is $w_j = z_j = 0$ potentially for some j. This needs to be accounted for in our expression for $\mathbb{E}[\bar{a}]$. We can rewrite our expressions for z_i from Proposition 1 to find a condition for which signals are ignored (see Appendix A for this short derivation). This yields that $zi > 0$ for $\xi_i < K_i$ with K_i defined as follows:

$$K_i \equiv \frac{1}{\sqrt{\frac{\partial C(z)}{\partial z_i} \sum_{j=1}^n \psi_j}}$$

Thus we have that a signal is not ignored if the clarity is not too smal. Signals with too little clarity can be ignored when looking for an equilibrium since they do not affect the average policy that is advocated. In particular assume, without loss of generality, that the following holds:

$$\xi_1 \leq \xi_2 \leq \cdots \leq \xi_m \leq \cdots \leq \xi_n$$

Assume further and without loss of generality that for some $m \leq n$ we have that $\xi_i < K_i$ for all $i \leq m$ and $\xi_i \geq K_i$ for all $i > m$, such that signal m is the signal with the smallest clarity that is not ignored. Then we have that $\mathbb{E}[\bar{a}] = \theta + \sum_i^m w_i b_i$. Thus we can ignore all $n - m$ remaining signals when looking for an equilibrium. As a result of that, in any equilibrium all the leaders that are ignored are once again indifferent with respect to their choice b_i. So each $b_i \in \Re$ is an equilibrium strategy for all leaders $i > m$.

Note that in order to derive the equilibrium strategy for all leaders $i \leq m$ we can make use of Lemma 1. To see this, set $\bar{a} = X$ and $\theta + b_i = a$ in the expression presented in the proof of Lemma 1 and then apply the lemma. From this it follows that for the expectation of the quadratic loss term to be minimized we need to have that $\mathbb{E}[\bar{a}] = \theta + \bar{b}_i$ for all $i \leq m$. This implies $\sum_{i=1}^m w_i b_i = \bar{b}_i$. But this condition must hold for all i simultaneously. Crucially, this implies the following:

$$\sum_{i=1}^m w_i b_i = \bar{b}_i, \ \forall i \ \Rightarrow \bar{b}_1 = \bar{b}_2 = \cdots \bar{b}_m. \tag{4.18}$$

As such, the equilibrium condition can only ever hold if all the policy preferences of the m leaders that receive attention are identical, that is $\bar{b}_i = \bar{b}$ for some $\bar{b} \in \Re$. To further see why this must be consider the opposite case, in which there is some i for which $\bar{b}_i \neq \bar{b}$. Then this i can always improve his utility by setting b_i according to the following best-response relation:

$$b_i = \frac{1}{w_i}\left(\bar{b}_i - \sum_{j \neq i} w_j b_j\right)$$

But now all other $j \neq i$ can do the same again acccording to the same rule. Leader i can reply to that again by using the same rule, and this process continues ad infinitum. The only stable outcome is if $\bar{b}_i = \bar{b}$ and if $b_j = \bar{b}$ for some constant $\bar{b} \in \Re$. In that case the rule gives the following:

$$b_i = \frac{1}{w_i}\left(\bar{b} - \sum_{j \neq i} w_j \bar{b}\right) = \frac{1}{w_i} w_i \bar{b} = \bar{b}$$

Thus the best response relation becomes self-affirming (i.e. the best response correspondance has a fixed point) and as such this constitutes an equilibrium. For any other distribution of $\bar{b}_i$ such a fixed point cannot exist by equation 4.18. Alternativey, repeated subsitution of the best-response condition does not lead to convergence for any other distribution, which follows from inspection. Thus we have the following proposition:

Proposition 2 *With completely naive party activists an equilibrium can only exist if for all leaders $i \leq m$ such that $z_i > 0$ it holds that $\bar{b}_1 = \bar{b}_2 = \cdots \bar{b}_m = \bar{b}$ for some $\bar{b} \in \Re$. All leaders i for which $z_i = 0$ are indifferent between any $b_i \in \Re$. As such, for a given $\bar{b}$ there are infintely many equilibria. Leaders only receive attention if $\xi_i < K_i$ with:*

$$K_i \equiv \frac{1}{\sqrt{\frac{\partial C(z)}{\partial z_i} \sum_{j=1}^{n} \psi_j}}$$

The equilibrium strategies of party activists are symmetric, unique, and independent of $\bar{b}$ and are given by the following:

$$w_{i\ell}^* = \frac{\psi_i}{\sum_j^n \psi_j},\ c_\ell^* = 0,\ z_{i\ell}^* = \frac{\xi_i w_{i\ell}^*}{\sqrt{C'(z)}}\ \textit{with}\ \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}}$$

Proof The proof of the activists' equilibrium strategies follows directly from Proposition 1 by setting $b_i = 0$ for all i. The proof of the leaders' equilibrium strategies follows directly from Lemmas 1 and 4 and the discussion above. ■

Note the implication of the above proposition: if an equilibrium exists, it is a faction dictatorship if $m > 1$ and an individual dicatorship if $m = 1$ (that is, one person receives all the attention). Whether or not an equilibrium can exist depends on if the best communicators all have the same policy preference (i.e. belong to the same faction). The remarkable thing is that this is independent of how accurately these leaders perceive the state of the world or how extreme their views are. That is to say, the equiilibrium is independent of the value of b_i. Consider for example $|\bar{b}| \gg 0$. With completely naive party activists this extreme view forms a perfectly acceptable equilibrium policy position as long as the best communicators (those that receive attention) all advocate it. Thus this proposition fundamentally shows that rational decision making can lead to extreme factions dominating a party if the party activists are sufficiently naive. Note however that it does trivially follow from Proposition 2 that party welfare is decreasing in $|\bar{b}|$.

The question now remains whether or not it is likely that this will happen. Under one condition this could be the case: if extreme policies are easier to advocate clearly. Although we will not model this in this thesis, one can imagine that extreme policies are generally less 'nuanced' and may as such be easy to communicate. Formally this could be respresented as $\xi = g(|\bar{b}|)$ where we have that $\frac{\partial g(|\bar{b}|)}{\partial |\bar{b}|} < 0$ for some function $g(\cdot)$. That is, clarity is inversely related to extremity of the policy. Proposition 2 may also provide some explanation for why historically extremist leaders/politicians have generally been charismatic and/or talented orators. Additionally, it also gives some indication of why such leaders typically rise in times of great political uncertainty, since such situations more closely approximate our setting of naivety. Simple models could be constructed based on our framework to test such hypotheses, but this is beyond the scope of this thesis. Instead, we will move on to some refinements of our two equilibria.

4.4 Equilibrium Refinements: Biasing Costs

In the previous sections, we assumed that biasing was costless. This may be realistic in cases where the underlying state of the world θ concerns an inherently subjective measure. In many cases however, θ may represent an underlying state that is by its nature objectively measurable. In such scenarios, it seems unlikely that biasing the signal is costless. Leaders may have to spend time and/or other resources in order to achieve a certain bias. This may include the production of reports, the writing of speeches etc. As such, in this section we will adapt our model slightly in order to accomodate costly biasing. This is captured in the following objective function for leaders:

$$v_i' = \bar{v} - (1-\gamma)(\bar{a} - \bar{b}_i - \theta)^2 - \gamma \int_0^1 (a_\ell - \bar{a})^2 d\ell - C(b_i) \tag{4.19}$$

Here our cost-function $C(\cdot)$ is a function of the bias b_i. We assume that costs are related to the magnitude not the direction of the bias. As such we have that the cost function is even such that $C(b_i) = C(-b_i)$ for all $b_i \in \Re$. Furthermore it holds that $C(b_i)$ is convex, continuous and everywhere differentiable with $C(0) = 0$. Costs are everywhere increasing in the magnitude of the bias such that $\frac{\partial C(b_i)}{\partial |b_i|} \geq 0$. Note that since $C(\cdot)$ is everywhere differentiable and has a global minimum at $b_i = 0$ we have by Fermat's theorem that $C'(0) = 0$.[3] Intuitively this implies that the cost function "transitions smoothly" from positive to negative biases through the origin. An example of such a function is $C(b_i) = b_i^2$.

Let us first consider our first equilibrium under perfect information. Since the objective function for activists remains unchanged we only need to consider party activists. In the proof of Proposition 1 it was noted that all elements in the leader's objective function are independent of b_i under perfect information. For our new objective function this still holds for all terms except for the cost function. Therefore it follows trivially that utility maximization is equivalent to cost minimization. By the properties of the cost function we therefore see that we now have a unique equilibrium, as is captured in the following proposition:

Proposition 3 *With perfect information and costly biasing there exists a unique equilibrium. In this equilibrium, the activists' strategies $\{z_\ell, w_\ell, c_\ell\}$ are symmetric and given by the following:*

$$w_{i\ell}^* = \frac{\psi_i}{\sum_j^n \psi_j}, \; c_\ell^* = -\sum_i^n w_{i\ell}^* b_i^*, \; z_{i\ell}^* = \frac{\xi_i w_{i\ell}^*}{\sqrt{C'(z)}} \text{ with } \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}}$$

Leaders minimize costs in equilibrium. As such, leaders strategies $\{b_i\}$ are unique and symmetric with $b_i^* = 0, \; \forall i$.

Proof By the proof of Proposition 1, v_i was independent of b_i under perfect information. The alternative objective function v_i' can be written as $v_i' = v_i - C(b_i)$. Therefore only $C(b_i)$ is dependent on b_i in v_i' and as such utility maximization is equivalent to minimization of $C(b_i)$. From its properties we see that $C(b_i)$ is minimized when $b_i = 0$ such that $C(0) = 0$. The objective function for activists remains unchanged and therefore their equilibrium strategies follow directly from Proposition 1. ■

Proposition 3 follows naturally and should be intuitively obvious. With perfect information, the bias is in many ways irrelevant, as it is perfectly compensated for by the party activists. With no costs choosing a bias was therefore of no consequence to the party leaders and as such they were indifferent between which bias to select. Now that biasing is costly, refraining from biasing is the only equilibrium since leaders now actively avoid the cost of doing so.

[3] Note that in order to formally apply Fermat's theorem we need to limit the domain our cost function to some finite interval. Since our cost function is monotonically increasing we have that $C(b_i) \to \infty$ as $|b_i| \to \infty$. This implies that for some $M \gg 0$ we can restrict the domain to the interval $(-M, M)$. By applying Fermat's theorem to our cost function over this interval $C'(0) = 0$ follows directly. Note that assumption of global differentiability is crucial here.

With imperfect information the situation is different since biasing directly affects the utility those leaders that receive attention (the first m leaders). Before analyzing their strategies let us first consider the last $n-m$ leaders. Obviously their choice is analogous to those under perfect information and as such their equilbrium strategy $\{b_i\}$ is also to set $b_i^* = 0$ for $i > m$. To see this note that these leaders are the worst communicators and are therefore ignored. As such their utility is also not affected by their choice of b_i and thus they are also cost minimizing. Now consider the remaining m leaders that do receive attention. Since the party activists' equilibrium strategies have not changed Lemma 4 still applies. This time however, we also need to account for costs. As such, the leaders choose b_i as to minimize the following:

$$(1-\gamma)\mathbb{E}\left[\left(\bar{a}-\theta-\bar{b}_i\right)^2\right] + C(b_i) \text{ given } A_\ell(x_\ell) = a_\ell = \sum_i^n w_{i\ell}(\theta+\eta_i+\epsilon_{i\ell}+b_i)$$

We see that the introduction of costs no longer allows us to find a solution by applying Lemma 1. As such, we need to find an alternative expression for the first term, just as we did when deriving the activists' strategies. To so, note that we can use the expression for $\bar{a}$ we derived earlier once more, this time using $\bar{c} = c_\ell = 0$. From this we see that $\bar{a} = \theta + \sum_i^m w_i (\eta_i + b_i)$. As such we get that $\bar{a} - \theta - \bar{b}_i = \sum_i^m w_i (\eta_i + b_i) - \bar{b}_i$. Squaring and taking expectations yields:

$$\mathbb{E}\left[\left(\bar{a}-\theta-\bar{b}_i\right)^2\right] = \sum_i^m w_i^2\kappa_i^2 + \sum_i^m w_i^2 b_i^2 + \sum_{j\neq i} w_j b_j w_i b_i - 2\bar{b}_i \sum_i^m w_i b_i + \bar{b}_i^2 \tag{4.20}$$

Now we can take first order conditions as follows:

$$\frac{\partial \mathbb{E}\left[\left(\bar{a}-\theta-\bar{b}_i\right)^2\right]}{\partial b_i} = 2b_i w_i^2 + 2w_i \sum_{j\neq i} w_j b_j - 2\bar{b}_i w_i = 0$$

And by rearranging we reach the best-response condition that was previously derived alternatively:

$$b_i w_i^2 = w_i\left(\bar{b}_i - \sum_{j\neq i} w_j b_j\right) \Rightarrow b_i = \frac{1}{w_i}\left(\bar{b}_i - \sum_{j\neq i} w_j b_j\right)$$

Note however that now the first order conditions are different, since we still have to include costs. This gives us the following condition:

$$(1-\gamma)\left[2b_i w_i^2 + 2w_i \sum_{j\neq i} w_j b_j - 2\bar{b}_i w_i\right] + C'(b_i) = 0$$

Rearranging yields that the following must hold for all $i \leq m$ (making use of the fact that $C(b_i) = C(-b_i)$ so that for $b_i < 0$ we can write $C'(b_i) = C'(-b_i) = -C'(b_i)$):

$$b_i = \begin{cases} \frac{1}{w_i}\left(\bar{b}_i - \sum_{j\neq i} w_j b_j\right) - \underbrace{\frac{1}{2w_i^2(1-\gamma)}C'(b_i)}_{U_i}, & b_i \geq 0 \\ \frac{1}{w_i}\left(\bar{b}_i - \sum_{j\neq i} w_j b_j\right) + \underbrace{\frac{1}{2w_i^2(1-\gamma)}C'(b_i)}_{U_i}, & b_i < 0 \end{cases} \tag{4.21}$$

Here the term labelled U_i is the *underbiasing factor*, for reasons that will soon be obvious. The implications of equation 4.21 are captured in the following proposition:

Proposition 4 *With biasing costs, the faction equilibrium from Proposition 2 no longer generally exists. Instead, leaders have an incentive to add less bias than their optimal policy preference.The incentive to understate their bias is larger if the call for unity is larger, if the leader is a worse communicator and/or technocrat, and if biasing costs are more sensitive to increases in the bias. That is,* $\frac{\partial U_i}{\partial \gamma} \leq 0, \frac{\partial U_i}{\partial \kappa_i} \leq 0, \frac{\partial U_i}{\partial \xi_i} \leq 0,$ *and* $\frac{\partial U_i}{\partial C'(b_i)} \leq 0.$

Proof We consider here the case $b_i > 0$. For $b_i < 0$ the argument is analogous. To see the first part of the propostion consider the case $b_j = \bar{b}_i = \bar{b}$ for all $j \neq i$. Then we have:

$$b_i = \frac{1}{w_i}\left(\bar{b} - (1-w_i)\bar{b}\right) - \frac{1}{2w_i^2(1-\gamma)}C'(b_i) = \bar{b} - \frac{1}{2w_i^2(1-\gamma)}C'(b_i) = \bar{b} - U_i$$

Using the fact that $C'(b_i) \geq 0$ we see that $b_i \leq \bar{b}$. Thus we see that we have $U_i \geq 0$ which implies that there is an incentive to underbias relative to the policy preference. Furthermore we see that $b_i = \bar{b}$ is no longer a best response to $b_j = \bar{b} \forall i \neq j$ and as such the faction equilibrium from Proposition 2 is no longer valid. The remainder of the proposition follows by substituting in our expression for w_i in U_i and taking derivatives. ■

The effect of the different factors on the degree of underbiasing is particularly interesting. First of all we see that the if the call for unity is larger, the underbiasing effect is larger. This follows from the fact that with a larger call for unity, the importance of being close to the preferred policy position is smaller which is captured by the factor $(1-\gamma)$. As such the cost component of the objective function becomes relatively more important and thus there is an incentive to decrease costs by applying a smaller bias. Furthermore we see that the applied bias gets smaller when costs are more sensitive to larger biases. This effect is obvious: since biasing costs are everywhere increasing in (the absolute value of) the bias we see that for an increase in sensitivity biasing costs increase for any given bias. This creates an incentive to reduce the introduced bias.

The effects of the communication and technocrat factors ξ_i and κ_i are not obvious. We see that these factors are introduced here via the weight w_i. The weight w_i is decreasing in both κ_i and ξ_i since high values for these factors imply low degrees of accuracy and clarity. A small value of w_i in turn implies that the particular leader has relatively liimited influence on the average party line since his signal is factored in to a lesser extent. As such, the bias introduced by this leader only has a small effect. The costs for this bias are however the same for this leader, independent of the influence of this bias. As such we see that the marginal benefit of biasing is small for a small w_i and therefore such leaders have an incentive to underbias.

Note that Proposition 4 only claimed that faction equilibria no longer *generally* exist. That is, it is no longer true that there is an equilibrium for all $\bar{b} \in \Re$ as long as $\bar{b} = \bar{b}_i$ for all $i \leq m$. It turns out that such an equilibrium can still exist for a particular value of $\bar{b}$. The following proposition states that such a faction equilibrium (that is an equilibrium in which $\bar{b}_i = \bar{b}$ for all $i \leq m$ and some $\bar{b} \in \Re$) can only exist in one special case:

Proposition 5 *With naive party activists and costly biasing there uniquely exists a faction equilibrium if and only if* $\bar{b}_1 = \bar{b}_2 = \cdots = \bar{b}_m = 0$*. In this equilibrium the activists strategies* $\{w_i, c_i, z_i\}$ *are as before, that is:*

$$w_{i\ell}^* = \frac{\psi_i}{\sum_j^n \psi_j}, \; c_\ell^* = 0, \; z_{i\ell}^* = \frac{\xi_i w_{i\ell}^*}{\sqrt{C'(z)}} \text{ with } \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}}$$

The party leaders' equilibrium strategies$\{b_i\}$ *are to set* $b_i = 0$ *for all* i.

Proof In the proof to Proposition 4 we saw that in a faction equilibrium it holds that $b_i = \bar{b} - U_i$. Set $\bar{b} = 0$ and $b_i = 0$. Given that $C'(0) = 0$ we have that for $b_i = 0$ it holds that $U_i = 0$. Thus $b_i = 0 + 0 = 0$ so the best response relation is satisfied for all $i \leq m$. Furthermore for all $i > m$ the choice of b_i is irrelevant since they are ignored. Therefore they minimize costs thus setting $b_i = 0$. Thus all leaders set $b_i = 0$ and this forms an equilibrium if and only if $\bar{b}_i = 0$ for all $i \leq m$ by the best response relation in equation 4.21. ■

Note that Proposition 5 makes no further claims about the existence of any other equilibria. The only claim that is made that a faction equilbrium uniquely exists if and only if $\bar{b}_i = 0$ for all $i \leq m$. This is in contrast with our findings in section 4.3 and Proposition 2. In particular, Proposition 5 has one major implication: the only (faction) dictatorship that can exist is the efficient one, namely the one which is completely unbiased. As such, the introduction of biasing costs has drastically changed the conclusion of our model. Without biasing costs any faction could be a dictatorship in the sense that the policy direction was completely determined by this faction, under the condition that the best communicators (all those that receive attention) beonged to the same faction. We observed that this could include arbitrarily 'extreme' factions (those with $|b_i| \gg 0$). The introduction of costs has eliminated this possibility: if a dictatorship exists, we have the efficient equilibrium where all leaders are unbiased. The communication constraints imposed upon the leaders in this equilibrium remain unchanged. Note that whilst this result is superficially similar to the one derived for perfect information, the efficiency implications of this finding are much greater since under perfect information any equilibrium is efficient, independent of the biases introduced by the leaders.

4.5 Equilibrium: Naive & Informed Activists

So far we have considered the two extreme cases in which all activists either posses perfect information or are completely naive. In this section we allow for the mass of activstis to consist of both types. Whilst the distinction into two classes of information is still heavily stylized, making this subdivision into two groups allows us to gain some intuition into how activists and leaders deal with heterogeneity within the group of activists.

We model this situation as follows. Consider two groups of activists. The first group of activists is perfectly informed and is denoted by I. The activists in this group are perfectly aware of the bias introduced by the party leaders. Furthermore, they are aware of the characteristics of the groups of activists. That is, they know the sizes of both groups and the information available to members of the groups. The second groups consists of naive activists and is denoted by N. The activists within this group are unaware of the biases introduced by the party leaders and are also unaware of the fact that activsts in the group differ with respect to their available information. That is, the naive activists think the entire unit mass consists of identical naive activists. Informed activsts are unable to communicate the bias to naive activists.

We denote the size of the informed group by π, where $\pi \in [0, 1]$. The size of the naive group is therefore $(1 - \pi)$. Activists' objective functions are identical to previous sections and we consider a case where biasing costs are non-existent (we use objective function v_i instead of v_i', in last section's notation). For convenience, note that the following holds in this scenario:

$$\bar{a} = \int_0^1 a_\ell d\ell = \int_0^\pi a_\ell d\ell + \int_\pi^1 a_\ell d\ell = \pi \bar{a}_I + (1 - \pi)\bar{a}_N \tag{4.22}$$

Here $\bar{a}_I$ and $\bar{a}_N$ denote the average policy directions within the two respective groups. That the same holds for the average weights w_i and average bias $\bar{c}_i$:

$$\bar{c}_i = \pi \bar{c}_{iI} + (1 - \pi)\bar{c}_{iN} \tag{4.23}$$

$$w_i = \pi w_{iI} + (1-\pi) w_{iN} \tag{4.24}$$

With this established, we can move on to establishing the equilibrium strategies. First consider the naive activsts. Since they are unaware of the distinction between the two groups their optimization problem is identical to the one solved in section 4.3. As such, equilibrium strategies directly follow from Proposition 2 and are given as follows (we surpress the ℓ in the subscript throughout this section for convenience:

$$w_{iN}^* = \frac{\psi_i}{\sum_j^n \psi_j},\ c_N^* = 0,\ z_{iN}^* = \frac{\xi_i w_{iN}^*}{\sqrt{C'(z)}}\ \textit{with}\ \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{iN}}} \tag{4.25}$$

Equilibrium strategies for informed activists are not trivial to derive. From section 4.2 we know they minimize the following:

$$\mathbb{E}(u_I) = \bar{u} - \mathbb{E}\left[(a_I - \theta)^2\right] - \mathbb{E}\left[(a_I - \bar{a})^2\right] - C(z)$$

Note that the first quadratic loss term is independent of the strategies of other activists, and as such the expectation is identical to the one derived for the proof of Proposition 1 (Equation 4.2):

$$\mathbb{E}\left[(a_I - \theta)^2\right] = \sum_i^n w_{iI}^2 \left(\kappa_i^2 + \frac{\xi_i^2}{z_{iI}}\right) + \left(\sum_i^n w_{iI} b_i + c_I\right)^2$$

The second term is affected by the change in composition of the activist mass. Earlier we had that $\bar{a} = \theta + \sum_i^n w_i (\eta_i + b_i) + \bar{c}$. Using equations 4.23 and equations 4.24 and noting that $\bar{c}_N = 0$ we have:

$$\bar{a} = \theta + \sum_i^n \left[\pi w_{iI} + (1-\pi) w_{iN}\right](\eta_i + b_i) + \pi \bar{c}_{iI}$$

Now by invoking symmetry within the group of informed activists (that is $a_{\ell I} = \bar{a}_I$ for all $\ell \in [0, \pi]$) we get:

$$a_{\ell I} - \bar{a} = \sum_i^n w_{iI} \epsilon_{iI} + \sum_i^n \left[(1-\pi) w_{iI} - (1-\pi) w_{iN}\right](\eta_i + b_i) + (1-\pi) c_I \tag{4.26}$$

Squaring and taking expectations gives the following:

$$\mathbb{E}\left[(a_{\ell I} - \bar{a})^2\right] = \sum_i^n w_{iI}^2 \frac{\xi_i^2}{z_{iI}} + \sum_i^n \left(\left[(1-\pi) w_{iI} - (1-\pi) w_{iN}\right]^2 (\kappa_i^2 + b_i^2)\right) + (1-\pi)^2 c_I^2 + 2(1-\pi)^2 c_I \sum_i^n \Big(b_i \left[(w_{iI} - w_{iN}\right]\Big) \tag{4.27}$$

Combining Equations 4.26 and 4.27 gives the final expression:

$$\mathbb{E}(u_\ell) = \bar{u} - \underbrace{\sum_i^n w_{iI}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{iI}}\right] + (1-\gamma)\left(\sum_i^n w_{iI} b_i + c_I\right)^2}_{L_1(z_I, w_I, c_I)} + \underbrace{\gamma\left[\sum_i^n \left(\left[(1-\pi) w_{iI} - (1-\pi) w_{iN}\right]^2 (\kappa_i^2 + b_i^2)\right) + 2(1-\pi)^2 c_I \sum_i^n \Big(b_i \left[w_{iI} - w_{iN}\right]\Big) + (1-\pi)^2 c_I^2\right]}_{L_2(w_I, w_N, c_I)} - C(z_I) \tag{4.28}$$

This leads naturally to the following lemma:

Lemma 5 *An informed activists strategy* $\{w_I, z_I, c_I\}$ *forms an equilibrium if and only if it solves:*

$$min\ L_1(z_I, w_I, c_I) + L_2(w_I, w_N, c_I) + C(z_I)\ subject\ to\ \sum_i^n w_{iI} = 1$$

Here $L_1(z_I, w_I, c_I)$ *and* $L_2(w_I, w_N, c_I)$ *are defined as in Equation 4.28. Furthermore the following holds:*

$$w^*_{iN} = \frac{\psi_i}{\sum_j^n \psi_j},\ c^*_N = 0,\ z^*_{iN} = \frac{\xi_i w^*_{iN}}{\sqrt{C'(z)}}\ with\ \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{iN}}}$$

Proof The exact expression follows from the discussion above. Given that $C(z_I)$ is convex (which it is by assumption) the loss function is convex in all its arguments and the unique (global) solution to this optimization problem is thus obtained from the relevant first order conditions. ■

Note that this time the 'coordination' term L_2 cannot be ignored. This is because the distinction between the two types has introduced an asymmetry. As a result of this, activists from the different groups no longer necessarily play the same strategy. Naive activists are however unaware of this and as such their strategy remains unchanged. Informed activsts do posses this information and as such will try to act on it, taking into account the asymmetry.

The introduction of this assymetry causes the solution of this minimization problem to be much more involved. In the proof of the following proposition we will see that the solution involves solving a system of n linear equations. Unfortunately, we are unable to arrive at a generaly analytical expression for such a solution. Therefore we will consider the simplest case where there are two leaders ($n = 2$). From there we arrive at the following solution:

Proposition 6 *With informed and uninformed activists the unique equilibrium informed activists strategy* $\{z_I, w_I, c_I\}$ *is given by the following:*

$$z^*_{iI} = \frac{\xi_i w^*_{iI}}{\sqrt{C'(z)}},\ c^*_I = \omega - \left(\sum_i^n w_{iI} b_i\right)\ with\ \omega = \left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] \tag{4.29}$$

The last element of the strategy w_I *is given by the solution* w^* *to the system* $\Phi w = \Omega$ *which is as defined in the proof of the proposition. For the case* $n = 2$ *this solution is given by:*

$$w^*_{1I} = \frac{\phi_1\phi_2^2(1-\pi)^2 b_1 b_2 + \phi_1^2 + \phi_1\phi_2 + \phi_1^2\phi_2^2(1-\pi)^2 b_1 b_2 + (\phi_1^2\phi_2 + \phi_1\phi_2^2)[\gamma(1-\pi)(\kappa_2^2 - \kappa_1^2 + b_2^2 - b_1^2) + (1-\pi)^2(b_2 - b_1)\omega]}{\phi_1^2 + \phi_2^2 + 2\phi_1\phi_2 + 4\phi_1\phi_2(\phi_1+\phi_2)(1-\pi)^2 b_1 b_2} \tag{4.30}$$

Where ϕ_i *is defined as follows:*

$$\phi_i = \frac{1}{\left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{iI}} - (1-\pi)^2 b_i\right]} \tag{4.31}$$

Proof The full proof of the proposition can be found in Appendix B.

Let us analyze the expressions in Proposition 6 further. First of all, note that in the case of no uninformed activists ($\pi = 1$) we have $\omega = 0$. Furthermore, all the $(1-\pi)$ terms drop out and thus our expressions reduce to those from Proposition 1. This is obvious for c_I and z_I. For w_I it can be seen by observing that for $\pi = 1$ we have:

$$w_{iI} = \frac{\phi_1^2 + \phi_1\phi_2}{(\phi_1+\phi_2)^2} = \frac{\phi_1}{\phi_1+\phi_2}$$

With $\pi = 1$ we also have that $\phi_i = \psi_i$ and thus the claim follows. We thus see that Proposition 6 generalizes Proposition 1. As such, it is interesting to contrast our general results from Proposition 6 with the specific results for when $\pi = 1$ given in Proposition 1.

First of all, we analyze the expected policy advocated by informed activists. Since they advocate according to an affine strategy the advocated strategy is equivalent to $a_{\ell I} = w_I x_\ell + c_I$. Under the equilibrium strategy this yields the following:

$$\mathbb{E}(a_{\ell I}) = \mathbb{E}(w_I x_\ell + c_I) = \theta + \left(\sum_i^n w_{iI} b_i\right) i + \omega - \left(\sum_i^n w_{iI} b_i\right) = \theta + \omega$$

Thus we see that informed activsts completely eliminate the bias introduced by their own information acquisition and weighting policies. They do however also introduce a certain bias which is dependent on the bias introduced by uninformed activists. This bias is given by ω. Inspecting the full expression for ω provides some insights into the nature of this bias:

$$\omega = \left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]} \frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] = \left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]} \sum_i^n w_{iN} b_i\right]$$

Thus we see that the informed activists' strategy involves introducing an optimal bias that is a fraction of the bias introduced by the naive activsts. This fraction depends on both the preference for unity γ and the percentage of informed activists π. This behaviour is captured in figure 1 below. Here p indicates the proportion of the bias that is introduced by naive activists that is also introduced by informed activists:

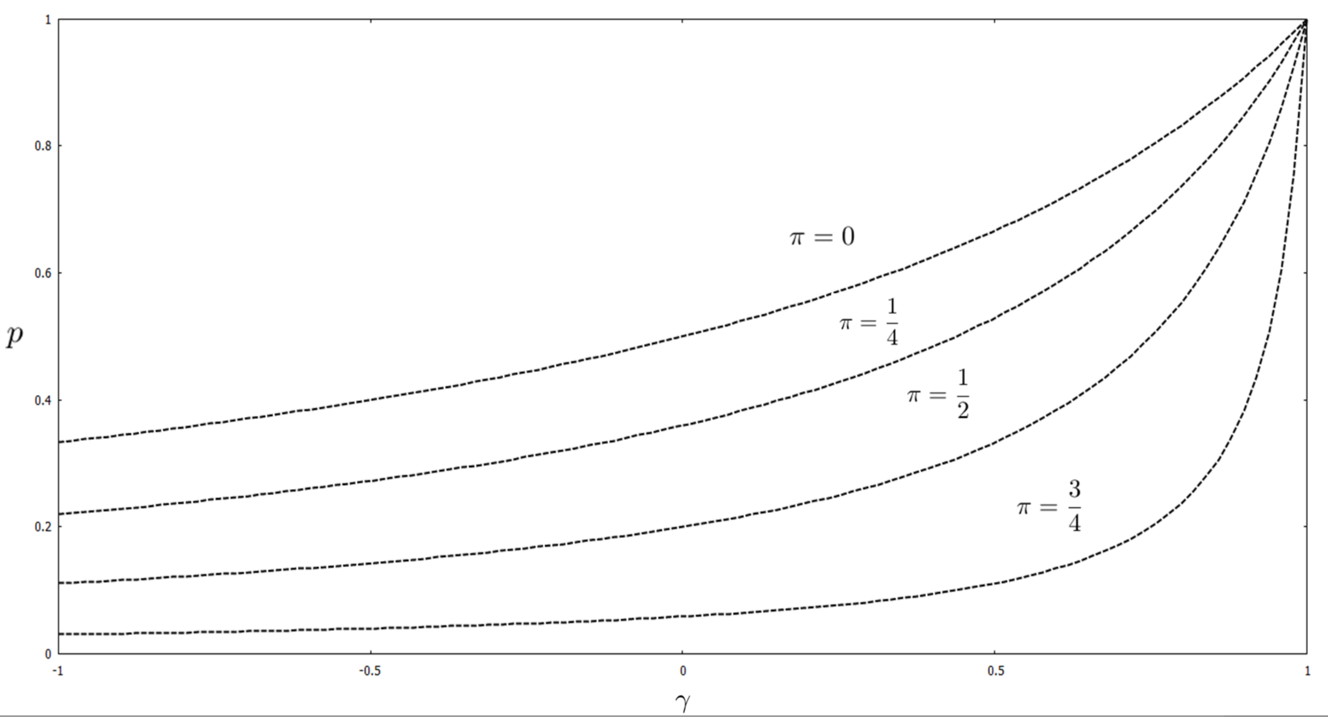

Figure 1 *Proportion of the bias introduced by naive activists that is also introduced by informed activists as a function of concern for unity γ plotted for various levels of π. For $\pi = 1$ the proportion is 0 for all values of γ.*

In particular, consider the situation where $\gamma = 1$. In that case there is only a concern for unity and no concern for the average policy advocated. As such we get that $\omega = \sum_i^n w_{iN} b_i$. That is, informed activists introduce a bias identical to the one introduced by individual activists. As γ decreases, the fraction of the bias taken over by informed activists decreases. This is because as γ decreases unity becomes less important whilst the concern for advocating the correct policy increases. As π increases, the proportion of the bias included by informed activists decreases towards 0. This is because for higher levels of π the average policy is influenced by a proportionally larger mass of informed activists. So each individual informed activists is closer to the average for a higher level of π. As such, the marginal disutility avoided by adjusting to the naive activists is smaller for π large, whilst the marginal disutility of deviation from the optimal policy θ is unchanged. As a result it is optimal for informed activists to advocate a policy with a smaller proportional bias. All in all it is striking that the coordination motive generally dominates the party position motive, as we see that the bias introduced by naive activists is never completely filtered out by informed activsists. This follows from our plot where we see that $p \geq 0$

One final and crucial implication of Proposition 6 is that the weights placed on individual signals and therefore also the information acquisition policy are no longer independent of the biases introduced by individual leaders. This follows from the proposition by noting that both w_{iI} and thus also z_{iI} are functions of b_i. The effect of b_i on w_{iI} is not obvious. Taking derivates shows that this effect can be negative and positive depending on the particular values of b_j for $j \neq i$, π and γ. Still, this finding stands in stark contrast to our results for the completely naive and fully informed cases considered in previous sections. The distinction between these scenarios and the mixed scenario considered here of course is the breakdown of symmetry we observe in this particular case.

The intuition behind this finding is simple: in a symmetric equilibrium all activists weight the received signals equally and thus biases are also weighted identically for all activists. In the asymmetrical case informed activists weight the signals differently from naive activists. Thus the bias is transferred differently for these types of activsts. Informed activists must take this into account when trying to maintain unity in the party, and thus the individual biases start to directly affect the weigthing process. It is therefore by the heterogeneity of the activists that biases affect the information acquisition and weighting process. Note however that the weights introduced by the naive activists crucially determine the party line, since informed activsts perfectly filter out their own bias and take over some proportion of the activists' bias.

Finally, we consider the party leaders' optimal bias. In order to do so we first derive the expression to be maximized by the party leaders. They maximize expected utility which is given by:

$$\mathbb{E}\left[v_i\right] = \bar{v} - \mathbb{E}\left[(1-\gamma)(\bar{a} - \bar{b}_i - \theta)^2\right] - \mathbb{E}\left[\gamma \int_0^1 (a_\ell - \bar{a})^2 d\ell\right]$$

Note that the error terms dissapear by the law of large numbers. By plugging in our expression for c_I^* and w_{iN}^* we have the following expression for $\bar{a}$:

$$\bar{a} = \theta + \left[(1-\pi) + \frac{\pi\left[1-\pi\right]^2}{\left[(1-\gamma)+(1-\pi)^2\right]}\right] \frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}$$

Squaring and taking expectations we thus get:

$$\mathbb{E}\left[(1-\gamma)(\bar{a} - \bar{b}_i - \theta)^2\right] = (1-\gamma)\left(\left[(1-\pi) + \frac{\pi\left[1-\pi\right]^2}{\left[(1-\gamma)+(1-\pi)^2\right]}\right] \frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i} - \bar{b}_i\right)^2 \tag{4.32}$$

Next we rewrite the second quadratic loss term as follows (using the fact that within the two subgroups the expectation term is independent of the particular label ℓ):

$$\mathbb{E}\left[\gamma\int_0^1 (a_\ell-\bar{a})^2 d\ell\right] = \gamma\pi\mathbb{E}\left[(a_{\ell I}-\bar{a})^2\right] + \gamma(1-\pi)E\left[(a_{\ell N}-\bar{a})^2\right]$$

Now we have that:

$$a_{\ell I}-\bar{a} = (1-\pi)a_{\ell I} - (1-\pi)a_{\ell N} = (1-\pi)\left[\omega - \frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] = (1-\pi)\left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]} - 1\right]\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}$$

Squaring and taking expectations once more yields (after some rewriting):

$$\mathbb{E}\left[(a_{\ell I}-\bar{a})^2\right] = \left[\frac{(1-\pi)(1-\gamma)}{[(1-\gamma)+(1-\pi)^2]}\right]^2 \left(\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right)^2$$

Performing the same steps for the naive activists yields:

$$\mathbb{E}\left[(a_{\ell N}-\bar{a})^2\right] = \left[\frac{\pi(1-\gamma)}{[(1-\gamma)+(1-\pi)^2]}\right]^2 \left(\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right)^2$$

Combining these last two expressions yields the final quadratic loss term:

$$\mathbb{E}\left[\gamma\int_0^1 (a_\ell-\bar{a})^2 d\ell\right] = \gamma\left[\frac{\pi(1-\pi)^2(1-\gamma)^2 + (1-\pi)\pi^2(1-\gamma)^2}{[(1-\gamma)+(1-\pi)^2]^2}\right]\left(\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right)^2 \tag{4.33}$$

Finally combining Equation 4.32 and 4.33 we get our final expression for the expected utility to be maximized:

$$\mathbb{E}[v_i] = \bar{v} - \underbrace{(1-\gamma)\left(\left[(1-\pi) + \frac{\pi[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\right]\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i} - \bar{b}_i\right)^2}_{L_1(b_i,b_j)} - \underbrace{\gamma\left[\frac{\pi(1-\pi)^2(1-\gamma)^2 + (1-\pi)\pi^2(1-\gamma)^2}{[(1-\gamma)+(1-\pi)^2]^2}\right]\left(\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right)^2}_{L_2(b_i,b_j)} \tag{4.34}$$

Note that the weights for the informed activists do not appear in this expression. This is because the effect of the weights on the bias is cancelled out by the affine term c_I in the informed activists' strategies. Simultaneously, the individual noise terms ϵ_i dissapear by the law of large numbers. As such, the individual equilibrium weights for informed activsts have no effect on the average party line and/or party disunity directly. We can now straightforwardly solve for the optimal b_i as by the following lemma.

Lemma 6 *A party leaders strategy* $\{b\}$ *forms and equilibrium strategy if and only if it solves:*

$$min\ L_1(b_i,b_j) + L_2(b_i,b_j)\ s.t.\ \forall i$$

Proof The expression follows from the discussion above. L_1 and L_2 are convex in b_i and thus the solution to the optimization problem is derived from the first-order conditions. ■

The following proposition follows directly from the above lemma:

Proposition 7 *With informed and uninformed activists the only equilibrium is a faction dictatorship. That is, for all leaders $i \leq m$ for which $z_i > 0$ we have that $\bar{b}_i = \bar{b}$ for some $\bar{b} \in \Re$. All leaders $i > m$ are indifferent between any $b_i \in \Re$. For all leaders $i \leq m$ the equilibrium bias is given by the following:*

$$b_i^* = \frac{\Lambda_1}{\Lambda_1 + \Lambda_2} \bar{b} \tag{4.35}$$

Here Λ_1 and Λ_2 are defined in the proof of the proposition.

Proof The proof is given in Appendix C.

As we can see from Proposition 6, the coordination motive for leaders prevents them from introducing a bias equal to that of their optimal policy preference. Specifically, they set a bias that is a proportion of their optimal policy preference. This proportion is given by the fraction $\frac{\Lambda_1}{\Lambda_1+\Lambda_2}$. As such, the proportion is fully determined by the call for unity γ and the proportion of informed activists π. Figure 2 below shows how this proportion varies with γ for various levels of π:

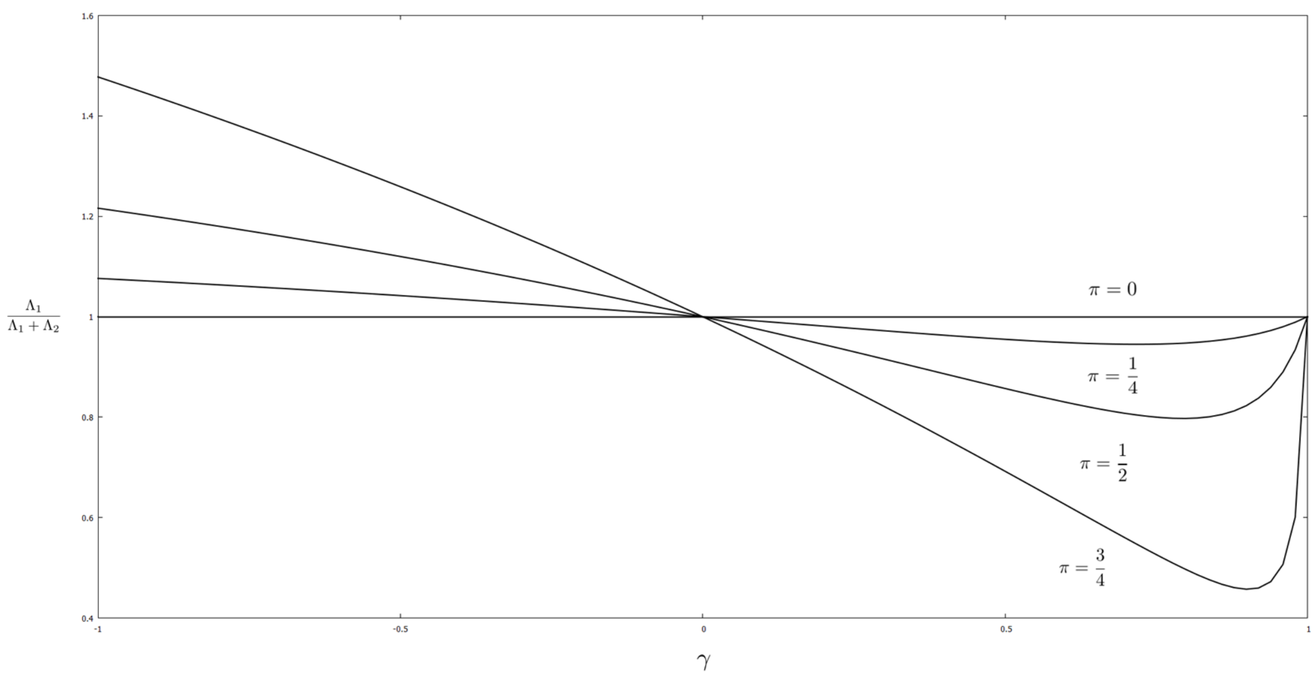

Figure 2 *Equilibrium bias b_i expressed as a proportion $\frac{\Lambda_1}{\Lambda_1+\Lambda_2}$ of the preferred policy position $\bar{b}_i$ as a function of concern for unity γ plotted for various levels of π. For $\pi = 1$ the proportion is not defined for any $\gamma \in [0, 1]$.*

As we can see, for each level of π (except for $\pi = 0$) we see a similar relation between the proportion of $\bar{b}$ that is included and γ, albeit more extreme for larger values of π. In particular, for $\gamma < 0$ we see that a larger bias is included. That is, $\frac{\Lambda_1}{\Lambda_1+\Lambda_2} > 1$. This is because over this range utility is decreasing in unity. For any values of γ and π the bias introduced by informed activists is smaller than that introduced by uninformed activists, and this difference is increasing in the size of the bias. As such, the larger the bias, the smaller the unity since there is a larger difference in biases between the two types. As a result introducing a larger bias increases disunity and thus utility. On the other hand, for $\gamma < 0$ getting the average party line close to the preferred policy $\theta + \bar{b}$ is of larger importance to the party leaders. As such, the bias cannot become too large. This trade-off results in an optimal bias that is found somewhere in the range $[1, 1.5]$ for $\gamma < 0$.

For $\gamma = 0$ party unity is of no importance, and utility is fully determined by the closeness of the advocated policy to the preferred policy. As such, party leaders directly communicate their preferred position and $b_i = \bar{b}$. For $\gamma > 0$, leader utility is increasing in party unity whilst advocating the perfect policy position becomes less important. As such, we see that the bias starts to decrease initially as the unity term is decreasing in the size of the bias. As $\gamma \to 1$ unity becomes more important, and in the limit only unity matters with the distance term dropping out. We see however from figure 1 that as $\gamma \to 1$ the bias added by informed and naive activists come closer together with them being identical in the limit. Party leaders react optimally to this by bringing the bias closer to their optimal position.This is possible since for a certain γ the marginal benefit of having an average policy position closer to the preferred policy starts to dominate the marginal utility gained from maintaining unity. This happens due to the convergence in the biases of naive and informed leaders as γ increases. The point at which this occurs is the one where $\frac{\Lambda_1}{\Lambda_1+\Lambda_2} > 1$ increases again. This point can be found for a particular value of π by solving the following:

$$\frac{\partial \left[\frac{\Lambda_1}{\Lambda_1+\Lambda_2}\right]}{\partial \gamma} = 0$$

The effect of π on these dynamics is obvious. A higher level of informed activists requires the leaders to introduce a larger bias, since informed activists partially compensate for the bias. As such, a larger bias is optimal to counteract this. This applies to $\gamma < 0$ and $\gamma > 1$. Note in particular from figure 1 that the convergence in the biases for informed and naive activists is weaker for π large, and as such the point at which $\frac{\Lambda_1}{\Lambda_1+\Lambda_2}$ starts increasing again occurs at a higher level of γ. The $\gamma = 0$ case naturally remains unaffected. For $\pi = 0$ we have the case of section 4.2 where we only had naive activists. We saw there that in that case leaders introduce a bias exactly equal to their optimal bias. This is exactly what we see from the graph for $\pi = 0$ as the proportion is everywhere equal to 1.

Proposition 7 concludes our analysis of the mixed case with both informed and naive activists. Proposition 6 and 7 together fully describe the equilibrium under these circumstances. This section also marks the last setting of our analysis, and we now move on to a short summary and discussion of the results as well as a conclusion.

Chapter 5

Discussion & Conclusion

In this thesis we have analyzed an endogenous information acquisition process under several informational assumptions. We found that under perfect information, clarity of communication is still more important than accuracy. That is, the communication skill of a leader is more important than his skills as a technocrat. In fact, whether or not a leader is ignored is fully determined by his clarity of communication. Furthermore we have found that perfectly informed activists perfectly compensate for the bias introduced by individual leaders, thus eliminating it completely. As a result, leaders are indifferent between any bias and thus we have infinitely many equilibria. If we introduce biasing costs we limit this to a unique equilibrium in which leaders choose not to bias at all. As such we see that the outcomes of this scenario are identical to those found in Myatt and Wallce (2011). This implies that their comparative statics results also apply here.

With naive activists we find that the information acquisition and weighting choice is identical to the one under perfect information. In this scenario however, party activists do not compensate for the bias at all and as such the bias is perfectly reflected in the advocated policy position. As a result of this, party leaders can now directly affect the average policy. We then found that an equilibrium can only exists if all leaders that receive attention (those with the highest clarity) share the same optimal policy position. As such, the only possible equilibrium is a faction dictatorship where only one faction receives attention and fully determines the policy to be advocated. When biasing costs are introduced the only faction equilibrium that can still exist is the one in which the bias is equal to zero. However, it is possible that other equilibria might exist. In these equilibria the biasing costs cause party leaders to underbias with respect to their optimal policy position.

With both naive and informed activists, naive activists have the exact same behaviour as before. For informed activists, equilibrium behaviour is in fact highly different. In particular, we see that, unlike in the previous two settings, the information acquisition and weighting process is no longer independent of the biases introduced by the leaders. In particular, each individual weight placed on an information source is a function of both the bias of the leader that is weighted and the other leaders. How this weight is affected by the bias depends on the importance of unity and the proportion of activists that are informed. Another difference is found in the way they compensate for the included bias. They still compensate completely for the bias that they themselves introduce through their information acquisition process, however informed activists now also introduce a bias which is a proportion of the bias added by the naive activists. This is a result of wanting to maintain unity within the party, and we do indeed see that the proportion is increasing in the desire for unity. For the leaders we see that the only equilibrium remains a faction equilibrium where all leaders that receive attention share the same optimal position. They over or understate their optimal policy position, depending on the concern for unity and the proportion of informed activists. This is because these parameters affect the trade-off between coordination and advocating the optimal position.

We would like to highlight the implications of two specific results. First of all, we have found that in the last two cases the only equilibrium is a faction dictatorship. A crucial feature of these equilibria is that the existence of these equilibria is independent of the extremity of the faction's policy position $\bar{b}$. Thus we find that a situation in which the optimal bias is very large is just as valid of an equilibrium as one with a small bias. In fact, the only condition is that these extreme leaders are also the best communicators such that they receive all the attention. This thus implies that an extremist dictatorship can result from rational decision making as long as the leaders of the extremist faction are charismatic and effective communicators. Whilst such a result might be expected with completely naive activists, it is interesting that this result is maintained even if we have a large proportion of informed activists. This can be seen directly from Figures 1 and 2 (consider for example the graphs for $\pi = \frac{3}{4}$ in both). This effect is dampened somewhat for large values of γ, as leaders then optimally reduce their bias to maintain unity within the party. This effect only occurs over a limited range for γ though, as we see that for γ close to 1 we once again see large biases in the faction equilibrium. This result thus indicates that even a small group of naive activists can cause extremist factions to dominate a party.

The second result that we wish to discuss is the relation between clarity of communication and bias and how they determine whether or not a leader receives attention. In our first two cases with only naive and only perfectly informed activists we saw that the attention paid to a certain leader was not affected by the bias introduced by this particular leader. This result is intuitively obvious: in the perfectly informed case, activists can completely compensate for biases and as such they acquire information as if there were no biases. In the other case, activists are unaware of these biases and thus it is impossible for them to influence the information acquisition process. As we mentioned above, this independence breaks down in our heterogenous final case. Here naive activists still do not take biases into account in the information acquisition process (after all, these activists are still unaware of their existence). Informed activists however, do take biases into account when acquiring information. This is clear from the fact that w_I, and thus z_I, are now functions of b_i and b_j. However, as informed activists perfectly compensate for the bias that are introduced through their own weights, we see that the weights that determine the bias are those introduced by naive activists. As such, the overall bias is eventually fully determined only by these weights. Since these weights are functions only of the clarity of the individual leaders, the independence is to a certain extent maintained. In particular, it means that the overall party line is still only determined by clarity of communication, the concern for unity, and the proportion of informed activists.

Finally, we would like to highlight some possible extentions that may improve on some of the limitations of this thesis and may thus be fruitful directions for future research. In particular, we would like to emphasize the mayor weakness of our model. In our analysis, we have only allowed for party activists to be either completely naive or perfectly informed. Naturally, we can imagine an entire spectrum in between. For example, an activist might be aware that a bias exists without knowing specifically what the bias may be. Alternatively, activists may be able to perceive the bias only with a certain noise. As such, we can imagine that instead of having only groups of perfectly informed and naive activists, we have a certain distribution over our unit mass of activists. This distribution assigns to each activists some kind of accuracy with which they can perceive the bias. Perfectly informed activists have perfect accuracy whilst naive activists only perceive complete noise. This setting can be used to generalize any clustering of groups with specific informational attributes. Additionally one could introduce other dynamics, such as that the amount of attention that is paid to a leader increases the accuracy with which that leader's individual bias can be perceived. Such a model would remove much of the artfificiality from our current model and allow for many interesting generalizations. As such, it would be an interesting direction for further research.

References

Literature

- Allen, F., Morris, S., & Shin, H. (2006). *Beauty Contests and Iterated Expectations in Asset Markets.* Review of Financial Studies, 720752.

- Amato, J. Morris, S., & Shin, H (2002). *Communication and Monetary Policy.* Oxford Review of Monetary Policy, 671-688.

- Angeletos, A., & Pavan, A. (2007). *Efficient Use of Information and Social Value of Information.* Econometrica, 1103-1142.

- Angeletos, G. M., & Pavan, A. (2004). *Transparency of Information and Coordination in Economies with Investment Complementaries.* American Economic Review, 91-98.

- Cover, T. M., & Thomas J. (2006). *Elements of Information Theory.* 2nd Ed, New York: Wiley.

- Cremer, J. (1990). *Common Knowledge and the Co-ordination of Economic Activities.* in Aoki, M., Gustafsson B. & Williamson, O. E. *The Firm as a Nexus of Treaties*, Chapter 3, London: Sage, 53-76.

- Cremer, J. (1993), *Corporate Culture and Shared Knowledge.* Industrial and Corporate Change, 351-386.

- Dewan, T., & Myatt, D. P. (2008). *The Qualities of Leadership: Direction, Communication, and Obfuscation.* American Political Science Review, 351-368.

- Dominguez-Martinez, S., Swank O. H., & Visser, B., (2008), *In Defense of Boards, Journal of Economics and Management Strategy.*, 667-682.

- Hellwig, C. (2005). *Heterogeneous Information and the Welfare Effects of Public Information Disclosure.* Unpublished Manuscript, University of California, Los Angeles .

- Hellwig, C., & Veldkamp, L. (2009). *Knowing What Others Know: Coordination Motives in Information Acquisition.* Review of Economic Studies, 223-251.

- Keynes, J. M. (1936). *The State of Long-Term Expectations.* In J. M. Keynes, *The General Theory of Employment, Interest and Money.*147-165. New York: Harcourt Brace and Co.

- Mackay, D. J. C., (2003), *Information Theory, Inference, and Learning Algorithms.* Cambridge: Cambridge University Press.

- Mackowiak, B. & Wiederholt, M. (2009), *Optimal Sticky Prices under Rational Inattention.* American EconomicReview, 769803.

- Marschak, J. & Radner, R. (1972), *Economic Theory of Teams.* New Haven: Yale University Press

- Morris, S., & Shin, H. S. (2002). *Social Value of Public Information.* American Economic Review , 15211534.
- Myatt, D. P., & Wallace, C. (2011). *Endogenous Information Acquisition in Coordination Games.* Review of Economic Studies, 340-374.
- Persico, N. (2004) *Committee Design with Endogenous Information.*, Review of Economic Studies, 165-91.
- Sims, C. A. (1998), *Stickiness.*, Carnegie-Rochester Conference Series on Public Policy, 317-356.
- Sims, C. A. (2003). *Implications of Rational Inattention.* Journal of Monetary Economics, 665-690.
- Sims, C. A. (2005), *Rational Inattention: A Research Agenda.* Unpublished Manuscript, Princeton University.
- Sims, C. A. (2006), *Rational Inattention: Beyond the Linear-Quadratic Case.* American Economic Review: AEA Papers and Proceedings, 158-163.
- Sims, C. A. (2011), *Rational Inattention and Monetary Economics.* in Friedman, B. A. and Woodford, M. *Handbook of Monetary Economics, Vol 3A* North-Holland: Amsterdam: Elsevier 155-182
- Swank, J., Swank, O.H., Visser, B. (2008) *Committees of Experts Interact with the Outside World: Some Theory, and Evidence from the FOMC.* Journal of European Economic Association, 478-486.

Software

- Maxima.Sourceforge.Net. Maxima, A Computer Algebra System
  Version 5.31.2 (2013). http://maxima.sourceforge.net/
- Williams, T. and Kelley, C. Gnuplot, An Interactive Plotting Program
  Version 4.65 (2014). http://gnuplot.info/

Appendices

Appendix A: Derivation of No Attention Condition

In section 4.3 we claim that a leader is ignored if $\xi_i \geq K_i$ with K_i equal to the following:

$$K_i \equiv \frac{1}{\sqrt{\frac{\partial C(z)}{\partial z_i} \sum_{j=1}^{n} \psi_j}}$$

We now derive this condition. From Proposition 1 we get the following:

$$w_{i\ell}^* = \frac{\psi_i}{\sum_j^n \psi_j},\ z_{i\ell}^* = \frac{\xi_i w_{i\ell}^*}{\sqrt{C'(z)}}\ \textit{with}\ \psi_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}}$$

Plugging the expression for $w_{i\ell}$ into the expression for $z_{i\ell}$ yields the following:

$$z_{i\ell} = \frac{\xi_i \psi_i}{\sqrt{C'(z)} \sum_j^n \psi_j} = \xi_i \psi_i K_i$$

Rewriting:

$$\frac{z_{i\ell}}{\psi_i} = z_{i\ell}(1-\gamma)\kappa_i^2 + \xi_i^2 = \xi_i K_i \Rightarrow z_{i\ell} = \frac{\xi_i\left(K_i - \xi_i\right)}{(1-\gamma)\kappa_i^2}$$

Thus we have that $z_i \leq 0$ for $\xi_i \geq K_i$. Since we do not allow for negative attention we set $z_i = 0$ for all such x_i instead. This shows that the condition mentioned in the text holds.

Note that, interestingly, this expression shows that leaders can compensate for being bad technocrats (or κ_i large) by being good communicators (ξ_i small). To see this observe that we have that $K_i \to 0$ as $\kappa_i \to \infty$. So consider a situation in which $\kappa_i \gg 0$ but still finite. For any finite but large value of κ_i a leader still receives attention if $\xi_i < K_i$. So as long as ξ_i remains sufficiently small the worst technocrat can still receive a large amount of attention. The argument is symmetric in the sense that technically the same can be said in reverse, however the effect is dampened through the square of κ_i and the factor $(1-\gamma)$.

Appendix B: Proof of Proposition 6

Proof We once again solve this minimization problem using the method of Lagrange multipliers. This yields the following Lagrangian:

$$\mathcal{L}(w_I, z_I, c_I, \lambda) = L_1(z_I, w_I, c_I) + L_2(w_I, w_N, c_I) + C(z_I) + \lambda\left(\sum_i^n w_{iI} - 1 = 0\right)$$

Taking partial derivatives gives the following set of first-order conditions:

$$\frac{\partial \mathcal{L}(w_I, z_I, c_I)}{\partial w_{iI}} = 2w_{iI}\left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{iI}}\right] + 2(1-\gamma)\left(\sum_i^n w_{iI}b_i + c_I\right) + 2\gamma(1-\pi)(\kappa_i^2 + b_i^2) + 2(1-\pi)^2 c_I b_i + \lambda = 0$$

$$\frac{\partial \mathcal{L}(w_I, z_I, c_I)}{\partial z_{iI}} = -\frac{\xi_i^2 w_{iI}^2}{z_{iI}^2} + \frac{\partial C(z_I)}{\partial z_{iI}} = 0$$

$$\frac{\partial \mathcal{L}(w_I, z_I, c_I)}{\partial c_I} = 2(1-\gamma)\left(\sum_i^n w_{iI}b_i + c_I\right) + 2(1-\pi)^2\sum_i^n\Big(b_i\left[w_{iI} - w_{iN}\right]\Big) + 2(1-\pi)^2 c_I = 0$$

$$\frac{\partial \mathcal{L}(w_I, z_I, c_I)}{\partial \lambda} = \sum_i^n w_{iI} - 1 = 0$$

Note that the first order condition for z_I is identical to the one found in the proof of Proposition 1. As such we find that z_I is defined identically as before (noting that the w_I term that appears in the expression for z_I is possibly different).This yields the first component of the proposition. Equation 3 yields, upon rearranging:

$$c_I = \left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\left(\sum_i^n w_{iN}b_i\right)\right] - \left(\sum_i^n w_{iI}b_i\right) \tag{5.1}$$

Substituting the expression for w_{iN} yields the second part of the proposition:

$$c_I = \left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] - \left(\sum_i^n w_{iI}b_i\right) \tag{5.2}$$

By substituting this expression in the first condition above and rewriting we arrive at the following, where K is some scalar:

$$\begin{aligned} w_{iI}\left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{iI}} - (1-\pi)^2 b_i\right] + (1-\gamma)\left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] + \\ \gamma(1-\pi)(\kappa_i^2+b_i^2) + (1-\pi)^2 b_i\left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] - (1-\pi)^2 b_i\left(\sum_{i\neq j}^n w_{jI}b_j\right) = K \end{aligned} \tag{5.3}$$

Let us now define the following:

$$\phi_i = \frac{1}{\left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{iI}} - (1-\pi)^2 b_i\right]} \tag{5.4}$$

$$\Delta_i = (1-\gamma)\left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] + \gamma(1-\pi)(\kappa_i^2+b_i^2) + (1-\pi)^2 b_i\left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right] \tag{5.5}$$

$$\zeta_i = (1-\pi)^2 b_i \tag{5.6}$$

Substituting these expressions in equation 5.3 yields the following:

$$w_{iI}\frac{1}{\phi_i} + \Delta_i - \zeta_i\left(\sum_{i\neq j}^n w_{jI}b_j\right) = K \tag{5.7}$$

Rearranging to arrive at an expression for w_{iI} we thus get:

$$w_{iI} = \phi_i \left[K - \Delta_i + \zeta_i \left(\sum_{i \neq j}^{n} w_{jI} b_j \right) \right] \tag{5.8}$$

Now by our final condition we have that $\sum_i^n w_{iI} = 1$ such that the following holds:

$$K \sum_i^n \phi_i - \sum_i^n \Delta_i \phi_i + \sum_i^n \left[\phi_i \zeta_i \left(\sum_{i \neq j}^{n} w_{jI} b_j \right) \right] = 1$$

Solving for K we thus get:

$$K = \frac{1 + \sum_i^n \Delta_i \phi_i - \sum_i^n \left[\phi_i \zeta_i \left(\sum_{i \neq j}^{n} w_{jI} b_j \right) \right]}{\sum_i^n \phi_i}$$

Substituting back for K in our expression for w_{iI} we get:

$$w_{iI} = \frac{\phi_i}{\sum_i^n \phi_i} + \frac{\phi_i \sum_i^n \Delta_i \phi_i}{\sum_i^n \phi_i} - \frac{\phi_i \sum_i^n \left[\phi_i \zeta_i \left(\sum_{i \neq j}^{n} w_{jI} b_j \right) \right]}{\sum_i^n \phi_i} - \phi_i \Delta_i + \phi_i \zeta_i \left(\sum_{i \neq j}^{n} w_{jI} b_j \right) \tag{5.9}$$

For simplicity we write the following for the term in Equation 5.9 above:

$$\frac{\phi_i \sum_i^n \left[\phi_i \zeta_i \left(\sum_{i \neq j}^{n} w_{jI} b_j \right) \right]}{\sum_i^n \phi_i} = \frac{\phi_i w_{i_I} b_i \left(\sum_{j \neq i}^{n} \phi_j \zeta_j \right)}{\sum_i^n \phi_i} + \frac{\phi_i \sum_i^n \phi_i \zeta_i}{\sum_i^n \phi_i} \left(\sum_{i \neq j}^{n} w_{jI} b_j \right)$$

Rewriting equation 5.9 using this form and solving for w_{iI} once more yields the following specification for w_{iI}:

$$w_{iI} = \frac{\sum_i^n \phi_i}{\sum_i^n \phi_i + \phi_i b_i \left(\sum_{j \neq i}^{n} \phi_j \zeta_j \right)} \left[\frac{\phi_i}{\sum_i^n \phi_i} + \frac{\phi_i \sum_i^n \Delta_i \phi_i}{\sum_i^n \phi_i} - \phi_i \Delta_i + \left(\phi_i \zeta_i - \frac{\phi_i \sum_i^n \phi_i \zeta_i}{\sum_i^n \phi_i} \right) \left(\sum_{i \neq j}^{n} w_{jI} b_j \right) \right] \tag{5.10}$$

Note that this equation determines a system of n equations in our n unknown weights w_{iI}. After making the appropriate substitutions each equation in this system can be written in the following form:

$$w_{iI} = \Omega_i + \Phi_i b_1 w_1 + \Phi_i b_2 w_2 + \cdots + \Phi_i b_{i-1} w_{i-1} + \Phi_i b_{i+1} w_{i+1} + \cdots + \Phi_i b_n w_n \tag{5.11}$$

In matrix form, the system has the form $\Phi w = \Omega$ where w is our vector of weights and Ω is an $n \times 1$ vector with the ith entry equal to $-\Omega_i$. Furthermore, Φ is an $n \times n$ matrix which is as defined as follows:

$$\Phi = \begin{pmatrix} -1 & \Phi_1 b_2 & \cdots & \Phi_1 b_n \\ \Phi_2 b_1 & -1 & \cdots & \Phi_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_n b_1 & \Phi_n b_2 & \cdots & -1 \end{pmatrix} \tag{5.12}$$

In principle, this sytem of equation can be solved using Gauss/Gauss-Jordan elimination or by finding the inverse of the above matrix. In that case the solution would be $w = \Phi^{-1}\Omega$. For example, for $n = 2$ our solutions are given below.[1] For $n > 2$ the solutions quickly become a lot less tractable

$$w_{1I} = \frac{\Phi_i b_2 \Omega_2 + \Omega_1}{\Phi_1 \Phi_2 b_1 b_2 - 1}, \quad w_{2I} = \frac{\Phi_2 b_1 \Omega_1 + \Omega_2}{\Phi_1 \Phi_2 b_1 b_2 - 1}$$

[1] Maxima was used to compute this solution (see software list for further details).

Reversing our substitutions for the $n = 2$ case and simplifying yields:

$$w_{1I} = \frac{\phi_1\phi_2^2\zeta_1 b_2 + \phi_1^2 + \phi_1\phi_2 + \phi_1^2\phi_2^2\zeta_1 b_2 + (\phi_1^2\phi_2 + \phi_1\phi_2^2)(\Delta_2 - \Delta_1)}{\phi_1^2 + \phi_2^2 + 2\phi_1\phi_2\left[1 + \phi_1\zeta_2 b_1 + \phi_2\zeta_2 b_1 + \phi_1\zeta_1 b_2 + \phi_2\zeta_1 b_2\right]} \tag{5.13}$$

We now resubstitute our expressions for Δ_i and ζ_i and we make the following substitution:

$$\omega = \left[\frac{[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i}\right]$$

If we do so we arrive at the following expression:

$$w_{1I} = \frac{\phi_1\phi_2^2(1-\pi)^2 b_1 b_2 + \phi_1^2 + \phi_1\phi_2 + \phi_1^2\phi_2^2(1-\pi)^2 b_1 b_2 + (\phi_1^2\phi_2 + \phi_1\phi_2^2)[\gamma(1-\pi)(\kappa_2^2 - \kappa_1^2 + b_2^2 - b_1^2) + (1-\pi)^2(b_2 - b_1)\omega]}{\phi_1^2 + \phi_2^2 + 2\phi_1\phi_2 + 4\phi_1\phi_2(\phi_1+\phi_2)(1-\pi)^2 b_1 b_2} \tag{5.14}$$

Which is the final part of the proposition, thus completing the proof. ■

Appendix C : Proof of Proposition 7

Proof For simplicity we first make the following substitution:

$$\frac{\sum_i^n \psi_i b_i}{\sum_i^n \psi_i} = \sum_i^n w_{iN} b_i$$

This yields the following:

$$L_1(b_i, b_j) + L_2(b_i, b_j) = (1-\gamma)\left(\left[(1-\pi) + \frac{\pi\,[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\right]\sum_i^n w_{iN} b_i - \bar{b}_i\right)^2 + \gamma\left[\frac{\pi(1-\pi)^2(1-\gamma)^2 + (1-\pi)\pi^2(1-\gamma)^2}{[(1-\gamma)+(1-\pi)^2]^2}\right]\left(\sum_i^n w_{iN} b_i\right)^2 \tag{5.15}$$

By Lemma 6 the first-order conditions will yield the solution to our minimization problem, thus we take derivatives with respect to b_i:

$$\frac{\partial L_1}{\partial b_i} = 2(1-\gamma)\left(\left[(1-\pi) + \frac{\pi\,[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\right]\sum_i^n w_{iN} b_i - \bar{b}_i\right)$$

$$\frac{\partial L_2}{\partial b_i} = 2\gamma\left[\frac{\pi(1-\pi)^2(1-\gamma)^2 + (1-\pi)\pi^2(1-\gamma)^2}{[(1-\gamma)+(1-\pi)^2]^2}\right]\left(\sum_i^n w_{iN} b_i\right)$$

Now let us make the following substitutions:

$$\Lambda_1 = (1-\gamma)\left[(1-\pi) + \frac{\pi\,[1-\pi]^2}{[(1-\gamma)+(1-\pi)^2]}\right]$$

$$\Lambda_2 = \gamma\left[\frac{\pi(1-\pi)^2(1-\gamma)^2 + (1-\pi)\pi^2(1-\gamma)^2}{[(1-\gamma)+(1-\pi)^2]^2}\right]$$

Taking these substitutions in mind, we arrive at the following condition:

$$(\Lambda_1 + \Lambda_2)\left(\sum_i^n w_{iN} b_i\right) = \Lambda_1 \bar{b}_i \Rightarrow \left(1 + \frac{\Lambda_2}{\Lambda_1}\right)\left(\sum_i^n w_{iN} b_i\right) = \bar{b}_i,\ \forall i$$

But in particular, this implies the following:

$$\left(1 + \frac{\Lambda_2}{\Lambda_1}\right)\left(\sum_i^n w_{iN} b_i\right) = \bar{b}_1 = \bar{b}_2 = \cdots = \bar{b}_n$$

Thus we once again have that only a faction equilibrium can exist. In particular, assume again and without loss of generality that there exists some $m \leq n$ such that for all $i \leq m$ we have $z_i > 0$ and for all $i > m$ we have $z_i = 0$ (note that the same condition must hold as previously derived, that is if $z_i > 0$ we have $\xi_i < K_i$ where ξ_i and K_i are as previously defined). Thus only these m leaders receive attention. Then we have that $\bar{b}_i = \bar{b}$ for some $\bar{b} \in \Re$ for all $i \leq m$ for an equilibrium to exist. In particular, we have:

$$\left(1 + \frac{\Lambda_2}{\Lambda_1}\right)\left(\sum_i^n w_{iN} b_i\right) = \bar{b} \text{ for some } \bar{b} \in \Re$$

But this faction equilibrium in turn implies that b_i is identical for all i and as such the equilbrium is symmetric. This implies in turn that the following holds:

$$\left(1 + \frac{\Lambda_2}{\Lambda_1}\right)\left(\sum_i^n w_{iN} b_i\right) = \left(1 + \frac{\Lambda_2}{\Lambda_1}\right)\left(b_i \sum_i^n w_{iN}\right) = \left(1 + \frac{\Lambda_2}{\Lambda_1}\right) b_i = \bar{b}$$

And thus we arrive at our equilibrium bias b_i^*:

$$b_i^* = \frac{\Lambda_1}{\Lambda_1 + \Lambda_2}\bar{b}\ ,\quad \forall\ i \leq m \tag{5.16}$$

Which is the expression for b_i^* as given in the proposition, thus completing the proof. ■