What makes an opinion leader: expertise versus popularity*

Theodoros Rapanos†

July 4, 2015

Abstract

This paper studies learning by using information obtained through social or professional networks. Based on the method first proposed by DeGroot (1974), agents form their beliefs by weighting the information acquired from their peers. The innovation lies in the introduction of dynamically updated weights. This allows agents to weight a contact with poor information little at first, but more later, if that contact has accumulated second-hand information from other, more knowledgeable agents. The main finding is that the social influence of each agent will depend on both their popularity (as captured by eigenvector centrality) and their expertise (as captured by statistical precision) in a simple and intuitively appealing way. It is moreover shown that even completely uninformed agents can affect public opinion. The paper also studies how the relationship between expertise and popularity in a network affects the speed and the efficiency of the learning process.

*Preliminary and incomplete version. Kindly do not quote.

†Department of Economics, Stockholm University, 10691 Stockholm, Sweden. theodoros.rapanos@ne.su.se.
I would like to thank Yves Zenou for his very helpful guidance and comments. Jakob Almerud, Ana Maria Čeh, Paola Di Casola, and Spyros Sichlimiris provided me with useful feedback that helped improve this paper. I would also like to acknowledge the financial support from Svenska Handelsbanken's grant H2012-0361:1.

1 Introduction

Individuals, firms, organizations and governments make use of information in order to form opinions and take decisions. Most of the times, however, they do not have access to the primary source of information, and hence the largest share of this information comes to them through contacts, who themselves may have in turn acquired this information indirectly. Social and professional networks play thus a key role in the transmission of information in the society.

Apart from the acquisition and transmission of information, networks may affect the formation of opinions. Consumers often ask the opinion of friends who have a deeper knowledge or a better understanding of the relevant area before deciding on buying a new computer or car. Firm executives consult experts before making strategic decisions. In some occasions even politicians ask for the opinion of persons considered specialists in a field before they attempt to implement what they have in mind.

Although the importance of networks as channels of transmission of both information and ideas is widely accepted (see, for example, Jackson, 2008; Moretti, 2011; Algan, Do, Le Chapelain, and Zenou, 2015, or Jackson, Rogers, and Zenou, 2015 for a recent survey) it is less clear how the agents aggregate the information or the opinions the get through their networks. There are two main benchmarks, often referred to as *fully rational* and *naïve* learning respectively. In practice though this distinction is not always clear, and in fact most models encompass elements of both approaches.

The present work follows the paradigm of the so-called *average-based updating process*, first introduced by DeGroot (1974). In this seminal study the author suggests a method for how a group of individuals, such as a committee, exchange opinions and update their beliefs about the value of some unknown parameter they wish to estimate. The process proposed is simple and intuitively appealing: Each group member assigns some weight (a degree of "trust") to each other member, and in every period revises his opinion by taking the weighted average of their beliefs. These weights can be subjective, and thus not all members need to agree on them. Moreover attention may not be necessarily reciprocal, and in fact some member may even completely disregard the opinion of other members. It follows thus that that the communication structure can be captured by a directed and weighted network. In the paper it is shown that if information can flow throughout the whole network without going in circles, a consensus among the members of the group will be attained.

Based on DeGroot's model, DeMarzo, Vayanos, and Zwiebel (2003) study the formation of (political) opinions in a society in the presence of what they refer to as *persuasion bias*. People, either because they are not aware of the structure of the network, or because they do not lack the ability or time to fully track the flow of information through the network, fail to account for the repetition of the opinions they receive. As result, their beliefs may be driven away from both the true value of the parameter and the society initial average due to the influence of some prominently positioned individuals. This result follows immediately from the fact that in DeGroot's model, agents are assumed to behave in a *naïve* rather than a fully rational way. Another interesting finding is that in the medium-run, that is, after some rounds of communication, yet in finite time and before a consensus is reached, each individual's beliefs will have become *unidimensional*. In other words, even if an individual starts having different views on a series of (possibly non-related) issues, his beliefs on all of them will sooner or later move towards the same direction (e.g. towards the left or the right end of the political spectrum) due to his exposure to persuasion bias.

Golub and Jackson (2010) use the same setup but introduce rigorous networks terminology, and a more standard networks-based approach. They find that in large societies the effect of persuasion bias can be neutralised only if the relative influence of the prominent individuals grows to zero, and the highly influential groups give increasing attention to individuals outside the group as the number of individuals becomes arbitrarily large.

The present study proposes and analyses a variant of the the DeGroot method in which individuals can adjust the degree of trust they assign to their peers based on the precision of any second-hand information these individuals came across after their last communication. This allows individuals to place a low initial weight to one of their peers, but in the next round increase that weight if they know that this peer has contacts who have more accurate information.

The main result establishes that, in the limit, each individual's influence will be determined in a simple and straightforward way by both his position in the network and the precision of the information he possesses. This suggests that the distribution of access to reliable information over the individuals in a network will have implications about the efficiency of the learning process. More specifically, learning is more efficient in networks where the individuals with better information occupy a quite central position in the network. Nevertheless the views of other individuals should not be completely ignored, since they contain at least some positive informational value. Next we also examine the speed of the learning process. Moreover, individuals without information or knowledge of their own will be in a position to affect public opinion by propagating information or opinions they acquired from other agents. This is in contrast with the findings of the canonical DeGroot model, where such individuals are ignored by their peers, and thus do not affect the resulting beliefs.

Before proceeding with the technicalities of the model, it would be useful to provide some guidelines about the notation used in this paper. Matrices shall be denoted with bold capital letters, for example $\mathbf{X}$, and vectors with bold lowercase letters, for example $\mathbf{y}$. The (i,j)-th element of matrix $\mathbf{X}$ shall be denoted with x_{ij}, and the i-th element of vector $\mathbf{y}$ with y_i. All vectors are defined as column vectors, and thus its transpose, $\mathbf{y}^\top$, will be a row vector. Finally, if $\mathbf{y}$ in an n-dimensional vector, *diag*($\mathbf{y}$) shall denote an $n \times n$ diagonal matrix with the elements of vector $\mathbf{y}$ on its main diagonal, with the rest of its elements equal to zero.

2 The setup

2.1 The agents

Consider a society consisting of a finite number of individuals who would like to gather more information about a parameter of interest or form an opinion regarding an issue they need to make a decision on. The pattern of communication among the agent is captured by a network $\mathcal{G}$.

As in the rest of the literature on average-based updating, it will be assumed that the agents are only interested in estimating the true value of the unknown parameter, and stick to the stipulated updating process; they do seek to maximize their social influence, nor can they have something to gain from propagating a particular belief.

2.2 The network

A network is modelled as a -potentially directed- graph $\mathcal{G} = \langle \mathcal{N}, \mathcal{E} \rangle$, where the set of *vertices* or *nodes*, $\mathcal{N} := \{1, 2, \dots, n\}$, represents a set of agents who can potentially interact with each other, and the set of *edges* $\mathcal{E} \subseteq \mathcal{N}^2$ represents the links among them. In many applications of network theory it is more convenient to represent a network using a matrix $\mathbf{G} := [g_{ij}]_{(i,j)\in\mathcal{N}^2} \in \{0, 1\}^{n\times n}$, where $g_{ij} := 1$ if there is a directed edge from node i to node j (i.e. agent i is linked to agent j), and $g_{ij} := 0$ otherwise. In network theory terminology, matrix $\mathbf{G}$ is referred to as the *adjacency matrix* of network $\mathcal{G}$.

In the present model, the adjacency matrix $\mathbf{G}$ captures the pattern of communication and transmission of opinions across the network. Consider any two agents $i, j \in \mathcal{N}$. A link from agent i to agent j, $g_{ij} = 1$, has the interpretation that agent i has access to agent j's belief. It shall be then said that agent i *observes*, *pays attention to*, or *listens to* agent j, or equivalently that agent j is an *out-neighbour* of agent i.

Two important remarks are in order at this point. First, as the above discussion suggests, attention may not be reciprocal: the fact that agent i can observe agent j's belief does not necessarily imply that j is in a position to observe i's belief. Hence the adjacency matrix **G** can, but need not be symmetric. Second, it is reasonable to assume that every agent can observe himself.[1] The diagonal of **G** will thus consist of ones, that is, $g_{ii} = 1$ for all $i \in \mathcal{N}$. This assumption will be maintained throughout this paper and will not be stated explicitly again.

The set of all agents observed by agent i (that is, all the out-neighbours of agent i) in network $\mathcal{G}$ constitutes the *out-neighbourhood* of agent i, and is denoted with $\mathcal{D}_{\mathcal{G}}(i)$. Using mathematical notation, for any $i \in \mathcal{N}$

$$\mathcal{D}_{\mathcal{G}}(i) := \{j \in \mathcal{N} | g_{ij} = 1\}$$

Notice that the out-neighbourhood of every agent is a non-empty set, since $i \in \mathcal{D}_{\mathcal{G}}(i)$ for every $i \in \mathcal{N}$.

The following terms are common in the networks literature, and will be used throughout our analysis.

Definition 1: Some Networks Terminology

- A *directed walk* in a network $\mathcal{G}$ is a sequence of (potentially repeated) nodes that are sequentially connected via directed links. Formally, it is a sequence of nodes $\langle i_1, i_2, \ldots, i_{H-1}, i_H \rangle$ in $\mathcal{G}$ such that $g_{i_h i_{h+1}} = 1$ for all $h \in \{1, 2, \ldots, H-1\}$.
- A *directed path* from node $i_1 \in \mathcal{N}$ to another node $i_H \in \mathcal{N}$ in a network $\mathcal{G}$ is a directed walk consisting of *distinct* nodes. Formally, it is a sequence of nodes $\langle i_1, i_2, \ldots, i_{H-1}, i_H \rangle$ in $\mathcal{G}$, with $i_k \neq i_l$ for $k \neq l$, $k, l \in \{1, 2, \ldots, H\}$, such that $g_{i_h i_{h+1}} = 1$ for all $h \in \{1, 2, \ldots, H-1\}$.
- A *simple cycle* of length K in a network $\mathcal{G}$ is a closed walk consisting of H distinct nodes. Formally, it is a sequence of nodes $\langle i_1, i_2, \ldots, i_{H-1}, i_H \rangle$ in $\mathcal{G}$ such that $g_{i_h i_{h+1}} = 1$ for $h \in \{1, 2, \ldots, H\}$, with $i_1 = i_H$ and $i_k \neq i_l$ for all other $k, l \in \{1, 2, \ldots, H\}$ with $k \neq l$.
- A network $\mathcal{G}$ is said to be *strongly connected* if there exists a directed path from any node to any other node in $\mathcal{G}$
- The *period* of a strongly connected network $\mathcal{G}$ is defined as the greatest common divisor of the lengths of its simple cycles.

[1] In the terminology introduced above, this implies that each agent is an out-neighbour of himself.

- A strongly connected network is called *aperiodic* if its period is equal to 1, otherwise it is called *periodic*.

2.3 The eigenvector centrality measure

Eigenvector centrality was first proposed by Bonacich (1972) as a measure of influence, prestige, or popularity in a network. It captures the idea that what makes an agent important in a network is how well-connected he is to other important agents. More specifically, each agent's eigenvector centrality is a weighted average of the eigenvector centralities of his neighbours. A rigorous definition is provided below.

DEFINITION 2: EIGENVECTOR CENTRALITY

Consider a strongly connected network $\mathcal{G} = \langle \mathcal{N}, \mathcal{E} \rangle$ with adjacency matrix $\mathbf{G}$. The *eigenvector centrality profile* of network $\mathcal{G}$ is defined as the positive left eigenvector of $\mathbf{G}$, that is, as a vector $\mathbf{c} := [c_i]_{i \in \mathcal{N}} > \mathbf{0}$ satisfying

$$\mathbf{c}^\top \mathbf{G} = \rho_G \mathbf{c}^\top \tag{1}$$

normalised so that

$$|| \mathbf{c} ||_1 := \sum_{i=1}^{n} |c_i| = 1 \tag{2}$$

where ρ_G is the spectral radius of adjacency matrix $\mathbf{G}$, and $|| \cdot ||_1$ denotes the vector 1-norm. The *eigenvector centrality* of agent $i \in \mathcal{N}$ is given by the element $c_i \in [0, 1]$.

Eigenvector centrality could be potentially problematic as a measure of influence since it is self-referential, and as such, it may not be always well-defined.[2] Nevertheless, the assumption of strong connectedness is sufficient to guarantee there is one and only one positive eigenvector associated with matrix $\mathbf{G}$ (see Section B.1 in the Appendix for a proof).[3] Bonacich and Lloyd (2001) and Jackson (2008) provide a motivation for the eigenvector centrality measure, and propose alternative measures that can be used in the cases that the former is not well-defined. An

[2] To see this, notice that if $\rho_G^{-1}\mathbf{G}$ is interpreted as a linear mapping, then $\mathbf{c}$ can be seen as a fixed point. There are mappings with no real (non-zero) fixed points, mappings with real but non-positive fixed points, and mappings with more than one non-negative fixed points. Any of the above would be problematic as a measure of centrality.

[3] Recall that eigenvectors are unique up to the relative magnitude of their entries. Since $\mathbf{c}$ is an eigenvector of matrix $\mathbf{G}$, any positive multiple of $\mathbf{c}$ is also an eigenvector of $\mathbf{G}$, and contains exactly the same information about $\mathcal{G}$ as $\mathbf{c}$ does; hence it could be also considered a vector of eigenvector centralities. Normalisation (2) serves only to uniquely pin down agents' centralities, and to facilitate the definition of some measures introduced in the sections that follow. Normalising eigenvector $\mathbf{c}$ with respect to its 2-norm, so that $|| \mathbf{c} ||_2 := \sqrt{\sum_{i=1}^{n} |c_i|^2} = 1$, is also quite common. The results though are not affected by that particular choice.

algorithm based on a variant of eigenvector centrality that does not presuppose strong connectedness, known as *PageRank*, was used in the first versions of Google search engine to determine the order of appearance of the search results (see Page, Brin, Motwani, and Winograd, 1999, especially sections 2.4, 2.5 and 6).

2.4 Beliefs

The first part of the present paper studies the evolution of the beliefs of the agents in a network through communication. The term *belief* shall be used to refer to an agent's opinion or accumulated information at a period of reference rather than to some (Bayesian) posterior. This constitutes an abuse of terminology, since what will be referred to as "belief" in this paper is technically a *statistic* for the accumulated information. This term, however, has been extensively used, and has become standard in the average-based updating literature in the last decade (see, for example, DeMarzo, Vayanos, and Zwiebel, 2003; Jackson, 2008; Golub and Jackson, 2010 and 2012; and Acemoğlu, Como, Fagnani, and Ozdaglar, 2013). Thus for conformity reasons, and in order not to cause confusion, the [ab]use of the term *belief* is maintained in the present work as well.

The choice to follow the path set by the existing literature, and not to involve priors in the analysis, should not be taken as a direct or indirect statement that priors are unimportant or that they should not be a part of a statistical updating process. This is done because the purpose of the present work is to study how information is transmitted and accounted for through a network, rather how this information is incorporated into the existing, prior beliefs of the agents. These initial beliefs may have been formed based on past observations, information from other sources, accumulated knowledge or experience, or even be arbitrary. Furthermore, nothing prevents agents in an average-based updating model from using the information eventually accumulated through this process to update their existing priors in some boundedly Bayesian or semi-Bayesian way.

The reader may have noticed that no rigorous definition of beliefs has been given so far, nor has it been specified how they are represented mathematically. Depending on the context, beliefs can be represented as a percentage, a value (expressed as a number), a set of values (expressed as a vector) or even as more general objects, such as functions or probability distributions. In fact, the model discussed here can admit as beliefs any objects that are members of some convex subset $\mathcal{B}$ of a linear space.[4]

[4] A constructive approach can be helpful in understanding why this requirement is sufficient for the process to be well-defined. First, the space of beliefs $\mathcal{B}$ must have a basic structure, in order for the process of aggregation of beliefs to be both well-defined from a technical point of view, and conceptually meaningful. This structure can be imposed by assuming that this space is abelian under addition. Yet, since the updating process to be discussed below will be based on the averaging of beliefs, with the weights being real numbers, the operation of scalar multiplication needs to be defined; hence the linear space. The convexity assumption simply guarantees that the object resulting from the

Golub and Jackson (2010) examine the relationship between the influence of individuals or small groups in large societies and the efficiency of the learning process. Since beliefs *per se* are not directly the focus of the paper, they assume for simplicity that beliefs express the subjective probability of an event occurring, and can be represented by a number in the unit interval, $\mathcal{B} = [0, 1]$. For the purpose of the paper by DeMarzo, Vayanos, and Zwiebel (2003), who study the phenomenon of unidimensional opinions, it makes better sense for agent's beliefs to be expressed as vectors whose elements represent their view or opinion on a series of m issues; in that case $\mathcal{B} = \mathbb{R}^m$ for some $m \in \mathbb{N}^*$. In the seminal work by DeGroot (1974) the objects being updated are subjective probability distributions for the value of some parameter of interest, and hence $\mathcal{B}$ is a space of probability distributions. It becomes hence apparent then that the type of updating process discussed in this paper is quite versatile, and can be adapted to various setups. Note also that the findings of the above papers (consensus, speed of convergence, wisdom of the crowds) do not depend on the representation of the beliefs, apart of course from those that directly concern the structure of the beliefs *per se*.

2.5 Precisions

It will be also assumed that agents assign a *precision* to their beliefs, expressed as a non-negative number. This precision can be based on some arbitrary assessment of the agent, or it could be an objectively defined statistic. Consider, for example, the case in which the initial belief of each agent i, to be denoted with $b_i(0)$, is equal to some noisy signal $s_i \in \mathcal{B}$ they receive about the true value of the parameter in question. Then the precision associated with it, $\pi_i(0)$, can be a sufficient statistic for the variance of agent i's signal-generating distribution.

EXAMPLE 1.
Assume that the agents are interested in estimating as accurately as possible the value v^* of some unknown parameter. This could be, among other, the future price of an asset, the net worth of some company, the return to pursuing a university degree or the quality of a product. Although they cannot directly observe v^*, each agent i receives a noisy signal $s_i = v^* + \varepsilon_i$, where $\varepsilon_i \sim N\left(0, \sigma_i^2\right)$ and $\varepsilon_i \perp \varepsilon_j$ for $i \neq j$. It follows that agents' signals are independent but not necessarily identically distributed. Smaller variances reflect that some agents may be specialists in the topic in question, or that they may have access to better or more precise information than others do.

updating process will still be a belief.

3 The model

Agents start with some initial beliefs, $\mathbf{b}(0)$, which they update through communication with their neighbours. In each round, agents ask their neighbours for their beliefs, as well as an assessment of how precise or accurate these beliefs are. Then they update their belief by weighting the reports they get based on the precision assessment reported. The belief of agent i after t rounds of communication, where $t \in \{0, 1, 2, \ldots\}$, will be denoted with $b_i(t) \in \mathcal{B}$. As discussed above, the space of beliefs $\mathcal{B}$ is assumed to be a convex set in some linear (vector) space. The beliefs of all agents in $\mathcal{G}$ after t rounds of communication can be expressed as a vector $\mathbf{b}(t) \in \mathcal{B}^n$, which will be called the *belief profile* in period t. Vector $\mathbf{b}(0)$ shall thus denote the agents' beliefs before the first round of the communication process.

Similarly to beliefs, the initial precisions of all agents can be stacked into a vector $\boldsymbol{\pi}^{\mathbf{s}} \in \mathbb{R}^n_+$. It shall be assumed that $\pi_i^s > 0$ for at least one agent $i \in \mathcal{N}$, in order for the communication process described below to be meaningful.

EXAMPLE 1.
In the example introduced above, initial beliefs can be assumed to be equal to the observed signal values, $b_i(0) := s_i$, and the corresponding initial precisions can de defined as the inverses of the variances of the signals received by each agent, $\pi_i^s := \frac{1}{\sigma_i^2}$.

3.1 The canonical average-based updating process

It is useful to begin by presenting the canonical average-based updating process due to DeGroot (1974). As described above, agents start with some initial beliefs which they update by consulting with their out-neighbours. Before the communication process begins, each agent $i \in \mathcal{N}$ assigns a *weight* $\overline{\gamma}_{ij} \in [0, 1]$ to each out-neighbour $j \in \mathcal{D}_{\mathcal{G}}(i)$ so that $\sum_{j=1}^{n} \overline{\gamma}_{ij} = 1$. These weights reflect the relative importance or trustworthiness of the beliefs of each neighbour. They may be derived from an objective formula or simply be some -potentially subjective- assessment of informational value contained in each belief. In DeMarzo, Vayanos, and Zwiebel (2003) the weight $\overline{\gamma}_{ij}$ is referred to as the *direct influence* of agent j on agent i. If $j \notin \mathcal{D}_{\mathcal{G}}(i)$, meaning that agent i cannot observe agent j, the corresponding weight is set equal to zero: $\overline{\gamma}_{ij} := 0$.

Of particular interest is the case in which the weights $\overline{\gamma}_{ij}$ that agents assign to their out-neighbours are objectively correct, in the sense that they are equal to relative precisions of their neighbours' signal-generating distributions. Although the implications of the model are not qualitatively affected, this assumption will be maintained throughout our analysis for reasons of convenience.[5] In the context of Example 1 weights $\overline{\gamma}_{ij}$ can be calculated based on precisions as follows

$$\overline{\gamma}_{ij} = \frac{g_{ij}\pi_j^s}{\sum_{k=1}^{n} g_{ik}\pi_k^s}$$

In that case, agent i's belief after the first round of communication, $b_i(1)$, will be a sufficient statistic for v^* given the neighbours' beliefs the agent has access to.

Once the weights have been set, the communication process begins. In every period $t \in \{1, 2, \ldots\}$ each agent i observes the beliefs of his out-neigbhours $j \in \mathcal{D}_{\mathcal{G}}(i)$, and revises his beliefs accordingly. In particular, the new belief of agent i will be a weighted average of the period $t-1$ beliefs reported by his out-neighbours

$$b_i(t) = \sum_{j=1}^{n} \overline{\gamma}_{ij} b_j(t-1)$$

or, using matrix notation

$$\mathbf{b}(t) = \overline{\mathbf{\Gamma}}\,\mathbf{b}(t-1). \tag{3}$$

where $\overline{\mathbf{\Gamma}} := [\overline{\gamma}_{ij}]_{(i,j)\in\mathcal{N}^2}$ is the matrix of weights. It follows then directly from its definition that $\mathbf{\Gamma}(t)$ is a row stochastic matrix[6]

By iterating on process (3) we can express the belief profile in period t as a function of the initial beliefs

$$\mathbf{b}(t) = \overline{\mathbf{\Gamma}}^t\,\mathbf{b}(0).$$

It follows that the *cumulative weight* assigned by agent i to agent j following t rounds of communication, will be given by the (i,j)-th element of matrix $\overline{\mathbf{\Gamma}}^t$, denoted with $\overline{\gamma}_{ij}(t)$. The limiting belief profile can be calculated then as a function of the matrix of weights, $\overline{\mathbf{\Gamma}}$, and the initial belief profile, $\mathbf{b}(0)$

$$\lim_{t\to+\infty} \mathbf{b}(t) = \lim_{t\to+\infty} \overline{\mathbf{\Gamma}}^t\,\mathbf{b}(0) \tag{4}$$

The limiting beliefs of agent i will be therefore given by

$$\lim_{t\to+\infty} b_i(t) = \sum_{j=1}^{n} \lim_{t\to+\infty} \overline{\gamma}_{ij}(t)\, b_j(0) \tag{5}$$

---

[5] The canonical average-based updating process presented in this section can admit any generic weights $\overline{\gamma}_{ij} \in [0, 1]$ and its results do not hinge on this assumption. This will however facilitate the presentation of the dynamic average-based updating process introduced in the section that follows.

[6] A non-negative square matrix $\mathbf{A} \in \mathbb{R}_+^{n\times n}$ is said to be *row stochastic* if the elements of each of its rows sum up to 1, that is, if $\mathbf{A}\mathbb{1}_n = \mathbb{1}_n$, where $\mathbb{1}_n$ is an n-dimensional vector of ones. This is why such matrices are often called *right stochastic*. Similarly, a non-negative square matrix $\mathbf{A}$ is called *column stochastic* or *left stochastic* if its columns sum up to 1, that is, if $\mathbb{1}_n^\top \mathbf{A} = \mathbb{1}_n^\top$.

A version of the main result in DeGroot (1974), adapted to the context of the present paper, is given below.

Proposition 1: Reaching a Consensus (DeGroot 1974)

Assume that $\mathcal{G}$ is strongly connected, and agents follow the average-based updating process described by expression (3). Then, for all $i, j \in \mathcal{N}$, the cumulative weight assigned by agent i to agent j in the limit is given by

$$\overline{\gamma}_j^{(\infty)} = \lim_{t \to +\infty} \overline{\gamma}_{ij}(t) \tag{6}$$

where $\overline{\gamma}_j^{(\infty)}$ is the j-th element of the left eigenvector $\overline{\boldsymbol{\gamma}}^{(\infty)}$ of matrix $\overline{\boldsymbol{\Gamma}}$. In DeMarzo, Vayanos, and Zwiebel (2003), $\overline{\gamma}_j^{(\infty)}$ is referred to as the *social influence* of agent j.

The above proposition readily gives rise to three remarks. First, if the stipulated conditions are met, in the limit each agent i will have the same influence on *all* other agents in the network, and hence the term *social influence* of agent i. Indeed, as expression (6) suggests

$$\lim_{t \to \infty} \overline{\gamma}_{ij}(t) = \lim_{t \to \infty} \overline{\gamma}_{ih}(t) \qquad \text{for all } i, j, h \in \mathcal{N}$$

Notice though that, in general, different agents will have different social influences, $\overline{\gamma}_i^{(\infty)} \neq \overline{\gamma}_j^{(\infty)}$ for $i \neq j$. In terms of the matrix of weights

$$\lim_{t \to +\infty} \overline{\boldsymbol{\Gamma}}^t = \mathbb{1}_n \left(\overline{\boldsymbol{\gamma}}^{(\infty)}\right)^\top$$

that is, all rows of the matrix of weights will the identical in the limit.

Second, it follows from expressions (5) and (6) that all agents will have the same limiting beliefs, or as it shall be said hence forth, a consensus will be reached.[7] In particular, for any $i \in \mathcal{N}$ it will hold

$$\lim_{t \to +\infty} b_i(t) = \sum_{j=1}^{n} \overline{\gamma}_j^{(\infty)} b_j(0)$$

Third, from a technical (although not a conceptual) point of view, the vector of social influences $\overline{\boldsymbol{\gamma}}^{(\infty)}$ can be seen as the stationary distribution of a homogeneous and aperiodic Markov chain with transition matrix $\overline{\boldsymbol{\Gamma}}$.[8]

---

[7] For a formal definition of *consensus*, see Definition 4.

[8] This a standard result about Markov chains, and it shall not be analysed further since it is out of of the scope of this paper. For a more in-depth discussion, see, for example, Section 5 in DeGroot (1974) and the references therein. Moreover, the technicalities of the dynamic average-based updating process discussed in detail in Section 4 are very similar to those for the canonical DeGroot process.

3.2 The dynamic average-based updating process

3.2.1 Motivation and preliminaries

In some cases though it would be more reasonable to assume that the weights agents' assign to their their neighbours are not constant but rather change based on how reliable the second-hand information the latter have access to is. Consider, for example, a person who has to decide whether to accept or turn down a job offer, and asks the opinions of his friends and colleagues. It may the case that one of them used to work in the past for the firm making the offer, and she has thus some partial but perhaps outdated information about it. Yet she may still be in contact with her former colleagues at the firm, whom she may contact. It would make therefore sense for the prospective employee to place some rather moderate weight on the opinion of his friend, but increase after the latter has consulted with her more informed contacts. Similar arguments could apply, among other, in the case of a prospective buyer of a house in a neighbourhood he has never lived in, a student who has to decide whether to pursue post-graduate education or enter the market, and a first-time traveller to a holiday destination.

More generally, assume that an agent $i \in \mathcal{N}$ has to decide how to weight each of his out-neighbours $j \in \mathcal{D}_{\mathcal{G}}(i)$. It could be the case that one of them, agent j, is less well-informed relatively to the other out-neighbours of agent i, but she is able to observe a third agent, k, who is an expert in the issue in question, and whom agent i cannot directly observe. In that case it would make sense for agent i to assign a small weight to j in the first round of communication, but a larger one in the subsequent round, since by then agent j's belief will have incorporated that of the better-informed agent k. Analysing this problem using the canonical average-based updating process discussed above is not possible since matrix $\overline{\mathbf{\Gamma}}$ has been assumed to be fixed.

Another implication of the fixed-weights assumption is that agents with no reliable first-hand information (that is, zero precision) will be completely ignored, and thus will play no role in shaping public opinion. It would be reasonable though to consider that such agents can have a significant, albeit indirect, influence by simply spreading information they acquired from their out-neighbours, especially if they occupy a prominent position in the network.

The aforementioned issues could in principle be addressed within the canonical average-based updating model, for example by letting agents weight their neighbours based on the precision of the information the latter are expected to receive in future periods. Such an approach though would be highly problematic. Not only would the distortions due to inappropriate weighting of information be larger, but also the weights used would have to be quite arbitrary; using some "correct" or "objective" weights would require advance knowledge of the information precision of ones's neighbours, and that of the neighbours of their neighbours, and so on, which would effectively translate into a requirement for full knowledge of the network structure. The updating process to be introduced in this section, instead, addresses the above problems by allowing for weights that vary over time.

Another important question that the model proposed here can help us answer is what makes an agent influential in a network. Within the average-based updating framework used in our analysis there can be two sources of influence: network position and information precision. It is not however straightforward how these attributes combine to determine the social influence of each agent. Under what conditions would a relatively badly informed or non-expert, yet centrally positioned agent, be more influential than an expert who is not in the spotlight? How much more influential would the former be? Up to what extent can people with a high degree of knowledge or specialisation in an area rely on their expertise to stir public opinion, disregarding network effects? Although the canonical model does not provide direct answers to the above questions, the dynamic version provides a much more suitable framework to study these issues.

In order for this to be achieved, we need to take a step backwards, and study how the weights of neighbours opinions, or the the direct influences $\overline{\gamma}_{ij}$ as called above, are pinned down in the first place. That is, we need to decompose the matrix of direct influences $\overline{\mathbf{\Gamma}}$ into a part determined only by the network structure, and a part detrained only by the distribution of information, or knowledge, among the agents. The following assumption simplifies the analysis to follow.

Assumption 1. For every agent $i \in \mathcal{N}$ there exists some agent $j \in \mathcal{D}_{\mathcal{G}}(i)$ such that $\pi_j^s > 0$.

Assumption 1 states that every agent i has at least one out-neighbour j (who could potentially be himself) who receives an informative signal ($\pi_j^s > 0$). The purpose it serves is to keep technicalities and notation at a minimum, and does not qualitatively affect our results. From a practical point of view it is not a very restrictive assumption, since in most applications it would be reasonable to assume that agents have some direct or indirect access to some information, even arbitrarily inaccurate, about the value of the unknown parameter. A sufficient, although not necessary, condition for this to hold is that every agent places positive precision to his initial belief, that is, $\boldsymbol{\pi^s} > \mathbf{0}_n$. This is satisfied if it is assumed that every agent receives some signal with finite variance about the true value of the unknown parameter. Assumption 1 will be assumed to hold for the remainder of this paper, and it will not be explicitly reiterated in the theorems and propositions to follow. Yet, for the sake of completeness of the analysis, the case in which the assumption fails is discussed in Section B.2 of the Appendix.

3.2.2 The process

With the technicalities in order, the model can be now introduced. At the beginning of each period $t \in \{1, 2, \dots\}$, every agent $i \in \mathcal{N}$ collects from each of out-neighbour $j \in \mathcal{D}_{\mathcal{G}}(i)$ a report $\big(b_j(t-1), \pi_j(t-1)\big)$ consisting of that neighbour's previous-period *belief* $b_j(t-1)$ and the corresponding *aggregate precision* $\pi_j(t-1)$. Then, based of these reports, agent i updates his belief and the precision he assigns to it. The updating process, and how these beliefs and precisions are formed, are described below.

As discussed above in this section, in the initial period, $t = 0$, agents hold some beliefs captured by $\mathbf{b}(0)$, with the corresponding precisions being $\boldsymbol{\pi}(0)$. Under Assumption 1, time-zero beliefs can be simply assumed to be agents' signals, $\mathbf{b}(0) = \boldsymbol{\pi}^{\boldsymbol{s}}$, and time-zero precisions to be the precisions of their signal-generating distributions, $\boldsymbol{\pi}(0) = \boldsymbol{\pi}^{\boldsymbol{s}}$.[9] In Example 1, that would be the inverse variances of their signal-generating distributions. In the first round of communication, they exchange information in the form $\big(b_j(0), \pi_j(0)\big)$ in the manner described above. Agent i's updated belief in period 1, $b_i(1)$, will then be a weighted average of the beliefs reported by his neighbours, with the weight $\gamma_j(1)$ assigned to each belief being equal to its relative initial precision. In the notation introduced above

$$b_i(1) = \sum_{j=0}^{n} \frac{g_{ij}\pi_j(0)}{\sum_{k=0}^{n} g_{ik}\pi_k(0)} b_j(0).$$

Up to this point the process is almost identical to the standard average-based updating process à la DeGroot presented in Section 3.1. The difference lies in the assumption that, under the current process, agents update not just their beliefs *per se*, but also the corresponding precisions. The precision $\pi_i(1)$ that agent i places to his updated belief after the first round of communication, $b_i(1)$, will be assumed to be simply the sum of the initial precisions reported by his out-neighbours, himself included

$$\pi_i(1) = \sum_{j=0}^{n} g_{ij}\pi_j(0).$$

Then the updating process is repeated. In the second round of communication, agent i inquires with his out-neighbours $j \in \mathcal{D}_{\mathcal{G}}(i)$ about their new beliefs and precisions, $\big(b_j(1), \pi_j(1)\big)$, and based on these reports he revises his belief once more. The new weights assigned to each belief reported are calculated based on his neighbours' updated precisions, $\pi_j(1)$; hence the precision

[9] If Assumption 1 does not hold, that is, if there are agents who neither them or their neighbours observe some signal with positive precision, then setting $\boldsymbol{\pi}(0) = \boldsymbol{\pi}^{\boldsymbol{s}}$ may render the process introduced in this section no longer well-defined. This should not be a concern, since the process will work for $\boldsymbol{\pi}(0) = \boldsymbol{\pi}^{\mathbf{o}}$, for some $\boldsymbol{\pi}^{\mathbf{o}} \neq \boldsymbol{\pi}^{\boldsymbol{s}}$. Intuitively, since the network has been assumed to be strongly connected, $\boldsymbol{\pi}^{\mathbf{o}}$ can be taken to be the vector of agents' precisions once some information with positive precision has reached the out-neighbourhood of every agent. A formal proof of this case is provided in the Appendix.

given to the new belief $b_i(2)$ will be the sum of these precisions.

In general, the belief of any agent $i \in \mathcal{N}$ in period $t \in \{1, 2, \ldots\}$ (that is, after t rounds of communication) will be given by

$$b_i(t) = \sum_{j=0}^{n} \gamma_{ij}(t) b_j(t-1) \tag{7}$$

where

$$\gamma_{ij}(t) := \frac{g_{ij}\pi_j(t-1)}{\sum_{k=0}^{n} g_{ik}\pi_k(t-1)} \tag{8}$$

denotes the relative weight that agent $i \in \mathcal{N}$ places on the belief reported by agent $j \in \mathcal{N}$ in the t-th round of communication. Following the terminology introduced by DeMarzo, Vayanos, and Zwiebel (2003), $\gamma_{ij}(t)$ will be referred to as the *direct influence* of agent j on agent i in period t. Unlike, however, the standard average-based updating process, it should be apparent from expression (8) that in the model introduced here the direct influences will not be constant over time, and hence the time index t. Notice of course that if agent i does not observe agent j, then $\gamma_{ij}(t) = 0$ for every $t \in \mathbb{N}$ since $g_{ij} := 0$.

Dynamic weights $\gamma_{ij}(t)$ are a result of precisions being updated every period. The *aggregate precision* attached by each agent i to his new belief in period t will be assumed to be the sum of the precisions of the beliefs reported by his neighbours that period

$$\pi_i(t) = \sum_{j=0}^{n} g_{ij}\pi_j(t-1). \tag{9}$$

Updating rule (7) can be expressed in matrix format as

$$\mathbf{b}(t) = \mathbf{\Gamma}(t)\,\mathbf{b}(t-1) \tag{10}$$

where $\mathbf{\Gamma}(t) := \left[\gamma_{ij}(t)\right]_{(i,j)\in\mathcal{N}^2}$ is the matrix of direct influences in period t.

Similarly, aggregate precisions in period t can be written as a vector as follows

$$\boldsymbol{\pi}(t) = \mathbf{G}\,\boldsymbol{\pi}(t-1)$$

or, as a function of the initial precisions $\boldsymbol{\pi}(0)$,

$$\boldsymbol{\pi}(t) = \mathbf{G}^t\boldsymbol{\pi}(0).$$

as a vector-valued function $\boldsymbol{\pi}(t)$, with $\boldsymbol{\pi} : \mathbb{N} \to \mathbb{R}^n_+$.

Now a formal definition of the process can be given.

DEFINITION 3: THE DYNAMIC AVERAGE-BASED UPDATING PROCESS

Agents will be said to follow the *dynamic average-based updating process* if their updated beliefs after each round of communication equal a weighted average of the beliefs reported by their out-neighbours (including themselves), where the weight assigned to each belief is equal to the relative aggregate precision associated with it. Using mathematical notation, agents' beliefs in period $t \in \{1, 2, \ldots\}$ (that is, after t rounds of communication) can be written as

$$\mathbf{b}(t) = \mathbf{\Gamma}(t)\,\mathbf{b}(t-1) \tag{11}$$

where the matrix of agents' direct influences is given by

$$\mathbf{\Gamma}(t) = \left[(\mathbf{G}^t \boldsymbol{\pi}(0)\mathbb{1}_n) \circ \mathbb{I}\right]^{-1}\left[\mathbf{G} \circ \mathbb{1}_n(\mathbf{G}^{t-1}\boldsymbol{\pi}(0))^\top\right] \tag{12}$$

and $\mathbf{A} \circ \mathbf{B}$ denotes the Hadamard product operation between matrices $\mathbf{A}$ and $\mathbf{B}$.[10]

It would be useful at this point to express the belief profile $\mathbf{b}(t)$ in any period $t \in \{1, 2, \ldots\}$ as a function of the initial belief profile $\mathbf{b}(0)$. By recursive backwards substitutions, expression (11) can be written as

$$\mathbf{b}(t) = \mathbf{\Gamma}(t)\,\mathbf{\Gamma}(t-1)\,\cdots\,\mathbf{\Gamma}(1)\,\mathbf{b}(0) \tag{13a}$$

or, in more compact form[11]

$$\mathbf{b}(t) = \prod_{\tau=1}^{t} \mathbf{\Gamma}(\tau)\,\mathbf{b}(0) \tag{13b}$$

The *cumulative influence*, or simply *influence* $w_{ij}(t)$ of agent j on agent i after t rounds of commu-

[10] The Hadamard product of two conformable matrices is the matrix of elementwise products of their respective elements. For a formal definition as well as some properties that will be used below, see Section A.1 in the Appendix. Expression (12) is motivated and explained in Section A.4.

[11] Since matrix multiplication is generally non-commutative, the order of multiplications in the product prescribed by the product operator (∏) is not uniquely defined. In this paper, however, it shall be used to denote the so-called *backwards matrix product*, that is

$$\prod_{\tau=1}^{t} \mathbf{X}(\tau) := \mathbf{X}(t)\,\mathbf{X}(t\text{-}1)\,\cdots\,\mathbf{X}(2)\,\mathbf{X}(1)$$

for some $\mathbf{X} \in \mathbb{R}_+^{n\times n}$. A brief discussion of these products, as well as some additional references, are provided in Section 3.3.

nication is defined as the (i,j)-th element of matrix $\mathbf{W}(t)$

$$w_{ij}(t) := \left[\mathbf{W}(t)\right]_{ij}$$

where

$$\mathbf{W}(t) := \prod_{\tau=1}^{t} \mathbf{\Gamma}(\tau). \tag{14}$$

Hence the belief updating process given by expressions (13a, 13b) can be written as

$$\mathbf{b}(t) = \mathbf{W}(t)\,\mathbf{b}(0) \tag{15}$$

EXAMPLE 2.
Consider an individual who wishes to assess the probability of an event that could affect his investment decisions, for example, a possible break-up of the Euro Area. Assume that the true probability is p^* is unknown.

Expression (12) simply suggests the weight an agent assigns to the belief of each neighbour every period equals the aggregate precision reported by that neighbour, normalised by the sum of the precision reported by all his neighbours that period.

3.2.3 Timing

At this point it would be useful to summarize the timing of the dynamic average-based updating process. Each agent $i \in \mathcal{N}$ starts with initial belief $b_i(0)$ to which he assigns precision $\pi_i(0)$. The belief profile of the agents at the beginning of each period $t \in \{1, 2, \ldots\}$ is denoted with $\mathbf{b}(t\text{-}1)$, and the corresponding precisions with $\boldsymbol{\pi}(t\text{-}1)$. The timing of the updating process that takes place every period t is the following:

[1] The t-th round of communication takes place. Agent i collects from each his out-neighbours j a report of their previous period beliefs and precision, that is, a pair $\left(b_j(t\text{-}1),\ \pi_j(t\text{-}1)\right) \in \mathcal{B} \times \mathbb{R}_+$.

[2] Agent i adjusts the weight he assigns to each neighbour j (i.e. the direct influence of agent j on agent i) to $\gamma_{ij}(t)$ as given by expression (8)

[3] Agent i updates his belief $b_j(t)$ as stipulated by (7): His new belief is the weighted average of the beliefs $b_j(t\text{-}1)$ reported by his out-neighbours, using the new weights $\gamma_{ij}(t)$ calculated in stage [2] above.

[4] Agent i calculates the precision of his updated belief as shown in expression (9). The new precision is simply the sum of precisions reported to him by his neighbours (including himself) in stage [1] above.

Hence $\mathbf{b}(t)$ and $\boldsymbol{\pi}(t)$ will respectively denote the belief and precision profiles after the period-t updating process has been completed (or, more compactly, belief and precision profiles *in* period t).

3.3 A note on backward matrix products

Note that although the updating rule stipulated by expressions (13) may be reminiscent of an inhomogeneous, or as it is sometimes called, a non-stationary Markov chain, in fact it is a different process. First, from a conceptual point of view, the process described in this paper is very different from an inhomogeneous Markov chain. Observe that, unlike a Markov chain, the dynamic average-based updating process is entirely deterministic. Moreover, the elements of the matrix of direct influences. $\boldsymbol{\Gamma}(t)$. represent weights, and not transition probabilities, as the elements of a Markov matrix $\mathbf{M}(t)$ do. Consequently, the object being updated is a vector of beliefs $\mathbf{b}(t)$, not a probability distribution $\mathbf{p}(t)$ as in the case of a Markov chain. Hence, although $\mathbf{p}(t)$ is by definition a stochastic vector, this will not be true in general for belief profile $\mathbf{b}(t)$.

Second, from a technical perspective, recall that the dynamics of an inhomogeneous Markov chain are captured by what is often referred to as a *forward product* of stochastic matrices, that is, a product of the form $\mathbf{M}(1)\,\mathbf{M}(2)\,\cdots\,\mathbf{M}(t)$. The distribution in period t will be thus given by

$$\mathbf{p}(t) = \mathbf{p}(0)\,\mathbf{M}(1)\,\mathbf{M}(2)\,\cdots\,\mathbf{M}(t).$$

Recall, however, from expression (13a), that the dynamic average-based updating process is described by a *backwards product*

$$\mathbf{b}(t) = \boldsymbol{\Gamma}(t)\,\boldsymbol{\Gamma}(t\text{-}1)\,\cdots\,\boldsymbol{\Gamma}(1)\,\mathbf{b}(0)$$

As non-commutativity of matrix multiplication would suggest, these two processes are different both in dynamics and in asymptotics. In general it will be $\mathbf{b}(t) \neq \mathbf{p}(t)$ for $t \in \{1, 2, \ldots\}$, even if $\mathbf{b}(0) = \mathbf{p}(0)$ and $\boldsymbol{\Gamma}(t) := \mathbf{M}(t)$. Backwards products, moreover, will be more depend more heavily on the first matrix in the product, $\boldsymbol{\Gamma}(1)$, than forward products are on $\mathbf{M}(1)$.[12]

[12] Unfortunately, in contrast to the quite rich literature on forward products, there is a dearth of studies on backwards products, perhaps due their more limited applications (namely studying aspects of Markov Decision Processes, and distributed algorithms, aside from DeGroot-type updating). Chatterjee and Seneta (1977) and Leizarowitz (1992) provide some sufficient conditions for convergence; Anthonisse and Tijms (1977) and Federgruen (1981) study the rate of convergence of such sequences. The author is not aware of any work providing an explicit formula for the limit of convergent sequences of backward products, analogous to those we have for forward products (see, for example, Isaacson and Madsen, 1976, Theorem V.4.7). The proofs in the present paper are based on results derived in the

Technically, the canonical average-based updating process a' la DeGroot is also described by a backwards product (and should be thought of as such). Since, however, the matrix of direct influences is constant over time, $\mathbf{\Gamma}(t) := \overline{\mathbf{\Gamma}}$, and every square matrix commutes with itself, the results from (homogeneous) Markov chain theory can be used in that case.

4 Information exchange dynamics and convergence of beliefs

4.1 Reaching a consensus in the dynamic model

This section studies the conditions under which a common belief arises in the network. Although the analysis is asymptotic, it may still be a very good approximation of the finite belief and influences dynamics, especially in the cases where the speed of convergence is high (see Section 5 for a discussion of the speed of convergence).

DEFINITION 4: CONSENSUS

It will be said that the agents in a network $\mathcal{G}$ reach a *consensus* if for any initial belief profile $\mathbf{b}(0) \in \mathcal{B}^n$, and any vector of initial precisions $\boldsymbol{\pi}(0) \in \mathbb{R}^n_+$, it holds

$$\lim_{t \to +\infty} \big(b_i(t) - b_j(t)\big) = 0 \qquad \text{for all } (i,\ j) \in \mathcal{N}^2. \tag{16}$$

If moreover there exists some belief $b^{(\infty)} \in \mathcal{B}$ such that

$$\lim_{t \to +\infty} b_i(t) = b^{(\infty)} \qquad \text{for all } i \in \mathcal{N} \tag{17}$$

the consensus shall be called *definitive*, and $b^{(\infty)}$ will be referred to as the *consensus belief*. Otherwise, the consensus will be called *oscillatory*.

As expression (16) suggests, a consensus is reached if after any -potentially arbitrarily large- number of communication rounds, all agents end up holding the same beliefs with each other. Notice that this does not automatically imply that these beliefs will be constant over time; it could be the case that all agents change their beliefs synchronously every period, or more accurately, keep oscillating indefinitely among a number of different beliefs. Interestingly enough though, it turns out that the dynamic average-based updating process discussed here cannot lead to oscillatory consensuses. The following result is an immediate application of Theorem 1 in Chatterjee and

aforementioned papers, as well as on certain results from the Markov chains literature that do not depend on the direction of the matrix product.

Seneta (1977).

PROPOSITION 2: STABLE BELIEFS IN THE LIMIT

Consider a strongly connected network $\mathcal{G}$, and suppose that agents $\mathcal{N}$ reach a consensus by following the dynamic average-based updating process stipulated in Definition 3. Then this consensus must be definitive.

Hence, if beliefs end up being identical across agents, they will also be constant over time. For the remainder of the paper, the qualifier *definitive* will be omitted; since there can be no oscillatory consensuses in the current setup, it should be clear that the term *consensus* will refer to definitive consensuses.

The result below follows directly from Theorem 3 in Chatterjee and Seneta (1977).[13]

PROPOSITION 3: CONVERGENCE

Consider a strongly connected network $\mathcal{G} = \langle \mathcal{N}, \mathcal{E} \rangle$, and assume that agents follow the dynamic average-based updating process. A consensus is reached if and only if there exists a unique vector $\mathbf{w}^{(\infty)} := \left[w_1^{(\infty)}, w_2^{(\infty)}, \ldots, w_n^{(\infty)}\right]^\top \in [0, 1]^n$ with $\|\mathbf{w}^{(\infty)}\|_1 = 1$ such that

$$\lim_{t \to +\infty} w_{ij}(t) = w_j^{(\infty)} \tag{18}$$

for all $i, j \in \mathcal{N}$. In that case $w_j^{(\infty)}$ shall be called the *social influence* of agent j, and the consensus belief $b^{(\infty)} \in \mathcal{B}$ shared by all agents will be given by

$$b^{(\infty)} = \left(\mathbf{w}^{(\infty)}\right)^\top \mathbf{b}(0) = \sum_{i=1}^{n} w_i^{(\infty)} b_i(0) \tag{19}$$

Although the aforementioned propositions in this section provide a characterisation of the consensus, they do not say anything about whether such a consensus is reached. Yet aperiodicity of network $\mathcal{G}$ together with the Corollary of Theorem 5 in Chatterjee and Seneta (1977) guarantee that a consensus will be reached. Although, as discussed above, there is no general formula for the limit of such a sequence, in this case the consensus beliefs can be computed directly using some "direction-free" results from matrix algebra.

[13] Notice that the *if* part of Proposition 3 is an immediate corollary of Proposition 2 and Definition 4. Proving the *only if* is more challenging. For a detailed proof see Seneta (1981), Chapter 4.6.

Theorem 1: The Opinion Leaders

Assume that $\mathcal{G}$ is strongly connected, and agents follow the dynamic average-based updating process stipulated in Definition 3. Then the social influence of each agent is equal to the product of his eigenvector centrality and his initial precision, adjusted by a network effects multiplier

$$\mathbf{w}^{(\infty)} = \alpha\, \mathbf{c} \circ \tilde{\boldsymbol{\pi}}(0) \tag{20}$$

where

$\mathbf{c}$ is the eigenvector centrality profile of network $\mathcal{G}$

$\tilde{\boldsymbol{\pi}}(0) := \frac{\boldsymbol{\pi}(0)}{\mathbb{1}_n \boldsymbol{\pi}(0)} = \frac{\boldsymbol{\pi}(0)}{\sum_{i=1}^{n} \pi_i^S}$ are their *relative initial precisions*, and

$\alpha = \alpha(\mathbf{G}, \boldsymbol{\pi}(0)) := \frac{\sum_{i=1}^{n} \pi_i(0)}{\sum_{i=1}^{n} c_i \pi_i(0)}$ is a scalar that captures the *network effects* or the distortion in agents' influences induced by the network.

This result is quite interesting since it shows that the social influence of each agent under limit of the dynamic average-based updating process only depends on his position in the network, as captured by his eigenvector centrality, and on how precise his information is relative to that of the other agents. It also disentangles these two effects in a clear and straightforward way.

Corollary 1. *The social influence of each agent is a function of his relative precision with respect to all other agents in the network, not just his neighbours.*

The first part of the above statement suggests that it is relative, not absolute precision that matters. This is not surprising since, as discussed above, consensus beliefs in the present setup constitute some statistic about the agents' information rather than some proper posterior. At a second stage, this statistic could be used to update some prior belief in a (boundedly) Bayesian fashion. The second part states that, in the limit, it is the agents' information and position relative to everybody in the network, not only their immediate neighbours, that determines their influence. This is an immediate consequence of the assumption of strong connectedness of the network.

The result in Theorem 1 is interesting from both a theoretical and an empirical perspective. Not only it is straightforward to see whether important agents derive their influence from their position in the network or the information they possess, but it is also easy to see how a small change in the information precision of some agent, or a rewiring of his links, would affect his social influence as well as the consensus beliefs. This could have direct implications on how some social planner could intervene in order to facilitate or disrupt the flow of information in a network.

4.2 A role for completely ignorant agents

Another interesting observation is that, under the dynamic average-based updating process, even agents without any credible initial information can affect consensus beliefs. This is because although their initial precision may well be zero, they can affect the beliefs of their neighbours by passing on second-hand information. Hence, although their own social influence will be zero, they will affect the social influences of the other agents, possibly asymmetrically, through their effect on the profile of eigenvector centralities **c**. This is not the case in the canonical average based updating process followed in the literature, since agents with no information would optimally receive zero weights from their neighbours; they could therefore be ignored, without this affecting the consensus beliefs or the social influence of the other agents.

4.3 Information allocation and efficiency of the learning process

Notice that (20) can be rewritten in the form

$$\mathbf{w}^{(\infty)} = \widehat{\mathbf{c}} \circ \boldsymbol{\pi}(0)$$

where $\widehat{\mathbf{c}} := \frac{1}{\sum_{i=1}^{n} c_i \pi_i(0)} \mathbf{c}$ is an alternative normalisation of the vector of eigenvector centralities. Nevertheless, the formulation in (20) is preferable for two reasons. First, since both precision and centrality measures are normalised to sum up to 1, it is much clearer to see how they interact and how much each of these contributes to the agents' social influence. The breakdown of each agent's social influence into a position-driven and an information driven part is discussed In the next section.

Second, the scalar $\alpha(\mathbf{G}, \boldsymbol{\pi}(0))$ has a nice intuitive interpretation: it is a coarse measure of the direction and the degree that agents' social influence deviates from their initial relative precision. To see this, it would be more convenient to normalise by the number of agents n in the network.

$$\begin{aligned} \overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) &= \frac{\alpha(\mathbf{G}, \boldsymbol{\pi}(0))}{n} \\ &= \frac{\frac{1}{n} \sum_{i=1}^{n} \pi_i(0)}{\sum_{i=1}^{n} c_i \pi_i(0)} \\ &= \frac{\overline{\pi_i}(0)}{\sum_{i=1}^{n} c_i \pi_i(0)} \end{aligned} \tag{21}$$

It can be thus seen that $\overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0))$ is the ratio of the unweighted average precision $\overline{\pi_i}(0)$ of the agents in $\mathcal{G}$ over their centrality-weighted precision. Hence the larger the value of $\overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0))$, the lower is on average the initial precision of the more centrally positioned agents. In fact, it can be shown that the relationship between $\overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0))$ and the covariance $Cov(\mathbf{c}, \boldsymbol{\pi}(0))$ between eigenvector centrality and initial precision is the following:

$$\left.\begin{matrix} Cov(\mathbf{c}, \boldsymbol{\pi}(0)) > 0 \\ Cov(\mathbf{c}, \boldsymbol{\pi}(0)) = 0 \\ Cov(\mathbf{c}, \boldsymbol{\pi}(0)) < 0 \end{matrix}\right\} \Longleftrightarrow \begin{cases} \overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) < 1 \\ \overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) = 1 \\ \overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) > 1 \end{cases} \tag{22}$$

Case A $\overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) < 1$: An $\overline{\alpha}$ smaller than 1 suggests that more centrally positioned agents will on average possess more precise information. An example is a star network with the agent in the centre having higher precision than the other agents. In that sense $\overline{\alpha}$ can be a measure on the network's effect on "inequality": networks with smaller $\overline{\alpha}$'s increase the influence of agents who would anyway be influential due to their high precision.

Case B $\overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) = 1$: There is no correlation between agents' position in the network and the precision of their beliefs. This will be the case if all agents have the same centrality, such as for example, in a circle, line, or a complete network. Another extreme case that induces $\overline{\alpha} = 1$ is that in which all agents have the same initial precision irrespectively of the structure of the network. Note that $\overline{\alpha} = 1$ can arise even in cases in which agents differ from each other in both centrality and precision.

Case C $\overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) > 1$: In this case it is the less central agents who possess on average more precise information. A star network with the central agent having information with lower precision than the network average is such an example.

PROPOSITION 4: EFFICIENCY OF LEARNING

The consensus beliefs $\mathbf{b}^{(\infty)}$ under the Dynamic Average-Based Updating Process will be optimal, in the sense that they will constitute a sufficient statistic of all information available in the network, if and only if all agents are equally central. In a generic network $\mathcal{G}$ though, this process will not be optimal. Given network $\mathcal{G}$, information aggregation will be more efficient in networks with $\overline{\alpha} \geq 1$.

According to the above proposition, similarly to the canonical DeGroot model, learning under the Dynamic Average-Based Updating Process will be, in general, suboptimal. It can be shown that in the DeGroot model, the consensus beliefs are optimal if and only if the matrix of direct influences.

$\overline{\mathbf{\Gamma}}$, is balanced, that is, if and only if

$$\sum_{j=1}^{n} g_{ij}\overline{\gamma}_{jj} = 1$$

for all agents $i \in \mathcal{N}$ (see DeMarzo, Vayanos, and Zwiebel (2003), Theorem 2). Under the dynamic model introduced above, it suffices that all agents are equally centrally located in the network.

Furthermore, the present paper provides a characterisation of the networks in which the "second-best" consensus beliefs are reached. This is achieved in networks with $\overline{\alpha}(\mathbf{G}, \boldsymbol{\pi}(0)) > 1$: , that is, networks in which the less prominent agents are those who possess higher precisions. The intuition behind this result is that, since central agents' opinions get to be heard more, more peripheral agents need to be more trustworthy to, at least partially, counter this persuasion bias.

5 The speed of convergence

The results of the previous section were based on asymptotic analysis. This suggests that they will not hold for in general in finite time. In cases, however, where the speed of convergence in sufficiently high, they will be very good approximations. This section studies this more thoroughly.

section incomplete...

References

ACEMOĞLU, D., G. COMO, F. FAGNANI, AND A. OZDAGLAR (2013): "Opinion fluctuations and disagreement in social networks," *Mathematics of Operations Research*, 38(1), 1–27, 🌐.

ALGAN, Y., Q.-A. DO, A. LE CHAPELAIN, AND Y. ZENOU (2015): "How social networks shape our values? A natural experiment among future French politicians," Unpublished manuscript, Science Po, Paris.

ANTHONISSE, J. M., AND H. TIJMS (1977): "Exponential convergence of products of stochastic matrices," *Journal of Mathematical Analysis and Applications*, 59(2), 360–364, 🌐.

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

BONACICH, P., AND P. LLOYD (2001): "Eigenvector-like measures of centrality for asymmetric relations," *Social Networks*, 23(3), 191–201, 🌐.

CHATTERJEE, S., AND E. SENETA (1977): "Towards consensus: Some convergence theorems on repeated averaging," *Journal of Applied Probability*, 14(1), 89–97, 🌐.

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

DEMARZO, P. M., D. VAYANOS, AND J. ZWIEBEL (2003): "Persuasion bias, social influence, and unidimensional opinions," *The Quarterly Journal of Economics*, 118(3), 909–968, 🌐.

FEDERGRUEN, A. (1981): "The rate of convergence for backwards products of a convergent sequence of finite Markov matrices," *Stochastic Processes and their Applications*, 11(2), 187–192, 🌐.

GOLUB, B., AND M. O. JACKSON (2010): "Naïve learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, 2(1), 112–149, 🌐.

——— (2012): "How homophily affects the speed of learning and best-response dynamics," *The Quarterly Journal of Economics*, 127(3), 1287–1338, 🌐.

HORN, R. A., AND C. R. JOHNSON (1991): *Topics in matrix analysis*. Cambridge University Press, New York, NY.

ISAACSON, D. L., AND R. W. MADSEN (1976): *Markov chains: Theory and applications*. Wiley, New York, NY.

JACKSON, M. O. (2008): *Social and economic networks*. Princeton University Press, Princeton, NJ.

JACKSON, M. O., B. ROGERS, AND Y. ZENOU (2015): "The economic consequences of social network structure," Available at SSRN 🌐.

LEIZAROWITZ, A. (1992): "On infinite products of stochastic matrices," *Linear Algebra and its Applications*, 168, 189–219, 🌐.

MEYER, C. D. (2001): *Matrix analysis and applied linear algebra*. American Mathematical Society, Philadelphia, PA.

MORETTI, E. (2011): "Social learning and peer effects in consumption: Evidence from movie sales," *The Review of Economic Studies*, 78(1), 356–393, 🌐.

PAGE, L., S. BRIN, R. MOTWANI, AND T. WINOGRAD (1999): "The PageRank citation ranking: Bringing order to the Web," Discussion Paper 1999-66, Stanford InfoLab, 🌐.

PERKINS, P. (1961): "A theorem on regular matrices," *Pacific Journal of Mathematics*, 11(1), 1529–1533, 🌐.

SENETA, E. (1981): *Non-negative matrices and Markov chains*. Springer Science+Business Media, New York, NY, 2nd edn.

APPENDIX

A Mathematical appendix

A.1 The Hadamard product

This section discusses shortly the Hadamard product matrix operation and some basic results related to it that are used in the present analysis.

DEFINITION 5: THE HADAMARD PRODUCT

Consider matrices $\mathbf{A} = [a_{ij}] \in \mathbb{C}^{m\times n}$ and $\mathbf{B} = [b_{ij}] \in \mathbb{C}^{m\times n}$ with $m, n \in \mathbb{N}^*$. The *Hadamard product*[14] of $\mathbf{A}$ and $\mathbf{B}$, denoted with $\mathbf{A} \circ \mathbf{B}$, is defined as the matrix of the scalar products of their corresponding elements

$$\mathbf{A} \circ \mathbf{B} := \left[a_{ij} b_{ij}\right]_{(i,j)\in\mathcal{M}\times\mathcal{N}} \in \mathbb{C}^{m\times n}$$

where $\mathcal{M} := \{1, \ldots, m\}$ and $\mathcal{N} := \{1, \ldots, n\}$.

Let $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{C}^{m\times n}$, and consider a conformable matrix of ones, $\mathbf{J}_{m\times n} := \mathbb{1}_{\mathrm{m}}\mathbb{1}_{\mathrm{n}}^{\top}$, and a scalar $\kappa \in \mathbb{C}$. The Hadamard product possesses the following properties:

[H.1]	*Commutativity*:	$\mathbf{A} \circ \mathbf{B} = \mathbf{B} \circ \mathbf{A}$
[H.2]	*Associativity*:	$\mathbf{A} \circ (\mathbf{B} \circ \mathbf{C}) = (\mathbf{A} \circ \mathbf{B}) \circ \mathbf{C}$
[H.3]	*Distributivity*:	$\mathbf{A} \circ (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \circ \mathbf{B}) + (\mathbf{A} \circ \mathbf{C})$
[H.4]	*Identity element*:	$\mathbf{A} \circ \mathbf{J}_{m\times n} = \mathbf{A}$
[H.5]	*Distributive transposition*:	$(\mathbf{A} \circ \mathbf{B})^{\top} = \mathbf{A}^{\top} \circ \mathbf{B}^{\top}$
[H.6]	*Compatibility with scalar multiplication*:	$\kappa(\mathbf{A} \circ \mathbf{B}) = (\kappa\mathbf{A}) \circ \mathbf{B} = \mathbf{A} \circ (\kappa\mathbf{B})$

In addition, consider vectors $\mathbf{x} \in \mathbb{C}^m$, $\mathbf{y} \in \mathbb{C}^n$, and define the diagonal matrices $\mathbf{D_x} := diag(x_1, \ldots, x_m)$ and $\mathbf{D_y} := diag(y_1, \ldots, y_n)$. The following observations will be particularly useful in our analysis :

- *Pre-multiplying* matrix $\mathbf{A}$ by a conformable diagonal matrix $\mathbf{D_x}$ multiplies every element of each *row* i of $\mathbf{A}$ by the corresponding element x_i of vector $\mathbf{x}$, that is

$$\mathbf{D_x}\,\mathbf{A} = \left[a_{ij} x_i\right]_{i\in\mathcal{M}, j\in\mathcal{N}} \in \mathbb{C}^{m\times n} \tag{23}$$

---

[14] Named after French mathematician Jacques Salomon Hadamard (1865-1963). The terms *elemntwise*, *entrywise*, or *Schur product* are also encountered in the literature.

- *Post-multiplying* matrix **A** by a conformable diagonal matrix $\mathbf{D_y}$ multiplies every element of each *column* j of **A** by the corresponding element y_j of vector **y**, that is

$$\mathbf{A}\,\mathbf{D_y} = \left[a_{ij}y_j\right]_{i\in\mathcal{M},j\in\mathcal{N}} \in \mathbb{C}^{m\times n} \tag{24}$$

The above observation gives rise to the following properties

[H.7] [*Multiply row i by x_i*] $\mathbf{D_x}\,\mathbf{A} = \mathbf{A} \circ \left(\mathbf{x}\mathbb{1}_{\mathrm{m}}^{\top}\right) = \left[\left(\mathbf{x}\mathbb{1}_{\mathrm{m}}^{\top}\right) \circ \mathbb{I}_m\right]\,\mathbf{A}$

[H.8] [*Multiply colm. j by y_j*] $\mathbf{A}\,\mathbf{D_y} = \mathbf{A} \circ \left(\mathbb{1}_{\mathrm{n}}\mathbf{y}^{\top}\right) = \mathbf{A}\,\left[\left(\mathbb{1}_{\mathrm{n}}\mathbf{y}^{\top}\right) \circ \mathbb{I}_n\right]$

Finally, the results below from Horn and Johnson (1991) are useful in the proofs in part B of this appendix.

[H.9] $\mathbf{D_x}(\mathbf{A} \circ \mathbf{B})\mathbf{D_y} = (\mathbf{D_x}\,\mathbf{A}) \circ (\mathbf{B}\,\mathbf{D_y}) = (\mathbf{A}\,\mathbf{D_y}) \circ (\mathbf{D_x}\,\mathbf{B}) = \mathbf{A} \circ (\mathbf{D_x}\mathbf{B}\,\mathbf{D_y})$

[H.10] $\left[(\mathbf{A} \circ \mathbf{B})\,\mathbf{y}\mathbb{1}_{\mathrm{m}}^{\top}\right] \circ \mathbb{I}_m = \left(\mathbf{A}\,\mathbf{D_y}\mathbf{B}^{\top}\right) \circ \mathbb{I}_m$

Property [[H.9]] is Lemma 5.1.2 in Chapter 5, while property [[H.10]] follows immediately from Lemma 5.1.3 and the definition of the Hadamard product.

A.2 Non-negative matrices and networks

For the sake of convenience, some standard elements of matrix algebra are presented below.

DEFINITION 6: IRREDUCIBLE AND PRIMITIVE MATRICES

- A matrix $\mathbf{P} \in \{0, 1\}^{n\times n}$ is called a *permutation matrix* if in each row and in each column there exists exactly one entry equal to 1, with all other entries being equal to 0.
- A matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is said to be a *reducible* matrix if there exists a permutation matrix **P** such that

$$\mathbf{P}^{\top}\mathbf{A}\,\mathbf{P} = \begin{bmatrix} \mathbf{X}_{k\times k} & \mathbf{Y}_{k\times n\text{-}k} \\ \mathbb{O}_{n\text{-}k\times k} & \mathbf{Z}_{n\text{-}k\times n\text{-}k} \end{bmatrix} \tag{25}$$

where **X**, **Z** are square matrices, and $\mathbb{O}$ is a matrix of zeros.

- A square matrix is called *irreducible* if it is not reducible.

- A non-negative irreducible matrix is said to be a *primitive* matrix if only one of its eigenvalues lies on its spectral circle.

Note that in some older papers, the terms *indecomposable* and *regular* are used respectively instead of irreducible and primitive. The following lemma establishes a relationship between the properties of an adjacency matrix and the structure of the network it represents.

Lemma 1: Matrix Properties and Network Structure

- A network $\mathcal{G}$ is strongly connected if and only if its adjacency matrix $\mathbf{G}$ is irreducible.
- A strongly connected network $\mathcal{G}$ is aperiodic if and only if its adjacency matrix $\mathbf{G}$ is primitive.

Proof. To prove the first statement, it suffices to show that its contrapositive holds true. In other words, it is equivalent to proving the following:

> ***Statement [CP].*** *A network $\mathcal{G} = \langle \mathcal{N}, \mathcal{E} \rangle$ is not strongly connected if and only if its adjacency matrix* $\mathbf{G}$ *is reducible.*

First, notice that pre- and post-multiplying a square matrix respectively by a permutation matrix $\mathbf{P}$ and its transpose $\mathbf{P}^{\mathsf{T}}$ reorders the rows and the columns of matrix in the same way. Hence the transformation

$$\widetilde{\mathbf{G}} := \mathbf{P}^{\mathsf{T}} \mathbf{G} \mathbf{P} \tag{26}$$

simply relabels the agents in $\mathcal{N}$ without essentially changing the structure of the network; matrix $\widetilde{\mathbf{G}}$ represents the same network as $\mathbf{G}$, but with the agents relabelled. More formally, transformation (26) can be seen as applying a bijection $f_N : \mathcal{N} \to \mathcal{N}$ from the set of nodes to itself, and a corresponding bijection $f_E : \mathcal{N}^2 \to \mathcal{N}^2$ with $f_E(i,j) := \big(f_N(i),\, f_N(j)\big)$ from the set of edges of $\mathcal{G}$ to itself. Except for the node labels, network $\widetilde{\mathcal{G}} = \langle f_N(\mathcal{N}),\, f_E(\mathcal{E})\rangle$, represented by adjacency matrix $\widetilde{\mathbf{G}}$, will be identical to $\mathcal{G}$. An equivalent, therefore, to Statement [CP] is that $\widetilde{\mathbf{G}}$ is not strongly connected if and only if $\mathbf{G}$ is reducible. For convenience, denote the new labels of the nodes of the transformed network with l_i, so that $f_N(\mathcal{N}) = \{l_1,\, l_2,\, \ldots,\, l_n\}$.

Notice that if $\mathbf{G}$ is reducible, then by (25) matrix $\mathbf{P}$ can chosen so that

$$\widetilde{\mathbf{G}} = \begin{bmatrix} \mathbf{X}_{k \times k} & \mathbf{Y}_{k \times n\text{-}k} \\ \mathbb{O}_{n\text{-}k \times k} & \mathbf{Z}_{n\text{-}k \times n\text{-}k} \end{bmatrix}$$

It can now easily be seen from $\widetilde{\mathbf{G}}$ that there does not exist a directed path from any of the nodes in $\mathcal{N}_{\text{isol}} := \{l_{k+1}, l_{k+2}, \ldots, l_n\}$ to any of the nodes in $\mathcal{N}_{\text{main}} := \{l_1, l_2, \ldots, l_k\}$ since the former pay attention only to agents within $\mathcal{N}_{\text{isol}}$. This suggests that network $\widetilde{\mathcal{G}}$, and hence network $\mathcal{G}$, is *not* strongly connected, which proves the *if* part of Statement [CP].

To prove the *only if* part of Statement [CP], assume that network $\mathcal{G}$ is not strongly connected. Then there must exist (at least) two nodes $i, j \in \mathcal{N}$ such that there are no directed paths from node i to node j. Consider the set of nodes $\mathcal{M}(j) \subset \mathcal{N}$ consisting of those and only those nodes $h \in \mathcal{N}$ such that there is a directed path from node h to node j, and denote its cardinality with m, where $1 \leq m \leq n\text{-}1$.[15] Since there is no directed path from node i to node j, it must be that $i \in \mathcal{N} \setminus \mathcal{M}$. Now consider a transformation similar to (26) that assigns labels from l_1 to l_m to the nodes in $\mathcal{M}(j)$. This can be implemented through a bijection f_N such as the one described above, with $f_N(\mathcal{M}) = \{l_1, \ldots, l_m\}$ and $f_N(\mathcal{N} \setminus \mathcal{M}) = \{l_{m+1}, \ldots, l_n\}$, together with the corresponding bijection f_E. Then the transformed matrix can be written as

$$\widetilde{\mathbf{G}} = \mathbf{P}^{\top} \mathbf{G} \, \mathbf{P} = \begin{bmatrix} \widetilde{\mathbf{X}}_{m \times m} & \widetilde{\mathbf{Y}}_{m \times n\text{-}m} \\ \widetilde{\mathbf{W}}_{n\text{-}m \times m} & \widetilde{\mathbf{Z}}_{n\text{-}m \times n\text{-}m} \end{bmatrix}$$

where block $\widetilde{\mathbf{X}}$ captures the edges among the nodes in $\mathcal{M}$, block $\widetilde{\mathbf{Z}}$ the edges among the nodes in $\mathcal{N} \setminus \mathcal{M}$, block $\widetilde{\mathbf{Y}}$ the edges from nodes in $\mathcal{M}$ to nodes in $\mathcal{N} \setminus \mathcal{M}$, and block $\widetilde{\mathbf{W}}$ the edges from nodes in $\mathcal{N} \setminus \mathcal{M}$ to nodes in $\mathcal{M}$. Observe, however, that there should not exist any edges from nodes not in $\mathcal{M}$ towards nodes in $\mathcal{M}$. Suppose towards a contradiction that there existed such an edge, emanating from node $q \in \mathcal{N} \setminus \mathcal{M}$. This would imply that there is a directed path from node q to node j, and hence $q \in \mathcal{M}$ by the definition of $\mathcal{M}$. Since no such edges exists, it must be that $\widetilde{\mathbf{W}} = \mathbb{O}_{n\text{-}m \times m}$, suggesting that $\mathbf{G}$ is a reducible matrix. This completes the proof of the first statement in Lemma 1.

The second statement follows directly from Theorem 1 in Perkins (1961) and the definition of an aperiodic network (see Definition 1). The form presented above is the one provided by Lemma 2 in Golub and Jackson (2010). $\diamondsuit$

A.3 The Perron-Frobenius theorem

The following statement of the Perron-Frobenius theorem,[16] is based on Meyer (2001), Chapter 8.3., nd has been adapted to the context of the present paper.

[15] Notice that $\mathcal{M}(j)$ will be non-empty since $j \in \mathcal{M}(j)$.

[16] A first version of the theorem, applying to positive matrices, was proved by German mathematician Oskar Perron in 1907. Five years later his colleague Ferdinand Georg Frobenius showed that most of Perron's results carry over to non-negative matrices, provided that they are irreducible.

PROPOSITION 5: THE PERRON-FROBENIUS THEOREM

Let $\mathbf{G} \in \mathbb{R}_+^{n\times n}$ be a non-negative, irreducible square matrix, and denote its spectral radius with ρ_G. The following statements hold true.

[PF.1] There exists a simple eigenvalue of $\mathbf{G}$ equal to ρ_G.

[PF.2] There exists a positive stochastic eigenvector corresponding to ρ_G, that is, a vector $\mathbf{p} > \mathbf{0}_n$ such that $\mathbf{G}\,\mathbf{p} = \rho_G \mathbf{p}$ and $\|\mathbf{p}\|_1 = 1$. This is called the *Perron vector* of $\mathbf{G}$.

[PF.3] The Perron vector is the only non-negative eigenvector of $\mathbf{G}$ up to a positive multiple.

A.4 Motivating matrix $\Gamma(t)$

Explain why $\boldsymbol{\Gamma}(t) = \left[(\mathbf{G}^t\boldsymbol{\pi}(0)\mathbb{1}_n) \circ \mathbb{I}\right]^{-1}\left[\mathbf{G} \circ \mathbb{1}_n(\mathbf{G}^{t-1}\boldsymbol{\pi}(0))^\top\right] = diag\left(\mathbf{G}^t\boldsymbol{\pi}(0)\right)^{-1} \mathbf{G} diag\left(\mathbf{G}^{t-1}\boldsymbol{\pi}(0)\right)$ and why it describes the process that is relevant here.

B Proofs

B.1 Existence and uniqueness of eigenvector centrality

We begin by establishing that adjacency matrix $\mathbf{G}$, and hence its transpose, $\mathbf{G}^\top$, is non-negative and irreducible, and thus the Perron-Frobenius theorem applies (see Section A.3). Non-negativity holds true by definition, since $\mathbf{G} \in \{0, 1\}^{n\times n}$, while irreducibility follows from Lemma 1 and the assumption that $\mathcal{G}$ is strongly connected.

It can now be readily shown that eigenvector centrality is a well defined measure, that is, it exists and it is unique in any strongly connected network $\mathcal{G}$. To establish existence, notice that [PF.1] suggests that ρ_G will be an eigenvalue $\mathbf{G}^\top$, and hence, by [PF.2], $\mathbf{c}$ will be the Perron vector of $\mathbf{G}^\top$. It will be therefore a positive vector, and thus meaningful as a measure of centrality, since it will not contain any negative or non-real entries. Uniqueness follows from the fact that the Perron vector is the only positive eigenvector of $\mathbf{G}$ (see [PF.3]).

B.2 The case that Assumption 1 fails

Write proof formally: Strongly connected $\mathcal{G}$ implies that info will reach (at least) one new neighbourhood per period of communication. Hence at most in $n - 1$ periods we will be at the baseline case. Write formula for new precisions.

B.3 Proof of Theorem 1

$$
\begin{aligned}
\mathbf{W}(t) &= \prod_{\kappa=1}^{t} \mathbf{\Gamma}(\kappa) \\
&= \prod_{\kappa=1}^{t} \left[\left(\mathbf{G}^{\kappa}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\left[\mathbf{G}\circ\mathbb{1}_n\left(\mathbf{G}^{\kappa-1}\boldsymbol{\pi}(0)\right)^{\top}\right] \\
&= \prod_{\kappa=1}^{t} \left[\left(\mathbf{G}^{\kappa}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\mathbf{G}\left[\left(\mathbf{G}^{\kappa-1}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right] && \text{by [H.8]} \\
&= \left[\left(\mathbf{G}^{t}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\mathbf{G}\left[\left(\mathbf{G}^{t-1}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]\left[\left(\mathbf{G}^{t-1}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\mathbf{G}\left[\left(\mathbf{G}^{t-2}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]\cdots \\
&\qquad \cdots\left[\left(\mathbf{G}^{2}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\mathbf{G}\left[\left(\mathbf{G}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]\left[\left(\mathbf{G}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\mathbf{G}\left[\left(\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right] \\
&= \left[\left(\mathbf{G}^{t}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\underbrace{\mathbf{G}\ldots\mathbf{G}}_{t-1\text{ terms}}\;\mathbf{G}\left[\left(\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right] \\
&= \left[\left(\mathbf{G}^{t}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\mathbf{G}^{t}\left[\left(\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right] \\
&= \left[\left(\mathbf{G}^{t}\boldsymbol{\pi}(0)\mathbb{1}_n^{\top}\right)\circ\mathbb{I}\right]^{-1}\mathbf{G}^{t}\,\mathbf{D}_{\boldsymbol{\pi}(0)} && (27)
\end{aligned}
$$

where we have used the properties of Hadamard product discussed in Section A.1 of the Appendix, and $\mathbf{D}_{\boldsymbol{\pi}(0)} := diag(\pi_1(0), \pi_2(0), \ldots, \pi_n(0))$ is a diagonal matrix with the elements of vector $\boldsymbol{\pi}(0)$ on its main diagonal.

PROPOSITION 6

Let $\mathcal{G}$ be a strongly connected, aperiodic network with adjacency matrix $\mathbf{G}$. Denote its spectral radius by ρ_G, and its Perron vector by $\mathbf{p}$. Then

$$
\lim_{t\to+\infty}\left(\frac{\mathbf{G}}{\rho_G}\right)^{t} = \frac{\mathbf{p}\,\mathbf{c}^{\top}}{\mathbf{c}^{\top}\mathbf{p}} \tag{28}
$$

PROOF. We know that a network $\mathcal{G}$ is aperiodic, and hence strongly connected, if and only if its adjacency matrix $\mathbf{G}$ is primitive, and hence irreducible (Lemma 1). Recall that eigenvector centrality is defined as the left eigenvector of $\mathbf{G}$ (Definition 2). Then (28) follows directly from the theorem on primitive matrices and expression (8.3.10) in Meyer (2001). ♢

Now we can use Proposition 6 to obtain an expression for the social influence of the agents. From (28) it follows that

$$\begin{aligned}
\lim_{t\to+\infty} \mathbf{W}(t) &= \lim_{t\to+\infty} \left\{ \left[\left(\left(\frac{\mathbf{G}}{\rho_G} \right)^t \boldsymbol{\pi}(0)\, \mathbb{1}_n^\top \right) \circ \mathbb{I} \right]^{-1} \left(\frac{\mathbf{G}}{\rho_G} \right)^t \mathbf{D}_{\boldsymbol{\pi}(0)} \right\} \\
&= \left[\left(\lim_{t\to+\infty} \left(\frac{\mathbf{G}}{\rho_G} \right)^t \boldsymbol{\pi}(0)\, \mathbb{1}_n^\top \right) \circ \mathbb{I} \right]^{-1} \lim_{t\to+\infty} \left(\frac{\mathbf{G}}{\rho_G} \right)^t \mathbf{D}_{\boldsymbol{\pi}(0)} \\
&= \left[\left(\frac{\mathbf{p}\,\mathbf{c}^\top}{\mathbf{c}^\top \mathbf{p}} \boldsymbol{\pi}(0)\, \mathbb{1}_n^\top \right) \circ \mathbb{I} \right]^{-1} \frac{\mathbf{p}\,\mathbf{c}^\top}{\mathbf{c}^\top \mathbf{p}} \mathbf{D}_{\boldsymbol{\pi}(0)} \\
&= \left[\left(\mathbf{p}\,\mathbf{c}^\top \boldsymbol{\pi}(0)\, \mathbb{1}_n^\top \right) \circ \mathbb{I} \right]^{-1} \mathbf{p}\,\mathbf{c}^\top \mathbf{D}_{\boldsymbol{\pi}(0)} && \text{by [H.6]} \\
&= \left(\mathbf{c}^\top \boldsymbol{\pi}(0) \right)^{-1} \left[\left(\mathbf{p} \mathbb{1}_n^\top \right) \circ \mathbb{I} \right]^{-1} \mathbf{p}\,\mathbf{c}^\top \mathbf{D}_{\boldsymbol{\pi}(0)} && \text{by [H.6]} \\
&= \left(\mathbf{c}^\top \boldsymbol{\pi}(0) \right)^{-1} \mathbf{D}_{\mathbf{p}}^{-1} \mathbf{p}\,\mathbf{c}^\top \mathbf{D}_{\boldsymbol{\pi}(0)} \\
&= \left(\mathbf{c}^\top \boldsymbol{\pi}(0) \right)^{-1} \mathbb{1}_n \mathbf{c}^\top \mathbf{D}_{\boldsymbol{\pi}(0)} \\
&= \left(\mathbf{c}^\top \boldsymbol{\pi}(0) \right)^{-1} \mathbb{1}_n \left(\mathbf{c}^\top \circ \boldsymbol{\pi}(0)^\top \right) && \text{by [H.8]} \\
&= \mathbb{1}_n \left(\mathbf{c}^\top \boldsymbol{\pi}(0) \right)^{-1} \left(\mathbf{c} \circ \boldsymbol{\pi}(0) \right)^\top && \text{by [H.5]} \\
&= \mathbb{1}_n \left(\frac{\mathbb{1}_n^\top \boldsymbol{\pi}(0)}{\mathbf{c}^\top \boldsymbol{\pi}(0)} \right) \left(\mathbf{c} \circ \frac{\boldsymbol{\pi}(0)}{\mathbb{1}_n^\top \boldsymbol{\pi}(0)} \right)^\top && \text{by [H.6]} \\
&= \mathbb{1}_n\, \alpha(\mathbf{G}, \boldsymbol{\pi}(0)) \left(\mathbf{c} \circ \tilde{\boldsymbol{\pi}}(0) \right)^\top
\end{aligned}$$

where $\mathbf{D}_{\mathbf{p}} := diag(p)$, $\alpha(\mathbf{G}, \boldsymbol{\pi}(0)) := \frac{\mathbb{1}^\top \boldsymbol{\pi}(0)}{\mathbf{c}^\top \boldsymbol{\pi}(0)} = \frac{\sum_{i=1}^n \pi_i(0)}{\sum_{i=1}^n c_i \pi_i(0)}$ is a measure that captures the effects of the network on social influence, and $\tilde{\boldsymbol{\pi}}(0) := \frac{\boldsymbol{\pi}(0)}{\mathbb{1}^\top \boldsymbol{\pi}(0)}$ is the vector of relative initial precisions of the agents in network $\mathcal{G}$.