Working Papers

Center for
Mathematical
Economics

469

August 2013

Opinion Dynamics and Wisdom under Conformity

Berno Buechel, Tim Hellmann and Stefan Klößner

IMW · Bielefeld University
Postfach 100131
33501 Bielefeld · Germany

email: imw@uni-bielefeld.de
http://www.imw.uni-bielefeld.de/wp/2013.php
ISSN: 0931-6558

Opinion Dynamics and Wisdom under Conformity

Berno Buechel*
Department of Economics
University of Hamburg
berno.buechel@uni-hamburg.de

Tim Hellmann
Institute of Mathematical Economics
Bielefeld University
tim.hellmann@uni-bielefeld.de

Stefan Klößner
Statistics and Econometrics
Saarland University
s.kloessner@mx.uni-saarland.de

August 12, 2013

Abstract

We study a dynamic model of opinion formation in social networks. In our model, boundedly rational agents update opinions by averaging over their neighbors' expressed opinions, but may misrepresent their own opinion by conforming or counter-conforming with their neighbors. We show that an agent's social influence on the long-run group opinion is increasing in network centrality and decreasing in conformity. For efficiency of information aggregation ("wisdom"), misrepresentation of opinions need not undermine wisdom. Given the network, we provide the optimal distribution of conformity levels in the society and show which agents should be more conforming in order to increase wisdom.

Keywords: opinion leadership, wisdom of crowds, consensus, social networks, conformity, eigenvector centrality
JEL: C72, D83, D85, Z13

*Corresponding author. Address: Department of Economics, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany. Phone number: +49 / (0)40 42838 5573.

1 Introduction

Understanding *opinions* is important because they crucially shape economic behavior. Consumers' demand for a product depends on the opinions about the quality of this product and about the integrity of its producing company. Majority opinions on political issues set the political course. Moreover, opinions on the relative importance of issues decide upon the agenda of actions or on the allocation of a budget – be it within a company, within a government, or within some other group of decision makers. *When asked for a personal opinion, however, people are often tempted to misrepresent what they actually think, e.g. because disagreement would make them feel uncomfortable.* Abstracting from this issue, models of opinion formation have worked with the assumption that people do not misrepresent their opinions. They provide conditions for the emergence of consensus of opinions (e.g. DeGroot, 1974), identify opinion leaders (e.g. DeMarzo et al., 2003), and even show that large societies can be "wise" in a well defined sense (e.g. Golub and Jackson, 2010). We challenge these results by incorporating the possibility that stated opinions differ from true opinions in a conforming or counter-conforming way. This requires additional conditions to guarantee consensus, it affects who is an opinion leader, and it can undermine or foster the wisdom of a society, as we will show.

If individuals are fully rational and completely informed, then the social network (of personal relationships) does not affect long-run opinions.[1] However, in most settings it is not realistic to assume that individuals know the whole social network which determines the communication structure. Moreover, as it has been shown recently in laboratory experiments, even in small social networks where the network is made common knowledge, people fail to properly account for repetitions of information (Corazzini et al., 2012). Therefore, models of more naïve social learning are used to describe the process of opinion formation (DeMarzo et al., 2003; Golub and Jackson, 2010; Acemoglu et al., 2010; Corazzini et al., 2012). The common idea of these approaches is to assume that agents update their opinion according to a weighted average of the current opinions (cf. DeGroot, 1974). In this so-called DeGroot model the weights of averaging are collected in an exogenously given learning matrix, which has the interpretation of a social network.

While the assumption of this form of non-Bayesian updating has been extensively discussed and motivated (Friedkin and Johnsen, 1990; DeMarzo et al., 2003; Acemoglu and Ozdaglar, 2011; Corazzini et al., 2012), this is not true to the same extent for another crucial assumption of the DeGroot model framework: it is assumed that actors do not misrepresent their opinion; in other words, stated opinions are assumed to coincide with true opinions. DeMarzo et al. (2003) argue that this assumption is problematic in contexts

[1] Indeed, among equally informed agents within a strongly connected social network that is common knowledge, Bayesian updating leads to convergence of all agents' opinions to their initial opinions' average (DeMarzo et al., 2003, theorem 3).

of persuasion, where actors have a material interest in influencing others' opinions. But even if there is no material incentive to persuade, people often misrepresent their opinions. In the famous study of Asch (1955), subjects wrongly judged the length of a stick after some other, allegedly neutral, participants had placed the same wrong judgment. Follow-up studies revealed that this effect is weaker if the subjects do not have to report their judgments publicly (Deutsch and Gerard, 1955). The authors argue that two forms of social influence can be observed in this study. While *informational social influence* describes the updating of opinions according to what others have said, *normative social influence* describes the behavior of stating an opinion that fits to the group norm.[2] Thus, one motive to misrepresent the true opinion is a preference for conformity in the sense of "getting a utility gain by simply making the same choice as one's reference group" (Zafar, 2011, p. 774). Incentives to conform can be derived from desires for social status (Bernheim, 1994) and are embodied in a utility component that depends on the difference of the behavior of the focal actor and the behavior of some peer group (Jones, 1984), also called "reference group" (Hayakawa and Venieris, 1977). Meanwhile, the concepts of informational and normative social influence have become a cornerstone in analyzing social influence, e.g. Ariely and Levav (2000, p. 279) call it the "primary paradigm".[3] In terms of this paradigm, the DeGroot model of opinion formation and its variations are models of informational social influence, but not of normative social influence.

In this work, we present a model that incorporates both informational and normative social influence. The model consists of a sequence of discussion rounds among naïve and boundedly rational agents. In each discussion round agents express an opinion depending on their true opinion and on their preferences for conformity. We consider agents with preferences for conformity, counter-conformity, and honest agents.[4] From one discussion round to the next, learning takes place in the sense that agents update their opinion according to a learning matrix. In the special case where every agent is honest, our model coincides with the classic DeGroot model studied by DeMarzo et al. (2003); Golub and Jackson (2010); Corazzini et al. (2012). Allowing agents to misrepresent their opinion in a conforming or counter-conforming way, we investigate *how opinions evolve*. We first analyze the two-agent case which illustrates that dynamics can diverge, converge, or cycle.

[2]Deutsch and Gerard (1955, p. 629) further explain: "Commonly these two types of influence are found together. However, it is possible to conform behaviorally with the expectations of others and say things which one disbelieves but which agree with the beliefs of others. Also, it is possible that one will accept an opponent's beliefs as evidence about reality even though one has no motivation to agree with him, per se."

[3]However, this paradigm did not explicitly enter economic models. The terms 'social influence' and 'conformity' do usually not clarify whether social or normative influence is at work. We will be more explicit on this distinction and only refer to conformity as a form of normative social influence.

[4]This is consistent with the psychological theory on normative social influence, which considers identification, non-identification and disidentification with the peer group as the sources of conformity, honesty (independence), and counter-conformity (Hogg and Abrams, 1988).

It turns out that a sufficiently conforming agent will reach consensus with any other agent. We then show more generally that excluding counter-conforming agents is sufficient to guarantee convergence of opinions to local consensus.

Focusing on convergence, we then ask *how opinion leadership depends on conformity.* Understanding the determinants of opinion leadership is important because opinion leaders have the potential to mislead others. In our model opinion leadership or *power* of any agent can be measured by the influence of her initial opinion on the long-run (consensus) opinion of her group. As one of the main results, we show how power is determined not only by eigenvector centrality (Bonacich, 1972; Friedkin, 1991) with respect to the learning matrix, but also by the distribution of conformity in the society. Comparative statics reveal that an agent's power is decreasing in own level of conformity, increasing in other agents' level of conformity and increasing in own network centrality. These results show that a strong position in the social network is not sufficient to become an opinion leader. Another requirement is that an agent is not more conforming than the other members of the society. Thus, our model provides a theoretical explanation for the empirical finding that opinion leaders are characterized by a low degree of conformity (Chan and Misra, 1990).

Finally, we consider a context where there is a true state of nature and the individuals' initial opinions are independent, unbiased, noisy signals which may differ with respect to signal precision (the inverse of the variance). The question is *how the misrepresentation of opinions affects the accuracy of information aggregation (the society's "wisdom").* A negative effect might be expected since stated opinions become even less reliable signals about the truth. Our results show that this conjecture does not hold in general. First, if the society is homogeneous with respect to conformity, then information aggregation is neither worse nor better than in the DeGroot model (i.e. when all individuals are honest). Moreover, heterogeneous levels of conformity foster wisdom if they balance the power of agents with their signal precision, while an unbalanced distribution can lead to lower wisdom. Using comparative statics we observe that for the goal of higher accuracy of the consensus opinion it would be helpful if people with a low signal precision (relative to their power) were more conforming, while people with a high signal precision (relative to their power) should be less conforming, or in more poetic words: "The whole problem with the world is that fools and fanatics are always so certain of themselves, but wiser people so full of doubts."[5]

[5] Credit for this quote is often given to the philosopher and mathematician Bertrand Russell. Although the origin of the quote is actually unknown, it is at least confirmed that Russell made a similar statement in his essay "The Triumph of Stupidity," which can be found in *Mortals and Others: Bertrand Russell's American Essays, 1931-1935.*

Related Models The non-Bayesian approach to learning in social networks roots in the pioneer work of French (1956), Harary (1959), and DeGroot (1974). Friedkin and Johnsen (1990) provide a framework that subsumes former models as special cases. A particular feature of Friedkin and Johnsen (1990) is that opinions can be updated in every period not only according to the current profile of opinions but also according to the own initial opinion. Another variation of the classic model is to let agents only be affected by opinions that are not too different from the own opinion (Hegselmann and Krause, 2002). Moreover, Lorenz (2005) allows the learning matrix to vary over time and identifies general conditions for convergence. Under some conditions, convergence to consensus is also robust if updating is noisy, as Mueller-Frank (2011) shows. In the seminal contribution of DeMarzo et al. (2003) self confidence is allowed to vary over time. There are also studies which extend the model by DeGroot (1974) to allow for adaption of learning weights, e.g. in Pan (2010) the influence weights are updated over time and Flache and Torenvlied (2004) study a variation of the classic model where actors anticipate the difference between own opinion and group decision and adapt learning weights accordingly. The recent model by Foerster et al. (2013) studies agents who increase the learning weights others have for them. In a context of cultural transmission of traits, Buechel et al. (2011) introduce strategic interaction for the DeGroot model in an OLG framework.

Finally, Corazzini et al. (2012) assess real opinion updating in a laboratory experiment. They reject the hypothesis that participants are fully Bayesian and conclude that the network structure does matter, which is in line with DeMarzo et al. (2003) who assume that agents do not properly account for repetition, but are subject to persuasion bias. Corazzini et al. (2012) suggest a specification of the DeGroot model, which is also a special case of our model.

In the context of binary opinions Condorcet's Jury theorem is a famous example of how aggregation of individual opinions can be efficient in the limit. In the framework of the DeGroot model, a similar phenomenon, coined "the wisdom of crowds," is studied. Golub and Jackson (2010) provide conditions for wisdom in the sense that the consensus opinion of a society comes arbitrarily close to the truth when letting the size of the society grow. Our approach to assess efficiency of information aggregation differs in that we analyze the accuracy of information aggregation for a society of fixed size. This is in line with Acemoglu et al. (2010) who assess the wisdom of a society as the difference between optimal information aggregation and the consensus opinion that emerges in their model. However, similar to Golub and Jackson (2010) we find that naïve agents can be remarkably wise.

Besides these highly related works, there are several contributions to social influence in the context of discrete choices of actions, such as the choice of one out of two technologies. While their discussion is beyond the scope of this paper, we refer the reader to the following few prominent examples: models of social learning (Bikhchandani et al., 1992; Ellison and Fudenberg, 1993, 1995; Bala and Goyal, 1998, 2001), cooperative models of social influence (Grabisch and Rusinowska, 2010, 2011), and a model of strategic influence (Galeotti and Goyal, 2009). Moreover, there is a recent literature on communication in social networks when agents are strategic (Hagenbach and Koessler, 2010; Anderlini et al., 2012; Ambrus et al., 2013). Those models of strategic information transmission clearly apply to different contexts than the models of naïve social learning.

The rest of this paper is organized into five sections. In Section 2 we introduce the model. Before we present the main results (in Section 4), we discuss the two-agent case (Section 3). Section 5 addresses the wisdom of a society and in Section 6 we conclude, while proofs are relegated to the appendix.

2 Model

2.1 Informational Social Influence

There is a set of agents/players $\mathcal{N} = \{1, 2, ..., n\}$ who interact with each other. A learning structure is given by a $n \times n$ row stochastic matrix G, i.e. $g_{ij} \geq 0$ for all $i, j \in \mathcal{N}$ and $\sum_{j=1}^{n} g_{ij} = 1$ for all $i \in \mathcal{N}$. This learning matrix represents the extent to which agents listen to other agents and it can be interpreted as a weighted and directed social network. We say that there is a directed path from i to j in this network if there exists $i_0, ..., i_k \in \mathcal{N}$ such that $i_0 = i$ and $i_k = j$ and $g_{i_l i_{l+1}} > 0$ for all $l = 0, ..., k-1$, which is equivalent to $(G^k)_{ij} > 0$.[6] Moreover, we assume that $g_{ii} < 1$ for all i to assure that all agents update their opinion.

We study a dynamic model where time is discrete $t = 0, 1, 2...$ and initially each agent has a predefined opinion $x_i(0)$ concerning some topic. The opinions of all agents at time t are collected in $x(t) \in \mathbb{R}^n$. In every period, agents talk to each other and finally update their opinions according to the matrix G. In the classical DeGroot model agents exchange opinions such that the opinions in period $t+1$ are formed by $x(t+1) = Gx(t) = G^{t+1}x(0)$ (DeGroot, 1974). The motivation for such a model is that agents always report their true opinions and suffer from persuasion bias when the next period's opinion is formed as a weighted average of own and others' opinions according to the social network G. Concerning the assumption of honesty in opinion formation, DeMarzo et al. (2003) note:

> "For simplicity, we assume that agents report their beliefs truthfully."[7]

[6] We follow the convention of Jackson (2008) and DeMarzo et al. (2003) that a directed link from agent i to agent j indicates that i listens to j, i.e. $g_{ij} > 0$, while the opposite convention is used by Corazzini et al. (2012).

[7] DeMarzo et al. (2003, p. 3, footnote 9).

We relax this assumption: an agent $i \in \mathcal{N}$ expresses some opinion $s_i(t) \in \mathbb{R}$ which need not coincide with her true opinion $x_i(t)$.[8]

A central assumption of our approach is that an agent cannot observe the true opinions of the others but only their stated opinions. Since each agent knows her own true opinion $x_i(t)$, we get that agent i's next period's opinion is formed by $x_i(t+1) = g_{ii}x_i(t) + \sum_{j \neq i} g_{ij}s_j(t)$, where the weights g_{ij} are the individual learning weights as in the classical model by DeGroot (1974). This holds for all agents $i \in \mathcal{N}$ and, thus, the updating process becomes

$$x(t+1) = Dx(t) + (G - D)s(t), \tag{1}$$

where D is the $n \times n$ diagonal matrix containing the diagonal of G.

2.2 Normative Social Influence

Misrepresenting the own opinion (i.e. being dishonest) might cause discomfort (e.g. Festinger, 1957). However, there are various motives to misrepresent the own opinion. Not only strategic considerations of persuasion play a role, but also personality traits or emotional motives. There is ample evidence that many people feel discomfort from stating an opinion that is different from their peer group's opinion (e.g. Deutsch and Gerard, 1955). While certainly many people feel this type of normative social influence, this need not be true for all people – there are even some who prefer to state an opinion that is far away from what others say.[9] We focus on these two motives for the misrepresentation of opinions: conformity and counter-conformity.

To formalize these ideas, consider an agent i who is confronted with some group opinion q_i, while her own opinion on this topic is x_i. In the spirit of the model of Bernheim (1994) we consider a utility function that depends on an intrinsic part – this will be the incentive to be honest – and a social part – this will be the incentive to conform/counter-conform. Additionally, we assume that utility of an agent is additively separable into these two parts and that for each part disutility takes a quadratic form.

Thus, the utility of agent i depends on the distance of true opinion x_i to stated opinion s_i as well as on the distance of stated opinion s_i to group opinion q_i in the following way:

$$u_i(s_i|x_i) := -(1-\delta_i)\big(s_i - x_i\big)^2 - \delta_i\big(s_i - q_i\big)^2, \tag{2}$$

[8]The incentive to state an opinion different from true opinion will be based on preferences for conformity or counter-conformity (cf. Subection 2.2). Moreover, agents adapt their stated opinions faster than true opinions such that $s(t)$ is given by Proposition 1.

[9]For instance, Hornsey et al. (2003) conducted a laboratory experiment where subjects reported their willingness to privately or publicly express and support their opinion. For subjects with a strong moral basis on the topic, the treatment of suggesting that a majority of the other subjects disagreed slightly increased the willingness to publicly express the opinion.

where $\delta_i \in (-1, +1)$ displays the relative importance of the preference for conformity in relation to the preference for honesty. The preference peak (or "bliss point," Bernheim, 1994) for such an agent is given by $s_i = (1-\delta_i)x_i(t)+\delta_i q_i(t)$. This assumption is illustrated in Figure 1. For $\delta_i \in (0, 1)$ the agent faces a trade-off between conforming and being honest such that her preference peak lies within the interval (x_i, q_i). For $\delta_i \in (-1, 0)$, a similar trade-off can be seen counter-conforming and being honest. In that case the preference peak lies within the interval $(x_i - (q_i - x_i), x_i)$. We assume that $\delta_i > -1$ to restrict counter-conformity to a certain bound which seems weak enough to cover all reasonable cases, but keeps the analysis tractable.

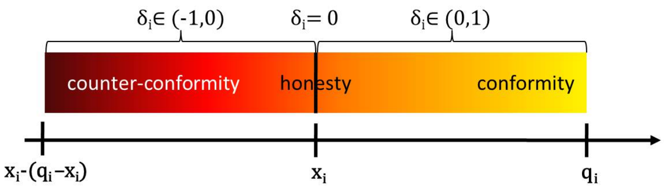

Figure 1: Preferences for conformity, counter-conformity, and honesty.

A stylized fact on normative social influence is that people are heterogeneous in the way and their degree of being influenced. The degree of conformity can hence be considered a personality trait, but it might also depend on the topic under discussion. Let Δ denote the $n \times n$ diagonal matrix with entries $\delta_i \in (-1, 1)$ on the diagonal representing the levels of conformity in the society.

While updating of true opinions, i.e. learning, is a relatively slow cognitive process, stated opinions can be adapted in a much more fluid way. Our model reflects this fact by considering an adaption process of stated opinions which takes place within a time period t, while true opinions are updated from one period to the next.[10] Thus, suppose that within each period $t \in \mathbb{N}$, there is a fast time scale $\tau \in \mathbb{N}$ such that at each time step τ one or more agents speak. The (possibly random) set of agents who are selected to state their opinions at time step τ (of period t) is denoted by $A^\tau(t)$. Let $s^\tau(t)$ be the vector of stated opinions. Agents who are not selected to revise keep the stated opinion of the previous time step, i.e. $s_i^\tau(t) = s_i^{\tau-1}(t)$ if $i \in \mathcal{N} \setminus A^\tau(t)$. Agents, who are selected to speak and thereby revise their stated opinion, observe last time step's stated opinions of their neighbors. These are perceived as a reference opinion $q_i^{\tau-1}(t)$, which is the average of the stated opinions with weights according to the listening matrix G, i.e.

$$q_i^\tau(t) = \sum_{j \neq i} \frac{g_{ij}}{1 - g_{ii}} s_j^\tau(t). \tag{3}$$

[10] An interpretation for this assumption is that each period is a discussion round within which stated opinions are adjusted, while learning takes place between discussion rounds.

In line with our assumption that agents are naïve when updating, we also assume that agents are boundedly rational when revising their stated opinions. Upon revision opportunity, i.e. $i \in A^\tau(t)$, an agent i myopically chooses a stated opinion which maximizes her current utility given by (2), i.e.

$$s_i^\tau(t) = (1 - \delta_i)x_i(t) + \delta_i q_i^{\tau-1}(t), \tag{4}$$

for any true opinion $x_i(t)$ and any reference opinion $q_i^{\tau-1}(t)$.[11] Hence, the stated opinion given by myopic best response differs from the true opinion proportionally to the difference of reference opinion and true opinion, and the proportion is determined by the preference parameter δ_i. The parameter δ_i can thus be directly interpreted as the degree of conformity of agent $i's$ behavior (cf. Figure 1). A conforming agent, characterized by $\delta_i \in (0, 1)$, states an opinion between the true opinion $x_i(t)$ and perceived opinion $q_i^{\tau-1}(t)$. A counter-conforming agent, characterized by $\delta_i \in (-1, 0)$, states an opinion that is more extreme than the true opinion $x_i(t)$ (with respect to the perceived opinion $q_i^{\tau-1}(t)$). Finally, an honest agent, characterized by $\delta_i = 0$, straight-forwardly states the true opinion, i.e. $s_i^\tau(t) = x_i(t)$ for all $\tau \in \mathbb{N}$.

To ensure that every agent takes part in opinion exchange in period t, we assume that for each agent i, the set $\{\tau \in \mathbb{N} : i \in A^\tau(t)\}$ is (almost surely) infinite, reflecting the idea that no agent will stay forever with a stated opinion that is not in line with her preferences. This assumption is satisfied if e.g. at each time step τ agents are randomly selected to speak according to some probability distribution with full support on $\mathcal{N}$.

It turns out that such a myopic best reply process within period $t \in \mathbb{N}$ inevitably leads to one specific profile of stated opinions $s(t)$ which only depends on the network G and the conformity parameters Δ, but not on the starting stated opinions $s^0(t)$.

Proposition 1. *Given the assumptions above, the within-period dynamics $s^\tau(t)$ converge for $\tau \to \infty$ to*

$$s(t) := [I - \Delta(I - D)^{-1}(G - D)]^{-1}(I - \Delta)x(t). \tag{5}$$

The proof of Proposition 1 as well as all proofs of the following propositions are relegated to an appendix. Proposition 1 shows that agents who revise opinions by conforming or counter-conforming to what their neighbors last said, finally state (or *express*) the opinions given by (5).

It is worth noting that considering the action sets $S_i(t) = \mathbb{R}$ and utility functions $u_i(s_i(t)|x_i(t))$ given by (2) implies that $s(t)$ obtained by Proposition 1 is the unique Nash equilibrium of the normal form game $(\mathcal{N}, S(t), u(\cdot|x(t))$ for each $t \in \mathbb{N}$. Note that the process that leads into this Nash equilibrium within period t neither requires complete

[11]Myopic maximizing is a common assumption in such models (see, e.g. Corazzini et al., 2012).

information (e.g. on the network structure G), nor high degrees of rationality, nor some sort of common knowledge.

2.3 Model Summary

In our model each period $t \in \mathbb{N}$ can be viewed as a discussion round within which agents express opinions and then learn from one discussion round to the next. Proposition 1 determines which opinions are finally stated in a given period as a function of the true opinions $x(t)$. These stated opinions $s(t)$ determine the vector of reference opinions $q(t)$ by (3) and are then a crucial ingredient of the updating process.[12] Since opinions of period $t+1$ are formed by (1) and the stated opinions of each period can be calculated as in Proposition 1, we conclude that the opinion profile in period $t+1$ depends on the opinion profile in period t in the following way:

$$x(t+1) = Mx(t), \tag{6}$$

where $M := \Big[D+(G-D)[I-\Delta(I-D)^{-1}(G-D)]^{-1}(I-\Delta)\Big]$. Note that the transformation from $x(t)$ to $x(t+1)$, i.e. the matrix M, is independent of $x(t)$. Thus, the opinion dynamics is fully described by the power series M^t, since $x(t+1) = Mx(t) = M^2x(t-1) = ... = M^{t+1}x(0)$.[13] The relation to the classical DeGroot model becomes apparent in this expression when recalling $x(t+1) = Gx(t) = G^{t+1}x(0)$. In that light the misrepresentation of opinions leads to a transformation of the matrix G into the matrix M. If every agent is honest, i.e. $\delta_i = 0$ for any $i \in \mathcal{N}$, then $M = G$ and, hence, we are back in the classical case of DeGroot (1974).

Let us illustrate the model introduced above by an example with three agents.

Example 1. *Suppose there are three agents. Player 1 (black) starts with an opinion* $x_1(0) = 0$*, Player 2 (red) and Player 3 (blue) have initial opinions of* $x_2(0) = 50$ *and* $x_3(0) = 100$*. Player 2 is an honest agent, i.e.* $\delta_2 = 0$*, Player 3 is a conforming agent, i.e.* $\delta_3 = .5 > 0$*, and Player 1 is a counter-conforming agent, i.e.* $\delta_1 = -.5 < 0$*. To illustrate the implications of the different degrees of conformity, we let the players be in a symmetric network position. In particular, let the interaction structure be given by*
$$G = \begin{pmatrix} .6 & .2 & .2 \\ .2 & .6 & .2 \\ .2 & .2 & .6 \end{pmatrix}.$$

The dynamics of opinions across periods are displayed in Figure 2, where the solid lines indicate the dynamics of true opinions $x(t)$*, the dashed lines display the stated opinions*

[12]Since one interpretation for $q_i(t)$ is that this is the society's opinion at time t as perceived by agent i, we also call it i's *perceived* opinion.

[13]The simple linear structure is of course implied by our assumption of quadratic utility.

$s(t)$ at the end of each period, and the dotted lines the perceived opinions $q(t)$. For better readability, we abstract from within-period dynamics and simply connect the opinions at time t and $t+1$ by straight lines.

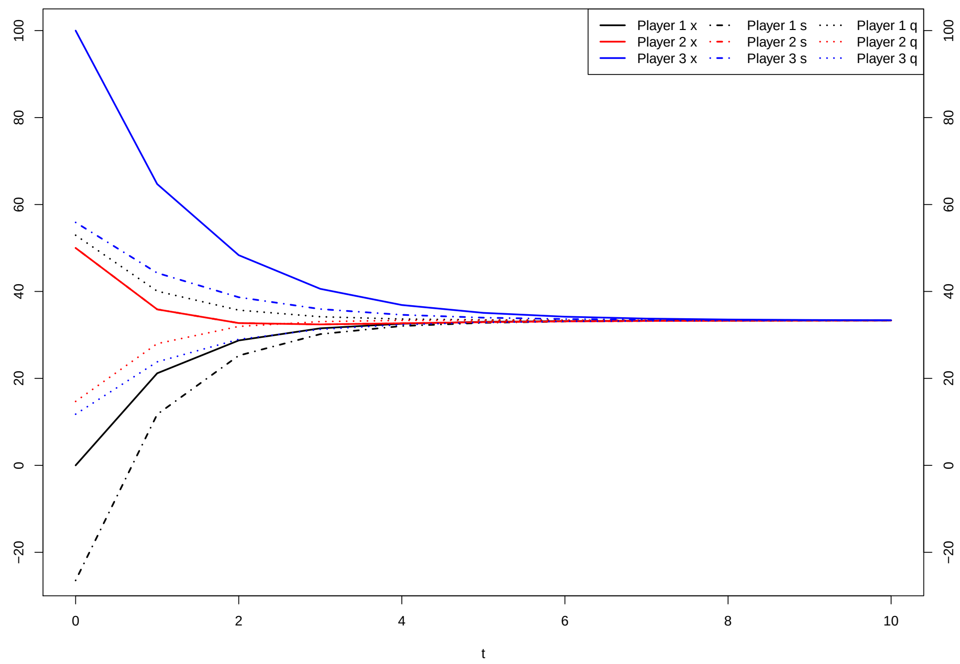

Figure 2: A three-agent example with one honest, one conforming, and one counter-conforming agent.

Since Player 2 (red) is honest, her stated opinion will always be equal to her true opinion. Therefore, those functions (red dashed line and red solid line) coincide. Player 3 is a conforming agent, she always expresses an opinion (dashed blue line) that is a convex combination of the perceived opinion of others (blue dotted line) and her true opinion (blue solid line). Player 1 is a counter-conforming agent. With respect to the perceived opinion (black dotted line), she always expresses an opinion (black dashed line) that is more extreme than her true opinion (black solid line).

The opinion dynamics in this simple example are such that stated, true, and perceived opinions of each agent become more and more similar and approach the value 33.3 in the long-run. Thus, the long-run opinions are closer to the initial opinion of the counter-conforming Player 1 than to the initial opinion of the conforming Player 3. If every agent was honest, i.e. ($\delta_i = 0$, $i = 1, 2, 3$), opinions in this simple example would approach a value of 50.

The dynamics of Example 1 highlights several features, the generality of which we discuss in the subsequent sections.

3 Two-Agent Case

Let us begin with the analysis of the two-agent case. In this case, closed form solutions are easy to obtain and, still, it is possible to observe several important properties of the opinion dynamics. (The n-agent case is presented in Section 4.)

Let $n = 2$. Then we can write G as

$$G = \begin{pmatrix} 1 - g_{12} & g_{12} \\ g_{21} & 1 - g_{21} \end{pmatrix}$$

with $g_{12}, g_{21} \in (0, 1)$. With only two agents, the relevant group average for one agent is simply the stated opinion of the other agent, i.e. $q_1(t) = s_2(t)$ and $q_2(t) = s_1(t)$. Plugging in the variables for G into (6) yields

$$M = \begin{pmatrix} 1 - m_{12} & m_{12} \\ m_{21} & 1 - m_{21} \end{pmatrix} = \begin{pmatrix} 1 - g_{12}\dfrac{1-\delta_2}{1-\delta_1\delta_2} & g_{12}\dfrac{1-\delta_2}{1-\delta_1\delta_2} \\ g_{21}\dfrac{1-\delta_1}{1-\delta_1\delta_2} & 1 - g_{21}\dfrac{1-\delta_1}{1-\delta_1\delta_2} \end{pmatrix}.$$

Since $x(t+1) = Mx(t)$, an entry m_{ij} gives the importance of Player j on the one-period opinion change of Player i. From $\frac{\partial m_{12}}{\partial \delta_2} = -g_{12}\frac{1-\delta_1}{(1-\delta_1\delta_2)^2}$, we see the following comparative static effect: higher conformity of Player 2 reduces her one-period influence on Player 1 (m_{12}), which vanishes ($m_{12} \to 0$) when Player 2's conformity approaches 1. Thus, in the short run, conformity results in a reduction of influence. To investigate long-run effects, we examine the power series M^t since $x(t) = M^t x(0)$. By induction one can easily see that M^t can be rewritten as follows:

$$M^t = \frac{1}{m_{12} + m_{21}} \begin{pmatrix} m_{21} + m_{12}(1 - m_{12} - m_{21})^t & m_{12} - m_{12}(1 - m_{12} - m_{21})^t \\ m_{21} - m_{21}(1 - m_{12} - m_{21})^t & m_{12} + m_{21}(1 - m_{12} - m_{21})^t \end{pmatrix}. \quad (7)$$

From (7), we observe that the decisive quantity for the (speed of) convergence of M^t is

$$\lambda := 1 - m_{12} - m_{21} = 1 - \frac{g_{12}(1-\delta_2) + g_{21}(1-\delta_1)}{1 - \delta_1\delta_2} < 1,$$

which is the second (largest) eigenvalue of M (the other eigenvalue of M is always 1). In particular, M^t converges if $|\lambda| < 1$ and, moreover, the smaller $|\lambda|$, the higher the speed of convergence. Before discussing the issue of convergence in more detail, let us have a brief look at the limit of M^t in case of convergence: with the help of (7), we have

$$M^\infty = \lim_{t\to\infty} M^t = \begin{pmatrix} \dfrac{m_{21}}{m_{12}+m_{21}} & \dfrac{m_{12}}{m_{12}+m_{21}} \\ \dfrac{m_{21}}{m_{12}+m_{21}} & \dfrac{m_{12}}{m_{12}+m_{21}} \end{pmatrix}$$

such that, in the long run, the two agents will reach a consensus because $x(\infty) = M^\infty x(0)$. Player 1's and Player 2's initial opinions enter this consensus opinion with weights $\frac{m_{21}}{m_{12}+m_{21}}$ and $\frac{m_{12}}{m_{12}+m_{21}}$, respectively. Since $\frac{m_{12}}{m_{12}+m_{21}} = \frac{g_{12}(1-\delta_2)}{g_{12}(1-\delta_2)+g_{21}(1-\delta_1)} = 1 - \frac{g_{21}(1-\delta_1)}{g_{12}(1-\delta_2)+g_{21}(1-\delta_1)}$, Player 2's influence in the long-run is decreasing in δ_2. Therefore, increasing conformity not only decreases the short-run importance of an agent, but also the long-term impact of this agent's initial opinion.

To study the effect of conformity/counter-conformity on convergence, we will first consider the special case $\delta_1 = \delta_2 =: \delta$ which simplifies λ to

$$\lambda = 1 - \frac{1}{1+\delta}(g_{12} + g_{21}).^{14} \tag{8}$$

Since $\lambda < 1$, the decisive thresholds for λ are $\lambda = 0$ and $\lambda = -1$: for $\lambda = 0$, convergence will be fastest (one-step convergence due to $M = M^2 = ... = M^\infty$), while $\lambda = -1$ marks the case of cycling M^t (M^t will alternate between $M^1 = M^3 = ...$ and $M^2 = M^4 = ...$). Figure 3 exemplifies the corresponding dynamics for $G = \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$ and initial opinions $x(0) = (0, 100)'$. Notice, in particular, that the speed of convergence of true opinions $x(t)$ is not monotone in δ: when δ decreases from 0.5 to -0.4, speed increases and eventually reaches one-step convergence; however, further reducing δ first leads to slower, alternating dynamics, cycling, and finally divergent behavior.[15] It might be surprising that higher levels of conformity can decrease the speed of convergence. The intuition for this effect can be gained by comparing cases (a) and (b). Under conformity, i.e. in case (a), stated opinions $s(t)$ are closer to each other in the first time periods such that agents' true opinions $x(t)$ are less swayed to the center compared with case (b) where agents are honest.[16]

If we relax the assumption of equal conformity ($\delta_1 = \delta_2$), the necessary and sufficient

[14] λ and $|\lambda|$ as a function of δ are depicted in part (0) of Figure 3.

[15] Another aspect that can be observed in Figure 3 is that, under convergence, i.e. in cases (a)-(e), the dynamics converges to the same limit independently of δ. We will show later on that this observation is not a coincidence and that it is induced by setting $\delta_1 = \delta_2 = \delta$.

[16] Recall that agents know their own true opinion and are thus resistant against their own misrepresentation.

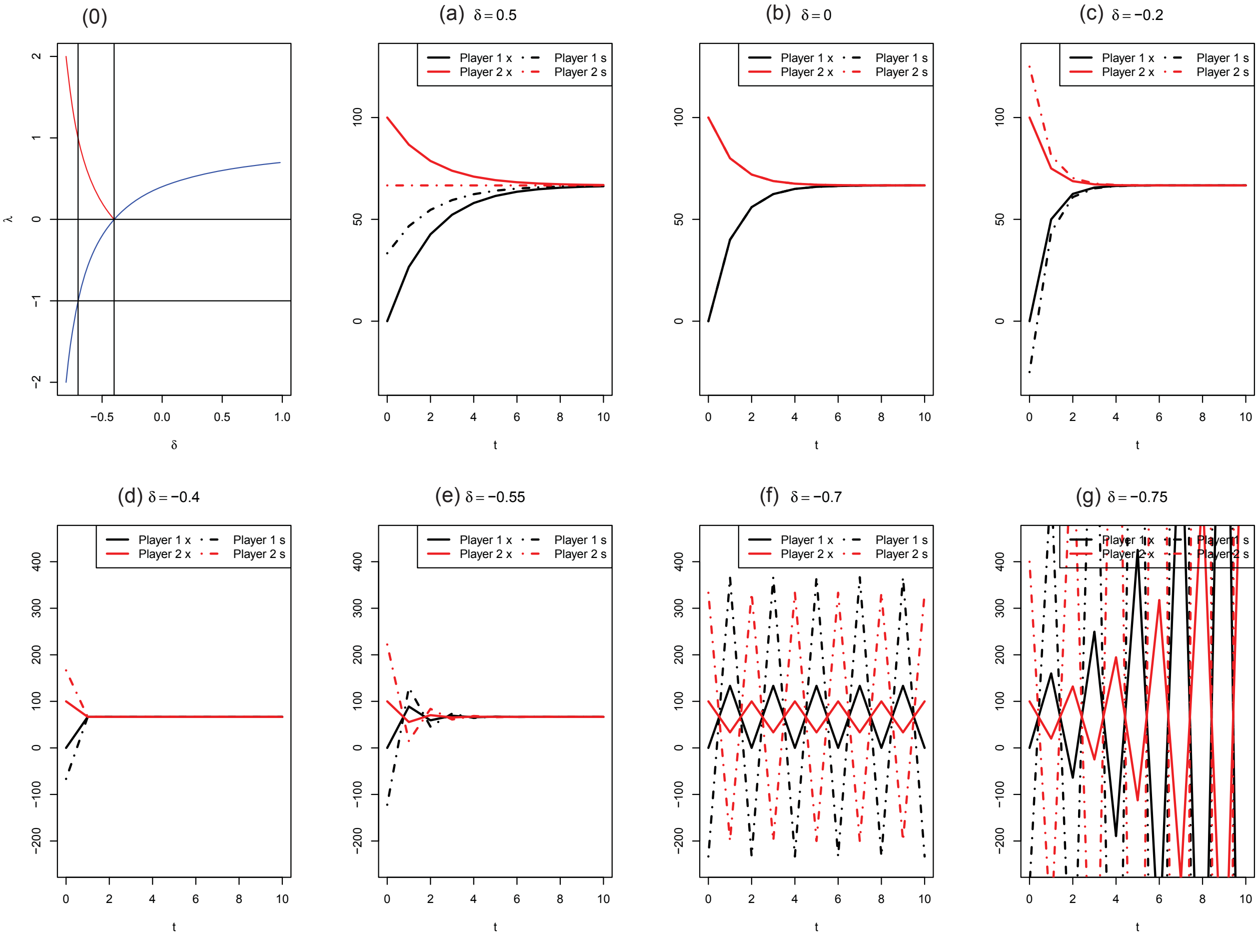

Figure 3: Seven cases of two-agent dynamics for $\delta_1 = \delta_2 = \delta$. Solid lines represent true opinions and dashed lines display stated opinions. (0) Shape of λ. (a) $\delta > 0$, conformity. (b) $\delta = 0$, honesty. (c) $-0.4 < \delta < 0$, smooth convergence under counter-conformity. (d) $\delta = -0.4$, one-step convergence. (e) $\delta < -0.4$, alternating dynamics with convergence. (f) $\delta = -0.7$, alternating dynamics ($\lambda = -1$). (g) $\delta < -0.7$, divergence.

condition for convergence of M^t ($\lambda > -1$) is equivalent to

$$g_{12}\frac{1-\delta_2}{1-\delta_1\delta_2} + g_{21}\frac{1-\delta_1}{1-\delta_1\delta_2} < 2. \tag{9}$$

To interpret this condition in terms of individual conformity parameters, let us distinguish two cases:[17]

(i) If $\delta_2 \leq \frac{2g_{21}+g_{12}-2}{2+g_{12}}$, then M^t converges if and only if $\delta_1 > \frac{g_{12}(1-\delta_2)+g_{21}-2}{g_{21}-2\delta_2}$.

(ii) If $\delta_2 > \frac{2g_{21}+g_{12}-2}{2+g_{12}}$, then M^t converges for any $\delta_1 \in (-1, +1)$.

Thus, if Player 2 has a relatively low degree of conformity (case (i)), then Player 1 must be

[17] It can be checked that the threshold which defines the two cases is always in $(-1, \frac{1}{3})$. Additionally, given that (i) holds, the threshold for δ_1 is below 1.

sufficiently conforming in order to assure convergence. However, if Player 2's conformity is above some threshold, then we will have convergence for any conformity level of Player 1. In fact, $\delta_2 > \frac{1}{3}$ is sufficient for (ii) to hold. Since similar arguments can be made by exchanging the players' labels, in the two-agent case we always have convergence if there is an agent with $\delta_i > \frac{1}{3}$. Thus, a sufficiently conforming agent will reach consensus with any other agent.

4 Opinion Dynamics

To study the dynamics of opinions of n agents, we first elaborate on the properties of steady states and the relation of true, perceived, and stated opinion. We then provide conditions for convergence of opinions and finally determine where opinions converge to. We establish these necessary results in a general and formal way in Sections 4.1–4.4 before we turn to the interpretation, in particular of the main result (Theorem 1), in Section 4.5.

4.1 Perceived, True, and Stated Opinions in Steady States

The dynamics of stated opinions $s(t)$ can be derived from the dynamics of $x(t) = M^t x(0)$ since Proposition 1 determines $s(t)$ in dependence of $x(t)$. Also, the dynamics of $q(t)$ are determined by $s(t)$ since each perceived opinion $q_i(t)$ is exogenously defined as some weighted average of $s(t)$.

If it is the case that dynamics eventually settle down, we have $x(t+1) = x(t)$, which is equivalent to $Mx(t) = x(t)$. In general, we define $z \in \mathbb{R}^n$ to be a steady state of the opinion dynamics if $Mz = z$, i.e. if it is a (right-hand) eigenvector of M corresponding to the eigenvalue 1. Considering the characteristic equation $\det(I - M) = 0$, we can rewrite its argument with use of (6) as follows:

$$I - M = \left[I - (G - D)\Delta(I - D)^{-1}\right]^{-1}(I - G), \tag{10}$$

as shown in the Appendix A.2. Proposition 2 uses this expression to clarify the relation between perceived, stated, and true opinions in a steady state.

Proposition 2 (Steady States). *1. The following statements are equivalent:*

(a) x is a steady state, i.e. $Mx = x$,

(b) $Gx = x$,

(c) perceived and true opinions coincide, i.e. $q = x$,

(d) perceived and stated opinions coincide, i.e. $q = s$.

2. *If* $s = x$*, then* $\delta_i(Gx - x)_i = 0$ *for all agents* $i \in \mathcal{N}$*. If* $\delta_i \neq 0$ *for all agents* $i \in \mathcal{N}$*, then* $s = x$ *implies that* x *is a steady state.*

The equivalence between $Mx = x$ and $Gx = x$ should not be misinterpreted. It does not mean that both dynamics $M^t x(0)$ and $G^t x(0)$ converge to the same vector of opinions. What this condition really means can best be seen when G is irreducible, i.e. every agent interacts (at least indirectly) with everybody else. Then, since G is row stochastic, $Gx = x$ is equivalent to $x_i = x_j$ for all $i, j \in \mathcal{N}$. In this case, all those opinion profiles are steady states of G, where every agent has the same opinion. We call this a *consensus*. Only consensus opinions can then be "steady states of G" (i.e. $Gx = x$) in case of irreducibility of G and hence of M. Thus, the opinion dynamics in our model (according to M) only lead to steady states that are also steady states of the special case with $\delta_i = 0$ for all i (i.e. the classic DeGroot model), but they do in general not lead to the same vector of opinions when starting with some vector $x(0)$. Further, the equivalence between $Mx = x$ and $Gx = x$ implies that the rows of M always sum up to 1. This is true since G is row stochastic and hence $G\mathbb{1} = \mathbb{1}$ (where $\mathbb{1}$ is the vector of ones) and thus by the above result $M\mathbb{1} = \mathbb{1}$. Note however that, in contrast to G, M may have negative entries or even entries larger than 1.

Proposition 2 part 1 also shows that in a steady state true opinion, stated opinion and perceived opinion of any agent agree (since $x = q = s$). This is only true in a steady state. However, the fact that true opinion x and stated opinion s coincide is not sufficient for a steady state. The reason is simply that an honest agent ($\delta_i = 0$) always reports her opinion truthfully no matter of being in a steady state or not. In Part 2, however, we show that if agents are dishonest ($\delta_i \neq 0$ for all $i \in \mathcal{N}$), then all opinions are reported truthfully ($x = s$) only in a steady state.

In the following we study the long-run dynamics. Since $x(t) = M^t x(0)$ it is straightforward to see that the opinion dynamics $x(t)$ converges to a steady state (for any given initial opinion profile $x(0)$) if and only if M^t converges. From Proposition 2 it follows that in this case also $q(t)$ and $s(t)$ converge. Note that we may also have convergence of opinions $x(t)$ if M^t diverges. This can be most easily seen if every agent starts with the same opinion (i.e. $x_i(0) = x_j(0)$ for all $i, j \in \mathcal{N}$). Then from Proposition 2 we get one-step convergence of $x(t)$. This may also happen in the classical DeGroot model, i.e. such that $\delta_i = 0$ for all i.[18] However, in any case – whether or not M^t converges – it is possible to show that in our model the true opinions $x(t)$ converge if and only if the stated opinions $s(t)$ converge, which is equivalent to convergence of perceived opinions $q(t)$ (see Appendix A.3, Lemma A.2). Moreover, all converge to the same limit. Therefore, throughout the paper, we restrict our analysis to the dynamics of true opinions $x(t)$.

[18] See Berger (1981) for a necessary and sufficient condition for convergence of opinions in the DeGroot model.

4.2 The Structure of the Society

While some of the intuition gained in the two-agent case will generalize, there are features of larger networks that cannot appear between only two agents: in the two-agent case, both agents necessarily interact with one another since we assume that $g_{ii} < 1$ for all $i \in \mathcal{N}$. When considering opinion dynamics with n agents, there can be agents who are not influenced at all by one another, or where the influence is only one-way. We thus consider a partition of the agent set $\mathcal{N}$ such that the agents are ordered into groups which are determined by the interaction patterns, i.e. the paths in the network implied by G.

Definition 1. *Let $\Pi(\mathcal{N}, G) = \{\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_K, \mathcal{R}\}$ be a partition of $\mathcal{N}$ into $K(\geq 1)$ groups and the (possibly empty) rest of the world $\mathcal{R}$ such that:*

- *Each group $\mathcal{C}_k$ is strongly connected, i.e. $\forall i, j \in \mathcal{C}_k$ there exists $l \in \mathbb{N}$ such that $(G^l)_{ij} > 0$.*
- *Each group $\mathcal{C}_k$ is closed, i.e. $\forall i \in \mathcal{C}_k$, $G_{ij} > 0$ implies $j \in \mathcal{C}_k$.*
- *The (possibly empty) rest of the world consists of the agents who do not belong to any group, i.e. $\mathcal{R} = \mathcal{N} \setminus \bigcup_{k=1}^{K} \mathcal{C}_k$.*

With a suitable renumeration, the matrix G can be organized into blocks which correspond to the groups of the partition $\Pi(\mathcal{N}, G)$:

$$G = \begin{pmatrix} G_{11} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & G_{KK} & 0 \\ G_{\mathcal{R}1} & \cdots & \cdots & G_{\mathcal{R}K} & G_{\mathcal{R}\mathcal{R}} \end{pmatrix} \tag{11}$$

with $G_{kk} = G|_{\mathcal{C}_k}$, $G_{\mathcal{R}\mathcal{R}} = G|_{\mathcal{R}}$, and $G_{\mathcal{R}k}$ consisting of the rows of G belonging to $\mathcal{R}$ and the columns of G belonging to $\mathcal{C}_k$. This kind of organizing the agents into groups and organizing the matrix into blocks is standard in the literature based on the DeGroot model (e.g. DeMarzo et al., 2003; Golub and Jackson, 2010). Proposition 3 explicitly shows that M – and in fact M^t, for all $t \in \mathbb{N}$ – has the same block structure as G.[19] Moreover, it characterizes M^t.

Proposition 3 (Blocks)**.** *Let G be given as in* (11)*, i.e. organized into blocks according*

[19] This result is not self-evident. It crucially depends on the definition of the reference opinion $q_i^{\tau}(t)$.

to the partition $\Pi(\mathcal{N}, G) = \{\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_K, \mathcal{R}\}$. *Then for every* $t = 1, 2, ...$ *we have*

$$M^t = \begin{pmatrix} M_{11}^t & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & M_{KK}^t & 0 \\ (M^t)_{\mathcal{R}1} & \cdots & \cdots & (M^t)_{\mathcal{R}K} & M_{\mathcal{R}\mathcal{R}}^t \end{pmatrix}$$

with

$$M_{kk}^t = [I - \left(I - (G_{kk} - D_{kk})\Delta_{kk}(I - D_{kk})^{-1}\right)^{-1}(I - G_{kk})]^t$$

for all $k = 1, \ldots, K, \mathcal{R}$, *and*

$$(M^t)_{\mathcal{R}k} = \sum_{l=0}^{t-1} M_{\mathcal{R}\mathcal{R}}^l M_{\mathcal{R}k} M_{kk}^{t-1-l},$$

where $M_{\mathcal{R}k} = (I - (G_{\mathcal{R}\mathcal{R}} - D_{\mathcal{R}\mathcal{R}})\Delta_{\mathcal{R}\mathcal{R}}(I - D_{\mathcal{R}\mathcal{R}})^{-1})^{-1}G_{\mathcal{R}k}[(I - \Delta_{kk}(I - D_{kk})^{-1}(G_{kk} - D_{kk}))^{-1}(I - \Delta_{kk})]$ *for all* $k = 1, \ldots, K$.

Concerning the block structure of M^t and considering that $x(t) = M^t x(0)$, Proposition 3 shows that the opinion dynamics of each group $\mathcal{C}_k$ can be studied independently. Only for agents in $\mathcal{R}$ multiple groups may matter. The agents in $\mathcal{R}$, on the other hand, do not affect the dynamics within groups. More importantly, Proposition 3 provides an explicit expression for M^t and thus for the sequence of true opinions (since $x(t) = M^t x(0)$). Let us now investigate the limit of this sequence.

4.3 Conditions for Convergence

From Proposition 3 it becomes apparent that the dynamics of the different closed and strongly connected groups are independent. Therefore, it is necessary for convergence of M^t that for any group $\mathcal{C}_k$ the relevant block M_{kk}^t converges for $t \to \infty$. To see that this is not sufficient, consider the following example.

Example 2. *Suppose there are four agents such that* $G = \begin{bmatrix} 0.7 & 0.3 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0.085 & 0.085 & 0.49 & 0.34 \\ 0.085 & 0.085 & 0.34 & 0.49 \end{bmatrix}$.
Thus Players 1 *and* 2 *form a closed and strongly connected group* $\mathcal{C}_1$, *while Players* 3 *and* 4 *are the rest of the world* $\mathcal{R}$. *Let the conformity parameter* δ *be given by* $\delta = (0, 0, \delta_{ROTW}, \delta_{ROTW})$. *Figure 4 shows the opinion dynamics for the cases* $\delta_{ROTW} = -.75$ *and* $\delta_{ROTW} = -.9$. *While convergence within the closed and strongly connected group is*

guaranteed, the rest of the world (ROTW) may cause divergence of M^t for $t \to \infty$.[20]

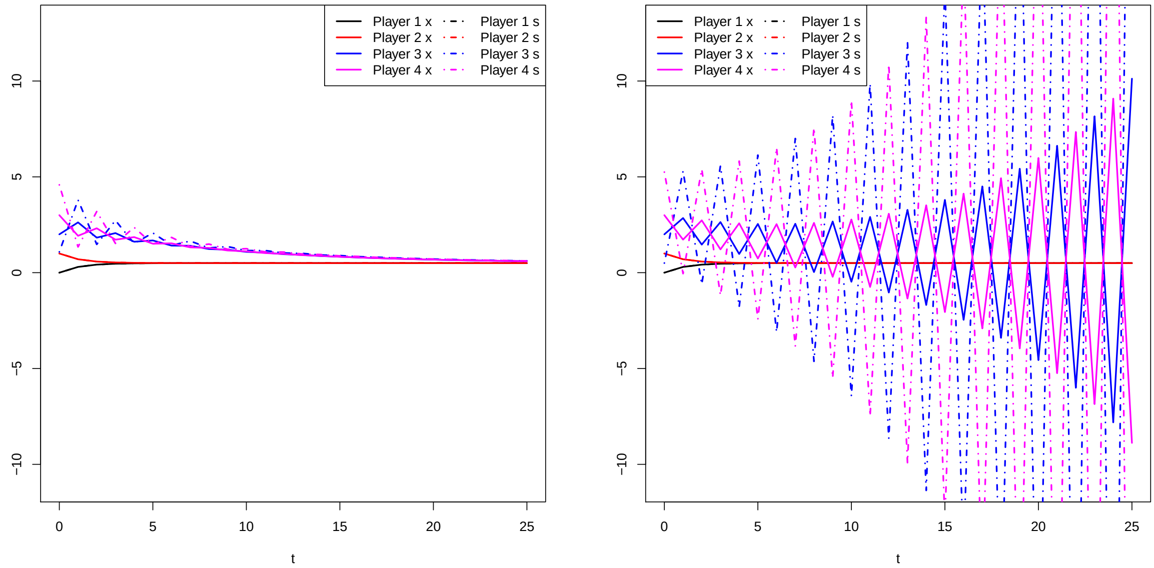

Figure 4: The opinion dynamics of Example 2 for (a) $\delta_{ROTW} = -.75$ and (b) $\delta_{ROTW} = -.9$.

Thus, convergence of all closed and strongly connected groups M_{kk}^t is not sufficient for convergence of M^t. In Proposition 4, we identify the additional condition on the rest of the world such that M^t converges.

Proposition 4 (First convergence result)**.** *Let the block structure of M be given as in Proposition 3. M^t converges for $t \to \infty$ if and only if M_{kk}^t converges for all $k = 1, \ldots, K$ and $M_{\mathcal{RR}}^t$ converges to 0.*

Proposition 4 presents a necessary and sufficient condition for convergence of M in terms of the block structure. In Example 2 the condition that $M_{\mathcal{RR}}$ converges to 0 fails since strong counter-conformity of two agents leads to eigenvalues with high absolute value to the extent that $|\lambda_{\mathcal{RR}}| > 1$, for some eigenvalue of $M_{\mathcal{RR}}$. A similar violation of the necessary condition for convergence occurs if counter-conformity of agents in the closed and strongly connected groups is too strong. Thus, one can derive the intuition that strong counter-conformity may cause divergence. The following result presents simple conditions on the degree of conformity and the interaction structure that ensure convergence of the opinion dynamics.

[20] Notice that, for the latter case, M not only has negative entries but also entries larger than unity:

$$M = \begin{pmatrix} 0.7 & 0.3 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0.053125 & 0.053125 & -0.115625 & 1.009375 \\ 0.053125 & 0.053125 & 1.009375 & -0.115625 \end{pmatrix}.$$

Proposition 5 (Second convergence result)**.** *M^t converges for $t \to \infty$ if $\forall i \in \mathcal{N}$ we have $g_{ii} > 0$ and $\delta_i \geq 0$.*

The condition presented here is fairly weak. If we exclude counter-conformity ($\delta_i \geq 0$), and every individual has at least some self-confidence, then the opinion dynamics converges. Although all cases of conformity are covered by Proposition 5, it is important to emphasize that this condition is not necessary for convergence. Examples of convergence which include counter-conforming agents are given in Examples 1, 2 and in Section 3.

4.4 Long-run Opinions

For the remainder, we now assume that the power series M^t converges. Although conformity is sufficient for convergence, we do not explicitly assume this.

We are now left to address where opinions converge to (in the long-run) when starting with some opinion profile $x(0)$. The answer to this question depends on the learning matrix G and the conformity parameters δ_i. We are particularly interested in the influence of each agent's initial opinion on the long-run opinion given her network position and her degree of conformity. The following result characterizes the long-run opinions explicitly (conditional on convergence). We present it first in a formal way and turn to its interpretation in Section 4.5.

Theorem 1. *Let G and M be organized as in Proposition 3. We denote by $w, v \in \mathbb{R}^n$ the vectors that fulfill the following: for each closed and strongly connected group $\mathcal{C}_k \in \Pi(\mathcal{N}, G)$, $w_{|\mathcal{C}_k}$ is the left unit eigenvector of G_{kk} with $\sum_{i \in \mathcal{C}_k} w_i = 1$, while $v_{|\mathcal{C}_k}$ is left unit eigenvector of M_{kk} with $\sum_{i \in \mathcal{C}_k} w_i = 1$. If M^t converges for $t \to \infty$ to some matrix M^∞, then the following holds:*

$$M^\infty = \begin{pmatrix} M_{11}^\infty & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & M_{KK}^\infty & 0 \\ M_{\mathcal{R}1}^\infty & \cdots & \cdots & M_{\mathcal{R}K}^\infty & 0 \end{pmatrix}$$

with

$$M_{kk}^\infty = \mathbb{1}_{|\mathcal{C}_k} v'_{|\mathcal{C}_k} = \mathbb{1}_{|\mathcal{C}_k} w'_{|\mathcal{C}_k} \frac{I - \Delta_{kk}}{\mathbb{1}'_{|\mathcal{C}_k}(I - \Delta_{kk}) w_{|\mathcal{C}_k}}, \tag{12}$$

and

$$M_{\mathcal{R}k}^\infty = (I - G_{\mathcal{R}\mathcal{R}})^{-1} G_{\mathcal{R}k} M_{kk}^\infty \tag{13}$$

for all $k = 1, \ldots, K$.

Theorem 1, the proof of which can be found in Appendix A.6, fully characterizes the long-run dynamics of (true) opinions given convergence since $x(\infty) = M^{\infty}x(0)$.[21] For the interpretation of the result, we distinguish again between the closed and strongly connected groups $\mathcal{C}_k$ and the rest of the world $\mathcal{R}$.

We can first observe that the long-run opinions may differ across groups, but each closed and strongly connected group $\mathcal{C}_k$ reaches a consensus $c_k \in \mathbb{R}$ as each block M^t_{kk} of M^t converges to a matrix of rank 1. Each row of M^{∞}_{kk} is given by the left-hand unit eigenvector $v'_{|\mathcal{C}_k}$, implying

$$c_k := x_i(\infty) = x_j(\infty) = v'_{|\mathcal{C}_k} x(0)_{|\mathcal{C}_k} \tag{14}$$

for all agents i, j in group $\mathcal{C}_k$. The left-hand normalized unit eigenvector $v'_{|\mathcal{C}_k}$ thus displays the extent to which the initial opinion of each agent i matters for consensus within group $\mathcal{C}_k$. Moreover, $v'_{|\mathcal{C}_k}$ is a function of $w'_{|\mathcal{C}_K}$, the left-hand unit eigenvector of G_{kk}, and the conformity parameters within the group, Δ_{kk}. We delay the interpretation of this result and its comparative statics to the next subsection.

The long-run opinion of an agent in the ROTW $\mathcal{R}$ is simply some weighted average of the long-run opinions $c_1, \dots, c_K$ within the groups $1, ..., K$.[22] To see this, consider the matrix

$$\Gamma := (I - G_{\mathcal{R}\mathcal{R}})^{-1}(G_{\mathcal{R}1}\mathbb{1}_{|\mathcal{C}_1}, \dots, G_{\mathcal{R}K}\mathbb{1}_{|\mathcal{C}_K}),$$

which is easily seen to be row-stochastic. Γ enables translating (13) into

$$x(\infty)_{|\mathcal{R}} = \Gamma c \tag{15}$$

combining the long-run opinions of the closed and strongly connected groups denoted by the K-dimensional vector $c = (c_1, \dots, c_K)'$. Thus, the initial opinion of some agent in the ROTW does not affect the long-run opinion profile $x(\infty)$ since the ROTW agents end up with a weighted average of the consensus opinions of the closed and strongly connected groups, which in turn are dependent on the initial opinions within those groups. Moreover, the weights of averaging depend on G but not on the conformity parameters δ_i for $i \in \mathcal{R}$. Consequently, the long-run opinion of an agent in the ROTW neither depends on an initial opinion nor on the conformity parameter of any agent within the ROTW (including herself). Since each agent in the ROTW may average differently between consent opinions of the closed and strongly connected groups, the agents in the ROTW need not reach a consensus if there is more than just one closed and strongly connected group. The important contribution of Theorem 1 lies in the characterization of v as a

[21] The dynamics collapses to the well-known DeGroot dynamics if every agent i is honest, i.e. Δ is a matrix of zeros.

[22] This result is fully analogous to theorem 10 in DeMarzo et al. (2003).

function of w and Δ, as we will discuss next.

4.5 Opinion Leadership

To simplify the discussion, let us now restrict attention to one closed and strongly connected group by assuming that there is only one such group, i.e. $\Pi(\mathcal{N}, G) = \mathcal{N}$. For this purpose it is sufficient to assume that G is strongly connected or, equivalently, that $\mathrm{rk}(I - G) = n - 1$.

From (14), we get that $x(\infty) = \mathbb{1}v'x(0)$ and hence $x_j(\infty) = v'x(0) = \sum_{i\in\mathcal{N}} v_i x_i(0)$. Thus, an entry v_i of v determines the weight of the initial opinion of agent i on the long-run consensus opinion of her group. This is a very intuitive formalization of opinion leadership: v measures the *power* of each agent in the group.

Note that for $\delta_i = 0$ for all $i \in \mathcal{N}$, (12) yields $v = w$, i.e. opinion leadership is fully determined by the unit eigenvector of G. w is a well-studied object in network science: it is known as eigenvector centrality of the transposed social network G' (Bonacich, 1972; Friedkin, 1991).[23]

When relaxing the assumption that every agent is honest, then the following Corollary of Theorem 1 shows how opinion leadership is not only determined by eigenvector centrality, but also by the degree of conformity.

Corollary 1. *Let* $\mathrm{rk}(I - G) = n - 1$. *Let w and v be the normalized left-hand unit eigenvectors of G and M, respectively. Then we have for any $i \in \mathcal{N}$*

$$v_i = \frac{(1-\delta_i)w_i}{\sum_{j\in\mathcal{N}}(1-\delta_j)w_j}. \tag{16}$$

Moreover,

$$\frac{\partial v_i}{\partial \delta_k} = \frac{w_k}{\sum\limits_{j=1}^{n} w_j(1-\delta_j)} \left(\frac{w_i(1-\delta_i)}{\sum\limits_{j=1}^{n} w_j(1-\delta_j)} - 1_{i=k} \right) = \frac{w_k}{\sum\limits_{j=1}^{n} w_j(1-\delta_j)} \left(v_i - 1_{i=k}\right). \tag{17}$$

As it becomes apparent from (16) opinion leadership (power) v_i of some agent i is determined by the *combination* of her network centrality in G (w_i) and the individual conformity δ_i divided by the sum of these values over all agents. Thus, there is a complementary relationship between network centrality and $1 - \delta_i$: power becomes minimal ($v_i \to 0$) if either i's network centrality approaches zero or if i is fully conform ($\delta_i \to 1$).

Taking the network G as given, we can observe the comparative statics with respect

[23]This index of centrality in a social network is recurrently defined via the rows of G' (i.e. via the columns of G): An agent's centrality is the weighted sum of centralities of the agents who listen to her.

to δ_i. From (17) we get for all $i \in \mathcal{N}$ that opinion leadership is decreasing in "own" conformity δ_i and increasing in other agents' conformity δ_k, $k \neq i$, since $w_j \in [0,1]$ and $1-\delta_j \geq 0$ for all $j \in \mathcal{N}$. Thus, low own conformity fosters opinion leadership. The same is true if other agents are more conforming. We also may use (17) to examine which agent's power changes most in response to a marginal increase in her own conformity. From (17), we calculate that

$$\left|\frac{\partial v_i}{\partial \delta_i}\right| < \left|\frac{\partial v_j}{\partial \delta_j}\right| \quad \Leftrightarrow \quad w_j^2(1-\delta_j) - w_i^2(1-\delta_i) < (w_j - w_i)\sum_{k=1}^{n} w_k(1-\delta_k). \tag{18}$$

Thus, if two agents have the same network centrality $w_i = w_j$, then by (18), $\left|\frac{\partial v_i}{\partial \delta_i}\right| < \left|\frac{\partial v_j}{\partial \delta_j}\right|$ if and only if $\delta_i < \delta_j$. In other words, the agent with the already higher degree of conformity and thus lower power loses even more power in response to a marginal increase in conformity compared with an agent with low conformity. Holding $\delta_i = \delta_j$, we get $\left|\frac{\partial v_i}{\partial \delta_i}\right| < \left|\frac{\partial v_j}{\partial \delta_j}\right|$ if and only if $w_i < w_j$, which implies that for two agents with equal conformity the agent with the higher network centrality loses more power when increasing own conformity.

We can also use Corollary 1 to compare opinion leadership in our model, v, with opinion leadership in the classic DeGroot model, w, (i.e. with the special case of our model where every agent i is honest, $\delta_i = 0$). For this purpose consider first a society where all agents are characterized by the same trait, i.e. $\delta_j = \bar{\delta}$ for all $j \in \mathcal{N}$. Then (16) yields $v = w$: opinion leadership is not affected by conformity when all agents are characterized by the same level of conformity. More generally, we have $v_i \geq w_i$ if and only if $\delta_i \leq \sum_{j \neq i} \frac{w_j}{\sum_{k \neq i} w_k}\delta_j$, i.e. an agent's power in our model compared to the classic DeGroot model is fostered if δ_i is below some average of the others' conformity parameters. This is illustrated in Figure 5 which depicts v_i as a function of δ_i for two different cases.

(a) Here we reconsider the learning matrix G as given in Example 1 with $\delta_1 = -0.5$, $\delta_3 = 0.5$, and study the effect of Player 2's conformity level δ_2 on her power v_2. If Player 2 is honest, her initial opinion's impact on the long-run consensus is 1/3, it completely vanishes for Player 2's conformity level approaching 1, while counter-conformity allows Player 2 to become more important, eventually approaching $v_2 = 0.5$ when δ_2 approaches -1.[24]

(b) Here we reconsider Example 2 with $\delta_1 = -0.7$: in this case, Player 2's ability to gain power is further bounded by the fact that too strong counter-conformity ($\delta_2 \leq -0.7$) leads to divergence of opinions.

[24] One can show that the power gain by counter-conforming is bounded by $v_i(\delta_i) \leq (2 - w_i)v_i(0)$.

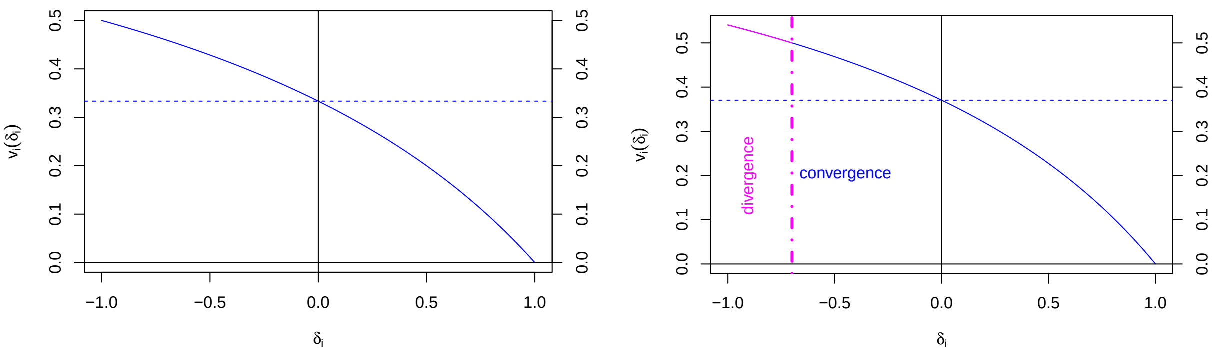

Figure 5: Power as a function of own conformity level.

5 Wisdom

The discussion so far applies to any continuous opinion including those for which no true value can be determined. In some applications, however, agents' opinions are more or less accurate with respect to some objective truth. As in the discrete context of Condorcet's Jury theorem, the question whether agents aggregate information in an efficient way is also of interest in the context of continuous opinions (Golub and Jackson, 2010; Acemoglu et al., 2010).

Therefore we assume that there is some true value $\mu \in \mathbb{R}$ and that all agents of the society receive independent unbiased signals about μ with individual precision (i.e. inverse of the variance) which constitute the agents' initial opinions. Formally, for all $i \in \mathcal{N}$, agent i's initial opinion $x_i(0)$ is a random variable with expected value μ and some individual variance σ_i^2, and all $x_i(0)$ are uncorrelated random variables. Assuming that opinion dynamics converge, a very natural question to ask is how close the different steady state opinions will be to the true, but to the agents unknown, value μ.[25] To measure this difference between μ and an estimate $\hat{\mu}$, we use the mean squared error (MSE), which is defined as $E((\hat{\mu}-\mu)^2)$.[26] The MSE can be decomposed into the squared bias $(E(\hat{\mu}-\mu))^2$ and the estimator's variance $\text{Var}(\hat{\mu})$:

$$E((\hat{\mu}-\mu)^2) = (E(\hat{\mu}-\mu))^2 + \text{Var}(\hat{\mu}).$$

As $x(\infty) = M^\infty x(0)$ and $M^\infty \mathbb{1} = \mathbb{1}$, it is obvious that $E(x(\infty)) = \mu\mathbb{1}$, i.e. all agents' long-run opinions are unbiased estimates for μ. Denoting by Σ the covariance matrix of $x(0)$, the corresponding MSEs are therefore given by the entries on the diagonal of

[25] Recall that in a steady state true opinions and stated opinions coincide and there is consensus within groups.

[26] The mean squared error as a measure of wisdom has also been used by Rauhut and Lorenz (2010).

$M^\infty \Sigma (M^\infty)'$. To study the effects of conformity on wisdom, we begin with an illustrative example.

5.1 Wisdom: an Example

Let $n = 10$, $(\sigma_1^2, \ldots, \sigma_{10}^2)$=(6, 4, 8, 7, 6, 3, 10, 12, 14, 16), and

$$G = \begin{pmatrix} 0.9 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.4 & 0.6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.3 & 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0.1 & 0 & 0 & 0.1 & 0 & 0 & 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0 & 0.2 & 0.6 \end{pmatrix}.$$

In this situation, we have $K = 3$ closed and strongly connected groups, $\mathcal{C}_1 = \{1, 2\}$, $\mathcal{C}_2 = \{3, 4\}$, and $\mathcal{C}_3 = \{5, 6\}$, while Players 7 to 10 form the rest of the world. If all agents report their opinions truthfully ($\Delta = 0$), we find the MSEs equal to (4, 4, 4, 4, 2.25, 2.25, 4, 4, 2, 1.0625). There are several notable features of this observation. First, due to the fact that their long-run opinions are equal, all agents within a given closed and strongly connected group share the same level of wisdom. Comparing the first two groups, we note that the MSEs of these two groups are 4 each, although the first group enjoys significantly better initial signals (of variances 6 and 4), while the second group seems to combine their less precise signals (of variances 8 and 7) much more effectively. It is also remarkable that Player 2, by communicating with Player 1, ends up with exactly the same MSE of 4 that she would reach if she used only her own signal. With respect to the rest of the world, notice that these agents typically have different MSEs. Furthermore, Players 7 and 8 each end up with the same MSE as the first two groups, while Players 9 and 10 achieve MSEs better than all members of the closed and strongly connected groups.

Now suppose that Players 2, 3, and 5 are conforming with $\delta_2 = 5/9$, $\delta_3 = 2/3$, and $\delta_5 = 1/2$ (and $\delta_i = 0$ for all other players). Then wisdom levels can be calculated to be (4.9, 4.9, 4, 4, 2, 2, 4.9, 4, 2.225, 1.05625). Thus, increasing conformity can lead to a decrease in wisdom (as the first group's MSE becomes larger), the same wisdom (as the second group's MSE does not change), or an increase in wisdom (as the third group's MSE becomes smaller). We also find that the agents in the rest of the world are affected by the changes in conformity of the agents in the closed and strongly connected groups:

the MSE of Players 7 and 9 increases, while Player 10's MSE decreases slightly. It still holds that Player 7 and 8's MSEs equal that of the first and second group, respectively.

We will now proceed by systematically analyzing the principles underlying the distribution of wisdom within the society.

5.2 Wisdom of Groups

Due to (14), a group $\mathcal{C}_k$ will, given convergence, eventually end up reaching a consensus where all agents' opinions are equal to $c_k = v'_{|C_k} x(0)_{|C_k} =: \hat{\mu}_k$. Hence, we can directly derive group $\mathcal{C}_k$'s wisdom as the MSE of $\hat{\mu}_k$.

Lemma 1. *The MSE of $\hat{\mu}_k$ is given by*

$$\mathrm{MSE}_k := E((\hat{\mu}_k - \mu)^2) = \sum_{i \in \mathcal{C}_k} v_i^2 \sigma_i^2 = \sum_{i \in \mathcal{C}_k} \left(\frac{(1-\delta_i) w_i}{\sum\limits_{j \in \mathcal{C}_k} (1-\delta_j) w_j} \right)^2 \sigma_i^2 .$$

We may use Lemma 1 to identify the individual contributions to the MSE in a given group $\mathcal{C}_k$. First, from Lemma 1 it follows directly that

$$\mathrm{MSE}_k = \sum_{i \in \mathcal{C}_k} v_i^2 \sigma_i^2 \leq \sum_{i \in \mathcal{C}_k} v_i \sigma_i^2 \leq \max_{i \in \mathcal{C}_k} \sigma_i^2, \tag{19}$$

since $v_i^2 \leq v_i$ due to $v_i \in (0,1]$ for all agents i. Thus, group $\mathcal{C}_k$'s long-run opinion is on average at least as close to the true value μ as that of the agent with the least precise signal. This worst case is given when both inequalities in (19) become equalities, which is the case for $v_i \in \{0,1\}$ for all $i \in \mathcal{C}_k$ (first inequality) and $v_i = 0$ for all i with $\sigma_i^2 < \max_{j \in \mathcal{C}_k} \sigma_j^2$ (second inequality). Therefore, information updating within group $\mathcal{C}_k$ is worst when importance is given to only one agent whose signal is most imprecise. This case would be approached if all other agents were close to full conformity, i.e. δ_i close to 1. We now consider the comparative static effect of one agent's conformity on the wisdom of her group.

Proposition 6. *The wisdom of a closed and strongly connected group C_k is increasing in the conformity level of a group member i if and only if i's product of signal variance and power is larger than the group's MSE, i.e.*

$$\frac{\partial \, \mathrm{MSE}_k}{\partial \delta_i} \leq 0 \;\Leftrightarrow\; v_i \sigma_i^2 \geq \mathrm{MSE}_k \,.$$

To give an interpretation for Proposition 6, let us rewrite $v_i \sigma_i^2 = \frac{v_i}{1/\sigma_i^2}$ and $\mathrm{MSE}_k = \sum\limits_{j \in \mathcal{C}_k} v_j \frac{v_j}{1/\sigma_j^2}$. This shows that it is not a person's expertise alone which is decisive for the question of how this person can increase the group's wisdom, rather, it is the ratio of power over signal precision, $\frac{v_i}{1/\sigma_i^2}$: if agents with a high ratio as compared to the group's average are more conforming, then this will reduce their power within the group, decrease the group's MSE, and thereby increase its wisdom. Vice versa, agents who are not powerful enough in relation to their signal precision will increase the group's wisdom if they are less conforming, because this will increase their power, decrease the group's MSE, and foster its wisdom.[27]

The above discussion implies that in the best possible case, the ratio of power over signal precision is constant within a group: $v_i\sigma_i^2 = v_j\sigma_j^2$ for all $i, j \in \mathcal{C}_k$. This is formalized in the following corollary of Proposition 6.

Corollary 2. *For the wisdom of group $\mathcal{C}_k$ as measured by* MSE_k*, we have*

$$\mathrm{MSE}_k \geq \frac{1}{\sum\limits_{j\in\mathcal{C}_k} \frac{1}{\sigma_j^2}} =: \mathrm{MSE}_k^*, \tag{20}$$

with equality in (20) *if and only if $v_i\sigma_i^2 = v_j\sigma_j^2$ for all $i, j \in \mathcal{C}_k$. The latter condition is equivalent to*

$$\delta_i = 1 - a\,\frac{1}{\sigma_i^2 w_i \sum\limits_{j\in\mathcal{C}_k} \frac{1}{\sigma_j^2}} \quad \textit{for all } i \in \mathcal{C}_k \tag{21}$$

for some constant $a \in (0, 2\sum\limits_{j\in\mathcal{C}_k} \frac{1}{\sigma_j^2} \min\limits_{j\in\mathcal{C}_k} w_j\sigma_j^2)$.

Corollary 2 delivers the analogue to (19). While (19) describes the worst case with respect to wisdom, Corollary 2 considers the best scenario: all agents within the same closed and strongly connected group share the same ratio of power over signal precision, and this case can always be constructed if the agents' conformity is distributed suitably. In particular, choosing $a \in (0, \sum\limits_{j\in\mathcal{C}_k} \frac{1}{\sigma_j^2} \min\limits_{j\in\mathcal{C}_k} w_j\sigma_j^2]$ in (21) ensures $\delta_i \geq 0$ for all $i \in \mathcal{C}_k$ and therefore by Proposition 5 guarantees convergence of the opinions in $\mathcal{C}_k$ to the best possible consensus $\hat{\mu}_k$. Notice also that the optimal MSE is smaller than individual signal variance σ_i^2 for all agents i in group $\mathcal{C}_k$, as is easily seen from (20). Therefore, under optimal conformity all agents within $\mathcal{C}_k$ benefit from communication.

Reconsidering the example discussed in Subsection 5.1, we find the network centralities (the left-hand unit eigenvectors of G) to be $w_1 = 0.8$, $w_2 = 0.2$, $w_3 = 0.6$, $w_4 = 0.4$, $w_5 = 0.5$, and $w_6 = 0.4$. Therefore, in (21) the constant a can be chosen in $(0, 2/3)$ (group 1) and $(0, 3/2)$ (groups 2 and 3). Choosing $a = 1/3$ (group 1) and $a = 3/4$ (groups 2 and 3) delivers $\delta_1 = 5/6$, $\delta_3 = 5/12$, and $\delta_5 = 1/2$ (and $\delta_i = 0$ for all other agents). Thus, choosing the agents' degrees of conformity according to these values ensures the

[27] An analogous discussion can be already found in DeMarzo et al. (2003) for the case where agents are honest.

optimal wisdom within the respective groups, given by (2.4, 2.4, $3.7\overline{3}$, $3.7\overline{3}$, 2, 2, 2.4, $3.7\overline{3}$, $1.5\overline{3}$, $0.88\overline{3}$). The same level could also be reached for other conformity levels, for instance, choosing $a = 1/4$ (first group), $a = 3/7$ (second group), and $a = 3/8$ (third group) in (21), we find that the conformity levels $\delta_{1:6} = (7/8,\ 1/4,\ 2/3,\ 3/7,\ 3/4,\ 1/2)$ also lead to the optimal wisdom. Notice that, as in Golub and Jackson (2010), wisdom thus is independent of the speed of convergence, as we have two examples with the same optimal wisdom but different speeds of convergence (the last-mentioned conformity levels lead to slightly slower convergence than the earlier mentioned ones).

5.3 Wisdom within the Rest of the World

Let us recall that agents in the rest of the world do not necessarily share a consensus opinion in the long-run, so that we will typically have individual wisdom levels. Due to (15), we have the following formula for the long-run opinions within the rest of the world: $x(\infty)_{|\mathcal{R}} = \Gamma\hat{\mu}$, with $\hat{\mu} := (\hat{\mu}_1, \dots, \hat{\mu}_K)'$. Therefore, the wisdom levels in the rest of the world depend on the conformity levels of the agents in the closed and strongly connected groups as these affect the consensus opinions $\hat{\mu}_k$ of these groups. On the other hand, as neither the initial signals nor the conformity levels of the agents in the rest of the world play any role for their long-run opinions, these agents' wisdom is independent of their conformity levels as well as of their initial signals. In other words, if the rest of the world is non-empty, information processing in the society is necessarily inefficient as the information contained in these agents' initial signals is inevitably lost. Assuming convergence, let $\gamma_{i,k}$ denote the long-term weight of the group $\mathcal{C}_k$ on the opinion of agent $i \in \mathcal{R}$, i.e. $x_i(\infty) = \sum_{k=1}^{K} \gamma_{i,k}\hat{\mu}_k$ (cf. (15)). This immediately translates into the wisdom of an agent $i \in \mathcal{R}$ as follows:

$$E((x_i(\infty) - \mu)^2) = \sum_{k=1}^{K} \gamma_{i,k}^2 \,\mathrm{MSE}_k \leq \max_{k=1,\dots,K} \mathrm{MSE}_k\,. \tag{22}$$

The wisdom of an agent in the rest of the world depends on the wisdom within the closed and strongly connected groups. More precisely, an agent i's wisdom only depends on the wisdom of groups $\mathcal{C}_k$ to which there is a directed path in the network G because this corresponds to $\gamma_{i,k} > 0$. The worst case for an agent in the rest of the world is to be influenced only by agents of one closed and strongly connected group with maximal MSE. With regard to the example discussed in subsection 5.1 this is the case for Players 7 and 8 who have directed paths only into group 1 and group 2, respectively, such that they share their MSEs of 4. Player 9, however, who has directed paths into both groups with MSE of 4 reaches an MSE of 2 since the long-term weights $\gamma_{9,1} = 0.5$ and $\gamma_{9,2} = 0.5$ are squared in

(22). Finally, Player 10 has directed paths into these groups via Player 9 and, moreover, has a directed path into group 3. Player 10 therefore is able to combine MSEs of 4, 4, and 2.25 into an MSE as low as 1.0625. It is intuitive that for maximal wisdom of an agent in the rest of the world, all groups' signals have to be accessed with some kind of balanced group weights. The following proposition confirms this intuition.

Proposition 7. *For agents $i \in \mathcal{R}$, we have:*

$$E((x_i(\infty) - \mu)^2) \geq \frac{1}{\sum_{k=1}^{K} \frac{1}{\mathrm{MSE}_k}}, \tag{23}$$

with equality if and only if $\gamma_{i,k} = \frac{1}{\mathrm{MSE}_k \sum_{l=1}^{K} \frac{1}{\mathrm{MSE}_l}}$ *for all* $k = 1, \ldots, K$.

Therefore, the highest wisdom is achieved if an agent in the rest of the world averages the different groups' opinions in such a way that the product of weight put on a group and its MSE is constant for all groups: the better a group's estimate, the more weight it should get. Nevertheless, as all the optimal weights are positive, this optimum can only be achieved if from agent i there is a directed path into all the closed and strongly connected groups. Notice also that the optimal weights depend on the groups' MSEs such that an agent in the rest of the world who is initially characterized by optimal weights would no longer average the groups' opinions optimally if conformity levels within the groups were to change.

It is remarkable that an agent in the rest of the world who is connected to multiple groups can reach a significantly lower MSE than the best informed agents from those groups. Thus, the fact that agents in the rest of the world are absolutely powerless does not imply that they are not wise.

6 Concluding Remarks

So far, the literature on opinion dynamics has focused on truthful opinion representation either with a Bayesian approach (Banerjee, 1992; Bikhchandani et al., 1992; Smith and Sorensen, 2000; Gale and Kariv, 2003; Acemoglu et al., 2011) or assuming naïve updating according to a learning matrix (DeGroot, 1974; DeMarzo et al., 2003; Golub and Jackson, 2010; Acemoglu et al., 2010). Despite some disputable assumptions in both approaches, as Acemoglu and Ozdaglar (2011) point out, these models serve well to study conditions under which societies will eventually reach a state of agreement, i.e. consensus. Moreover, in both contexts the aggregation of initial opinions may, but need not, be "asymptotically efficient," in the sense that social learning leads to a high accuracy of information in the long-run. One basic force fostering efficient information aggregation even among naïve agents is a statistical effect of growing sample size (which is also called “the wisdom of crowds”) such as in Condorcet’s Jury Theorem. On the other hand, prominent agents or opinion leaders might reduce the accuracy of information aggregation by superseding valuable opinions of others.

To our best knowledge, this paper is the first contribution to incorporate misrepresentation of opinions. We assume that individuals depart from their true opinion by conforming or counter-conforming with their peer group which is a well documented phenomenon (Deutsch and Gerard, 1955; Jones, 1984; Zafar, 2011). While we follow the literature based on DeGroot (1974) in modeling *informational social influence* as naïve updating of opinions through the network, we, thus, also model *normative social influence* by including conforming/counter-conforming behavior. In order to study the effects of conformity on long-run opinions and information aggregation, we characterize sufficient conditions for convergence and characterize the long-run opinions in this dynamic framework. When all agents are conforming or honest, then opinions converge (Proposition 5).

Assuming convergence, we then characterize the long-run (consensus) opinion in each closed and strongly connected group under conformity (Theorem 1). Thereby, we are in a position to study the impact of the individual levels of conformity on opinion leadership and on wisdom of the society. Opinion leaders are those whose initial opinion has a high impact on consensus. We find that this influence is increasing in network centrality (as in the DeGroot model), but moreover decreasing in the individual level of conformity (Corollary 1). Thus, taking the network as given, we conclude that low conformity fosters opinion leadership while high conformity undermines opinion leadership. This result is fully in line with empirical evidence that opinion leaders are characterized by a higher inclination to “publicly individuate” themselves (Chan and Misra, 1990). Therefore, counter-conformity might be interpreted as a persuasion device since not only the connected agents’ opinions of next period are swayed towards own opinion but a higher impact on the consensus opinion is achieved.

The effect of heterogeneous levels of conformity on wisdom of the society is ambiguous. Here, wisdom is defined as the mean squared error (MSE) of the consensus opinion where agents’ initial opinions are noisy but unbiased signals about some true state of the world with heterogeneous signal precision. Increasing conformity of a given individual need not undermine the wisdom of the society, but can also enhance it or leave it unchanged. We find that increasing conformity of agents with high power and low signal precision increases the group’s wisdom (Proposition 6). In particular, optimal wisdom within a given closed and strongly connected group is achieved if distribution of conformity levels is such that ratio of power over signal precision is balanced across agents (Corollary 2). This result resembles the fact that reducing prominence of individuals – in particular prominence of uninformed agents – increases the accuracy of information aggregation. While in the previous literature reduction of prominence is achieved by increasing population size (see e.g. Golub and Jackson, 2010), in our model this can be achieved by conformity and therefore also holds for small groups. Finally, when considering agents in the rest of the world, we find that their levels of conformity have no influence on wisdom. Although powerless, individuals in the rest of the world can be quite wise since they may aggregate information from different groups.

The model presented here contains some simplifying assumptions which may be relaxed in future research. First, we assumed that the social network is exogenous and stays fixed over time. In the literature we can find models where the network structure may vary over time such that only agents with "close opinions" are listened to (Hegselmann and Krause, 2002), self confidence varies (DeMarzo et al., 2003), and general changes are possible (Lorenz, 2005). It would be interesting to see how changes in the learning structure, either exogenously or endogenously, affect our results. Second, we assumed that interaction neighborhood equals observation neighborhood in the sense that agents conform or counter-conform with those agents they listen to. If this assumption is relaxed, the group structure may no longer be preserved and interesting applications to lobbying (addressing a certain group) become possible. We leave these ideas and possible other extensions to future research.

A Appendix: Proofs

A.1 Expressed Opinions

Proof of Proposition 1

First, notice that $s(t)$ by construction satisfies $s(t) = (I - \Delta)x(t) + \Delta Y s(t)$ with $Y := (I - D)^{-1}(G - D)$ and that for all $i \in A^\tau(t)$, $s_i^\tau(t)$ is the i-th component of $(I - \Delta)x(t) + \Delta Y s^{\tau-1}(t)$. For all $i \in A^\tau(t)$, we therefore find $s_i^\tau(t) - s_i(t)$ as the i-th component of $\Delta Y\left(s^{\tau-1}(t) - s(t)\right)$. As Y is obviously a row-stochastic matrix, we immediately have $|s_i^\tau(t) - s_i(t)| \leq \delta^* ||s^{\tau-1}(t) - s(t)||_\infty$ for all $i \in A^\tau(t)$, with $\delta^* := \max_{i \in \mathcal{N}} |\delta_i| < 1$, while we have $|s_i^\tau(t) - s_i(t)| = |s_i^{\tau-1}(t) - s_i(t)| \leq ||s^{\tau-1}(t) - s(t)||_\infty$ for all $i \notin A^\tau(t)$. Together, we therefore have that $||s^\tau(t) - s(t)||_\infty \leq ||s^{\tau-1}(t) - s(t)||_\infty$ for all τ, showing that the distance between $s^\tau(t)$ and $s(t)$ measured using the $|| \cdot ||_\infty$-norm is a non-increasing and therefore converging sequence.

Now, let $U_i(t) := \{\tau \in \mathbb{N} : i \in A^\tau(t)\}$, for each agent i. Using the assumption that every agent i belongs almost surely to infinitely many $A^\tau(t)$, we define $\tau_1 := \min\{\tau \in$

$\mathbb{N} : (\forall i \in \mathcal{N})(U_i(t) \cap \{1, \ldots, \tau\} \neq \varnothing)\}$ as the first time-step where every agent has at least once been satisfied with her stated opinion.[28] Given the above, it is easy to see that $||s^{\tau_1}(t) - s(t)||_\infty \leq \delta^* ||s^0(t) - s(t)||_\infty$. Proceeding in the same way by recursively defining $\tau_{k+1} := \min\{\tau > \tau_k : (\forall i \in \mathcal{N})(U_i(t) \cap \{\tau_k + 1, \ldots, \tau\} \neq \varnothing)\}$ as the first time-step after τ_k such that all agents have at least been once satisfied with their stated opinion, we then have $||s^{\tau_k}(t) - s(t)||_\infty \leq (\delta^*)^k ||s^0(t) - s(t)||_\infty$, yielding that $||s^{\tau_k}(t) - s(t)||_\infty$ and therefore also $||s^\tau(t) - s(t)||_\infty$ converges to 0. □

A.2 Rewriting I-M

Lemma A.1 (I-M)**.** $I - M = \left(I - (G - D)\Delta(I - D)^{-1}\right)^{-1} (I - G)$.

Proof of Lemma A.1 (I-M)

First, we can rewrite M, given by (6), to obtain

$$M = G - (G - D)(I - \Delta(I - D)^{-1}(G - D))^{-1}\Delta(I - (I - D)^{-1}(G - D)).$$

This can be verified by the following calculation.

$$\begin{aligned} M &= D + (G - D)(I - \Delta(I - D)^{-1}(G - D))^{-1}(I - \Delta) \\ &= D + (G - D)\big[I - \Delta(I - D)^{-1}(G - D)\big]^{-1}\big[I - \Delta(I - D)^{-1}(G - D) \\ &\quad + \Delta(I - D)^{-1}(G - D) - \Delta\big] \\ &= D + (G - D)(I + \big[I - \Delta(I - D)^{-1}(G - D)\big]^{-1}\big[\Delta(I - D)^{-1}(G - D) - \Delta)\big] \\ &= G - (G - D)\big[I - \Delta(I - D)^{-1}(G - D)\big]^{-1}\Delta\big[I - (I - D)^{-1}(G - D)\big]. \end{aligned}$$

Thus,

$$\begin{aligned} I - M &= I - G + (G - D)\big[I - \Delta(I - D)^{-1}(G - D)\big]^{-1}\Delta(I - D)^{-1}(I - G) \\ &= \Big(I + (G - D)\left[I - \Delta(I - D)^{-1}(G - D)\right]^{-1}\Delta(I - D)^{-1}\Big)(I - G). \qquad \text{(A.1)} \end{aligned}$$

Now, note that for any $n \times m$-matrix A and any $m \times n$-matrix B, with I_k the k-dimensional identity matrix ($k \in \{n, m\}$), we have that $I_n - AB$ is invertible if and only if $I_m - BA$ is invertible, and then $(I_n - AB)^{-1} = I_n + A(I_m - BA)^{-1}B$, since $(I_n + A(I_m - BA)^{-1}B)(I_n - AB) = I_n - AB + A(I_m - BA)^{-1}B - A(I_m - BA)^{-1}BAB = I_n - AB + A(I_m - BA)^{-1}(I_m - BA)B = I_n$. Taking $A = G - D$ and $B = \Delta(I - D)^{-1}$ in (A.1) then gives $I - M = \left(I - (G - D)\Delta(I - D)^{-1}\right)^{-1}(I - G)$. □

[28] The assumption that all $U_i(t)$ are almost surely infinite guarantees that $\tau_1, \tau_2, \ldots$ are almost surely well-defined.

A.3 Steady states

Proof of Proposition 2

1. x is a "steady state of G," i.e. $Gx = x \Leftrightarrow (I-G)x = 0 \Leftrightarrow \left[I-(G-D)\Delta(I-D)^{-1}\right]^{-1}(I-G)x = 0$, since by Lemma A.1 $\left[I-(G-D)\Delta(I-D)^{-1}\right]$ is invertible. Thus by Lemma A.1, $Gx = x$ if and only if $Mx = x$.

   It therefore suffices to show that $Mx = x \Rightarrow q = x \Rightarrow q = s \Rightarrow Mx = x$.

   (a) $x = Mx = Dx + (G-D)s = Dx + (I-D)q$ implies $(I-D)x = (I-D)q$, thus $q = x$.

   (b) $q = x$ implies $s = (I-\Delta)x + \Delta q = (I-\Delta)q + \Delta q = q$.

   (c) $q = s$ implies $s = (I-\Delta)x + \Delta q = (I-\Delta)x + \Delta s$ and therefore $(I-\Delta)s = (I-\Delta)x$ and $s = q = x$, from which we find $Mx = Dx + (I-D)q = Dx + (I-D)x = x$.

2. Suppose $x = s$. Note that $s = \left(I-\Delta(I-D)^{-1}(G-D)\right)^{-1}(I-\Delta)x$ by Proposition 1. Thus,

$$\begin{aligned} x = s \;&\Leftrightarrow\; \left(I-\Delta(I-D)^{-1}(G-D)\right)x = (I-\Delta)x \\ &\Leftrightarrow\; \Delta\left(I-(I-D)^{-1}(G-D)\right)x = 0 \\ &\Leftrightarrow\; \Delta(I-D)^{-1}\left(I-D-(G-D)\right)x = 0 \\ &\overset{(*)}{\Leftrightarrow}\; (I-D)^{-1}\Delta\left(I-D-(G-D)\right)x = 0 \\ &\Leftrightarrow\; \Delta\left(I-G\right)x = 0, \end{aligned}$$

   where (*) holds since $(I-D)^{-1}$ and Δ are diagonal. □

Lemma A.2. *The following statements are equivalent:*

1. *True opinions $x(t)$ converge for $t \to \infty$.*
2. *Stated opinions $s(t)$ converge for $t \to \infty$.*
3. *Perceived opinions $q(t)$ converge for $t \to \infty$.*

Moreover, if the true, stated, and perceived opinions converge, then the limits coincide: $\lim_{t\to\infty} x(t) = \lim_{t\to\infty} s(t) = \lim_{t\to\infty} q(t)$.

Proof of Lemma A.2

From Proposition 1, we get that $s(t) = \left(I - \Delta(I-D)^{-1}(G-D)\right)^{-1}(I-\Delta)x(t)$. Thus convergence of $x(t)$ implies convergence of $s(t)$. By definition we have that $q(t) = (I-D)^{-1}(G-D)s(t)$, and hence convergence of $s(t)$ implies convergence of $q(t)$. To see that convergence of $q(t)$ implies convergence of $x(t)$, we use that $x(t+1) = Dx(t) + (G-D)s(t) = Dx(t) + (I-D)q(t)$. For all $t \geq 0$, this implies $x(t) = D^t x(0) + \sum_{l=0}^{t-1} D^{t-1-l}(I-D)q(l)$, the first part of which converges to 0 because all elements of the diagonal matrix D belong to $[0,1)$. The limit of $x(t)$ therefore equals

$$\begin{aligned}\lim_{t\to\infty}\sum_{l=0}^{t-1} D^{t-1-l}(I-D)q(l) &= \lim_{t\to\infty}\sum_{l=0}^{t-1} D^{t-1-l}(I-D)\left(q(l)-q(\infty)\right)\\ &\quad+\lim_{t\to\infty}\sum_{l=0}^{t-1} D^{t-1-l}(I-D)q(\infty).\end{aligned}$$

First, note that the second limit equals $q(\infty)$, because $\sum_{l=0}^{\infty} D^l = (I-D)^{-1}$. For the first limit, note that for any $\varepsilon > 0$, we can find an index l_ε such that we have $||q(l) - q(\infty)|| < \varepsilon$ for all $l > l_\varepsilon$. Splitting the sum into small l ($l \leq l_\varepsilon$) and large l ($l > l_\varepsilon$), we then easily see that the first term converges to 0. Therefore, $x(t)$ converges to $q(\infty)$. Since $s(t) = (I-\Delta)x(t) + \Delta q(t)$, $s(t)$ also shares the same limit. □

A.4 Block structure

Proof of Proposition 3

Let $Z := [I - \Delta(I-D)^{-1}(G-D)]^{-1}(I-\Delta)$ to simplify $s = Zx$ and $M = D + (G-D)Z$. We now proceed in three steps: we first characterize Z, then M, and finally M^t. Let G be given as in (11). Then simple but tedious block matrix algebra together with Lemma A.1 yields:

1. $$Z = \begin{pmatrix} Z_{11} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & Z_{KK} & 0 \\ Z_{\mathcal{R}1} & \cdots & \cdots & Z_{\mathcal{R}K} & Z_{\mathcal{R}\mathcal{R}} \end{pmatrix}$$

   with

   $$Z_{kk} = (I - \Delta_{kk}(I-D_{kk})^{-1}(G_{kk}-D_{kk}))^{-1}(I-\Delta_{kk}),$$

   $$Z_{\mathcal{R}k} = Z_{\mathcal{R}\mathcal{R}}(I-\Delta_{\mathcal{R}\mathcal{R}})^{-1}\Delta_{\mathcal{R}\mathcal{R}}(I-D_{\mathcal{R}\mathcal{R}})^{-1}G_{\mathcal{R}k}Z_{kk}$$

for all $k = 1, \dots, K$, and

$$Z_{\mathcal{R}\mathcal{R}} = (I - \Delta_{\mathcal{R}\mathcal{R}}(I - D_{\mathcal{R}\mathcal{R}})^{-1}(G_{\mathcal{R}\mathcal{R}} - D_{\mathcal{R}\mathcal{R}}))^{-1}(I - \Delta_{\mathcal{R}\mathcal{R}}).$$

2. For $M = D + (G - D)Z = I - (I - (G - D)\Delta(I - D)^{-1})^{-1}(I - G)$, we get

$$M = \begin{pmatrix} M_{11} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & M_{KK} & 0 \\ M_{\mathcal{R}1} & \cdots & \cdots & M_{\mathcal{R}K} & M_{\mathcal{R}\mathcal{R}} \end{pmatrix}$$

with

$$\begin{aligned} M_{kk} &= D_{kk} + (G_{kk} - D_{kk})(I - \Delta_{kk}(I - D_{kk})^{-1}(G_{kk} - D_{kk}))^{-1}(I - \Delta_{kk}) \\ &= I - \left(I - (G_{kk} - D_{kk})\Delta_{kk}(I - D_{kk})^{-1}\right)^{-1}(I - G_{kk}), \end{aligned}$$

$$\begin{aligned} M_{\mathcal{R}k} &= G_{\mathcal{R}k}Z_{kk} + (G_{\mathcal{R}\mathcal{R}} - D_{\mathcal{R}\mathcal{R}})Z_{\mathcal{R}k} \\ &= (I - (G_{\mathcal{R}\mathcal{R}} - D_{\mathcal{R}\mathcal{R}})\Delta_{\mathcal{R}\mathcal{R}}(I - D_{\mathcal{R}\mathcal{R}})^{-1})^{-1}G_{\mathcal{R}k}Z_{kk} \end{aligned}$$

for all $k = 1, \dots, K$, and

$$\begin{aligned} M_{\mathcal{R}\mathcal{R}} &= D_{\mathcal{R}\mathcal{R}} + (G_{\mathcal{R}\mathcal{R}} - D_{\mathcal{R}\mathcal{R}})(I - \Delta_{\mathcal{R}\mathcal{R}}(I - D_{\mathcal{R}\mathcal{R}})^{-1}(G_{\mathcal{R}\mathcal{R}} - D_{\mathcal{R}\mathcal{R}}))^{-1}(I - \Delta_{\mathcal{R}\mathcal{R}}) \\ &= I - \left(I - (G_{\mathcal{R}\mathcal{R}} - D_{\mathcal{R}\mathcal{R}})\Delta_{\mathcal{R}\mathcal{R}}(I - D_{\mathcal{R}\mathcal{R}})^{-1}\right)^{-1}(I - G_{\mathcal{R}\mathcal{R}}). \end{aligned}$$

3. Finally, we claim that for every $t \in \mathbb{N} \setminus \{0\}$,

$$M^t = \begin{pmatrix} M_{11}^t & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & M_{KK}^t & 0 \\ (M^t)_{\mathcal{R}1} & \cdots & \cdots & (M^t)_{\mathcal{R}K} & M_{\mathcal{R}\mathcal{R}}^t \end{pmatrix}$$

with $(M^t)_{\mathcal{R}k} = \sum_{l=0}^{t-1} M_{\mathcal{R}\mathcal{R}}^l M_{\mathcal{R}k} M_{kk}^{t-1-l}$ for all $k = 1, \dots, K$.

The assertion for the diagonal elements $M_{11}^t, \dots, M_{KK}^t$ and $M_{\mathcal{R}\mathcal{R}}^t$ is trivial. We prove the formula for $M_{\mathcal{R}k}^t$ by induction.

- For $t = 1$, the assertion is trivial.
- $t \mapsto t+1$: first, we have $(M^{t+1})_{\mathcal{R}k} = (M^t M)_{\mathcal{R}k} = (M^t)_{\mathcal{R}k} M_{kk} + M^t_{\mathcal{R}\mathcal{R}} M_{\mathcal{R}k}$ by simple matrix multiplication. Inserting $(M^t)_{\mathcal{R}k} = \sum_{l=0}^{t-1} M^l_{\mathcal{R}\mathcal{R}} M_{\mathcal{R}k} M_{kk}^{t-1-l}$, we find

$$
\begin{aligned}
(M^{t+1})_{\mathcal{R}k} &= \left(\sum_{l=0}^{t-1} M^l_{\mathcal{R}\mathcal{R}} M_{\mathcal{R}k} M_{kk}^{t-1-l} \right) M_{kk} + M^t_{\mathcal{R}\mathcal{R}} M_{\mathcal{R}k} \\
&= \sum_{l=0}^{t+1-1} M^l_{\mathcal{R}\mathcal{R}} M_{\mathcal{R}k} M_{kk}^{t+1-1-l},
\end{aligned}
$$

which concludes the proof. □

A.5 Convergence

Proof of Proposition 4

1. 'Only if': this is proven in the first part of the proof of Theorem 1.
2. 'If': Suppose each M^t_{kk} converges and $M^t_{\mathcal{R}\mathcal{R}}$ converges to 0. First, since M^t_{kk} converges, its only eigenvalue with $|\lambda| \geq 1$ is $\lambda = 1$ with algebraic and geometric multiplicity equal to 1 for every $k = 1, \ldots, K$. On the other hand, $M^t_{\mathcal{R}\mathcal{R}} \to 0$ implies that the eigenvalues of $M_{\mathcal{R}\mathcal{R}}$ are all smaller than 1 in absolute value and, thus, $M_{\mathcal{R}\mathcal{R}} - \lambda I$ is invertible for all complex numbers λ with $|\lambda| \geq 1$.

   Now, let the complex number $\tilde{\lambda}$ be either outside of the unit circle ($|\tilde{\lambda}| > 1$) or exactly on the unit circle ($|\tilde{\lambda}| = 1$), but different from 1. Taking into account the block structure of M, we easily see that any solution of $(M - \tilde{\lambda} I)x = 0$ must satisfy $x_{|\mathcal{C}_1} = 0, \ldots, x_{|\mathcal{C}_K} = 0$, and therefore also $x_{|\mathcal{C}_\mathcal{R}} = 0$, so that we can conclude that $\lambda = 1$ is the only possible eigenvalue of M with $|\lambda| \geq 1$.

   In order to show convergence of M^t, we therefore have to show that algebraic and geometric multiplicity of $\lambda = 1$ coincide. With regard to algebraic multiplicity, the block structure of M implies $\det(M - \lambda I) = \prod_{k=1}^{K} \det(M_{kk} - \lambda I) \det(M_{\mathcal{R}\mathcal{R}} - \lambda I)$, such that the algebraic multiplicity of $\lambda = 1$ is the sum of the algebraic multiplicities of $M_{11}, \ldots, M_{KK}$ and $M_{\mathcal{R}\mathcal{R}}$, which are given by 1 and 0, respectively, since M_{kk} is by definition irreducible for all $k = 1, ..., K$. Consequently, the algebraic multiplicity equals K. With regard to geometric multiplicity, the block structure of M implies that for any real numbers $c_1, \ldots, c_K$, the vector x of the form $x_{|\mathcal{C}_k} = c_k \mathbb{1}_{|\mathcal{C}_k}$ ($k = 1, \ldots, K$) and $x_{|\mathcal{C}_\mathcal{R}} = (I - M_{\mathcal{R}\mathcal{R}})^{-1} \sum_{k=1}^{n} c_k M_{\mathcal{R}k} \mathbb{1}_{|\mathcal{C}_k}$ is an eigenvector to M for $\lambda = 1$, implying that the geometric multiplicity is at least K, thereby concluding the proof. $\square$

Proof of Proposition 5

Denote $Y := (I-D)^{-1}(G-D)$ which is row stochastic. Thus, as $|\delta_i| < 1$ for all $i \in \mathcal{N}$, we have that $I - \Delta Y$ is invertible and $(I-\Delta Y)^{-1} = \sum_{k=0}^{\infty}(\Delta Y)^k$. Moreover, if $\delta_i \geq 0$ for all $i \in \mathcal{N}$, the sum $\sum_{k=0}^{\infty}(\Delta Y)^k$ is a sum of non-negative matrices, implying that $(I-\Delta Y)^{-1}$ has only non-negative entries. Hence $M = D+(G-D)[I-\Delta Y]^{-1}(I-\Delta)$ is non-negative since it is the product of non-negative matrices (since $0 < g_{ii} < 1$) added to D, which is a diagonal matrix with strictly positive entries ($0 < g_{ii}$). Finally, since $M\mathbb{1} = \mathbb{1}$ by Lemma A.1, we get that M is row stochastic. Since the diagonal of D is strictly positive, we get that the diagonal of M is strictly positive, $m_{ii} > 0$, implying aperiodicity of M. Thus M^t converges. $\square$

A.6 Long–run

To prove Theorem 1, the following Lemma is helpful.

Lemma A.3 (Convergence to Eigenvector). *Let A be an $n \times n$-matrix with $A\mathbb{1} = \mathbb{1}$ and* $\text{rk}(I-A) = n-1$. *If A^t converges to A^∞ for $t \to \infty$, then $A^\infty = \mathbb{1}w'$, with w' the unique normalized left eigenvector of A associated with the eigenvalue* 1.

Proof of Lemma A.3

Obviously, $AA^\infty = A^\infty = A^\infty A$. This implies that

- the columns of A^∞ must be multiples of $\mathbb{1}$,
- the rows of A^∞ must be multiples of w',

from which we find $A^\infty = r\,\mathbb{1}w'$ for some real number r which is found to be equal to 1 as $\mathbb{1} = A^\infty\mathbb{1} = r\,\mathbb{1}w'\mathbb{1} = r\,\mathbb{1}$. $\square$

Proof of Theorem 1

We first derive the formula for M^∞_{kk}. Then we will turn to $M^\infty_{\mathcal{R}\mathcal{R}}$ and $M^\infty_{\mathcal{R}k}$.

Assume for the moment that $\text{rk}(I-G) = n-1$. Then, as $v'(M-I) = 0$, we have due to Lemma A.1

$$0 = v'(I-M) = v'\left(I-(G-D)\Delta(I-D)^{-1}\right)^{-1}(I-G),$$

implying

$$v'\left(I-(G-D)\Delta(I-D)^{-1}\right)^{-1}=r\,w'$$

for some real number r. Using $w'G=w'$, we then find

$$v'=r\,w'\left(I-(G-D)\Delta(I-D)^{-1}\right)=r\,w'\left(I-(I-D)\Delta(I-D)^{-1}\right)=r\,w'(I-\Delta).$$

The normalization of v then entails $r=\frac{1}{w'(I-\Delta)\mathbb{1}}$, which shows that $v=\dfrac{(I-\Delta)w}{\mathbb{1}'(I-\Delta)w}$. Now, relaxing the assumption $\mathrm{rk}(I-G)=n-1$, the formula for M^∞_{kk} follows.

Furthermore, $MM^\infty x=M^\infty x$ and therefore due to Proposition 2, $GM^\infty x=M^\infty x$ for all n-dimensional vectors x, delivering $GM^\infty=M^\infty$. This implies

- $M^\infty_{\mathcal{R}\mathcal{R}}=G_{\mathcal{R}\mathcal{R}}M^\infty_{\mathcal{R}\mathcal{R}}$ and therefore $(I-G_{\mathcal{R}\mathcal{R}})M^\infty_{\mathcal{R}\mathcal{R}}=0$, entailing $M^\infty_{\mathcal{R}\mathcal{R}}=0$ because $I-G_{\mathcal{R}\mathcal{R}}$ is invertible,
- $M^\infty_{\mathcal{R}k}=G_{\mathcal{R}k}M^\infty_{kk}+G_{\mathcal{R}\mathcal{R}}M^\infty_{\mathcal{R}k}$, and therefore $M^\infty_{\mathcal{R}k}=(I-G_{\mathcal{R}\mathcal{R}})^{-1}G_{\mathcal{R}k}M^\infty_{kk}$. □

A.7 Wisdom

Proof of Lemma 1

First, $\hat{\mu}_k$ is easily seen to be unbiased for μ because $\sum_{i\in\mathcal{C}_k} v_i=1$. Therefore, its MSE equals its variance which is given by $\sum_{i\in\mathcal{C}_k} v_i^2\sigma_i^2$ as the $x_i(0)$ are uncorrelated. □

Proof of Proposition 6

$$\frac{\partial\,\mathrm{MSE}_k}{\partial\delta_i}=\frac{\partial\sum_{j\in\mathcal{C}_k}v_j^2\sigma_j^2}{\partial\delta_i}=\sum_{j\in\mathcal{C}_k}2\sigma_j^2v_j\frac{\partial v_j}{\partial\delta_i}\overset{(17)}{=}\frac{2w_i}{\sum_{j\in\mathcal{C}_k}w_j(1-\delta_j)}\sum_{j\in\mathcal{C}_k}\sigma_j^2v_j\left(v_j-1_{j=i}\right).$$

The assertion follows easily noting that $\mathrm{MSE}_k=\sum_{j\in\mathcal{C}_k}v_jv_j\sigma_j^2$. □

Proof of Proposition 7 First, notice that $E((x_i(\infty)-\mu)^2)=\sum_{k=1}^{K}\gamma_{i,k}^2\,\mathrm{MSE}_k$, with $\sum_{k=1}^{K}\gamma_{i,k}=1$ for all $i\in\mathcal{R}$. By the Cauchy-Schwarz inequality, we have

$$1=\sum_{k=1}^{K}\gamma_{i,k}=\sum_{k=1}^{K}\left(\gamma_{i,k}\sqrt{\mathrm{MSE}_k}\right)\frac{1}{\sqrt{\mathrm{MSE}_k}}\le\sqrt{\sum_{k=1}^{K}\gamma_{i,k}^2\,\mathrm{MSE}_k}\sqrt{\sum_{k=1}^{K}\frac{1}{\mathrm{MSE}_k}},$$

with equality if and only if there exists some (necessarily positive) constant a such that $\gamma_{i,k}\sqrt{\mathrm{MSE}_k} = a\,\frac{1}{\sqrt{\mathrm{MSE}_k}}$ for all k. We therefore have $\sum_{k=1}^{K} \gamma_{i,k}^2\,\mathrm{MSE}_k \geq \frac{1}{\sum_{k=1}^{K}\frac{1}{\mathrm{MSE}_k}}$, with equality if and only if $\gamma_{i,k} = \frac{1}{\mathrm{MSE}_k \sum_{l=1}^{K}\frac{1}{\mathrm{MSE}_l}}$ for all k. □

Acknowledgments

We thank Leonie Baumann, Roland Bénabou, Matthew Jackson, Michael Maes, Christian Martin, Gerd Muehlheusser, Agnieszka Rusinowska, Károly Takács, and Fernando Vega-Redondo for helpful comments and discussions. Moreover, we are thankful to the participants of the conferences "SING7" in Paris, "Game Theory and Society" in Zurich, "SING8" in Budapest, "Sunbelt XXXIII" in Hamburg, and the participants of the research seminars in Belfast, Bielefeld, Florence, Hamburg, Milano (at Bocconi), Saarbruecken, and Zurich.

References

Acemoglu, D., Dahleh, M. A., Lobel, I., and Ozdaglar, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236.

Acemoglu, D. and Ozdaglar, A. E. (2011). Opinion dynamics and learning in social networks. *International Review of Economics*, 1(1):3–49.

Acemoglu, D., Ozdaglar, A. E., and ParandehGheibi, A. (2010). Spread of (mis)information in social networks. *Games and Economic Behavior*, 70(2):194–227.

Ambrus, A., Azevedo, E. M., and Kamada, Y. (2013). Hierarchical cheap talk. *Theoretical Economics*, 8(1):233–261.

Anderlini, L., Gerardi, D., and Lagunoff, R. (2012). Communication and learning. *Review of Economic Studies*, 79(2):419–450.

Ariely, D. and Levav, J. (2000). Sequential choice in group settings: Taking the road less traveled and less enjoyed. *Journal of Consumer Research*, 27(3):279–90.

Asch, S. (1955). Opinions and social pressure. *Scientific American*, 193(5):31–35.

Bala, V. and Goyal, S. (1998). Learning from neighbours. *Review of Economic Studies*, 65(3):595–621.

Bala, V. and Goyal, S. (2001). Conformism and diversity under social learning. *Economic Theory*, 17(1):101–120.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Berger, R. L. (1981). A Necessary and Sufficient Condition for Reaching Consensus Using DeGroot's Method. *Journal of the American Statistical Association*, 76:415–418.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, 102(5):841–77.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy*, 100(5):992–1026.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, 2(1):113–120.

Buechel, B., Hellmann, T., and Pichler, M. M. (2011). The dynamics of continuous cultural traits in social networks. Working Paper 457, Bielefeld University, Institute of Mathematical Economics.

Chan, K. K. and Misra, S. (1990). Characteristics of the opinion leader: A new dimension. *Journal of Advertising*, 19(3):pp. 53–60.

Corazzini, L., Pavesi, F., Petrovich, B., and Stanca, L. (2012). Influential listeners: An experiment on persuasion bias in social networks. *European Economic Review*, 56(6):1276–1288.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118(3):909–968.

Deutsch, M. and Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgement. *Journal of Abnormal Psychology*, 51(3):629–636.

Ellison, G. and Fudenberg, D. (1993). Rules of thumb for social learning. *Journal of Political Economy*, 101(4):612–43.

Ellison, G. and Fudenberg, D. (1995). Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, 110(1):93–125.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Flache, A. and Torenvlied, R. (2004). When Will They Ever Make up Their Minds? The Social Structure of Unstable Decision Making. *Journal of Mathematical Sociology*, 28:171–196.

Foerster, M., Mauleon, A., and Vannetelbosch, V. (2013). Trust and manipulation in social networks. mimeo, University of Louvain, Center for Operations Research and Econometrics.

French, J. (1956). A formal theory of social power. *Psychological Review*, 63:181–194.

Friedkin, N. E. (1991). Theoretical foundations for centrality measures. *The American Journal of Sociology*.

Friedkin, N. E. and Johnsen, E. C. (1990). Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193–206.

Gale, D. and Kariv, S. (2003). Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346.

Galeotti, A. and Goyal, S. (2009). Influencing the influencers: a theory of strategic diffusion. *The RAND Journal of Economics*, 40(3):509–532.

Golub, B. and Jackson, M. O. (2010). Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–49.

Grabisch, M. and Rusinowska, A. (2010). A model of influence in a social network. *Theory and Decision*, 69(1):69–96.

Grabisch, M. and Rusinowska, A. (2011). Influence functions, followers and command games. *Games and Economic Behavior*, 72(1):123–138.

Hagenbach, J. and Koessler, F. (2010). Strategic communication networks. *The Review of Economic Studies*, 77(3):1072–1099.

Harary, F. (1959). Status and contrastatus. *Sociometry*, 22:23–43.

Hayakawa, H. and Venieris, Y. (1977). Consumer interdependence via reference groups. *Journal of Political Economy*, 85(3):pp. 599–616.

Hegselmann and Krause (2002). Opinion dynamics and bounded confidence, models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3):1–33.

Hogg, M. and Abrams, D. (1988). *Social Identifications: A Social Psychology of Intergroup Relations and Group Processes*. University paperbacks. Routledge.

Hornsey, M. J., Majkut, L., Terry, D. J., and McKimmie, B. M. (2003). On being loud and proud: Non-conformity and counter-conformity to group norms. *British Journal of Social Psychology*, 42(3):319–335.

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press.

Jones, S. R. G. (1984). *The economics of conformism*. B. Blackwell, Oxford; New York, NY.

Lorenz, J. (2005). A stabilization theorem for dynamics of continuous opinions. *Physica A: Statistical Mechanics and its Applications*, 355(1):217–223.

Mueller-Frank, M. (2011). The Robustness of Consensus in Network-Based Non-Bayesian Learning Models. Technical report.

Pan, Z. (2010). Trust, influence, and convergence of behavior in social networks. *Mathematical Social Sciences*, 60(1):69–78.

Rauhut, H. and Lorenz, J. (2010). The wisdom of crowds in one mind: How individuals can simulate the knowledge of diverse societies to reach better decisions. *Journal of Mathematical Psychology*.

Smith, L. and Sorensen, P. (2000). Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398.

Zafar, B. (2011). An experimental investigation of why individuals conform. *European Economic Review*, 55(6):774–798.