Diffusion of Multiple Information*

Nicole Tabasso†

January 30, 2015

Preliminary Version

Abstract

We introduce two pieces of information into a diffusion process in which information is transmitted when individuals meet and forgotten at an exogenous rate. At most one information can be transmitted at a meeting, which introduces opportunity costs in the process. Individuals differ according to which information they find more interesting, and that is the one they transmit if they face a choice. We find that both pieces of information survive under the same parameter values, and that relative interest is the main determinant in the number of people informed of one piece in the long run. Thus our model is able to explain the variety of information in the public sphere. Next, we apply our framework to answer questions relating to segregation according to information interests. We find that in a segregated group, only the preferred information survives, i.e., segregation leads to polarization. Segregation also reduces the overall number of people informed in the long run. Finally, we ask when does segregation endogenously occur. We find that it is more likely if information preferences are extreme, and/or the number of individuals interested in different pieces of information is very varied. Its likelihood decreases as information transmission itself is facilitated.

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

JEL Classification: D83, D85

*Many thanks for valuable comments, discussions and suggestions 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.

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

1 Introduction

Social networks are a crucial factor in the diffusion of information in a society, as highlighted already in in Lazarsfeld et al. (1968) or Katz and Lazarsfeld (1970). For information transmission through casual contacts, a standard way to model this diffusion is to treat "being informed" as a state that spreads like a contagion on a network (see Jackson and Rogers (2007), López-Pintado (2008), Jackson and Yariv (2010), Jackson and López-Pintado (2013)). So far, these models have assumed that there exists a unique state that spreads through the population. In the context of information though, it appears obvious that many distinct pieces of information diffuse simultaneously. This creates competition among them. In Twitter data, e.g., Leskovec et al. (2009) have shown that the arrival of a particularly popular hashtag does not lead to more tweets. The total volume of tweets stays roughly constant over time. Rather, it coincides with reduced tweets of other hashtags. Different pieces of information do appear to crowd each other out,[1] a fact that the standard model of information contagion does not accommodate.

The present paper introduces a model in which two pieces of information, $l = \{A, B\}$, simultaneously diffuse on a network, in which nodes denote agents and links meetings. The basic diffusion process is the *Susceptible-Infected-Susceptible* (SIS) framework, as employed in Jackson and Rogers (2007), López-Pintado (2008), Jackson and Yariv (2010), and Jackson and López-Pintado (2013), among others.

Our model exhibits three main features. First, as in the standard SIS framework, we consider a population in which agents transition between being susceptible to information (state S), or infected with it (state I). I.e., they are either uninformed, or informed. Agents transition between states either if an agent susceptible to information l becomes informed of it during a meeting, or if he forgets it. We focus on the existence and properties of the steady-state (*prevalence*) of each information. The introduction of two pieces of information increases the set of infectious states. Compared to the standard model, in which I is a single state, our model incorporates a set of infectious states of $I = \{I_{A\setminus B}, I_{B\setminus A}, I_{AB}\}$.

The second main feature, which is entirely novel, is the introduction of opportunity costs, through the assumption that the two pieces of information share an agent's limited communi-

[1] This fact has been recognized also long before the recent increase in online communication. A prime example is the accusation that policy makers might try to "bury" unfavorable news through a manipulation of their release date.

cation time. Whenever two agents meet, each can communicate at most one piece.

Finally, we relate the choice of which information to communicate to intrinsic preferences of agents. We assume that a subset of the population is intrinsically more interested in information A, and the complement in 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. This assumption captures in a parsimonious way the tendency of individuals to chat mainly about things they find interesting.

Despite being highly stylized, the model captures a basic trade-off that exists in casual information diffusion, both on- and offline. It predicts a number of observed phenomena. Among them is a general resilience of information that remains although crowding out can be significant. It also highlights the importance of relative, not absolute, interest in an information for its survival in a population. Furthermore, the introduction of information preferences allows us to analyze both the occurrence of *homophily*, and its impact on the prevalence of information. We find that its occurrence is influenced by various parameters of the diffusion process, and relative information interests. If the population indeed becomes segregated according to interests, information prevalence is reduced, and the population becomes polarized.

1.1 Related Literature

Within economics, the literature we are most closely related to is the network literature on diffusion that builds on the SIS framework, such as Jackson and Rogers (2007), López-Pintado (2008), or Jackson and Yariv (2010), but also Galeotti and Rogers (2013a), and Galeotti and Rogers (2013b). 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 processes of learning, best response dynamics, or explicit adoption decisions. These processes however differ significantly from the SIS model we employ. See, e.g., Jackson (2008) for an excellent introduction to the literature. While we share the basic methodology of the SIS framework with this literature, in all of the above papers the focus is on the diffusion of a unique state.

Diffusion of competing products or innovations instead has been analyzed in models of influence maximization, e.g., by Bharathi et al. (2007), Borodin et al. (2010), Dubey et al. (2006)

and Goyal and Kearns (2012). These models differ significantly from an SIS diffusion process, both with respect to the modeling characteristics, and the questions that they wish to answer. The above papers 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, which is unlikely to be the case for information. Consequently, while this literature also considers competing diffusion processes, the present results are of a complementary nature.

The impact that homophily has on information prevalence in our model is based entirely on agents choosing which information to pass on, which differentiates it from other work in that area, such as Golub and Jackson (2012) or Granovetter (1973). Indeed, our results on the impact of segregation on polarization are complementary to those of Baccara and Yariv (2008), Flaxman et al. (2013), Gentzkow and Shapiro (2010), Rosenblat and Mobius (2004), or Sunstein (2009) 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.

We are more closely related to contagion models that study the diffusion of multiple states, such as Beutel et al. (2012), Karrer and Newman (2011), Pathak et al. (2010) and Prakash et al. (2012). In contrast to the present paper, in these models infection with one virus/state provides full or partial immunity against the other. Such immunity introduces a tendency for the more virulent state to be the only one that survives in the population. These results are reminiscent of those of influence maximization processes. The fact that the present paper finds that information is resilient highlights the importance of the stage at which competition takes place.

Finally, both Myers and Leskovec (2012) and Weng et al. (2012) are close in spirit to our paper, as they aim to determine the interaction of multiple pieces of information, making use of Twitter data. In both papers, nodes are allowed to be infected with more than one information at a time. The competition among information in Weng et al. (2012) stems again from limited capabilities of nodes to store information. Myers and Leskovec (2012) in turn focus on estimating interaction patterns. While close in spirit, the framework and insights of these papers differ significantly from our own.

The rest of the paper is organized as follows. Section 2 presents the model and derives the prevalence of each information, while Section 3 quantifies crowding out and relates it to network characteristics. Section 4 investigates the impact of homophily and derives the conditions under which agents themselves wish to segregate according to information interests. Section 5 concludes.

2 The Model

2.1 Propagation Mechanism

We consider a population in which two pieces of information, A and B, are of interest to agents. It is worthwhile to take a moment to fix ideas about what A and B represent. The simplicity of our model allows us to be quite broad about this. They can be verifiable facts, opinions, or even unverifiable rumors. Any of these types of information appears to travel though chit-chat both on- and offline. We exclude only what might be termed "obvious lies", i.e., statements that are factually incorrect and directly verifiable by an agent.[2] Thus one can think of A as a piece of celebrity gossip and B as some political news. Else, A and B may both relate to the same topic, such as arguments for and against the severity of climate change. Yet again, they might be different ideological viewpoints on the same issue.

The population consists of an infinite number of agents, who represent nodes on a network. The links of the network denote meetings between agents. Following the SIS framework, we assume that this network is realized every period. We also assume that the network is regular, such that each agent meets k others at any time t, and we solve for the *meanfield approximation* of the system, assuming that t is continuous. The set of possible states in which an agent can be is $\{S, I_{A\setminus B}, I_{B\setminus A}, I_{AB}\}$. 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

[2] Such as "the sky is green".

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*.[3]

A central assumption is that at each meeting, at most one information can be communicated, as communication time is limited. We endow agents with preferences over the two pieces of 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.[4] Agents who prefer A are in group A and agents who prefer B are members of group 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. In particular, individuals are non-strategic in the way they pass on information, i.e., they neither distort the information they possess, nor do they strategically choose not to transmit an information.[5]

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,

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

[3]If agents never forgot, all information would eventually be known by everybody, which does not appear to be a relevant case. Much of the information that is transmitted as chit-chat is not immediately pay-off relevant, and often it is unknown whether it ever will be. Such information may be a prime target to be forgotten if memory limitations exist.

[4]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.

[5]For models of strategic information transmission on a network, see, e.g., Galeotti et al. (2013), Hagenbach and Koessler (2010), or recently Bloch et al. (2014).

This is (weakly) less than ρ_l, as not everybody aware of an information will necessarily 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 this rate approximates 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.[6]

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 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.[7]

[6] In essence, this assumption implies that at most one information is forgotten at any t. This assumption simplifies the analytical derivations, but also seems reasonable for the short time frames we approximate with this model. Finally, as for small ν infection with both information at t has zero probability, our assumption of a similarly small δ ensures that the setup is not exogenously biased against information survival.

[7] 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.

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 $l \in \{A, B\}$, and its characteristics. To analyze these, we adapt the following definition 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.[8]

Furthermore, we are interested in how the diffusion threshold and the prevalence of either information l compare to the case in which information l is 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 (2008) 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 $\lambda_A^* = \lambda_B^* = \lambda_d = 1$.*
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. 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

[8] López-Pintado (2008) actually defines a *critical threshold* above which a positive steady-state exists, and a *diffusion threshold* above which this positive steady-state is stable. In the present setting, these two thresholds coincide.

"catch" information from an informed node, i.e., $\lambda = R$. In the present model, $R_l \leq \lambda$, 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

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

Lemma 1. *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 Lemma 1 are derived in Appendix E.[9] 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 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[10] 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 opinions. Assume that $\lambda > 1$. Then any piece of information that is deemed the most interesting by some individuals, will survive on this network, no matter how much of a niche topic it might be. The Proposition also provides insights into the practice of "burying news". In the model, media coverage plays the role of planting the initial seed of information in the population. The prevalence of information is entirely independent of this. While news might be buried if they are released simultaneously with other major events, this is the result of differential interest in the population only, and independent of differences in news coverage. We now go beyond the qualitative results of information prevalence and quantify the extent to which two information crowd each other out.

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

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 information prevalence.

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 a piece of information as follows:

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

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 pieces of information diffusing simultaneously indeed causes them to (partially) crowd each other out.

Lemma 2. *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*

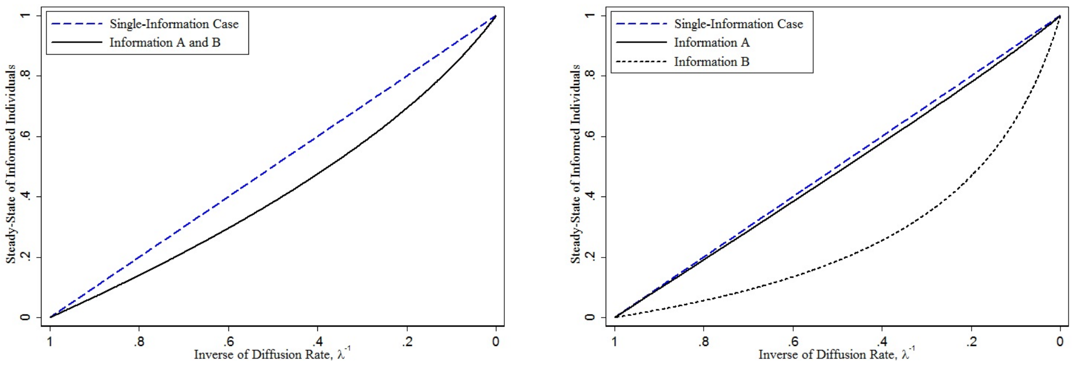
$$0 < \rho_B \leq \rho_A < \tilde{\rho}$$

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. □

Lemma 2 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 pieces of information demands that fewer people are aware of each of them. 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 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. 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%.[11] These are large differences, especially as we consider ordinal preferences over information pieces. 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.[12]

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

[12]This interpretation would differ if we had considered ν_l to be the likelihood with which each agent com-

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

Proof. See Appendix D. □

The first part of Proposition 2 establishes that the loss of information due to to crowding out is maximized for "intermediate" values of the diffusion rate. An increase in λ always increases the prevalence of any information. But if multiple information diffuse simultaneously, this increase is smaller than in the one-information case for $\lambda < \lambda_c$ and larger for $\lambda > \lambda_c$. *Relative* crowding out, in turn, depends only on relative interest. It is particularly interesting that differences in the sizes of groups A and B have magnified effects on crowding out. E.g., if group A is three times as large as B then the extent of crowding out of B is nine times as large as that of A. Finally, the prevalence growth rate as λ is increased is always larger for the information preferred by the minority. It might be argued that technological advances such as OSNs represent increases in the diffusion rate λ, as they might increase the number of interactions per period (k) or the likelihood that communication occurs at a meeting (ν). Proposition 2 predicts that if this is the

municates information l at a meeting. Under such an interpretation, the difference between ν_A and ν_B should be treated as cardinal. 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.

case, it will (i) increase overall information prevalence, (ii) have an ambiguous effect on crowding out of information, and (iii) will benefit disproportionally the prevalence of information that is of minority interest.

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*.[13]

In the present framework, homophily will determine the likelihood that an individual of group A meets an individual of group B. In particular, we will focus on the difference in information prevalence in an integrated society (groups A and B interact randomly with each other) as opposed to a segregated society (all interactions are within the same group). Our first result arises as a Corollary of Proposition 1.

Corollary 1. *Assume that the society is 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 agents are informed of at most one point of view. This occurs even if initial news consumption is entirely unbiased, and it is 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 in a segregated society than in an integrated one.

[13] One of the earliest work on this is Lazarsfeld et al. (1954). See also the survey by McPherson et al. (2001).

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 □

The loss of information due to segregation goes beyond the potentially polarizing impact on society of having no agent informed about both A and B in the long run. Indeed, there are many instances in which information A and B are entirely unrelated, in which case being informed about both might not have perceivable benefits. Nevertheless, even in such a case, segregation has an impact. Of course, whether or not the loss of information is considered harmful or beneficial for society depends entirely on the type of information.

A noteworthy aspect is the distinction between overall information prevalence and information prevalence among each group. If, e.g., A is a piece of celebrity gossip and B a piece of political news, the value that individuals in group A put on being informed about B (and *vice versa*) might be limited. That is, while overall information is lost due to segregation, it maximizes prevalence among those that value the respective information more. This leads us to question under which conditions agents themselves have incentives to segregate, which we address now.

4.2 Endogenous Segregation

To address the question of endogenous segregation, we need to impose some additional structure on the utility agents gain from being informed. To keep the analysis as simple as possible, we assume that agents derive utility directly from being informed about A and/or B. We assume that an agent in group l receives a flow utility of h while he is informed about l and a flow utility of s while he is informed about $-l$, where $h \geq s \geq 0$.[14] The prevalence of information l, ρ_l,

[14] 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

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.*

More generally, members of group l prefer a segregated society if

$$\frac{s}{h} < \frac{\tilde{\rho} - \rho_l}{\rho_{-l}}, \tag{14}$$

which leads to the following Proposition.

Proposition 4. *For any $\nu_l \in (0, 1)$ and finite $\lambda > 1$, a decrease in $\frac{s}{h}$ makes it more likely that a segregated society will emerge.*

Proof. Obvious given equation (14). □

I.e., the more extreme information preferences are, the more likely it is that we observe a segregated society. We state Proposition 4 as a likelihood that segregation occurs, as the exact value of $\frac{s}{h}$ at which individuals are indifferent between segregation and integration depends on the values of ρ_l and ρ_{-l}, which in turn depend on λ and ν_l.

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

Proof. See Appendix F. □

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 whose pay-off depends on the state. In either case, the expected utility of an agent would be increasing in the amount of time he is informed, which is captured in our simple utility function.

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 the 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. Arguably, they have made it easier for individuals to find others that share similar interests. Hence, they might offer the possibility to segregate that did not exist in the offline world. Given our results, we would expect that groups that make use of this possibility are those that (i) are particularly interested in niche or very specialized pieces of information, and (ii) are extreme in their valuation of information, i.e., they care little about information apart from their preferred one. On the other hand, if their rise indeed represents an improvement in the diffusion rate (an increase in λ), this will translate into lower incentives to segregate for any group.

5 Conclusion

The present paper models the diffusion of two pieces of information under communication constraints in a simple and standard diffusion process. Individuals may be aware of one or both pieces of information, but if they are aware of both, they will communicate the piece that they find more interesting. The type of information diffusion that is best captured by our model is casual chit-chat, which most individuals engage in on a daily basis.

Our results contribute to the understanding of how communication constraints affect the diffusion of information. They are able to shed light on the mechanism that underlies the possibility of "burying" information, and to explain why so many pieces of information, many very obscure, seem to survive. The model also offers strong predictions on the impact of the Internet and OSNs on information diffusion: If they represent increases in the diffusion rate, they will be particularly beneficial for "niche" information, and decrease the likelihood that individuals segregate according to their information interests.

Finally, our model predicts that segregation according to interests will unambiguously lead to polarization, and a loss of information. If offered, segregation will become more likely as information preferences become more extreme, and will be chosen predominantly by individuals whose interest lies in niche topics.

To the best of our knowledge, communication costs of any type have not before been analyzed in an SIS framework. Indeed, we have kept our model deliberately simple to highlight the strong implications that opportunity costs have for its predictions. Polarization due to segregation in particular has often been analyzed. Our model highlights a complementary channel in this process that has not received much attention, if any. That is, even without biased messages, or news consumption, information that is of relatively little interest has no chance of surviving in a segregated group. As a consequence, campaigns to introduce "competing" information into segregated groups will not have any long term impact, only a reduction of segregation will.

A richer framework could allow for a more general interaction structure of individuals, or indeed a fixed network. Although the second route promises to be interesting, previous work in this area has shown that it is a problem of substantial complexity. Another promising area is the question of how individuals choose which information to communicate. While we believe our assumption to link this to intrinsic preferences is a valid starting point, there are numerous other factors that might contribute to this decision. It might, e.g., depend on how likely it is that the information is "new" information to the other party. This is an area that we are currently working on, but that we believe goes beyond the scope of the present paper.

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,

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

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

and

$$H'(0) = \lambda\nu_A + \nu_B, \tag{25}$$

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

$$\theta_A = \frac{\lambda\nu_A\theta_A}{1+\lambda\theta_A} + \frac{\nu_B\theta_A}{1+\nu_B\lambda\theta_A}. \tag{26}$$

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

$$\begin{aligned}
(1+\lambda\theta_A)(1+\lambda\nu_B\theta_A) &= \lambda\nu_A(1+\lambda\nu_B\theta_A) + \nu_B(1+\lambda\theta_A) \\
1+\lambda\theta_A(1+\nu_B) + \lambda^2\nu_B\theta_A^2 &= \lambda\nu_A + \lambda^2\nu_A\nu_B\theta_A + \nu_B + \lambda\nu_B\theta_A \\
\nu_A(1-\lambda) + \lambda^2\nu_B\theta_A^2 + \lambda\theta_A(1-\lambda\nu_A\nu_B) &= 0
\end{aligned}$$

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

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

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 2, 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 2, 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 2, 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

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

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

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

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,

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

QED.

References

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

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

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

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

Bloch, F., G. Demange and R. Kranton, "Rumors and Social Networks," (2014).

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

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

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

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

Galeotti, A., C. Ghiglino and F. Squintani, "Strategic information transmission networks," *Journal of Economic Theory* 148 (2013), 1751–1769.

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

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

Gentzkow, M. and J. M. Shapiro, "Ideological segregation online and offline," Technical Report, National Bureau of Economic Research, 2010.

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

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

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

Hagenbach, J. and F. Koessler, "Strategic communication networks," *The Review of Economic Studies* 77 (2010), 1072–1099.

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

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

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

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

Karrer, B. and M. Newman, “Competing epidemics on complex networks,” *Physical Review E* 84 (2011), 036106.

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

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

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

Leskovec, J., L. Backstrom and J. Kleinberg, “Meme-tracking and the dynamics of the news cycle,” in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM, 2009), 497–506.

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

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

Myers, S. A. and J. Leskovec, “Clash of the Contagions: Cooperation and Competition in Information Diffusion.,” in *ICDM* volume 12 (Citeseer, 2012), 539–548.

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

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

———, “Immunization of complex networks,” *Phys. Rev. E* 65 (Feb 2002), 036104.

Pathak, N., A. Banerjee and J. Srivastava, “A generalized linear threshold model for multiple cascades,” in *Data Mining (ICDM), 2010 IEEE 10th International Conference on* (IEEE, 2010), 965–970.

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

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

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

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

Weng, L., A. Flammini, A. Vespignani and F. Menczer, “Competition among memes in a world with limited attention,” *Scientific Reports* 2 (2012).