



Opinion dynamics and wisdom under conformity



Berno Buechel ^{a,*}, Tim Hellmann ^b, Stefan Klößner ^c

^a Department of Economics, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany

^b Institute of Mathematical Economics, Bielefeld University, Germany

^c Statistics and Econometrics, Saarland University, Germany

ARTICLE INFO

Article history:

Received 13 May 2014

Received in revised form

6 December 2014

Accepted 18 December 2014

Available online 29 December 2014

JEL classification:

C72

D83

D85

Z13

Keywords:

Opinion leadership

Wisdom of crowds

Consensus

Social networks

Conformity

Eigenvector centrality

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. Concerning efficiency of information aggregation or "wisdom" of the society, it turns out that misrepresentation of opinions need not undermine wisdom, but may even enhance it. 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.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Opinions crucially shape individual behavior and affect economic decisions and outcomes.¹ For instance, opinions on political issues set the political course, opinions about a product's quality and the integrity of its producer influence demand, and opinions about an economy's growth determine investment decisions. The formation and evolution of opinions are often carried by day-to-day interactions of individuals, i.e. the opinions are determined by exchange in a social network.

We model the formation of opinions through communication in a given social network such that individuals are influenced by the opinions stated by others: individuals update their opinion in a naïve way by taking a weighted average of others' stated opinions (as in the literature on naïve learning, see e.g. DeGroot, 1974; DeMarzo et al., 2003; Golub and Jackson, 2010; Acemoglu et al., 2010). However, influence often goes beyond this simple updating of opinions. When asked for a personal opinion, people usually do not straightforwardly state what they truly think, *rather they are tempted to*

* Corresponding author. Tel.: +49 40 42838 5573.

E-mail addresses: berno.buechel@uni-hamburg.de (B. Buechel), tim.hellmann@uni-bielefeld.de (T. Hellmann), s.kloessner@mx.uni-saarland.de (S. Klößner).

URL: <http://www.berno.info> (B. Buechel).

¹ Under the term *opinions* we subsume also beliefs, judgments, and estimations – depending on the application.

misrepresent their opinion to conform to their friends since disagreement entails uncomfortable feelings (see Zafar, 2011, for empirical evidence). In this paper, we consequently allow that not only the own opinion is influenced by what others say, but also the statement itself. In other words, some individuals tweak their stated opinions to conform to what their social contacts say.

In such a framework, we study the long-run dynamics of opinions and particularly focus on *how conformity affects opinion leadership* and *how it affects the quality of information aggregation* (“wisdom”). Concerning opinion leadership, we obtain a closed-form solution for the influence of an agent on the long-run consensus opinion (i.e. her power). It is determined not only by each individual's position in the network, given by eigenvector centrality (Bonacich, 1972; Friedkin, 1991; DeMarzo et al., 2003), but also by the distribution of conformity in the society. Comparative statics reveal that an agent's power is increasing in own network centrality, increasing in other agents' level of conformity, but decreasing in own level of conformity. The last result, hence, explains the empirical finding that opinion leaders are often characterized by low conformity.²

Concerning quality of information aggregation, we interpret initial opinions as unbiased noisy signals about some true state of nature, and study the precision of the consensus beliefs. First, in a finite population setup, we are able to exactly determine the quality of information aggregation depending on network position, initial signal quality and the distribution of conformity. It turns out that heterogeneous levels of conformity foster precision of the consensus beliefs if conformity is distributed such that it balances the power of agents with their signal quality. Agents who are connected to multiple groups and, thereby, have access to multiple sources of information, end up being well informed although their initial information is lost completely. In large populations, we similarly obtain that agents learn the true state with probability one (i.e. wise crowds emerge, cf. Golub and Jackson, 2010) if too central agents are very conform.

We allow for conformity in an opinion formation framework since there is substantial empirical evidence that individuals conform to the actions of others when these actions are observable (as stated opinions are). For instance in the famous study by Asch (1955), subjects wrongly judged the length of a line after other participants of the experiment (conceived as neutral by the subjects, but being collaborators) 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). In the study by Asch (1955), subjects were asked for the reasons of their wrong judgment. Some said they were convinced of the wrong answer by the collaborators; others said that they knew that their answer was wrong, but felt uncomfortable by not conforming to what the collaborators said (see Asch, 1955, p. 21). Deutsch and Gerard (1955), hence, distinguish two forms of social influence that can be observed in this study. While *informational social influence* describes the updating of (true) opinions according to what others have said, *normative social influence* describes the behavior of stating an opinion that fits the group norm.³

Normative social influence is also documented with respect to other publicly observable behavior. In an experiment on charitable giving, Zafar (2011) shows that individuals adjust more to the contributions of their neighbors (and hence conform more by reducing respectively increasing their contribution), the more their donations are observable, supporting the findings by Asch (1955) and Deutsch and Gerard (1955). Moreover, subjects in Zafar's experiment mainly conform to the actions of participants who are their friends outside the lab. Hence, normative social influence is determined by the social network itself. Zafar (2011) concludes that individuals experience “a utility gain by simply making the same choice as [their] 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).

While normative social influence affects the choice of stated opinions, informational social influence embodies the updating of the true opinions. We assume that individuals update their true opinions naïvely rather than sophisticatedly since empirical evidence strongly suggests that individuals in these settings behave boundedly rational (Corazzini et al., 2012; Grimm and Mengel, 2013; Battiston and Stanca, 2014). If individuals were fully rational, they would perfectly account for repetition of information (for some references on Bayesian learning in opinion formation, see Gale and Kariv, 2003; Acemoglu et al., 2011; Sethi and Yildiz, 2012; Mueller-Frank, 2013; Sethi and Yildiz, 2013). In fact, evidence from laboratory experiments shows that even in small social networks (of only four people) where the network is made common knowledge, people fail to properly account for repetitions of information (Corazzini et al., 2012; Battiston and Stanca, 2014). In more complex networks, other studies (Chandrasekhar et al., 2012; Grimm and Mengel, 2013) also confirm that learning in the lab is very well approximated by the naïve learning approach.

Hence, we model informational social influence by assuming that individuals learn naïvely from what others say (see also DeMarzo et al., 2003; Golub and Jackson, 2010; Acemoglu et al., 2010). In view of the substantial empirical evidence, we enrich the naïve learning model by studying the effects of individuals who have a desire to adjust their behavior (i.e. their stated opinion) to the behavior of their friends (i.e. their friends' stated opinions). In the words of psychology, this corresponds to modeling normative social influence. Remarkably, this type of influence has not been studied in a theoretical

² A personality trait that has been found to discriminate opinion leaders from followers is called ‘public individuation’ (Chan and Misra, 1990). It measures by a list of questions the extent to which “people choose to act differently than others” (Maslach et al., 1985).

³ 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.”

model of opinion dynamics despite the large empirical evidence.⁴ The main conceptual contribution of this work is hence to fill this gap by studying a model incorporating both informational and normative social influence. We focus on two motives for the misrepresentation of opinions: conformity and counter-conformity, while we also allow for honest agents.⁵ The desire to relate own stated opinions to the stated opinions of friends is given by an additional utility component parameterized by a preference parameter which we call the agent's degree of conformity. If positive, an agent is of conforming type and states an opinion which is a convex combination of own true opinion and other agents' stated opinions. If negative, an agent is counter-conforming and will state a more extreme opinion. This overstatement of opinions is similar to, but less extreme than, the behavior in [Kalai and Kalai \(2001\)](#). Finally, if an agent's degree of conformity is zero, the agent is honest, i.e. behaving like agents in the standard DeGroot model.

We consider heterogeneous societies whose agents are connected by a social network and endowed with individual levels of (counter-)conformity. In contrast to the DeGroot model, the law of motion in our model may have negative entries, due to counter-conformity of some agents.⁶ We study opinion dynamics in this model with a focus on opinion leadership and wisdom of the society.

Related models: There is a growing body of literature that studies naïve learning in social networks. [DeMarzo et al. \(2003\)](#) introduce this approach into the economics literature arguing that people are often unable to properly account for repetition of information. The underlying assumption of a “persuasion bias” is helpful to understand different empirical phenomena such as the importance of airtime in political discussions and it has also found empirical support in the laboratory ([Corazzini et al., 2012](#); [Grimm and Mengel, 2013](#); [Battiston and Stanca, 2014](#)) and in the field ([Chandrasekhar et al., 2012](#); [Breza et al., 2014](#)). Among naïve agents the social network becomes vital in the sense that not only accuracy of information but also network centrality determines an agent's influence on her group ([DeMarzo et al., 2003](#)). This form of social influence makes naïve agents prone to be misled by powerful actors such as community leaders or lobbyists ([Acemoglu et al., 2010](#)). On the other hand, dispersed pieces of information can also be efficiently aggregated among naïve agents if the influence of each individual is vanishingly small ([Golub and Jackson, 2010](#)). The crucial question is hence under which conditions exchange of opinions among naïve agents leads to efficient information aggregation which is also called wisdom. While in [Golub and Jackson \(2010\)](#) the result holds asymptotically for large societies, even finitely many naïve agents can be wise, if all individuals take their personal signals into account in a Bayesian way ([Jadbabaie et al., 2012](#)), or if one fully Bayesian agent exists ([Mueller-Frank, 2014](#)). Our model takes the examination of the questions of power and wisdom to a further level since it incorporates not only the social network structure but also individual degrees of conformity.

The modeling approach of the above literature roots in the pioneer work of [French \(1956\)](#), [Harary \(1959\)](#), [DeGroot \(1974\)](#) and [Friedkin and Johnsen \(1990\)](#).⁷ Several variations of the naïve learning approach can be found with respect to arrival of new information in every period ([Jadbabaie et al., 2013](#)), exogenously varying influence weights ([Lorenz, 2005](#)), adaption of influence weights ([Hegselmann and Krause, 2002](#); [Flache and Torenvlied, 2004](#); [Pan, 2010](#)), or even manipulation thereof ([Foerster et al., 2013](#)). The focus of many of these models is to provide conditions for convergence, or to determine opinion leadership. We contribute to this literature by allowing agents to misrepresent their opinion and study the effect on opinion leadership and wisdom. In a context of cultural transmission of traits, [Buechel et al. \(2014\)](#) introduce strategic interaction for the DeGroot model in an OLG framework. While this resembles counter-conforming misrepresentation of opinions, their model differs with respect to the optimization problem of individuals, the updating rule, and the resulting dynamics.

Besides these highly related works, there are several contributions to somewhat similar research questions, but with respect to different settings. While their discussion is beyond the scope of this paper, we refer the reader to the following few prominent examples: other models of social learning ([Bikhchandani et al., 1992](#); [Ellison and Fudenberg, 1993, 1995](#); [Bala and Goyal, 1998, 2001](#); [Sethi and Yildiz, 2012](#)), cooperative models of social influence ([Grabisch and Rusinowska, 2010, 2011](#)), a model of strategic influence ([Galeotti and Goyal, 2009](#)), a model on rumors ([Merlone and Radi, 2014](#)) and a fairly general approach to dynamic games with informational externalities ([Rosenberg et al., 2009](#)). Most of these models investigate social influence on a discrete choice of actions, such as the choice of one out of two technologies, as opposed to continuous opinions. The framework of [Rosenberg et al. \(2009\)](#) goes beyond discrete action choice. In fact, [Rosenberg et al. \(2009\)](#) subsumes strategic misrepresentation of opinions under Bayesian learning as a special case and provides conditions such that a consensus is reached.

The remainder of this paper is organized as follows. In [Section 2](#) we introduce the model. After presenting the main results (in [Section 3](#)), [Section 4](#) addresses the wisdom of the society and in [Section 5](#) we conclude, while proofs are relegated to the appendix.

⁴ 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”. 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. 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.

⁵ This is consistent with the psychological theory where identification, non-identification and disidentification lead to these three types of normative social influence ([Hogg and Abrams, 1988](#)).

⁶ Negative entries are not only remarkable because of the different dynamics they induce, but also because of their interpretation as a negative relation between two agents: Although only positive weights are put on each other's opinions, an agent may negatively incorporate a peer's opinion.

⁷ We adopt their assumptions on naïve learning to model informational social influence.

2. Model

2.1. Informational social influence

There is a set of agents $\mathcal{N} = \{1, 2, \dots, n\}$ 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 $i_0, \dots, i_k \in \mathcal{N}$ such that $i_0 = i$ and $i_k = j$ and $g_{i_l i_{l+1}} > 0$ for all $l = 0, \dots, k-1$, which is equivalent to $(G^k)_{ij} > 0$.⁸ 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 $t = 0, 1, 2, \dots$ 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.”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 $x_i(t)$.¹⁰

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.¹¹ 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:

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

where $\delta_i \in (-1, +1)$ displays the relative importance of the preference for (counter-)conformity in relation to the preference for honesty. Negative δ_i captures a preference for counter-conformity.¹² The degree of conformity can 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.

⁸ 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).

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

¹⁰ The incentive to state an opinion different from true opinion will be based on preferences for conformity or counter-conformity (cf. Section 2.2). Moreover, agents adapt their stated opinions faster than true opinions such that $s(t)$ is given by Proposition 1.

¹¹ 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.

¹² 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.

Now, we want to determine each agent's stated opinion without assuming that the network structure and the individual types are common knowledge and without assuming that agents are sophisticated in anticipating the consequences of their behavior. For this purpose we consider 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.¹³ 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.

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

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.

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

for any true opinion $x_i(t)$ and any reference opinion $q_i^{\tau-1}(t)$. 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. 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 $\delta_i = 0$, straightforwardly 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 $\tau \rightarrow \infty$ to*

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

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 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 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.¹⁴ 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:

$$x(t+1) = Mx(t), \quad (6)$$

where $M := D + (G - D)[I - \Delta(I - D)^{-1}(G - D)]^{-1}(I - \Delta)$. 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 are fully described by the power series M^t , since $x(t+1) = Mx(t) = M^2x(t-1) =$

¹³ 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.

¹⁴ 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.

... = $M^{t+1}x(0)$.¹⁵ 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 standard case of DeGroot (1974).

3. Opinion dynamics

In this section we study opinion dynamics in the long run. We first elaborate on conditions of convergence of opinions (Section 3.1). Then we determine where opinions converge to for strongly connected societies (Section 3.2) and for unconnected societies (Section 3.3).

3.1. Convergence

By convergence we mean that opinions settle down in a steady state, but not necessarily that a consensus in the society is reached. In the standard DeGroot model, convergence of opinions is obtained under very mild conditions, which basically exclude cycling dynamics (Golub and Jackson, 2010). In our more general model, opinions may not only converge or cycle, but also diverge, as we will show below.

Mathematically, convergence of opinions is driven by convergence of M^t . Indeed, since $x(t) = M^t x(0)$, true opinions $x(t)$ converge for arbitrary starting opinions $x(0)$ if and only if M^t converges. Counter-conforming agents can lead to negative entries of matrix M which may but need not make M^t divergent. Conversely, honest and conforming agents do not induce negative entries of M such that convergence can be guaranteed by standard results.

Proposition 2 (Convergence). *If we have $g_{ii} > 0$ and $\delta_i \geq 0$ for all $i \in \mathcal{N}$, then M^t , true opinions $x(t)$, stated opinions $s(t)$, and perceived opinions $q(t)$ converge for $t \rightarrow \infty$.*

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 converge. The assumption of positive self-confidence thereby only serves to assure aperiodicity of matrix M which could also be generated by weaker assumptions. Although all cases of conformity are covered by Proposition 2, it is important to emphasize that conformity is not necessary for convergence.

For instance, consider the case of $n=2$ agents with identical level of (counter-)conformity $\delta_1 = \delta_2 = \delta$ and weights $g_{12}, g_{21} \in (0, 1)$ of listening to each other. In this case, we find that the eigenvalues of M are given by unity and

$$\lambda := 1 - \frac{1}{1+\delta}(g_{12} + g_{21}) < 1. \tag{7}$$

Opinion dynamics therefore converge if and only if $\lambda > -1$, i.e. if and only if conformity δ exceeds the obviously negative threshold $(g_{12} + g_{21})/2 - 1$, showing that convergence is possible under counter-conformity. Fig. 1 exemplifies opinion dynamics for $g_{12} = 0.4, g_{21} = 0.2, x(0) = (0, 100)^t$, varying δ from conformity ($\delta = 0.5$) via honesty ($\delta = 0$) to counter-conformity with convergence ($\delta = -0.6$) and eventually counter-conformity strong enough to induce divergence ($\delta = -0.75$). Notice that, in this example, speed of convergence is faster under honesty ($\delta = 0$) than both under conformity ($\delta = 0.5$) and counter-conformity ($\delta = -0.6$). In general, therefore, speed of convergence is not monotonous in conformity.

While for the case of only two agents it is easily possible to characterize the speed of convergence (this is also true for the case $\delta_1 \neq \delta_2$) and to compare it with the DeGroot model by setting $\delta = 0$, the results do not simply generalize to larger societies, even when assuming homogeneous conformity and self-confidence ($\delta_i = \delta, g_{ii} = d$ for all agents). For this most simple special case, eigenvalues λ_M of M and λ_G of G are related to each other via

$$\lambda_M = d + (1-d)(1-\delta) \frac{\lambda_G - d}{1-d-\delta(\lambda_G - d)}.$$

This relation implies that the dynamics of M^t are governed by

$$\kappa_M := \max_{\lambda_G \text{ eigenvalue of } G} \left| d + (1-d)(1-\delta) \frac{\lambda_G - d}{1-d-\delta(\lambda_G - d)} \right|. \tag{8}$$

If $\kappa_M < 1$, then M^t will converge at rate κ_M^t , if $\kappa_M = 1$, the system will cycle, and if $\kappa_M > 1$, it will diverge. The crux with (8) is that κ_M in general cannot be written as a function of $\kappa_G := \max\{|\lambda_G| : \lambda_G \text{ eigenvalue of } G\}$ because, typically, different values of δ will lead to different eigenvalues λ_G of G being the maximizers in (8).

As we have seen, although tight convergence conditions on the matrix M are known, it is, even in special cases, difficult to trace these conditions back to the model fundamentals, namely the network G and the distribution of conformity Δ . For every numerical example, however, it is an easy computational exercise to determine M and M^t and thereby establish the dynamic properties including whether opinions converge or not. Therefore, we now assume for the remainder that the

¹⁵ The simple linear structure is of course implied by our assumption of quadratic utility.

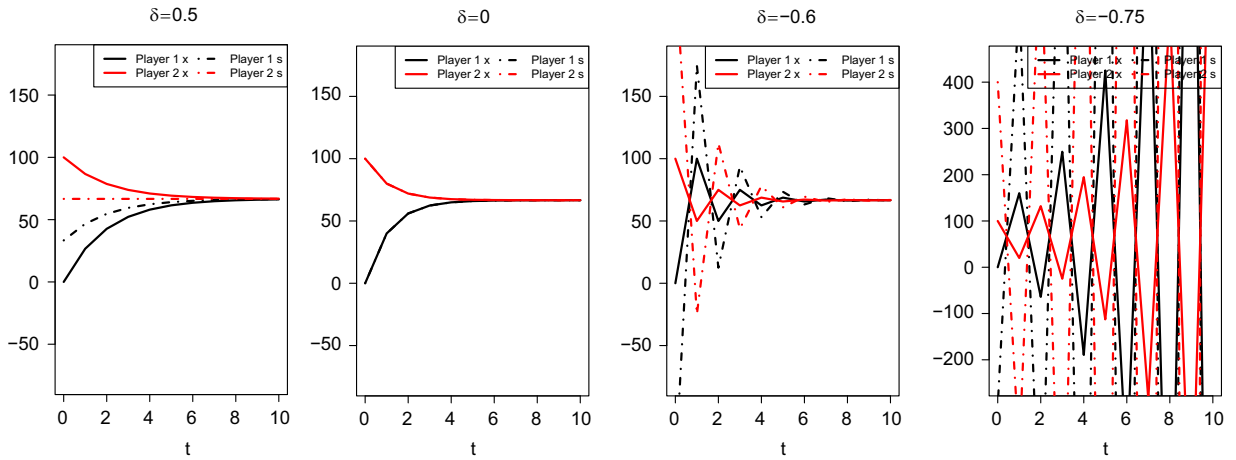


Fig. 1. Four cases of two-agent dynamics for $\delta_1 = \delta_2 = \delta$. Solid lines represent true opinions while dashed lines display stated opinions.

power series M^t converges. Again, we emphasize that this does not preclude the presence of counter-conforming agents or negative entries in M .

3.2. Long-run opinions and opinion leadership

Given convergence, we now address where opinions converge to (in the long run) when starting with some opinion profile $x(0)$. We are particularly interested in the influence of each agent’s initial opinion on the long-run opinion given her position in the network G and her degree of conformity δ_i . To simplify the discussion, let us focus on the case that G is strongly connected, which is equivalent to the assumption that $\text{rk}(I - G) = n - 1$, where rk denotes the rank of a matrix (cf. Golub and Jackson, 2010). The interpretation of this assumption is that each agent is at least indirectly influenced by every other agent. While we have mentioned above that true, stated, and perceived opinions converge under the same conditions, the following lemma shows that these opinions all converge to one point.

Lemma 1. *Let the society be strongly connected and denote by v the left unit eigenvector of M with normalization $\sum_{j \in \mathcal{N}} v_j = 1$. If opinions converge, i.e. $\lim_{t \rightarrow \infty} M^t$ for $t \rightarrow \infty$ exists, then for every $i \in \mathcal{N}$,*

$$x_i(\infty) = s_i(\infty) = q_i(\infty) = \sum_{j \in \mathcal{N}} v_j x_j(0). \tag{9}$$

Since Lemma 1 establishes that true, stated, and perceived opinions always converge to the same limit, we can restrict our analysis of the long run to the dynamics of true opinions. As it is common in the literature on naïve learning, strongly connected societies converge to a consensus opinion which is a mixture of the agents’ initial true opinions. From Eq. (9), we see that v_j captures the weight of agent j ’s initial opinion on the consensus opinion, which is a very intuitive formalization of opinion leadership: v measures the power of each agent. We now investigate how power is determined by the network G and the degrees of conformity δ_i .

Suppose first that $\delta_i = 0$ for all $i \in \mathcal{N}$. Then $M = G$ and Lemma 1 yields that opinion leadership is fully determined by the left unit eigenvector of G , which we will denote by w . This vector (w) is a well-studied object in network science: it is known as eigenvector centrality and captures how well-connected each agent is in the social network (Bonacich, 1972; Friedkin, 1991).¹⁶ Relaxing the assumption that every agent is honest, the following result shows how opinion leadership is not only determined by eigenvector centrality, but also by the degree of conformity.

Theorem 1. *Let the society be strongly connected and denote by w and v the normalized left unit eigenvectors of G and M , respectively. If opinions converge, i.e. $\lim_{t \rightarrow \infty} M^t$ for $t \rightarrow \infty$ exists, then power of every agent $i \in \mathcal{N}$ is given by*

$$v_i = \frac{(1 - \delta_i) w_i}{\sum_{j \in \mathcal{N}} (1 - \delta_j) w_j}. \tag{10}$$

¹⁶ More precisely, this index of centrality is recurrently defined via the rows of G' (i.e. via the columns of G) such that an agent’s centrality is the weighted sum of centralities of the agents who listen to her.

Theorem 1 establishes that 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. As it becomes apparent from (10), there is a complementary relationship between network centrality and $1 - \delta_i$: power is minimal (v_i close to zero) if either i 's network centrality w_i is very low or if i is fully conform (δ_i close to 1). In the same sense, power is maximal if all other agents' power is minimal, relative to own power. Taking the network as given, we can directly derive from **Theorem 1** the comparative statics with respect to the level of conformity of any agent:

$$\frac{\partial v_i}{\partial \delta_k} = \frac{w_k}{\sum_{j=1}^n w_j(1 - \delta_j)}(v_i - 1_{i=k}). \tag{11}$$

From (11) 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. This finding is consistent with empirical evidence that opinion leaders are characterized by low levels of conformity (Chan and Misra, 1990). While a negative relation between conformity and opinion leadership is intuitive, **Theorem 1** goes beyond this observation by providing a direction of causality and by quantifying the effect. Importantly, it shows how conformity interacts with the network position and with the conformity levels of the other members of the society.

In particular, consider a society where all agents are characterized by the same trait, i.e. $\delta_j = \bar{\delta}$ for all $j \in \mathcal{N}$. Then (10) yields $v = w$: opinion leadership is not affected by conformity when all agents are characterized by the same level of conformity. Relaxing the assumption of homogeneous conformity, we have $v_i \geq w_i$ if and only if $\delta_i \leq \sum_{j \neq i} (w_j / \sum_{k \neq i} w_k) \delta_j$, i.e. an agent "gains" power in a comparative statics sense compared to the case of homogeneous conformity if her level of conformity is below some average of the others' conformity levels.

The following corollary of **Theorem 1** shows which agent's power changes most in response to a marginal increase in her own conformity.

Corollary 1. Under the assumptions of **Theorem 1**, we have

$$\left| \frac{\partial v_i}{\partial \delta_i} \right| < \left| \frac{\partial v_j}{\partial \delta_j} \right| \Leftrightarrow w_j^2(1 - \delta_j) - w_i^2(1 - \delta_i) < (w_j - w_i) \sum_{k=1}^n w_k(1 - \delta_k). \tag{12}$$

Thus, if two agents have the same network centrality $w_i = w_j$, then by (12), $|\partial v_i / \partial \delta_i| < |\partial v_j / \partial \delta_j|$ 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 $|\partial v_i / \partial \delta_i| < |\partial v_j / \partial \delta_j|$ 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.

3.3. Beyond strongly connected networks

So far, we have assumed that the society is strongly connected such that G is irreducible. Here, we show briefly that the results extend to non-connected societies. For non-connected societies we introduce the following notation:

Definition 1. Let $\Pi(\mathcal{N}, G) = \{C_1, C_2, \dots, C_K, \mathcal{R}\}$ 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 C_k is strongly connected, i.e. for all $i, j \in C_k$ there exists $l \in \mathbb{N}$ such that $(G^l)_{ij} > 0$.
- Each group C_k is closed, i.e. for all $i \in C_k$, $G_{ij} > 0$ implies $j \in C_k$.
- The (possibly empty) rest of the world (ROTW) consists of the agents who do not belong to any closed and strongly connected set, i.e. $\mathcal{R} = \mathcal{N} \setminus \bigcup_{k=1}^K C_k$.

With a suitable reenumeration, the matrix G can be organized into blocks which correspond to the groups of the partition $\Pi(\mathcal{N}, G)$. Since individuals both listen and react only to their neighbors in the network G , they are not influenced by individuals to whom there does not exist a (directed) path in the network. Hence, the block structure of the law of motion M is the same as the block structure of G .

$$G = \begin{pmatrix} G_{11} & & & 0 \\ & \ddots & & \\ 0 & & G_{KK} & \\ G_{\mathcal{R}1} & \dots & G_{\mathcal{R}K} & G_{\mathcal{R}\mathcal{R}} \end{pmatrix} \Rightarrow M = \begin{pmatrix} M_{11} & & & 0 \\ & \ddots & & \\ 0 & & M_{KK} & \\ M_{\mathcal{R}1} & \dots & M_{\mathcal{R}K} & M_{\mathcal{R}\mathcal{R}} \end{pmatrix} \tag{13}$$

with $G_{kk} = G|_{C_k}$, $G_{\mathcal{R}\mathcal{R}} = G|_{\mathcal{R}}$, and $G_{\mathcal{R}k}$ consisting of the rows of G belonging to \mathcal{R} and the columns of G belonging to C_k and analogously for M .

The dynamics of M^t can then be studied independently in each strongly connected group C_k . Thus, the results of Section 3 carry over straightforwardly to the closed and strongly connected subgroups by restricting to each submatrix M_{kk} . Only in the ROTW different phenomena may occur.

In particular, the long-run opinions of the strongly connected and closed groups C_k converge to consensus and can be calculated as in (10) by restricting to the respective submatrices M_{kk} . Note that the groups may reach different consensuses. Thus by Theorem 1, each row of M_{kk}^∞ is given by the left unit eigenvector $v'_{|C_k}$, implying

$$c_k := x_i(\infty) = x_j(\infty) = v'_{|C_k} x(0)_{|C_k} \quad \forall i, j \in C_k, \tag{14}$$

where $v'_{|C_k}$ is a combination of network centrality and conformity related to the sum of the group's centrality and conformity given by (10).

As in the standard model, the ROTW cannot influence the long-run opinion of anyone including themselves since the groups C_k do not pay attention to them, and, thus, their own opinions get washed out in the long-run as $|\lambda_{\mathcal{R}\mathcal{R}}| \in (-1, 1)$ for all eigenvalues $\lambda_{\mathcal{R}\mathcal{R}}$ of the submatrix $M_{\mathcal{R}\mathcal{R}}$. The long-run influence of the respective groups C_k on the ROTW can be calculated to be (see Proposition A.1 in Appendix A.4)

$$M_{\mathcal{R}k}^\infty = (I - G_{\mathcal{R}\mathcal{R}})^{-1} G_{\mathcal{R}k} M_{kk}^\infty \tag{15}$$

for all $k = 1, \dots, K$. Since $\Gamma := (I - G_{\mathcal{R}\mathcal{R}})^{-1} (G_{\mathcal{R}1} \mathbb{1}_{|C_1}, \dots, G_{\mathcal{R}K} \mathbb{1}_{|C_K})$ is easily seen to be row-stochastic, the long-run opinion of individuals in the ROTW $i \in \mathcal{R}$ are simply weighted averages of the consensuses of the groups C_k to which they are connected by G , i.e.

$$x_i(\infty) = e_i' \Gamma c = \gamma_{i,1} c_1 + \dots + \gamma_{i,k} c_k, \tag{16}$$

where e_i is the i -th unit vector. Since the weights γ_{ij} only depend on the network G , agents in the ROTW average between consensuses of the groups exactly as in the standard DeGroot model. Consequently, the long-run opinion of an agent in the ROTW does not depend on an initial opinion of any agent within the ROTW (including herself). Thus, the only way conformity of agents in the ROTW can affect long-run opinions is to induce divergence, since extreme counter-conformity can induce divergence just as in the closed and strongly connected groups. The agents in the ROTW do not necessarily reach a consensus if there is more than just one closed and strongly connected group, since each agent in the ROTW may average differently between consent opinions of the closed and strongly connected groups.

Thus, the previous analysis extends straightforwardly to multiple groups with the difference that multiple consensuses may emerge. The ROTW features the phenomenon that individuals are not influential and hence their conformity does not matter (except for convergence).

4. Wisdom

“The whole problem with the world is that fools and fanatics are always so certain of themselves, but wiser people so full of doubts.”¹⁷

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). We will now introduce the concept of wisdom to our model (Section 4.1) and study its properties for members of closed and strongly-connected groups (Section 4.2) and for members of the rest of the world (Section 4.3). Finally, we turn to asymptotical wisdom in large societies (Section 4.4).

4.1. Framework

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, agent i 's initial opinion $x_i(0)$ is a random variable with expected value μ and some individual variance σ_i^2 , and for all $i \in \mathcal{N}$, $x_i(0)$ are uncorrelated random variables. Assuming that opinion dynamics converge, a natural question to ask is how close the different long-run opinions will be to the true, but to the agents unknown, value μ .¹⁸ To this end, let us denote the long-run consensus opinion of some group by $\hat{\mu}$, with $\hat{\mu}$ being – ex ante – random due to the random initial values $x_1(0), \dots, x_n(0)$. The difference $\hat{\mu} - \mu$ then describes by how much the long-run consensus opinion deviates from the true value μ , and $(\hat{\mu} - \mu)^2$ is a measure for the amount of this estimation error. For measuring the ex ante expected magnitude of the estimation error, we use the expected value of $(\hat{\mu} - \mu)^2$, i.e. the mean squared error (MSE) of $\hat{\mu}$ as an estimator of μ , defined as $E((\hat{\mu} - \mu)^2)$.¹⁹

¹⁷ Credit for this quote is often given to 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” (10 May 1933), which can be found on pp. 203–204 in the collection of essays “Mortals and Others”.

¹⁸ Recall that in the long run true opinions and stated opinions coincide and there is consensus within groups.

¹⁹ The mean squared error as a measure of wisdom has also been used by Rauhut and Lorenz (2011).

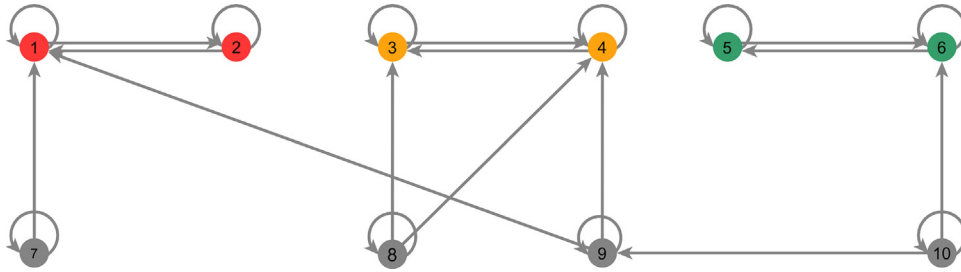


Fig. 2. Network for wisdom example. This society consists of three closed and strongly connected groups of size two and four agents in the rest of the world.

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 $M^\infty \mathbf{1} = \mathbf{1}$, it is obvious that $E(x(\infty)) = \mu \mathbf{1}$, 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 $M^\infty \Sigma (M^\infty)'$. To illustrate the effects of conformity on wisdom, we study the following example, to which we will return in the next two subsections.

Example. Let $n = 10$, $(\sigma_1^2, \dots, \sigma_{10}^2) = (6, 4, 8, 7, 6, 3, 10, 12, 14, 16)$, 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, $C_1 = \{1, 2\}$, $C_2 = \{3, 4\}$, and $C_3 = \{5, 6\}$, while Players 7–10 form the rest of the world, as also illustrated in Fig. 2. 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.

4.2. Wisdom of groups

Due to (14), a group C_k will, given convergence, eventually end up reaching a consensus where all agents' opinions are equal to $c_k = v'_{C_k} x(0)_{C_k} = \hat{\mu}_k$.²⁰ Hence, we can directly derive group C_k 's wisdom as the MSE of $\hat{\mu}_k$.

²⁰ The discussion of groups also applies to a strongly connected society, since such a society simply consists of one single group.

Lemma 2. The MSE of $\hat{\mu}_k$ is given by

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

We may use Lemma 2 to identify the individual contributions to the MSE in a given group C_k . First, from Lemma 2 it follows directly that

$$\text{MSE}_k = \sum_{i \in C_k} v_i^2 \sigma_i^2 \leq \sum_{i \in C_k} v_i \sigma_i^2 \leq \max_{i \in C_k} \sigma_i^2, \tag{17}$$

since $v_i^2 \leq v_i$ due to $v_i \in (0, 1]$ for all agents i . Thus, group 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 (17) become equalities, which is the case for $v_i \in \{0, 1\}$ for all $i \in C_k$ (first inequality) and $v_i = 0$ for all i with $\sigma_i^2 < \max_{j \in C_k} \sigma_j^2$ (second inequality). Therefore, information updating within group 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 statics effect of one agent's conformity on the wisdom of her group.

Proposition 3. 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.

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

To give an interpretation for Proposition 3, let us rewrite $v_i \sigma_i^2 = v_i / (1 / \sigma_i^2)$ and $\text{MSE}_k = \sum_{j \in C_k} v_j (v_j / (1 / \sigma_j^2))$. 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, $v_i / (1 / \sigma_i^2)$: 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.²¹

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 $i, j \in C_k$. This is formalized in the following corollary of Proposition 3.

Corollary 2. For the wisdom of group C_k as measured by MSE_k , we have

$$\text{MSE}_k \geq \frac{1}{\sum_{j \in C_k} \frac{1}{\sigma_j^2}} =: \text{MSE}_k^*, \tag{18}$$

with equality in (18) if and only if $v_i \sigma_i^2 = v_j \sigma_j^2$ for all $i, j \in C_k$. The latter condition is equivalent to

$$\delta_i = 1 - a \frac{1}{\sigma_i^2 w_i \sum_{j \in C_k} \frac{1}{\sigma_j^2}} \text{ for all } i \in C_k \tag{19}$$

for some constant $a \in (0, 2 \sum_{j \in C_k} (1 / \sigma_j^2) \min_{j \in C_k} w_j \sigma_j^2)$.

Corollary 2 delivers the analogue to (17). While (17) describes the worst case with respect to wisdom, Corollary 2 considers the best scenario: all agents within a 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 $a \in (0, \sum_{j \in C_k} (1 / \sigma_j^2) \min_{j \in C_k} w_j \sigma_j^2]$ in (19) ensures $\delta_i \geq 0$ for all $i \in C_k$ and therefore by Proposition 2 guarantees convergence of the opinions in 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 C_k , as is easily seen from (18). Therefore, under optimal conformity all agents within C_k benefit from communication.

Reconsidering the example discussed in Section 4.1, we find the network centralities (the left 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 (19) the constant a can be chosen in the intervals $(0, 2/3)$ (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, setting the agents' degrees of conformity according to these values ensures the optimal wisdom within the respective groups, given by $(2.4, 2.4, 3.7\bar{3}, 3.7\bar{3}, 2, 2, 2.4, 3.7\bar{3}, 1.5\bar{3}, 0.88\bar{3})$. The same level could also be reached for other conformity levels, for instance, choosing $a = 1/4$ (first group), $a = 3/7$ (second

²¹ An analogous discussion can be already found in DeMarzo et al. (2003) for the case where agents are honest.

group), and $a = 3/8$ (third group) in (19), 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).

4.3. 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 (16), 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 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. (16)). This immediately translates into the wisdom of an agent $i \in \mathcal{R}$ as follows:

$$E((x_i(\infty) - \mu)^2) = \sum_{k=1}^K \gamma_{i,k}^2 \text{MSE}_k \leq \max_{k=1, \dots, K} \text{MSE}_k. \tag{20}$$

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 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 Section 4.1 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 (cf. Fig. 2). 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 (20). 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 4. For agents $i \in \mathcal{R}$, we have

$$E((x_i(\infty) - \mu)^2) \geq \frac{1}{\sum_{k=1}^K \frac{1}{\text{MSE}_k}}, \tag{21}$$

with equality if and only if

$$\gamma_{i,k} = \frac{1}{\text{MSE}_k \sum_{l=1}^K \frac{1}{\text{MSE}_l}}$$

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

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.

4.4. Asymptotic learning

So far, we have studied learning in the framework of opinion dynamics under conformity with respect to a finite population. Since information aggregation is simply a weighted average of initial signals, individuals will never learn the true state of the world μ with probability 1, a commonly observed finding in models based on naïve learning. Golub and Jackson (2010) instead look at growing societies and determine conditions for wise crowds such that the obtained consensus converges in probability to the unknown state of the world.

To show how our approach to learning extends to large crowds, denote by $N(n)$ the population set depending on the population size n and assume that there is a sequence of aperiodic and strongly connected networks $[G(n)]_{n=1}^{\infty}$ as well as a sequence of conformity values $[\delta_{i,n}]_{n=1}^{\infty}$ for all $i \in N(n)$ such that each law of motion M in the sequence $[M(n)]_{n=1}^{\infty}$ is

convergent.²² As before, let each individual $i \in N(n)$ receive an unbiased signal $x_{i,n}(0)$ which is drawn from an independent distribution with mean μ and bounded variance $\sigma_{i,n}^2$ such that $0 < \underline{\sigma}^2 \leq \sigma_{i,n}^2 \leq \bar{\sigma}^2 < \infty$. We say that the society $N(n)$ is asymptotically wise if the MSE tends to zero, i.e. $E((\hat{\mu}(n) - \mu)^2) \rightarrow 0$ for $n \rightarrow \infty$, where the estimator $\hat{\mu}(n)$ is given by the consensus opinion $x_n(\infty)$, as before. Since initial signals are unbiased and variance is bounded, this is easily seen to be equivalent to $\hat{\mu}(n)$ converging to μ in probability which is exactly what Golub and Jackson (2010) define to be wise crowds (cf. Golub and Jackson, 2010, Definition 3).

Since by Lemma 2 (see also Golub and Jackson, 2010, Proposition 2), $\max_{i=1, \dots, n} v_{i,n} \rightarrow 0$ for $n \rightarrow \infty$ is necessary and sufficient for (asymptotically) wise crowds, we get that the society is asymptotically wise if and only if

$$\frac{\max_{i=1, \dots, n} w_{i,n}(1 - \delta_{i,n})}{\sum_{i=1}^n w_{i,n}(1 - \delta_{i,n})} \xrightarrow{n \rightarrow \infty} 0. \tag{22}$$

Now, we can study the effect of conformity on asymptotical wisdom. First, suppose that $\delta_{i,n}$ is bounded away from 1, i.e. conformity is bounded.

Corollary 3. *Let $\delta_{i,n} \leq \bar{\delta} < 1$. The society is asymptotically wise if and only if $\max_{i=1, \dots, n} w_{i,n} \rightarrow 0$.*

Thus, for bounded conformity, we arrive exactly at the same result as Golub and Jackson (2010): the society is asymptotically wise if and only if no individual has excessive influence on others in terms of a prominent network position.

However, even if some members of the society become too prominent, which undermines asymptotical wisdom under honesty (cf. Golub and Jackson, 2010, Proposition 3), high conformity of these prominent families may still result in asymptotical wisdom, as we will show.

Assume for example that we have $w_{i,n} > 0$ for $n \rightarrow \infty$ such that i asymptotically becomes too important with respect to the network position. Then, for crowds to be wise, we require $\delta_{i,n} \rightarrow 1$, i.e. prominent individuals have to be asymptotically highly conform. However, this is not sufficient as we need to ensure that $\delta_{i,n}$ converges to 1 at the right speed. Too slow convergence to 1 may not be sufficient for their prominence to vanish and too fast convergence may cause non-prominent agents with vanishing network centrality ($w_{j,n} \xrightarrow{n \rightarrow \infty} 0$) to emerge as prominent in terms of $v_{j,n}$, causing asymptotical wisdom to fail. The following result presents tight conditions on conformity for the society to be asymptotically wise.

Proposition 5. *The society is asymptotically wise if and only if $\delta_{i,n}$ is of the form*

$$\delta_{i,n} = 1 - \beta_n \frac{\alpha_{i,n}}{w_{i,n}}$$

with positive sequences $[\beta_n]_{n=1}^\infty$ and $[\alpha_{i,n}]_{n=1}^\infty$ such that $M(n)^t$ converges in t (for all n), $\beta_n, \alpha_{i,n} > 0$, $\sum_{i=1, \dots, n} \alpha_{i,n} = 1$, $\beta_n < 2 \min_{i=1, \dots, n} w_{i,n} / \alpha_{i,n}$, and $\max_{i=1, \dots, n} \alpha_{i,n} \xrightarrow{n \rightarrow \infty} 0$.

To give some intuition to the above conditions, note that $1 - \delta_{i,n}$ being proportional to $\alpha_{i,n} / w_{i,n}$ ensures that power $v_{i,n}$ is balanced in an asymptotical sense. Non-vanishing network centrality then yields vanishing power since $\alpha_{i,n} \rightarrow 0$. In other words, those with high network centrality, need to be highly conforming. The sequence β_n ensures that all $\delta_{i,n}$ are distributed within $(-1, 1)$. In particular, choosing $\beta_n < 1 \cdot \min_{i=1, \dots, n} w_{i,n} / \alpha_{i,n}$ would exclude counter-conformity and thereby imply convergence for all $\delta_{i,n}$ by Proposition 2. In societies where the underlying sequence of networks is already wise (see Golub and Jackson, 2010), i.e. where honesty ($\delta_{i,n} = 0$ for all $i \in N$, $n \in \mathbb{N}$) is sufficient for wisdom, $\alpha_{i,n} = w_{i,n}$ and $\beta_n = 1$ for all $i \in N$, $n \in \mathbb{N}$ fulfill the conditions of Proposition 5.

It is quite easy to see that for any sequence of connected networks many sequences $[\beta_n]_{n=1}^\infty$ and $[\alpha_{i,n}]_{n=1}^\infty$ exist that lead to asymptotic wisdom. As a particular example, consider the conformity distribution such that network centrality and signal precision are balanced as in Corollary 2, i.e. for all $n \in \mathbb{N}$ conformity is such that

$$\delta_{i,n} = 1 - \frac{a_n}{\sigma_{i,n}^2 w_{i,n} \sum_{j=1}^n \frac{1}{\sigma_{j,n}^2}}$$

with $a_n \in (0, \sum_{j=1}^n (1 / \sigma_{j,n}^2) \min_{j=1, \dots, n} w_{j,n} \sigma_{j,n}^2)$. Then it does not only hold that for any $n \in \mathbb{N}$ the MSE is minimized, but also that the society is asymptotically wise since $\delta_{i,n}$ satisfies the required conditions by choosing

$$\alpha_{i,n} = \frac{1}{\sigma_{i,n}^2 \sum_{j=1}^n \frac{1}{\sigma_{j,n}^2}}$$

and $\beta_n = a_n$. Another example which always ensures wisdom is to set $\alpha_{i,n} = 1/n$ and $\beta_n < n \cdot \min_{i=1}^n w_{i,n}$ for all $i \in N$, $n \in \mathbb{N}$. Then power is evenly distributed, $v_{i,n} = 1/n$ for all $n \in \mathbb{N}$.

To sum up, bounded conformity has no effect on asymptotical wisdom. However, even if wisdom fails under honesty by e.g. having prominent families, we still get wisdom if these families are sufficiently conform. Hence, if the sequence of

²² A sequence of convergent $[M(n)]_{n=1}^\infty$ can be achieved e.g. by only allowing non-negative values of $\delta_{i,n}$ (see Proposition 2).

conformity is allowed to converge to 1, then this affects wisdom, allowing societies to obtain asymptotical wisdom even if they are not wise under honesty.²³

5. Concluding remarks

So far, the literature on opinion dynamics has mainly 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 help us understand 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 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 among naïve agents. 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). Following 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 discuss 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 2).

Assuming convergence, we then characterize the long-run (consensus) opinion (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. Thus, taking the network as given, we conclude that low conformity fosters opinion leadership while high conformity undermines opinion leadership so that counter-conformity might be interpreted as a persuasion device. 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), but goes beyond it by showing how this effect depends on the network position and on the other conformity levels in the society.

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 in a finite population setup 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 3). 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). For large societies, we find necessary and sufficient conditions on the sequence of conformity parameters such that asymptotic wisdom emerges, i.e. such that the MSE converges to zero (Proposition 5). Hence, asymptotic wisdom is obtainable due to conformity, even if an honest society is not wise (cf. Golub and Jackson, 2010). Such a sequence of conformity exists for any sequence of networks and always reduces the prominence of influential agents by increasing their conformity which resembles the intuition gained for finite societies. 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.

²³ Of course, it is also possible to construct examples where the opposite effect emerges such that wise societies lose wisdom under conformity.

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 and workshops “SING7” in Paris, “Game Theory and Society” in Zurich, “SING8” in Budapest, “Sunbelt XXXIII” in Hamburg, “CTN 2014” in Brussels, “EEA-ESEM 2014” in Toulouse, and the participants of the research seminars in Belfast, Bern, Bielefeld, Florence, Hamburg, Milano (at Bocconi), Saarbruecken, St. Gallen, and Zurich.

Appendix A. Mathematical appendix

A.1. Expressed opinions

Proof of Proposition 1. First, notice that $s(t)$ by construction satisfies $s(t) = (I - \Delta)x(t) + \Delta Ys(t)$ 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 $(I - \Delta)x(t) + \Delta Ys^{\tau-1}(t)$. For all $i \in A^\tau(t)$, we therefore find $s_i^\tau(t) - s_i(t)$ as the i -th component of $\Delta Y(s^{\tau-1}(t) - s(t))$. We now study this difference by the maximum norm ($\|\cdot\|_\infty$), which is also called infinity norm. 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 $s(t)$ 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 $\tau_1 := \min\{\tau \in \mathbb{N} : (\forall i \in \mathcal{N})(U_i(t) \cap \{1, \dots, \tau\} \neq \emptyset)\}$ as the first time-step where every agent has at least once been satisfied with her stated opinion.²⁴ 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 $\tau_{k+1} := \min\{\tau > \tau_k : (\forall i \in \mathcal{N})(U_i(t) \cap \{\tau_k + 1, \dots, \tau\} \neq \emptyset)\}$ 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. \square

A.2. Convergence

To prove Proposition 2 (and Lemma 1 and Theorem 1) the following lemma is helpful.

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

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.

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

Thus,

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

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

Proof of Proposition 2. 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

²⁴ The assumption that all $U_i(t)$ are almost surely infinite guarantees that τ_1, τ_2, \dots are almost surely well-defined.

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.

This implies that $x(t)$ converges since $x(t) = M^t x(0)$. From Proposition 1, we get that $s(t) = (I - \Delta(I - D)^{-1}(G - D))^{-1}(I - \Delta)x(t)$. 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 $q(t)$. \square

A.3. Long-run opinions and opinion leadership

Proof of Lemma 1. We first show that $x(\infty) = s(\infty) = q(\infty)$ and then turn to $x_i(\infty) = \sum_{j \in \mathcal{N}} v_j x_j(0)$ for all $i \in \mathcal{N}$.

From $x(t + 1) = Dx(t) + (G - D)s(t) = Dx(t) + (I - D)q(t)$, we can infer for all $t \geq 0$ that $x(t) = D^t x(0) + \sum_{l=0}^{t-1} D^{t-1-l}(I - D)q(l)$, 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

$$\lim_{t \rightarrow \infty} x(t) = \lim_{t \rightarrow \infty} \sum_{l=0}^{t-1} D^{t-1-l}(I - D)q(l) = \lim_{t \rightarrow \infty} \sum_{l=0}^{t-1} D^{t-1-l}(I - D)(q(l) - q(\infty)) + \lim_{t \rightarrow \infty} \sum_{l=0}^{t-1} D^{t-1-l}(I - D)q(\infty).$$

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.

For the second part, we use the assumption that the society is strongly connected, which is equivalent to $\text{rk}(I - G) = n - 1$. This implies $\text{rk}(I - M) = n - 1$ due to Lemma A.1 which also delivers $M\mathbb{1} = \mathbb{1}$.

Observe that $MM^\infty = M^\infty = M^\infty M$. This implies that

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

from which we find $M^\infty = r \mathbb{1} v'$ for some real number r which is found to be equal to 1 as $\mathbb{1} = M^\infty \mathbb{1} = r \mathbb{1} v' \mathbb{1} = r \mathbb{1}$. Hence, $x(\infty) = M^\infty x(0) = \mathbb{1} v' x(0)$. \square

Proof of Theorem 1. As $\text{rk}(I - G) = n - 1$ (by strong connectedness) and $v'(M - I) = 0$, 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 that the left unit eigenvector w of G and v relate to each other by

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

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 $r = 1/w'(I - \Delta)\mathbb{1}$, which shows that $v = (I - \Delta)w/\mathbb{1}'(I - \Delta)w$. This yields (10). \square

A.4. Beyond strongly connected networks

Proposition A.1 (M^∞ for ROTW). If M^t converges, we get

$$M_{\mathcal{R}\mathcal{R}}^\infty = 0 \quad \text{and} \quad M_{\mathcal{R}k}^\infty = (I - G_{\mathcal{R}\mathcal{R}})^{-1} G_{\mathcal{R}k} M_{kk}^\infty. \tag{A.2}$$

Proof of Proposition A.1. Suppose that M^t converges to M^∞ . To determine the formulas for $M_{\mathcal{R}\mathcal{R}}^\infty$ and $M_{\mathcal{R}k}^\infty$, we first establish that $GM^\infty = M^\infty$. We have $Gx = x \Leftrightarrow (I - G)x = 0 \Leftrightarrow [I - (G - D)\Delta(I - D)^{-1}]^{-1}(I - G)x = 0$, since by Lemma A.1 $[I - (G - D)\Delta(I - D)^{-1}]$ is invertible. Thus by Lemma A.1, $Gx = x$ if and only if $Mx = x$. Furthermore, $MM^\infty x = M^\infty x$ and therefore $GM^\infty x = M^\infty x$ for all n -dimensional vectors x , delivering $GM^\infty = M^\infty$. Since the block structure of G and M is the same, this implies

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

A.5. Wisdom of groups

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

Proof of Proposition 3.

Please refer to Eq. (11).

$$\frac{\partial \text{MSE}_k}{\partial \delta_i} = \frac{\partial \sum_{j \in C_k} v_j^2 \sigma_j^2}{\partial \delta_i} = \sum_{j \in C_k} 2\sigma_j^2 v_j \frac{\partial v_j}{\partial \delta_i} = \frac{2w_i}{\sum_{j \in C_k} w_j(1-\delta_j)} \sum_{j \in C_k} \sigma_j^2 v_j (v_j - 1_{j=i}).$$

The assertion follows easily noting that $\text{MSE}_k = \sum_{j \in C_k} v_j v_j \sigma_j^2$. \square

A.6. Wisdom within the Rest of the World

Proof of Proposition 4. First, notice that $E((x_i(\infty) - \mu)^2) = \sum_{k=1}^K \gamma_{i,k}^2 \text{MSE}_k$, with $\sum_{k=1}^K \gamma_{i,k} = 1$ for all $i \in \mathcal{R}$. By the Cauchy-Schwarz inequality, we have

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

with equality if and only if there exists some (necessarily positive) constant a such that

$$\gamma_{i,k} \sqrt{\text{MSE}_k} = a \frac{1}{\sqrt{\text{MSE}_k}}$$

for all k . We therefore have

$$\sum_{k=1}^K \gamma_{i,k}^2 \text{MSE}_k \geq \frac{1}{\sum_{k=1}^K \frac{1}{\text{MSE}_k}},$$

with equality if and only if

$$\gamma_{i,k} = \frac{1}{\text{MSE}_k \sum_{l=1}^K \frac{1}{\text{MSE}_l}}$$

for all k . \square

A.7. Asymptotic learning

Proof of Corollary 3. From $-1 < \delta_{i,n} \leq \bar{\delta}$, we find that $1 - \bar{\delta} \leq 1 - \delta_{i,n} < 2$. The denominator of the ratio

$$\frac{\max_{i=1, \dots, n} w_{i,n} (1 - \delta_{i,n})}{\sum_{i=1}^n w_{i,n} (1 - \delta_{i,n})} \xrightarrow{n \rightarrow \infty} 0$$

is therefore always between the positive numbers $1 - \bar{\delta}$ and 2 such that the ratio converges to zero if and only if its numerator converges to zero. With respect to the numerator, as the factor $1 - \delta_{i,n}$ is both bounded from above by 2 and bounded away from zero by $1 - \bar{\delta} > 0$, the numerator converges to zero if and only if $\max_{i=1, \dots, n} w_{i,n} \xrightarrow{n \rightarrow \infty} 0$ converges to zero, which completes the proof. \square

Proof of Proposition 5.

1. If $\delta_{i,n} = 1 - \beta_n \frac{\alpha_{i,n}}{w_{i,n}}$, then $w_{i,n} (1 - \delta_{i,n}) = \beta_n \alpha_{i,n}$ and

$$\frac{\max_{i=1, \dots, n} w_{i,n} (1 - \delta_{i,n})}{\sum_{i=1}^n w_{i,n} (1 - \delta_{i,n})} = \frac{\max_{i=1, \dots, n} \beta_n \alpha_{i,n}}{\sum_{i=1}^n \beta_n \alpha_{i,n}} = \max_{i=1, \dots, n} \alpha_{i,n} \xrightarrow{n \rightarrow \infty} 0,$$

i.e. the society is asymptotically wise.

2. Let the society be asymptotically wise. Set

$$\alpha_{i,n} := \frac{w_{i,n} (1 - \delta_{i,n})}{\sum_{j=1}^n w_{j,n} (1 - \delta_{j,n})}$$

and $\beta_n := \sum_{j=1}^n w_{j,n}(1 - \delta_{j,n})$. Then $\beta_n, \alpha_{i,n} > 0$, $\beta_n < 2\min_{i=1}^n w_{i,n}/\alpha_{i,n}$, and $\sum_{i=1}^n \alpha_{i,n} = 1$. Furthermore, $w_{i,n}(1 - \delta_{i,n}) = \beta_n \alpha_{i,n}$, entailing

$$\max_{i=1}^n \alpha_{i,n} = \max_{i=1}^n \frac{\alpha_{i,n}}{\sum_{j=1}^n \alpha_{j,n}} = \max_{i=1}^n \frac{\beta_n \alpha_{i,n}}{\sum_{j=1}^n \beta_n \alpha_{j,n}} = \max_{i=1}^n \frac{w_{i,n}(1 - \delta_{i,n})}{\sum_{j=1}^n w_{j,n}(1 - \delta_{j,n})} \xrightarrow{n \rightarrow \infty} 0.$$

□

References

- Acemoglu, D., Dahleh, M.A., Lobel, I., Ozdaglar, A., 2011. Bayesian learning in social networks. *Rev. Econ. Stud.* 78 (4), 1201–1236.
- Acemoglu, D., Ozdaglar, A.E., 2011. Opinion dynamics and learning in social networks. *Int. Rev. Econ.* 1 (1), 3–49.
- Acemoglu, D., Ozdaglar, A.E., ParandehGheibi, A., 2010. Spread of (mis)information in social networks. *Games Econ. Behav.* 70 (2), 194–227.
- Ariely, D., Levav, J., 2000. Sequential choice in group settings: taking the road less traveled and less enjoyed. *J. Consum. Res.* 27 (3), 279–290.
- Asch, S., 1955. Opinions and social pressure. *Sci. Am.* 193 (5), 31–35.
- Bala, V., Goyal, S., 1998. Learning from neighbours. *Rev. Econ. Stud.* 65 (3), 595–621.
- Bala, V., Goyal, S., 2001. Conformism and diversity under social learning. *Econ. Theory* 17 (1), 101–120.
- Banerjee, A.V., 1992. A simple model of herd behavior. *Q. J. Econ.* 107 (3), 797–817.
- Battiston, P., Stanca, L., 2014. Boundedly Rational Opinion Dynamics in Directed Social Networks: Theory and Experimental Evidence. Working Papers 267, University of Milano-Bicocca, Department of Economics.
- Bernheim, B.D., 1994. A theory of conformity. *J. Polit. Econ.* 102 (5), 841–877.
- Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change in informational cascades. *J. Polit. Econ.* 100 (5), 992–1026.
- Bonacich, P., 1972. Factoring and weighting approaches to status scores and clique identification. *J. Math. Sociol.* 2 (1), 113–120.
- Breza, E., Chandrasekhar, A.G., Larreguy, H., 2014. Social Structure and Institutional Design: Evidence from a Lab Experiment in the Field. Technical Report.
- Buechel, B., Hellmann, T., Pichler, M.M., 2014. The Dynamics of Continuous Cultural Traits in Social Networks. *J. Econ. Theory* 154, 274–309, <http://dx.doi.org/10.1016/j.jet.2014.09.008>.
- Chan, K.K., Misra, S., 1990. Characteristics of the opinion leader: a new dimension. *J. Advert.* 19 (3), 53–60.
- Chandrasekhar, A.G., Larreguy, H., Xandari, J.P., 2012. Testing Models of Social Learning on Networks: Evidence from a Framed Field Experiment. Mimeo.
- Corazzini, L., Pavesi, F., Petrovich, B., Stanca, L., 2012. Influential listeners: an experiment on persuasion bias in social networks. *Eur. Econ. Rev.* 56 (6), 1276–1288.
- DeGroot, M.H., 1974. Reaching a consensus. *J. Am. Stat. Assoc.* 69 (345), 118–121.
- DeMarzo, P.M., Vayanos, D., Zwiebel, J., 2003. Persuasion bias, social influence, and unidimensional opinions. *Q. J. Econ.* 118 (3), 909–968.
- Deutsch, M., Gerard, H.B., 1955. A study of normative and informational social influences upon individual judgement. *J. Abnorm. Psychol.* 51 (3), 629–636.
- Ellison, G., Fudenberg, D., 1993. Rules of thumb for social learning. *J. Polit. Econ.* 101 (4), 612–643.
- Ellison, G., Fudenberg, D., 1995. Word-of-mouth communication and social learning. *Q. J. Econ.* 110 (1), 93–125.
- Festinger, L., 1957. *A Theory of Cognitive Dissonance*. Row, Peterson, Evanston, IL.
- Flache, A., Torenvlied, R., 2004. When will they ever make up their minds? The social structure of unstable decision making. *J. Math. Sociol.* 28, 171–196.
- Foerster, M., Mauleon, A., 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. *Psychol. Rev.* 63, 181–194.
- Friedkin, N.E., 1991. Theoretical foundations for centrality measures. *Am. J. Sociol.* 96 (6), 1478–1504.
- Friedkin, N.E., Johnsen, E.C., 1990. Social influence and opinions. *J. Math. Sociol.* 15 (3–4), 193–206.
- Gale, D., Kariv, S., 2003. Bayesian learning in social networks. *Games Econ. Behav.* 45 (2), 329–346.
- Galeotti, A., Goyal, S., 2009. Influencing the influencers: a theory of strategic diffusion. *RAND J. Econ.* 40 (3), 509–532.
- Golub, B., Jackson, M.O., 2010. Naïve learning in social networks and the wisdom of crowds. *Am. Econ. J.: Microecon.* 2 (1), 112–149.
- Grabisch, M., Rusinowska, A., 2010. A model of influence in a social network. *Theory Decis.* 69 (1), 69–96.
- Grabisch, M., Rusinowska, A., 2011. Influence functions, followers and command games. *Games Econ. Behav.* 72 (1), 123–138.
- Grimm, V., Mengel, F., 2013. An Experiment on Belief Formation in Networks. Available at SSRN 2361007.
- Harary, F., 1959. Status and contrastatus. *Sociometry* 22, 23–43.
- Hegselmann, Krause, 2002. Opinion dynamics and bounded confidence, models, analysis and simulation. *J. Artif. Soc. Soc. Simul.* 5 (3), 1–33.
- Hogg, M., 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., McKimmie, B.M., 2003. On being loud and proud: non-conformity and counter-conformity to group norms. *Br. J. Soc. Psychol.* 42 (3), 319–335.
- Jackson, M.O., 2008. *Social and Economic Networks*. Princeton University Press.
- Jadbabaie, A., Molavi, P., Sandroni, A., Tahbaz-Salehi, A., 2012. Non-Bayesian social learning. *Games Econ. Behav.* 76 (1), 210–225.
- Jadbabaie, A., Molavi, P., Tahbaz-Salehi, A., 2013. Information Heterogeneity and the Speed of Learning in Social Networks. Technical Report 13–28, Columbia Business School.
- Jones, S.R.G., 1984. *The economics of conformism*. Blackwell, Oxford, New York, NY.
- Kalai, A., Kalai, E., 2001. Strategic polarization. *J. Math. Psychol.* 45, 656–663.
- Lorenz, J., 2005. A stabilization theorem for dynamics of continuous opinions. *Physica A: Stat. Mech. Appl.* 355 (1), 217–223.
- Maslach, C., Stapp, J., Santee, R.T., 1985. Individuation: conceptual analysis and assessment. *J. Personal. Soc. Psychol.* 49, 729–738.
- Merlone, U., Radi, D., 2014. Reaching consensus on rumors. *Physica A: Stat. Mech. Appl.* 406 (0), 260–271.
- Mueller-Frank, M., 2013. A general framework for rational learning in social networks. *Theor. Econ.* 8 (1), 1–40.
- Mueller-Frank, M., 2014. Does one Bayesian make a difference? *J. Econ. Theory* 154, 423–452, <http://dx.doi.org/10.1016/j.jet.2014.09.005>.
- Pan, Z., 2010. Trust, influence, and convergence of behavior in social networks. *Math. Soc. Sci.* 60 (1), 69–78.
- Rauhut, H., Lorenz, J., 2011. The wisdom of crowds in one mind: how individuals can simulate the knowledge of diverse societies to reach better decisions. *J. Math. Psychol.* 55 (2), 191–197.
- Rosenberg, D., Solan, E., Vieille, N., 2009. Informational externalities and emergence of consensus. *Games Econ. Behav.* 66 (2), 979–994.
- Sethi, R., Yildiz, M., 2012. Public disagreement. *Am. Econ. J.: Microecon.* 4 (3), 57–95.
- Sethi, R., Yildiz, M., 2013. Perspectives, Opinions, and Information Flows. Technical Report, Department of Economics, Massachusetts Institute of Technology, Cambridge, MA.
- Smith, L., Sorensen, P., 2000. Pathological outcomes of observational learning. *Econometrica* 68 (2), 371–398.
- Zafar, B., 2011. An experimental investigation of why individuals conform. *Eur. Econ. Rev.* 55 (6), 774–798.