Discussion Papers in Economics

DIFFUSION OF MULTIPLE INFORMATION

By

Nicole Tabasso
(University of Surrey)

DP 09/14

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

Diffusion of Multiple Information*

Nicole Tabasso†

December 16, 2014

Preliminary Version

Abstract

We model the diffusion of two types of information through a population under the assumption that communication time is limited. When a meeting between individuals occurs, at most one information can be communicated. Preferences over information types divide the population into two groups, and if a choice has to be made about which information to communicate, members of either group will choose their preferred information. We find that crowding out of information does occur, but information is rarely eradicated entirely. Somewhat surprisingly, the parameter values under which a unique information would survive in the population are sufficient for both information to survive. Only if information preferences in the entire population are aligned, i.e., every individual prefers to communicate the same information, does the second information die out. We apply our framework to answer questions on the impact that segregation has on information diffusion and polarization. We find that segregation unambiguously increases polarization and decreases the proportions of informed individuals, and derive the conditions under which agents endogenously choose to segregate.

Keywords: Social Networks, Information Transmission, Multiple States, Segregation

JEL Classification: D83, D85

*Thanks go to Rabah Amir, Irem Bozbay, Andrea Galeotti, Alex Gershkov, Christian Ghiglino, Alberto Vesperoni, as well as seminar participants at the Copenhagen Business School, Surrey School of Economics, APET 2014, EEA 2014 and UECE 2014 for useful comments and discussions.

†University of Surrey - School of Economics, Guildford, GU2 7XH, UK. n.tabasso@surrey.ac.uk

1 Introduction

Social networks are acknowledged as a crucial factor in the diffusion of information in a society, as highlighted in seminal work of Lazarsfeld, Berelson, and Gaudet (1968) and Katz and Lazarsfeld (1970). Given the importance of information in the choices that agents make, or in how they form beliefs about the state of the world, a large literature has studied the diffusion of information on social networks.[1] However, so far the literature has focused almost exclusively on the diffusion of one unique information (or, in general, a unique state).[2] Instead, it seems more plausible that whenever two agents meet, the set of topics they could potentially discuss is larger than the set of topics they talk about. Given that communication time, or the number of mutually interesting topics, is limited, agents *choose* what to talk about. This creates opportunity costs in information transmission, as talking about one topic implies not to talk about another. How this will affect the diffusion process of information is not immediately obvious.

In the present paper, we address this question by introducing two distinct types of information, or topics (A and B), into a standard diffusion process, the *Susceptible-Infected-Susceptible* (SIS) framework.[3] This allows us to answer a number of questions relating to the diffusion of information, in particular, the extent to which two information crowd each other out, how this is related to the network characteristics, and how it depends on the relative interest in the population for either topic.

Our model exhibits three main features. First, as in the standard SIS framework, we consider a population of an infinite number of agents, in which individuals transition between being uninformed (susceptible, S) or informed (infected, I). Agents transition between states either if an agent uninformed of a topic becomes informed about it during a meeting, or if he forgets the topic. While forgetting occurs at a constant, exogenous rate, how likely agents are to become informed about a topic will depend on the proportion of the population informed

[1] These studies include questions of how information is aggregated, the impact of social networks on labor market outcomes, voting, product or technology adoption, and many more. It is not our aim to provide an exhaustive list here. For an excellent overview of the literature, see, e.g., Goyal (2012) or Jackson (2008b) and the references therein.

[2] Notable exceptions are a few papers that discuss the impact of competition in models of influence maximization, e.g., Goyal and Kearns (2012), which we discuss in more detail below. Their diffusion processes and the questions they pose differ significantly from ours.

[3] This framework has been developed in an epidemiological context, see Bailey et al. (1975), or Pastor-Satorras and Vespignani (2001a,b, 2002). Within economics, it has been adapted in recent years to study the diffusion of a harmful state, information, or a behavior, by, e.g., Jackson and Rogers (2007), López-Pintado (2008), Jackson and Yariv (2010), Jackson and López-Pintado (2013), Galeotti and Rogers (2013b), or Galeotti and Rogers (2013a).

of it. In the long run, the proportion of the population in each state will remain constant, and our focus is this steady-state, or *prevalence*, of information. By introducing two types of information, we increase the set of infectious states to $I = \{I_{A\setminus B}, I_{B\setminus A}, I_{AB}\}$, but otherwise the process is standard.

The second main feature, which goes beyond the standard model, is an introduction of opportunity costs into the model, which we achieve by assuming that information A and B have to share an agent's limited communication time. Whenever two agents meet, each can communicate at most one information.

Finally, we relate the choice of which information to communicate to intrinsic preferences of agents. We assume that some agents, group A, prefer topic A, and the complement, group B, prefer topic B. If an agent is in state I_{AB}, he will communicate information A only if he belongs to group A, and information B otherwise.

While the model is highly stylized, it captures a basic trade-off that exists in information diffusion, and its simplicity allows us to derive clean and stark results. Our first set of results (section 2.1) shows that information is extremely resilient. This is in line with casual observation that so many obscure bits of information survive.[4] As long as a strictly positive fraction of the population prefers either information l, both of them exhibit a strictly positive steady-state (in which a positive fraction of the population is informed about l), under the exact same network parameter values that are necessary for a single information to exhibit a positive steady-state. While crowding out does occur, and can be quantitatively significant, it is only ever complete if the entire population strictly prefers one topic over another. In this case, there are no parameter values that would allow the second information to survive in the population.

We analyze the extent of crowding out in section 3. Our results confirm the intuitive assertions that the information preferred by more agents will also have a larger prevalence, which is always bounded above by the prevalence in the one-information case. A less obvious result is that crowding out depends not only on the measure of interest a given information receives, but also on the underlying parameters of the diffusion process. *Relative* crowding out of the two information is unaffected by changes in these parameters, and only depends on relative interest. However, the prevalence of information A relative to the prevalence of B depends on the

[4]It is also in line with the general observation that once something is made public, it seems to be virtually impossible to retract it.

network parameters in such a way that fostering the spread of information overall reduces the relative prevalence of A over B if A is preferred by the majority of the population. While both information benefit from improvements in the information transmission process, the minority information gains relatively more.

These results highlight the importance of interest in the population for the possibility and extent to which policy makers might be able to suppress news items.[5] Conversely, the difficulties in eradicating (mis-)information.

Our final set of results relates information diffusion to homophily, the tendency of individuals to interact relatively more with others that are similar to themselves. We address this question in section 4, in which we compare a society in which agents interact randomly, as in the earlier sections, with one in which agents are perfectly segregated according to their information interests. As a corollary from our earlier results, we find that if groups are segregated, within each group, only the preferred information exhibits a strictly positive prevalence. In addition, in an integrated society, the prevalence of *both* information is larger than under segregation. On the other hand, fewer agents are informed about their preferred topic in an integrated society. We go beyond the simple diffusion of information and explicitly endow agents with utilities from being informed, and being informed of the preferred topic provides a higher utility. We use this framework to derive the conditions under which agents themselves would prefer to segregate. In particular, we find that (i) segregation is more likely to occur if the ratio of the preferred over the non-preferred information is larger, that (ii) the minority group prefers segregation for smaller values of this ratio, and that (iii) segregation is less likely the denser the network.

The rise of the Internet and, more recently, Online Social Networks (OSNs) has arguably made it easier for individuals to segregate. This has sparked a discussion on how such segregation might bias news/information consumption and learning and increase polarization in a population (see, e.g., Baccara and Yariv (2008), Rosenblat and Mobius (2004), or Sunstein (2009)). While empirical results suggest that online news consumption, while biased, does not have a significant impact on polarization (Flaxman, Goel, and Rao (2013), Gentzkow and Shapiro (2010)), our results suggest that even with an unbiased consumption of news, or without any additional

[5]It seem taken for granted that policy makers might be induced to try and "bury" unfavorable news by releasing them simultaneously with other major events. In the UK, one high-profile example of this was the case of spin doctor Jo Moore. She resigned her position after emails of hers were leaked in which she suggested on 11 September 2001 that it was "a very good day to get out anything we want to bury". Our model shows that if this practice works, it is unconnected to media coverage, but purely due to differential interest in the population.

assumptions on biased learning/updating, segregation might well lead to polarization due to the fact that the less preferred information is simply not talked about as much in a segregated group. In this sense, OSNs might indeed increase polarization. On the other hand, online technology might also increase the density of the social network. This in turn will lower the likelihood that the population segregates into interest groups.

Related Literature

Within economics, the literature we are most closely related to is the network literature on diffusion processes that build on the SIS framework, such as Jackson and Rogers (2007), López-Pintado (2008), or Jackson and Yariv (2010), but also Galeotti and Rogers (2013b), and Galeotti and Rogers (2013a).[6] While we share the basic methodology of the SIS framework with this literature, in all of the above papers the focus is the diffusion of a unique state.

Diffusion of competing products or innovations has been analyzed in models of influence maximization, in particular by Bharathi, Kempe, and Salek (2007), Borodin, Filmus, and Oren (2010), Dubey, Garg, and De Meyer (2006) and Goyal and Kearns (2012). Models of influence maximization differ significantly from an SIS diffusion process, both with respect to the modeling characteristics, and the questions that they wish to answer. They are based on threshold models, in which contagion occurs on a fixed network and nodes never recover. The central question in this strand of literature is which nodes a player with a fixed budget would choose to infect to maximize the contagion of his product (in Goyal and Kearns (2012), the focus is on how the efficiency of a "seeding" strategy depends on the precise diffusion process and its interaction with the network structure) rather than questions of prevalence, or crowding out, such as we consider. In all of the above papers, being infected with one product precludes infection with another. Consequently, while this literature also considers competing diffusion processes, our own results are complementary to the findings in this literature.

To the best of our knowledge, ours is the first paper that studies how multiple information diffuse simultaneously on a network under communication constraints. The intrinsic trade-off between information in the present paper is essential to study questions such as the resilience of

[6]This literature itself builds on work on epidemiological models in the natural sciences, such as Bailey et al. (1975), Dodds and Watts (2004) Pastor-Satorras and Vespignani (2001b,a), Pastor-Satorras and Vespignani (2002), or Watts (2002). More broadly, the paper is also related to network diffusion processes such as, e.g., learning, best response dynamics, or explicit adoption decisions. These processes however differ significantly from the SIS model we employ. See, e.g., Jackson (2008b) for an excellent introduction to the literature.

information, or the impact of differential media attention on information prevalence. Similarly, our results on the impact of homophily on information diffusion rests exclusively on the fact that agents choose which information to pass on, which is a very different, and complementary channel to earlier results on the impact of homophily on information transmission, such as in Golub and Jackson (2012) or even the seminal work of Granovetter (1973). Indeed, our results on the impact of segregation on polarization are complementary to the work of Baccara and Yariv (2008), Flaxman, Goel, and Rao (2013), Gentzkow and Shapiro (2010), or Rosenblat and Mobius (2004), who study the impact of biased news/information consumption on polarization. In our model, agents in each group are initially informed to the same amount of both information, and the non-preferred information is not strategically withheld. Polarization arises from the rather intuitive assumption that agents will choose to communicate their preferred information.

We are most closely related to Prakash, Beutel, Rosenfeld, and Faloutsos (2012) and Beutel, Prakash, Rosenfeld, and Faloutsos (2012), who model the diffusion of two competing products on a graph in an SIS framework. They find that the stronger product, which is defined as the product that has a larger infection rate, is the only product that survives if competition is "strong enough". This is in stark contrast to our case, in which for almost all parameter values both information survive. The difference to our results stems from the fact that in Prakash, Beutel, Rosenfeld, and Faloutsos (2012) and Beutel, Prakash, Rosenfeld, and Faloutsos (2012), the harshness of competition is determined by the likelihood that an agent may use more than one product. In particular, Prakash, Beutel, Rosenfeld, and Faloutsos (2012) establishes that if the use of one product perfectly "immunizes" against a second product, the stronger product is the only one surviving. I.e., their model disregards one of the potential states in which an individual can be in our model, that of being aware of *both* information. Their results are a natural analogue to our corner case in which one information is deemed more interesting by the entire population. The extension in Beutel, Prakash, Rosenfeld, and Faloutsos (2012) shows that the result of Prakash, Beutel, Rosenfeld, and Faloutsos (2012) continues to hold if immunization is not complete, but above a certain threshold. Again, in contrast to our model, competition is at the level of the infection, not at the level of the diffusion, which is a crucial distinction. While using at most one of two competing products is likely a natural assumption for products (or in some cases diseases), we believe that for information, the state of possessing multiple information at the same time is a more reasonable assumption.

2 Diffusion Process

2.1 Propagation Mechanism

The model of information transmission in the present paper builds on the SIS (Susceptible-Infected-Susceptible) model of the propagation of a disease or (harmful) state. We consider a population with an infinite number of agents. In this population, two topics A and B, are of interest to agents. Each agent can either be informed or uninformed about each topic. Then, an agent can be in one of four states: He may be uninformed of both A and B (which we call Susceptible, S), or in one of three "infection" states: informed of A but not B ($I_{A\setminus B}$); informed of B but not A ($I_{B\setminus A}$); or informed of both A and B (I_{AB}). The set of possible states is then $\{S, I_{A\setminus B}, I_{B\setminus A}, I_{AB}\}$. At any time t, each agent meets k others. If an agent susceptible to information $l \in \{A, B\}$ gets informed about it at a meeting, or when he forgets information $l \in \{A, B\}$, he transitions between states. We denote by ν the rate at which information is transmitted at a meeting and by δ the rate at which it is forgotten. In line with the previous literature and the epidemiological roots of the model, we refer to ν as the (per contact) *infection rate* and δ as the *recovery rate.*

At each meeting, at most one information can be communicated. This assumption is central and embodies the fact that communication time between individuals is limited; only a finite number of topics can be discussed at any meeting.[7] This constraint is only relevant for individuals in state I_{AB}, who have to choose which information to pass on. We endow agents with preferences over the two information and assume that the preferred information is the one that is communicated, conditional on communication taking place at all. A proportion $\nu_A \in [0, 1]$ of the population prefers A, while the remaining proportion $\nu_B = 1 - \nu_A$ prefers information B.[8] Agents who prefer A are denoted group A and agents who prefer B are members of group

[7]Equivalently, one can think of the constraint being on the number of topics that may be of interest to the parties at a meeting. We call the constraint communication time for convenience and because we believe that it possesses the clearest interpretation within our model.

[8]This preference can be interpreted as an agent finding a particular topic simply more interesting. Our results will not change if we instead assume that ν_l is the probability that a single agent in state I_{AB} passes on information l. This assumption would not allow us to investigate questions of the effect of segregation according to information interests, though.

B. Note that if agents are either in state $I_{A\setminus B}$ or $I_{B\setminus A}$, their information preferences will not matter for the rate at which they pass on information l.

Formally, we define $\rho_{A\setminus B}$, $\rho_{B\setminus A}$ and ρ_{AB} as the proportion of the population in the three infection states, $I_{A\setminus B}$, $I_{B\setminus A}$, and I_{AB}, respectively. By definition, the following relationships hold

$$\begin{aligned}\rho_A &= \rho_{A\setminus B} + \rho_{AB}\\ \rho_B &= \rho_{B\setminus A} + \rho_{AB} \\ \rho &= \rho_{A\setminus B} + \rho_{B\setminus A} + \rho_{AB}.\end{aligned} \tag{1}$$

Denote by θ_l the probability that, conditional on information being communicated, a randomly encountered individual will transmit information l. In the one-information case, θ_l would be every individual aware of information l, but in the present model it is given by

$$\begin{aligned}\theta_A &= \rho_{A\setminus B} + \nu_A \rho_{AB} = \rho_A - \nu_B \rho_{AB},\\ \theta_B &= \rho_{B\setminus A} + \nu_B \rho_{AB} = \rho_B - \nu_A \rho_{AB},\end{aligned} \tag{2}$$

as not everybody aware of an information will also pass it on. This is the essence through which the existence of a second information $-l$ imposes an externality on the diffusion of l.

The rate at which a susceptible individual becomes infected with information l is $k\nu\theta_l$. We assume that the infection rate ν is sufficiently small that these rates approximate the chance that an individual becomes informed of l through his k independent interactions at t. Similarly, we assume that the recovery rate δ is sufficiently small such that δ approximates the probability that an agent forgets a particular topic at time t.[9]We assume that A and B diffuse through the population independently of each other. Knowledge of one does not make knowledge of the other any more or less likely. The information propagation process exhibits a steady-state if the

[9]In essence, this assumption implies that at most one information is forgotten at any t, as the probability that both information are forgotten within a small time frame, $\Delta\delta^2$, goes to zero as Δ goes to zero. This assumption simplifies the analytical derivations considerably, but also ensures that no bias against information is introduced: As at most one information can be transmitted at a meeting, to allow more than one information to be forgotten at any t would introduce an additional negative impact on the survival of information.

following three differential equations are satisfied,

$$\frac{\partial \rho_A}{\partial t} = (1-\rho_A)k\nu\theta_A - \rho_A\delta = 0, \tag{3}$$

$$\frac{\partial \rho_B}{\partial t} = (1-\rho_B)k\nu\theta_B - \rho_B\delta = 0, \tag{4}$$

$$\frac{\partial \rho_{AB}}{\partial t} = (\rho_A-\rho_{AB})k\nu\theta_B + (\rho_B-\rho_{AB})k\nu\theta_A - 2\rho_{AB}\delta = 0, \tag{5}$$

i.e., the proportion of agents who become aware of an information at t equals the proportion of agents who forget it.[10]

2.2 Steady-States and Diffusion Threshold

Define $\lambda = \frac{k\nu}{\delta}$ as the *diffusion rate* of information. The (implicit) steady-states of ρ_A, ρ_B, and ρ_{AB} are

$$\rho_A = \frac{\lambda\theta_A}{1+\lambda\theta_A}, \tag{6}$$

$$\rho_B = \frac{\lambda\theta_B}{1+\lambda\theta_B}, \tag{7}$$

$$\rho_{AB} = \frac{\lambda^2\theta_A\theta_B}{(1+\lambda\theta_A)(1+\lambda\theta_B)} = \rho_A\rho_B, \tag{8}$$

and we denote the values of ρ_A and ρ_B that solve equations (6)-(8) as the *prevalence* of information A and B. Due to the inherent symmetry of the model, in the remainder of the paper we focus, without loss of generality, on the case in which $\nu_A \geq \nu_B$.

Remark 1. For any given diffusion rate $\lambda \geq 0$, there exists a steady-state in which $\rho_l = 0$ for $l \in \{A, B\}$.

The existence of a steady-state in which nobody is informed is trivial. If the initial conditions are such that no agent is informed about a topic, nobody ever will be. Questions of interest instead concern the existence of a steady-state in which $\rho_l > 0$ for either or each $l \in \{A, B\}$,

[10] We assume that δ is the unique rate at which both A and B are forgotten. There are numerous alternative ways to model forgetting, e.g., the preferred information might be forgotten at a lower rate, or being aware of multiple pieces of information increases the rate at which all of them are forgotten. On the other hand, it might also be the complexity of an information that is the determining factor in forgetting, something that is entirely exogenous to the model. The unique value of δ allows us to derive very cleanly the impact that the existence of a second information has on the diffusion process, without additional complications.

and its characteristics. To analyze these, we first define the following concept, which is adapted from López-Pintado (2008).

Definition 1. For each $l \in \{A, B\}$, let λ_l^* be such that the following two conditions are satisfied for all $\lambda > \lambda_l^*$:

1. There exists a positive steady-state for information l, i.e., a steady-state in which a strictly positive fraction of the population is informed about it. For all $\lambda \leq \lambda_l^*$, such a positive steady-state does not exist for information l.

2. The positive steady-state is globally stable. That is, starting from any strictly positive fraction of agents informed about l, the dynamics converge to the positive steady-state. For all $\lambda \leq \lambda_l^*$, the dynamics converge to a steady-state in which no agent is informed about l.

We call λ_l^* the *diffusion threshold* of information l.[11]

Furthermore, we are interested in how the diffusion threshold and the prevalence of either information l compare to the case in which information l was the unique information diffusing on the network. We therefore define the following concepts.

Definition 2. Let λ_d be the diffusion threshold of information in case a unique information diffuses through the network.

Definition 3. Let $\tilde{\rho}$ denote the positive steady-state of an information if it is the unique information that diffuses through the network.

For the present diffusion process, it has been established (see, e.g., López-Pintado (2008) or Jackson (2008b) and the reference therein) that $\lambda_d = 1$ and $\tilde{\rho} = 1 - \lambda^{-1}$. We are now in a position to state our first set of results regarding the existence and stability of positive steady-states for either information $l \in \{A, B\}$.

Proposition 1. *The diffusion threshold λ_l^* depends on the value of ν_l:*

1. *If $\nu_l \in (0, 1)$ for each $l \in \{A, B\}$, then $\lambda_A^* = \lambda_B^* = \lambda_d = 1$.*

[11] López-Pintado (2008) actually defines a *critical threshold* that determines whether a positive steady-state exists, and a *diffusion threshold* which determines the stability properties of this positive steady-state. In the present setting, in which every agent has the same number of meetings, the two thresholds always coincide, allowing us to define the diffusion threshold over both the existence and the stability of a positive steady-state.

2. *If $\nu_l = 0$ for some $l \in \{A, B\}$, there exists no finite value of λ_l^*.*

3. *If $\nu_l = 1$ for some $l \in \{A, B\}$, $\lambda_l^* = \lambda_d = 1$. For $\lambda > 1$, $\rho_l = \tilde{\rho}$.*

Independent of the value of ν_l, there exists at most one positive steady-state.

Proof. See Appendix A. □

Our result that λ_l^* is identical to λ_d for almost all values of ν_l highlights the enormous resilience of information. Any parameter values of infection rate ν, recovery rate δ, and density of the network (the number of meetings k) that are sufficient for a positive prevalence of one information are also sufficient for positive prevalences of two information, as long as a strictly positive proportion of the population prefers either information.

The independence of the diffusion threshold from the number of topics that propagate through the network is striking and deserves some closer attention. In epidemiology, it is a well-established result that an infection exhibits a positive prevalence if its *basic reproduction number* (R), the number of agents to which an infected agent spreads the disease on average, is larger than 1. This in turn explains why for a unique information, the diffusion threshold is $\lambda_d = 1$. As $\lambda = \frac{\nu k}{\delta}$, it is exactly the average number of nodes that "catch" information from an informed node, i.e., $\lambda = R$. In the present model, though, $R_l \leq \lambda$, as not every agent that is given the chance to communicate l will do so. Instead, out of all nodes that are aware of A, a proportion ρ_B are also aware of B, and of those, a proportion ν_B will never communicate A at a meeting. Hence, only a proportion $1 - \nu_B \rho_B$ of all agents aware of A would ever communicate it. I.e., the basic reproduction number for information l is

$$R_l = \lambda(1 - \nu_{-l}\rho_{-l}). \tag{9}$$

Remark 2. For either $l \in \{A, B\}$, the following is true for R_l: (i) If $\nu_l \in (0, 1)$, then $R_l > 1$ if and only if $\lambda > 1$. (ii) If $\nu_l = 1$, then $R_l = \lambda$. (iii) If $\nu_l = 0$, then $R_l = 1$, independent of the value of λ.

The results in Remark 2 are derived in Appendix E.[12] It is true that, while the basic reproduction number R_l is in general different from λ (except for the case of $\nu_l = 1$), it is

[12] The argument makes use of the explicit solutions of ρ_A and ρ_B, which are derived in Appendix B.

nevertheless *larger* than 1 only if $\lambda > 1$. Intuitively, this is because an increase in λ increases both ρ_A and ρ_B: If λ is just slightly above 1, information l has only a very low diffusion rate, but so does information $-l$, which implies a low value of ρ_{-l}, in which case the proportion of agents aware of l and willing to spread it is almost 1. The positive, direct, effect of an increase in λ on l through increasing its diffusion rate is always larger than the indirect, negative, effect it has on l through increasing $\nu_{-l}\rho_{-l}$. The two effects are identical and offset each other if $\nu_{-l} = 1$. In this case, it can be shown[13] that $\rho_{-l} = 1 - \lambda^{-1}$, which leads to $R_l = 1$, independently of λ.

Proposition 1 explains the longevity of diverse rumors, gossips, or beliefs: As long as *some* people care about these topics enough to talk about them, once the seed exists, the information will never die out. Media coverage matters for prevalence only insofar as it needs to provide some initial seed. As the positive steady-states are unique and globally stable, the prevalence of an information is, *per se*, unrelated to the size of any media coverage it receives. In other words, if people are interested in a piece of news, it can never be buried completely. However, in the proceeding section we show that the magnitude of an information's prevalence depends positively on the overall interest that exists in the population, ν_l. While quite intuitive in itself, this result also implies that if the amount of media coverage adequately reflects interest in the population, we will observe that news that have received higher media coverage will also exhibit a larger prevalence.

3 Crowding Out of Information

Independently of ν_l, if $\lambda \leq 1$, no positive steady-state exists, neither in the single information case nor in the case of two information diffusing. Hence, in the remainder of the paper, we focus on the more interesting case of $\lambda > 1$ and on positive prevalences.

We know from Proposition 1 that if $\nu_l = 1$, $\rho_l = 1 - \lambda^{-1}$, identical to the one-information case. This identity is not surprising: As $\nu_l = 1$, information l does not in fact face any competition from $-l$. We define crowding out of an information as follows:

Definition 4. For any $\nu_l \in [0, 1]$, crowding out of information l is given by $\tilde{\rho} - \rho_l$.

[13] see Appendix A

Note that for $\nu_l = 1$, crowding out of information l is zero, while it is complete for information $-l$. To analyze crowding out of information for values of $\nu_l \in (0,1)$, we derive the positive prevalence of A and B. Let

$$C \equiv -(1-\lambda\nu_A\nu_B) + \left[(1-\lambda\nu_A\nu_B)^2 - 4\nu_A\nu_B(1-\lambda)\right]^{\frac{1}{2}},$$

then we can express the positive steady-states of A and B as:

$$\theta_A = \tfrac{C}{2\lambda\nu_B} \quad \Rightarrow \quad \rho_A = \tfrac{C}{2\nu_B + C} \tag{10}$$

$$\theta_B = \tfrac{C}{2\lambda\nu_A} \quad \Rightarrow \quad \rho_B = \tfrac{C}{2\nu_A + C}. \tag{11}$$

The derivations of (10) and (11) can be found in Appendix B. The following Lemma formalizes that the existence of two information diffusing simultaneously indeed causes them to (partially) crowd each other out, and establishes the dependence of the steady-states on ν_l and λ.

Lemma 1. 1. *For any finite $\lambda > 1$ and any $\nu_l \in (0,1)$, ρ_l is strictly increasing in ν_l. If, in addition, $\nu_A \geq \nu_B$, then*

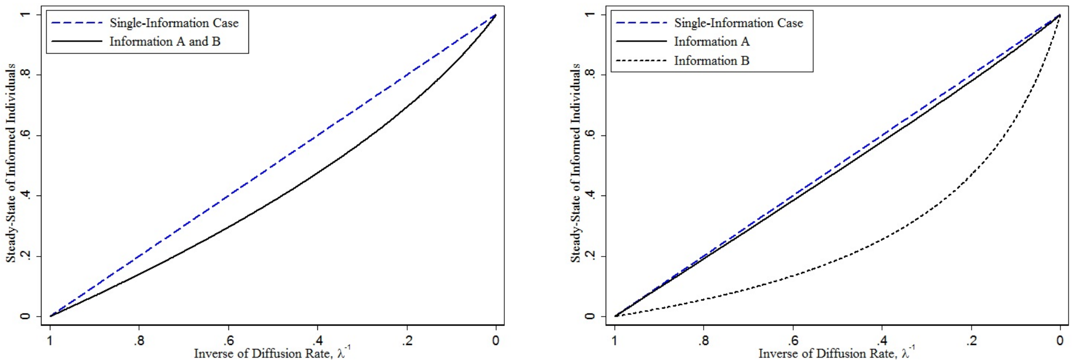
$$0 < \rho_B \leq \rho_A < \tilde{\rho}$$

holds, with strict inequality if $\nu_A > \nu_B$.

2. *For any $\nu_l \in (0,1)$, ρ_l is strictly increasing in λ.*

Proof. See Appendix C. □

As stated in the preceding section, it is indeed interest in an information that determines its prevalence in the population. Lemma also 1 shows that, although crowding out is never complete for interior values of ν_l, it is also never zero. The fact that communication time is fixed and has to be split between different topics by itself demands that fewer people are aware of each information in the two-information scenario than in the single-information case. In fact, crowding out may be substantial. In figure 1, we compare the positive steady-states for ρ_l for varying values of ν_l and λ to the corresponding values of $\tilde{\rho}$. The horizontal axes plot λ^{-1} ranging from 1 to 0, while the vertical axes give $\tilde{\rho}$, ρ_A, and ρ_B. In panel 1a, we set $\nu_A = \nu_B = 0.5$, while in panel 1b, the values are $\nu_A = 0.8$ and $\nu_B = 0.2$.



(a) $\nu_A = \nu_B = 0.5$ (b) $\nu_A = 0.8$, $\nu_B = 0.2$

Figure 1: Population information rates in steady-state as functions of the inverse of the diffusion rate, λ^{-1}.

The effect on ρ_l in moving from $\nu_B = 0.5$ to $\nu_B = 0.2$ is indeed substantial. As an example, a value of $\lambda = 2$ is sufficient for 50% of the population to be informed about B if it was the only information diffusing. The same value of λ leads to roughly 38.2% of the population informed about B if both A and B are preferred by half the population. However, if 80% of the population prefer to talk about A, a value of $\lambda = 2$ implies that the prevalence of B drops to 18.78% of the population.[14] These are large differences, especially considering that we consider preferences over topics to be ordinal. The fact that 80% of the population prefer to talk about topic A does not imply that these agents have no interest in B whatsoever. The difference in valuation might be small.[15]

Figure 1 highlights a second interesting aspect of the crowding out process. The impact of λ^{-1} on ρ_l is convex, i.e., crowding out depends on the exact parameters of the diffusion process. This is due to the two effects that an increase in λ has on ρ_l, discussed before: It directly increases ρ_l as information diffuses more easily, but as it also increases ρ_{-l} there exists

[14]Conversely, to inform half of the population about B if $\nu_B = 0.2$ would require a value of $\lambda \approx 5.55$.

[15]This interpretation would differ if we had considered ν_l to be the likelihood with which each agent communicates information l at a meeting. Under such an interpretation, the difference between ν_A and ν_B should be treated as cardinal. Also in our current interpretation, a cardinal interpretation of $\nu_A - \nu_B$ is possible: It might be plausible that a vast majority of the population will prefer one information over another only if there is also a large difference in interest. However, our present setup does not exclude the possibility that small differences in information valuation might lead to big differences in the proportion of informed agents. We return to implications of cardinal differences in information valuation in section 4.

a negative, indirect, effect on ρ_l. The following Proposition establishes how crowding out and relative prevalence depend on network characteristics.

Proposition 2. *For any $\nu_l \in (0,1)$ and finite $\lambda > 1$, the following holds:*

1. *There exists a unique value of λ, denoted λ_c that maximizes crowding out of both information,*
$$\lambda_c = \frac{1}{1-\nu_A\nu_B}\left[1+\frac{1-2\nu_A\nu_B}{(\nu_A\nu_B)^{1/2}}\right],$$
which is decreasing in $\nu_A\nu_B$.

2. *The ratio of crowding out of A relative to B is independent of λ, $\frac{\tilde{\rho}-\rho_A}{\tilde{\rho}-\rho_B} = \left(\frac{\nu_B}{\nu_A}\right)^2$.*

3. *If, in addition, $\nu_A > \nu_B$, the ratio $\frac{\rho_A}{\rho_B}$ is decreasing in λ.*

Proof. See Appendix D. □

The first part of Proposition 2 establishes that the existence of $-l$ harms the prevalence of l most for "intermediate" values of the diffusion rate. An increase in λ always increases the prevalence of any information, independent of whether it is the unique information or not. But, starting from $\lambda = 1$, initially an increase in λ would have a larger effect on the prevalence of a unique information, as compared to multiple information. This order is reversed above λ_c. Instead, as part two of the Proposition shows, the relative harm that information impose on each other depends entirely on relative interest in the population. Small changes in ν_l can lead to big changes in relative crowding out; the information that is preferred by the majority of the population imposes a more than proportional externality on the minority information. Finally, while an increase in λ never reverses the order of ρ_A and ρ_B, we find that the information preferred by the minority gains relatively more from such an increase.

4 Segregation and Integration

4.1 Information Prevalence

In the preceding analysis, agents of groups A and B interact randomly with each other, irrespective of group membership. It is however a well-documented fact that individuals have a tendency to interact relatively more with others that are similar to them, i.e., interaction patterns exhibit *homophily*.[16] Questions of how homophily might impact the diffusion of various states or ideas have gained further prominence with the rise of the Internet and Online Social Networks (OSNs), as these arguably make it easier for individuals to segregate according to their background or interests (see, e.g., Jackson (2008a), Jackson and López-Pintado (2013), Golub and Jackson (2012) for questions on how homophily impacts diffusion of a state or the speed of learning, or Sunstein (2009), Flaxman, Goel, and Rao (2013), or Gentzkow and Shapiro (2010) for the impact of homophily on news consumption). So far, results point in the direction that increased segregation does not lead to more segregated news consumption, and that homophily may both help the diffusion of a state (if groups have difference diffusion rates, λ, see Jackson and López-Pintado (2013)), but that it also can slow down learning (in a model in which agents update their beliefs by taking averages over their neighbors' beliefs, see Golub and Jackson (2012)).

Our framework focuses on the impact that homophily has on the long run prevalence of information. It highlights a complementary channel to those discussed in the previous literature: Homophily can affect information prevalence in the long run not only because of its (possible) effect on news consumption, or through biases in updating rules, but because it can affect the existence of information in the population. Our first result compares the prevalence of information in a segregated society (agents only meet others of their own group) to that in a fully integrated society (agents randomly meet, irrespective of group membership). It arises as a Corollary of Proposition 1.

Corollary 1. *Assume that the society is fully segregated according to interest groups. Then, for any finite $\lambda > 1$, the prevalence of information l among members of group l is $\tilde{\rho}$, while the prevalence of information $-l$ in group l is zero.*

The implications of Corollary 1 are stark. Independent of the amount of media coverage, or the diffusion rate of information λ, information B will never become endemic in group A and *vice versa*. This in itself gives credence to the idea that segregation might lead to polarization. If we interpret A and B as two alternative points of view on the same topic, segregation immediately implies polarization: Under full segregation, there exists no positive steady-state for ρ_{AB} and

[16] One of the earliest work on this is Lazarsfeld, Merton, et al. (1954). See also the survey by McPherson, Smith-Lovin, and Cook (2001).

agents are informed of at most one point of view. This occurs even if initial news consumption is entirely unbiased, and is also independent of biases in messages, updating, or bounded memory.

While the potential for polarization due to segregation is in itself important, our next result establishes that the prevalence of *either* information is lower under segregation compared to an integrated society.

Proposition 3. *For $\nu_l \in (0,1)$ and finite $\lambda > 1$, the prevalence of information l is ρ_l in an integrated society and it is $\nu_l \tilde{\rho}$ in a segregated society. We find that:*

1. *The prevalence of information l is larger in an integrated society than in a segregate society.*

2. *The prevalence of information l among group l, $\nu_l \rho_l$, is larger in a segregated society than in an integrated society.*

Proof. See Appendix F □

Proposition 3 shows that not only does segregation potentially lead to polarization, it also reduces the overall proportion of the population informed about either topic. This effect goes beyond the potentially negative impact on society of having no agent informed about both A and B in the steady-state. Indeed, there are many instances in which information A and B could be entirely unrelated (such as "music" and "world news"), in which case there is no immediate perceived benefit to society to have individuals informed about both topics at the same time. Nevertheless, even in such a case, the impact of segregation is to harm information prevalence. In a segregated society, there are fewer agents informed about *either* "music" *and* "world news". However, we also find that the prevalence of "music" *among agents who prefer "music"* is larger in a segregated society, the same for agents who prefer "world news". This leads to question us under which conditions agents themselves have incentives to segregate, which we address now.

4.2 Endogenous Segregation

We need to impose some additional structure on the utility agents gain from information, to address questions of endogenous segregation. To keep the analysis as simple as possible, we assume that agents derive utility directly from being informed about topics A and B. We assume that an agent in group l receives a flow utility of h if he is informed about topic l and a flow utility of s if he is informed about topic $-l$, where $h \geq s \geq 0$.[17] The prevalence of information l, ρ_l, denotes the time that an agent spends being informed about l in steady-state. We also assume that agents care only about the steady-state values of ρ_l and ρ_{-l}. The utility of an agent in group l in an integrated and a segregated society is then

$$U_{l|int} = \rho_l h + \rho_{-l} s, \quad \text{and} \tag{12}$$

$$U_{l|seg} = \tilde{\rho} h. \tag{13}$$

Corollary 2 follows immediately from these utilities and the fact that $\rho_l < \tilde{\rho}$.

Corollary 2. *Assume that $\lambda > 1$ and finite. If $h > 0$ and $s = 0$, all agents prefer segregation over integration. If $s = h > 0$, all agents prefer integration.*

Given the results on the prevalence of information under segregation and integration, the decision to segregate depends, quite naturally, on how valuable the less preferred information is to agents. In particular, if agents have extreme preferences for just one unique information, they will always choose to segregate. It is noteworthy that agents of either group, A and B, value a segregated society equally. The utility that an agent receives in an integrated society, on the other hand, depends on the values of ρ_l and ρ_{-l}: The larger the group size of l, ν_l, the larger will be ρ_l and the smaller ρ_{-l}. In general, group l prefers segregation over integration if

$$\frac{\tilde{\rho} - \rho_l}{\rho_{-l}} > \frac{s}{h}. \tag{14}$$

By definition, $h > s$, which implies that $\frac{s}{h} \in [0, 1]$. Equation (14) allows us to strengthen the result of Corollary 2 to all interior values of $\frac{s}{h}$.

Proposition 4. *For any $\nu_l \in (0, 1)$ and finite $\lambda > 1$, a decrease in $\frac{s}{h}$ makes segregation more attractive to either group.*

[17] Such utility flows could arise if agents truly valued information in itself, but also if they value it because there is the possibility that it will be useful at an uncertain, future, date. E.g., agents might value to be informed about current events / history / politics, not so much because it provides them with any benefit as such, but because there is a chance that these topics might be discussed in their presence, and not being informed would brand them as ignorant. Alternatively, the information might pertain to the state of the world and an agent knows that at an uncertain point in the (distant) future he will have to take an action to match that state. In either case, the expected utility of an agent would be increasing in the amount of time he is informed.

Proof. Obvious given equation (14). □

I.e., the more extreme information preferences are, the more likely it is that we observe a segregated society. The exact value of $\frac{s}{h}$ for which segregation will occur, on the other hand, depends on the parameters on the network. In particular, it depends on ν_l and λ, as these are the parameters that determine the value of $\frac{\tilde{\rho}-\rho_l}{\rho_{-l}}$. Their impact is analyzed in Proposition 5.

Proposition 5. *For all finite* $\lambda > 1$ *and* $h > s > 0$, $\frac{\tilde{\rho}-\rho_l}{\rho_{-l}}$ *is decreasing in both* ν_l *and* λ.

Proof. See Appendix F. □

That is, as we decrease the size of the group that prefers information l, segregation is preferred for larger values of $\frac{s}{h}$. As segregation is driven by the group that has a lower valuation for an integrated society, this result implies that a society in which interest in information A and B is evenly split, $\nu_A = \nu_B$, is least likely to segregate into separate groups. On the other hand, for both groups the incentive to segregate is decreasing in λ. Propositions 4 and 5 paint a mixed picture of the effect of the Internet and OSNs on segregation. Given our results, we would expect that groups that make use of the Internet to segregate themselves are those that (i) are particularly interested in niche or very specialized topics, that are of little interest in the population in general, and (ii) care about little else outside their topic of interest. On the other hand, if the change in communication technology has translated into a denser network (an increase in k that results in an increase in λ), all groups will experience lower incentives to segregate into their respective groups of interest.

5 Conclusion

The present paper models the diffusion process of two information through a population. Information may be passed between individuals whenever a meeting occurs. Each individual has a given number of meetings per period, and at each meeting at most one information can be communicated. Individuals may be aware of either one or both information, and if they are aware of both, we assume that they pass on the information they find more interesting.

We find that, as long as the population's views on which is the more interesting information are not perfectly aligned, both information will be endemic under exactly the same model parameters as in the case of only one information diffusing through the population. This result provides an insight into how it is possible that even obscure (or even incorrect) pieces of information survive in a society, and why it appears so difficult to eradicate information completely once it has been made public. This resilience of information notwithstanding, the two information impose (potentially significant) negative externalities on each other, and neither information has a prevalence as high as it would have if it was the unique information diffusing. As intuition predicts, the information that is considered more interesting by the larger proportion of the population will have a larger prevalence, and the prevalence of either information is increasing in the overall diffusion rate. In relative terms, however, the information preferred by the minority of the population will gain more from an increase in the diffusion rate than the "stronger" information.

One application of our model is the question of how (if at all) different news stories may crowd each other out. It is indeed true that a news story receive differential media coverage, however, this should only affect the initial information seed, i.e., how many individuals are initially informed about a story. In our model, this will not have an impact on the prevalence of a particular story. If, on the other hand, the relative importance of different news stories relates to the likelihood with which individuals spread these stories, we will expect that in the long run, the less important story will be known by fewer individuals, independent of initial media coverage.

A second application relates to the effect that segregation according to topics of interest will have on information prevalence and polarization in a society. Our model predicts that if agents communicate exclusively with others that share their preferred information, we will observe both polarization, and an overall decrease in the number of informed agents. It is indeed impossible to ensure that a non-preferred information has a positive prevalence in a group that is entirely segregated from the rest of society. This result is particularly stark as it does not rely on a biased consumption of news, or any bias in how agents update information or learn. Segregation in turn tends to be preferred by agents in the minority interest group, and those that have larger differences in their valuation of the two information. Both of these results have an intuitive appeal. OSNs and the Internet in general, by making it easier for agents to segregate according to their interests, might therefore indeed lead not only to polarization, but also to an overall loss in information, driven primarily by minority groups and agents with extreme information preferences. On the other hand, we find that a denser network will reduce the incentives to segregate. Consequently, with respect to segregation, polarization, and overall information prevalence, the Internet has an *ex ante* ambiguous impact.

While our model is obviously highly stylized, our two main assumptions appear to us to be rather innocuous, especially for the type of information that we model, such as chit-chat and gossip: (i) Agents face constraints in their communication time, relative to the amount of information they are aware of; and (ii) if faced with such constraints, agents are more likely to talk about information they care about more. The resulting loss of information prevalence in the population can be seen as complementary to theories of loss of information based on cognitive constraints. We do not impose any limits on the amount of information that agents can store or process, for example.

Our results on the resilience of information are in stark contrast to earlier results by Prakash, Beutel, Rosenfeld, and Faloutsos (2012) and Beutel, Prakash, Rosenfeld, and Faloutsos (2012), who investigated competing viruses or products. These differences stem predominantly from the difference in which we treat agents that are infected with both information/viruses. In our model, the competition between information affects the transmission likelihood *from* an infected node, while in Prakash, Beutel, Rosenfeld, and Faloutsos (2012) and Beutel, Prakash, Rosenfeld, and Faloutsos (2012), competition affects the infection likelihood *to* a susceptible node. Indeed, in Prakash, Beutel, Rosenfeld, and Faloutsos (2012), no node can be infected with both viruses at the same time. The fact that such a small difference in the transmission process yields such divergent results highlights the need to exhibit caution in the modeling of the diffusion process. The SIS framework has been used to model the diffusion of viruses, products, behavior, information, and more. But we believe that for information transmission, our assumptions are a better description of the diffusion process of multiple information than for, e.g., products or viruses, where complete or partial immunity is much more credible.

Our main results on information prevalence are strongly reminiscent of the SIS model with only a unique state, e.g., the existence of an identical diffusion threshold and the uniqueness of a positive steady-state. This leads us to conjecture that these results will hold even if we increase the number of information above two, although we leave the formal proof of this for further research. Similarly, we believe that further work on different network structures has the potential to yield interesting insights into more complex diffusion processes.

A Diffusion Thresholds and Steady-States

Proof of Proposition 1

In steady-state,

$$\rho_A = \frac{\lambda\theta_A}{1+\lambda\theta_A} \tag{15}$$

$$\rho_B = \frac{\lambda\theta_B}{1+\lambda\theta_B} \tag{16}$$

with

$$\theta_A = \rho_A[1-\nu_B\rho_B] \tag{17}$$

$$\theta_B = \rho_B[1-\nu_A\rho_A] \tag{18}$$

where we made use of the fact that $\rho_{AB} = \rho_A\rho_B$. Due to the symmetry of information A and B, note that we can change the labels of the information to apply any arguments that we make about A also for B. We will therefore prove proposition 1 for information A, without loss of generality.

First, by substituting equation (17) into (15) and (18) into (16), it is immediate that the steady-state $\rho_A = \rho_B = 0$ always exists. Furthermore, if $\nu_A = 1$, then (15) becomes

$$\rho_A = \frac{\lambda\rho_A}{1+\lambda\rho_A} \tag{19}$$

which is identical to the steady-state condition in the one-information case, i.e., the uniquely positive steady-state of A is $\rho_A = 1-\lambda^{-1}$, and it is globally stable. This proves the final part of proposition 1.

We constrain ourselves to look now for the existence and stability properties of steady-states in which $\rho_l > 0$ for both l. If $\rho_B > 0$, we can rewrite equation (16) as a function of ρ_A and parameters only:

$$\rho_B = 1 - \frac{1}{\lambda(1-\nu_A\rho_A)} \tag{20}$$

Substituting equations (15) and (20) into equation (17), we can express θ_A as

$$\theta_A = H(\theta_A) = \frac{\lambda\theta_A}{1+\lambda\theta_A}\left[\nu_A + \frac{\nu_B}{\lambda\left(1-\nu_A\frac{\lambda\theta_A}{1+\lambda\theta_A}\right)}\right], \tag{21}$$

i.e.,

$$H(\theta_A) = \frac{\lambda\nu_A\theta_A}{1+\lambda\theta_A} + \frac{\nu_B\theta_A}{1+\nu_B\lambda\theta_A} \tag{22}$$

And the steady-state of θ_A is the fixed point of (22). Following the argument put forward in López-Pintado (2008), note that $H(0) = 0$ and that

$$H(1) = \frac{\lambda\nu_A}{1+\lambda} + \frac{\nu_B}{1+\lambda\nu_B},$$

which some manipulation shows to be strictly below 1. I.e., if $H'(\theta_A) > 0$, $H'(0) > 1$ and $H''(\theta_A) < 0$, then any $\theta_A^* > 0$ that solves equation (22) is unique and globally stable. Consequently, so is any ρ_A^* relating to θ_A^*.

Indeed,

$$H'(\theta_A) = \frac{\lambda\nu_A}{(1+\lambda\theta_A)^2} + \frac{\nu_B}{(1+\lambda\nu_B\theta_A)^2} > 0, \quad (23)$$

$$H''(\theta_A) = -2\lambda\left[\lambda\nu_A(1+\lambda\theta_A)^{-3} + \nu_B^2(1+\lambda\nu_B\theta_A)^{-3}\right] < 0, \quad (24)$$

and

$$H'(0) = \lambda\nu_A + \nu_B, \quad (25)$$

which is larger than 1 for $\nu_l \in (0,1)$ if and only if $\lambda > 1$. This completes the proof that for $\nu_l \in (0,1)$, the diffusion thresholds for both A and B are identical and equal to 1. If, however, $\nu_A = 0$, we see that independently of the value of λ, $H'(0) = 1$, i.e., there exists no positive steady-state for θ_A, consequently not for ρ_A either. This completes the proof that for $\nu_l = 0$, information l does not have a strictly positive steady-state.

B Derivation of steady-state θ_l

From equation (22), we know that a steady-state of the θ_A is such that

$$\theta_A = \frac{\lambda\nu_A\theta_A}{1+\lambda\theta_A} + \frac{\nu_B\theta_A}{1+\nu_B\lambda\theta_A}. \quad (26)$$

Note again that $\theta_A = 0$ is always a solution. For $\theta_A > 0$, re-arranging of (26) yields

$$(1+\lambda\theta_A)(1+\lambda\nu_B\theta_A) = \lambda\nu_A(1+\lambda\nu_B\theta_A) + \nu_B(1+\lambda\theta_A)$$
$$1 + \lambda\theta_A(1+\nu_B) + \lambda^2\nu_B\theta_A^2 = \lambda\nu_A + \lambda^2\nu_A\nu_B\theta_A + \nu_B + \lambda\nu_B\theta_A$$
$$\nu_A(1-\lambda) + \lambda^2\nu_B\theta_A^2 + \lambda\theta_A(1-\lambda\nu_A\nu_B) = 0$$

which in turn implies that

$$\theta_{A_{1,2}} = \frac{1}{2\lambda\nu_B}\left\{-(1-\lambda\nu_A\nu_B) \pm \left[(1-\lambda\nu_A\nu_B)^2 - 4\nu_A\nu_B(1-\lambda)\right]^{\frac{1}{2}}\right\}.$$

Note that for all $\lambda > 1$, the square-root is larger than $(1-\lambda\nu_A\nu_B)$, which implies that there exists a unique positive steady-state for θ_A^*, with

$$\theta_A^* = \frac{1}{2\lambda\nu_B}\left\{-(1-\lambda\nu_A\nu_B) + \left[(1-\lambda\nu_A\nu_B)^2 - 4\nu_A\nu_B(1-\lambda)\right]^{\frac{1}{2}}\right\} = \frac{C}{2\lambda\nu_B}, \quad (27)$$

as stated in equation (10). Finally, as $\rho_A = \frac{\lambda\theta_A}{1+\lambda\theta_A}$, by equation (6), plugging this value of θ_A in, we arrive at the result that $\rho_A^* = \frac{C}{2\nu_B + C}$. The derivations of θ_B^* and ρ_B^* are identical.

C Properties of ρ_l

Proof of Lemma 1, Point 1

The ranking of ρ_B, ρ_A, and $\tilde{\rho}$ is established straightforwardly: The fact that $\rho_A > \rho_B$ if and only if $\nu_A > \nu_B$ and that $\rho_A = \rho_B$ if and only if $\nu_A = \nu_B$ is obvious from equations (10) and (11). Also, note that for $\rho_A > 0$, we know that

$$\rho_A = 1 - \frac{1}{\lambda(1-\nu_B\rho_B)}$$

which can be used to show that

$$\begin{aligned}\tilde{\rho} - \rho_A &= 1 - \frac{1}{\lambda} - 1 + \frac{1}{\lambda(1-\nu_B\rho_B)} \\ &= \frac{\nu_B\rho_B}{\lambda(1-\nu_B\rho_B)}\end{aligned} \tag{28}$$

which is strictly positive whenever $\rho_B > 0$ and zero if $\rho_B = 0$. This establishes that $0 < \rho_B \leq \rho_A < \tilde{\rho}$.

As $\rho_A = \frac{\lambda\theta_A}{1+\lambda\theta_A}$, it is increasing in ν_A whenever θ_A is increasing in ν_A, i.e., if

$$\frac{d\theta_A}{d\nu_A} = \frac{1}{2\lambda\nu_B^2}\left[C + \frac{dC}{d\nu_A}\nu_B\right] > 0, \tag{29}$$

which in turn holds if $C + \frac{dC}{d\nu_A}\nu_B > 0$. Given the definition of C, we work with a change of variable of $x = \nu_A\nu_B$, which means that we can calculate $\frac{dC}{d\nu_A} = \frac{dC}{dx}(1-2\nu_A)$, where

$$\frac{dC}{dx} = \lambda + \frac{-2+\lambda(1+\lambda x)}{c^{1/2}}$$

with $c = (1-\lambda x)^2 - 4x(1-\lambda)$. This is obviously positive if $-2+\lambda(1+\lambda x) > 0$ (which holds for sure for all $\lambda > 2$). If $-2+\lambda(1+\lambda x) < 0$, some re-arranging yields the same result, i.e., $\frac{dC}{dx} > 0$ if

$$\begin{aligned}\lambda c^{1/2} &> 2 - \lambda(1+\lambda x) \\ \lambda^2(1-\lambda x)^2 - 4x\lambda^2(1-\lambda) &> 4 - 4\lambda - 4\lambda^2 x + \lambda^2 + 2\lambda^3 x + \lambda^4 x^2 \\ 0 &> 4(1-\lambda)\end{aligned}$$

which is true for all $\lambda > 1$. The sign of $\frac{dC}{d\nu_A}$ is hence ambiguous; it is positive if $\nu_A < \frac{1}{2}$, zero if $\nu_A = \frac{1}{2}$, and negative otherwise. Hence, for $\nu_A > \frac{1}{2}$, the sign of $\frac{d\theta}{d\nu_A}$ is not immediately obvious. In full, $\frac{d\theta}{d\nu_A} > 0$ if

$$\begin{aligned}-(1-\lambda x) + c^{1/2} + [\lambda c^{-1/2}(-2+\lambda(1+\lambda x))](\nu_B - 2x) &> 0 \\ -(1-\lambda x)c^{1/2} + c + [\lambda c^{1/2} - 2 + \lambda(1+\lambda x)](\nu_B - 2x) &> 0 \\ 1 + 2\lambda x + \lambda^2 x^2 - 2\nu_B + \lambda\nu_B(1+\lambda x) - 2\lambda x(1+\lambda x) &> -c^{1/2}(-1+\lambda\nu_B - \lambda x) \\ (1+\lambda x)(1+\lambda\nu_B^2) &> c^{1/2}(1-\lambda\nu_B^2).\end{aligned} \tag{30}$$

Again, we consider two cases, if $1-\lambda\nu_B^2 < 0$, this condition is always satisfied. If $1-\lambda\nu_B^2 > 0$, we define $D \equiv \frac{1-\lambda\nu_B^2}{1+\lambda\nu_B^2}$ and can write

$$\begin{aligned}(1+\lambda x) &> c^{1/2}D \\ (1+\lambda x)^2 &> (1-\lambda x)^2 D^2 - 4x(1-\lambda)D^2 \\ (1+\lambda x)^2 &> (1+\lambda x)^2 D^2 - 4xD^2\end{aligned} \tag{31}$$

which is always true as $D < 1$. Consequently, $\frac{d\rho_A}{d\nu_A} > 0$ for all values of ν_A and $\lambda > 1$.

Proof of Lemma 1, Point 2

Note that ρ_l is increasing in C (strictly if $\nu_l \in (0,1)$). To prove the dependence of ρ_l on λ, it suffices to show that C is increasing in λ. Indeed,

$$\frac{dC}{d\lambda} = x + \frac{x(1+\lambda x)}{c^{1/2}} > 0 \tag{32}$$

Q.E.D.

D Crowding out

Note that, as the ratio of crowding out for $\nu_l \in (0,1)$ is independent of λ, a unique value of λ_c must maximize crowding out of both A and B. At least one finite λ_c must exist, as for both $\lambda = 1$ and $\lambda \to \infty$, crowding out of either information is zero, while for finite $\lambda > 1$, by Lemma 1, crowding out is positive. We proceed to show that λ_c is unique, i.e., it maximizes crowding out, and is as stated in Proposition 3. We work again with ρ_A. It is convenient to work with $\rho_A = 1 - \frac{1}{\lambda(1-\nu_B\rho_B)}$. In this case,

$$\frac{d(\tilde{\rho}-\rho_A)}{d\lambda} = \nu_B \frac{\lambda\frac{d\rho_B}{d\lambda} - \rho_B(1-\nu_B\rho_B)}{\lambda^2(1-\nu_B\rho_B)^2} \tag{33}$$

which is equal to zero if

$$\lambda\frac{d\rho_B}{d\lambda} = \rho_B(1-\nu_B\rho_B) \tag{34}$$

substituting $\rho_B = \frac{C}{2\nu_A + C}$ and $\frac{d\rho_B}{d\lambda} = \frac{2\nu_A\frac{dC}{d\lambda}}{(2\nu_A+C)^2}$, this simplifies to

$$2\lambda\frac{dC}{d\lambda} = C(2+C). \tag{35}$$

Substituting the values for C from the main text and for $\frac{dC}{d\lambda}$ from equation (32), this simplifies to

$$\begin{aligned}
2\lambda x\frac{c^{1/2}+1+\lambda x}{c^{1/2}} &= [-(1-\lambda x)+c^{1/2}](1+\lambda x+c^{1/2}) \\
\left[1+\lambda^2x^2-4x\right]^2 &= (1-\lambda x)^2 c \\
\left[(1-\lambda x)^2+2x(\lambda-2)\right]^2 &= (1-\lambda x)^4 - 4x(1-\lambda)(1-\lambda x)^2 \\
x(\lambda-2)^2 &= (1-\lambda x)^2 \\
\lambda^2x(1-x) - 2\lambda x - 1 + 4x &= 0
\end{aligned} \tag{36}$$

and the solution of

$$\lambda_c = \frac{1}{1-x}\left[1 \pm \frac{1-2x}{x^{1/2}}\right]. \tag{37}$$

It is straightforward to show that only $\lambda_c = \frac{1}{1-x}\left[1+\frac{1-2x}{x^{1/2}}\right]$ satisfies the additional condition that $\lambda_c > 1$. Substituting $x = \nu_A\nu_B$ yields the expression in Proposition 2. Finally, taking the derivative of λ_c with respect to x yields

$$\begin{aligned}
\frac{d\lambda_c}{dx} &= \frac{1}{(1-x)^2} - \frac{1}{x(1-x)}\left[\frac{3}{2}x^{1/2} + \frac{1}{2}x^{-1/2}\right] \\
&= \frac{1}{2x^{3/2}(1-x)^2}\left[2x^{3/2} - x^2 - 1\right] < 0
\end{aligned} \tag{38}$$

which is negative as the maximum value that x can take is $\frac{1}{4}$, i.e., the term in square brackets, which is itself increasing in x, can never be positive.

Relative crowding out, however, does not depend on network characteristics. Using the expression of crowding out of information l established in equation (28), we can express relative crowding out as

$$\frac{\tilde{\rho} - \rho_A}{\tilde{\rho} - \rho_B} = \frac{\nu_B}{\nu_A} \frac{\rho_B(1 - \nu_A \rho_A)}{\rho_A(1 - \nu_B \rho_B)} \tag{39}$$

and substituting equations (6) and (7), as well as noting that $\theta_l = \frac{C}{2\lambda\nu_l}$, into this expression yields the result,

$$\begin{aligned} \frac{\tilde{\rho} - \rho_A}{\tilde{\rho} - \rho_B} &= \frac{\nu_B}{\nu_A} \frac{\theta_B(1 + \nu_B \lambda \theta_A)}{\theta_A(1 + \nu_A \lambda \theta_B)} \\ &= \left(\frac{\nu_B}{\nu_A}\right)^2. \end{aligned} \tag{40}$$

Finally, as $\rho_l = \frac{C}{2\nu_l + C}$, we have $\frac{\rho_A}{\rho_B} = \frac{2\nu_A + C}{2\nu_B + C}$. The derivative with respect to λ is straightforward,

$$\frac{d\frac{\rho_A}{\rho_B}}{d\lambda} = \frac{2\frac{dC}{d\lambda}}{(2\nu_B + C)^2}(\nu_B - \nu_A), \tag{41}$$

which is decreasing if $\nu_A > \nu_B$.

E Basic Reproduction Number

The basic reproduction number for information A is

$$R_A = \lambda(1 - \nu_B \rho_B). \tag{42}$$

As the prevalence of A can be written as

$$\rho_A = 1 - \frac{1}{\lambda(1 - \nu_B \rho_B)} = 1 - \frac{1}{R_A} \tag{43}$$

it is trivial that there exists a positive prevalence of A if and only if $R_A > 1$. It is also trivial that if either $\nu_B = 0$ or $\rho_B =$, then $R_A > 1 \Leftrightarrow \lambda > 1$. It remains to show that this is also true if $\rho_B > 0$ and $\nu_B \in (0, 1)$.

The positive prevalence of B is

$$\rho_B = \frac{C}{2\nu_A + C}. \tag{44}$$

Substituting this expression into equation (42), we can write

$$R_A = \lambda \frac{\nu_A(2 + C)}{2\nu_A + C} \tag{45}$$

and as $C = -(1 - \lambda\nu_A\nu_B) + \left[(1 - \lambda\nu_A\nu_B)^2 - 4\nu_A\nu_B(1 - \lambda)\right]^{\frac{1}{2}}$, we find that if $\lambda = 1$, then $C = 0$. In this case, equation (45) simplifies to

$$R_A = \lambda = 1. \tag{46}$$

To complete the proof, we show that R_A is strictly increasing in λ, i.e., for any value of $\lambda > 1$, $R_A > 1$ as well. Note that we have already established in Appendix C that $\frac{d\rho_A}{d\lambda} > 0$. But, since we can write $\rho_A = 1 - R_A^{-1}$, it is true that

$$\frac{d\rho_A}{d\lambda} = R_A^{-2} \frac{d\ R_A}{d\lambda} \tag{47}$$

which is positive only if $\frac{d\ R_A}{d\lambda} > 0$. QED.

F Segregation and Integration

The fact that in a fully segregated society more agents are in the long run informed about their preferred information follows trivially from the fact that for finite λ and $\nu_l \in (0,1)$, $\rho_l < \tilde{\rho}$. If the society is fully integrated, the proportion of agents informed about l who also care about l is simply $\nu_l \rho_l$. Under full segregation, society is split into two exclusive groups, of (proportionate) size ν_l and $1-\nu_l$ respectively. Hence, a total proportion of $\nu_l \tilde{\rho}$ will be informed of their preferred information in the long run. As $\tilde{\rho} > \rho_l$, so $\nu_l \tilde{\rho} > \nu_l \rho_l$.

Conversely, integration implies a larger proportion of the population being informed about l (irrespective of preferences) if $\rho_l > \nu_l \tilde{\rho}$. Working with $l = A$, this is true if

$$\nu_A \tilde{\rho} - \rho_A = \frac{2\nu_A \nu_B(\lambda - 1) - C(\nu_B \lambda + \nu_A}{\lambda(2\nu_B + C)} < 0, \tag{48}$$

i.e., if

$$\begin{aligned} (\nu_B\lambda + \nu_A)[-(1 - \lambda\nu_A\nu_B) + c^{1/2}] &> 2\nu_A\nu_B(\lambda - 1) \\ (\nu_B\lambda + \nu_A)c^{1/2} &> 2\nu_A\nu_B(\lambda - 1) + (\nu_B\lambda + \nu_A)(1 - \lambda x) \\ (\nu_B\lambda + \nu_A)^2 &> \nu_A\nu_B(\lambda - 1) + (\nu_B\lambda + \nu_A)(1 - \lambda\nu_A\nu_B) \\ (\nu_B\lambda + \nu_A)\nu_B[\lambda(1 + \nu_A) - 1] &> \nu_A\nu_B(\lambda - 1) \\ (1 + \nu_A)(\lambda\nu_B + \nu_A) &> 1 \end{aligned} \tag{49}$$

which is always true, as both terms on the left-hand side are strictly larger than 1 for any $\lambda > 1$ and any $\nu_l \in (0,1)$. As this expression also holds if we switch the labels of A and B, also $\nu_B \tilde{\rho} < \rho_B$, which completes the proof.

Proof of Proposition 5

We provide here the proof of Proposition 5 for the case of group A by showing that $\frac{d\frac{\tilde{\rho}-\rho_A}{\rho_B}}{d\nu_A} < 0$. The proof for group B is analogous. First, note that

$$\frac{\tilde{\rho} - \rho_A}{\rho_B} = \frac{\nu_B}{\lambda(1 - \nu_B\rho_B)} \tag{50}$$

Its first derivative with respect to ν_A is then

$$\frac{d\frac{\tilde{\rho}-\rho_A}{\rho_B}}{d\nu_A} = \frac{-\lambda(1-\nu_B\rho_B) - \nu_B\lambda\left(\rho_B - \nu_B\frac{d\rho_B}{d\nu_A}\right)}{\lambda^2(1-\nu_B\rho_B)^2}$$
$$= \frac{1}{\lambda(1-\nu_B\rho_B)^2}\left[-1+\nu_B^2\frac{d\rho_B}{d\nu_A}\right] < 0 \tag{51}$$

as $\frac{d\rho_B}{d\nu_A} < 0$. QED.

Similarly, to proof that $\frac{\tilde{\rho}-\rho_l}{\rho_{-l}}$ is decreasing in λ, we give only the proof that $\frac{\tilde{\rho}-\rho_A}{\rho_B}$ is decreasing in λ, as the derivations for $\frac{\tilde{\rho}-\rho_B}{\rho_A}$ proceed in identical fashion.

First, we know that

$$\tilde{\rho} - \rho_A = \frac{\nu_B\rho_B}{\lambda(1-\nu_B\rho_B)}.$$

In Appendix E we defined $\lambda(1-\nu_B\rho_B) = R_A$, and showed that R_A is increasing in λ. Consequently,

$$\frac{d\frac{\tilde{\rho}-\rho_A}{\rho_B}}{d\lambda} = \frac{d\frac{\nu_B}{R_A}}{d\lambda}$$
$$= -\frac{\nu_B}{R_A^2}\frac{dR_A}{d\lambda} < 0 \tag{52}$$

QED.

References

Baccara, M., and L. Yariv (2008): "Similarity and polarization in groups," *Available at SSRN 1244442*.

Bailey, N. T., et al. (1975): *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd.

Beutel, A., B. A. Prakash, R. Rosenfeld, and C. Faloutsos (2012): "Interacting viruses in networks: can both survive?," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 426–434. ACM.

Bharathi, S., D. Kempe, and M. Salek (2007): "Competitive influence maximization in social networks," *Internet and Network Economics*, pp. 306–311.

Borodin, A., Y. Filmus, and J. Oren (2010): "Threshold models for competitive influence in social networks," *Internet and network economics*, pp. 539–550.

Dodds, P. S., and D. J. Watts (2004): "Universal Behavior in a Generalized Model of Contagion," *Phys. Rev. Lett.*, 92, 218701.

Dubey, P., R. Garg, and B. De Meyer (2006): "Competing for customers in a social network: The quasi-linear case," *Internet and Network Economics*, pp. 162–173.

Flaxman, S., S. Goel, and J. M. Rao (2013): "Ideological Segregation and the Effects of Social Media on News Consumption," *Available at SSRN*.

Galeotti, A., and B. W. Rogers (2013a): "Diffusion and protection across a random graph," .

——— (2013b): "Strategic immunization and group structure," *American Economic Journal: Microeconomics*, 5(2), 1–32.

Gentzkow, M., and J. M. Shapiro (2010): "Ideological segregation online and offline," Discussion paper, National Bureau of Economic Research.

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

Goyal, S. (2012): *Connections: an introduction to the economics of networks*. Princeton University Press.

Goyal, S., and M. Kearns (2012): "Competitive contagion in networks," in *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pp. 759–774. ACM.

Granovetter, M. S. (1973): "The strength of weak ties," *American journal of sociology*, pp. 1360–1380.

Jackson, M. O. (2008a): "Average distance, diameter, and clustering in social networks with homophily," *Internet and Network Economics*, pp. 4–11.

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

Jackson, M. O., and D. López-Pintado (2013): "Diffusion and contagion in networks with heterogeneous agents and homophily," *Network Science*, 1(01), 49–67.

Jackson, M. O., and B. W. Rogers (2007): "Relating network structure to diffusion properties through stochastic dominance," *The BE Journal of Theoretical Economics*, 7(1).

Jackson, M. O., and L. Yariv (2010): "Diffusion, strategic interaction, and social structure," *in Handbook of Social Economics, edited by J. Benhabib, A. Bisin and M. Jackson.*

Katz, E., and P. F. Lazarsfeld (1970): *Personal Influence, The part played by people in the flow of mass communications.* Transaction Publishers.

Lazarsfeld, P. F., B. Berelson, and H. Gaudet (1968): *The peoples choice: how the voter makes up his mind in a presidential campaign.* New York Columbia University Press.

Lazarsfeld, P. F., R. K. Merton, et al. (1954): "Friendship as a social process: A substantive and methodological analysis," *Freedom and control in modern society*, 18(1), 18–66.

López-Pintado, D. (2008): "Diffusion in complex social networks," *Games and Economic Behavior*, 62(2), 573 – 590.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001): "Birds of a feather: Homophily in social networks," *Annual review of sociology*, pp. 415–444.

Pastor-Satorras, R., and A. Vespignani (2001a): "Epidemic dynamics and endemic states in complex networks," *Phys. Rev. E*, 63, 066117.

——— (2001b): "Epidemic Spreading in Scale-Free Networks," *Phys. Rev. Lett.*, 86, 3200–3203.

——— (2002): "Immunization of complex networks," *Phys. Rev. E*, 65, 036104.

Prakash, B. A., A. Beutel, R. Rosenfeld, and C. Faloutsos (2012): "Winner takes all: competing viruses or ideas on fair-play networks," in *Proceedings of the 21st international conference on World Wide Web*, pp. 1037–1046. ACM.

Rosenblat, T. S., and M. M. Mobius (2004): "Getting closer or drifting apart?," *The Quarterly Journal of Economics*, pp. 971–1009.

Sunstein, C. R. (2009): *Republic. com 2.0.* Princeton University Press.

Watts, D. J. (2002): "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences*, 99(9), 5766–5771.