Rumors and Social Networks

Francis Bloch, Gabrielle Demange, and Rachel Kranton *

November 30, 2016

Abstract: Why do people spread rumors? This paper studies the transmission of possibly false information—by rational agents who seek the truth. Unbiased agents earn payoffs when a collective decision is correct in that it matches the true state of the world. Biased agents desire a particular decision regardless of the true state. One agent possibly learns the true underlying state and chooses whether to send a true or false message to her friends and neighbors, who then decide whether or not to transmit it further. The paper finds that a social network can serve as a filter: unbiased agents block messages from parts of the network that contain many biased agents, so that sufficiently credible information circulates. In order to maximize transmission of messages, it can serve biased agents to limit their number and spread themselves throughout the network.

Keywords: Bayesian updating, rumors, misinformation, social networks.

JEL Classification: C72, D83.

*Francis Bloch: Paris School of Economics-Paris I, Gabrielle Demange: Paris School of Economics-EHESS, Rachel Kranton: Duke University. We thank Aaron Kolb, Margaux Luflade and Richard Richardson for invaluable research assistance, three anonymous referees, the editor, and seminar participants at various universities and conferences for comments. Francis Bloch is supported by the grant ANR 13BSH1-0010, Gabrielle Demange is supported by the grant NET from ANR. Rachel Kranton thanks Chaire Blaise Pascal/Paris School of Economics, the National Science Foundation, and the Canadian Institute for Advanced Research for support.

I Introduction

Why do people spread rumors? Rumors, in the dictionary definition, are opinions spread from person to person with uncertain veracity and possibly no discernible source.[1] In a prominent book Cass Sunstein (2009) documents the pervasiveness of rumors, their public benefits, and their perils. Rumors abound concerning the efficacy of vaccines, the birthplace of presidential candidates, the propriety of politicians, the fabrication of data in academic research, and the integrity of local and national elections. This paper studies why rumors are spread—by rational agents who seek the truth.

The paper studies strategic communication of rumors in a network. Agents' individual payoffs depend on a collective decision, such as election of a candidate or authorizing the use of a new technology. Collective-decision making is modeled as a stylized "vote" that reflects each agent's expected utility from the decision. Some agents are unbiased and prefer that the decision correctly matches the true state of the world. Other agents are biased and prefer a particular decision regardless of the true state. (Such agents might personally benefit, say, from the decision.) Agents have prior beliefs as to the true state. One agent, selected at random, possibly receives precise information about the true state. This agent, whose identity is not known, can create a false or true message—a rumor—of the state of the world. Biased agents have the incentive to create a false message. Agents who receive a message make an inference as to the veracity of the message and decide whether or not to pass it along in order to influence how others will vote on the collective outcome.

The paper derives conditions for a *full communication equilibrium*, where all biased and unbiased agents transmit messages and, therefore, spread possibly false rumors. They do so because there is a sufficiently large probability the rumor is true. The equilibrium conditions rely on the number and distribution of biased and unbiased agents in the population. In an network, for any agent, the set of possible senders of a message must contain sufficiently few biased agents.

When this condition fails so that full communication is not possible, there is an equilibrium, called *maximal communication equilibrium*, in which communication is maximized: in any equilibrium, information flows on an edge only if it flows in this maximal communication equilibrium. We construct an algorithm (which runs in finite time) that precisely identifies the edges of the

[1] Webster's English dictionary and Oxford English dictionary.

network through which communication takes place at the maximal communication equilibrium. A main feature of this equilibrium is that information can flow from one part of the network to another but not in the reverse direction. Unbiased agents maintain the credibility of messages by blocking those that come from a part of the network that contains too many biased agents. This same agent, however, will transmit messages coming from another direction. These maximal communication equilibria yield the highest expected payoffs of all perfect Bayesian equilibria of the game.

We have three main economic insights. First, networks can serve as a filter and aid communication. We contrast the network outcomes to a situation where agents can communicate to everyone simultaneously. In this *public broadcast model*, there are only two equilibrium outcomes: one with full communication and one with no communication. Full communication arises if and only if there are sufficiently few biased agents in the population. The network can replicate the full communication outcome when biased agents are evenly distributed in the network. The network, however, can also allow partial communication when no communication occurs in the public broadcast model. In a network, agents can block messages that originate in parts of the network that contain many biased agents. The messages that do circulate contain sufficient information for agents to take them into account when voting on the collective decision.

Second, biased agents wishing to influence a population could be better off limiting their numbers. As unbiased agents are strategic, they block the transmission of opinions that originate in a part of the network that contains many biased agents. Hence, it can serve biased agents to limit their numbers and to spread themselves throughout the network, so as to maximize message transmission.

Third, homophily – the tendency for biased agents to cluster together in the same part of the social network – has a non-monotonic influence on the spread of rumors. When homophily is low, biased agents are dispersed and do not hamper communication though the network; when homophily is high, biased agents are clustered in some areas of the network, and rumors spread without difficulties in the other areas which only contain unbiased agents. Communication is reduced when homophily reaches intermediate levels, and small islands of biased agents can be found throughout the network.

The paper sheds light on why false beliefs can be widely held within a population, such as the belief that President Obama was not born in the United States, the belief that there is a connection between measles-mumps-rubella (MMR) vaccine and austism, and the belief in parts of India that Muslims are responsible for killing cows. Each case fits the basic assumptions of the model. Biased agents desire to spread the false belief, and rumors are initiated or spread by person-to-person communication such as email and conversations. The model indicates, for the first two examples that the belief was spread or maintained thanks to sufficient dispersion of biased agents within the United States population, at the both the right and the left of the political spectrum.[2] As for the third example, the population's initial priors are sufficiently high that even the hint of a cow-killing event sparks the rumor mill (Barstow and Raj (2005)).

The rest of the paper is organized as follows. Section II discusses the paper's contribution to the literature. Section III specifies the basic model of utility and preferences, the two benchmark models of communication, and the possibility of communication in each. Section IV studies maximal communication equilibria in networks, building the algorithm that yields the maximal paths along which unbiased agents are willing to transmit messages. This section also shows that maximal communication yields the highest expected utility for both biased and unbiased agents. Section V studies the number and distribution of biased agents in a network, including how biased agents should best be placed to advance their agenda and how homophily would affect communication. Section VI considers the robustness of the network communication results to incomplete information and to general networks, and Section VII concludes.

II Contribution to the Literature

Relative to previous literature, the innovation of this paper is to study the strategic decision to create and transmit rumors in order to influence public opinion. In its foundation, the model combines two classic elements of information games: "cheap talk" (Crawford & Sobel (1982)) in the decision of the initial receiver of the signal as to whether or not to create a truthful message, and "persuasion" (Milgrom (1981), Milgrom & Roberts (1986)) in the decision of agents who

[2]Fact finding organizations report that the belief gained traction during the primaries for the 2008 United States presidential election, when emails circulated among both Hillary Clinton supporters and Republicans (Farley (2015)), though no smoking gun has been found (to extent we could discern from these organizations' reports). Like the Obama case, the continued anti-vax movement with its various beliefs as to the dangers of this and other vaccines is possibly explained by the dispersion of biased agents (see e.g., Bouie (2015)). On the political right and left, individuals oppose government-mandated vaccinations in the name of personal liberty and primacy of personal health decisions, as well as impugn the profit motives of pharmaceutical firms.

subsequently choose whether to transmit the message, which they cannot transform. We draw on insights from both in the analysis. In our model, there are multiple equilibria, along the lines of cheap talk games. However, as in persuasion games, at the transmission stage agents have an incentive to pass on credible information to other agents. Our analysis focuses on network conditions that allow fully revealing strategies by unbiased agents and identifies the paths in a network along which agents are willing to listen to messages and persuade others.

In a large previous literature, agents somewhat mechanically adopt the opinions of their neighbors and eventually the population converges on a set of beliefs, which could be unduly influenced by a set of well-located biased agents. In one set of models, opinions spread like diseases; i.e., individuals become infected (adopt an opinion) by contact with another agent with that disease (see e.g. Chapter 7 of Jackson (2008)). Such diffusion processes are also studied in computer science, statistical physics, and sociology.[3] In such models, biased agents are always better off when there are more biased agents, in contrast with the present paper. In a second set of models, opinion formation in social networks builds on DeGroot (1974). Agents, with possibly different initial priors, repeatedly "exchange" their beliefs with their neighbors and adopt some statistic (the weighted average, say) of their neighbors' opinions. Such agents fail to take into account the repetition of information that can propagate through a network, leading to a persuasion bias as referred to by DeMarzo, Vayanos, & Zweibel (2003). Golub & Jackson (2010) find sufficient network conditions under which such a naive rule leads to convergence to the truth—there can be no prominent groups, for example, that have disproportionate influence. Research on Bayesian learning in networks (e.g. Gale & Kariv (2003), Bala & Goyal (1998), Acemoglu, Dahleh, Lobel & Ozdaglar (2011)) characterizes convergence or not to common opinions for different network architectures. In our model, there is a single unknown source of information and agents are Bayesian, but due to differences in their preferences and the possibility of falsification and blocking, they may end up with different beliefs and choose different actions.

A new literature studies individuals' incentives to communicate private information to others. Niehaus (2011) adds a cost to sharing information; an agent will weigh the benefits to her friends and neighbors against the personal cost. Other papers study influence in networks; agents all

[3]For a review article in statistical physics, see for example, Castellano, Fortunato & Loreto (2009) For complex contagion models where agents need multiple exposure to become infected see Centola & Macy (2007) and Romero, Meeder & Kleinberg (2011).

have private information and have an incentive to share their information since, for example, agents benefit when others' adopt the same action (Calvó-Armengol, de Martí & Prat (2015), Hagenbach & Koessler (2011), Galeotti, Ghiglino, and Squintani (2013)). Perhaps the closest paper to ours is the recent paper by Chatterjee and Dutta (2016) on credibility and messages in social networks. They also focus attention on the credibility of messages received by agents in a social network when the message can be false. However, their setting is very different from ours. False messages are sent by a single agent, a firm, which lies about the product quality, and messages are mechanically spread by agents who buy the product when the product is of good quality. In contrast to all this work, the present paper features a situation in which information is not widely held and unbiased agents strategically spread information so that a correct public decision is taken. When deciding whether or not to transmit a message, unbiased agents rationally take into account the presence of possibly many biased agents who could have strategically created misinformation or blocked communication of messages that do not match their bias.

A large economic literature also studies the transmission and communication of information through the observation of other agents' actions. Observation helps discern the true state of the world. Knowledge or information costlessly spreads (Banerjee (1992, 1993), Bikhchandani, Hirshleifer & Welch (1992), or spills over, to others, as occurs when people observe others' use of a new technology (e.g., Foster & Rosenzweig (1995), Conley & Udry (2010)). In these models, though individuals influence others through their actions, they derive no benefit in influencing them and, contrary to this paper, any decision to communicate is not strategic.

III Benchmark Models of Communication

A Utility, Beliefs, Strategies and Equilibrium

There is a population of $|N| = n$ agents, and two possible states of nature, $\theta \in \{0, 1\}$. Individual agents earn payoffs from a collective decision, or outcome, which can be understood, for example, as a public policy, a verdict, or election of a particular candidate.[4] Let $x \in \{0, 1\}$ denote the outcome. There are two types of agents, with different preferences. *Unbiased agents*, set $\mathcal{U}$, prefer

[4] This is one particular model in which agents have an incentive to communicate; another simple possibility is that agents take individual actions, such as to vaccinate their children, and unbiased agents are altruistic in sharing their information.

the outcome to match the state of nature and have utility

$$w(x, \theta) = -(x - \theta)^2.$$

Biased agents, set $\mathcal{B}$, prefer outcome $x = 1$ to be implemented, regardless of the state of nature. The utility for a biased agent is

$$v(x, \theta) = -(x - 1)^2.$$

The number of biased and unbiased agents in the population is common knowledge. For any subset of agents S, b_S denotes the fraction of biased agents in S and u_S the fraction of unbiased agents, where necessarily $b_S + u_S = 1$. For any unbiased individual, let $b \equiv \frac{|\mathcal{B}|}{|N|-1}$denote the fraction of biased agents in the remainder of the population.

Agents have a common prior belief that $\theta = 1$ with probability π. We assume $\pi < 1/2$ so that agents initially believe the true state is 0 with higher probability. With this initial prior, agents are particularly interested in credible information that the outcome is 1.

Interaction between agents is divided into three phases: (i) a message creation phase, (ii) a communication phase and (iii) a collective vote phase.

Message creation: With probability $p < 1$, a perfect signal of the true state, $s \in \{0, 1\}$, is generated by nature, and each agent has an equal chance of receiving this signal. The recipient agent–and this agent only–has the opportunity to create a message $m \in \{0, 1\}$. The strategy of the recipient is thus a mapping M_i from the set of states $\{0, 1\}$ to the set of messages $\{0, 1, \emptyset\}$. The empty message $\emptyset$ corresponds to the situation where the recipient chooses not to create any message. With probability $(1 - p) > 0$, no signal is generated.[5] These assumptions ensure that only one message circulates within a population, which allows us to obtain precise and insightful results.[6]

[5]With this set up, the probability p is only relevant for the posterior belief of an agent who does not receive the signal and does not receive a message. For such an agent there are two possibilities. With probability $(1 - p)$ no signal was generated. With probability p a signal was generated but (a) no message was created or (b) a message was created but was "blocked" by other agents. For the agent who receives the signal or who receives a message, p is not relevant since the posterior is conditional on the fact that a signal must have been generated.

[6]Agents receiving multiple, competing messages would have a complex inference problem, but a problem that would still be based on the basic forces highlight in this model. For example, biased agents might want to increase the number of messages they send, but only to the extent unbiased agents believe the messages. Too many messages could lead unbiased agents to discount the messages. Similarly, transforming messages would only be useful to the extent that unbiased agents are convinced of the veracity of the message.

Communication: We consider two benchmark models of communication to be discussed in more detail below. In the *public broadcast* model, the recipient agent broadcasts his message directly to all other agents. In the *network communication model*, agents are organized along a social network. Agent i receives a message $m(j)$ from one of his neighbors j and chooses whether to transmit his message to all other neighbors or not.[7] A transmission strategy t_i maps the message $m(j)$ into $t_i(m(j)) \in \{m(j), \emptyset\}$. When $t_i(m(j)) = \emptyset$ agent i does not transmit, or we say "blocks" message m received from agent j. We let ρ_i denotes agent i's posterior belief that $\theta = 1$ after communication has taken place.

Collective vote: After all possible communication has taken place, agents vote between two alternatives, 0 and 1. We consider a probabilistic voting model, where the probability that an alternative is chosen is increasing in the number of votes received. Let $f(z)$ be the probability that outcome 1 is implemented when z agents vote for outcome 1, with $1 - f(z)$ the probability that outcome 0 is implemented. Given that z other agents vote for $x = 1$, an unbiased agent i with posterior ρ_i has expected utility from voting for $x = 1$ of

$$Ew(x, \theta) = -\rho_i \left(1 - f(z+1)\right) - (1 - \rho_i) f(z+1),$$

where the first term gives the utility loss when the true state is 1 but outcome 0 is implemented, and the second term gives the utility loss when the true state is 0 but outcome 1 is implemented. Similarly, the expected utility from voting for $x = 0$ is

$$Ew(x, \theta) = -\rho_i \left(1 - f(z)\right) - (1 - \rho_i) f(z).$$

We consider pure-strategy perfect Bayesian equilibria of the game. In both communication games, there are possibly many equilibria where agents do not communicate or communication does not contain any information.[8] We are concerned with equilibria where agents communicate as much as possible as defined precisely below. We also show that this maximal communication corresponds to equilibria which maximize the welfare of the agents.

[7] In the benchmark model where the social network is a tree, this assumption of multi-cast communication is made without loss of generality. When an agent has an incentive to communicate with one neighbor, she has an incentive to communicate with all other neighbors. We also note that if an agent only has one neighbor, then he cannot transmit in the communication stage.

[8] See Section V for a discussion of multiplicity of equilibria in our game.

We solve the game by backward induction and start by analyzing the last phase of interaction where agents vote between two alternatives.

B Collective decision

We suppose that the probabilistic voting function is linear, $f(z) = z/n$, so that agents do not vote strategically voting as shown in Lemma 1.[9] Each biased agent votes for outcome 1 regardless of his posterior. Each unbiased agent votes for outcome $x = 1$ if $\rho_i > 1/2$, and votes for outcome 0 if $\rho_i < \frac{1}{2}$, and is indifferent in her vote for $\rho_i = \frac{1}{2}$. We have the following preliminary result:

Lemma 1 *With probabilistic voting, $f(z) = \frac{z}{n}$, it is optimal for unbiased agents to vote according to their beliefs: A Nash equilibrium of the voting game consists of the following strategies: each unbiased agent i votes for outcome $x = 1$ if $\rho_i > 1/2$, votes for outcome $x = 0$ if $\rho_i < \frac{1}{2}$ and votes for 0 and 1 with equal probability if $\rho_i = \frac{1}{2}$. Each unbiased agent votes for outcome $x = 1$.*

In the analysis of communication and message creation which follows, we presume that agents reach the collective decision in such a manner and therefore have an incentive to communicate information that influences others' posteriors and hence their "votes." Unbiased agents' prior beliefs are that state 0 is more likely and therefore vote for outcome 0 if there is no possibility of communication. A particular benefit of communication is then learning that 1 is more likely the true state.

C Communication and Message Creation

As discussed in the Introduction, this paper considers two communication models, which represent two different institutional settings. In keeping with definition of a rumor, in both communication models information circulates but the original source is not known. The *public broadcast model* represents an environment where individuals communicate through broad-based media, such as websites and newspapers. While the message is communicated, the actual identity of the individual who creates the message is hidden and his or her type is not known. The *network model*

[9]The same behavior holds under more general increasing functions f if one assumes agents to be 'naive,' meaning that they do not account for the possible correlation between others' vote and their information on the state: for $\rho_i > 1/2$ ($\rho_i < 1/2$) agent i's utility is larger when he votes for 1 instead of 0 (0 instead of 1) for a fixed z, hence also for any distribution of z provided the correlation between this distribution and the true state is neglected. Such correlation of course matters with common values. For example, this correlation is the basis of the winner's curse in common values auctions or of the strategic behavior of a pivotal voter in the Condorcet jury setting.

represents an environment where individuals communicate privately to colleagues or friends, such as in texts, emails, and phone calls. People know each others' types and can pass on information, but again the original source of the information is unknown (except in one special case described below). We analyze public broadcast first.

C.1 Public Broadcast

In the public broadcast model, agents' types are private information. Any message created is seen by all other agents, but the source is not named and his type not known.[10] An equilibrium consists of message creation strategies M_i for each biased and unbiased agents and posterior beliefs ρ_i for each unbiased agent i such that each agent's strategy is sequentially rational given the beliefs and strategies of others, and beliefs are formed using Bayes rule from the strategies whenever possible. As in cheap talk games, there exist many equilibria, including babbling equilibria where messages do not contain any information about the true state and thus no agents update their priors. Our main interest is equilibria in which unbiased agents, upon receiving the signal, create truthful messages - their messages directly correspond to the true state of the world. We ask, in particular, when full communication is possible - all unbiased agents send messages that match the truth.

Consider the following strategies and beliefs. Strategies: For each unbiased agent i, $M_i(s) = s$, and for each biased agent i $M_i(s) = 1$. Beliefs: For each unbiased agent i, $\rho_i(0) = 0$,

$$\rho_i(1) = \frac{\pi}{b + (1-b)\pi}.$$

and $\rho_i(\varnothing) = \pi$.

It is easy to see that these strategies and beliefs constitute an equilibrium if and only if $\rho_i(1) = \frac{\pi}{b+(1-b)\pi} \geq \frac{1}{2}$ so that agents who receive message 1 believe that state 1 is more likely than state 0 and therefore vote for outcome 1 rather than outcome 0.

Working backwards, first consider the agents' posteriors beliefs.: On the equilibrium path, the posteriors are formed by Bayes' rule; $m = 0$ is created only by unbiased agents, the probability that the originator of message $m = 1$ is biased is b, and receiving no message occurs in the event

[10]Equivalently, agents know the identity of the creator of the message but do not condition their posterior on the agent's identity.

no agent receives a signal, hence the prior is maintained. Off the equilibrium path, no message is received in which case we set the posterior to be equal to the prior.

Turning to agents' strategies to create messages, given these posterior beliefs, no unbiased agent that receives the signal would choose to send an untrue message $M(s) \neq s$, since this action will decrease the number of agents that vote for the outcome corresponding to the true state. No biased agent has an incentive to deviate and choose $m = \varnothing$ or $m = 0$, since these actions will decrease the posterior belief that 1 is the true state.

It is also easy to see that no other equilibrium can yield higher expected utility for unbiased agents and that no partial communication equilibria exist. Consider the possibility that in equilibrium a strict subset of unbiased agents send truthful messages. One of the unbiased agents who do not send a message or an untruthful one would have an incentive to deviate and send a truthful message, since it increases the likelihood of the correct outcome.

These arguments give us our first result concerning communication. Rewriting the condition on the posterior $\rho_i(1)$ we see that communication occurs in the pubic broadcast model if and only if the proportion of biased agents in the population if sufficiently low.

Proposition 1 *In a public broadcast game, an equilibrium exists where all unbiased agents broadcast truthful messages if and only if*

$$\frac{\pi}{1-\pi} \geq b. \tag{1}$$

This equilibrium maximizes unbiased agents' expected payoffs. It is an equilibrium for no unbiased agents to broadcast messages, but there is no equilibrium where a strict subset of unbiased agents broadcast truthful messages.

Proof. Proofs of all results are provided in the Appendix.

The network model, which we examine next, differs from public broadcast in two ways. First, agents communicate pairwise and, second, agents' types are common knowledge. The combination of these two features can allow communication under conditions in which no truthful communication is possible in public broadcast. With pairwise transmission of messages and knowledge of where agents of different types are located, agents can judge the veracity of a message transmitted through a particular part of the network and block it from reaching others. This blocking then increases the veracity of other messages which then circulate.[11]

C.2 Network Communication

In the *network model* a pair of agents i and j have a link, denoted ij, if they have the potential to communicate, and we say agent i is agent j's *neighbor* and vice versa. The communication could occur in either, both, or neither direction. To indicate the direction of any communication, let (i, j) denote the directed link from i to j, (j, i) denote the directed link from j to i, and let G denotes the set of all directed links. G and individual agents' types are common knowledge. We assume that the network is connected; i.e., every agent has at least one neighbor.

To generate insights and analytical results, we suppose throughout the paper that agents are connected in such a way that a message can reach any individual through only one route. That is, the network is a *tree*, where there is a unique path from any agent i to any agent j. With a tree, we can neatly parse the network and study agents' posterior beliefs as to the veracity of a received message. Section VI discusses general networks.

An equilibrium in the network model consists of message creation strategies, transmission strategies, and beliefs (M_i, t_i, ρ_i) for each agent i such that each agent's strategy is sequentially rational given the beliefs and strategies of others, and beliefs are formed using Bayes rule from the strategies whenever possible.

We first study *full communication equilibria*, where all unbiased agents create truthful messages, $M_i(s) = s$ and transmit any message received from a neighbor, $t_i(m(j)) = m(j)$ (when they have more than one neighbor). Full communication is not always an equilibrium. The following simple example illustrates the impact of the prior and of the placement of biased agents on the existence of full communication equilibria..

Example 1 *Consider 5 agents in a line, as shown at the top of Figure 1, where the links allow communication in either direction, with 4 unbiased agents and 1 biased agent in the middle.* $\pi = 0.3$*. Assume that unbiased agents create true messages upon receiving a signal and transmit any message they receive. Can these strategies form an equilibrium? Message 0 does not raise*

[11]The network model-but with incomplete information-could be interpreted as the public broadcast model with information flow for each agent; that is, any agent only knows the source of the message is someone who precedes him in the information flow. With incomplete information, the proportion of biased agents in that flow is the same of the proportion of biased agents in the population. With complete information, the agent can adjust his posterior given the proportion of biased agents for any particular segment of the population through which information flows.

a difficulty since it can be created only by unbiased agents, hence reveals that the state is 0 so that unbiased agents transmit it. Similarly, message 1 received by $U2$ *from* $U1$ *reveals that the state is 1 since* $U2$ *knows that* $U1$ *is unbiased. Consider now message 1 received by* $U4$ *from* $B3$*. The message may have been created by* $B3$ *(whatever the signal he has received); it may also have been created by* $U2$ *if he has received signal 1, and then transmitted by* $B3$*, or finally it may have been created by* $U1$ *if he has received signal 1, and then transmitted by* $U2$ *and* $B3$*. Though* $U4$ *receives message 1 from* $B3$ *who he knows is biased,* $U4$ *receives some relevant information: applying Bayes rule,* $U4$*'s posterior belief is equal to* $\frac{3\pi}{2\pi+1}$ *which is larger than 1/2 (*$\pi = 0.3$*):* $U4$ *believes that the state 1 is more likely than state 0, hence finds it optimal to vote for 1 and to transmit the message to* $U5$*. The argument works in all other situations. Hence these strategies in which all unbiased agents transmit all messages form an equilibrium. This equilibrium will be called a Full Communication Equilibrium (FCE).*

Now consider the network in which $U4$ *and* $B3$*'s positions are exchanged, as shown at the bottom of Figure 1. Now* $U4$*'s posterior belief upon receiving message 1 from* $B3$ *is equal to* $\frac{2\pi}{\pi+1}$*, which is smaller than 1/2:* $U4$ *blocks the message: there is no FCE. This shows that both the prior* π *and the placement of agents influence the existence of a FCE.*

Observe that communication, though partial, may still exist; for example let $U1$*,* $U2$ *and* $U4$ *create truthful messages and transmit messages. In that case, when they receive message 1 from either of them, they are certain that the state is 1. Furthermore, when* $U5$ *receives message 1 he believes that state 1 is more likely (with a posterior equal to* $\frac{4\pi}{3\pi+1}$*). It is easy to check that an equilibrium is obtained. Furthermore, there is maximal communication, in a sense made precise in Theorem 2.*

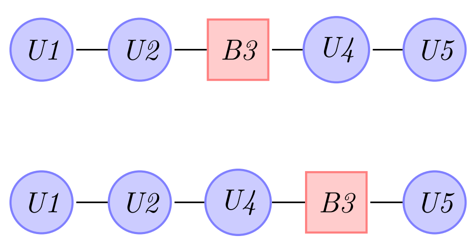


Figure 1: Five Agent Line with One Biased Agent

Before proceeding with specifying particular strategies, notice that for any strategies, the event $\varnothing$ occurs with positive probability since no signal is received by any agent with probability

$(1-p) > 0$. Notice also that if an agent receives a message, only in one special case can the agent directly infer who received the signal; if i receives a message from a neighbor j who has no other neighbors, then i can infer that j received the signal. Otherwise, the source is not known.

In the general case, consider the following strategies.

Strategies: Upon receipt of the signal, every biased agent i creates a message that matches their bias, i.e., $M_i(s) = 1$. Every biased agent only transmit a message if the message is 1; i.e., $t_i(0) = \varnothing$, $t_i(1) = 1$. Every unbiased agent i creates a true message upon receiving a signal; i.e., $M_i(s) = s$, and transmits any received message, i.e., $t_i(m) = m$.

To construct the posterior beliefs $\widetilde{\rho}_i(m(j))$, we define the following elements of a network. Consider the directed edge (j, i). Since the network is a tree, agents in the network can be divided into two disjoint subsets, with one subset on either side of the edge. Let $S_i(j)$ be the set of agents whose messages can reach i by going through j (this set includes j). The set $S_i(j)$ corresponds to the nodes in the oriented subgraph of G flowing toward i and ending with the directed edge (j, i); we denote this oriented subgraph $G_i(j)$. Figure 2 illustrates. The other set $S_j(i)$ is the set of agents whose messages can reach j by going through i (this set includes i). $G_j(i)$ is defined similarly.

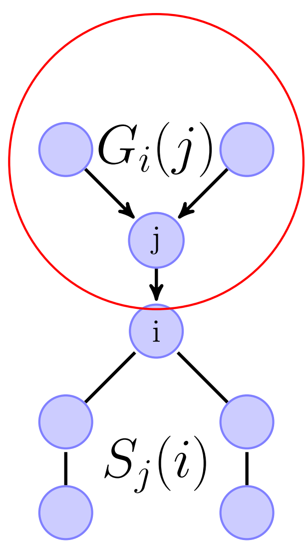


Figure 2: Decomposition of the tree

Beliefs: (1) For an agent i who has received a message $m = 0$ from an unbiased neighbor j,

$\widetilde{\rho}_i(0(j)) = 0$. (2) For an agent i who has received a message $m = 1$,

$$\widetilde{\rho}_i(1(j)) = \frac{\pi}{b_{S_i(j)} + u_{S_i(j)}\pi}, \tag{2}$$

where $b_{S_i(j)}$ is the proportion of biased agents in $S_i(j)$ and $u_{S_i(j)}$ is the proportion of unbiased agents. (3) For an agent i who receives no message, $\rho_i(\varnothing) \leq \pi$. (4) For an agent i who receives a message 0 from a biased neighbor j, $\rho_i(0(j)) = \pi$ for all $j \in \mathcal{B}$.

These strategies and beliefs constitute an equilibrium of the network game when the posterior belief of an agent receiving message $m(j) = 1$ is willing to pass it on; ie., when $\widetilde{\rho}_i(1(j)) \geq 1/2$. To see this, consider first beliefs: (1) and (2) derive from Bayes' rule. For belief (3), receipt of no message occurs on the equilibrium path with probabiity $(1 - p) > 0$ and would occur off the equilibrium path if a biased agent fails to pass on a message 1 or an unbiased agent fails to pass on a message. A posterior of $\rho_i(\varnothing) = \pi$ is consistent with all these scenarios. Second, given agents' beliefs, unbiased agents increase their expected utility by passing on message 0 and passing message 1 as long as $\widetilde{\rho}_i(1(j)) \geq 1/2$. Biased agents increase their expected utility by increasing the posterior for outcome 1, and hence would not benefit from passing on message 0 or from blocking a message 1.

Rewriting the key condition for the full communication equilibrium, we have the following result:

Theorem 1 *In the network model, a full communication equilibrium (FCE) exists if and only if for each unbiased agent i and each of his neighbors j:*

$$\frac{\pi}{1 - \pi} \geq b_{S_i(j)}. \tag{3}$$

This condition holds only when there are sufficiently few biased agents in all subgraphs of the network $S_i(j)$ for all unbiased agents i with neighbor j.[12] Roughly speaking, a full communication equilibrium exists if biased agents are few in number and dispersed through the network. We explore this insight further in Section V.

[12] The condition obviously fails, for example, when an unbiased agent is connected to a "leaf" of the network since then $b_{S_i(j)} = 1$.

IV Maximal Communication in Networks

In this section we construct strategies allowing for maximal communication among unbiased agents and prove that they form an equilibrium. Of course, these strategies coincide with those of the full communication equilibrium when it exists.

Here, we parse the network and construct an algorithm to find the subgraphs within which communication can occur. These subgraphs are directed and represent paths along which agents are willing to transmit messages, since the agents believe the message with sufficiently high probability. The algorithm eliminates directed edges from G, and we denote the remaining set of directed edges G^*. In the strategies constructed below, unbiased agents transmit messages from a neighbor j if and only if the directed link (j, i) is contained in G^*. We show that these strategies maximize the possible communication in the graph at an equilibrium and yield the highest possible expected utility for unbiased agents.

A Algorithm to Identify Subgraphs of Transmission

When a full communication equilibrium does not exist, $b_{S_i(j)} > \frac{\pi}{1-\pi}$ for at least one unbiased agent i and directed edge (j, i) (Theorem 1). In this case, let V be the (non-empty) set of all directed edges in G that violate the condition $b_{S_i(j)} \leq \frac{\pi}{1-\pi}$. The algorithm will eliminate a subset of V: In the process, some edges in V may become non-violating and will not be eliminated; on the other hand all non-violating edges will remain non-violating, so that V is the maximal set of edges that can be eliminated.

Consider the following algorithm. A directed edge $(j, i) \in V$ is said to be of *level 1* in G if there is no directed edge (k, l) such that $(k, l) \neq (j, i)$ in $V \cap G_i(j)$.[13]

Pick one level 1 edge $(j, i) \in V$. Remove (j, i) from G and let

$$\begin{aligned} G^1 &= G \setminus (j, i), \\ \Gamma^1 &= G \setminus G_i(j). \end{aligned}$$

For each unbiased agent l and directed edge (k, l) in Γ^1, let $S_l^1(k)$ be the set of agents whose message 1 can reach l through k in Γ^1. Compute the proportion $b_{S_l^1(k)}$ of biased agents in that

[13] There is always at least one level 1 edge because the graph G is finite.

set, $b_{S_l^1(k)} = \frac{|B \cap S_l^1(k)|}{|S_l^1(k)|}$, and define V^1 to be the set of directed edges in Γ^1 such that $b_{S_l^1(k)} \geq \frac{\pi}{1-\pi}$. If the set V^1 is empty, the algorithm stops.

Otherwise, pick a directed edge (k, l) of V^1 which is of level 1 in the graph Γ^1. Remove (k, l) from G^1 to obtain G^2 and let $\Gamma^2 = \Gamma^1 \setminus G_l(k)$, and search for violating edges (if any) in Γ^2.

In the general step $t + 1$ of the algorithm, given directed graphs G^t, Γ^t and a non-empty set V^t of violating edges in Γ^t, pick a level 1 violating edge (a, b). Eliminate this edge from G^t to obtain G^{t+1}. Define $\Gamma^{t+1} = \Gamma^t \setminus G_b(a)$ and accordingly the proportions $b_{S_b^{t+1}(a)}$ of biased agents in Γ^{t+1}. Let V^{t+1} be the set of edges (j, i) in G^{t+1} such that $b_{S_i^{t+1}(j)} > \frac{\pi}{1-\pi}$.

If the set V^{t+1} is empty, the algorithm stops and define $G^* = G^{t+1}$. Let W be the set of directed edges that have been eliminated by the algorithm, $G^* = G \setminus W$.

Lemmas in the Appendix show (i) that the sequence $\{V^t\}$ is decreasing, (ii) the set W does not depend on the order links are chosen to be eliminated and (iii) for any $(j, i) \in W$, j is biased (and i is unbiased by definition of a violating edge). That is, the algorithm parses the network into directed links of biased and unbiased agents. We illustrate the output of the algorithm in the following example:

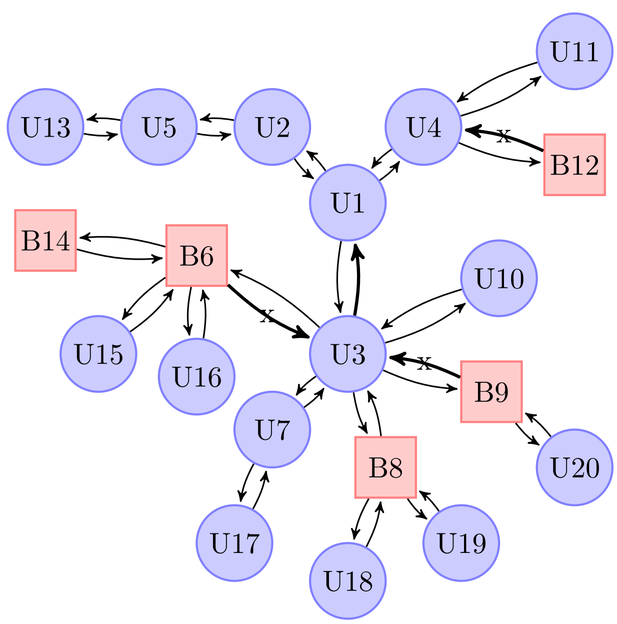


Figure 3: Complex Network and Algorithm

Example 2 *Subgraphs of Communication. The network in Figure 3 was generated by a random* *process for 20 agents, with an overall target fraction of 0.3 biased agents. Let the initial belief be $\pi = \frac{1}{7}$. Given π, the threshold for the proportion of biased agents is $b = \frac{\pi}{1-\pi} = \frac{1}{6}$. The edges in bold are the violating edges in G. They are all of level 1 except (U3-U1) which is of level 2. The edges which are crossed out are the edges in W which are eliminated by the algorithm. G^* does not contain the edges (B9-U3) (B12-U4) and (B6-U3), since messages flowing from these biased agents would not be believed. The edge (U3-U1) is violating in the original graph but not in the final graph because the proportion of biased agents whose messages would flow through U3 to the rest of the graph decreases from $\frac{4}{13}$ to $\frac{1}{7}$.*

B Maximal Communication Equilibrium: Strategies and Beliefs

We next construct strategies and beliefs that constitute an equilibrium in which communication flows along all edges except those in W; i.e. along edges in G^*.

Strategies: Biased agents, upon receipt of the signal, create a message that matches their bias, i.e., $M(s) = 1$. Biased agents only transmit messages that match their bias, i.e., $t(0) = \oslash$, $t(1) = 1$. Unbiased agents, upon receipt of a signal, create true messages; i.e., $M(s) = s$. All unbiased agents i transmit message 1 received from agent j if $(j, i) \in G^*$, otherwise agent i does not transmit the message. All unbiased agents transmit messages $m = 0$ received from any agent.

Beliefs: For information sets which are reached with positive probability, beliefs follow Bayes rule consistent with the above strategies: (1) For an agent i who has received a message $m = 0$ from an unbiased neighbor j, $\widetilde{\rho}_i(0(j)) = 0$, since only unbiased agents create and transmit $m = 0$, and they create truthful messages. (2) For an agent i who has received a message $m = 1$ from a neighbor j, her beliefs reflect the strategies of unbiased agents to only submit messages along edges in G^* and to not transmit messages otherwise. Posteriors are then given by $\widetilde{\rho}_i(1(j)) \geq 1/2$ for $(j, i) \in G^*$ and $\widetilde{\rho}_i(1(j)) < 1/2$ for $(j, i) \notin G^*$. (3) For an agent i who receives no message, her posterior beliefs take into account the probability that no signal has been generated and the fact that biased and unbiased agents block messages that originate in particular parts of the network. These posteriors are surely less than π.[14]

[14]To see this, consider first the impact of an agent i who does not transmit a message 1 from j for $(j, i) \notin G^*$. Recall that j is biased. So, not only does i not transmit $m = 1$ received from j, but i never transmits $m = 0$ from j because j, being biased, does not create or transmit 0: all the signals received by an agent in $S_i(j)$, be them 0 or 1, are lost for the other agents, those in $N - S_i(j)$. As for biased agents, they block $m = 0$, which, by the strategies, is circulated only when the true state is 0. This implies that the posterior can only be lower than π: $\rho_i(\varnothing) < \frac{1}{2}$.

The only event for which beliefs need to be specified is when an agent receives a message zero from a biased agent.[15] As previously, we suppose i's posterior belief is equal to his prior in this case; i.e., $\rho_i(0(j)) = \pi$, for all $j \in \mathcal{B}$.

These strategies constitute an equilibrium of the network game. Furthermore, communication is maximal among all equilibria. The intuition is as follows. If (j,i) is a violating edge of level 1, whatever behavior of the unbiased agents in $S_i(j)$ at an equilibrium, the proportion of biased agents whose message 1 can reach i through j can only be equal or larger than $b_{S_i(j)}$. Hence i's posterior belief is smaller than 1/2,[16] and in no equilibrium communication flows from j to i. Inspection of the algorithm shows that all violating edges of level 1 in G are eliminated. Recursively, the edges in W—eliminated by the algorithm—are edges along which communication surely never flows in any equilibrium.

Theorem 2 *The above strategies and beliefs form an equilibrium of the network game. We call this equilibrium the "maximal communication equilibrium"(MCE) as communication is maximal among all equilibria in the following sense: in any equilibrium, if $(j,i) \notin G^*$ (equivalently $(j,i) \in W$), then j is biased and i does not transmit $m = 1$ received from j.*

C The network as a filter: public broadcasting vs. network communication

The above analysis shows that when full communication is not possible in the public broadcast model, communication is still possible in a network. In the network, unbiased agents block messages from certain parts of the network, limiting the influence of localized biased agents. The network serves as a filter, allowing for credible communication in the rest of the network. In particular, two unbiased agents always communicate to each other in an MCE since the corresponding edges are not in W. We have the following result, comparing communication with public broadcast and in a network:

Proposition 2
- *If $\frac{\pi}{1-\pi} \geq \max_{(j,i)} b_{S_i(j)}$ full communication is an equilibrium in both the public broadcast and the network models.*

[15] The event in which an unbiased agent receives a message from an unbiased neighbor has positive probability: Even if the neighbor's strategy is to block messages from all his neighbors he still has the strategy to create a true message if he receives the signal.

[16] Consider an equilibrium. Let S be the subset of agents whose message 1 can reach i through j. S is a subset of $S_i(j)$ and i's posterior upon the receipt of message 1 from j is equal to $\frac{\pi}{b_S + u_S \pi}$. We show that if (j,i) is a violating edge of level 1 in G, the proportion of biased agents b_S can only be larger than $b_{S_i(j)}$. Hence i's posterior is less than $\frac{\pi}{b_{S_i(j)} + u_{S_i(j)} \pi}$, which is strictly less than 1/2 by definition of V.

- *If* $\max_{(j,i)} b_{S_i(j)} \geq \frac{\pi}{1-\pi} \geq b$, *full communication is an equilibrium in the public broadcast model, but not in a network.*

- *If* $b \geq \frac{\pi}{1-\pi}$, *no communication occurs in equilibrium in the public broadcast model whereas partial communication exists in equilibrium in the network game as long as at least one unbiased agent is linked to another unbiased agent.*

D Multiplicity of equilibria and optimality property of an MCE

This section discusses other equilibria in our network model, relating the analysis to cheap talk and persuasion games. We show that the MCE is Pareto optimal for the unbiased agents and provides a refinement criterion, referred to as *activity* that distinguishes among equilibria.

First, as in cheap talk games, there are babbling equilibria in which no valuable information is created. Suppose each unbiased agent who has not received the signal takes the same action independent of any message received, and votes for 0 according to his prior. In this case, all unbiased agents are indifferent between all actions: creating, or not, true or false messages and transmitting, or not, messages. A simple equilibrium then consists of the following strategies: Unbiased agents never create or transmit any messages, and biased agents always create $m = 1$ upon receipt of the signal, and transmit any $m = 1$, but no other message.[17] The only messages that are generated are those from the biased agents, and hence they are not informative. These strategies form an equilibrium supported by (consistent) posterior beliefs equal to the prior, except for the agent who has received the signal.

Second, there are equilibria where unbiased agents create truthful messages but do not transmit credible messages. These equilibria involve a coordination failure and cannot easily be eliminated using standard selection arguments.The standard perfection argument which generates transmission in a persuasion game does not hold in our model. Because of the presence of biased agents, messages are not perfectly informative and it may be rational not to transmit message 1. This is illustrated by the following example.

Example 3 *Consider 5 agents in a line again, as shown in the top of Figure 1 with agents 1 and 2 unbiased, 3 biased, and 4 and 5 unbiased. Consider* $\pi \geq 1/4$, *in which case there is a full*

[17] More formally consider the following strategies. For message creation, biased agents adopt the strategy $M(s) = 1$ and unbiased agents adopt the strategy $M(s) = \varnothing$. For transmission, biased agents adopt the strategy $t(m) = \varnothing$ for $m = 0$ and $t(m) = 1$ for $m = 1$. Unbiased agents have the strategy $t(m) = \varnothing$ for all m.

communication equilibrium. (The largest proportion of biased to unbiased agents in any subgraph $S_i(j)$ is $1/3$.) Change the strategies of the FCE as follows: $U2$ does not transmit message $m = 1$ received from $U1$ to $B3$; $U4$ does not transmit any message from $B3$. Note that all unbiased agents still create truthful messages. It is easy to check that these strategies form an equilibrium for $\pi \leq 1/3$: When $U4$ receives $m = 1$ from $B3$, the proportion of biased agents among the initiators is $1/2$ (instead of $1/3$ in the FCE), so the posterior on the true state being state 1 is lower than 1/2. $U2$ has no incentive to transmit $m = 1$ received from $U1$ since it will not influence the vote of $U4$, who maintains his prior upon receipt of any message from $B3$. Since for $U4$ receiving $m = 1$ from $B3$ is on the equilibrium path, a perturbation argument does not destabilize this equilibrium. This equilibrium exhibits a coordination failure, and the equilibrium payoffs of unbiased agents are lower than the payoffs in an FCE. In the FCE, if a signal that the state is 1 is received by any agent, all agents receive a message $m = 1$ and all agents vote for 1. If a signal that the state is 0 is received by any unbiased agent, an unbiased agent either receives the signal, receives a message $m = 0$, or receives no message (as it is blocked by B3). Hence all unbiased agents vote for 0 and the biased agent votes for 1. The only mismatch between the circulated message and the state occurs when the state is 0 and a signal is received by B3; all agents vote for 1 in this case. The expected loss (for $p \to 1$) for unbiased agents is therefore $\frac{4}{5}(1-\pi)\frac{1}{5} + \frac{1}{5}(1-\pi)\frac{5}{5}$. In the above equilibrium with coordination failure, unbiased agents $U4$ and $U5$ do not update their priors in the events that a signal $s = 1$ is received by agents U1, U2, or B3. They vote for 0 in these events, and their votes do not match the state. On the other hand, $U4$ and $U5$ also do not change their prior in the event $s = 0$ is received by B3 (who then sends message $m = 1$). The expected loss (for $p \to 1$) for unbiased agents is therefore $\frac{3}{5}\pi\frac{2}{5} + \frac{4}{5}(1-\pi)\frac{1}{5} + \frac{1}{5}(1-\pi)\frac{3}{5}$. As $\pi \geq \frac{1}{4}$, the expected payoff of all unbiased agents is higher in the FCE than in the alternative equilibrium.

To refine the equilibrium set, consider restricting attention to the following simple strategies. A biased agent is *active* if and only if she creates message $M(s) = 1$ and only transmits message 1. An unbiased agent is *active* if and only if she creates a message that matches the signal and transmits message m (when she has more than one neighbor) if she thinks the probability that the true state is m is higher than $\frac{1}{2}$. This refinement allows us to single out the MCE.

Proposition 3 *The MCE is the only equilibrium where all agents are active.*

In an equilibrium where all agents are active, coordination failures are ruled out both at the message creation and transmission stages. This results in the highest expected payoff for the unbiased agents.

Theorem 3 *Among all equilibria, the MCE yields the highest expected payoffs for unbiased agents.*

The utility of unbiased agents in the equilibria can be partially ranked. By the arguments of Theorem 3, the expected utility of unbiased agents is the highest in the maximal communication equilibrium (i.e., in the FCE when it exists) and is the lowest in an equilibrium without communication. Since the prior is lower than 1/2, without communication unbiased agents vote for outcome 0. They pass along sufficiently credible information that leads them to vote for outcome 1 rather than outcome 0. Therefore, equilibria with communication involve greater probability of implementation of outcome 0. As a result, biased agents also prefer the maximal communication equilibrium to any equilibrium with partial communication, and any equilibrium with partial communication to an equilibrium without communication. Furthermore, both biased and unbiased agents would prefer network communication to public broadcast for the lower values of π, shown in Proposition 2, where communication does not occur in public broadcast

V Number and Distribution of Biased Agents in a Network

In this section we apply our analysis to questions concerning the number and distribution of biased agents in a network. Our first exercise is a comparative static on the number of biased agents; we replace one biased agent with one unbiased agent and consider the effect of this replacement on individual welfare. Our second exercise takes the point of view of biased agents and considers the optimal number and placement of biased agents to maximize the probability that outcome 1 is implemented.[18] Finally, we consider how the distribution of agents affects communication; in particular we study homophily, the extent to which biased agents are more likely to be linked to each other than to biased agents.

[18]Chatterjee and Dutta (2016) study a firm's decision to locate one biased agent ("an implant") in a network, who communicates possibly false information about, say, product quality, to his neighbors. If this information is sufficiently credible, it will be spread among other agents, who are all unbiased and rational. The optimal placement of the implant depends on the quality of the firm's product and the shape of the network.

A Number of Biased Agents in a Network

The replacement of an unbiased agent by a biased agent j increases the utility of unbiased agents but could increase or decrease the expected utility of biased agents. There are three effects from such a replacement: (1) a direct effect on the number of votes for collective action 1, (2) a direct effect on information transmission because a message $m = 1$ is always created when the signal is received by agent j, and (3) an indirect effect on information transmission as messages $m = 1$ are more likely to be blocked by unbiased agents, since the message is less likely to be credible.

For unbiased agents who receive the same message as the unbiased agent whose status has switched, all effects concur to reduce expected utility.

Proposition 4 *Consider two assignments of biased and unbiased agents in the network, σ and σ' such that one unbiased agent under σ is replaced by a biased agent in σ'. Then the expected utility of any unbiased agent in the MCE in σ' is lower than at the MCE under σ.*

For biased agents, there is a tradeoff. Both direct effects result in an increase in expected utility, but the indirect effect may induce a decrease in the number of unbiased agents who receive and believe message $m = 1$. As the following example shows, the indirect effect may dominate the two direct effects so that the replacement of an unbiased agent by a biased agent may reduce the utility of biased agents.

Example 4 *Placement of Biased Agents on a Line. As in Figure 2, consider eight agents arranged on a line. Under the assignment σ, agents $1, 2, 3, 4, 6, 7, 8$ are unbiased and agent 5 is biased. Under the assignment σ', agents $1, 2, 3$ and $6, 7, 8$ are unbiased and agents $4, 5$ are biased. Suppose that the prior π satisfies $\frac{1}{5} \leq \pi < \frac{2}{7}$. Then, in the MCE under σ, message $m = 1$ is believed and transmitted by all unbiased agents – the MCE is an FCE, whereas in the MCE under σ', message $m = 1$ received from agent 4 is not believed by agent 3 and message $m = 1$ received from agent 5 is not believed by agent 6. A simple computation shows that the expected loss of a biased agent under σ (as $p \to 1$) is*

$$\mathcal{L} = \frac{7}{8}\frac{7(1-\pi)}{8} = \frac{49(1-\pi)}{64},$$

whereas the expected loss of a biased agent under σ' is

$$\mathcal{L}' = \frac{3}{8}\frac{6\pi}{8} + \frac{6}{8}\frac{2+6(1-\pi)}{8} = \frac{48-18\pi}{64}.$$

For values $\pi \in [\frac{1}{5}, \frac{2}{7})$, $\mathcal{L}' > \mathcal{L}$.

The negative effect of adding biased agents stands in sharp contrast to models of rumors and opinion formation based on fixed laws of diffusion or adoption. In such models, it is always beneficial for biased agents to increase their numbers. Here, where agents strategically transmit messages from others, the introduction of a biased agents can reduce their expected utility, depending on where the agent is located in the network.

B Optimal Placement of Biased Agents

The above observation in turn raises the following question: If biased agents could coordinate and place k biased agents in the network in order to maximize the expected probability that collective action 1 is taken, what is the optimal number of agents and where should they be placed? This question is computationally complex in a general tree network, but we can analyze this problem when the network is a line. Consider n agents located in a line, and let $k^* = \frac{n\pi}{1-\pi} + 2\pi - 1$ so that, in the MCE, unbiased agents will transmit a message from a subset of agents that contains at most k^* biased agents. If $k \leq k^*$, there is a full communication equilibrium when biased agents are spaced evenly at locations: $\{\llcorner \frac{n-k}{k+1} + 1 \lrcorner, \llcorner \frac{2(n-k)}{k+1} + 2 \lrcorner, ..., \llcorner \frac{k(n-k)}{k+1} + k \lrcorner\}$. If $k > k^*$, there is a maximal equilibrium with partial communication when $k - k^*$ biased agents are located at the end of the line, and the remaining k^* biased agents are spaced evenly along the line at locations $\{k - k^* + \llcorner \frac{n-k}{k+1} + 1 \lrcorner, k - k^* + \llcorner \frac{2(n-k)}{k+1} + 2 \lrcorner, k - k^* + \llcorner \frac{k(n-k)}{k+1} + k \lrcorner\}$.

The above analysis provides an upper bound on the number of biased agents on a line for full communication, and characterizes uniform spacing as an optimal way for an operator to implant biased agents in the network. The uniform spacing may not be the only optimal location strategy of the operator. For example, if $k = 1$ and π is close to $\frac{1}{2}$, all unbiased agents will transmit a message that could have originated from the biased agent wherever she is located, except at the end of the line. But, as π decreases, unbiased agents are less likely to transmit a message that could have been created by a biased agent, and in the end, the only way to guarantee that the biased agent's message is transmitted is to locate the biased agent exactly in the middle of the line.

C Homophily Simulations

This section studies the possibility of homophily, where biased agents are more likely to be linked to biased agents. While the theory does not capture this feature, simulations indicate that homophily has non-monotonic effects on communication in networks. When homophily is low, biased agents are spread in the population and communication occurs in most of the network. Communication is high, but a significant fraction of messages could originate from biased agents and could be false. As homophily increases, unbiased agents can filter messages from parts of the network where there are many biased agents and ultimately, when biased agents are completely clustered together, no messages are transmitted from those branches of the network.

Figure 4 is produced from simulations where we create a set of approximately 450 random twenty-agent trees, by starting with one node and successively attaching agents to existing nodes. The degree of each node is drawn from a Poisson distribution with parameter 3. For the simulations depicted in the Figure, each agent that is added is biased with probability $b = 0.4$ (the realized proportion of biased agents could be higher or lower). In the realized network, homophily for each agent is simply the number of agents of the same type divided by the number of the agent's neighbors. Averaging across all agents yields the "average homophily" of the network, on the x-axis of the Figure.[19] To measure communication for each network, we successively pick one unbiased agent to be the signal recipient and use the algorithm above to compute the total number of unbiased agents who would receive a truthful message. Averaging over all unbiased agents in the tree gives the number on the y-axis of the Figure.

The estimation of quadratics reveals a U-shaped relation between homophily and communication. The curves give the best fitting quadratic and the quadratics formed by the estimated parameters plus one standard deviation, and the estimated parameters minus one standard deviation for each of the parameters. For an homophily level between 0.4 and 0.8, communication is very weak. However, at the highest levels of homophily, biased agents are highly linked to each other in the network and there are clusters of biased and unbiased agents. Unbiased agents then cut communication from their biased neighbors but transmit the messages received from their unbiased neighbors. Thus, highly homophilous societies are likely to have greater communication

[19] As this random generation method will typically produce networks with intermediate levels of average homophily, we also use a method to generate trees with extreme homophily by fixing a target homophily level which is used each time a new agent is linked to an existing node.

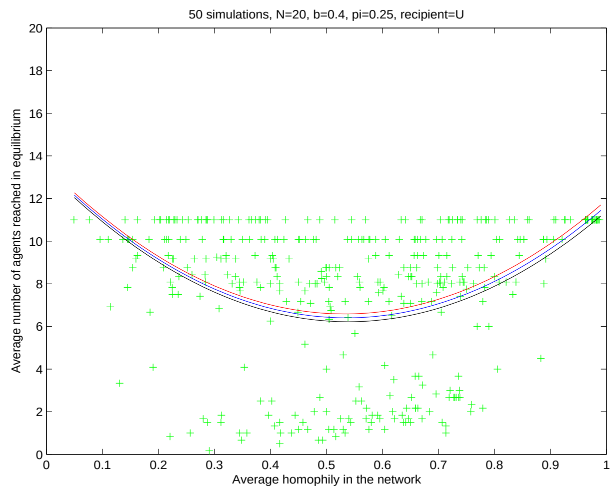
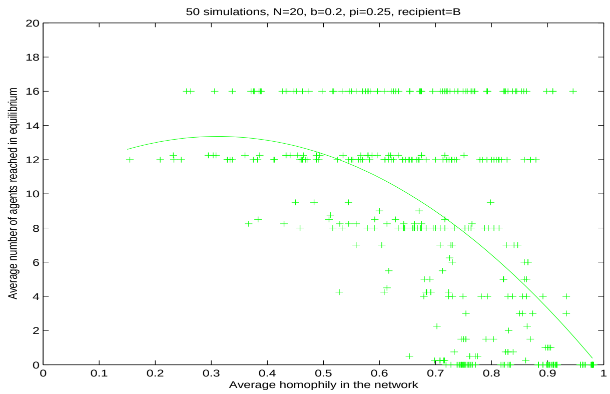
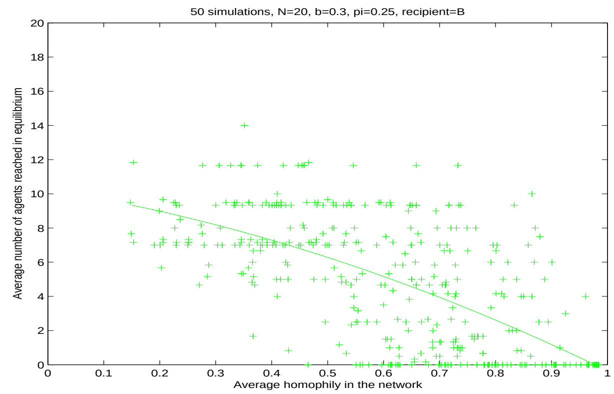


Figure 4: Homophily and communication

and fewer false rumors.

Figure 5 represents similar simulations when the signal (0 or 1) is instead received by a biased agent and illustrates the fact that the biased agents can be hurt when their proportion increases. The three plots correspond to the three values of b: 0.2, 0.3, and 0.4 and show that the biased agents are less successful when their proportion increases. At a homophily level of 0.4, for example, as b increases from 0.2 to 0.3 to 0.4, the number of U recipients decrease from 13 to 7 to 4. Correspondingly, the proportion of U agents receiving the message falls from 13/16, to 1/2, to 1/3.





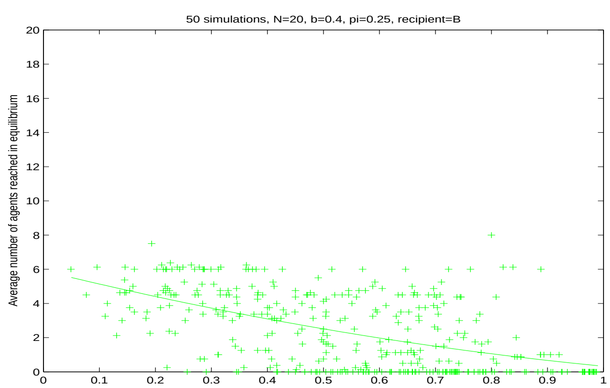


Figure 5: Homophily and communication when B receives the signal

VI Robustness of the Network Model Results

In this section we consider the robustness of our results to relaxing key assumptions in our network model. First, we consider incomplete information about agents' types. We then consider general networks, that is, networks that are not trees and thus can contain possibly many paths from one agent to another along which messages can travel.

A Incomplete Information

The above analysis assumes that agents' types and the network structure are common knowledge. The assumptions allow us to obtain precise analytical results but they are strong and unlikely to be satisfied in many applications. Here we consider the possibility that agents only have local information about the network; for example agents know the identity and types of their neighbors but they know only the proportion of biased and unbiased agents in different branches of the network. In this case, agents would have private information about their own types and private information about types in their own neighborhood, but this private information is not independently distributed. Thus, the game has a possibly complex information structure; it becomes a game of incomplete information where agents have correlated private information.

One information structure, however, is simple to analyze: Suppose that agents only know the fraction of biased and unbiased agents who are connected to them through any of their neighbors. When the conditions for an FCE are satisfied, this information is sufficient to guarantee that all agents have an incentive to transmit information when all other agents transmit information as well, i.e. that an FCE exists.

Theorem 4 *Suppose that agent i only knows his neighborhood N_i and the fraction of biased agents accessing i through any nieghbor j, $b_{S_i(j)}$. Then a full communication equilibrium (FCE) exists if and only if for each unbiased agent i and each of his neighbors j:*

$$\frac{\pi}{1-\pi} \geq b_{S_i(j)}. \tag{4}$$

The extension to an MCE is more complex. Theorem 1 does not require complete knowledge of the network and of the distribution of biased and unbiased agents in the network. Agents can condition their strategy on a simple statistic: the proportion of biased agents in different parts of the social network. But if there exists one violating edge and communication is limited, agents need to update their beliefs based on the strategies of other players. This updating requires an understanding of the messages blocked by other players and a computation of the proportion of biased players in subtrees. Since the strategy of an agent depends on the identity of the neighbor who sent him the message, the description of a strategy profile requires a complete list of all the neighbors of all the players, i.e. a description of the entire social network. If some agents block communication and others do not, an agent must be able to compute the proportion of biased agents in subtrees, which requires some knowledge of the distribution of biased and unbiased agents in the network. Hence the computation of the posterior probability of the state when receiving message 1 in an MCE requires perfect knowledge of the network and precise information of the distribution of biased and unbiased agents in the network.

B General Networks

This section considers which results extends to general network structures. The possibility of cycles introduces a host of modeling issues since messages could be received from multiple neighbors, passed on to different neighbors with different paths to others in the network, and the time it could take for a message to travel along shorter or longer paths could be important to inference. We identify three key simplifying assumptions that render the analysis of the general network similar to the analysis of a tree. First, if agents transmit message, they send the message to all their neighbors (except the one from whom they have receives a message). Second, the time it takes to message to travel along a path is proportional to the length of the path. Third, agents only decide whether or not to transmit the message the first time they receive it; that is, while a message could reach an agent along several paths and an agent could therefore receive several (necessarily identical) messages, they ignore all messages after receiving one message.

Under these three assumptions, we can generalize Theorem 1 on Full Communication Equilibrium. As in a tree, an agent i conditions her beliefs and her behavior only on the identity j of the agent from whom she received the first message. This result requires to redefine the set of agents who could have been the source of the message. Now, the message can be first received simultaneously from several neighbors, which gives information as to the source of the message. So we need to distinguish from which subset of neighbors the message has been first received.

Consider first the case where an agent i receives the message from a single neighbor. For any i and any j in N_i, we consider the set, denoted by $T_i(j)$, of nodes k such that the shortest paths between k and i all go through j – if the source was k, the message would necessarily reach i through j. To define $T_i(j)$, consider for any pair of agents (i, k), the set $P(i, k)$ of i's neighbors on a shortest path between i and k in G. Formally

$$\text{For any } i, \text{ any } j \in N_i, : T_i(j) = \{k | P(i, k) = \{j\}\}.$$

This set plays the same role as the set $S_i(j)$ in the benchmark tree model.

Consider now the case where i first receives the message simultaneously from all the neighbors in the subset J. In this case, the source k is in the set denoted by $T_i(J)$ such that the shortest paths between k and i go through any of the neighbors $j \in J$ – if the source was k, then the message would necessarily reach i simultaneously through all the neighbors $j \in J$. Formally,

$$\text{For any } i, \text{ any } J \subset N_i \ : T_i(J) = \{k | P(i, k) = J\}.$$

Of course, $T_i(J) = T_i(j)$ if J is the singleton $\{j\}$. By construction, the sets $T_i(J)$ for different J do not intersect. Furthermore, as the graph is connected, their union cover all nodes except i: the family $\{T_i(J), J \subset N_i\}$ form a partition of $G \setminus i$. Figure 6 illustrates this situation. The sets $T_0(1)$, $T_0(2)$ and $T_0(3)$ are the sets of agents who could have originated the rumor when agent 0 hears it first from agents 1, 2 and 3. If agent 0 receives simultaneously the rumor from agents 2 and 3, he will infer that the rumor originated from some agent in the set $T_0(\{2,3\})$.

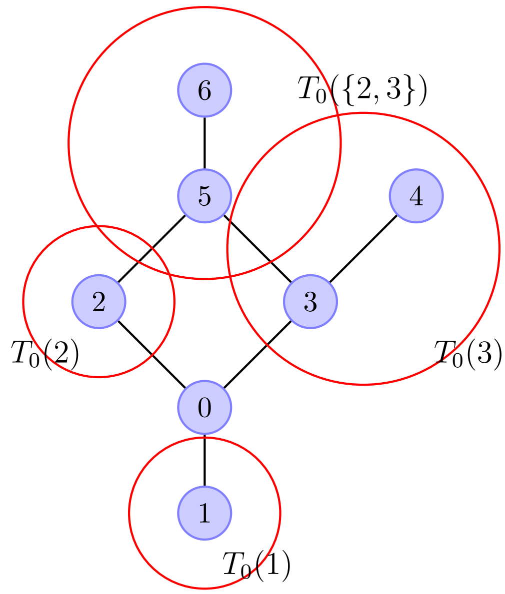


Figure 6: Receipt of Messages and Message Sources in General Networks

We can now state:

Theorem 5 *In the network model with general networks, with multi-cast messages and where agents consider messages that travel on shortest paths, the full communication equilibrium exists if and only if for each unbiased agent i, each of his neighbors j and each subset of neighbors $J, |J| \geq 2$ such that $T_i(J) \neq \emptyset$,*

$$\frac{\pi}{1-\pi} \geq b_{T_i(j)} \text{ and } \frac{\pi}{1-\pi} \geq b_{T_i(J)}.$$

However, we observe that, even under the three assumptions on agents' behavior, Maximal Communication Equilibria might not be defined when the network contains cycles. Consider the following counterexample.

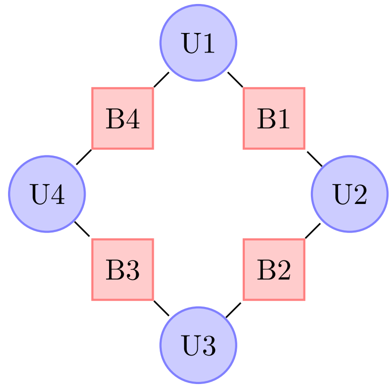


Figure 7: Circle Network

An equilibrium with maximal communication (in the sense defined in Theorem 2) is such that for any directed link (i, j), if a message is transmitted along link (i, j) in any equilibrium, it is also transmitted along link (i, j) in the MCE. We will show that an MCE does not exist in the circle. Let π satisfy $1/2 < \frac{\pi}{1-\pi} < 2/3$. With these values, when an unbiased agent receives $m = 1$ and the proportion of biased agents who could be the source is 2/3, the posterior belief of state 1 is smaller than 1/2: it is optimal for the unbiased agent to vote for 0 and not to transmit the message 1 if it increases the number of votes in favor of the less likely outcome 1. When the proportion instead is 1/2, the posterior belief is higher than 1/2: it is optimal for the unbiased agent to vote for 1 and to transmit the message if this increases the number of votes in favor of the more likely outcome 1.

We first show that there is no equilibrium with full communication. Suppose by contradiction that full communication arises. When U_1 receives the message $m = 1$ from his left neighbor B_1 but not from the right one, the message has been created by B_1, U_4 or B_3, i.e., $T_{U_1}(B_1) = \{B_1, U_2, B_2\}$. Hence the proportion of biased creators is 2/3: it is not optimal for U_1 to transmit the message and hence the condition for existence of a FCE is violated).

Next we show that any directed link (i, j) can be used to transmit information in an equilibrium. To do this, we use the symmetry of the circle to construct a series of equilibria. We first consider agent $U1$ and build a (non-symmetric) equilibrium by assuming that U_1 is the only unbiased agent who creates messages. The strategies are as follows: U_1 creates truthful messages and transmits no one. U_2: U_2 creates no message. When U_2 receives $m = 1$ from B_1, he votes for 1 and transmits it to B_2. Otherwise, U_2 votes for 0 and transmits no message if he receives one. U_3: U_3 creates no message. When U_3 receives $m = 1$ simultaneously from both B_2 and B_3, she votes for 1, and otherwise she votes for 0. U_4: the strategy is defined similarly as for U_2. In particular he votes for 1 only when he receives $m = 1$ from B_4, in which case he transmits it to B_3. We check that these strategies are optimal: U_1: U_1 can receive messages created by biased agents, so does not benefit from transmitting it; when he creates $m = 1$, the message induces everyone to vote for 1, the true state. U_2: When he receives $m = 1$ from B_1, the message has been created by B_1 or U_1: the proportion of biased creators is 1/2, hence it is optimal for U_2 to vote for 1 and to transmit the message to B_2 (strictly optimal, because U_3 will get perfect information whatever situation, see below). When U_2 receives a message from B_2, it has been created by biased B_3, so U_2 keeps the prior π. Finally U_2 has no incentive to create a message since it is either blocked by his neighbor (if it is $m = 0$) or would not been believed by U_1 nor U_3. U_3: When U_3 receives $m = 1$ both from B_2 and B_3, she know that U_1 has created the message, hence is sure that the signal is 1. When she receives $m = 1$ from B_2 only, it has been created by biased B_2, so she does not believe in it. Finally U_3 has no incentive to create a message by the same argument used for U_2. U_4: By symmetry, his strategy is optimal applying the same argument as for U_2. In this equilibrium, information is transmitted along the links $(U1, B1), (U1, B4), (U2, B3), (U4, B3)$. Choosing a different unbiased agent as the anchor of the equilibrium, any directed link in the circle can be used to transmit messages, and as the FCE does not exit, no maximal communication equilibrium exists.

We next discuss relaxing the assumption of multi-cast transmission. The key condition for the existence of a full communication equilibrium is now sufficient, but not necessary. By targeting communication, biased agents can channel their messages through a subgraph where the proportion of biased agents is small. First, when there is no violating edge, a FCE always exists:

Proposition 5 *In a model of targeted transmission, if for each unbiased agent i, each of his neighbors j and each subset of neighbors $J, |J| \geq 2$ such that $T_i(J) \neq \emptyset$,*

$$\frac{\pi}{1-\pi} \geq b_{T_i(j)} \text{ and } \frac{\pi}{1-\pi} \geq b_{T_i(J)}.$$

then an FCE exists.

However, when agents can choose to target neighbors rather than send their messages multicast, full communication can arise even when there are violating edges. This possibility is illustrated by the following example where an agent targets his message through a particular path in the network:

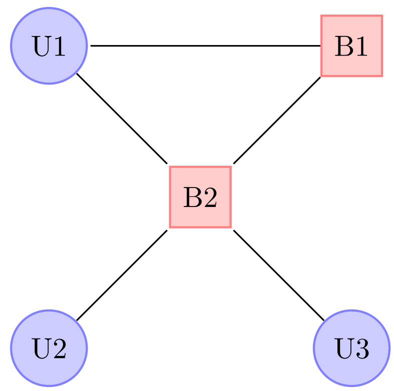


Figure 8: Targeted communication in cycles

Suppose that $\pi = \frac{1}{3}$ so that we need $b \leq \frac{1}{2}$ for the posterior belief to be larger than $\frac{1}{2}$. Notice that the $T_{U1}(B1) = \{B1\}, T_{U1}(B2) = \{B2, U2, U3\}$. Hence the edge $B1 \to U1$ is violating. In a multicast model, the message sent by $B1$ would not be believed by $U1$. Suppose however that $B1$ can choose to send his message only along the link $B1 \to B2$. Then the message 1 will be transmitted by B2 and U1 would update his beliefs to $T'_{U1}(B2) = \{B1, B2, U2, U3\}$ and believe any message received through $B2$. Hence, $B1$ has an incentive to target communication and send his message 1 only along the path through $B2$.

VII Conclusion

This paper studies why agents purposefully and rationally spread rumors. Biased agents desire a particular outcome to occur; unbiased agents want the outcome to match the true state of nature. Both types of agents create and transmit messages in order to influence the common outcome. The analysis compares two benchmark models of communication. In a public broadcast setting, agents send a message about the state of nature to all other agents. In a network setting, agents send a message to their friends and neighbors, who can then transmit the message to their friends and neighbors.

When agents are less sure about the true state, information is valuable and in both structures full communication is possible. When agents have more confidence in the true state, however, they are less willing to believe messages that could come from a biased source. In the public broadcast model, no communication becomes the only outcome, since agents cannot discern the veracity of messages from different people. In a network, however, agents can discriminate among messages received from different neighbors. They can choose to not transmit messages that originate in parts of the network that are heavily populated with biased agents. We construct an algorithm to identify, in any network, the paths along which messages can flow in an equilibrium.

A feature that emerges in the network maximal equilibria is one-way flow of information. A message can flow from a part of the network to another, but not in the opposite direction, since the proportion of biased agents on either side of the link determine the credibility of the message. Thus, studies of the spread of information and rumors in networks should consider that links are not always used and not always used in both directions.

We also find that in order to influence outcomes, biased agents might prefer to limit their numbers and to spread themselves within the population. If there are too many biased agents, unbiased agents do not believe any messages and do not send them along to their neighbors. With fewer biased agents located sporadically in the network, unbiased agents transmit all messages, since the likelihood of the messages being false is sufficiently small. Hence biased agents are better off since, if they have the opportunity, they can create a false message which then spreads.

Future research would further consider situations where agents have incomplete information on the network and information is received through different channels. For example, following Galeotti et al. (2010), we could assume that agents only know their neighbors and their types, the degree distribution in the network, and some measure of homophily in the network. Agents would then choose their strategies as a function of their degree and the status of the neighbor who transmits the message. Allowing agents to condition their behavior on the number of times they receive messages in general networks would also open new possibilities. By observing the sequence in which messages are received and computing the delay in the reception of messages from different neighbors, agents could improve their inference on the source of the rumor. This possibility would give rise to a model where agents use calendar time and choose when to transmit messages to their neighbors. Introducing such elements in a model of message creation and rumor spreading is a difficult but important direction for future work.

Appendix

Proof of Lemma 1. Let us consider unbiased agent i at the end of the transmission stage and $\mathcal{I}$ her information: $\mathcal{I}$ includes others' strategies as well as the message/signal i may have received and from whom, or the absence of message. Others' votes depend on their own information but not on i's vote. Let us denote their number by $\tilde{z}$ and by $Proba(\theta, z|\mathcal{I})$ their joint probability with the state as is perceived by i. Thus i's expected utility from voting for a, $a = 0, 1$, is

$$E[w(\tilde{x}, \tilde{\theta})|a, \mathcal{I}] = \sum_{z,\theta}[w(1,\theta)f(z+a) + w(0,\theta)(1 - f(z+a))]Proba(\theta, z|\mathcal{I}) \tag{5}$$

The incentive to vote for 1 instead of 0 are thus determined by the sign of

$$\sum_{z,\theta}[w(1,\theta) - w(0,\theta)](f(z+1) - f(z))Proba(\theta, z|\mathcal{I})$$

For $f(z) = z/n$, this expression writes

$$\frac{1}{n}\sum_{z,\theta}[w(1,\theta) - w(0,\theta)]Proba(\theta, z|\mathcal{I}).$$

Since $[w(1,\theta) - w(0,\theta)]$ is equal to 1 for $\theta = 1$ and to -1 for $\theta = 0$ the above expression is equal to

$$\frac{1}{n}[2\sum_{z} Proba(\theta = 1, z|\mathcal{I}) - 1]$$

As $\sum_z Proba(\theta = 1, z|\mathcal{I}) = Proba(\theta = 1|\mathcal{I})$, we finally obtain that the incentives to vote for 1 or 0 only depends on the sign of $2Proba(\theta = 1|\mathcal{I}) - 1$, i.e. on $2\rho_i - 1$ where ρ_i is the posterior on the state being 1 at the time of the vote. ■

Proof of Proposition 2. The proof is immediate. ■

Proof of Theorem 1.
Sufficiency. Given $b_{S_i(j)} \leq \frac{\pi}{1-\pi}$, we consider possible deviations from the specified strategies.
Biased agents. For message strategies, a biased agent's expected payoff cannot increase by adopting the strategy $M(s) = 0$ or $M(s) = \varnothing$. Given the strategies of other agents and beliefs consistent with these strategies, creating message $m = 0$ rather than $m = 1$ would decrease the probability that unbiased agents receive a message $m = 1$ and thus decrease the number of agents that vote for outcome $x = 1$. For transmission strategies, the same argument applies.
Unbiased agents. For message strategies, an unbiased agent who receives a signal s believes with probability 1 that the true state is s. Given other agents' strategies, creating any message other than s then lowers the expected number of agents who will vote for outcome $x = s$. For transmission strategies, all unbiased agents believe with greater than probability 1/2 that the true state is 1(0) upon receiving a message $m = 1 (m = 0)$. Since any unbiased agent's expected utility is increasing in the number of agents who share his beliefs, an agent cannot benefit by not transmitting a message or blocking a message, given other agents' transmission strategies and beliefs.

Necessity. Suppose $b_{S_i(j)} > \frac{\pi}{1-\pi}$ for some unbiased agent i and one of his neighbors j. This agent would have an incentive to deviate from the specified transmission strategy and adopt the strategy to block a message $m = 1$ received from neighbor j. In this case, agent i holds that state 1 is less likely than state 0, despite having received the message $m = 1$. Given other agents' strategies and beliefs, agent i can improve his expected payoffs by not transmitting the message.■

Lemma 2 *A If $V^t \neq \varnothing$, then $V^{t+1} \subset V^t$. Hence, there exists a step T such that $V^T = \varnothing$ and $V^{T-1} \neq \varnothing$.*

Proof of Lemma 2.

Consider V^t and let (j, i) be the level 1 directed edge in V^t that is eliminated in this step. Consider a directed edge $(k, l) \in G^{t+1}$ and suppose that (k, l) is not in V^t. We show that it is not in V^{t+1}. As $(k, l) \in G^{t+1}$, (k, l) cannot belong to $G_i(j)$. Now, consider two possibilities: either $(j, i) \in G_l(k)$ or not.

If $(j, i) \notin G_l(k)$ when k receives message $m = 1$ from l, the message cannot have traveled through (j, i): the elimination of (i, j) does not affect $G_l^t(k)$ nor $S_l^t(k)$. Hence,

$$b_{S_l^{t+1}(k)} = b_{S_l^t(k)} \leq \frac{\pi}{1-\pi},$$

so that (k, l) does not belong to V^{t+1}.

If $(j, i) \in G_l(k)$, then $G_i(j) \subset G_l(k)$. Hence, $S_i(j)^t \subset S_l(k)^t$. Following the elimination of (j, i), $S_i(j)^t$ has been withdrawn from $S_l(k)^t$, hence $S_l(k)^{t+1} = S_l(k)^t - S_i(j)^t$ and

$$b_{S_k^{t+1}(l)} = \frac{|B \cap S_l(k)^{t+1}|}{|S_l(k)^{t+1}|} = \frac{b^t_{S_l(k)}|S_l(k)^t| - b^t_{S_i(j)}|S_i(j)^t|}{|S_l(k)^t| - |S_i(j)^t|}. \tag{6}$$

As $b_{S_k^t(l)} \leq \frac{\pi}{1-\pi} < b_{S_i^t(j)}$,

$$\frac{b^t_{S_l(k)}|S_l(k)^t| - b^t_{S_i(j)}|S_i(j)^t|}{|S_l(k)^t| - |S_i(j)^t|} < \frac{b^t_{S_l(k)}|S_l(k)^t|}{|S_l(k)^t|},$$

so that

$$b_{S_k^{t+1}(l)} < b_{S_k^t(l)} \leq \frac{\pi}{1-\pi},$$

(k, l) is not in V^{t+1}, concluding the proof of the lemma.■

Lemma 3 *A The set W of eliminated edges is independent of the order in which directed edges are chosen at each step of the algorithm.*

Proof of Lemma 3.

The proof is by induction on the *initial* levels of edges in the set V, that is their levels in G^0.

All level 1 directed edges $(j, i) \in V$ are always eliminated in the algorithm, irrespective of the order in which edges are chosen: as the subgraph $G_i(j)$ contains no edge in V, the proportion $b^t_{S_i(j)}$ stays constant and any level 1 edge $(j, i) \in V$ remains level 1.

Suppose the induction assumption holds for all directed edges (k, l) of initial levels smaller than ℓ: either (k, l) is eliminated by the algorithm in all possible orders or never. Consider a directed edge (j, i) of level ℓ. Since $G_i(j)$ only contains edges of initial levels smaller than ℓ, the final graph $G_i^T(j)$ obtained when the algorithm stops is independent of the order. As a result, the proportion of biased agents in $G_i^T(j)$ is independent of the order in which edges are chosen, and we can unambiguously determine whether (j,i) is eliminated or not, proving the induction step. ■

Lemma 4 *A For any $(j,i) \in W$, j is biased and i is unbiased.*

Proof of Lemma 4.

Recall that i is unbiased by definition of directed edges in V. As for j, j must be biased for any level 1 edge in V (otherwise there would be a violating edge in $G_i(j)$. The same argument holds for any edge in W because each edge in W is a level 1 edge of V^t in G^t at the step t it is eliminated. ■

Proof of Theorem 2.

Consider first the behavior of a biased agent. A biased agent does not have an incentive to deviate and either create $m = 0$ upon receipt of a signal, transmit $m = 0$, or not transmit a $m = 1$. Given agents' beliefs, any of these action would (weakly) increase the probability that an agent votes for outcome 0 instead of outcome 1.

Consider next unbiased agents. An agent i who receives $s = 0$ or $m = 0$ has the belief that the true state is 0. She then does not have an incentive to deviate and create or transmit $m = 1$, since this action will (weakly) increase the probability that more agents vote for outcome 1. An agent i who receives $s = 1$ knows that the true state is 1. She does not have an incentive to deviate and create $m = 0$ since, given the beliefs, this will (weakly) increase the number of agents who vote for outcome 0. For transmission of $m = 1$, an unbiased agent i who receives $m = 1$ and places sufficiently high probability that the true state is 1, cannot gain by blocking the message. For $(j,i) \in G^*$, then, an agent i who receives $m = 1$ from j cannot gain by blocking the message. If on the other hand, she receives $m = 1$ from a neighbor j where $(j,i) \notin G^*$, she cannot gain by transmitting the message: her beliefs are $\widetilde{\rho}_i(1(j)) < \frac{1}{2}$, and given the strategies of others, more agents will then vote for 1 and lower her expected utility.

We now show that there cannot be an equilibrium where $m = 1$ is transmitted along a directed edge (j,i) not in G^*. (j,i) is in W hence in V. The proof is by induction on the level of (j,i) in the initial graph G^0.

First suppose that (j,i) is a level 1 edge. In the specified strategies, for any edge (k,l) with l unbiased in $G_i(j)$ l transmits $m = 1$ when he receives it from k (this is also the case for l biased by assumption). Consider an alternative equilibrium.

If for any edge (k,l) in $G_i(j)$, with l unbiased, l behaves as in the original equilibrium, then i's posterior $\widetilde{\rho}_i(1(j))$ is the same as in the original equilibrium, hence is lower than 1/2: i must block the message.

Otherwise, there are edges (k,l) in $G_i(j)$, with l unbiased, for which l does not transmit $m = 1$ received from k. Call such an edge deviating and denote D the set of deviating edges. The subgraph $G_i'(j)$ along which $m = 1$ can reach i through j in the alternative equilibrium is made of all the paths to i in $G_i(j)$ that contains no edge in D. We show that the proportion of biased agents in $S_i'(j)$ (with obvious notation) is larger than $\frac{\pi}{1-\pi}$.

Let (k,l) be in D such that $G_l(k)$ contains no edge in D (such a (k,l) surely exists). Since (k,l) is not in V (because (j,i) is of level 1), the proportion of biased agents in $S_l(k)$ is not larger than $\frac{\pi}{1-\pi}$. Since

(i,j) is in V the proportion of biased agents in $S_l(k)$ is larger than $\frac{\pi}{1-\pi}$. Hence the proportion of biased agents in $S_i(j) - S_l(k)$ is strictly larger than $\frac{\pi}{1-\pi}$.[20]

Consider the directed tree $G_i(j) - G_k(l)$. If it contains no element in D, $G'_i(j) = G_i(j) - G_k(l)$. The set of nodes of $G_i(j) - G_k(l)$ is $S_i(j) - S_l(k)$, which has a proportion of biased agents strictly larger than $\frac{\pi}{1-\pi}$: i's posterior $\widetilde{\rho'}_i(1(j))$ is lower than 1/2: i must block the message.

If the directed tree $G_i(j) - G_k(l)$ contains an element in D we can use the previous argument to that tree (i.e. replacing $G_i(j)$ by $G_i(j) - G_k(l)$) and obtain a sub-tree by deleting an element of D. Continuing this way, we obtain a decreasing sequence of subgraphs whose nodes have a proportion of biased agents larger than $\frac{\pi}{1-\pi}$ by deleting at each step a subgraph containing an edge in D. The process stops when all deviating elements have been eliminated and the tree $G'_i(j)$ is reached; this proves that $S'_i(j)$ has surely a proportion of biased agents larger than $\frac{\pi}{1-\pi}$ so that i's posterior $\widetilde{\rho}_i(1(j))$ is lower than 1/2: i must block message 1 from j at the alternative equilibrium.

Next, at the induction step, suppose that all unbiased agents l with $(k,l) \notin G^*$ of level smaller than ℓ block message 1. Consider a level ℓ edge $(i,j) \notin G^*$. Consider the directed subtree $G^*_i(j)$ of $G_i(j)$. $G^*_i(j)$ is the subgraph along which $m = 1$ travels i through j in the original equilibrium and contains violating edges of level smaller than ℓ. Therefore, by the induction assumption, at an alternative equilibrium, $m = 1$ can reach i through j only along a path included in $G^*_i(j)$. We can therefore apply exactly the same argument as above, replacing the tree $G_i(j)$ by $G^*_i(j)$.■

Proof of Proposition 2. The argument is provided in the text.■

Proof of Proposition 3. As the behavior of biased agents and of unbiased agents at the initial stage are fixed, we only need to consider the transmission of unbiased agents at edges in G^*. If the message is 0, unbiased agents surely want to transmit the message because they know it is thruthful. Suppose that the message is 1. We show that the activity rule shows that the message will also be believed and transmitted.

We implement the following coloring of directed edges in G^*. Start by coloring edge (j,i) in green if i is biased and in white if i is unbiased. Each white edge will be colored in blue if i believes the message with probability grater than $\frac{1}{2}$. By the activity rule, this implies that i transmits the message as in the MCE. Uniqueness is proved by coloring all white edges in blue.

Consider all white dangling edge (j,i) with j as a leaf. As (j,i) is in G^*, i believes with probability greater than $\frac{1}{2}$ that the message is truthful (in fact the posterior probability is 1). Color the directed edge (j,i) in blue. At the end of the first step, all dangling edges are green or blue. Furthermore, there are no other blue edges in the graph. This initial step shows that whenever the diameter of $G_i(j)$ is equal to one, all edges are colored in blue or green.

Suppose now that in all graphs $G_i(j)$ with diameter smaller than d, all directed edges are colored in blue or green. Pick a white edge (j,i) such that the diameter of $G_i(j)$ is equal to d. As all paths in strict subgraphs of $G_i(j)$ are either blue or green, agents in $S_i(j)$ behave as in the MCE, and the posterior belief

[20] By a computation similar to (6)

$$b_{S_i(j)-S_l(k)} = \frac{b_{S_i(j)}|S_i(j)| - b_{S_l(k)}|S_l(k)|}{|S_i(j)| - |S_l(k)|}. \tag{7}$$

As $b_{S_l(k)} \leq \frac{\pi}{1-\pi}$ and $b_{S_i(j)} > \frac{\pi}{1-\pi}$, we obtain $b_{S_i(j)-S_l(k)} > \frac{\pi}{1-\pi}$.

of agent i receiving message 1 from j is the same as in the MCE. As this posterior is larger than $\frac{1}{2}$, by the activity rule he transmits the message. Hence, the edge (j, i) is colored in blue.■

Proof of Theorem 3. We provide the proof in the case where a biased agent always creates message $m = 1$ when he receives the signal. The proof in the general case is more involved, and is available on request.

We compare the expected utility, or, its opposite, the expected loss of an unbiased agent in the MCE and in an alternative equilibrium, denoted $'$.

A strategy profile determines the number of votes $\tilde{z}$ in each state given who receives the signal or whether no signal is sent. As each (ex ante) individual loss depends only on these votes $\tilde{z}$, all unbiased agents derive the same ex ante loss (even if they do not always vote identically as they may not have the same information). For $f(z) = z/n$, this common expected loss is directly related to the total number of votes not matching the state; up to the factor $\frac{1}{n}$ it is equal to

$$\pi \times [\text{number of votes for } 0 \ |\theta = 1] + (1 - \pi) \times [\text{number of votes for } 1 \ |\theta = 0]$$

Since biased agents always vote for 1, this expression is equal to the sum of the constant $(1 - \pi)b_N$ and

$$\pi \times [\text{number of U-votes for } 0 \ |\theta = 1] + (1 - \pi) \times [\text{number of U-votes for } 1 \ |\theta = 0].$$

which writes by disaggregating over all U-agents

$$\sum_{i \in U} \pi[\text{number of } i\text{-votes for } 0 \ |\theta = 1] + (1 - \pi)[\text{number of } i\text{-votes for } 1 \ |\theta = 0].$$

i's term inside the square brackets is the probability she casts a vote that does not match the state. In what follows, we show that this probability is minimized at the MCE relative to other equilibria. This will prove the theorem.

To simplify the presentation we take $p = 1$. Consider an unbiased agent i. Given a strategy profile, i's vote is a function of who receives the signal and the state, i.e. the value of the signal (this does not assume that i has this information). Let $v(j, \theta) \in \{0, 1\}$ denote i's vote when the signal lands on agent j in N and the state is θ, i.e. j has received signal θ, where we omit index i for simplification. The probability that i's vote does not match the state is

$$\mathcal{L}^i = \frac{1}{n} \sum_{j \in N \setminus i} [\pi \mathbf{1}_{v(j,1)=0} + (1 - \pi)\mathbf{1}_{v(j,0)=1}].$$

where $\mathbf{1}$ is the indicator function.

We will need the following lemma, which follows from the construction and properties of the MCE.

Lemma 5 *There is a collection of disjoint sets $S_k(\ell)$ where (ℓ, k) is in W such that if the signal lands on an agent in one of these sets, i receives no message at any equilibrium. Let us denote by N_{-i} the union of these sets.*

Proof of Lemma 5. Let (ℓ, k) be a directed edge in W where k is between ℓ and i. We know from

Theorem 2 that k does not transmit $m = 1$ from ℓ neither at the MCE nor at any equilibrium. Furthermore ℓ is biased and k is unbiased, so k never receives $m = 0$ from ℓ. We thus obtain that no message goes through the directed edge (ℓ, k) in W. This proves that i receives no message at any equilibrium when the signal lands on an agent in one of these sets. As the sets are either disjoint or included into one another, we may pick up the maximal sets and obtain a partition of N_{-i}. (Each maximal set is such that the path from ℓ to i contains no edge in W.) ■

In the MCE, if unbiased agents receive message m they vote for m ; if they receive message $\emptyset$ they vote for 0. Hence, there are two sources of incorrect votes:

1. Agent i does not receive any message and $\theta = 1$
2. Agent i receives message $m = 1$, the signal is received by $\ell \in B$ and $\theta = 0$.

Consider another equilibrium denoted by $'$. We show that if i votes for the correct outcome when i does not in a MCE, i must vote for the wrong outcome in a sufficiently large number of situations where he is correct at the MCE so that the incurred loss overweighs the benefit. We deal with each source of incorrect votes in turn.

1. Agent i does not receive any message and $\theta = 1$.

At the MCE, everyone creates $m = 1$ upon the receipt of signal $\theta = 1$. Hence the signal has been received by some j who has sent message $m = 1$ but the message has not reached i.[21] Thus there must exist a directed edge (ℓ, k) in W on the path from j to i: j belongs to N_{-i}.

At the $'$ equilibrium, no message can go through (ℓ, k) (lemma 5), and i votes for the same outcome whenever $j \in S_k(\ell)$ (even if i does not know that $j \in S_k(\ell)$). In the MCE, i votes for 0. In the $'$ equilibrium, i might vote for 1. As (ℓ, k) is in W, $b_{S_k(\ell)} > \frac{\pi}{1-\pi}$, so that voting for 0 matches the state more often than voting for 1 when $j \in S_k(\ell)$.

As the argument works for any set $S_k(\ell)$ in the partition of N_{-i}, the probability that i's vote matches the state when the signal is received by $j \in N_{-i}$ is at least as high in the MCE as in any other equilibrium.

2. The signal is received by ℓ in B and $v(\ell, 0) = 1$ in the MCE equilibrium.
ℓ is surely in $N_i = N - N_{-i}$. As biased ℓ creates message $m = 1$ when he receives the signal 0, either (a) $m = 1$ does not reach i and i votes for 0 when $m = \varnothing$ or (b) $m = 1$ reaches i through an agent j; she votes for 0 because the posterior is less than 1/2.

Assume that in the alternative $'$ equilibrium, $v'(\ell, 0) = 0$ so that the expected loss when ℓ receives the signal is less at this $'$ equilibrium than at the MCE. Consider the edge (ℓ, k) on the path from ℓ to i ($k = i$ is possible). Let $S^*_k(\ell)$ denote the set of agents from which i can receive a message transiting through (ℓ, k) in the MCE.[22] In the $'$ equilibrium, i possibly receives a message transiting through (ℓ, k) from these agents only. We show that the probability that i's vote matches the state when the signal is received by $j \in S^*_k(\ell)$ is at least as high in the MCE than in any other equilibrium.

Assume first that k is unbiased. As $m = 0$ does not travel through ℓ, i can receive only $m = 1$ or $m = \varnothing$. If i receives $m = 1$ it is through j (case (b) above) and this triggers vote 0. If $m = \varnothing$, there are two possible cases :

[21] Or the signal has not been received, which is impossible for $p = 1$.

[22] N_i is the component to which i belongs in the graph where all the elements W have been dropped, i.e. it is the component of the graph Γ obtained at the end of the algorithm. $S^*_k(\ell)$ is the set corresponding to (ℓ, k) in Γ.

1- i votes for 0 when $\varnothing$. Then, whatever the signal received in $S_k^*(\ell)$, i votes for 0 at the $'$ equilibrium. At the MCE, k is better off following the messages from strategy than voting constantly for 0.

2- i votes for 1 when $\varnothing$. Then, i votes for the wrong outcome when the signal 0 is received by an unbiased agent in $S_k^*(\ell)$. Furthermore when a biased agent receives the signal, the vote is constant and the minimal probability of a wrong outcome is π. Hence, denoting $b = b_{S_k^*(\ell)}$, the probability of an incorrect vote in the $'$ equilibrium when the signal lands on $S_k^*(\ell)$ is at least $(1-b)(1-\pi)+b\pi$. This has to be compared with $b(1-\pi)$ which is the expected loss at the MCE. By construction of the MCE, we have $b < \pi/(1-\pi)$; it is easy to check that this implies $(1-b)(1-\pi)+b\pi > b(1-\pi)$.[23] Hence, the probability that i's vote matches the state when the signal is received by $j \in S_k^*(\ell)$ is at least as large in the MCE than in any other equilibrium.

Assume now k to be biased. Consider the first unbiased agent k' on the directed path from ℓ to i ($k' = i$ is possible): there is a biased agent ℓ' on this path linked to k' and all agents between ℓ and ℓ' are biased. Thus when ℓ sends 1, all these biased agents transmit it. It implies that when ℓ' sends $m = 1$ (either transmits or creates), agent i is in the same situation as when ℓ sends $m = 1$: either (a) message 1 does not reach i or (b) $m = 1$ reaches i through j. As i cannot distinguish whether ℓ or ℓ' have sent the message, he votes 0. We thus have a pair (k', ℓ') where k' is unbiased, ℓ' is biased and i votes for 0 when ℓ' receives the signal and we can apply the previous result.

To conclude, in each situation in which agent j receives the signal and i's vote does not match the state in the MCE, there is a set of agents that contains j such that when the signal lands on this set, the expected number of i's incorrect votes in the $'$ equilibrium is at least as high as in the MCE, and furthermore the sets are disjoint.■

Proof of Proposition 4. Let σ be the initial assignment and σ' obtained by making agent j's status switch from unbiased to biased. As the set of biased agents has increased, the set of violating edges eliminated under σ', W' is a superset of the set of violating edges eliminated under σ, W and $G^{*'} \subseteq G^*$. For an unbiased agent i, we decompose the difference in elected utility in the MCE under σ' and under σ into three terms:

$$EU_i' - EU_i = (EU_i' - EU_i^2) + (EU_i^2 - EU_i^i) + (EU_i^1 - EU_i),$$

where

- EU_i^1 is the expected utility of agent i when all agents' strategies are the same as in the MCE under σ except that agent j always votes for 1
- EU_i^2 is the expected utility of agent i when all agents strategies are the same as in the MCE under σ, except that agent j always votes for 1 and always creates message $m = 1$.

First consider the difference $EU_i^1 - EU_i$. The only difference in expected utilities arises when agents j and i vote for zero under the original assignment. Consider any realization of an extended state where this happens. As the strategies of the players at the message creation and transmission stage have not

[23] $(1-b)(1-\pi)+b\pi > b(1-\pi)$ is equivalent to $b \leq (1-\pi)/(2-3\pi)$. Now $\pi/(1-\pi) \leq (1-\pi)/(2-3\pi)$ (as this is equivalent to $(1-2\pi)^2 \geq 0$, hence $b \leq \pi/(1-\pi)$, implies $b \leq (1-\pi)/(2-3\pi)$.

changed, posterior beliefs (and the voting behavior) of all players but j have not changed. Let z_1 be the number of unbiased players voting for 1 and b the number of biased players. The expected utility of player i in the original assignment is

$$-\rho_i(1 - f(z_1 + b)) - (1 - \rho_i)f(z_1 + b),$$

and in the new assignment

$$-\rho_i(1 - f(z_1 + b + 1)) - (1 - \rho_i)f(z_1 + b + 1),$$

the difference in expected utility between the old and new assignment is thus

$$(2\rho_i - 1)[f(z_1 + b + 1) - f(z_1 + b) < 0,$$

where the last inequality stems from the fact that $\rho_i < \frac{1}{2}$. Taking expectations over all extended states where agents i and j vote for 0 in the MCE for the assignment σ, we obtain $EU_i^1 - EU_i < 0$.

Next, to compute $EU_i^2 - EU_i^1$, let z_1 be the number of unbiased agents who vote for 1 when j receives the signal and creates message $m = 1$, and z_0 the number of unbiased agents who vote for 0 when j receives the signal and creates message $m = 0$. Because in a MCE, an unbiased agent votes for 0 when she receives no signal, z_0 is the total number of unbiased agents in σ'. Let b denote the number of biased agents in σ'. By construction, the only difference in expected utilities EU_i^1 and EU_i^2 arises when agent j receives the signal $s = 0$ and transmits $m = 1$. We thus compute:

$$EU_i^2 - EU_i^1 = -\frac{p}{n}(1 - \pi)[f(z_1 + b) - f(n - z_0)] = -\frac{p}{n}(1 - \pi)[f(z_1 + b) - f(b)] \leq 0. \tag{8}$$

Finally, to compute the difference $EU_i' - EU_i^2$ notice that the only difference between the two expected utilities stems from the transmission behavior of unbiased agents under the new assignment σ'. As $G^{*'} \subseteq G^*$, messages $m = 1$ may be blocked more often in the MCE under σ' than under σ. Because $G^{*'} \subseteq G^*$, the extended state sets must satisfy:

$$\begin{aligned} E_i'^0 &\subseteq E_i^0, \\ E_i'^1 &\subseteq E_i^1, \end{aligned}$$

Following the same steps as in the proof of Theorem 3, we conclude that $EU_i' - EU_i^2 \leq 0$. This shows that $EU_i' - EU_i < 0$, completing the proof of the Proposition. ■

Optimal placement of agents on a line. We first show that, in a maximal communication equilibrium, no unbiased agent can believe message $m = 1$ received from more than k^* biased agents. Suppose by contradiction that there exists an unbiased agent and $k > k^*$ biased agents such that, whenever message $m = 1$ is originated by one of the k biased agents, the unbiased agent switches his prior to $\rho_i > \frac{1}{2}$. First note that, as $k \geq k^* + 1$,

$$k \geq \frac{n\pi}{1 - \pi} + 2\pi - 1 + 1 = \frac{n\pi}{1 - \pi} + 2\pi.$$

Fix the unbiased agent i and consider two sets $S_i(j)$ and $S_i(j')$. Let $0 \leq l \leq k$ be the number of biased agents in $S_i(j)$ and $k-l$ the number of biased agents in $S_i(j')$. By Theorem 1,

$$\begin{aligned} l(1-\pi) &\leq |S_i(j)|\pi, \\ (k-l)(1-\pi) &\leq |S_i(j')|\pi. \end{aligned}$$

Summing up,

$$k \leq \frac{(|S_i(j)| + |S_i(j')|)\pi}{1-\pi} < \frac{n\pi}{1-\pi},$$

yielding a contradiction.

We now show that, in the optimal location described in the Proposition, all unbiased agents believe message $m = 1$ received from $\hat{k} = \min\{k, k^*\}$ biased agents. Pick an unbiased agent i and consider the set $S_i(j)$. By construction, if the set $S_i(j)$ contains l biased agents,

$$|S_i(j)| \geq l + l\frac{n-\hat{k}}{\hat{k}+1}$$

so that the fraction of biased agents in $S_i(j)$ satisfies:

$$b_{S_i(j)} = \frac{l}{|S_i(j)|} \leq \frac{\hat{k}+1}{n+1} \leq \frac{k^*+1}{n+1} \leq \frac{\pi}{1-\pi}.$$

By Theorem 1, this implies that a full communication equilibrium among $\hat{k}$ biased agents and the unbiased agents exists. ■

Proof of Theorem 5: The proof is identical to the proof of Theorem 1.■

Proof of Proposition 5: The proof is immediate.■

Proof of Theorem 4: The proof is identical to the proof of Theorem 1.■

REFERENCES

Acemoglu, Daron, Munther Dahleh, Ilan Lobel, and Asuman Ozdaglar. 2011. "Bayesian Learning in Social Networks," *Review of Economic Studies*, 78, pp. 1201-1236.

Bala, Venkatesh and Sanjeev Goyal, 1998. "Learning from Neighbors," *The Review of Economic Studies* 65 (3), pp. 595-621.

Banerjee, Abhijit. 1992. "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107(3), pp. 797-817.

Banerjee, Abhijit. 1993. "The Economics of Rumours," *Review of Economic Studies* 60, pp. 309-327.

Barstow, David and Suhasini Raj. 2015. "Mob Attack, Fueled by Rumors of Cow Slaughter has political overtones in India." *The New York Times*. October 4.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, pp. 992-1026.

Bouie, Jamelle. 2015. "Anti-Science Views are a Bipartisan Problem," *Slate* , February 4. http://www.slate.com/articles/news and politics/politics/2015/02/conservatives and liberals hold anti science views anti vaxxers are a bipartisan.html

Calvó-Armengol, Joan de Martí and Andrea Prat, "Communication and Influence" , *Theoretical Economics*, 10, 649-690.

Castellano, Claudio, Santo Fortunato, and Vittorio Loreto. 2009. "Statistical Physics of Socal Dynamics," *Review of Modern Physics* 81(2), pp. 591-646.

Centola, Damon and Michael Macy. 2007. "Complex Contagion and the Weakness of Long Ties in Social Networks," *American Journal of Sociology* 113(3), pp. 702-34.

Chatterjee, Kalyan and Bhaskar Dutta. 2016 "Credibility and Strategic Learning in Networks,", *International Economic Review*, 57(3), pp. 759-786.

Conley, T. and C. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100, pp. 35-69.

Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission". *Econometrica* 50 (6), pp. 1431-1451.

DeMarzo, Peter M. Dimitri Vayanos, and Jeffrey Zweibel. 2003. "Persuasion bias, social influence, and uni-dimensional opinions," *Quarterly Journal of Economics* 113(3), pp. 909-968.

DeGroot, Morris H. 1974. "Reaching a Consensus," *Journal of the American Statistical Association*, 69 (345) , pp. 118-121.

Farley, Robert. 2015. "Was Hillary Clinton the Original 'Birther'." *Factcheck.org* July 2. http://www.factcheck.org/2015/07/was-hillary-clinton-the-original-birther/

Foster, Andrew and Mark Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103, pp. 1176-1209.

Gale, Douglas and Shachar Kariv. 2003. "Bayesian Learning in Social Networks," *Games and Economic Behavior*, 45(2), pp. 329-346.

Galeotti, Andrea, S. Goyal, M.O. Jackson, F. Vega-Redondo, L. Yariv. 2010. "Network Games", *Review of Economic Studies*, 77(1), pp. 218-244.

Galeotti, Andrea, C. Ghiglino and F. Squintani. 2013. "Strategic Information in Networks ", *Journal of Economic Theory*, 148(5), pp. 1751-1769.

Golub, Benjamin and Matthew Jackson. 2010. "Naive Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2, pp. 112-149.

Hagenbach, Jeanne and Frédéric Koessler. 2010. "Strategic Communication in Networks," *Review of Economic Studies* 77(3), pp. 1072-1099.

Jackson, Matthew. 2008. *Social and Economic Networks*. Princeton: Princeton University Press.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), pp. 380–391.

Milgrom, Paul and John Roberts. 1986. "Relying on the Information of Interested Parties." *Rand Journal of Economics,* 17, pp. 18-32.

Niehaus, Paul. 2011. “Filtered Social Learning,” *Journal of Political Economy*, 119(4), pp. 686-720.

Romero, Daneil, Brendan Meeder, and Jon Kleinberg. 2011. "Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter.” *Proc. 20th International World Wide Web Conference*, 2011.

Sunstein, Cass R. 2009. *On Rumors: How Falsehoods Spread.* New York: Farrar, Straus and Giroux.