# Persuasion bias, social influence, and uni-dimensional opinions

Peter M. DeMarzo, Dimitri Vayanos and Jeffrey Zwiebel



**This is an electronic version of an Article published in the Quarterly journal of economics, 118 (3). pp. 909-968 © 2003 MIT Press.**
**http://mitpress.mit.edu/QJE**

# PERSUASION BIAS, SOCIAL INFLUENCE, AND UNIDIMENSIONAL OPINIONS

Peter M. DeMarzo
Dimitri Vayanos
Jeffrey Zwiebel*

April 2003

**Abstract**

We propose a boundedly-rational model of opinion formation in which individuals are subject to *persuasion bias*; that is, they fail to account for possible repetition in the information they receive. We show that persuasion bias implies the phenomenon of *social influence*, whereby one's influence on group opinions depends not only on accuracy, but also on how well-connected one is in the social network that determines communication. Persuasion bias also implies the phenomenon of *unidimensional opinions*; that is, individuals' opinions over a multidimensional set of issues converge to a single "left-right" spectrum. We explore the implications of our model in several natural settings, including political science and marketing, and we obtain a number of novel empirical implications.

*DeMarzo and Zwiebel: Graduate School of Business, Stanford University, Stanford CA 94305, Vayanos: MIT Sloan School of Management, 50 Memorial Drive E52-437, Cambridge MA 02142. This paper is an extensive revision of our paper, "A Model of Persuasion – With Implication for Financial Markets," (first draft, May 1997). We are grateful to Nick Barberis, Gary Becker, Jonathan Bendor, Larry Blume, Simon Board, Eddie Dekel, Stefano DellaVigna, Darrell Duffie, David Easley, Glenn Ellison, Simon Gervais, Ed Glaeser, Ken Judd, David Kreps, Edward Lazear, George Loewenstein, Lee Nelson, Anthony Neuberger, Matthew Rabin, José Scheinkman, Antoinette Schoar, Peter Sorenson, Pietro Veronesi, Richard Zeckhauser, three anonymous referees, and seminar participants at the American Finance Association Annual Meetings, Boston University, Cornell, Carnegie-Mellon, ESSEC, the European Summer Symposium in Financial Markets at Gerzensee, HEC, the Hoover Institution, Insead, MIT, the NBER Asset Pricing Conference, the Northwestern Theory Summer Workshop, NYU, the Stanford Institute for Theoretical Economics, Stanford, Texas A&M, UCLA, U.C. Berkeley, Université Libre de Bruxelles, University of Michigan, University of Texas at Austin, University of Tilburg, and the Utah Winter Finance Conference for helpful comments and discussions. All errors are our own.

# I Introduction

In this paper, we propose a model of opinion formation in which individuals are subject to *persuasion bias*, failing to adjust properly for possible repetitions of information they receive. We argue that persuasion bias provides a simple explanation for several important phenomena that are otherwise hard to rationalize, such as propaganda, censorship, marketing, and the importance of air-time. We show that persuasion bias implies two additional phenomena. First, that of *social influence*, whereby one's influence on group opinions depends not only on accuracy, but also on how well-connected one is in the social network according to which communication takes place. Second, that of *unidimensional opinions*, whereby individuals' opinions over a multidimensional set of issues can be represented by a single "left-right" spectrum.

To motivate persuasion bias, consider an individual who reads an article in a newspaper with a well-known political slant. Under full rationality, the individual should anticipate that the arguments presented in the article will reflect the newspaper's general political views. Moreover, the individual should have a prior assessment about how strong these arguments are likely to be. Upon reading the article, the individual should update his political beliefs in line with this assessment. In particular, the individual should be swayed towards the newspaper's views if the arguments presented in the article are stronger than expected, and away from them if the arguments are weaker than expected. On average, however, reading the article should have *no effect* on the individual's beliefs. Formally, under full rationality, beliefs must follow martingales.[1]

The martingale result seems, however, in contrast with casual observation. It seems, in particular, that newspapers do sway readers towards their views, even when these views are publicly known. A natural explanation of this phenomenon, that we pursue in this paper, is that individuals fail to adjust properly for repetitions of information. In the example above, repetition occurs because the article reflects the newspaper's general political views, expressed also in previous articles. An individual who fails to adjust for this repetition (by not discounting appropriately the arguments presented in the article), would be predictably swayed towards the newspaper's views, and the more so, the more articles he reads.[2] We refer to the failure to adjust properly for information repetitions as *persuasion bias*, to highlight that this bias is related to persuasive activity.

More generally, the failure to adjust for repetitions can apply not only to information coming from one source over time, but also to information coming from multiple sources connected through a social network. Suppose, for example, that two individuals speak to one another about an issue after having both spoken to a common third party on the issue. Then, if the two conferring individuals do not account for the fact that their counterpart's opinion is based on some of the same (third party)

[1] Suppose that, starting from $t = 0$, the individual reads the article at $t = 1$. Denote by $I_t$ the individual's information set at date $t$, and by $x_t$ the individual's political beliefs at that date, e.g., $x_t = \Pr(\text{Policy } A \text{ is good}|I_t)$. Then, the law of iterative expectations implies that $E[x_1|I_0] = x_0$, i.e., reading the article should have no expected effect on the individual's beliefs.

[2] Formally, we are arguing that the change in the individual's beliefs can be predicted based on the newspaper's political slant, even when this political slant is known to the individual, i.e., is included in $I_0$. For example, if the newspaper is pro-Policy A of footnote 1, then we are arguing that $E[x_1|I_0] > x_0$.

information as their own opinion, they will double-count the third party's opinion.

Our notion of persuasion bias can be viewed as a simple, boundedly-rational heuristic for dealing with a very complicated inference problem. Correctly adjusting for repetitions would require individuals to recount not only the source of all the information that has played a role in forming their beliefs, but also the information that led to the beliefs of those they listen to, of those who those they listen to listen to, and so on. This would become extremely complicated with just a few individuals and a few rounds of updating, to say nothing of a large population of individuals, interacting according to a social network, where beliefs are derived from multiple sources over an extended time period. Under persuasion bias instead, individuals update sensibly, as Bayesians, except that they do not account accurately for which components of the information they receive is new and which is repetition. This notion corresponds with social psychology theories of political opinion formation. In particular, Lodge [1995] develops an on-line model of information processing, where individuals receive messages from campaigns, and integrate their interpretation of the message into a "running tally" for the candidate, while forgetting the details (e.g., source) of the message. We elaborate on the running-tally interpretation of persuasion bias in Section II.

Persuasion bias is consistent with psychological evidence. Several studies document that the simple repetition of statements increases the subjects' belief in the statements' validity. The interpretation given by these studies is that repetition makes the statements more familiar, and familiarity serves as a cue to validity (that is, subjects are more likely to believe a statement if it "rings a bell").[3] Closely related to familiarity, are the notions of salience and availability: repetition can have an effect because it makes statements more salient/available in the subjects' memory.[4] Our analysis is consistent with such evidence. Indeed, while we emphasize the interpretation of persuasion bias as a boundedly-rational heuristic for an otherwise intractable problem, the effect of repetition in our model could alternatively arise from familiarity, salience, or limited memory.[5]

Persuasion bias yields a direct explanation for a number of important phenomena. Consider, for example, the issue of air-time in political campaigns and court trials. A political debate without equal time for both sides, or a criminal trial in which the defense was given less time to present its case than the prosecution, would generally be considered biased and unfair.[6] This seems at odds with a rational model. Indeed, listening to a political candidate should, in expectation, have no effect on a rational individual's opinion, and thus, the candidate's air-time should not matter. By contrast, under persuasion bias, the repetition of arguments made possible by more air-time

[3]The repetition-induced increase in validity has been labeled the "truth effect." See, for example, Hawkins and Hoch [1992], and the references therein. A somewhat related notion is that upon hearing a statement, individuals store it in their memory as "true," and then engage in costly processing to determine whether it is actually true or false. (See, for example, Gilbert [1991].) Under this notion, however, it is unclear why hearing a statement *repeatedly* would have a *cumulative* effect, i.e., why persuasion would be effective.

[4]On salience and availability bias see, for example, Tversky and Kahneman [1973], Nisbett and Ross [1980], and Fiske and Taylor [1984]. See also Zallner [1992] for memory-based theories of political opinion formation whereby beliefs depend on probabilistic memory with more likely recall for recent considerations.

[5]Additionally, our model does not preclude the possibility that the effect of repetition arises because individuals under-react to information initially, and then adjust over time.

[6]Indeed, inequities in fundraising seem to be such a contentious issue in political campaigns precisely because unequal funds can be used to purchase unequal air-time.

can have an effect.[7] Other phenomena that can be readily understood with persuasion bias are marketing, propaganda, and censorship. In all these cases, there seems to be a common notion that repeated exposures to an idea have a greater effect on the listener than a single exposure. More generally, persuasion bias can explain why individuals' beliefs often seem to evolve in a predictable manner towards the standard, and publicly known, views of groups with which they interact (be they professional, social, political, or geographical groups) – a phenomenon considered indisputable and foundational by most sociologists.[8] We elaborate on a number of these phenomena and discuss additional related applications in Section V.

While these phenomena follow immediately from our assumption of persuasion bias, the primary goal of our model is to analyze the dynamics of beliefs under this bias. This analysis yields several deeper implications of our notion of persuasion bias: in particular, the phenomena of social influence and unidimensional opinions. To describe these phenomena, we need to briefly describe our model.

The essence of our model is as follows. A set of agents start with imperfect information on an unknown parameter, and then communicate according to a social network. The network is represented as a directed graph indicating whether agent $i$ "listens to" agent $j$. This graph, or "listening structure," is exogenous and can correspond to geographical proximity (neighbors, co-workers), social relationships (friendship, club membership), hierarchical relationships (position in an organizational chart), etc. Communication occurs over multiple rounds. In each round, agents listen to the beliefs of others according to the listening structure, and then update their own beliefs. For simplicity, we assume that agents report their beliefs truthfully.[9] Our key assumption is that agents are subject to persuasion bias; that is, they fail to adjust for possible repetitions of information they receive, instead treating all information as new.

In this setting we show that with sufficient communication, the beliefs of all agents converge over time to a weighted average of initial beliefs. The weight associated to any given agent measures the agent's impact on the group's beliefs, and can naturally be interpreted as the agent's *social influence*. We show that an agent's social influence depends both on how many other agents listen to him, and on the endogenously determined social influence of those agents.

[7] A rational explanation for the importance of air-time, or persuasion more generally, is that the resources one spends on persuasive activity can signal the accuracy of one's opinion. This explanation can apply, however, only when persuasive activity is chosen endogenously. Yet, persuasive opportunities, such as air-time, are perceived to be important even when they are allocated exogenously by some prior mechanism such as courtroom or debate procedures. Uneven allocation of campaign financing funds or access to debates are contentious issues in political campaigns, even when this is done according to well-specified rules set in advance. Furthermore, parties engaged in persuasive activity often seem to try to downplay or hide the amount of resources they have expended in this activity. Politicians do not publicly boast about funds they have raised (rather, their opponents point this out), and marketers frequently look for subtle ways to advertise, where their influence will be hidden.

[8] For a classic case study, see Festinger, Schachter, and Back [1950] who study the evolution of beliefs among residents of MIT dormitories. This study argues that the nature of group influence is related to the underlying social network in the group, for which frequency of contact plays an important role (see Chapter 3 on the Spatial Ecology of Group Formation).

[9] We are thus ignoring issues of strategic communication. Ignoring these issues is reasonable in many settings (e.g., sharing of political opinions or stock market views among friends and colleagues). And as will be seen, our notion of persuasion will yield interesting biases in the evolution of beliefs, absent any strategic behavior. Nonetheless, in many persuasive settings (e.g., political campaigns and court trials) agents clearly do have incentives to strategically misreport their beliefs. We discuss the interaction of strategic behavior with our notion of persuasion in Section V.

Note that if instead the information of all agents in the group were aggregated optimally, one's influence on the group's beliefs would depend only on the accuracy of one's initial information. In our setting, however, influence is determined not only by accuracy, but also by network position. This seems quite realistic: well-connected individuals often seem to be very influential in a way that is not necessarily related to the accuracy of their information.[10] We show that persuasion bias is crucial for this phenomenon: if agents are fully rational, know the entire social network, and can consequently perform the complicated calculations to adjust for information repetitions, information is aggregated optimally, and influence is solely determined by accuracy.[11] Under persuasion bias instead, well-connected agents are influential because their information is repeated many times in the social network.

Persuasion bias implies an additional general phenomenon that we refer to as *unidimensional opinions*. Quite often, individuals' opinions over a multidimensional set of issues can be well approximated by a simple one-dimensional line, where an individual's position on the line (i.e., "left" or "right") determines the individual's position on all issues. For example, many individuals' opinions on a wide range of essentially unrelated issues, ranging from free trade to military spending to environmental regulation to abortion, can be characterized by the single measure of how conservative or liberal they are.[12]

The phenomenon of unidimensional opinions follows in our model by considering the long-run differences of opinion between agents. Even though beliefs converge over time, at any point in time there is disagreement. Obviously, for any single issue, we can characterize this disagreement by agents' relative positions along an interval. Quite interestingly, we show that with sufficient communication, and for a general class of networks, an agent's relative position on *any* issue will be the same. For example, if an agent has an extreme viewpoint on one issue, his viewpoint will be extreme on all issues. Thus, in the long run, a simple left-right characterization can be used to identify agents' beliefs on all issues. Moreover, the relative position of an agent, and the association of beliefs (i.e., which collection of beliefs define the "left" vs. "right" leanings), can be predicted based on the social network and the initial beliefs.

We explore the implications of our model in several particular settings. One natural setting is when agents live in neighborhoods (defined by geographical proximity, cultural proximity, social relationships, etc.) and listen to their neighbors. In this setting it is natural to presume that communication is bilateral, i.e., the listening relationship is symmetric. Under bilateral communication,

[10] As an example, consider the individuals participating in internet chat rooms on financial investments. Many of these individuals have inaccurate (and sometimes false) information. However, the fact that they have a large audience, and are thus "well-connected," can give them enough influence to affect market prices. For example, the Wall Street Journal (November 6, 2000) reports that "Preliminary figures show that market manipulation accounted for 8% of the roughly 500 cases the SEC brought in fiscal 2000, ended Sept. 30, up from 3% in fiscal 1999. 'Manipulation on the Internet is where the action is,' and appears to be replacing brokerage 'boiler rooms' of the past, said SEC enforcement-division director Richard Walker."

[11] For general results on communication between rational agents see, for example, Geanakoplos and Polemarchakis [1982] and Parikh and Krasucki [1990].

[12] That political opinions can be well approximated by a simple unidimensional structure is a well accepted principle in political science. See Section V for further discussion and references.

the relative social influence weights can be computed explicitly, and depend only on characteristics of each agent's local neighborhood. (For example, they depend on the number of the agents' neighbors but - somewhat surprisingly - not on the number of their neighbors' neighbors.) For the special case where neighborhoods have a linear structure, we explicitly compute long-run differences of opinion and show that these are smallest between neighbors. Hence, agents end up listening to those with similar beliefs - not because we have assumed that neighbors are chosen on the basis of common beliefs, but rather because neighbors' beliefs evolve to being similar through the social updating process.

We explore, more informally, a number of important applications of our model, including political discourse, court trials, and marketing. We argue that our model provides a useful way for thinking about opinion formation in these settings, and has a number of novel empirically testable predictions. One prediction, for example, is that political views should follow a unidimensional pattern, and that adherence to this pattern should be greater for issues that are frequently discussed (which can be proxied by media coverage). Additionally, the similarity between the views of different individuals should depend on how close these individuals are in the social network. Thus, the relative importance of different demographic factors in explaining individuals' political orientation should be related to the relative importance of such factors in determining social interaction.

This paper is related to several literatures. In addition to the psychology literatures on belief formation discussed earlier, there is a formal relation to a mathematical sociology literature on *social networks*. This literature considers measures of "social power," "centrality," and "status," and how they can be derived from the structure of the social network. In particular, French [1956] and Harary [1959] consider an abstract model of social power evolution which has dynamics similar to ours.[13] There are several important differences between our work and this literature. First, while the concepts of social power and centrality are somewhat abstract, we focus concretely on beliefs. This allows us to interpret our dynamics relative to rational updating, and makes clear the precise behavioral assumptions underlying our model. Second, we derive new results on the dynamics, including a characterization of long-run differences in beliefs. Third, we derive new analytical results and empirical predictions for a variety of applications, such as neighborhoods, political science, court trials, and marketing.

Also related is the theoretical literature on social learning, in which agents are assumed to choose actions over time based on information received from others.[14] As in the social learning literature, agents communicate with others and update their beliefs. Unlike that literature, however, the updating process is subject to persuasion bias. Because of persuasion bias, communication becomes a mechanism for exerting influence: well-connected individuals can influence others in a way that is

[13]More recent treatments include Wasserman and Faust [1994] and Bonacich and Lloyd [2001]. Dynamics similar to those in French and Harary are obtained in a statistics literature starting with DeGroot [1974]. See also Press [1978] and Berger [1981].

[14]See, for example, Banerjee [1992], Bikhchandani, Hirshleifer and Welch [1992], and Ellison and Fudenberg [1993], where agents receive information from the entire population, and Ellison and Fudenberg [1995], where agents receive information from their neighbors. See also Bala and Goyal [1998] for a general neighborhood structure, analogous to our network structure.

not warranted by the accuracy of their information.

Finally, this paper is related to an empirical literature which examines whether social networks influence individuals' outcomes.[15] The findings of this literature suggest that social networks do indeed have an effect: for example, controlling for many relevant characteristics, individuals are more likely to participate in welfare programs if their neighbors do, and are likely to invest in a similar way to their neighbors.

This paper is organized as follows. Section II presents the model. Section III analyzes the dynamics of agents' beliefs, and shows the phenomena of social influence and unidimensional opinions. Section IV applies the model to neighborhoods, and Section V to political discourse, court trials, and marketing. Section VI concludes, and all proofs are in the Appendix.

# II The Model

We consider a setting where a finite set of agents $\mathcal{N} = \{1, 2, \ldots, N\}$ wish to estimate an unknown parameter $\theta \in \mathbb{R}^L$ (we later interpret each dimension in $\mathbb{R}^L$ as a different "issue"). Agents start with some initial information on $\theta$. For simplicity of exposition, we assume here that this information consists of normally distributed noisy estimates of $\theta$, though below we observe that our analysis also applies to a very general, distribution-free setting of "observations." Specifically, agent $i \in \mathcal{N}$ starts with a noisy signal $x_i^0 = \theta + \epsilon_i$, where $\epsilon_i \in \mathbb{R}^L$ is an error term whose components are mean zero, independent across agents and dimensions, and normally distributed. We assume that agent $i$ assigns initial precision $\pi_{ij}^0$ to agent $j$, and this precision applies to all components of agent $j$'s error term, i.e., $\pi_{ij}^0 = Var_i[\epsilon_{j\ell}]^{-1}$, for $\ell = 1, \ldots, L$. (For generality, we allow $i$'s assessment of $j$'s precision to be subjective, but it could also be objectively correct for all agents.)

Upon receiving their initial information, agents communicate according to a social network. We describe the network as a directed graph indicating whether agent $i$ *listens* to agent $j$. We denote by $S(i) \subset \mathcal{N}$ the set of agents that agent $i$ listens to, and by $q_{ij} \in \{0, 1\}$ the indicator function of $S(i)$, i.e., $q_{ij} = 1$ if and only if $j \in S(i)$. We refer to the set $S(i)$ as the *listening set* of agent $i$, and to the function $S$ as the *listening structure*. Since agent $i$ knows his own information (he listens to himself), $i \in S(i)$.

Communication occurs over multiple rounds. In each round, agents listen to the beliefs of those in their listening set (where beliefs will be explained below), and then update their own beliefs. As discussed in the introduction and in greater length in Section V, we assume that agents report their beliefs truthfully.

We model our notion of persuasion bias by presuming that agents treat all information they receive as new, ignoring possible repetitions. With only one communication round, the information agents receive is genuinely new, since the error terms $\epsilon_i$ in agents' signals are independent. Hence,

[15]See, for example, Case and Katz [1991], Evans, Oates, and Schwab [1992], Borjas [1995], Glaeser, Sacerdote, and Scheinkman [1996], Katz, Kling, and Liebman [1999], Bertrand, Luttmer, and Mullainathan [2000], Duflo and Saez [2000], Madrian and Shea [2000], and Hong and Stein [2001].

there is no persuasion bias, and agents update their beliefs in an optimal manner, giving weight to others' information in proportion to the precision they ascribe to them. With multiple communication rounds, however, agents' information involves repetitions. These arise both because one listens to the same set of agents over time, and because those agents might be reporting information from overlapping sources. Agents, however, treat such repetitions as new information. Specifically, agents treat the information they hear in each round as new and independent, and ascribe the same relative precisions to those they listen to as in the first round (not adjusting for the fact that over time, the information of some agents might contain more repetitions than that of others).

Correctly adjusting for repetitions would impose a heavy computational burden on the agents. Suppose, for example, that an agent, say $i$, listens to two other agents, who themselves listen to overlapping sets of agents. Then, agent $i$ must know the extent to which the listening sets of the two agents overlap, in order to adjust for the correlation in the information he is receiving. Moreover, he must recall the information he has received from each of the two agents in the previous communication rounds, to disentangle old information from new. (For example, in the second round, the old information consists of the signals of the two agents, and the new information of the signals of those the two agents listen to.) With multiple communication rounds, such calculations would become quite laborious, even if the agent knew the entire social network. If instead, the agent did not know the entire network (as is likely to be the case in large networks), the agent would have to instead try to infer the sources of his sources' information from the reports he received. Even for a simple network, this appears to be an extremely complicated problem that does not beget an obvious solution.[16]

Persuasion bias can be viewed as a simple, boundedly-rational heuristic for dealing with the above complicated inference problem, when agents cannot determine (or recall) the source of all the information that has played a role in forming their beliefs. Agents might instead simply maintain a "running tally" summarizing their current beliefs, as in Lodge's [1995] conception of on-line information processing. Furthermore, upon hearing such a summary from someone else, agents would be unable to distinguish which part is new information, and which part is information already incorporated into their own beliefs. Agents might then simply integrate the information they hear into their running tally, using an updating rule that would be optimal if the information were new. This is precisely what agents do under persuasion bias.

Formally, communication and updating in our model are as follows. In the first communication round, agent $i$ learns the signals of the agents in $S(i)$. Given normality and agents' fixed assessment of the precision of others' information, a sufficient statistic for these signals is their weighted average, with weights given by the precisions. We denote this statistic by $x_i^1$, and refer to it as agent $i$'s *beliefs*

[16] Instead of requiring agents to know the entire social network, agents could alternatively communicate not only their beliefs, but also the sources of their information, the sources of their sources, etc. Such communication, however, would be extremely complicated: the information agents would need to recall and communicate would increase exponentially with the number of rounds.

after one round of updating:

$$x_i^1 = \sum_j \frac{q_{ij}\pi_{ij}^0}{\pi_{ii}^1} x_j^0, \tag{1}$$

where $\pi_{ii}^1 = \sum_j q_{ij}\pi_{ij}^0$ denotes the precision that agent $i$ assigns to his beliefs. We refer to equation (1) as agent $i$'s updating rule.[17]

In the updating rule (1), agent $j$ gets non-zero weight if $q_{ij} = 1$ ($i$ listens to $j$) and $\pi_{ij}^0 > 0$ ($i$ believes $j$'s information is useful). From now on, we assume that agent $i$ believes that the information of all agents he listens to is useful, i.e., $\pi_{ij}^0 > 0$ for all $j \in S(i)$. This is without loss of generality: if $i$ listens to an agent without useful information, then we can exclude that agent from $S(i)$ without changing that agent's (zero) weight. The agents who are not in $S(i)$ can thus be interpreted as those with whom $i$ has either no direct contact, or no faith in the validity of their information.

The updating rule (1) can be expressed more succinctly in vector notation. Denote by $x^t$ the matrix whose $i$th row is the vector $x_i^t$ of agent $i$'s beliefs in communication round $t$. Denote also by $\mathbf{T}$ the *listening matrix* with elements

$$\mathbf{T}_{ij} = q_{ij}\ \pi_{ij}^0/\pi_{ii}^1. \tag{2}$$

Then, the updating rule (1) can be expressed as

$$x^1 = \mathbf{T}x^0. \tag{3}$$

To illustrate the above notation, we consider the following example.

**Example 1** Suppose that $N = 4$, $S(1) = \{1, 2\}$, $S(2) = \{1, 2, 3\}$, $S(3) = \{1, 2, 3, 4\}$, and $S(4) = \{1, 3, 4\}$. This listening structure depicts a setting where four agents are ordered along a line, with everyone listening to their direct neighbors, themselves, and agent 1 even when agent 1 is not a direct neighbor. Then, if each agent believes that all agents he listens to have equal precision, the listening matrix $\mathbf{T}$ is given by

$$\mathbf{T} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

The directed graph corresponding to the listening structure $S$ is depicted in Figure 1. We use this listening structure in Section III as a running example.

Beliefs (i.e., a sufficient statistic for an agent's information) followed a linear updating rule in (1) by normality. Linearity, however, also holds for the following very general, distribution-free setting. Suppose that for each dimension $\theta_\ell$ of $\theta$, agents start with some "observations," rather than with a single signal. Suppose that observations are identically distributed, independent conditional on $\theta_\ell$,

[17] To be precise, this equation determines a sufficient statistic for agent $i$'s information (for a given precision), rather than $i$'s posterior estimate of $\theta$. The posterior estimate also depends upon $i$'s prior on $\theta$, and coincides with the sufficient statistic if the prior is diffuse. By assuming that agents communicate raw data rather than posteriors, we focus on how data is aggregated by the social network, absent any effect of the priors.

and with outcomes drawn from a finite set.[18] Then, for a given number of observations, a sufficient statistic of an agent's observations is the empirical frequency distribution of the observed outcomes. If agents in turn communicate this sufficient statistic, then they should update exactly as above: the empirical frequency distribution that agent $i$ computes after the first communication round will be given by equation (1), where now $\pi_{ij}^0$ represents agent $i$'s assessment of the number of observations that underlie agent $j$'s initial distribution.

Turning to multiple communication rounds, our persuasion assumption is that agents ignore repetitions of information. Rather, agents treat the information they hear in each round as new and independent, and ascribe the same relative weights to those they listen to as in the first round. We do allow, however, agents to vary the weight they give to their own beliefs relative to outsiders over time as they update more. Thus, agents' beliefs after round $t+1$ are given by

$$x^{t+1} = \mathbf{T}_t \, x^t, \tag{4}$$

where $\mathbf{T}_t$ denotes the matrix given by

$$\mathbf{T}_t = (1-\lambda_t)\mathbf{I} + \lambda_t \mathbf{T}, \tag{5}$$

where $\lambda_t \in (0,1]$. These belief dynamics, given by (4) and (5), are the focus of our analysis.[19]

# III Belief Dynamics

In this section, we study the dynamics of beliefs under persuasion bias. We determine the long-run limit to which agents' beliefs converge, and the nature of agents' long-run differences of opinion. Furthermore, we characterize these as a function of the social network according to which agents communicate.

The dynamics of beliefs are given by equations (4) and (5). Setting $\mathbf{T}(\lambda) \equiv ((1-\lambda)\mathbf{I} + \lambda\mathbf{T})$, we

[18] Interpreting agents' information as consisting of observations is sensible in many settings, ranging from formal scientific experiments to the experiences consumers have regarding the quality of a product.

[19] While we consider belief dynamics for a general process $\{\lambda_t\}_t$, it is worth noting that specific models of how agents perceive the information of others to evolve will give rise to specific processes. One simple model is that of *equal learning*, where agents believe that others learn at the same rate as they do, i.e., $\pi_{ij}^{t+1}/\pi_{ij}^t = \pi_{ii}^{t+1}/\pi_{ii}^t$ for all $i$, $j$, and $t$, where $\pi_{ij}^t$ denotes the precision that agent $i$ assigns to the information he hears from agent $j$ in round $t+1$. Under this model, the matrix $\mathbf{T}_t$ is constant over time and equal to $\mathbf{T}$, i.e., $\lambda_t = 1$.

Another natural, though slightly more complicated, model is that of *increasing self-confidence*, where agents assume that only they are becoming more informed over time, while others are not, i.e., $\pi_{ij}^t = \pi_{ij}^0$ for all $i$, $j \neq i$, and $t$. With the additional assumption that the increase in self-confidence is proportional to each agent's initial precision (i.e., $\pi_{ii}^{t+1} - \pi_{ii}^t = \sum_{j \neq i} q_{ij}\pi_{ij}^0 = k\pi_{ii}^0$ for all $i$), we have $\mathbf{T}_{ij} = q_{ij}\pi_{ij}^0/[\pi_{ii}^0(1+k)]$, and $\mathbf{T}_{t,ij} = q_{ij}\pi_{ij}^0/[\pi_{ii}^0(1+k(t+1))]$ for $j \neq i$. Consequently, the matrix $\mathbf{T}_t$ is again given by equation (5), with $\lambda_t = (1+\frac{k}{1+k}\,t)^{-1}$.

Note that equations (4) and (5) do not preclude the possibility that the effect of repetition arises because individuals under-react to their initial information, and then adjust over time. Such under-reaction corresponds to $\lambda_0 < 1$.

can write the beliefs after communication round $t$ as

$$x^t = \left[\prod_{s=0}^{t-1} \mathbf{T}(\lambda_s)\right] x^0.$$

Since each row of the matrix $\mathbf{T}$ sums to one, the same is true for $\mathbf{T}(\lambda)$ and hence for $\prod_{s=0}^{t-1} \mathbf{T}(\lambda_s)$.[20] Therefore, the beliefs of an agent $i$ after any communication round are a weighted average of all agents' initial beliefs. The weight of agent $j$ can naturally be interpreted as $j$'s influence on $i$.

**Definition 1** The *influence* of agent $j$ on agent $i$ after $t$ rounds of communication is

$$w_{ij}^t \equiv \left[\prod_{s=0}^{t-1} \mathbf{T}(\lambda_s)\right]_{ij},$$

and the *direct influence* of agent $j$ on agent $i$ is $\mathbf{T}_{ij}$.

The agents having a direct influence on $i$ are those in $i$'s listening set $S(i)$. These agents influence $i$ in the first communication round, since $w_{ij}^1 = \mathbf{T}(\lambda_0)_{ij}$ and $\lambda_0 > 0$. In the second communication round, $i$ is influenced by agents who influenced agents in $S(i)$ in the first round. That is, $w_{ij}^2$ is positive for agents $j$ such that $j \in S(k)$ and $k \in S(i)$. Continuing in this fashion, we see that for large enough $t$, $i$ is influenced by agents in the set

$$S^*(i) \equiv \{j : j \in S(k_1), k_1 \in S(k_2), \ldots, k_{r-1} \in S(k_r), k_r \in S(i)\}.$$

### III.A Belief Convergence and Social Influence

We next show that under general conditions, beliefs converge to a limit which is common to all agents. This means that in the long run agents reach a consensus. We characterize the consensus beliefs, and examine whether they aggregate optimally agents' initial information.

A necessary condition for consensus is that agents are not isolated from each other. If $j \notin S^*(i)$ (i.e., agent $j$ does not influence agent $i$) and $i \notin S^*(j)$ (i.e., $i$ does not influence $j$) then $j$'s beliefs are never incorporated into $i$'s, and vice versa. To rule this out, we assume that the set $\mathcal{N}$ of agents is strongly connected, in the following sense:

**Definition 2** A set $A$ of agents is *strongly connected* if for all $i \in A$, $A \subset S^*(i)$.

If $\mathcal{N}$ is strongly connected, then every agent influences every other agent, and thus agents are not isolated from each other.

Additionally, to reach consensus, agents must not become too fixed in their beliefs, i.e., $\lambda_t$ must not go to zero too quickly. Otherwise, agents would stop updating and consensus would not be reached. To rule this out, we make the following assumption:

[20] To see why, note that $\left[\prod_{s=0}^{t-1} \mathbf{T}(\lambda_s)\right] \mathbf{1} = \left[\prod_{s=0}^{t-2} \mathbf{T}(\lambda_s)\right] \mathbf{T}(\lambda_{t-1})\mathbf{1} = \left[\prod_{s=0}^{t-2} \mathbf{T}(\lambda_s)\right] \mathbf{1} = \ldots = \mathbf{1}$, where $\mathbf{1} = (1, \ldots, 1)'$.

**Assumption 1** $\sum_{t=0}^{\infty} \lambda_t = \infty$.

Assumption 1 is satisfied for many plausible specifications of $\lambda_t$, and in particular, those mentioned in footnote 19. We maintain Assumption 1 throughout our analysis.

**Theorem 1** *Suppose $\mathcal{N}$ is strongly connected, and Assumption 1 holds. Then there exists a vector of strictly positive weights $w = (w_1, \ldots, w_N)$ such that the influence of agent $j$ on each agent $i$ converges to $w_j$, i.e.,*

$$\lim_{t\to\infty} w_{ij}^t = w_j.$$

*Furthermore, the beliefs of each agent $i$ converge to the consensus beliefs $wx^0$, i.e.,*

$$\lim_{t\to\infty} x_i^t = wx^0.$$

*The vector $w$ is given by the unique solution to $w\mathbf{T} = w$. We refer to $w_j$ as the* **social influence** *of agent $j$.*

The intuition for this result can be seen by relating our model to the standard theory of finite Markov chains. To do this, consider a Markov chain whose set of states is the set of agents $\mathcal{N}$, and whose one-step transition probability from state $i$ to state $j$ is $\mathbf{T}_{ij}$. Then the $t$-step transition probability from $i$ to $j$ is $(\mathbf{T}^t)_{ij}$, i.e., the influence weight $w_{ij}^t$ in the case where $\lambda_t = 1$ for all $t$.

Our assumption that $\mathcal{N}$ is strongly connected implies that the Markov chain is irreducible. Furthermore, the Markov chain is aperiodic since $\mathbf{T}_{ii} > 0$. (We have $\mathbf{T}_{ii} > 0$ in our model, since agents listen to themselves, i.e., $i \in S(i)$.) It is well known that a finite, irreducible, aperiodic Markov chain has a *unique* stationary distribution, i.e., a unique probability distribution $w$ satisfying $w\mathbf{T} = w$. Moreover, starting from any state $i$, the probability of being in state $j$ at date $t$ converges to $w_j$ as $t \to \infty$.[21] This means that the influence weight $w_{ij}^t$ converges to $w_j$ (and thus beliefs converge to a consensus belief) if $\lambda_t = 1$. What remains to show is that one still gets convergence, and the limit is the same, for any process $\lambda_t$ satisfying Assumption 1. That the limit is the same (assuming that convergence occurs) follows from $\mathbf{T}(\lambda)$ being a weighted average of $\mathbf{T}$ and the identity matrix. Intuitively, if all agents uniformly put more weight on their own beliefs, this does not affect the path of convergence, but only the speed at which beliefs converge. “Enough” weight must be put on others’ beliefs, however, to ensure that convergence does occur. In the proof we show that “enough” weight is determined precisely by whether Assumption 1 holds or not.

The weight $w_i$ can be interpreted as the *social influence* that agent $i$ has on the consensus beliefs. To gain some intuition on how social influence depends on the listening structure, we consider Example 1. In this example, the set $\mathcal{N}$ is strongly connected, and the social influence weights are $w = \left(\frac{16}{42}, \frac{15}{42}, \frac{8}{42}, \frac{3}{42}\right)$. Not surprisingly, this example demonstrates that an agent is influential if he is listened to by many other agents. Thus, the most influential agent is agent 1, the only agent listened to by all other agents. Less obvious, however, is the insight that an agent’s influence depends not

[21] See, for example, Aldous and Fill [1999, Ch. 2] and the references therein.

only on the number of agents who listen to that agent, but also on whether those agents themselves are influential. Indeed, both agents 2 and 3 are listened to by the same number of agents. However, agent 2 is more influential than agent 3 because he is listened to by the influential agent 1 (in addition to agent 3) while agent 3 is listened to by the less influential agent 4 (in addition to agent 2). To formalize this intuition, recall that $\mathbf{T}_{ij}$ is the *direct influence of j on i.* Then, the condition $w\mathbf{T} = w$ can be restated as the following immediate corollary:[22]

**Corollary 1** *The social influence of agent i is the sum over j of the direct influence of i on j (i.e., $\mathbf{T}_{ji}$) times the social influence of j.*

Having characterized the consensus beliefs, we next examine whether these beliefs are correct in the sense of being optimal aggregates of agents' initial information. This question is most meaningful when agents hold correct beliefs about others' precision so we assume this for the remainder of this subsection: $\pi^0_{ij} = \hat{\pi}^0_j$ for all $i, j$, where $\hat{\pi}^0_i$ is the true precision of agent $i$'s initial information. Then, under optimal aggregation of information, beliefs would be

$$\hat{x} = \frac{\sum_i \hat{\pi}^0_i x^0_i}{\sum_i \hat{\pi}^0_i}.$$

In general, it is obvious that consensus beliefs under persuasion will not yield the optimal aggregation of information. In fact, defining the self-weight $\mathbf{T}_{jj}$ as agent $j$'s *self-importance*, Theorem 2 states that consensus beliefs are correct if and only if the listening structure is *balanced*, in the sense that the total self-importance of those who listen to each agent is equal to one.

**Theorem 2** *Suppose $\pi^0_{ij} = \hat{\pi}^0_j$ for all $i, j$. Then, the consensus beliefs are correct if and only if for all $i$,*

$$\sum_j q_{ji} \mathbf{T}_{jj} = 1. \tag{6}$$

Typically, listening structures are unbalanced. For example, any listening structure in which an agent is listened to by a subset of those listening to another agent is unbalanced. There are, however, some non-trivial balanced listening structures. For example, suppose that all agents have equal precision ($\hat{\pi}^0_i = \hat{\pi}^0_j$ for all $i, j$). Then, any listening structure where there exists an $n$ such that each agent listens to, and is listened to by, exactly $n$ agents is balanced. (An example is when agents are on a circle, and each agent listens to his two neighbors.)

Intuitively, for an unbalanced listening structure, consensus beliefs are incorrect because well-connected agents have influence in excess of the accuracy of their information. This excess influence derives from persuasion bias: the information of well-connected agents is repeated many times in the social network, and no adjustment is made for the repetitions. This notion is formalized in Theorem 3, which indicates that when agents are not subject to persuasion bias, and the social network is common knowledge, consensus beliefs are correct.

[22]Statements similar to that in Corollary 1 are often taken as defining "axioms" of (rather than derived results about) social power in the mathematical sociology literature. See, for example, Bonacich and Lloyd [2001].

**Theorem 3** *Suppose agents are fully rational (i.e., not subject to persuasion bias), the social network is common knowledge, $\mathcal{N}$ is strongly connected, and $\pi_{ij}^0 = \hat{\pi}_j^0$ for all $i, j$. Then, agents converge to the correct beliefs $\hat{x}$ after at most $N^2$ rounds of updating.*

The intuition for this result is quite simple. Recall from Section II that agents can adjust for repetitions if they know the structure of the social network, and can keep track of all the information they have learned in the previous rounds. The latter information consists of linear combinations of other agents' signals, and since agents know the structure of the social network, they know what these linear combinations are. With enough communication rounds, agents learn enough linear combinations as to deduce the signal of every agent in the network.

We should emphasize that the calculations that agents must perform even in this simple case where the network is common knowledge can be very complicated. In fact, the proof of the theorem is not constructive, i.e., we do not determine the linear combinations that an agent receives. Rather, we show that in each round where agents have not converged to the correct beliefs, at least one agent must learn a new linear combination.

## III.B Long-Run Differences of Opinion and Unidimensionality

Even though agents' beliefs converge over time, at any point in time there is disagreement. In this subsection we characterize the relative disagreement, or relative "differences of opinion." This characterization is important for applications, such as political science, where relative opinions are the most relevant.[23]

At any time $t$, let $\bar{x}^t \equiv \frac{1}{N}\sum_i x_i^t$ be the average of the population beliefs. To measure relative disagreement on "issue" (i.e., belief dimension) $\ell$, we take the difference between agent $i$'s beliefs on this issue, $x_{i\ell}^t$, and the average population beliefs. We also express this difference in relative terms, by normalizing by the average difference of beliefs in the population. That is, we define

$$d_{i\ell}^t \equiv \frac{x_{i\ell}^t - \bar{x}_\ell^t}{\frac{1}{N}\sum_j \|x_j^t - \bar{x}^t\|},$$

where $\|x\|$ denotes the Euclidean norm of the vector $x$.[24] The relative difference of opinion $d_{i\ell}^t$ will be positive or negative, depending upon whether agent $i$'s beliefs are to the "right" or "left" of average on issue $\ell$.

Initially, an agent might be to the right on some issues, and to the left on others. Our main result

[23] Indeed, it is not at all clear how to measure "absolute" disagreement in many contexts. One could argue that liberal and conservative views in the United States are very dissimilar; on the other hand, taken in the broad context of possible political views (fascism, communism, monarchy, anarchy, etc.), they can also be seen as extremely close.

[24] Our results do not depend on the particular choice of the weights used to compute the "average" population beliefs. Equal weights are a natural choice, but any other weighted average can be used without changing the results. (In fact, in the proof of Theorem 4, we consider a general weight vector $\hat{w}$.) Our results also do not depend on the particular choice of normalization. An alternative normalization, for example, is by the average difference of beliefs *on issue* $\ell$, i.e., $\sum_j |x_{j\ell}^t - \bar{x}_\ell^t|/N$, or by the size of the largest disagreement on that issue, i.e., $\max_{j,j'} |x_{j\ell}^t - x_{j'\ell}^t|$. We choose a normalization that is issue-independent to emphasize that this normalization plays no role in our results.

in this section is that with sufficient communication, and for a general class of networks, an agent's relative position on *any* issue will be the same. That is, in the long run, if agent $i$ is "moderately to the right" on one issue, then agent $i$ will be moderately to the right on *all* issues. This implies that long-run differences of opinion across all issues can be summarized by identifying agents on a simple "left-right" spectrum. In this case, we say that long-run differences of opinion are *unidimensional.*

To illustrate the unidimensionality result, consider the listening structure of Example 1. Figure 2 illustrates the dynamics of beliefs in this example for two distinct issues represented by the two axes. Initially ($t = 1$), beliefs are arbitrary. For example, agents 1 and 4 are similar with regard to the "vertical" issue, but at opposite extremes with regard to the "horizontal" issue. However, after 16 rounds of communication, beliefs appear to lie along a single line. Agent 1 is now at one extreme on *both* issues, agents 3 and 4 are at the other extreme on both issues, and agent 2 holds moderate views.[25]

We will show that the convergence to unidimensional beliefs is a quite typical result of communication under persuasion bias. Moreover, the nature of the long-run disagreement (i.e., the line along which beliefs differ) and the long-run positions of individual agents are determined from the listening structure.

Recall that the dynamics of beliefs are determined by the listening matrix $\mathbf{T}$ through equations (4) and (5). Since $\mathbf{T}$ can be interpreted as the transition matrix of a finite, irreducible, and aperiodic Markov chain, it has a unique eigenvalue equal to one, and all other eigenvalues with modulus smaller than one. The row eigenvector corresponding to the largest eigenvalue coincides with the social influence weights (since $w\mathbf{T} = w$), and thus characterizes the beliefs to which agents ultimately converge. As we show below, the row and column eigenvectors corresponding to the second largest eigenvalue characterize the long-run differences of opinion.

Defining the second largest eigenvalue is complicated because eigenvalues might be complex, and because they must be ranked according to the ranking of their counterparts in $\mathbf{T}(\lambda_t)$ for $t$ large. The ranking criterion is, in fact,

$$f_{\lambda^*}(\alpha) \equiv \|\alpha(\lambda^*)\|^{\frac{1}{\lambda^*}},$$

where $\alpha(\lambda) \equiv (1 - \lambda) + \lambda\alpha$, $\lambda^*$ is the limit of $\lambda_t$ when $t$ goes to $\infty$, and $\|\alpha\|$ the modulus of the complex number $\alpha$.[26] Based on this ranking, we let $\alpha_n$ be the $n$th largest eigenvalue, and $V_n^r$ and $V_n^c$ the $n$th row and column eigenvectors, respectively. Of course, ties in the ranking are possible. We assume, however, that except for complex conjugate pairs, there are no ties. This is true for generic matrices, and thus our result below applies to generic listening structures.[27]

We can now state our result regarding the unidimensionality of long-run differences of opinion.

[25] Thus, we might describe agent 1 as "right-wing," agents 3 and 4 as "left-wing," and agent 2 as "centrist."

[26] Note that $\alpha(\lambda)$ is the eigenvalue of $\mathbf{T}(\lambda)$ corresponding to the eigenvalue $\alpha$ of $\mathbf{T}$, and the modulus of the complex number $a + ib$ is $\|a + ib\| \equiv \sqrt{a^2 + b^2}$. Also, we define $f_0(\alpha)$ in the obvious way, by taking the limit: $f_0(\alpha) \equiv \lim_{\lambda \to 0} f_\lambda(\alpha) = \exp[\mathrm{Re}(\alpha) - 1]$, where $\mathrm{Re}(\alpha)$ denotes the real part of the complex number $\alpha$.

[27] In fact, genericity ensures not only the lack of ties, but also that the diagonalizability of the matrix $\mathbf{T}$.

**Theorem 4** *Suppose $\alpha_2$ is real. Then, for a generic listening matrix $T$, and for all $i$ and $\ell$,*

$$d_{i\ell}^{\infty} \equiv \lim_{t\to\infty} d_{i\ell}^{t} = \phi_i D_\ell,$$

*where $\phi_i = a + bV_2^c$ for some constants $a$ and $b > 0$, and $D_\ell$ denotes the $\ell$th component of the vector $V_2^r x^0$.*

This theorem states that in the long run, any given agent's position on one issue is perfectly correlated with his positions on all other issues. This is because agent $i$'s long-run difference of opinion on issue $\ell$ depends on $i$ only through the scalar $\phi_i$, which is the same across all issues.

This scalar $\phi_i$ can naturally be interpreted as the long-run position of agent $i$ on a unidimensional spectrum of disagreement. As Theorem 4 shows, agents' positions are determined by the second column eigenvector of $\mathbf{T}$, $V_2^c$. The line along which agents disagree is, in turn, given by the vector $D \equiv (D_1, \ldots, D_L) = V_2^r x_0$, which is determined by the second row eigenvector of $\mathbf{T}$, together with agents' initial beliefs.

To illustrate Theorem 4, consider again Example 1. In this example, $\alpha_2 = 1/3$, and so the second eigenvalue is real. The second column eigenvector is $V_2^c = \left(\frac{3}{2}, -\frac{1}{2}, -\frac{3}{2}, -\frac{3}{2}\right)$. Thus, in the long run and on every issue, agent 1 is at one extreme, agents 3 and 4 at the other, and agent 2 in between and closer to agents 3 and 4. This result follows from the listening dynamics, and is independent of initial beliefs and the number of dimensions.

The second row eigenvector of $\mathbf{T}$ in Example 1 is $V_2^r = (0, 1, 0, -1)$. This implies that the long-run differences of opinion are fully determined by the initial difference of opinion between agents 2 and 4.[28] This can be seen in Figure 2 by the fact that the line along which beliefs ultimately lie ($t = 16$) has the same slope as the line between agents 2 and 4 initially ($t = 1$). Again, this result is independent of initial beliefs and the number of dimensions.[29]

A technical intuition for our unidimensionality result is as follows. Generically, the listening matrix $\mathbf{T}$ has the diagonal decomposition,

$$\mathbf{T} = V^c A V^r, \tag{7}$$

where $A$ is the diagonal matrix of the eigenvalues of $\mathbf{T}$, and $V^c$ and $V^r$ are the matrices of the column

[28] In fact, the vector $V_2^r x_0$ can always be interpreted as a difference between weighted averages of initial beliefs of two groups of agents. This is because row and column eigenvectors corresponding to different eigenvalues are orthogonal, and $V_1^c = (1, \ldots, 1)'$. Therefore, $V_2^r$ can be normalized so that the positive and negative elements each sum to 1.

[29] A more general illustration of the unidimensionality result can be found on the website `http://faculty-gsb.stanford.edu/demarzo/papers/persuasion.xls`. This website provides an Excel macro that in each application randomly locates ten agents in geographic neighborhoods (which define the listening structure as in Section IV), and independent of location also randomly assigns two-dimensional beliefs. The application then animates the evolution of differences of opinion, demonstrating the convergence to a line for any initial beliefs and neighborhood network.

and row eigenvectors, respectively. Thus,

$$\mathbf{T}^t x_\ell^0 = V^c A^t V^r x_\ell^0 = \sum_n \left[\alpha_n^t (V_n^r x_\ell^0)\right] V_n^c.$$

We can interpret this as follows. After $t$ rounds of updating, the vector of agents' beliefs on any issue $\ell$ can be represented as a linear combination of the column eigenvectors of $\mathbf{T}$, where the coefficient of the $n$th eigenvector is initially given by $V_n^r x_\ell^0$, and then declines geometrically according to the $n$th eigenvalue. (For simplicity, we are assuming here that $\lambda_t = 1$ for all $t$.) The term $V_n^r x_\ell^0$ maps the agents' initial beliefs into the long-run consensus beliefs $wx_\ell^0$, for $n = 1$ (since $V_1^r = w$), and into "differences of opinion," for $n > 1$ (by Footnote 28). Thus, in the long run, since $\alpha_n^t \to 0$ for $n > 1$, and $V_1^c = (1, \ldots, 1)'$, the differences of opinion disappear and beliefs converge to $wx_\ell^0$. However, the differences of opinion corresponding to the second largest eigenvalue ($V_2^r x_\ell^0$) persist the longest, with agents' relative positions given by $V_2^c$.

It is worth remarking on the relationship between our unidimensionality result and persuasion. Unidimensionality obtains in Theorem 4 due to the linearity of updating, i.e., due to the repeated application of a combination of $\mathbf{T}$ and $\mathbf{I}$ to prior beliefs. In words, unidimensionality depends on the Markovian structure of updating (whereby old information does not affect updating in a given period save through the formation of prior beliefs entering into that period), and the constant relative weights that an agent gives others over time. But this is precisely what our assumption of persuasion entails: agents continue to treat reports from different individuals as independent from old information (thereby justifying Markovian updating), and in a constant manner over time (yielding constant relative weights).[30]

Theorem 4 relies on the second eigenvalue being real. It is possible that the second eigenvalue is complex. This occurs in the case when the listening structure is dominated by an one-way cycle (for example, agents are on a circle and listen to their clockwise neighbors). An analogous result to Theorem 4 can be derived for the complex case. There, it can be shown that agents' relative positions cycle prior to convergence, with the periodicity of the cycle determined by the angle of $\alpha_2$ in the complex plane.[31] By generating random listening structures, we have found numerically that the case where the second eigenvalue is complex is relatively rare. Moreover, in Section IV we will show that for a broad and important class of listening structures all eigenvalues are real.

## III.C Speed of Convergence

In the previous subsections we have demonstrated the convergence of agents' beliefs and differences of opinion. In this subsection, we briefly discuss the speed of convergence. Not surprisingly, the

[30] Other updating processes generally will not yield long-run linear differences of opinion. We have already seen in Theorem 3 that if agents know the network and are fully rational, they converge in a finite number of rounds to the same beliefs, and therefore there are no long-run differences of opinion. There is little reason to believe that other, more general, updating processes would yield long-run linear differences of opinion.

[31] This result is available upon request. For example, if agents are on a circle and listen to their clockwise neighbors, and if $\lambda_t = 1$, then the cycle has period length $2\pi / \arctan(\mathrm{I}m(\alpha_2)/\mathrm{Re}(\alpha_2)) = 2N$. Since each agent moves half-way towards his neighbor each period, relative positions will cycle after $2N$ communication rounds.

convergence rate is governed by the size of the eigenvalues of the listening matrix $\mathbf{T}$, as the following theorem makes explicit.

**Theorem 5** *Consider a listening matrix* $\mathbf{T}$ *with eigenvalues* $\alpha_n$ *ordered according to* $f_{\lambda^*}$*. Then*

$$\|x_i^t - wx^0\| = O\left(f_{\lambda^*}(\alpha_2)^{\sum_{s=0}^{t-1}\lambda_s}\right) \qquad \text{and} \qquad \|d_i^t - d_i^\infty\| = O\left(\left(f_{\lambda^*}(\alpha_3)/f_{\lambda^*}(\alpha_2)\right)^{\sum_{s=0}^{t-1}\lambda_s}\right).$$

The theorem makes clear that the rate of convergence of beliefs is governed by the size of the second eigenvalue of $\mathbf{T}$, whereas the rate of convergence of differences of opinion is governed by the relative size of the second and third eigenvalues. Thus, for example, if we compare two listening matrices, $\mathbf{T}$ and $\hat{\mathbf{T}}$ with associated eigenvalues $\alpha_2$ and $\hat{\alpha}_2$ such that $f_{\lambda^*}(\alpha_2) < f_{\lambda^*}(\hat{\alpha}_2)$, then beliefs will converge more quickly under $\mathbf{T}$ than under $\hat{\mathbf{T}}$.

## III.D Generalizations of the Model

There are a number of possible generalizations of the basic model and the results. Rather than fully elaborate on them here, we describe them briefly and refer interested readers to the Appendix.

**Individuals versus Groups:** In the basic model we have interpreted each agent as an individual. In many applications, it might be useful to consider groups of similar agents, all sharing the same listening weights and position in the social network. (This is especially useful when trying to model the "macro" structure of a social network; e.g., we might identify colleagues at the same research institute as one group, and focus on the links between institutes.) We show in Appendix VII.A that the basic model can be extended in this fashion. In this case, we interpret $\pi_{ij}^t$ as the *aggregate* precision individuals in group $i$ perceive for group $j$. In the special case where individuals have equal precision, we can therefore interpret $\pi_{ij}^t$ as proportional to the size of group $j$.

**Random Matching:** We also extend the model in Appendix VII.A to the case in which individuals are randomly matched with other individuals in the population each period. In this case, we show that expected beliefs can also be described by equation (5), where we now interpret $\pi_{ij}^t$ as the probability that $i$ listens to $j$ each period.

**A Graphical Characterization of Influence Weights:** Here we demonstrated that the influence weights $w$ can be interpreted as the stationary distribution associated with the listening matrix $T$. We show in Appendix VII.B an alternative method of characterizing the influence weights using directly the graph of the listening structure and "counting trees" consistent with this structure for which each agent is at the root. The advantage of this approach is it relates social influence directly to aspects of an individual's position in the directed graph.

**Non-Strongly Connected Listening Structures:** Our analysis has concentrated on the case in which the set of agents is strongly connected. In Appendix VII.C, we consider the case in which some groups of agents are isolated and do not listen to agents outside their own group. We show that the beliefs of agents in isolated groups converge, as in Section III.A, but with different isolated groups converging to different beliefs. The beliefs of agents who are not in an isolated group lie within the convex hull of the beliefs of the isolated groups. This setting may be applicable to situations where there are isolated groups of "extremists," with the general population distributed in between.

# IV Bilateral Communication and Neighborhoods

One natural listening structure in many applications is for individuals to listen to their neighbors. Depending on the application, neighborhoods can be defined by one of many criteria, such as geographical proximity, cultural proximity, social interaction, etc. In such settings, it is natural to assume that communication is bilateral: if an agent $i$ listens to an agent $j$ (because $j$ is $i$'s neighbor), then $j$ also listens to $i$ (because, by symmetry, $i$ is $j$'s neighbor), i.e., $q_{ij} = q_{ji}$ for all $i, j$.

We consider two natural cases regarding agents' assessments of precision. First, the *common precision* case, where agents agree on their assessments, i.e., $\pi^0_{ij} = \pi^0_j$ for all $i, j$. (Agents' common assessments may further coincide with the correct precision, in which case $\pi^0_j = \hat{\pi}^0_j$ for all $j$.) Second, the *symmetric precision* case, where agents can disagree on their assessments, but assessments are symmetric in the sense that $i$'s assessment of $j$'s precision equals $j$'s assessment of $i$'s precision, i.e., $\pi^0_{ij} = \pi^0_{ji}$ for all $i, j$. The symmetric precision case corresponds to settings where agents use different "models" of the world, and view those who share their model as having more accurate information. For example, traders relying on technical analysis (chartists) may attribute a high precision $\overline{\pi}^0$ to other chartists and a low precision $\underline{\pi}^0$ to traders analyzing fundamentals. Conversely, fundamentals analysts may attribute the high precision $\overline{\pi}^0$ to other such analysts and the low precision $\underline{\pi}^0$ to chartists.[32] We should note that both the common and the symmetric precision cases include the equal precision case ($\pi^0_{ij} = \pi^0$ for all $i, j$) as a special case.

## IV.A Social Influence in Neighborhoods

Under bilateral communication, the social influence weights take a particularly simple form.

**Theorem 6** *Suppose communication is bilateral. Then:*

*a. If $\pi^0_{ij} = \pi^0_j$ for all $i, j$ (common precision), then* $\frac{w_i}{w_j} = \frac{\pi^0_i}{\pi^0_j} \frac{\sum_k q_{ki} \pi^0_k}{\sum_k q_{kj} \pi^0_k}$,

*b. If $\pi^0_{ij} = \pi^0_{ji}$ for all $i, j$ (symmetric precision), then* $\frac{w_i}{w_j} = \frac{\sum_k q_{ki} \pi^0_{ki}}{\sum_k q_{kj} \pi^0_{kj}}$.

*In the equal precision case, both equations reduce to* $\frac{w_i}{w_j} = \frac{\#S(i)}{\#S(j)}$.

Thus, in the equal precision case, agent $i$'s relative social influence is simply given by the cardinality of $i$'s listening set, i.e., the number of $i$'s neighbors (where $i$ is also counted as a neighbor). Consequently, agent $i$ is influential if he has many neighbors. Intuitively, those with many neighbors are listened to by many agents. Somewhat surprisingly, however, $i$'s influence only depends on the

[32] The symmetric precision case also corresponds to the random matching model, sketched in Subsection III.D. Indeed, in that model, the assessments of precision are the matching probabilities, which are by definition symmetric.

number of his neighbors, and not on these neighbors' own influence. This is because there are two effects which exactly offset. On one hand, being listened to by an influential neighbor $j$, tends to increase $i$'s influence, holding the weight that $j$ gives to $i$ constant. On the other hand, if $j$ is influential (i.e., has many neighbors), then he listens to many agents, lessening the weight given to any one neighbor, such as $i$.

More generally, in the common precision case, agent $i$'s relative social influence is given by the product of $i$'s precision, times the aggregate precision of $i$'s neighbors. Accuracy increases $i$'s influence since $i$ is then given more weight by his neighbors. Similarly, having accurate neighbors increases $i$'s influence since accurate neighbors are also more influential. In the symmetric precision case, agent $i$'s relative social influence is given by the sum across neighbors of the precision these neighbors attribute to $i$.

An interesting feature of Theorem 6 is that an agent's relative social influence depends only on simple characteristics of the agent's *local* neighborhood (such as the number and precision of the agent's direct neighbors, and the precision these direct neighbors assign to the agent). This contrasts with the general case where social influence can depend in a rather complex manner on the entire listening structure.

## IV.B Long-Run Differences of Opinion in Neighborhoods

We next turn to agents' long-run differences in opinion. Bilateral communication imposes strong restrictions on the eigenvalues and eigenvectors of $\mathbf{T}$.

**Theorem 7** *Suppose communication is bilateral, and $\pi^0_{ij} = \pi^0_j$ for all $i, j$ (common precision) or $\pi^0_{ij} = \pi^0_{ji}$ for all $i, j$ (symmetric precision). Then the listening matrix $\mathbf{T}$ is diagonalizable, and all its eigenvalues are real. Furthermore, if $(v_1, \dots, v_N)'$ is a column eigenvector of $\mathbf{T}$, then a row eigenvector corresponding to the same eigenvalue is $(w_1 v_1, \dots, w_N v_N)$.*[33]

Since the eigenvalues of $\mathbf{T}$ are real, we can apply Theorem 4. Therefore, in the long run, agents hold the same relative positions on all issues. The agents holding the most extreme positions are those corresponding to the extreme components of the second column eigenvector. On the other hand, the agents whose initial beliefs contribute the most to the long-run disagreement are those corresponding to the extreme components of the second row eigenvector. It is natural to think that such agents will be those who have extreme views, and those who are influential. Theorem 7 shows that these two effects interact in a very simple manner. In particular, the component of the second row eigenvector associated with each agent is simply the product of the corresponding component of the second column eigenvector and the agent's social influence $w_i$.

[33] The results in Theorem 7 have a counterpart in the theory of Markov chains. Under bilateral communication, and common or symmetric precision, the Markov chain associated to $\mathbf{T}$ is *reversible*, i.e., satisfies the *detailed balance equations*: $w_i \mathbf{T}_{ij} = w_j \mathbf{T}_{ji}$ for all $i, j$. In our context, the detailed balance equations mean that $i$'s social influence times the direct influence of $j$ on $i$ (which can be interpreted as the social influence $j$ obtains "through" $i$) is equal to $j$'s social influence times the direct influence of $i$ on $j$. The results in Theorem 7 follow from this reversibility property.

To illustrate our results, we consider an example where agents are located on a two-dimensional space, listen to those within a given radius around themselves, and have equal precision. In Figure 3 we plot the two-dimensional "map" corresponding to the listening structure, indicating each agent by a node, and each pair of agents that listen to each other by an edge. Note that the influence weights are proportional to the size of each agent's listening set (number of edges plus the implicit edge from an agent to himself), as in Theorem 6. Note also that the long-run positions, $V_{i2}^c$, are correlated across neighbors, and the agents with the most extreme long-run views are far apart from each other in the graph.

### IV.C A Special Case – Agents Located on a Line

A simple and particularly appealing listening structure involving bilateral communication is when agents are located on a line and listen to their immediate neighbors. In this subsection we obtain specific results for this listening structure. For simplicity, we take all agents to have equal precision. We label the agent located in the middle of the line as agent 0, and the agent located $i$ positions to the right (left) as agent $i$ (agent $-i$). In addition, we set $N = 2M + 1$, so that the agents at the extremes are $M$ and $-M$.

A word of caution is in order here, as there are two distinct orderings of agents that are relevant in this subsection. As elsewhere in this paper, we take the listening structure to be fixed, and consequently agents' labels (i.e., their positions on the line which determines the listening structure) are fixed. We make no assumption here about the relationship between this "geographical" ordering and agents' initial beliefs. Thus, neighbors can start with very different beliefs. Rather strikingly, we show in Theorem 8 that the geographical ordering, and not initial beliefs, entirely determines the ordering of beliefs in the long run.

**Theorem 8** *Suppose that agents are located on a line, and listen to their immediate neighbors (and themselves). Suppose also* $\pi_{ij}^0 = \pi^0$ *for all* $i, j$ *(equal precision). Define* $\theta_2$ *as the unique solution in* $[0, \pi)$ *to*

$$M\theta_2 + \arctan\left[\frac{1}{3}\tan\left[\frac{\theta_2}{2}\right]\right] = \frac{\pi}{2}.$$

*(In particular,* $\frac{\pi}{2M+1} < \theta_2 < \frac{\pi}{2M}$*.) Then, the i'th component of* $V_2^c$ *is proportional to* $v_i^c = \sin[i\theta_2]$*, while the i'th component of* $V_2^r$ *is proportional to* $v_i^r$*, defined by* $v_i^r = 2\sin[i\theta_2]$ *for* $i = \pm M$*, and* $v_i^r = 3\sin[i\theta_2]$ *otherwise. When* $M \geq 2$*, the extreme values of* $v_i^r$ *are for* $i = \pm(M-1)$*.*

Since $\theta_2 < \frac{\pi}{2M}$, $v_i^c$ is increasing in $i$, and thus agents' geographical ordering does indeed coincide with the ordering of their positions in the long run. Figure 4 plots $v_i^c$ as a function of $i$. In addition to monotonicity, this figure indicates that the distance between agents' long-run positions is smallest at the extremes and largest at the center. In other words, the distance between agents $M$ and $M-1$ is smaller than between $M-1$ and $M-2$, which is smaller than between $M-2$ and $M-3$, and so on. Intuitively, agents at the extremes are "pulled" only in one direction, and thus their beliefs become close to those of their neighbors quite quickly. Likewise, agents close to the extremes are pulled mainly in one direction, since the beliefs of their more extreme neighbors are close to theirs.

The result that distance is smallest at the extremes can also be expressed in terms of the density of agents holding a given position. We plot this density in Figure 5. As shown in the figure, the density is largest at the extremes, since agents' positions are the closest. Thus, independent of the initial density, the long-run density exhibits a form of polarization.

According to Theorem 8, the agents whose difference in initial beliefs matters the most for long-run disagreement are $M-1$ and $-(M-1)$, i.e., the agents next closest to the extremes. This is because these agents are not only close to the extremes, but are also the only agents that those at the extremes are listening to.

Finally, it is interesting to examine the speed of convergence in this example. Recall from Theorem 5 that the speed at which beliefs converge to consensus is of order $f_{\lambda^*}(\alpha_2)^{\sum_{s=0}^{t-1}\lambda_s}$, while the speed at which differences of opinion become unidimensional is of order $(f_{\lambda^*}(\alpha_3)/f_{\lambda^*}(\alpha_2))^{\sum_{s=0}^{t-1}\lambda_s}$, where $\alpha_2$ and $\alpha_3$ are the second and third largest eigenvalues of $\mathbf{T}$. For large $M$, these eigenvalues are given by

$$\alpha_2 \approx 1 - \frac{\pi^2}{12M^2} \qquad \text{and} \qquad \alpha_3 \approx 1 - \frac{\pi^2}{3M^2}, \tag{8}$$

respectively.[34] Note also that for $\delta$ close to zero, $f_\lambda(1-k\delta) \approx [f_\lambda(1-\delta)]^k$. Together with (8), this implies that the speed of convergence of both beliefs and differences of opinion decreases with $M^2$; that is, doubling the length of the line causes the speed of convergence to slow by a factor of 4. Similarly, we can also compare the relative speeds of convergence. Because $f_{\lambda^*}(\alpha_3)/f_{\lambda^*}(\alpha_2) \approx [f_{\lambda^*}(\alpha_2)]^3$, the speed at which differences of opinion become unidimensional is three times faster than the speed at which beliefs converge to consensus.

# V Additional Applications

In this section, we explore more informally a number of other applications of our model. For concreteness, we pick and discuss in detail one area where persuasive activity is particularly notable: that of political discourse. We then briefly discuss applications to a number of other areas.

One important caveat is in order here. Our model presumes that agents report their information truthfully. This assumption might describe well communication between friends and colleagues sharing opinions, and some of the implications discussed below are consistent with this setting. For some other implications, however (e.g., propaganda, political spin, political campaigns, etc.), this assumption is not reasonable, as persuaders might have a lot to gain from influencing listeners, and therefore from misreporting their information. The extent of misreporting might, of course, be limited by reputational or legal concerns (e.g., politicians might be punished by the electorate for telling lies, false advertising might be restricted, etc.).

Persuasion bias could have an effect even in settings with misreporting. For example, listeners

[34] This follows from equations (23) and (28), by setting $m = 1$ and $m = 2$.

could be fully aware that only arguments in support of an issue will be presented. However, if they do not adjust for repetitions, they could be swayed by the number of arguments, not fully accounting for the fact that similar arguments are repeated.

In addition to misreporting, our notion of persuasion bias implies that strategic persuaders would have an incentive to repeat their information frequently, especially to those that many others listen to. This seems very realistic: it is hard to think of a setting where agents who would benefit from persuasion do not behave in this manner. As just one notable example, candidates for public office campaign for months or even years, with a message that can be clearly communicated in hours.[35]

## V.A Political Discourse

We begin by noting that our notion of persuasion bias yields an immediate explanation for several widespread political phenomena such as propaganda, censorship, political spin, and the importance of air-time. We then discuss a number of deeper implications and novel empirical predictions that follow from applying our analysis and results to various political science questions.

**Propaganda, Censorship, Political Spin and Air-Time:** Propaganda and censorship seem to depend critically on the notion that repeated exposures to a particular message have a cumulative effect on a listener over time. Indeed, one defining feature of propaganda campaigns is the frequent repetition of a message – i.e., the intent is for listeners to hear the message many times. Similarly, censorship often seems motivated by the notion that frequent exposure to an opposing opinion will sway beliefs in that direction, and therefore such exposure should be restricted.[36]

The phenomenon of political spin seems closely related. Political strategists often seem to be concerned with pushing a common message repeatedly. Political spokespersons coordinate on common lists of "talking-points" which they repeat in the media, seemingly under the belief that such repetition will give their viewpoint the most effective "spin."[37]

More generally, "air-time" is considered to have an important effect on beliefs. For example, a political debate without equal time for both sides, or a criminal trial in which the defense was given less time to present its case than the prosecution, would generally be considered biased and unfair. Indeed, common concerns about inequitable political fund-raising seem to be motivated by concerns

[35] The interaction between strategic behavior and persuasion bias yields a number of interesting questions for future research. One question, for example, concerns the competition between strategic political parties when voters are subject to persuasion bias.

[36] A distinction should be drawn here between the censor who believes the censored opinions are inaccurate, and the censor who believes they are accurate (but personally undesirable). Without a notion of persuasion, it is hard to understand censorship of the former type. Indeed, absent persuasion bias, arguments for incorrect opinions should, on average, lessen the appeal of these opinions. While, in contrast, persuasion bias is not necessary to understand the activity of a censor trying to silence opinions he believes are in fact correct, it seems clear that many censors do believe the views they are silencing are harmful and incorrect.

[37] Closely related are social psychology literatures on the efficacy of priming and agenda setting through the media in political campaigns. (See, for example, Iyengar and Kinder [1987] on priming and Cohen [1963] and Ehbring, Goldenberg and Miller [1980] on agenda setting.) Related to our model, Huckfeldt and Sprague [1995] find evidence that the specific form of social networks play an important role in influencing political beliefs. In particular, they find that frequent political conversation partners have an important effect on an individual's political beliefs after controlling for natural demographic and personal characteristics.

about the unequal air-time that the funds will subsequently provide. Note that this appears to be true even when the message is simple, and could seemingly be conveyed rather quickly, and when it has already been widely heard by the electorate. Here as well, such concerns can be easily understood under the presumption that repetition unduly affects beliefs.

**Unidimensional Opinions:** Perhaps the most notable implication of our model for political science is that according to Theorem 4, long-run differences of opinion should be "unidimensional."[38] In particular, representing an agent's beliefs over $L$ issues by a point in $\mathbb{R}^L$, Theorem 4 indicates that after "enough" updating, the beliefs of all agents will line up along a single line. Consequently, an agent's beliefs over all issues will be characterizable by a single scalar measure; for example, how far to the "left" or "right" he is. It is worth emphasizing that unidimensional long-run beliefs is a very general result of our model that follows without imposing any unidimensionality on the underlying listening structure or initial beliefs: i.e., it follows generally for any multidimensional initial beliefs and any strongly connected listening structure.[39] We are unaware of any other model of political belief formation or dynamics where unidimensional beliefs follow so generally.[40]

Our linearity result is especially notable given the attention the unidimensionality issue has received in political science. For example, a large literature has examined whether the voting records of Congress and Senate members can be explained by a unidimensional "liberal-conservative" model, and has found strong support for such a model.[41] The unidimensionality assumption is also prominent in theoretical political-science modeling. For example, the standard median-voter result relies on unidimensional political views, and does not easily generalize to multiple dimensions.[42]

One reason, undoubtedly, that unidimensional opinions have received so much attention in the political science literature is that such characterizations appear to be the norm in informal political discourse as well. That is, individuals' political views are typically depicted by the extent to which they are liberal or conservative. And while such unidimensional depictions might, at times, be a matter of convenience or simplicity, they frequently convey quite well what the individual believes

[38] We believe the focus on long-run differences of opinion, rather than on convergent beliefs, is more appropriate for political applications, in light of our discussion in footnote 23 above. That is, lacking a natural metric to measure the distance between political beliefs, or to compare across different issues, makes empirical tests of convergence problematic. In contrast, however, our unidimensionality result yields ordinal predictions about individuals' beliefs (i.e., the order of individuals will be identical across issues), which are testable without any measurement of distance between beliefs.

[39] This is, of course, subject to the second eigenvalue being real and not identical to the third. In many applications (such as settings where communication is bilateral) the second eigenvalue is real. Additionally, ties among real eigenvalues are non-generic.

[40] One theoretical model that purports to obtain unidimensional beliefs is Spector's [2000] cheap-talk model of opinion formation. However, with only two "types" in his model, beliefs by definition must be linear across the two types. Rather, he finds that the ultimate difference of opinion between his two types will be over a single issue. By contrast, our result is that the beliefs of many different agents will all line up along a line, where the line will generally be characterized by disagreements over all issues.

[41] See, for example, Converse [1964], Poole and Rosenthal [1991,1997], and the references therein.

[42] For the median voter result see, for example, Varian [1992], and for the multidimensional case see Caplin and Nalebuff [1991]. We should note that much of the political science literature has focused on preferences rather than beliefs. While we acknowledge this difference, and recognize that preferences may be more fundamental to political actions than beliefs, there are times when they will coincide. For example, if preferences over *outcomes* are similar, preferences over *policies* will coincide with beliefs about the outcomes that policies will induce.

about a wide range of issues.

**Grouping of Opinions:** An alternative explanation for unidimensional opinions is that all issues are naturally related (or share important common components), and therefore the issue space is effectively unidimensional. A comparison of this view with our notion, that unidimensional beliefs follow from the particular manner of social updating, reveals several implications that our model has for the grouping of opinions.

First, in contrast with the alternative view, our model suggests that many different opinions on unrelated issues will be grouped together in a common set of opinions. Consider for example the following set of contentious issues in American politics: free trade, abortion, flag-burning, the death penalty, school vouchers, and environmental regulation. To us, at least, debate on these issues seem to involve very different concepts and arguments. Nonetheless, it seems striking how far one can go in describing many individuals' beliefs across all these issues with a simple liberal/conservative characterization.

Furthermore, the two alternative explanations of unidimensional beliefs yield different implications for the consistency of groupings of opinions. If opinions are unidimensional because issues are inherently related, the set of opinions grouped together and associated with conservatives or liberals should be consistent over time and location. By contrast, under persuasion bias, the set of opinions grouped together will depend on the initial beliefs and the listening structure. Thus, changes in initial beliefs, or in the social network along which people communicate, could lead to different groupings of opinions. Notably, opinions that are considered "conservative" in some countries and at certain times are considered "liberal" elsewhere and at different times. For example, free trade has passed from a liberal to a conservative opinion in the United States over the last century. Similarly, favoring free trade is considered a liberal opinions in some countries (such as China and Vietnam), while in contrast, it is considered conservative in others (such as the United States and Mexico).[43]

Notably, such grouping of opinions yields a natural explanation for the existence and stability of political parties. Indeed, if individuals' opinions were instead independent across different issues, it is unclear how (or why) parties would stake out a position on all important issues. Furthermore, one would expect to see individuals varying their allegiance regularly from party to party, depending on the issue of the day. Indeed, the presence of individuals with similar beliefs across a wide range of issues, seems a critical condition for the existence of political parties. Persuasion bias, in turn, would then give individuals a motivation to support such parties, which could serve as an efficient mechanism for repeating opinions so as to convince others.[44]

[43] This is not simply a matter of different political labelings in different countries, but rather, indicative of different sets of opinions that seem to be grouped together. For example, in all these countries, being in favor of a strong military is considered a conservative opinion. Thus, in China and Vietnam, those in favor of a strong military are likely to be opposed to free trade, while in the United States and Mexico, the converse is likely to be true.

[44] An alternative explanation for the grouping of opinions is that individuals all choose to identify with the views of a combination of pre-existing political parties. It is not clear why individuals would do so, however, absent some notion of persuasion. Furthermore, such an explanation would not predict unidimensional beliefs across a population if there was more than two political parties, nor would it explain why many individuals have views more extreme on some issues than the positions of the major parties in their country.

**Demographic Explanations of Political Opinion:** Our results on neighborhood structures in Section IV imply that agents who are "near" one another in a social network are likely to acquire similar beliefs over time. Thus, to the extent that certain demographic characteristics such as age, race, wealth, profession, neighborhood, religion, etc., affect social interactions, we would expect these same characteristics to determine where on the political line one's views are located. Conversely, demographic characteristics that are not an important basis for social interaction would not be related to political beliefs. Insofar as the demographic characteristics that determine social interaction vary across places and time, this observation yields testable variation in determinants of political beliefs. For example, our model predicts that race or religion would play a larger role in determining someone's political beliefs in less racially or religiously integrated places and/or times.

**Closed Groups:** If there are groups of individuals who confine themselves to communicating with one another, their beliefs will converge faster to a common group belief, which will be very hard for outsiders to impact. Thus, political opinions should be more uniform in countries with fewer ties to outsiders than in countries where visitors are common. Similarly, opinions should be more uniform in small isolated towns than in cities. Or as an extreme example, the model would predict a high degree of conformity among closed religious communities and among religious cults.[45] Indeed, typical depictions of brainwashing techniques employed by cults – frequent bombardment with the group message, the forced severing of ties with outsiders, repeated group reinforcement of the message over time – can be seen as a straightforward manifestation of our notion of persuasion.

**Unidimensional Views and Media Coverage:** Since the speed of convergence depends on the frequency of communication, unidimensionality of beliefs should be more pronounced for important or timely issues that are discussed more than for trivial or non-topical issues. It seems reasonable to believe that media attention is likely to be related to the amount an issue is discussed.[46] Consequently, a novel prediction of our model is that the degree of media attention that an issue receives will be positively related to the degree that individuals' views on the issue are predictable from their general political convictions. Or, equivalently, media attention to topics will be associated with more polarization along standard party lines.

**Summary:** Our model yields a number of novel political science predictions. It predicts unidimensional long-run beliefs, with adherence to this pattern being greater for issues that are frequently discussed (which we posit is related to media coverage). Groupings of political issues along the line that defines unidimensional beliefs will depend on the social network, and consequently, is likely to be more similar across times and places with similar social interaction. "Social neighbors" will evolve to have similar views; and consequently, the importance of different demographic factors in explaining individuals' political orientation will be related to the importance of such factors in determining social interaction. Insofar as the latter is likely to vary over time and place, the model

[45] An example in line with this interpretation is the voting of certain Hasidic Jewish communities in New York. For example, one such community, New Square, voted 1400-12 in favor of Hillary Clinton (the Democrat) over Rick Lazio (the Republican) in the 2000 Senate race. Nor can this be attributable to a natural association of Hasidism with Democrats. On the contrary, other Hasidic communities strongly favored Lazio, and Alphonse D'Amato (the Republican) carried New Square by a similar margin in 1998 over Charles Schumer (the Democrat).

[46] This might be true either because the media can create public interest, or because it reflects public interest.

predicts demographic factors will vary in their importance in determining political beliefs. Additionally, communities that interact less with outsiders are likely to have homogeneous views that are hard for outsiders to change.

### V.B Other Applications - A Brief Overview

Here we briefly discuss applications to two other settings: court trials and marketing.

**Court Trials:** Persuasion bias justifies the common notion that air-time is important in a trial. Trials where only one side gets to present evidence, or where one side has a far greater opportunity to speak, are universally condemned as unfair. Additionally, our model predicts that disagreements within a deliberative body (such as a jury) about different pieces of evidence or different arguments are likely to be unidimensional after sufficient deliberations, but not beforehand. Thus, the members of a jury are likely to ultimately separate into two camps - "left" and "right" of the consensus - with those in the same camp being on the same side on the interpretation of all pieces of evidence.[47]

**Marketing:** Persuasion bias seems central to marketing activity. In particular, many marketing campaigns seem to be designed with the idea that repeated exposures to an advertisement will have a cumulative effect on consumer preferences. Additionally, our model predicts that consumer preferences should exhibit closer adherence to unidimensionality over goods that are discussed more frequently. Thus, if some goods ("social goods") are more enjoyable to discuss than others (e.g., discussing music, restaurants, and fashion is more enjoyable than discussing insurance, cleaning products, and hardware), one should observe more predictability in individuals' tastes for these goods. Additionally, the more related a demographic factor is to social networks, the more it should predict tastes; and such demographic factors should predict tastes better for social goods. For example, an individual's social circle is likely to be a better predictor of his taste in music and restaurants than his taste in insurance or cleaning products.

## VI Conclusion

We propose a boundedly rational model of opinion formation in which individuals are subject to persuasion bias. Specifically, individuals fail to adjust for possible repetitions of (or common sources in) information they receive. Persuasion bias can be viewed as a simple, boundedly rational heuristic, arising from the complexity of recounting the source of all the information that has played a role in forming one's beliefs. We argue that it provides an explanation for several important phenomena, such as the effectiveness of air-time, propaganda, censorship, political spin, marketing, etc. These phenomena suggest that persuasion bias is ubiquitous, and plays an important role in the process of social opinion formation.

[47] In line with our notion of persuasion, juries frequently deliberate for an extended period before reaching a consensus. Judges often instruct deadlocked juries to continue deliberating, despite the absence of new information, and such juries frequently do reach a consensus.

We explore the implications of persuasion bias in a model where individuals communicate according to a social network. We show that persuasion bias implies the phenomenon of social influence, whereby an individual's influence on group opinions depends not only on accuracy, but also on how well-connected the individual is in the network. Persuasion bias implies the additional phenomenon of unidimensional opinions, because individuals' opinions over a multidimensional set of issues converge to a single "left-right" spectrum.

We apply our model to several natural settings, including neighborhoods with bilateral communication, political discourse, court trials, and marketing. We argue that our model provides a useful way for thinking about opinion formation in these settings, and has a number of novel empirically testable predictions. For example, our model predicts that political beliefs should be unidimensional, with conformity to this pattern being greater for issues that receive media coverage. Additionally, the importance of different demographic factors in explaining political beliefs should depend on the role these factors play in social interaction.

An important extension of our model is to endogenize the listening structure. In some settings, for example, agents might believe that those expressing similar opinions to their own are more accurate, and might choose to listen only to such individuals. In a companion paper (DeMarzo, Vayanos, and Zwiebel [2001]), we consider persuasion and endogenous listening in an asset market setting. Agents communicate information about an asset across a social network, and trade on their information. Within any "trading episode" we hold the network fixed, but across trading episodes we allow agents to change who they listen to based on the trading profits they realize. Among our results, we characterize which listening structures are "stable" in the sense that no agent would choose to alter the agents he listens to given the listening choices of other agents. We find that provided that there is a sufficient amount of updating, stable listening structures generally take the form where all agents listen to the same "expert," or the same set of "experts." Furthermore, the experts listened to need not be the ones with the best information, and consequently, influence can be self-perpetuating independently of accuracy. We believe the framework provided by this model is well suited for other applications where the listening structure might be endogenous.

Finally, our model of persuasion is well suited for experimental testing. The model yields a number of unique predictions relating the listening structure to dynamic properties of beliefs, e.g., to social influence weights and agents' positions on the line of long-run disagreement. Although the precise listening structure might be difficult to measure in many empirical settings, it would be easy to create and modify in an experimental setting. In such a setting it would be possible to test many of the specific predictions of our model, as well as more "general" predictions, such as whether beliefs lie along a line in the long run. It is worth noting that the quantitative nature of our model's predictions goes well beyond that of many behavioral models. Consequently, we believe that testing some of these predictions experimentally would be quite interesting and would provide for a strong test of our framework. But perhaps this is because we have listened to one another too much, and have therefore been unduly persuaded about the merits of our ideas.

# Appendix

## VII Generalizations of the Model

### VII.A Groups and Random Matching

To model groups, suppose that there exists a partition $\mathcal{P}$ of the set $\mathcal{N}$ of agents, such that for all $I \in \mathcal{P}$, and all $i, i' \in I$, $j \in \mathcal{N}$, we have

a. The agents in $I$ employ equal listening weights: $\mathbf{T}_{ij} = \mathbf{T}_{i'j}$,

b. The agents in $I$ receive equal relative listening weight: either $(q_{ji} = q_{ji'} = 0)$ or $(q_{ji} = q_{ji'} = 1$ and $\mathbf{T}_{ji}/\mathbf{T}_{ji'}$ is independent of $j)$.

Then we can aggregate each group $I \in \mathcal{P}$ into a single representative agent as follows. Define a new listening matrix by $\mathbf{T}_{IJ} = \sum_{j \in J} \mathbf{T}_{ij}$ for $I, J \in \mathcal{P}$, and $i \in I$ arbitrary by condition (a) above. Also define the beliefs of each representative agent as the weighted average beliefs of the group,

$$x_I^t = \frac{\sum_{i \in I} \pi_{ji}^0 x_i^t}{\sum_{i \in I} \pi_{ji}^0},$$

for $j \in \mathcal{N}$ arbitrary (such that $q_{ji} = 1$) by condition (b) above. Then it is easy to check that the updating rule (3) holds for the representative agents as well. Moreover, $x_i^t = x_I^t$ for $i \in I$ and $t \geq 1$; that is, after one round of updating all agents within any given group have the same beliefs.

The importance of this observation is that it allows us to reinterpret any agent $j$ in a listening structure as representing a group of agents rather than one individual agent. The relationship $\mathbf{T}_{IJ} = \sum_{j \in J} \mathbf{T}_{ij}$ then is equivalent to interpreting $\pi_{ij}^0$ as the aggregate perceived precision of the group of agents represented by $j$. In the special case where individual group members have equal precision, this implies that $\pi_{ij}^0$ is proportional to the size of the group represented by $j$.

It is also possible to reinterpret our model in a setting where communication occurs through a random matching process. Suppose that each agent $i$ listens to only one agent $j \in S(i) \backslash \{i\}$, chosen at random. Let $\pi_{ii}^0$ be the *relative* precision that agent $i$ believes his information has compared to any other agent $j$. Then, if $i$ is matched with $j$,

$$x_i^1 = \frac{\pi_{ii}^0 x_i^0 + x_j^0}{\pi_{ii}^0 + 1}.$$

Denoting by $p_{ij}$ the probability that $i$ is matched with $j$, so that $\sum_{j \neq i} q_{ij} p_{ij} = 1$, we have

$$E[x^1] = \mathbf{T} x^0.$$

where $\mathbf{T}$ is the same matrix as in equation (2) but with $p_{ij} = \pi_{ij}^0$. This means that in the random matching model, the matrix $\mathbf{T}$ describes the evolution of *expected* beliefs, and the precisions can be interpreted as the matching probabilities. To make the matching model deterministic, one could further assume that each agent is a "type," consisting of a continuum of "identical" agents (as in the group structure discussed above). Then, the uncertainty of the random matching process is eliminated in the aggregate, and equation (3) describes the deterministic evolution of the average beliefs of each type.

## VII.B A Graphical Characterization of Influence Weights

The social influence weights $w$ can also be characterized in a way which is more directly related to the underlying listening structure. This is done by using a "tree-counting" method. Define a spanning tree $G$ as a directed graph connecting all agents in $\mathcal{N}$, where every agent save for one has a unique predecessor, there are no cycles, and one agent (the root) has no predecessors. The edge $(i, j) \in G$ if $i$ is the unique predecessor of $j$ in the tree. Define the influence score $\rho$ of the tree by

$$\rho(G) = \prod_{(i,j)\in G} \mathbf{T}_{ji};$$

that is, the product of the direct influence that flows from the root of the tree. Let $\mathcal{G}_i$ be the set of spanning trees with agent $i$ at the root and $\rho(G) > 0$. Then, the social influence weights can be characterized as follows:

**Theorem 9** *Suppose $\mathcal{N}$ is strongly connected, and Assumption 1 holds. Then the social influence weights satisfy*

$$\frac{w_i}{w_j} = \frac{\sum_{G\in\mathcal{G}_i} \rho(G)}{\sum_{G\in\mathcal{G}_j} \rho(G)}.$$

That is, the relative social influence of agent $i$ is given by summing over all spanning trees with $i$ at the root, the product of the direct influence within each tree. Theorem 9 follows from a classic result in Markov chain theory, sometimes referred to as the "Markov Chain Tree Formula."[48] In the special case of *equal precision* ($\pi^0_{ij} = \pi^0$ for all $i, j$), Theorem 9 simplifies as follows:

**Corollary 2** *Suppose the assumptions to Theorem 9 hold, and $\pi^0_{ij} = \pi^0$ for all $i, j$. Then the social influence weights satisfy*

$$\frac{w_i}{w_j} = \frac{\#\mathcal{G}_i \,\#S(i)}{\#\mathcal{G}_j \,\#S(j)}.$$

That is, to determine the relative social influence of agent $i$, we simply need to count the number of spanning trees with $i$ at the root and the property that each agent listens to his predecessor, and multiply this by the number of agents that $i$ listens to.

To illustrate the tree-counting method, consider the listening structure of Example 1. The vector corresponding to the number of spanning trees is $(8, 5, 2, 1)$. (The spanning trees for agents 3 and 4 are shown in Figure 6.) The vector corresponding to the cardinality of an agent's listening set is $(2, 3, 4, 3)$. Corollary 2 implies that the relative influence weights are $(16, 15, 8, 3)$. These are equal to the weights found in Subsection III.A, after normalization.

## VII.C Nonstrongly Connected Listening Structures

Our analysis can be extended to general listening structures, where agents may not necessarily be strongly connected. For a general listening structure, define a bilateral relationship $\mathcal{R}$ on $\mathcal{N}$ by

$$i \,\mathcal{R}\, j \Leftrightarrow \{i, j\} \text{ is strongly connected.}$$

The relationship $\mathcal{R}$ is reflexive, symmetric, and transitive. Therefore, it defines a partition of $\mathcal{N}$ into equivalence classes. We denote this partition by $\mathcal{P}$. The definition of $\mathcal{R}$ ensures that each element of $\mathcal{P}$ is strongly connected.

[48] See, for example, Aldous and Fill [1999, Ch. 9] and the references therein.

Suppose that for an element $I \in \mathcal{P}$, no agent in $I$ listens to agents outside $I$. Then the dynamics of beliefs of the agents in $I$ can be studied in isolation. This motivates the following definition:

**Definition 3** A set $I \in \mathcal{P}$ *is linked to* $J \in \mathcal{P}$ if there exists $i \in I$ and $j \in J$ such that $j \in S^*(i)$ (i.e., $j$ influences $i$). A set $I \in \mathcal{P}$ is *isolated* if it is not linked to any $J \neq I$.

Consider an isolated set $I \in \mathcal{P}$. Since the dynamics of beliefs of the agents in $I$ can be studied in isolation, and since $I$ is strongly connected, Theorem 1 applies. All agents in $I$ converge to a common belief, which is a weighted average of their initial beliefs. We denote the vector of initial beliefs restricted to $I$ by $x^0(I)$, and the social influence weights by $w(I)$.

Consider next a non-isolated set $J \in \mathcal{P}$. If $J$ is linked to only one isolated set $I$, then all agents in $J$ will converge to the limit beliefs of the agents in $I$, since they keep being "persuaded" by those agents. If, by contrast, $J$ is linked to multiple isolated sets, then the agents in $J$ will converge to a point in the convex hull of the limit beliefs corresponding to the isolated sets. Different agents in $J$ may, however, converge to different points in the convex hull, because they may be giving different weights to the agents in the isolated sets. Theorem 10 makes these intuitions precise:

**Theorem 10** *The collection $\mathcal{I}$ of isolated elements of $\mathcal{P}$ is nonempty, and each non-isolated element of $\mathcal{P}$ is linked to at least one isolated element. Furthermore:*

*a. If $I \in \mathcal{I}$, then the beliefs of all agents in $I$ converge to $w(I)x^0(I)$.*

*b. If $J \in \mathcal{P}\backslash\mathcal{I}$ is linked to a unique $I \in \mathcal{I}$, then the beliefs of all agents in $J$ converge to $w(I)x^0(I)$.*

*c. If $J \in \mathcal{P}\backslash\mathcal{I}$ is linked to multiple $I \in \mathcal{I}$, then the beliefs of each agent in $J$ converge to a point in the interior of the convex hull of the beliefs $w(I)x^0(I)$, for all such $I$.*

We illustrate Theorem 10 with the following example:

**Example 2** Suppose that $N = 5$, $S(1) = \{1\}$, $S(2) = \{1, 2, 3\}$, $S(3) = \{2, 3, 4\}$, $S(4) = S(5) = \{4, 5\}$, and agents attribute equal precision to those they listen to. Then $\mathcal{P} = \{\{1\}, \{2, 3\}, \{4, 5\}\}$, and the sets $\{1\}$ and $\{4, 5\}$ are isolated. The beliefs of agent 1 converge to $x_1^0$, and the beliefs of agents 4 and 5 converge to $(x_4^0 + x_5^0)/2$. Furthermore, for the non-isolated set $\{2, 3\}$, beliefs do not converge to the same point. The beliefs of agent 2 converge to $(2/3)x_1^0 + (1/6)x_4^0 + (1/6)x_5^0$, and those of agent 3 converge to $(1/3)x_1^0 + (1/3)x_4^0 + (1/3)x_5^0$. Agent 2 ends up with more weight on agent 1's initial beliefs than agent 3 does, since agent 2 listens directly to agent 1 while agent 3 does not. Conversely, agent 3 listens directly to an agent in the isolated set $\{4, 5\}$ (in particular, agent 4) while agent 2 does not, and therefore ends up with more weight than agent 2 on agents' 4 and 5 initial beliefs.

# VIII Proofs

**Proof of Theorem 1:** The matrix $\mathbf{T}$ is the transition matrix of a finite, irreducible, and aperiodic Markov chain. It is well known (see, for example, Aldous and Fill [1999, Ch. 2] and the references therein) that (i) such a Markov chain has a unique stationary distribution, i.e., a unique probability distribution $w$ satisfying $w\mathbf{T} = w$, (ii) $w_i > 0$ for any state $i$, and (iii) starting from any state $i$, the probability of being in state $j$ at date $t$, converges to $w_j$ as $t \to \infty$. Property (iii) implies that the matrix $\mathbf{T}^t$ converges to a limit $\mathbf{T}^\infty$, each row of which is equal to $w$.

To prove the Theorem, it remains to show that the matrix $\prod_{s=0}^{t-1}\mathbf{T}(\lambda_s)$ converges to $\mathbf{T}^\infty$. Define the random variable $\Lambda_t$ to be equal to 1 with probability $\lambda_t$ and 0 otherwise. Assume also that $\Lambda_t$ are independent over time. Define the (random) matrix $Z_t$ by

$$Z_t = \prod_{s=0}^{t-1}[(1-\Lambda_s)\mathbf{I} + \Lambda_s\mathbf{T}] = \mathbf{T}^{\sum_{s=0}^{t-1}\Lambda_s}.$$

Then

$$E(Z_t) = \prod_{s=0}^{t-1}\mathbf{T}(\lambda_s).$$

By the Borel-Cantelli lemma, if $\sum_{t=0}^{\infty}\lambda_t = \infty$, then

$$\Pr(\Lambda_t = 1 \text{ infinitely often}) = \Pr(\sum_{t=0}^{\infty}\Lambda_t = \infty) = 1.$$

Since the matrix $\mathbf{T}^t$ is bounded uniformly in $t$, the dominated convergence theorem implies that

$$\lim_{t\to\infty} E(Z_t) = \lim_{t\to\infty} E(\mathbf{T}^{\sum_{s=0}^{t-1}\Lambda_s}) = \mathbf{T}^\infty.$$

$\|$

**Proof of Corollary 1:** The corollary follows immediately from $w\mathbf{T} = w$. $\|$

**Proof of Theorem 2:** The consensus beliefs are correct if and only if

$$w_i = \frac{\hat{\pi}_i^0}{\sum_j \hat{\pi}_j^0},$$

i.e., if and only if the vector $\hat{\pi}^0 \equiv (\hat{\pi}_1^0, \ldots, \hat{\pi}_N^0)$ solves equation $\hat{\pi}^0\mathbf{T} = \hat{\pi}^0$. The $i$th component of this vector equation is

$$\sum_j \hat{\pi}_j^0 \mathbf{T}_{ji} = \hat{\pi}_i^0. \tag{9}$$

Since

$$\mathbf{T}_{ji} = \frac{q_{ji}\pi_{ji}^0}{\pi_{jj}^1} = \frac{q_{ji}\pi_{ji}^0}{\pi_{jj}^0}\frac{\pi_{jj}^0}{\pi_{jj}^1} = \frac{q_{ji}\pi_{ji}^0}{\pi_{jj}^0}\mathbf{T}_{jj} = \frac{q_{ji}\hat{\pi}_i^0}{\hat{\pi}_j^0}\mathbf{T}_{jj},$$

equation (9) is equivalent to equation (6). $\|$

**Proof of Theorem 3:** Given normality, a sufficient statistic of agent $i$'s information after communication round $t$ is a weighted average $v_i^t x^0$ of all agents' signals, for some vector $v_i^t$ of weights. (This can easily be proven by induction.)

Suppose that after communication round $t$, agents have not converged to the correct beliefs, i.e., there exists $i$ such that $v_i^t$ differs from the vector $\hat{w}$ of weights corresponding to the correct beliefs $\hat{x}$. Then, we will show that the beliefs of at least one agent must change in round $t$, i.e., there exists $i$ such that $v_i^t \neq v_i^{t-1}$. Suppose by contradiction that $v_i^t = v_i^{t-1}$ for all $i$. Agent $i$ does not change his beliefs in round $t$ only if $v_i^{t-1}x^0$ is a sufficient statistic for $\{v_j^{t-1}x^0\}_{j\in S(i)}$. The same applies to the agents in $S(i)$, and so on. Therefore, by strong connectedness, $v_i^{t-1}x^0$ is a sufficient statistic for $\{v_j^{t-1}x^0\}_{j\in\mathcal{N}}$, and thus for $\{v_j^{t'}x^0\}_{j\in\mathcal{N},t'=0,\ldots,t-1}$. This implies that $v_i^{t-1} = \hat{w}$ for all $i$.

To show that agents converge to the correct beliefs in at most $N^2$ rounds, we will show that each agent can change his beliefs in at most $N$ rounds. For this, we observe that if agent $i$ changes his beliefs in round $t$, then $v_i^t$ must be linearly independent of $\{v_i^{t'}\}_{t'=0,\ldots,t-1}$. (Otherwise, $v_i^t x^0$ would be known in $t-1$, and thus be equal to $v_i^{t-1}x^0$). The dimension of the subspace generated by $\{v_i^{t'}\}_{t'=0,\ldots,t-1}$ can, however, increase only $N$ times. ‖

**Proof of Theorem 4:** Generically, the matrix $\mathbf{T}$ is diagonalizable, and thus

$$\mathbf{T} = V^c A V^r,$$

where $A$ is the diagonal matrix of the eigenvalues of $\mathbf{T}$, and $V^c$ and $V^r$ are the matrices of the column and row eigenvectors, respectively. Since an eigenvector of $\mathbf{T}$ corresponding to the eigenvalue $\alpha$ is also an eigenvector of $\mathbf{T}(\lambda)$ corresponding to the eigenvalue $\alpha(\lambda)$, we have

$$\mathbf{T}(\lambda) = V^c A(\lambda) V^r,$$

where $A(\lambda) \equiv (1-\lambda)\mathbf{I}+\lambda A$. Applying this equation between 0 and $t-1$, and noting that $V^c = [V^r]^{-1}$, we get

$$\prod_{s=0}^{t-1} \mathbf{T}(\lambda_s) = \prod_{s=0}^{t-1} [V^c A(\lambda_s) V^r] = V^c \left[\prod_{s=0}^{t-1} A(\lambda_s)\right] V^r.$$

Noting that the matrix $\prod_{s=0}^{t-1} A(\lambda_s)$ is diagonal, we get

$$V^c \left[\prod_{s=0}^{t-1} A(\lambda_s)\right] V^r = \sum_{n=1}^{N} \left[\prod_{s=0}^{t-1} \alpha_n(\lambda_s)\right] V_n^c V_n^r.$$

Setting

$$k_n(t) \equiv \prod_{s=0}^{t-1} \alpha_n(\lambda_s),$$

and noting that for the eigenvalue $\alpha_1 = 1$, we have $k_1(t) = 1$, $V_1^r = w$, and $V_1^c = \mathbf{1}$, we get

$$\sum_{n=1}^{N} \left[\prod_{s=0}^{t-1} \alpha_n(\lambda_s)\right] V_n^c V_n^r = \mathbf{1}w + \sum_{n=2}^{N} k_n(t) V_n^c V_n^r.$$

We can thus write beliefs at round $t$ as

$$\begin{aligned} x^t &= \left[\prod_{s=0}^{t-1} \mathbf{T}(\lambda_s)\right] x^0 \\ &= \mathbf{1}wx^0 + \sum_{n=2}^{N} k_n(t) V_n^c V_n^r x^0. \end{aligned} \tag{10}$$

Multiplying equation (10) from the left by a weight vector $\hat{w}$ (i.e., a vector such that $\hat{w} \geq 0$ and $\sum_{i=1}^{N} w_i = 0$), we have

$$\begin{aligned} \hat{w}x^t &= \hat{w}\mathbf{1}wx^0 + \sum_{n=2}^{N} k_n(t) \hat{w} V_n^c V_n^r x^0 \\ &= wx^0 + \sum_{n=2}^{N} k_n(t) \hat{w} V_n^c V_n^r x^0. \end{aligned} \tag{11}$$

Multiplying equation (10) from the left by the row vector $e_i$ whose $i$th component is one and all others are zero, we similarly have

$$\begin{aligned} x_i^t &= e_i x^t \\ &= e_i \mathbf{1} w x^0 + \sum_{n=2}^{N} k_n(t) e_i V_n^c V_n^r x^0 \\ &= w x^0 + \sum_{n=2}^{N} k_n(t) V_{in}^c V_n^r x^0. \end{aligned} \tag{12}$$

Using equations (11) and (12), we can write the vector $d_i^t \equiv (d_{i1}^t, \ldots, d_{iL}^t)'$ as

$$d_i^t = \frac{\sum_{n=2}^{N} k_n(t) \left(V_{in}^c - \hat{w} V_n^c\right) V_n^r x^0}{\sum_j \hat{w}_j \left\| \sum_{n=2}^{N} k_n(t) \left(V_{jn}^c - \hat{w} V_n^c\right) V_n^r x^0 \right\|}.$$

Separating the terms corresponding to the second largest eigenvalue, dividing numerator and denominator by $k_2(t)$, and setting $D \equiv V_2^r x^0$, we have

$$d_i^t = \frac{\left(V_{i2}^c - \hat{w} V_2^c\right) D + \sum_{n=3}^{N} \frac{k_n(t)}{k_2(t)} \left(V_{in}^c - \hat{w} V_n^c\right) V_n^r x^0}{\sum_j \hat{w}_j \left\| \left(V_{j2}^c - \hat{w} V_2^c\right) D + \sum_{n=3}^{N} \frac{k_n(t)}{k_2(t)} \left(V_{jn}^c - \hat{w} V_n^c\right) V_n^r x^0 \right\|}.$$

The theorem will obviously follow if we show that $k_n(t)/k_2(t) \to 0$ for $n > 2$. Taking logs, we have

$$\log \left\| \frac{k_n(t)}{k_2(t)} \right\| = \sum_{s=0}^{t-1} \log \left\| \frac{\alpha_n(\lambda_s)}{\alpha_2(\lambda_s)} \right\| \equiv \sum_{s=0}^{t-1} y_s. \tag{13}$$

From the definition of $f_\lambda(\alpha)$, it follows that

$$\frac{1}{\lambda_s} \log \|\alpha_n(\lambda_s)\| = f_{\lambda_s}(\alpha_n).$$

Since $f_\lambda(\alpha)$ is continuous w.r.t. $\lambda$ (this is obvious for $\lambda > 0$, and holds by definition for $\lambda = 0$), we have

$$\lim_{s\to\infty} \left[ \frac{1}{\lambda_s} \log \|\alpha_n(\lambda_s)\| \right] = \log \left[ f_{\lambda^*}(\alpha_n) \right].$$

Using this fact, we have

$$\lim_{s\to\infty} \frac{y_s}{\lambda_s} = \log \left[ \frac{f_{\lambda^*}(\alpha_n)}{f_{\lambda^*}(\alpha_2)} \right] \equiv z_n < 0.$$

Therefore,

$$\lim_{s\to\infty} \frac{y_s}{\lambda_s z_n} = 1,$$

and thus the series (13) has the same behavior as the series

$$\sum_{s=0}^{t-1} \lambda_s z_n = z_n \sum_{s=0}^{t-1} \lambda_s,$$

which goes to $-\infty$ from Assumption 1. Therefore, $k_n(t)/k_2(t)$ goes to zero. $\|$

**Proof of Theorem 5:** From the proof of Theorem 4, it follows that $\|x_i^t - w x^0\|$ is of order $\|k_2(t)\|$

and $\|d_i^t - d_i^\infty\|$ of order $\|k_3(t)/k_2(t)\|$. From the same proof, it follows that $\log\|k_3(t)/k_2(t)\|$ is of order

$$z_3 \sum_{s=0}^{t-1} \lambda_s.$$

Therefore, $\|k_3(t)/k_2(t)\|$ is of order

$$\exp\left[z_3 \sum_{s=0}^{t-1} \lambda_s\right] = \left[\frac{f_{\lambda^*}(\alpha_3)}{f_{\lambda^*}(\alpha_2)}\right]^{\sum_{s=0}^{t-1} \lambda_s}.$$

The proof for $\|k_2(t)\|$ is similar. ‖

**Proof of Theorem 6:** To prove the theorem for the common precision case, it suffices to show that the vector

$$\Pi^c \equiv (\pi_1^0 \sum_k q_{k1}\pi_k^0, \ldots, \pi_N^0 \sum_k q_{kN}\pi_k^0)$$

satisfies equation $\Pi^c \mathbf{T} = \Pi^c$. The $i$th component of this vector equation is

$$\sum_j \left[\pi_j^0 \sum_k q_{kj}\pi_k^0\right] \mathbf{T}_{ji} = \pi_i^0 \sum_k q_{ki}\pi_k^0. \tag{14}$$

Since

$$\mathbf{T}_{ji} = \frac{q_{ji}\pi_{ji}^0}{\pi_{jj}^1} = \frac{q_{ji}\pi_{ji}^0}{\sum_k q_{jk}\pi_{jk}^0} = \frac{q_{ji}\pi_i^0}{\sum_k q_{jk}\pi_k^0} = \frac{q_{ji}\pi_i^0}{\sum_k q_{kj}\pi_k^0},$$

equation (14) is equivalent to

$$\sum_j \pi_j^0 q_{ji}\pi_i^0 = \pi_i^0 \sum_k q_{ki}\pi_k^0,$$

which obviously holds.

Similarly, for the symmetric precision case, it suffices to show that the vector

$$\Pi^s \equiv (\sum_k q_{k1}\pi_{k1}^0, \ldots, \sum_k q_{kN}\pi_{kN}^0)$$

satisfies equation $\Pi^s \mathbf{T} = \Pi^s$. The $i$th component of this vector equation is

$$\sum_j \left[\sum_k q_{kj}\pi_{kj}^0\right] \mathbf{T}_{ji} = \sum_k q_{ki}\pi_{ki}^0. \tag{15}$$

Since

$$\mathbf{T}_{ji} = \frac{q_{ji}\pi_{ji}^0}{\pi_{jj}^1} = \frac{q_{ji}\pi_{ji}^0}{\sum_k q_{jk}\pi_{jk}^0} = \frac{q_{ji}\pi_{ji}^0}{\sum_k q_{kj}\pi_{kj}^0},$$

equation (15) is equivalent to

$$\sum_j q_{ji}\pi_{ji}^0 = \sum_k q_{ki}\pi_{ki}^0,$$

which obviously holds. ‖

**Proof of Theorem 7:** We first show the detailed balance equations $w_i \mathbf{T}_{ij} = w_j \mathbf{T}_{ji}$ for all $i, j$. In the common precision case, we have

$$\Pi_j^c \mathbf{T}_{ji} = \left[\pi_j^0 \sum_k q_{kj}\pi_k^0\right] \mathbf{T}_{ji} = \left[\pi_j^0 \sum_k q_{kj}\pi_k^0\right] \frac{q_{ji}\pi_i^0}{\sum_k q_{kj}\pi_k^0} = \pi_j^0 q_{ji}\pi_i^0 = \pi_i^0 q_{ij}\pi_j^0 = \Pi_i^c \mathbf{T}_{ij}.$$

Similarly, in the symmetric precision case,

$$\Pi_j^s \mathbf{T}_{ji} = \left[\sum_k q_{kj}\pi_{kj}^0\right] \mathbf{T}_{ji} = \left[\sum_k q_{kj}\pi_{kj}^0\right] \frac{q_{ji}\pi_{ji}^0}{\sum_k q_{kj}\pi_{kj}^0} = q_{ji}\pi_{ji}^0 = q_{ij}\pi_{ij}^0 = \Pi_i^s \mathbf{T}_{ij}.$$

Therefore, in both cases, the detailed balance equations hold.

Defining the diagonal matrix $W$, with $i$'th diagonal element $\sqrt{w_i}$, we can write the detailed balance equations as

$$W^2\mathbf{T} = (W^2\mathbf{T})' \Rightarrow W\mathbf{T}W^{-1} = (W\mathbf{T}W^{-1})'.$$

Since the matrix $W\mathbf{T}W^{-1}$ is symmetric, it is diagonalizable with real eigenvalues and orthonormal eigenvectors. This means that there exists a diagonal matrix $A$ with real elements, and an invertible matrix $V$ such that $V^{-1} = V'$, satisfying

$$W\mathbf{T}W^{-1} = V^{-1}AV.$$

This equation implies that

$$\mathbf{T} = (VW)^{-1}AVW,$$

i.e., the matrix $\mathbf{T}$ is diagonalizable with real eigenvalues, row eigenvector matrix

$$V^r = VW,$$

and column eigenvector matrix

$$V^c = (V^r)^{-1} = W^{-1}V'.$$

Combining the last two equations, we get

$$V^r = (WV^c)'W = (V^c)'W^2.$$

Therefore, given a column eigenvector of $\mathbf{T}$ (i.e., a column of $V^c$) the corresponding row eigenvector (i.e., the corresponding row of $V^r$) can be deduced as in the theorem. ‖

**Proof of Theorem 8:** Suppose that $\alpha$ is an eigenvalue of $\mathbf{T}$, and $v = (v_{-M}, \ldots, v_M)$ a corresponding column eigenvector. Equation $\mathbf{T}v = \alpha v$ implies that

$$\frac{1}{2}v_{-M} + \frac{1}{2}v_{-(M-1)} = \alpha v_{-M}, \tag{16}$$

$$\frac{1}{3}v_{i-1} + \frac{1}{3}v_i + \frac{1}{3}v_{i+1} = \alpha v_i, \quad \text{for } -M < i < M, \tag{17}$$

$$\frac{1}{2}v_{M-1} + \frac{1}{2}v_M = \alpha v_M. \tag{18}$$

The solution to the difference equation (17) is of the form

$$v_i = \beta_1(\rho_1)^i + \beta_2(\rho_2)^i, \tag{19}$$

where $-M \leq i \leq M$, $\beta_1$ and $\beta_1$ are two constants to be determined, and $\rho_1$ and $\rho_2$ are the roots of the characteristic equation

$$\rho^2 + (1-3\alpha)\rho + 1 = 0. \tag{20}$$

The constants $\beta_1$ and $\beta_2$ can be determined by plugging the solution (19) into equations (16) and (18). Plugging into equation (16), we get

$$\beta_1(\rho_1)^{-M}\left[(1-2\alpha)+\rho_1\right] + \beta_2(\rho_2)^{-M}\left[(1-2\alpha)+\rho_2\right] = 0, \tag{21}$$

and plugging into equation (18), we get

$$\beta_1(\rho_1)^{M}\left[(1-2\alpha)+\frac{1}{\rho_1}\right] + \beta_2(\rho_2)^{M}\left[(1-2\alpha)+\frac{1}{\rho_2}\right] = 0. \tag{22}$$

Equations (21) and (22) form a $2 \times 2$ linear system in $\beta_1$ and $\beta_2$. For $\alpha$ to be an eigenvalue of $\mathbf{T}$, this system must have a non-zero solution, and thus its determinant has to be zero. This will imply an equation which will determine $\alpha$.

To derive this equation, we assume that $\alpha \in (-1/3, 1]$. (Under this assumption, we will obtain $2M+1 = N$ distinct eigenvalues, which is obviously the maximum number that $\mathbf{T}$ can have.) Setting

$$\cos\theta = \frac{3\alpha - 1}{2} \in (-1, 1], \tag{23}$$

we have $\rho_1 = e^{i\theta}$ and $\rho_2 = e^{-i\theta}$. The determinant of the system of (21) and (22) is

$$e^{-2iM\theta}\left[(1-2\alpha)+e^{i\theta}\right]^2 - e^{2iM\theta}\left[(1-2\alpha)+e^{-i\theta}\right]^2.$$

Setting this to zero, we get

$$e^{-iM\theta}\left[(1-2\alpha)+e^{i\theta}\right] = \pm e^{iM\theta}\left[(1-2\alpha)+e^{-i\theta}\right]. \tag{24}$$

The equation corresponding to the "plus" sign is

$$(1-2\alpha)\sin[M\theta] + \sin[(M-1)\theta] = 0, \tag{25}$$

and can be simplified into

$$\begin{aligned}
& (1-2\alpha+\cos\theta)\sin[M\theta] - \sin\theta\cos[M\theta] = 0 \\
\Leftrightarrow\quad & \frac{1}{3}(1-\cos\theta)\sin[M\theta] - \sin\theta\cos[M\theta] = 0 \\
\Leftrightarrow\quad & \sin[g(\theta)]\sin[M\theta] - \cos[g(\theta)]\cos[M\theta] = 0 \\
\Leftrightarrow\quad & \cos[M\theta + g(\theta)] = 0,
\end{aligned} \tag{26}$$

where we first used the addition of sines rule, then equation (23), and finally the function $g(\theta) \in [-\pi/2, \pi/2]$ defined by

$$\tan[g(\theta)] = \frac{1-\cos\theta}{3\sin\theta} = \frac{1}{3}\tan\left[\frac{\theta}{2}\right].$$

Likewise, the equation corresponding to the "minus" sign is

$$(1-2\alpha)\cos[M\theta] + \cos[(M-1)\theta] = 0,$$

and can be simplified into

$$\sin[M\theta + g(\theta)] = 0. \tag{27}$$

Any solution $\theta$ to equations (26) and (27) must satisfy

$$M\theta + g(\theta) = \frac{m\pi}{2}, \tag{28}$$

for some integer $m$. We will show that for each $m = 0, \dots, 2M$, equation (28) has a unique solution in $[0, \pi)$, which we will denote by $\theta_{m+1}$. Given this solution, we can deduce an eigenvalue of $\mathbf{T}$ from equation (23). The $2M + 1 = N$ eigenvalues so obtained are distinct since the solutions $\theta_{m+1}$ are distinct and in $[0, \pi)$.

Setting

$$G(\theta) = M\theta + g(\theta) - \frac{m\pi}{2},$$

it is easy to check that $G(\theta)$ is strictly increasing in $\theta \in [0, \pi)$, and that, given $0 \le m \le 2M$, $G(0) \le 0$ and $G(\pi) > 0$. Therefore, for each $m = 0, \dots, 2M$, equation $G(\theta) = 0$ has a unique solution $\theta_{m+1} \in [0, \pi)$.

For $m = 0$, $G(0) = 0$, and thus $\theta_1 = 0$. For $m = 1$, $G(\pi/2) > 0$, and thus $\theta_2 \in (0, \pi/2)$. Moreover, since $\theta_2 \in (0, \pi/2)$,

$$\begin{aligned}
& 0 < g(\theta_2) < \frac{\theta_2}{2} \\
\Rightarrow\quad & M\theta_2 - \frac{\pi}{2} < M\theta_2 + g(\theta_2) - \frac{\pi}{2} = G(\theta_2) = 0 < M\theta_2 + \frac{\theta_2}{2} - \frac{\pi}{2} \\
\Rightarrow\quad & \frac{\pi}{2M+1} < \theta_2 < \frac{\pi}{2M}.
\end{aligned}$$

The eigenvalue corresponding to $\theta_1 = 0$ is $\alpha_1 = 1$. We will show that the eigenvalue $\alpha$ corresponding to $\theta_2$ is the second largest. For this, it suffices to show that $\alpha$ exceeds all remaining eigenvalues (i) algebraically and (ii) in absolute value. For (i), note that implicit differentiation of $G(\theta_{m+1}) = 0$ implies that $\theta_{m+1}$ is strictly increasing in $m$. For (ii), note that since $\theta_2 < \pi/2$, $\alpha > 1/3$. Therefore, $\alpha$ is the second largest eigenvalue, i.e., $\alpha = \alpha_2$.

Consider next the second column eigenvector. Since $\theta_2$ solves equation (28) for $m = 1$, it is a solution of equation (26), and thus of the equation with the "plus" sign in (24). Equation (21) then implies that $\beta_1 = -\beta_2$, and equation (19) implies that

$$v_i = 2\beta_1 \sin(i\theta).$$

Therefore, the second column eigenvector is proportional to $v_i^c$.

Consider finally the second row eigenvector. Theorem 6 implies that the relative social influence weights are 2 for agents $\pm M$ and 3 for all other agents. Theorem 7 then implies that the second row eigenvector is proportional to $v_i^r$. Regarding the extreme components of $v_i^r$, we only need to show that for $M \ge 2$, $v_{M-1}^r \ge v_M^r$, i.e.,

$$3\sin[(M-1)\theta_2] \ge 2\sin[M\theta_2].$$

Since $\theta_2$ solves equation (28) for $m = 1$, it is a solution of equation (25). Using that equation to substitute $\sin[(M-1)\theta_2]$, we need to show that

$$(6\alpha_2 - 5)\sin[M\theta_2] \ge 0.$$

Since $M\theta_2 < \pi/2$, we have $\sin[M\theta_2] > 0$. Moreover, for $M = 2$, $\alpha_2 = 5/6$ (this can be checked numerically), and since $\alpha_2$ is strictly increasing in $M$, $\alpha_2 > 5/6$ for $M > 2$. ‖

**Proof of Theorem 9:** See Aldous and Fill [1999, Ch. 9] and the references therein. A sketch of the proof is as follows. We need to show that the vector

$$\mathcal{G} \equiv (\sum_{G \in \mathcal{G}_1} \rho(G), \ldots, \sum_{G \in \mathcal{G}_N} \rho(G)),$$

solves equation $\mathcal{G}\mathbf{T} = \mathcal{G}$. The $i$th component of this vector equation is

$$\begin{aligned} & \sum_j [(\sum_{G \in \mathcal{G}_j} \rho(G))\mathbf{T}_{ji}] = \sum_{G \in \mathcal{G}_i} \rho(G) \\ \Leftrightarrow \quad & \sum_{j \neq i} [(\sum_{G \in \mathcal{G}_j} \rho(G))\mathbf{T}_{ji}] = \sum_{G \in \mathcal{G}_i} \rho(G)(\sum_{j \neq i} \mathbf{T}_{ij}) \end{aligned} \tag{29}$$

Consider now the mapping $F$ that to a spanning tree $G$ with $i$ at the root, and to an agent $j \neq i$, associates the tree $F(G, j)$ constructed by taking the "subtree" of $G$ formed by $i$ and all of $i$'s subtrees except that containing $j$, and "pasting" it just below $j$. The range of $F$ is the full set of spanning trees with an agent other than $i$ at the root, and thus $F$ is one-to-one. Moreover, if $k$ is the root of $F(G, j)$, we have $\rho(G)\mathbf{T}_{ij} = \rho(F(G, j))\mathbf{T}_{ki}$. This implies equation (29). ‖

**Proof of Corollary 2:** In the equal precision case, the weight an agent $k$ gives to all agents he listens to is $1/\#S(k)$. Therefore, for all $G \in \mathcal{G}_i$, we have

$$\rho(G) = \prod_{k \neq i} \frac{1}{\#S(k)},$$

and thus

$$\sum_{G \in \mathcal{G}_i} \rho(G) =^{\#} \mathcal{G}_i \prod_{k \neq i} \frac{1}{\#S(k)}.$$

The corollary then follows from Theorem 9. ‖

**Proof of Theorem 10:** We first show that each non-isolated element of the partition $\mathcal{P}$ is linked to at least one isolated element. This will also establish that the set of isolated elements of $\mathcal{P}$ is nonempty. Consider a non-isolated element $J$ of $\mathcal{P}$ that is linked only to non-isolated elements. Then starting from $J$, we can construct a chain of linked, non-isolated elements. Since $\mathcal{P}$ contains a finite number of elements, this chain must contain a cycle consisting of more than one elements. But then all agents belonging to elements in this cycle are strongly connected, and therefore $\mathcal{P}$ is not the partition defined by strong connectedness, a contradiction.

To determine agents' limit beliefs, we first determine the matrix $\mathbf{T}^\infty$. By grouping agents according to the partition $\mathcal{P}$, we can write the matrix $\mathbf{T}$ as

$$\left( \begin{array}{ccccc} \mathbf{T}_1 & & & & 0 \\ & \mathbf{T}_2 & & & \\ & & . & & \\ 0 & & & \mathbf{T}_M & \\ R_1 & R_2 & . & R_M & Q \end{array} \right),$$

where the $m$th set of rows, $m = 1, \ldots, M$, corresponds to the agents in the $m$th isolated set, $I_m$, and the $M + 1$th set of rows corresponds to the agents in the non-isolated sets. The matrix $\mathbf{T}^t$ has the same form as $\mathbf{T}$, except that $\mathbf{T}_m$ is replaced by $\mathbf{T}_m^t$, $Q$ by $Q^t$, and $R_m$ by

$$R_{m,t} \equiv \sum_{s=0}^{t-1} Q^s R_m \mathbf{T}_m^{t-1-s}.$$

As in Theorem 1, the matrix $\mathbf{T}_m^t$ converges to a limit $\mathbf{T}_m^\infty$, each row of which is equal to $w(I_m)$. The matrix $Q^t$ converges to 0. Indeed, denoting by $q_{j,t}$ the maximum element in the $j$th column of $Q^t$, equation $Q^{t+1} = QQ^t$ implies that

$$(Q^{t+1})_{ij} = \sum_k Q_{ik}(Q^t)_{kj} \Rightarrow (Q^{t+1})_{ij} \leq \left(\sum_k Q_{ik}\right) q_{j,t} \Rightarrow q_{j,t+1} \leq \left(\sum_k Q_{ik}\right) q_{j,t} \Rightarrow q_{j,t+1} \leq q_{j,t},$$

where the second implication follows from the elements $Q$ being non-negative, and the last from each row of $Q$ summing to at most one. Since $q_{j,t}$ is decreasing and bounded below (by zero), it must converge to a limit $q_j$. To show that $q_j = 0$, we note that each agent in a non-isolated set is influenced by agents in isolated sets, after enough communication rounds. Therefore, there exists $t^*$ such that all rows of $Q^{t^*}$ sum to strictly less than 1, and thus less than $1-\epsilon$ for some $\epsilon > 0$. Using equation $Q^{t+t^*} = Q^{t^*}Q^t$, and proceeding as before, we get

$$q_{j,t+t^*} \leq (1-\epsilon) q_{j,t}.$$

Therefore, $q_j = 0$, and thus the matrix $Q^t$ converges to zero. Since $\mathbf{T}_m^t$ converges to $\mathbf{T}_m^\infty$ and $Q^t$ converges to zero, it is easy to show that $R_{m,t}$ converges to

$$R_{m,\infty} \equiv \sum_{s=0}^{\infty} Q^s R_m \mathbf{T}_m^\infty.$$

This fully determines the matrix $\mathbf{T}^\infty$. As in Theorem 1, $\mathbf{T}^\infty$ is also the limit of $\prod_{s=0}^{t-1} \mathbf{T}(\lambda_s)$ for all $\lambda_t$ satisfying Assumption 1.

The limit beliefs of an agent $i$ are given by the $i$th row of $\mathbf{T}^\infty x^0$. If $i$ belongs to the isolated set $I_m$, the $i$th row is $w(I_m)x^0(I_m)$. If $i$ belongs to a non isolated set $J$, the $i$th row is

$$\sum_{m=1}^{M} \left[R_{m,\infty} x^0(I_m)\right]_i = \sum_{m=1}^{M} \left[\sum_{s=0}^{\infty} Q^s R_m \mathbf{T}_m^\infty x^0(I_m)\right]_i.$$

Since $\mathbf{T}_m^\infty = \mathbf{1}w(I_m)$, where $\mathbf{1}$ is a column vector of ones corresponding to the set $I_m$, we can write the above as

$$\sum_{m=1}^{M} \left[\sum_{s=0}^{\infty} Q^s R_m \mathbf{1}\right]_i w(I_m)x^0(I_m).$$

Agent $i$'s limit beliefs are thus a weighted average of the limit beliefs of the agents in the isolated sets. The weight corresponding to the set $I_m$ is non zero if the $i$th row of the matrix $\sum_{s=0}^{\infty} Q^s R_m \mathbf{1}$ is non zero. This occurs precisely when $i$ is influenced by an agent in $I_m$ after enough communication rounds, i.e., when the set $J$ is linked to $I_m$. ∥

## References


**Aldous, David, and Jim Fill,** 1999, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, U.C. Berkeley.

**Bala, Venkatesh, and Sanjeev Goyal,** 1998, "Learning from Neighbors," *Review of Economic Studies*, Vol. 65, No. 3, pp. 595-621.

**Banerjee, Abhijit,** 1992, "A Simple Model of Herd Behavior," *Quarterly Journal of Economics,* Vol. 107, No. 3, pp. 797-817.

**Berger, Roger,** 1981, "A Necessary and Sufficient Condition for Reaching a Consensus Using DeGroot's Method," *Journal of the American Statistical Association*, Vol. 76, No. 374, pp. 415-418.

**Bertrand, Marianne, Erzo Luttmer, and Sendhil Mullainathan,** 2000, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, Vol. 115, No. 3, pp. 1019-1055.

**Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch,** 1992, "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy*, Vol. 100, No. 5, pp. 992-1026.

**Bonacich, Phillip, and Paulette Lloyd,** 2001, "Eigenvector-like Measures of Centrality for Asymmetric Relations," working paper, UCLA.

**Borjas, George,** 1995, "Ethnicity, Neighborhoods, and Human-Capital Externalities," *American Economic Review*, Vol. 85, No. 3, pp. 365-390.

**Caplin, Andrew, and Barry Nalebuff,** 1991, "Aggregation and Social Choice: A Mean Voter Theorem," *Econometrica,* Vol. 59, No. 1, pp. 1-23.

**Case, Anne, and Lawrence Katz,** 1991, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," NBER working paper.

**Cohen, Bernard,** 1963, *The Press and Foreign Policy*, Princeton University Press, Princeton.

**Converse, Philip,** 1964, "The Nature of Beliefs Systems in Mass Publics," in D. Apter (ed.), *Ideology and Discontent,* Free Press, New York.

**DeGroot, Morris,** 1974, "Reaching a Consensus," *Journal of the American Statistical Association*, Vol. 69, No. 345, pp. 118-121.

**DeMarzo, Peter, Dimitri Vayanos, and Jeffrey Zwiebel,** 2001, "Social Networks and Financial Markets," working paper, MIT and Stanford University.

**Duflo, Esther, and Emmanuel Saez,** 2000, "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices," *Journal of Public Economics*, forthcoming.

**Ellison, Glenn, and Drew Fudenberg,** 1993, "Rules of Thumb for Social Learning," *Journal of Political Economy*, Vol. 101, No. 4, pp. 612-643.

**Ellison, Glenn, and Drew Fudenberg,** 1995, "Word of Mouth Communication and Social Learning," *Quarterly Journal of Economics*, Vol. 110, No. 1, pp. 93-125.

**Erbring, Lutz, Edie Goldenberg, and Arthur Miller,** 1980, "Front-Page News and Real-World Cues: Another Look at Agenda-Setting by the Media," *American Political Science Review*, Vol. 24, pp. 16-49.

**Evans, William, Oates Wallace, and Robert Schwab,** 1992, "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy*, Vol. 100, pp. 966-991.

**Festinger, Leon, Stanley Schachter, and Kurt Back,** 1950, *Social Pressures in Informal Groups: A Study in Human Factors in Housing*, Stanford University Press, Stanford.

**Fiske, Susan, and Shelley Taylor,** 1984, *Social Cognition*, Addison-Wesley.

**French, John,** 1956, "A Formal Theory of Social Power," *Psychological Review*, Vol. 63, No. 3, pp. 181-194.

**Geanakoplos, John, and Heracles Polemarchakis,** 1982, "We Can't Disagree Forever," *Journal of Economic Theory*, Vol. 28, No. 1, pp. 192-200.

**Gilbert, Daniel,** 1991, "How Mental Systems Believe," *American Psychologist*, Vol. 46, No. 2, pp. 107-119.

**Glaeser, Edward, Bruce Sacerdote, and José Scheinkman,** 1996, "Crime and Social Interactions," *Quarterly Journal of Economics*, Vol. 111, No. 2, pp. 507-548.

**Harary, Frank,** 1959, "A Criterion for Unanimity in French's Theory of Social Power," in Cartwright, Dorwin, (eds.) *Studies in Social Power*, University of Michigan Press, Ann Arbor, pp. 168-182.

**Hawkins, Scott, and Stephen Hoch,** 1992, "Low-Involvement Learning: Memory without Evaluation," *Journal of Consumer Research*, Vol. 19, pp. 212-225.

**Hong, Harrison, and Jeremy Stein,** 2001, "Social Interaction and Stock-Market Participation," NBER Working Paper.

**Huckfledt, Robert, and John Sprague,** 1995, *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign*, Cambridge University Press, Cambridge.

**Iyengar, Shanto, and Donald Kinder,** 1987, *News That Matters: Television and American Opinion*, University of Chicago Press, Chicago.

**Katz, Lawrence, Jeffrey Kling, and Jeffrey Liebman,** 1999, "Moving to Opportunity in Boston: Early Impacts of a Housing Mobility Program," working paper, Princeton University.

**Lodge, Milton,** 1995, "Toward a Procedural Model of Candidate Evaluation," in M. Lodge and K.M. Mcgraw (eds.), *Political Judgement: Structure and Process*, University of Michigan Press, Ann Arbor.

**Madrian, Brigitte, and Dennis Shea,** 2000, "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," NBER working paper.

**Nisbett, Richard, and Lee Ross,** 1980, *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice Hall, Princeton.

**Parikh, Rohit, and Paul Krasucki,** 1990, "Communication, Consensus and Knowledge," *Journal of Economic Theory*, Vol. 52, No. 1, pp. 178-189.

**Poole, Keith, and Howard Rosenthal,** 1991, "Patterns of Congressional Voting," *American Journal of Political Science,* Vol. 35, pp. 228-278.

**Poole, Keith, and Howard Rosenthal,** 1997, *Congress; A Political-Economic Theory of Roll Call Voting*, Oxford University Press, Oxford.

**Press, James,** 1978, "Qualitative Controlled Feedback for Forming Group Judgments and Making Decisions," *Journal of American Statistical Association*, Vol. 73, No. 363, pp. 526-535.

**Spector, David,** 2000, "Rational Debate and One-Dimensional Conflict," *Quarterly Journal of Economics,* Vol. 115, No. 1, pp. 181-200.

**Tversky, Amos, and Daniel Kahneman,** 1973, "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, Vol. 5, No. 2, pp. 207-232.

**Varian, Hal,** 1992, *Microeconomic Analysis*, Norton, New York.

**Wasserman, Stanley, and Katherine Faust,** 1994, *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge.

**Zallner, J.,** 1992, *The Nature and Origins of Mass Opinion*, Cambridge University Press, Cambridge.

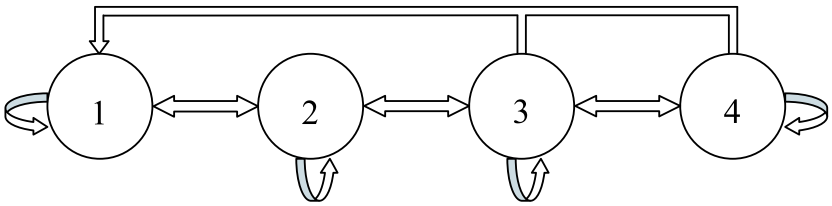


Figure 1: Listening Structure for Example 1.

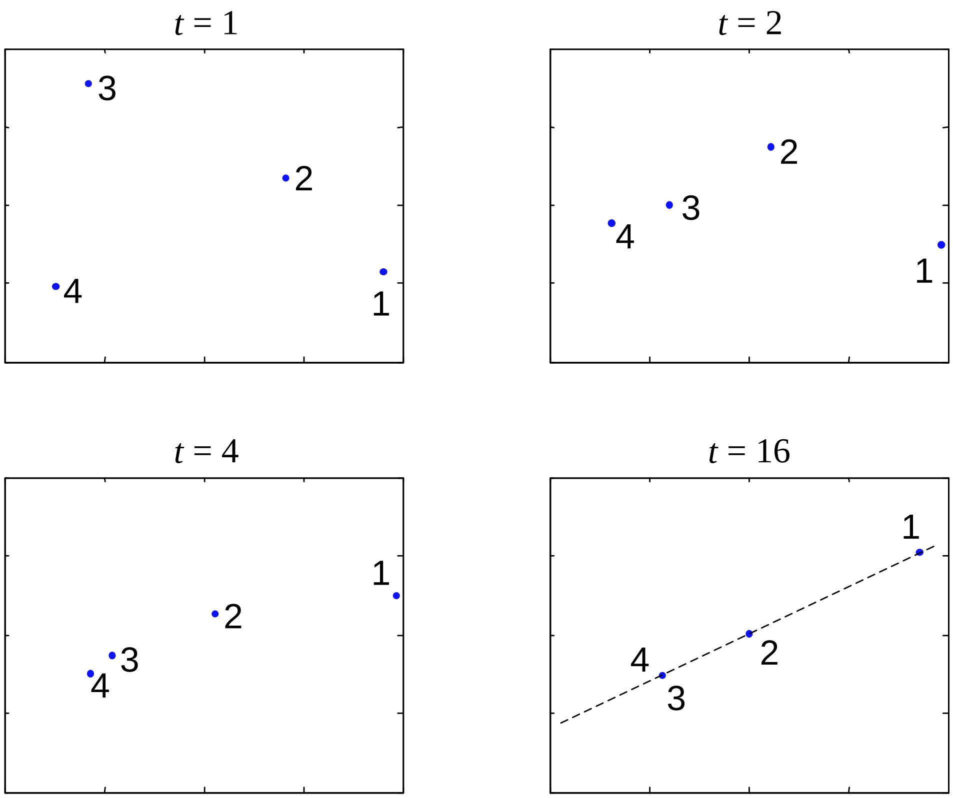


Figure 2: Dynamics of Beliefs for Example 1.

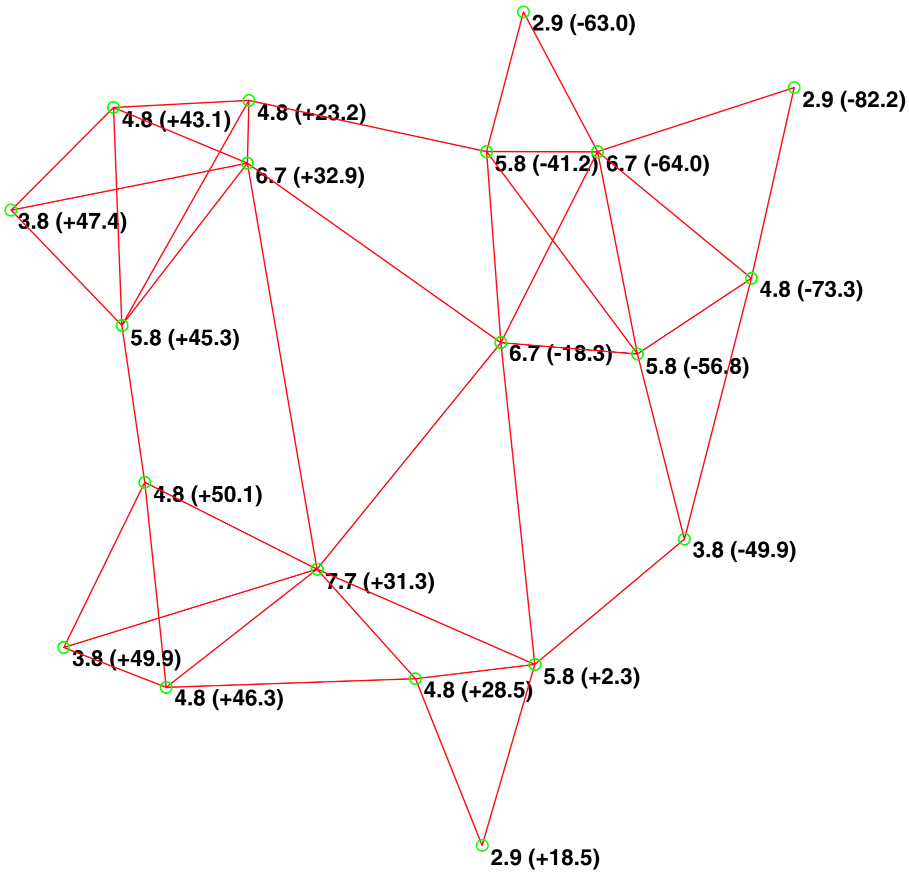


Figure 3: Two-Dimensional Neighborhood Example of Section IV.B. First Number gives Influence Weight (%), Second Number gives the Long-Run Position, $V_{i2}^c$.

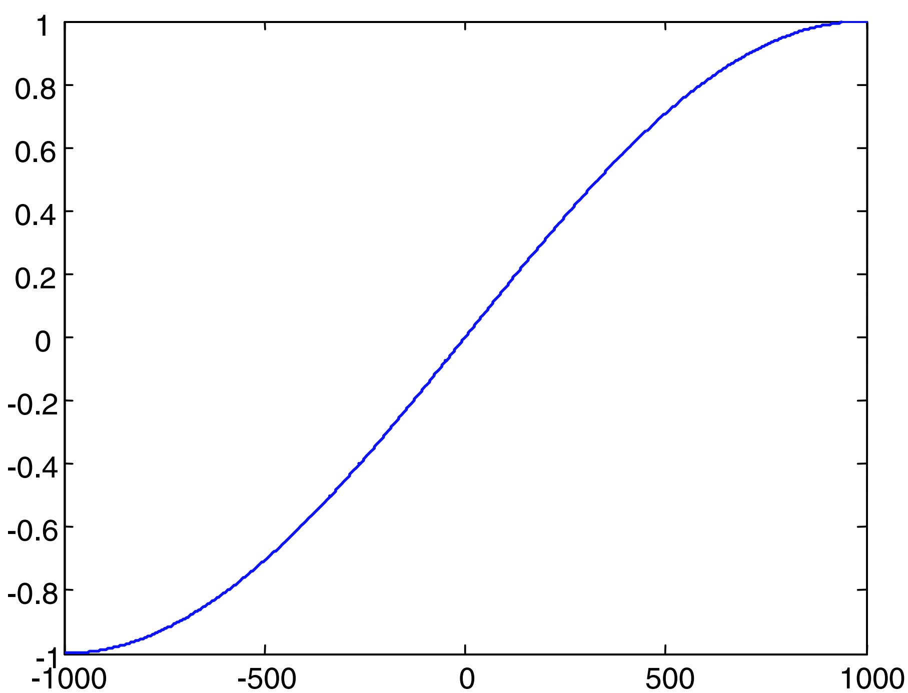


Figure 4: Long-Run Positions for the Linear Neighborhood Example of Section IV.C. ($M = 1000$)

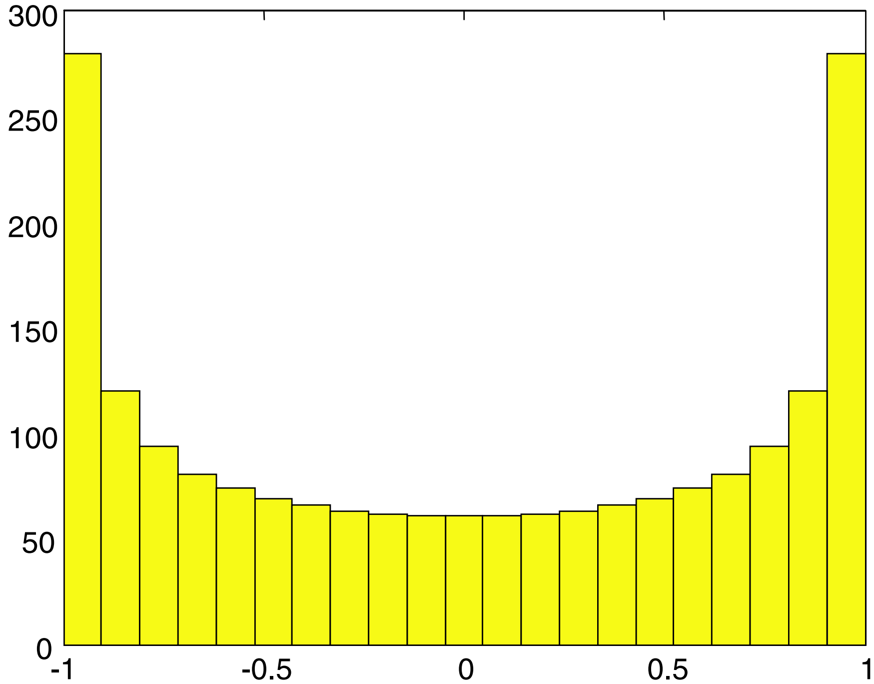


Figure 5: Density of Long-Run Disagreement Positions for the Linear Neighborhood Example of Section IV.C.

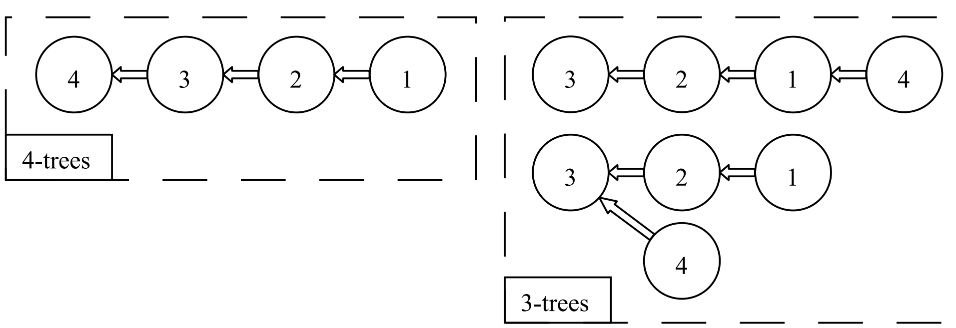


Figure 6: Spanning Trees for Agents 3 and 4 in Example 1.