Spreading Information via Social Networks: An Irrelevance Result*

Yu Awaya† and Vijay Krishna‡

February 5, 2024

Abstract

An informed planner wishes to spread information among a group of agents in order to induce efficient coordination—say the adoption of a new technology with positive externalities. The agents are connected via a social network. The planner informs a seed and then the information spreads via the network. While the structure of the network affects the rate of diffusion, we show that the rate of adoption is the *same* for all *acyclic* networks.

1 Introduction

Policymakers often wish to inform the public about various policies and technologies—a tax credit, a new seed variety, a new digital payment system—so that the public will avail of or adopt these. The simplest way to disseminate such information is to just broadcast a public service announcement (PSA) on television, radio and other media. But there have been some doubts about the effectiveness of PSAs. For various reasons, people may pay more attention to information coming from friends and neighbors rather than mass media.[1] Thus in many circumstances it is better to "seed" the information to a few individuals and then let it spread naturally via the existing social network. Of course, how quickly information diffuses depends on the network. At one extreme, information will spread very quickly in a "star" network—where one individual, say 1, is directly connected to all others who are directly connected only to 1. At the other extreme, it will spread very slowly in a "line" network—where

*The research reported here was supported by a grant from the National Science Foundation (SES-2048806). We thank Nageeb Ali and Rahul Krishna for helpful discussions.

†Department of Economics, University of Rochester, E-mail: yuawaya@gmail.com.

‡Department of Economics, Penn State University, E-mail: vkrishna@psu.edu.

[1] Banerjee, Breza, Chandrasekhar and Golub (2023) observe this in the "field" and provide some behavioral explanations why this might be the case.

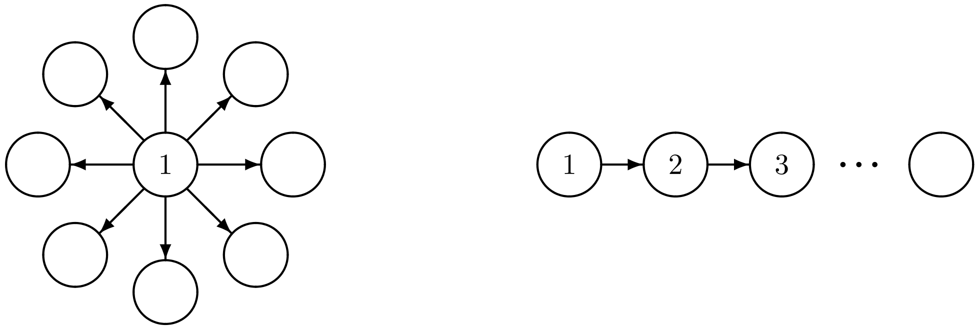

Figure 1: Star and Line Networks

individual 1 is connected only to individual 2, who is connected only to one other, say 3, etc. See Figure 1.

In this paper, we ask a different question. Suppose that the information concerns the benefits of a new technology—say, a new digital payment system—and the policymaker wishes to get the public to adopt the system. In many such situations, there are significant positive externalities—adopting a new digital payment system is useful only if other people do so as well.[2] Instead of asking how the network structure affects the rate of diffusion, we ask how it affects the rate of *adoption*.

Our main finding is that in the class of *acyclic* networks, the structure of the network and how it is seeded is irrelevant—the rate of adoption is the *same* for all such networks. In particular, the adoption rate when information diffuses quickly via the star network is the same as when it diffuses slowly via the line network. So while the structure of the network affects both the speed at which information is diffused and, as we will see, its quality, it does not affect the prospects of efficient coordinated behavior.

To illustrate our findings, we begin with an example.

1.1 Example

A new technology is of uncertain value—it may or may not be useful/viable. There are three agents who must simultaneously decide whether or not to adopt the new technology at a cost $c < 1$ per person. The gross payoff to an agent is \$1 if and only if the technology turns out to be useful and *all* three agents adopt the technology; otherwise, the gross payoff is zero. Thus, adoption has positive externalities. Let $\rho \in (0, 1)$ be the prior probability that the technology is useful.

A planner, agent 0, knows whether or not the technology is useful, and if it is, sends a message to the agents. If it is not useful, no message is sent. Thus, anyone who gets the message is sure that the technology is useful.

[2]See Crouzet, Gupta and Mezzanotti (2023) for evidence of such extenalities in the adoption of a digital wallet following the Indian demonetization in 2016.

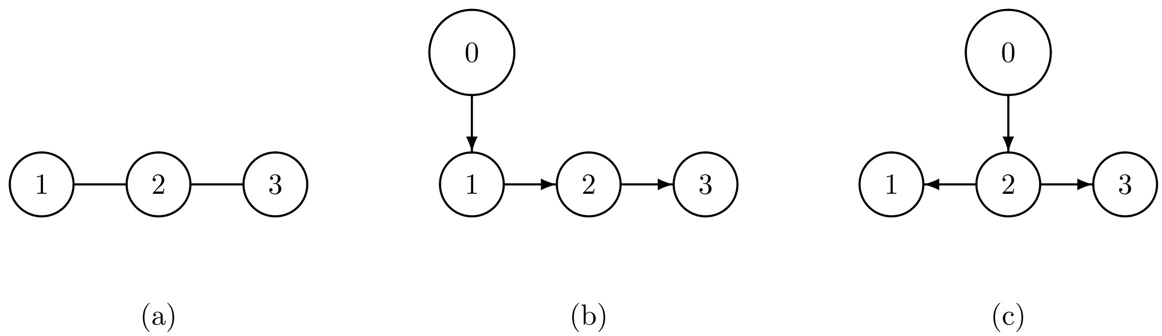

Figure 2: Seeding the Line Network

The three agents are part of single connected social network. Specifically, they are arranged along a line as in Figure 2 (a). Agents can receive messages from and pass these along to their neighbors. Message transmission is imperfect, however—at every stage there is a small probability $\varepsilon > 0$ that a message that is sent to a neighbor is lost. Thus, if the planner sends a message to 1, there is only a probability $1 - \varepsilon$ that 1 will in fact get the message. If 1 receives the message and sends it on to 2, then there is only a probability $1 - \varepsilon$ that 2 will get the message and so on. Transmission losses occur independently across links.

A1. Seeding the network via 1. Here, if the technology is useful, the planner sends a message to 1, which if received, is sent to 2, which if received, is then sent to 3 (this is depicted by the arrows in Figure 2 (b)). We claim that if the cost $c \leq (1 - \varepsilon)^2$, then every agent who is informed—gets a message—adopts.[3] And if $c > (1 - \varepsilon)^2$, then no agent, informed or not, adopts.

Case 1: $c \leq (1 - \varepsilon)^2$. First, consider agent 1 and suppose the others adopt if informed. If 1 gets a message, then she knows that the technology is useful and so her only worry is whether all other agents got the message as well. Since messages are only passed along the line, the probability that agents 2 and then 3 are also informed, and so will adopt, is just $(1 - \varepsilon)^2$ which is greater than the cost.[4] Thus, if informed, it is optimal for agent 1 to adopt.

Next, consider agent 2 and as above, suppose the others adopt if informed. If 2 gets a message, then she knows that 1 also got the message and that 3 got the message with probability $1 - \varepsilon$. Thus, it is optimal for agent 2 to adopt as long as $c \leq 1 - \varepsilon$, a weaker requirement than that for agent 1.

[3] For our purposes it is not necessary to specify the exact strategy—that is, what an agent does if she does not get the message. A detailed specification of the strategies is in Section 4.

[4] Since the gross payoff is 1 if the technology is useful and everyone adopts, and 0 otherwise, this probability is also the gross expected payoff.

Finally, consider agent 3 and again suppose others adopt if informed. If 3 gets a message, then she knows for sure that 1 and 2 also got the message and so she is willing to adopt for all $c \leq 1$.

Thus, if $c \leq (1-\varepsilon)^2$ there is an equilibrium in which every agent adopts if she gets the message. The probability that all agents adopt the technology when it is useful is just the probability that the message reaches 3, that is, $(1-\varepsilon)^3$.

Case 2: $c > (1-\varepsilon)^2$. In this case, the *unique* equilibrium is one in which no agent ever adopts.

To see why, note that if 1 is uninformed, the probability that she assigns to the event that the technology is useful is small—it is of order ε. This is because the only way this can happen is if the message from the planner to 1 was lost, an ε probability event. When ε is small, this probability is smaller than the cost. This means that it is dominated for an uninformed agent 1 to adopt.

Now from the argument above, 2 knows that 1 will not adopt if uninformed. If 2 does not get a message, her belief that 1 is informed is also of order ε because the event that 1 is informed while 2 is not can occur only if the message from 1 to 2 was lost, again an ε probability event. This means that it is (iteratively) dominated for an uninformed 2 to adopt.

Now 3 knows that 1 and 2 will not adopt if uninformed. If 3 does not get a message, then for similar reasons as above, her belief that both 1 and 2 got a message is again of order ε. So it is (iteratively) dominated for an uninformed 3 to adopt.

Thus we have argued that it is *iteratively* dominated for every agent to adopt if she does not get a message.

Now suppose agent 1 is informed. At best, the other agents will adopt only if informed and the chance of this is $(1-\varepsilon)^2$ and since c exceeds this, it is optimal for 1 to not adopt even when informed. Thus, agent 1 will never adopt.

But now if agent 1 never adopts, it is optimal for other agents to never adopt as well.

A2. Seeding the network via 2. Seeding the network via 1 seems inefficient since the information has to travel from 1 to 2 and then from 2 to 3. Suppose instead that the planner sends a message to the agent who is "central," that is, 2. This message, if received by 2, is forwarded to 1 and 3 simultaneously (this is depicted by the arrows in Figure 2 (c)). We claim that even though this seems like a better way to disseminate information, the prospects for efficient coordination are the *same* as when 1 is the seed. Again, if $c \leq (1-\varepsilon)^2$, then every informed agent adopts. And if $c > (1-\varepsilon)^2$, then no agent ever adopts.

Case 1: $c \leq (1-\varepsilon)^2$. Here an informed agent 2's belief that the others will get the message is just $(1-\varepsilon)^2$ which is greater than the cost. Thus, upon getting a message it is optimal for agent 2 to adopt if the others are doing so.

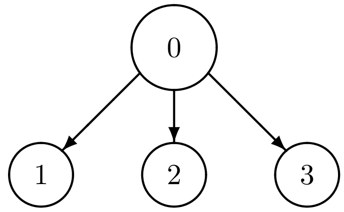

Figure 3: Broadcast

The same calculation for 1 shows that 1 is willing to adopt as long as $c \leq 1 - \varepsilon$. This is because 1 knows that 2 is informed for sure and that 3 is informed with probability $1 - \varepsilon$. Thus, it is optimal for an informed agent 1 to adopt as long as $c \leq 1 - \varepsilon$, a weaker requirement than for 2. The same is true for agent 3 since 1 and 3 are symmetrically placed.

Thus, it is an equilibrium for every agent to adopt if she is informed. When the technology is useful, the probability that the message reaches all the agents is just $(1 - \varepsilon)^3$, the same as in scenario A1.

Case 2: $c > (1 - \varepsilon)^2$. Again, the unique equilibrium is one in which no agent ever adopts.

To see why, note that if 2 does not get a message, the probability that she assigns to the event that the technology is useful is again of order ε (it is the same as that assigned by 1 when 1 was the seed in Scenario A1). Since c is greater than this probability, this implies that it is dominated for an informed 2 to adopt.

Now 1 knows that 2 will not adopt if uninformed. As above, if 1 does not get a message, her belief that 2 got a message is also of order ε. This belief is again smaller than the cost. This means that it is (iteratively) dominated for an uninformed 1 to adopt. As before, the same is true for 3 since 1 and 3 are symmetrically placed.

Now suppose agent 2 gets a message. The chance that the other agents will adopt is at most $(1 - \varepsilon)^2$ and since c exceeds this, it is optimal for 2 to not adopt even when she gets a message. Thus, agent 2 will not adopt whether or not she is informed.

But now if agent 2 never adopts, it is optimal for 1 and 3 to never adopt as well.

Thus, we see that in this example, seeding information to a more "central" agent—with more neighbors—does not improve the prospects for coordination.

B. Broadcasting. An alternative is to bypass the social network entirely and "broadcast" the message—it is sent privately to all agents simultaneously (as in Figure 3). Again, with probability ε, the message to any agent i is lost and so not heard by the agent. Lost messages occur independently across agents, each with probability $\varepsilon > 0$.

It seems intuitive that directly broadcasting is a better method of dissemination than letting the information trickle from agent to agent. In particular, with broadcasting the probability that any agent gets the information is $1-\varepsilon$ and so is the same for all agents. In contrast, when the network is seeded via 1, the probability that agent 2 gets the message is $(1-\varepsilon)^2$ and the probability that 3 gets the message is $(1-\varepsilon)^3$.

We will show that even though broadcasting provides better information about the usefulness of the technology, when it comes to engendering efficient coordination, it is *equivalent* to either of the indirect methods of dissemination A1 and A2.

Case 1: $c \leq (1-\varepsilon)^2$. Again, in this case there is an equilibrium in which every agent who gets the message adopts the technology. This follows from the fact that for any informed agent i, the probability that the other two agents are also informed is $(1-\varepsilon)^2$.

When the technology is useful, the probability that all agents get the message is just $(1-\varepsilon)^3$, the *same* as that when the network is seeded via 1 or 2.

Case 2: $c > (1-\varepsilon)^2$. In this case, again the unique equilibrium is again one in which no one ever adopts. If agent i does not get a message, her belief that the technology is useful is again of order ε which is less than c. This means it is dominated for an uninformed agent to adopt. If agent i does get a message, the probability that the other two agents will adopt is at most $(1-\varepsilon)^2$ and since $c > (1-\varepsilon)^2$, agent i will not adopt even if informed. This means that the unique equilibrium is for all agents to never adopt.

Thus, in this example we see that how information is disseminated—indirectly via seeding the network or directly sending it to each agent—does not affect the prospects for coordination.

In this paper we show that there is nothing special about the example. As in the example, we study information dissemination in social networks without cycles—that is, trees. Informally stated, our main result is:[5]

The prospects of efficient coordination are the same no matter how information is disseminated—it is independent both of the structure of the tree and how it is seeded.

Our main result says that if the goal of the planner is to induce efficient coordinated action—adopting a new technology or product—then the tree structure is irrelevant. Why is this? Efficient coordination requires not only that agents be informed about the fundamental uncertainty—whether or not the technology is useful—but also be informed whether other agents are informed. It is easy to see that one tree network may be better than another in conveying information about fundamentals. For

[5] The formal results (Theorem 1 below) applies not only to a single tree but also to networks which are collections of disjoint trees—*forests*.

instance, in the example above, let us compare broadcasting (Scenario B) to seeding 1 and then letting the information trickle from 1 to 2 to 3 (Scenario A1). Broadcasting is better than seeding in that every agent has better information about the fundamentals with broadcasting than seeding. But seeding is better than broadcasting in revealing whether others are informed—for instance, if agent 3 gets a message, she knows that 1 and 2 did so as well. [6]

Rather than consider a particular game—like the adoption game in the example—we derive and phrase our results by using the language of approximate common knowledge. Thus our main result says that the extent of approximate common knowledge is independent of the structure of the tree network and how it is seeded. The close connection between approximate common knowledge and equilibrium behavior in games is well-known (see Monderer and Samet, 1989, Kajii and Morris, 1997 and Oyama and Takahashi, 2020).

1.2 Related literature

The question of diffusion in social networks appears in many contexts—infectious diseases, product awareness, plans for a revolt, etc. In most of these situations, the planner is interested in the affecting the speed and extent of diffusion—either slowing it down in the case of disease or speeding it up in the other cases. Recently, the question has drawn the attention of development economists who are interested in conveying information about various policy initiatives and has been studied in various contexts—microfinance (Banerjee et al. 2013), immunizations (Banerjee, Chandrasekhar, Duflo and Jackson, 2019), planting techniques (Beaman, Ben Yishay, Magruder and Mobarak, 2021), demonetization (Banerjee et al. 2023)—by conducting randomized controlled trials (RCTs). But these issues are not confined to developing countries. Chetty, Friedman and Saez (2013) find that whether or not people optimally avail of the earned income tax credit (EITC)—a large US government transfer program—depends on the neighborhood they live in. In other words, information about the EITC spreads via a network.

One of the findings of this line of research is that spreading the information via existing social networks may be superior to "broadcasting" the information via media (Banerjee et al. 2023).

One is then naturally led to the question of how best to "seed" the information by conveying it to a few key agents who spread it via the social network. Clearly, if one is interested in spreading the information quickly and widely, the information should be seeded via agents that are well-connected—that is, central players. But identifying who is central is daunting task in any reasonable sized network. First, one has to determine the network—a difficult task itself—and second, to find the central players in the network. The latter problem is known to be computationally hard.

[6] A precise result along these lines is Proposition 3.4 below.

In a very interesting paper, Akbarpour, Malladi and Saberi (2023) have argued that instead of finding the optimal seed, it is better to choose multiple seeds randomly. The argument is that even if the information is seeded to non-central agents it will find its way to those that are central anyway—by definition, the central players are well-connected.

In all of this work, the focus is on the speed at which information spreads as well as the extent of diffusion. Implicit in this is the assumption that once an agent is informed, he/she will automatically adopt the new technology or avail of the policy initiative. This may be the case if the costs and benefits of adoption do not depend on whether others are doing so as well. But many new technologies/products are subject to complementarities in adoption/consumption—that is, network externalities. Crouzet et al. (2023) document the presence of such externalities in the adoption of a new digital payment platform in India. Naturally, adopting a digital payment platform is useful only if others adopt it as well. Such externalities are also a key feature of the theoretical diffusion model of Sadler (2020).

Our paper departs from the focus on the speed of diffusion. Rather, we are interested in the likelihood that the new technology—subject to network externalities—will be adopted once the information has spread. Put another way, to what extent will the public be able to coordinate adoption? As is well-known, efficient coordination requires that people know not only whether or not the digital wallet works and is safe—known as first-order uncertainty—but whether others know this as well and whether others know that others know, etc.—higher-order uncertainty. The importance of considering higher-order uncertainty is the main lesson of Rubinstein's (1989) E-mail game who shows that it can be a major cause of coordination failure.[7]

Field experiments by Gottlieb (2016) point to the importance of reducing higher-order uncertainy in elections in Mali. An interesting finding along the same lines is reported by George, Gupta and Neggers (2019). They measure how broadcasting—via text messages—the criminal records of candidates affects vote shares in Indian elections. In one treatment of an RCT, the texts just report the criminal records of the candidates and the effect on vote shares is negligible. In a second treatment, the texts not only report the criminal records but also say that others are also receiving the same text. Now the effect on vote shares is measurable—there is roughly a 3% shift in votes. Thus, it seems that people are willing to change their voting behavior only if they think that others will do so as well. In other words, reducing higher-order uncertainty matters.

The change in focus away from the speed of diffusion to efficient coordination is the key to our result. Once efficient coordination is the goal, we find that in acyclic networks, the network structure and who is the seed becomes irrelevant. In such

[7] Coles and Shorrer (2012) show that the extreme coordination failure in Rubinstein's two-player E-mail game can be mitigated in multi-player games where communication takes place in a hub-and-spoke network. In De Jaegher (2015) higher-order information is directly communicated to the agents.

networks, there is no need to ascertain the exact structure or the optimal person to choose as the seed.

Organization of the Paper The remainder of the paper is organized as follows. The next section outlines the model as well as the terminology of approximate common knowledge. Section 3 then derives the main result that in all acyclic networks the extent of approximate common knowledge is the same. Section 4 studies the technology adoption game from the Introduction, as well as other games, and shows that the approximate common knowledge results of Section 3 have exact counterparts concerning equilibria of these games. Our results rely on the assumptions that (a) the network is acyclic; and (b) each tree in the network has a single seed. In Section 5 we show that the irrelevance result does not extend if these assumptions are relaxed. We show by example that there are circumstances in which a cycle can make the situation worse in terms of the adoption rate. Also, there are circumstances in which a single seed is better than multiple seeds. Finally, we also consider the possibility of randomly chosen seeds. We show that choosing seeds at random can, in some cases, improve the situation. Appendices A and B contain some auxiliary results.

2 Model

There is an uncertain state of nature $\theta = g$ or b with prior probabilities $\rho \in (0,1)$ and $1-\rho$, respectively. A planner who knows θ wishes to convey this information to a set of agents $\mathcal{I} = \{1, 2, ..., I\}$—the public. The planner will be labeled as agent 0.

The agents in $\mathcal{I}$ constitute the nodes of a social network which is either a *tree* T—an undirected connected graph without cycles—or a disjoint union $T^1 \cup T^2 \cup ... \cup T^R$ of trees, that is, a *forest*. Let $F = \left(T^1, T^2, ..., T^R\right)$ denote the forest.[8]

In state g, the planner sends a private message to a *single* node s^r in each tree T^r—the *seed* of T^r. Let $s = \left(s^1, s^2, ..., s^R\right)$ denote a *seeding* of the forest. The message then spreads through each tree as follows.

Fix a particular tree, say T^1, and let agent 1 be the seed. If the seed gets a message, she forwards it to each i in the set of her neighbors, denoted by $\mathcal{N}(1)$. Each of 1's neighbors $i \in \mathcal{N}(1)$ then forwards the message to each of his neighbors $j \in \mathcal{N}(i)$ *except* 1. Each $j \in \mathcal{N}(i) \setminus \{1\}$ then forwards the message to each of her neighbors in $\mathcal{N}(j)$ except i and so on.

In this manner, information spreads throughout the tree. Notice that because (a) there are no cycles in the underlying undirected network; and (b) no agent sends a message back to the person she received a message from, it is the case that now the

[8]For formal definitions of these and other terms, we refer the reader to the excellent book by Jackson (2008).

tree becomes *directed*—there is single direction of flow of information from seed to all other nodes. Formally, for every node i, there is a unique agent that immediately precedes i in the tree and is the only source of information for i. Given a forest F and a seeding s, let $\mathcal{T}(F, s)$ denote the resulting directed tree with the planner, agent 0, as the root. We will refer to $\mathcal{T}$ as the (directed) *information tree*.

The top panel of Figure 4 depicts a forest consisting of two undirected trees. The other two panels show how the choice of different seeds results in different directed trees. In each case, the arrows depict the flow of information.

Messages can be lost, however. If i forwards a message to her neighbor j, then there is a probability $\varepsilon > 0$ that the message is lost and not received by j. Conditional on i being informed, the losses of i's messages to her neighbors are independent. Thus, if i sends messages to her neighbors j and k, then the probability that both will receive the message is $(1-\varepsilon)^2$. The same is true for messages from the planner to a seed—the probability that in state g, the seed receives the message is also $1-\varepsilon$.

Note that if the message from i to j is lost, then j cannot forward it to anyone and the flow of information to all the nodes that succeed j stops.

In state b, no messages are sent by the planner and so there is no flow of information.

Note that if i receives a message, then he knows for sure that (1) the state of nature is g; and (2) all agents j along the unique path from the seed to his immediate predecessor also received a message.

Common beliefs Let $x_i \in \{y, n\}$ denote the information available to $i \in \mathcal{I}$, where $x_i = y$ ("yes") denotes that i received a message and $x_i = n$ ("no") denotes that i did not.

The set of *states of the world* is

$$\Omega = \{g, b\} \times \{y, n\}^I$$

A state of the world $\omega \in \Omega$, then determines both the fundamental state, g or b, as well as which of the agents are informed. For instance, the state of the world (g, y, y, n) is one where the fundamental is g, agents 1 and 2 receive the message while agent 3 does not.

Different network structures and seedings lead to different probability distributions on Ω. For instance, if three agents are arranged in a line as in Figure 2, then state (g, n, y, y) is impossible if 1 is the seed, while it has a positive probability of occurring if 2 is the seed. In what follows, we will fix the forest $F = \left(T^1, T^2, ..., T^R\right)$ and a seeding $s = \left(s^1, s^2, ..., s^R\right)$ so also the resulting information tree $\mathcal{T}$. All probabilities will be calculated using the resulting probability distribution $\mathbb{P}_{\mathcal{T}}$ over Ω. To avoid notational clutter, in what follows, we will temporarily suppress the dependence of $\mathbb{P}_{\mathcal{T}}$ and other objects on $\mathcal{T}$.

Following Monderer and Samet (1989), given any event $E \subseteq \Omega$ and probability p, the event $B_i^p(E) \subseteq \Omega$ consists of states of the world ω in which E is *p-believed* by i

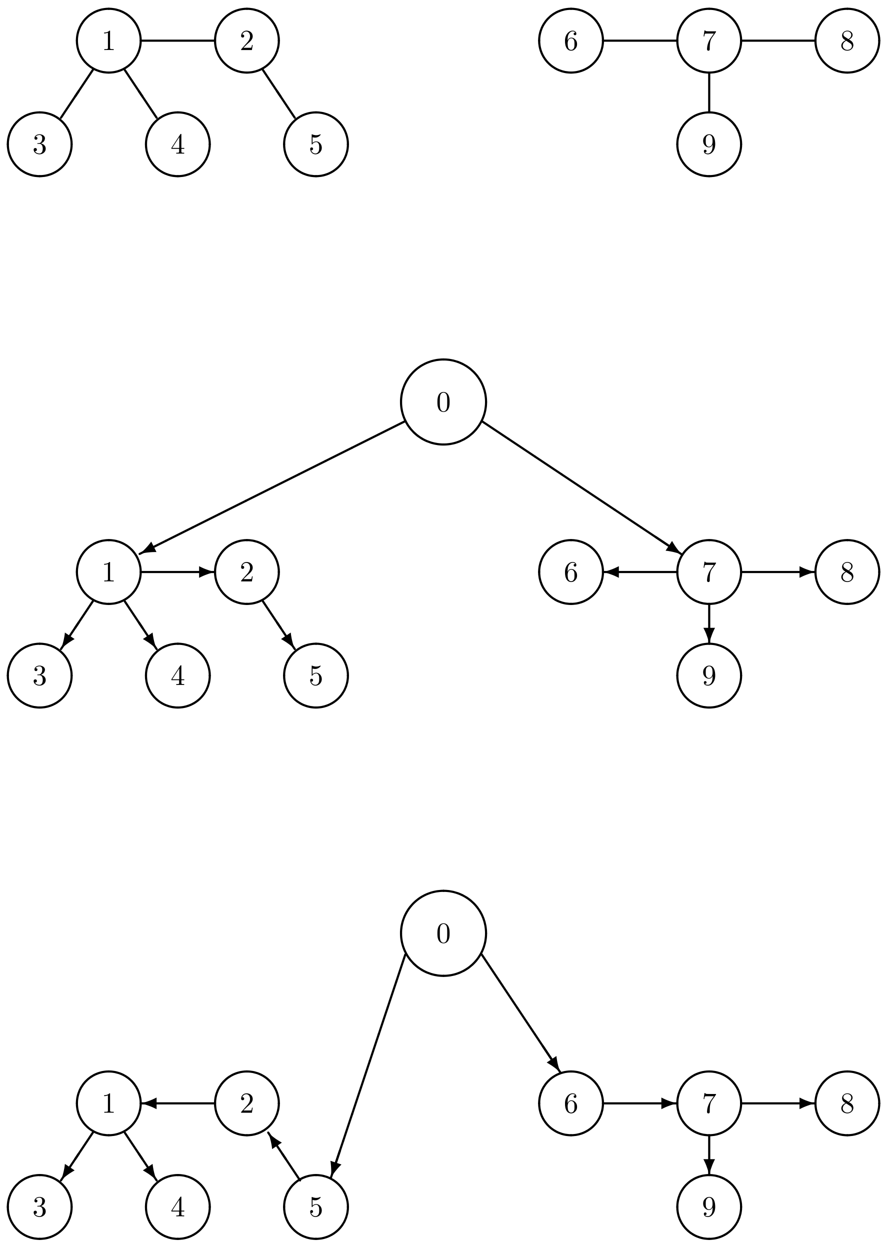

Figure 4: Undirected Forest to Directed Tree by Seeding

given the information $x_i(\omega) \in \{y, n\}$ available to her in state ω. Formally,

$$B_i^p(E) = \{\omega \in \Omega : \mathbb{P}[E \mid X_i = x_i(\omega)] \geq p\}$$

In other words, in any state $\omega \in B_i^p(E)$, i assigns probability exceeding p to the event E given her information $x_i(\omega)$. We write

$$B^p(E) = \cap_{i\in\mathcal{I}} B_i^p(E)$$

as the set of states in which E is p-believed by all the agents.

Now for $\ell = 1, 2, \ldots$ define $B^{p,\ell}$ recursively by

$$B^{p,\ell}(E) = B^p\left(B^{p,\ell-1}(E)\right)$$

where $B^{p,0}(E) = E$ and finally,

$$C^p(E) = \cap_{\ell\geq 1} B^{p,\ell}(E)$$

Thus, $C^p(E)$ is the set of states of the world in which E is *common p-believed*. In other words, (i) everyone assigns probability exceeding p to the event E, and also (ii) assigns probability exceeding p to the event that everyone assigns probability exceeding p to the event E, and also (iii) assigns probability exceeding p to the event that everyone assigns probability exceeding p to the event that everyone assigns probability exceeding p to the event E, and so on.

Note that B_i^p is a monotone mapping, that is, $E \subseteq E'$ implies that $B_i^p(E) \subseteq B_i^p(E')$. The same is then true of $B^{p,\ell}$ and C^p. Also, if for some ℓ it is the case that $B^{p,\ell+1}(E) = B^{p,\ell}(E)$, then $C^p(E) = B^{p,\ell}(E)$. Thus, $C^p(E)$ is a fixed point of B^p.

When $p = 1$, $C^1(E)$ is the set of states in which the event E is commonly known. For p close to 1, $C^p(E)$ is the set of states in which E is approximately commonly known.

We emphasize once again that since the probability distribution $\mathbb{P}$ over states depends on the underlying information tree $\mathcal{T}$, the sets $B_i^p(E)$, $B^p(E)$ and $C^p(E)$ also depend on $\mathcal{T}$. Later when we want to make this dependence explicit, we will write $\mathbb{P}_{\mathcal{T}}$ and $C_{\mathcal{T}}^p(E)$, for instance.

Notation Let

$$G = \{\omega \in \Omega : \theta = g\}$$

be the set of states of the world with $\theta = g$ and let

$$Y_i = \{\omega \in \Omega : x_i(\omega) = y\}$$

consist of states in which i is informed (gets a message). Since messages are sent only in state g, $Y_i \subset G$. Finally, let

$$Y^* = \cap_{i\in\mathcal{I}} Y_i$$

be the set of states in which every agent is informed. Since y is conclusive evidence that the $\theta = g$, in fact, Y^* consists of a single state $\omega^* = (g, y, y, \ldots, y)$.

3 Irrelevance of structure

Rather than considering a specific game, say the technology adoption game from the Introduction, we begin by showing that the set of common p-beliefs does not depend on the network or its seeding. It is well-known that the degree of approximate common knowledge is a fundamental determinant of equilibrium behavior in incomplete information games (Kajii and Morris, 1997).

Our main result is that the extent of approximate common knowledge is independent of the underlying information tree $\mathcal{T} = (F, s)$. It depends only on the number of agents I, the prior ρ and the error probability ε.

Theorem 1 *For any p, the event $C_{\mathcal{T}}^{p}(G)$ in which G is common p-believed does not depend on the information tree $\mathcal{T}$. Moreover, the probability $\mathbb{P}_{\mathcal{T}}[C_{\mathcal{T}}^{p}(G)]$ does not depend on $\mathcal{T}$ either.*

3.1 Proof of Theorem 1

The proof of Theorem 1 is divided into two parts—when the error probability ε is small and when it is large.

Let agent 1 be a seed of some tree in the forest and note that the probability of G given that 1 is uninformed is

$$\Pr[G \mid N_1] = \frac{\rho\varepsilon}{1 - \rho(1-\varepsilon)} \tag{1}$$

From Lemma A.3, the probability that all other agents are informed given that 1 is informed is

$$\Pr[Y^* \mid Y_1] = (1-\varepsilon)^{I-1} \tag{2}$$

Note that these probabilities are the same for *any* seed of *any* tree in the forest since all seeds receive information directly from the planner. Thus we simply write Pr to denote these rather than $\mathbb{P}_{\mathcal{T}}$.

Let $\bar{\varepsilon}$ be the unique value of ε that equates $\Pr[G \mid N_1]$ and $\Pr[Y^* \mid Y_1]$. Such a value exists and is unique since $\Pr[G \mid N_1]$ is an increasing function of ε while $\Pr[Y^* \mid Y_1]$ is a decreasing function.

3.1.1 Small ε

When $\varepsilon < \bar{\varepsilon}$, it is the case that

$$\Pr[G \mid N_1] = \frac{\rho\varepsilon}{1 - \rho(1-\varepsilon)} < (1-\varepsilon)^{I-1} = \Pr[Y^* \mid Y_1] \tag{3}$$

We then have

Proposition 3.1 *If* $0 < \varepsilon < \overline{\varepsilon}$, *then for any information tree* $\mathcal{T}$,

$$C^p_{\mathcal{T}}(G) = \begin{cases} \Omega & \text{if } p \leq \Pr[G \mid N_1] \\ Y^* & \text{if } \Pr[G \mid N_1] < p \leq \Pr[Y^* \mid Y_1] \\ \varnothing & \text{if } p > \Pr[Y^* \mid Y_1] \end{cases}$$

Proof We consider each range of p's separately.

Case 1: $p \leq \Pr[G \mid N_1]$. In this case, p is so low that even an uninformed agent 1 assigns greater probability than p to G.

Lemma A.1 now implies that for any agent i in the forest, seed or not,

$$\Pr[G \mid N_1] \leq \Pr[G \mid N_i]$$

and so for all i, $p \leq \Pr[G \mid N_i]$. This means that every agent, informed or not, assigns a probability of at least p to G. Formally,

$$B^p_i(G) = Y_i \cup N_i = \Omega$$

and since $B^p(G) = \cap_{i \in \mathcal{I}} B^p_i(G)$,

$$B^p(G) = \Omega$$

But since everyone assigns probability 1 to Ω, it follows that $C^p(G) = \Omega$.

Case 2: $\Pr[G \mid N_1] < p \leq \Pr[Y^* \mid Y_1]$. This case is broken up into two steps.

Step 1: $\Pr[G \mid N_1] < p$ implies that $C^p(G) \subseteq Y^*$.

To show this step we will argue that for any agent k, $C^p(G) \cap N_k = \varnothing$. In other words, the event that G is common p-believed cannot include any state in which an agent is uninformed.

Consider the unique path from 0 to k and suppose (after renaming, if necessary) that this path consists of agents $1, 2, ..., k$ such that the direct predecessor of k is $k-1$. Note that 0 is the direct predecessor of 1.

Then since $p > \Pr[G \mid N_1]$, $B^p_1(G) \cap N_1 = \varnothing$. But since $C^p(G) \subseteq B^p(G) \subseteq B^p_1(G)$, it is also the case that

$$C^p(G) \cap N_1 = \varnothing \tag{4}$$

This is because if an uninformed agent 1 does not assign probability p to G, then the event that G is common p-believed cannot include any state in which 1 is uninformed.

Now from Lemma A.2, $\Pr[Y_1 \mid N_2] < \Pr[G \mid N_1]$ which is less than p. So $B^p_2(Y_1) \cap N_2 = \varnothing$. Next (4) implies that $C^p(G) \subseteq Y_1$ and since B^p_2 is a monotone mapping, $B^p_2(C^p(G)) \subseteq B^p_2(Y_1)$. Finally, since $C^p(G)$ is a fixed point of B^p, $C^p(G) = B^p(C^p(G)) \subseteq B^p_2(C^p(G))$ and so $C^p(G) \subseteq B^p_2(Y_1)$. This implies that

$$C^p(G) \cap N_2 = \varnothing$$

Proceeding in this way we see that for all agents j along the path $1, 2, ..., k-1$, $C^p(G) \cap N_j = \varnothing$ and so

$$C^p(G) \cap N_k = \varnothing$$

In other words, the event that G is common p-believed cannot include any state in which k is uninformed.

Since k was arbitrary, we have shown that

$$\begin{aligned} C^p(G) &\subseteq \cap_{i\in\mathcal{I}} Y_i \\ &= Y^* \end{aligned}$$

Step 2: $p \leq \Pr[Y^* \mid Y_1]$ implies that $Y^* \subseteq C^p(G)$.

Since $p \leq \Pr[Y^* \mid Y_1]$, Lemma A.3 implies that for all i, $\Pr[Y^* \mid Y_1] < \Pr[Y^* \mid Y_i]$, we have that for all i,

$$B_i^p(Y^*) = Y_i$$

and taking intersections over i, $B^p(Y^*) = \cap_{i\in\mathcal{I}} Y_i = Y^*$ and so

$$C^p(Y^*) = Y^*$$

Now since $Y^* \subseteq G$, and the C^p operator is monotone, $C^p(Y^*) \subseteq C^p(G)$ and so

$$Y^* \subseteq C^p(G)$$

Case 3: $p > \Pr[Y^* \mid Y_1]$. Now p is so high that $B_1^p(Y^*) = \varnothing$ and so

$$C^p(Y^*) = \varnothing$$

as well.

From Step 1 of Case 2, we already know that $C^p(G) \subseteq Y^*$ and so

$$C^p(G) \subseteq C^p(Y^*) = \varnothing$$

This completes the proof. ■

3.1.2 Large ε

When $\varepsilon \geq \bar{\varepsilon}$, it is the case that

$$\Pr[G \mid N_1] \geq \Pr[Y^* \mid Y_1]$$

We then have

Proposition 3.2 *If $\varepsilon \geq \bar{\varepsilon}$, then for any information tree,*

$$C_{\mathcal{I}}^p(G) = \begin{cases} \Omega & \text{if } p \leq \Pr[G \mid N_1] \\ \varnothing & \text{if } p > \Pr[G \mid N_1] \end{cases}$$

Proof

Case 1: $p \leq \Pr[G \mid N_1]$. Here the proof is the same as in Case 1 of Proposition 3.1.

Case 2: $p > \Pr[G \mid N_1]$. As in Step 1 of Case 2 in the proof of Proposition 3.1,

$$C^p(G) \subseteq Y^*$$

Now since $\Pr[Y^* \mid Y_1] \leq \Pr[G \mid N_1] < p$, the probability that 1 assigns to Y^* is less than p and so $B_1^p(Y^*) = \varnothing$. It now follows that

$$C^p(Y^*) = \varnothing$$

This completes the proof. ■

Propositions 3.1 and 3.2 prove the first part of Theorem 1 since they show that $C_{\mathcal{T}}^p(G)$ depends only on $\Pr[G \mid N_1]$ and $\Pr[Y^* \mid Y_1]$, both probabilities that are independent of the information tree $\mathcal{T} = (F, s)$.

The second part of Theorem 1 now follows as a simple consequence of the two propositions. When $C_{\mathcal{T}}^p(G) = \Omega$ or $\varnothing$, the probabilities are obviously 1 or 0, respectively. When $C_{\mathcal{T}}^p(G) = Y^*$, Lemma A.4 implies that the probability is simply $(1-\varepsilon)^I$. So we have

Proposition 3.3 *The probability $\mathbb{P}_{\mathcal{T}}[C_{\mathcal{T}}^p(G)]$ does not depend on the information tree $\mathcal{T}$.*

3.2 An informativeness perspective

Some intuition for the irrelevance result can be gleaned by comparing different information trees using the informativeness criterion of Blackwell (1951).

Consider two information trees $\mathcal{T} = (F, s)$ and $\mathcal{T}' = (F', s')$. Let $d(i)$ denote the number of links between i and the root, agent 0, in the information tree $\mathcal{T}$ and let $d'(i)$ denote the analogous number in $\mathcal{T}'$.

It is then natural to say that $\mathcal{T}$ *diffuses information faster* than $\mathcal{T}'$, written $\mathcal{T} \succeq_{dif} \mathcal{T}'$, if there is a permutation of the names of the agents $\pi : \mathcal{I} \to \mathcal{I}$ such that for each $i \in \mathcal{I}$, $d(i) \leq d(\pi(i))$.

We will say that $\mathcal{T}$ is *first-order* more informative than $\mathcal{T}'$, written $\mathcal{T} \succeq_{FO} \mathcal{T}'$, if there is a permutation π such that for each $i \in \mathcal{I}$, i's information about G versus $\Omega \diagdown G$ in $\mathcal{T}$ is Blackwell more informative than $\pi(i)$'s information about G versus $\Omega \diagdown G$ in $\mathcal{T}'$.

Similarly, we will say that $\mathcal{T}$ is *second-order* more informative than $\mathcal{T}'$, written $\mathcal{T} \succeq_{SO} \mathcal{T}'$, if there is a permutation π such for each $i \in \mathcal{I}$, i's information about Y^*

versus $\Omega \diagdown Y^*$ in $\mathcal{T}$ is Blackwell more informative than $\pi(i)$'s information about Y^* versus $\Omega \diagdown Y^*$ in $\mathcal{T}'$. The terminology reflects the fact that Y^* is the event that all agents know that the state is G.

The following proposition shows that while the diffusion ordering $\succeq_{dif}$ is the same as the first-order $\succeq_{FO}$ ranking, the second-order $\succeq_{Y^*}$ ranking runs in the opposite direction. If $\mathcal{T}$ is better than $\mathcal{T}'$ in conveying first-order information, it is worse that $\mathcal{T}'$ in conveying second-order information (and vice versa).

Proposition 3.4 *For any two information trees $\mathcal{T}$ and $\mathcal{T}'$,*

(i) $\mathcal{T} \succeq_{dif} \mathcal{T}'$ if and only if $\mathcal{T} \succeq_{FO} \mathcal{T}'$

and

(ii) $\mathcal{T} \succeq_{FO} \mathcal{T}'$ if and only if $\mathcal{T} \preceq_{SO} \mathcal{T}'$

Proof. First, given any permutation π, note that $d(i) \leq d(\pi(i))$ if and only if i's information about G is Blackwell superior to $\pi(i)$'s information. This is because Lemma A.1 implies that $\Pr[G \mid N_i] \leq \Pr\left[G \mid N_{\pi(i)}\right]$ whereas $\Pr[G \mid Y_i] = \Pr\left[G \mid Y_{\pi(i)}\right] = 1$ since any Y_j is conclusive evidence that $\theta = g$. Thus, $d(i) \leq d(\pi(i))$ if and only if i's posterior beliefs about G are a mean-preserving spread of $\pi(i)$'s beliefs about G versus $\Omega \diagdown G$.[9] (i) now follows immediately.

Second, $d(i) \leq d(\pi(i))$ if and only if $\pi(i)$'s second-order information is Blackwell superior to i's information. In the latter case, the Blackwell experiment is well-defined since the agents have a common prior about Y^* given by Lemma A.4. Lemma A.3 and Lemma A.5 imply that $\Pr[Y^* \mid Y_i] < \Pr\left[Y^* \mid Y_{\pi(i)}\right]$ whereas, by definition, $\Pr[Y^* \mid N_i] = \Pr\left[Y^* \mid N_{\pi(i)}\right] = 0$. Thus, $\pi(i)$'s posterior beliefs about Y^* are a mean-preserving spread of i's beliefs about Y^* versus $\Omega \diagdown Y^*$.

Now from (i), $\mathcal{T} \succeq_{FO} \mathcal{T}'$ if and only if $\mathcal{T} \succeq_{dif} \mathcal{T}'$ and so there exists a permutation π such that for each i, $d(i) \leq d(\pi(i))$. Now the argument above shows that this is equivalent to $\mathcal{T} \preceq_{SO} \mathcal{T}'$. ■

The proposition establishes that there is a trade-off between the quality of information about G and the quality of information about Y^*. While this trade-off between first- and second-order information by itself is insufficient to establish our irrelevance result— the notion of common p-belief also employs third- and higher-order information—it does offer some intuition why it might hold.

4 Irrelevance result for games

In this section we show that for many games of interest, the results of the Section 3 concerning approximate common knowledge have *exact* counterparts concerning

[9] In experiments with two "states," G and $\Omega \diagdown G$, and two "signals," Y_i and N_i, this is sufficient for ranking the information in terms of the Blackwell criterion.

equilibria of certain games. We first return to the technology adoption game from the Introduction.

4.1 Technology adoption game

Recall that in the adoption game, each of I agents can decide to adopt a new technology (choose action $a_i = 1$) or not ($a_i = 0$). The cost of adoption is c per person. Adoption yields a gross payoff of 1 if everyone else adopts and the state is g. Otherwise, the gross payoff is zero. Let $\boldsymbol{a} = (a_i)_{i \in \mathcal{I}}$ denote the vector of actions of all the agents and let $\mathbf{1}$ denote the $I-$vector of 1s.

Consider a forest $F = (T^1, T^2, ..., T^R)$ and a seeding $s = \left(s^1, s^2, ..., s^R\right)$, where s^r is the unique seed of tree T^r. Again, denote by $\mathcal{T}$ the resulting (directed) information tree. Let $\mathcal{E}(\mathcal{T})$ be the set of (possibly mixed) equilibria of the adoption game in which the information to the agents comes via the network and seeding pair $\mathcal{T} = (F, s)$.

The counterpart of Proposition 3.1 is for any $\mathcal{T}$,

Proposition 4.1 *Let agent* 1 *be a seed of some tree. If* $0 < \varepsilon < \overline{\varepsilon}$, *then the highest equilibrium probability that everyone adopts is* $\mathbb{P}_{\mathcal{T}}\left[C^p_{\mathcal{I}}(G)\right]$ *where* $p = c$. *Precisely, for any* (F, s)

$$\max_{\sigma \in \mathcal{E}(\mathcal{T})} \mathbb{P}_{\mathcal{T}}\left[\boldsymbol{a} = \mathbf{1} \mid \sigma\right] = \begin{cases} 1 & \text{if } c \leq \Pr\left[G \mid N_1\right] \\ \Pr\left[Y^*\right] & \text{if } \Pr\left[G \mid N_1\right] < c \leq \Pr\left[Y^* \mid Y_1\right] \\ 0 & \text{if } c > \Pr\left[Y^* \mid Y_1\right] \end{cases}$$

Proof

Case 1: $c \leq \Pr\left[G \mid N_1\right]$. In this case, there is an equilibrium in which everyone adopts regardless of whether he is informed or not. To see this, note that if everyone but i always adopts, then the only uncertainty facing any agent is whether the fundamental state is g or b. Lemma A.1 implies that for all i, $\Pr\left[G \mid N_i\right] \geq \Pr\left[G \mid N_1\right] \geq c$, every agent is willing to adopt even if uninformed. Since everyone adopts regardless of information, the probability that everyone adopts is 1.

Case 2: $\Pr\left[G \mid N_1\right] < c \leq \Pr\left[Y^* \mid Y_1\right]$**.** In this case, there is an equilibrium in which everyone adopts if and only if informed. Moreover, there is no equilibrium in which an agent adopts with positive probability when uninformed.

There are two steps to the argument.

Step 1: $\Pr\left[G \mid N_1\right] < c$ implies that no uninformed agent adopts.

Consider the unique path from 0 to k and suppose (after renaming, if necessary) that this path consists of agents $1, 2, ..., k$ such that the direct predecessor of j is $j-1$.

Then since $c > \Pr\left[G \mid N_1\right]$, an uninformed agent 1 does not adopt even if everyone else adopts. In other words, adopting is dominated for an uninformed agent 1.

Now from Lemma A.2, $\Pr[Y_1 \mid N_2] < \Pr[G \mid N_1]$ which is less than c. In other words, adopting is iteratively dominated for an uninformed agent 2. Proceeding in this way we see that for all agents j along the path $1, 2, ..., k$, adopting is iteratively dominated for an uninformed agent j. Since k was arbitrary, we have shown that adopting is iteratively dominated for all agents.

Step 2: $c \leq \Pr[Y^* \mid Y_1]$ implies that if all other agents adopt when informed, it is a best response for an informed agent i to do so as well.

To see why, suppose all agents but i adopt when informed. Since $c \leq \Pr[Y^* \mid Y_1]$, Lemma A.3 implies that for all i, $c < \Pr[Y^* \mid Y_i]$, and so it is a best response for agent i to adopt as well. Thus, there exists an equilibrium in which everyone adopts if and only if informed.

In this case, the probability that everyone adopts is just the probability of Y^*. Because of Step 1, no equilibrium can involve adopting with positive probability when uninformed. Thus, the equilibrium in which every informed agent adopts gives the highest probability of adoption.

Case 3: $c > \Pr[Y^* \mid Y_1]$. In this case, the unique equilibrium is one in which no one ever adopts.

To see why, note that the cost is so high that even if everyone else adopts only if informed, it is a best response for an informed agent 1 to not adopt. Thus, agent 1 will never adopt.

This implies that no agent will ever adopt. Thus, the only equilibrium is one in which no one ever adopts. Of course, the probability of adoption is then 0.

This completes the proof. ■

Proposition 4.1 can also be derived as a consequence of Proposition 3.1 using some results from Kajii and Morris (1997) and Oyama and Takahashi (2020). We have chosen to provide a self-contained proof of the former as it emphasizes the parallel nature of the proofs of the two propositions. A counterpart of Proposition 3.2 can similarly be derived.

Payoffs What about agents' payoffs in the technology adoption game? It is easy to see that while the maximum equilibrium probability of adoption is independent of the information tree $\mathcal{T}$, agents' payoffs do depend on $\mathcal{T}$. This is easily verified in the three-agent example in the Introduction. Suppose ε is small and c is in the intermediate range. When 1 is the seed, the expected payoffs are $u_i = \rho(1-\varepsilon)^3 - (1-\varepsilon)^i c$. Note that $u_1 < u_2 < u_3$ and so the agent with the worst information about θ has the highest expected payoff. With broadcasting, all three agents have the same payoff $\rho(1-\varepsilon)^3 - (1-\varepsilon)c$.

It is also the case that the equilibria identified in Proposition 4.1 not only maximize the probability that everyone adopts, but also Pareto dominate all other equilibria in terms of payoffs.

Corollary 4.1 *In each case, the equilibrium from Proposition 4.1 that maximizes the probability that everyone adopts also Pareto dominates all other equilibria.*

Proof. To see that in each case the identified equilibrium Pareto dominates all other equilibria, first let (α_i, β_i) be the randomized strategy of i in which when informed, he adopts with probability α_i and when uninformed, adopts with probability β_i. Define $u_i\left(\alpha_i, \beta_{ii}, \boldsymbol{\alpha}_{-i}, \boldsymbol{\beta}_{-i}\right)$ to be i's expected payoff when he plays strategy (α_i, β_i) and the others play $\boldsymbol{\alpha}_{-i} = (\alpha_j)_{j\neq i}$ and $\boldsymbol{\beta}_{-i} = \left(\beta_j\right)_{j\neq i}$.

In Case 1, the fact that adopting exerts positive externalities implies that

$$u_i\left(\alpha_i, \beta_i, \boldsymbol{\alpha}_{-i}, \boldsymbol{\beta}_{-i}\right) \leq u_i\left(\alpha_i, \beta_i, \mathbf{1}_{-i}, \mathbf{1}_{-i}\right)$$

where $\boldsymbol{\alpha}_{-i} = \mathbf{1}_{-i}$ means that all $j \neq i$ play $\alpha_j = 1$ and the similarly for $\boldsymbol{\beta}_{-i}$. But since $(\alpha_i, \beta_i) = (1, 1)$ is a best-response to $(\mathbf{1}_{-i}, \mathbf{1}_{-i})$, we have that

$$u_i\left(\alpha_i, \beta_i, \boldsymbol{\alpha}_{-i}, \boldsymbol{\beta}_{-i}\right) \leq u_i\left(1, 1, \mathbf{1}_{-i}, \mathbf{1}_{-i}\right)$$

In Case 2, suppose that there is an equilibrium $\left(\alpha_j, \beta_j\right)$ for $j \in \mathcal{I}$. Since adopting is iteratively dominated when uninformed, it must be that in any equilibrium $\beta_j = 0$ for all j. Now again since adopting exerts positive externalities,

$$u_i\left(\alpha_i, 0, \boldsymbol{\alpha}_{-i}, \mathbf{0}_{-i}\right) \leq u_i\left(\alpha_i, 0, \mathbf{1}_{-i}, \mathbf{0}_{-i}\right)$$

where $\boldsymbol{\beta}_{-i} = \mathbf{0}_{-i}$ means that all $j \neq i$ play $\beta_j = 0$. Since $(\alpha_i, \beta_i) = (1, 0)$ is a best-response to $(\mathbf{1}_{-i}, \mathbf{0}_{-i})$,

$$u_i\left(\alpha_i, 0, \boldsymbol{\alpha}_{-i}, \mathbf{0}_{-i}\right) \leq u_i\left(1, 0, \mathbf{1}_{-i}, \mathbf{0}_{-i}\right)$$

In Case 3 the equilibrium is unique. ■

In each case of Proposition 3.1, the equilibrium that maximizes the probability that everyone adopts is symmetric—all agents play the same strategy. But not all equilibria are symmetric. Consider a situation in which 1 and 2 are connected and 3 is isolated (a trivial tree). Suppose 1 and 3 are the seeds for each of the two trees. Then if $\rho > \frac{1}{2}$ and ε is small enough, for an open set of c's there is an asymmetric equilibrium in which 1 and 2 adopt if and only if informed whereas 3 always adopts. Of course, the corollary above implies that this asymmetric equilibrium is Pareto dominated by one in which everyone always adopts.

In the technology adoption game, the players exert positive externalities on each other. The reader may then wonder if the irrelevance result relies on this feature. This is not the case. In the game below, the irrelevance result holds even though there are *negative* externalities.

4.2 Negative externalities

A group of I agents try to overthrow a repressive regime by protesting. If all I agents protest and the state is g, then the regime falls. Otherwise, it survives. But protests can turn destructive—by damaging public buildings, burning buses etc.. The greater the number of protesters, the greater the damage. Let $a_i = 1$ denote that i protests and $a_i = 0$ that he does not.

Suppose that in state g, the payoff functions are as follows:

$$\begin{aligned} u_i^g(1, \boldsymbol{a}_{-i}) &= \begin{cases} R - c - Id & \text{if } \sum_j a_j = I \\ -c - \left(\sum_j a_j\right) d & \text{if } \sum_j a_j < I \end{cases} \\ u_i^g(0, \boldsymbol{a}_{-i}) &= -\left(\sum_{j \neq i} a_j\right) d \end{aligned}$$

In state b, they are

$$\begin{aligned} u_i^b(1, \boldsymbol{a}_{-i}) &= -c - \left(\sum_j a_j\right) d \\ u_i^b(0, \boldsymbol{a}_{-i}) &= -\left(\sum_{j \neq i} a_j\right) d \end{aligned}$$

Here each player incurs a personal cost $c > 0$ of protesting (either an opportunity cost or one resulting from the risk of arrest or injury). Each additional protester also causes an additional $d > 0$ damage to public property and this damage reduces the payoff to everybody. If the regime is overthrown, then each player gets a reward $R > c + d$, the private plus public marginal cost of protesting. This last condition ensures that in state g, it is an equilibrium for everyone to protest.

In the protest game, players exert a *negative* externality on each other because of the destruction of public property.

Then it can be verified that the counterpart of Propositions 4.1 holds for the protest game as well. Simply replace "adopting" with "protesting" and the parameter c with the parameter $(c + d)/R$.

It should be noted, however, that Corollary 4.1 is special to the technology adoption game that exhibits positive externalities and does not extend to the destructive protest game with negative externalities.

4.3 Potential games

What is the general class of games for which the irrelevance result holds?

Consider an I-player *symmetric* game in which each i chooses an action $a_i \in A = \{0, 1\}$. Let $\boldsymbol{a} = (a_i)_{i \in \mathcal{I}}$ denote the vector of actions of all the players and let $\boldsymbol{a}_{-i} = (a_j)_{j \neq i}$ denote the vector of actions of all the players except i.

There are two possible states of nature $\theta \in \{g, b\}$ and the payoff functions in state θ, $u_i^\theta : A^I \to R$ are such that $u_i^\theta(a_i, \boldsymbol{a}_{-i})$ depends only on i's own action a_i and on the number of other players who play $a_j = 1$, $\sum_{j \neq i} a_j$.

As defined by Monderer and Shapley (1996), a game with payoffs $u_i^\theta : A^I \to R$ is a *potential* game if there is a *potential function* $v^\theta : A^I \to R$ such that for all i, a_i, a_i' and a_{-i},

$$u_i^\theta (a_i, \boldsymbol{a}_{-i}) - u_i^\theta (a_i', \boldsymbol{a}_{-i}) = v^\theta (a_i, \boldsymbol{a}_{-i}) - v^\theta (a_i', \boldsymbol{a}_{-i})$$

In other words, for each θ, the game is best-response equivalent to a game with common interests.

It is easily verified that *every* binary-action symmetric game is a potential game.

Consider games with potentials of the form

$$\begin{aligned} v^g (\boldsymbol{a}) &= \begin{cases} w & \text{if } \sum_j a_j = I \\ -\left(\sum_j a_j\right)\gamma & \text{if } \sum_j a_j < I \end{cases} \\ v^b (\boldsymbol{a}) &= -\left(\sum_j a_j\right)\gamma \end{aligned} \tag{5}$$

where w and $\gamma > 0$ are parameters that satisfy $w > -(I-1)\gamma$. Let $\mathcal{P}$ denote the class of potential games of the form in (5).

The requirement that $w > -(I-1)\gamma$ guarantees that in the game with common payoffs v^g, it is a strict Nash equilibrium for everyone to choose $a_i = 1$.

It is easy to verify that the technology adoption game from the previous subsection is a potential game from the class $\mathcal{P}$ where $\gamma = c$ and $w = 1 - I\gamma$, so that the last requirement that $w > -(I-1)\gamma$ reduces to $1 > c$.

Much along the lines of Proposition 4.1, it can be shown that if $\varepsilon < \bar{\varepsilon}$, then for *any* potential game in the class $\mathcal{P}$ the irrelevance result applies.

Proposition 4.2 *Let agent* 1 *be a seed of some tree. If* $0 < \varepsilon < \bar{\varepsilon}$, *then the highest equilibrium probability that everyone plays* $a_i = 1$ *is* $\mathbb{P}_\mathcal{T}\left[C_\mathcal{T}^p (G)\right]$ *where* $p = \frac{\gamma}{w+\gamma I}$. *Precisely, for any* $\mathcal{T} = (F, s)$

$$\max_{\sigma \in \mathcal{E}(\mathcal{T})} \mathbb{P}_\mathcal{T} [\boldsymbol{a} = \mathbf{1} \mid \sigma] = \begin{cases} 1 & \textit{if } \frac{\gamma}{w+\gamma I} \leq \Pr[G \mid N_1] \\ \Pr[Y^*] & \textit{if } \Pr[G \mid N_1] < \frac{\gamma}{w+\gamma I} \leq \Pr[Y^* \mid Y_1] \\ 0 & \textit{if } \frac{\gamma}{w+\gamma I} > \Pr[Y^* \mid Y_1] \end{cases}$$

We omit a proof of this result as it mimics the proof of Proposition 4.1.

5 Other networks and seedings

In this section we explore some limits to our results. Specifically, we show that the irrelevance result does not hold once the networks have cycles. Also, it is sensitive to the assumption that each tree has a single seed—multiple or random seedings can make the structure relevant.

In this section we state all of our findings using the technology adoption game. These can all be restated in terms of common p-beliefs as well.

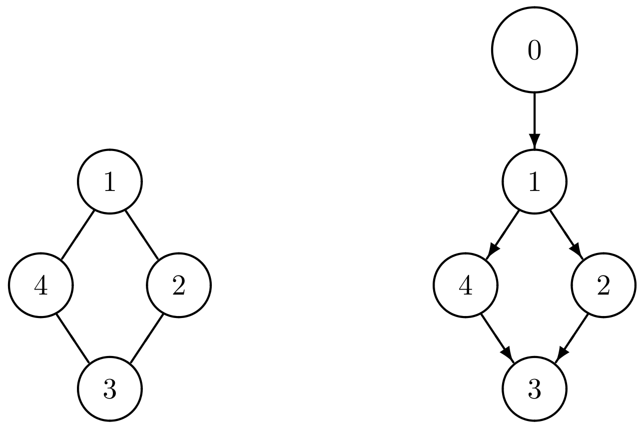

Figure 5: Cycle Network

5.1 Cycles

Our result that the set $C^p(G)$ of states in which G is common p-believed rests crucially on the assumption the underlying network is acyclic—a forest. Once there are cycles, the conclusion of our main result does not hold.

Consider, for example, a situation with 4 agents in the cyclic network depicted in Figure 5. Suppose that the planner seeds the network by sending a message to 1. The resulting information network, depicted on the right-hand side of the figure, is now not a directed tree. Here, 3 can receive messages from two sources—either from 2 or from 4. And of course, what 3 believes about whether agents 2 and 4 are informed depends on whom she hears from. If 3 receives a message only from 2, then she knows 1 and 2 are informed but is unsure about 4. If she hears from both 2 and 4, then she knows that everyone else is informed.

So let Y_3^2 denote the event that agent 3 heard a message *only* directly from 2 and similarly, let Y_3^4 denote the event that agent 3 heard only from 4. Finally, let $Y_3^{2\wedge 4}$ denote the event that she heard from both 2 and 4. As before, let N_3 denote the event that 3 did not hear from either source. The events Y_3^2, Y_3^4, $Y_3^{2\wedge 4}$ and N_3 are mutually exclusive and exhaustive.

Consider the adoption game with four players. We will compare equilibrium outcomes in the cycle network with those in the line network with four agents.

Claim 5.1 *Consider the four-agent technology adoption game. When ε is small, for an open set of costs c, there is an equilibrium with a line network seeded via* 1 *in which the probability that everyone adopts is greater than that from any equilibrium with the cycle network again seeded via* 1.

The proof of the claim is in Appendix B.1.

Figure 6: Mutiple Seeds

5.2 Multiple seeds

We have assumed that the planner sends information to a *single* agent in each tree of the forest network—that is, each tree has a single seed. Intuition suggests that it is better to send information to many agents at once—that is, to create multiple seeds. This will surely help reduce first-order uncertainty about the state. Here we show that the effect of multiple seeds on reducing higher-order uncertainty is ambiguous and in some circumstances a single seed is "better" than multiple seeds. In the technology adoption game from the introduction, there is a range of costs c for which agents' welfare is higher with a single seed than with multiple seeds.

Consider a simple network with two connected agents (as in left-hand panel of Figure 6). Now suppose the planner seeds *both* 1 and 2. This results in the information network depicted in the right-hand panel. Here if 1 gets a message, she passes it along to 2 and vice versa. Thus each agent has two sources of information—directly from the planner or indirectly from the other agent. And of course, what 1 believes about whether 2 is also informed depends on the channel by which she received a message. If 1 received a message from 2, then she knows that 2 is also informed. But if 1 received a message directly from 0, and not from 2, then she is unsure about whether 2 is informed.

Suppose that the agents play the two-player version of the technology adoption game from Section 4.

For the connected two-player network in Figure 6, we have

Claim 5.2 *Consider the two-agent technology adoption game. When ε is small, for an open set of costs c, there is an equilibrium with a single seed in which the probability that everyone adopts is greater than that from any equilibrium with two seeds.*

The proof of this claim is in Appendix B.2.

5.3 Random seeds

We have assumed that the planner chooses a single seed in each tree and the identity of the seed is known to all the agents. Does the irrelevance result still hold if the planner choose a single seed, but at *random*? The answer is no as the following example demonstrates.

Suppose there are only two agents, 1 and 2 and the network is connected. With probability $\frac{1}{2}$ the planner chooses 1 as the seed and with probability $\frac{1}{2}$ chooses 2 as the seed.

For the connected two-player network, we have

Claim 5.3 *Consider the two-agent technology adoption game. When ε is small, for an open set of costs c, there is an equilibrium with a random seeding in which the probability that everyone adopts is greater than that from any equilibrium with only one seed.*

The proof of this claim is in Appendix B.3.

6 Conclusion

People are interconnected in many ways. The same person may be part of a professional network, a family network or a leisure network and so have many sources of information. In such situations, the overall network may not have a tree-like structure and so our irrelevance result will not apply. Nevertheless, the point that there is a trade-off between disseminating information quickly and making the information commonly known can be applied more generally. When the goal of the policymaker is to engineer coordinated behavior, the latter is more important.

A Appendix: Agents' beliefs

This appendix derives agents' beliefs of different events used to prove the main result.

A.1 Beliefs along a path

Here we derive three results that compare the beliefs of agents who lie along the same path originating with the planner (agent 0). So let 1 be a seed and let K be a terminal node (a leaf). Suppose, after renaming, that the unique path from 0 to K consists of agents $k = 1, 2, .., K$ such that the direct predecessor of k is $k - 1$.

The first lemma simply says that the further an uninformed agent is from the seed, the more optimistic he is that the fundamental $\theta = g$. This is because an uninformed agent further from the seed ascribes a higher probability to the event that the fundamental is g and the message got lost somewhere along the way, than someone closer to the seed.

Lemma A.1 *The sequence*

$$\Pr[G \mid N_k]$$

is increasing in k*.*

Proof. Note that

$$\Pr[G \mid N_k] = \frac{\rho\left(1-(1-\varepsilon)^k\right)}{1-\rho(1-\varepsilon)^k} \tag{6}$$

since $\Pr[Y_k] = \rho(1-\varepsilon)^k$ and so $\Pr[N_k] = 1-\rho(1-\varepsilon)^k$.

The result then follows immediately. ■

The second lemma says that the further an uninformed agent is from the seed, the more pessimistic he is about the event that all his predecessors are informed. The intuition is that the further the agent is along the path, the greater the chance that the message was lost somewhere prior to reaching his immediate predecessor.

In what follows, it will be convenient to write

$$Y_0 = G \tag{7}$$

Lemma A.2 *The sequence*

$$\Pr[Y_{k-1} \mid N_k]$$

is decreasing in k *for* $k \geq 1$*.*

Proof. Note that since $Y_0 = G$, for any $k \geq 1$,

$$\Pr[Y_{k-1} \mid N_k] = \frac{\rho(1-\varepsilon)^{k-1}\varepsilon}{1-\rho(1-\varepsilon)^k}$$

The numerator is the probability of the joint event $Y_{k-1} \cap N_k$ which occurs if only if $k-1$ receives the message (probability $\rho(1-\varepsilon)^{k-1}$) and k does not (probability ε). The denominator is the probability of N_k which is just $1-\Pr[Y_k]$. Now Y_k occurs if and only if k gets the message (probability $\rho(1-\varepsilon)^k$).

It is now easy to verify that

$$\Pr[Y_{k-1} \mid N_k] > \Pr[Y_k \mid N_{k+1}]$$

■

The third lemma is also rather intuitive. It says that informed agents who are further along the path are increasingly optimistic that all agents, whether or not they are on the path, are informed as well.

Lemma A.3 *The sequence*

$$\Pr[Y^* \mid Y_k]$$

is increasing in k*.*

Proof. Since for all k,

$$\Pr[Y^* \mid Y_k] \times \Pr[Y_k] = \Pr[Y^*]$$

we have

$$\frac{\Pr[Y^* \mid Y_{k-1}]}{\Pr[Y^* \mid Y_k]} = \frac{\Pr[Y_k]}{\Pr[Y_{k-1}]}$$

And since $Y_k \subset Y_{k-1}$ and $k-1$ is the unique direct predecessor of k,

$$\begin{aligned} \Pr[Y_k] &= \Pr[Y_k \mid Y_{k-1}] \times \Pr[Y_{k-1}] \\ &= (1-\varepsilon) \times \Pr[Y_{k-1}] \end{aligned}$$

and so

$$\frac{\Pr[Y^* \mid Y_{k-1}]}{\Pr[Y^* \mid Y_k]} = 1 - \varepsilon \tag{8}$$

■

A.2 Probability of Y^*

Suppose the I agents are completely disconnected so that each agent is a "tree" with one node (as in Example 1 case B). Now the only way the information can get to all the agents is if every agent is a seed—that is, the planner "broadcasts" the message. Since each message is lost independently with probability ε, the probability that the information reaches all the agents is simply $(1-\varepsilon)^I$.

In an arbitrary tree (or more generally, a forest), if a message from i to j is lost so that j is uninformed, then this means that all agents in the sub-tree with j as the root are also uninformed. Thus, unlike in the case of a broadcast, whether or not i and j are informed are correlated. The next result shows that despite this, no matter what the structure of the forest is, the probability that all agents are informed is the same as when there is a broadcast.

Lemma A.4 *For any forest with I nodes,*

$$\Pr[Y^* \mid G] = (1-\varepsilon)^I$$

Proof. The proof is by induction on I.

For $I = 1$, clearly the probability $\Pr[Y^* \mid G] = 1 - \varepsilon$.

Now suppose that for any forest with $I - 1$ agents

$$\Pr\left[\cap_{i=1}^{I-1} Y_i \mid G\right] = (1-\varepsilon)^{I-1}$$

In the forest with I agents, let I be a leaf (a terminal node) of some tree in the forest and let the unique direct predecessor of I be $I-1$. Then since

$$\Pr\left[\cap_{i=1}^{I} Y_i \mid G\right] = \Pr\left[\cap_{i=1}^{I-1} Y_i \mid G\right] \times \Pr\left[Y_I \mid \cap_{i=1}^{I-1} Y_i, G\right]$$

and $\Pr\left[Y_I \mid \cap_{i=1}^{I-1} Y_i, G\right] = 1 - \varepsilon$, the claim is established. ■

A simple consequence of the previous result is

Lemma A.5 *For any forest,*

$$\Pr\left[Y^* \mid Y_1\right] = (1-\varepsilon)^{I-1}$$

Proof. The proof just mimics the proof of Lemma A.4. ■

Combined with the fact that for successive agents along the path from 1 to K,

$$\Pr\left[Y^* \mid Y_{k-1}\right] = (1-\varepsilon) \times \Pr\left[Y^* \mid Y_k\right]$$

(see (8)), Lemma A.5 implies that for all k,

$$\Pr\left[Y^* \mid Y_k\right] = (1-\varepsilon)^{I-k} \tag{9}$$

B Appendix: Other networks and seedings

B.1 Cycles

For the cyclical network of Section 5.1 we have

Claim B.1 *Suppose* $\frac{1-\varepsilon}{2-\varepsilon} < c < (1-\varepsilon)^4$ *and* $\frac{\rho\varepsilon}{\rho\varepsilon+1-\rho} < c$. *Then with the cycle network, there is an equilibrium in which* $i = 1, 2, 4$ *adopt if and only if informed and* 3 *adopts if only if he is informed via both* 2 *and* 4.

Proof. First, since $c > \Pr[G \mid N_1] = \frac{\rho\varepsilon}{\rho\varepsilon+1-\rho}$ it is iteratively dominated for any uninformed agent to adopt. Clearly, it is dominated for N_1 to adopt. Given this, Lemma A.2, it is iteratively dominated for N_2 and N_4 to adopt.

Second, given that N_1, N_2 and N_4 do not adopt, it is iteratively dominated for Y_3^2 to adopt. This is because if 3 got a message from 2, she is sure that both 1 and 2 are informed. The only uncertainty she faces concerns agent 4 and it is easily verified that $\Pr\left[Y_4 \mid Y_3^2\right] = (1-\varepsilon)/(2-\varepsilon)$. Since this is smaller than c, Y_3^2 should not adopt. By interchanging the roles of 2 and 4, we infer that Y_3^4 should not adopt either.

Third, given the specified strategies, it is a best-response for Y_1 to adopt since the probability that all others adopt is $\Pr\left[Y_2 \cap Y_4 \cap Y_3^{2\wedge 4} \mid Y_1\right] = (1-\varepsilon)^4$. To see this note that

$$\begin{aligned}\Pr\left[Y_1 \cap Y_2 \cap Y_4 \cap Y_3^{2\wedge 4}\right] &= \Pr\left[Y_1\right]\Pr\left[Y_2 \cap Y_4 \mid Y_1\right]\Pr\left[Y_3^{2\wedge 4} \mid Y_2 \cap Y_4\right] \qquad (10)\\ &= \Pr\left[Y_1\right] \times (1-\varepsilon)^2 \times (1-\varepsilon)^2\end{aligned}$$

Since $c < (1-\varepsilon)^4$, it is a best-response for Y_1 to adopt.

Given the strategies, Y_2, Y_4 and $Y_3^{2\wedge 4}$ are all more optimistic than Y_1 about the event that everyone will adopt. So they too will adopt. ■

From (10), the probability that everyone will adopt in the equilibrium described in the claim above is just $(1-\varepsilon)^5$ and this is the highest achievable since N_1, N_2, N_4, Y_3^2 and Y_3^4 do not adopt in any equilibrium.

In the line network (or any tree), the corresponding probability is $(1-\varepsilon)^4$.

B.2 Multiple seeds

When both 1 and 2 are seeds, let Y_i^0 denote the event that i heard only directly from the planner, Y_i^j the event that i heard only from agent $j = 3 - i$, and $Y_i^{0\wedge j}$ the event that i heard from both the planner and $j \neq i$. Finally, let N_i be the event that i hears from neither. Let Y_i denote the event that i heard from either source, that is, $Y_i = Y_i^0 \cup Y_i^j \cup Y_i^{0\wedge j}$.

In effect, there are now four types of each agent and thus the states of the world are more complicated since they specify not only whether or not i is informed but the source of her information. Let Ω_M be the states of the world for the example when there are multiple (two) seeds.

With multiple seeds,

$$\begin{aligned}\mathbb{P}_M\left[G \mid N_i\right] &= \frac{\rho\varepsilon^2 + \rho\varepsilon\left(1-\varepsilon\right)\varepsilon}{\rho\varepsilon^2 + \rho\varepsilon^2\left(1-\varepsilon\right) + 1 - \rho} \\ &= \frac{\rho\varepsilon^2\left(2-\varepsilon\right)}{\rho\varepsilon^2\left(2-\varepsilon\right) + 1 - \rho} \qquad (11)\end{aligned}$$

where the probabilities $\mathbb{P}_M$ are now determined in the network with multiple seeds. The numerator is the probability that if $\theta = g$, neither hears from 0 (probability ε^2) plus the probability that i does not hear from 0 (probability ε) but j does (probability $1-\varepsilon$) and then j's message is lost (probability ε). The denominator is just the probability that i hears from neither source.

First, note that if 1 hears directly only from 0, then her belief that 2 is informed is

$$\begin{aligned}\mathbb{P}_M\left[Y_2 \mid Y_1^0\right] &= \left(1-\varepsilon\right) + \varepsilon\left(1-\varepsilon\right) \\ &= 1 - \varepsilon^2\end{aligned}$$

and so from (11) it follows that for ε small enough,

$$\mathbb{P}_M\left[G \mid N_1\right] < \mathbb{P}_M\left[Y_2 \mid Y_1^0\right]$$

Claim B.2 *With multiple seeds, if* $\mathbb{P}_M\left[G \mid N_1\right] < c \leq \mathbb{P}_M\left[Y_2 \mid Y_1^0\right]$, *then there is an equilibrium in which agents adopt if and only if they get a message from either source. This equilibrium Pareto dominates all other equilibria.*

Proof. If i does not get a message, then his belief about the event G is smaller than the cost and so it is dominant to not adopt.

Suppose 2 adopts whenever she is informed. Now since, $c \leq \Pr\left[Y_2 \mid Y_1^0\right]$, if 1 gets a signal only from 0, he will adopt. And if 1 hears from 2, then he knows that 2 is also informed and so will also adopt. ■

For the range of costs in the claim above, with multiple seeds, the resulting equilibrium payoff can be calculated as follows.

The probability that both adopt in the event G is

$$\begin{aligned}\mathbb{P}_M [Y_1 \cap Y_2 \mid G] &= 2(1-\varepsilon)^2 \varepsilon + (1-\varepsilon)^2 \\ &= (1-\varepsilon)^2 (1+2\varepsilon)\end{aligned}$$

The probability that 1 adopts and 2 doesn't is

$$\mathbb{P}_M [Y_1 \cap N_2 \mid G] = (1-\varepsilon)\varepsilon^2$$

Since adoption occurs only in the event G, the ex ante equilibrium payoff of either agent when both are seeds is

$$\pi_M = \rho(1-\varepsilon)^2 (1+2\varepsilon)(1-c) + \rho(1-\varepsilon)\varepsilon^2(-c) \tag{12}$$

Single seed If 1 is the only seed, then we are in the situation studied in Section 4 and let $\mathbb{P}$ denote the resulting probability distribution over Ω. Proposition 3.1 (Case 1) implies that

Claim B.3 *With a single seed, if* $c < \mathbb{P}[G \mid N_1]$, *then there is an equilibrium in which everyone adopts regardless of information.*

The equilibrium payoff of an agent in the equilibrium of the claim is

$$\pi = \rho - c \tag{13}$$

Finally, note that when ε is small,

$$\mathbb{P}_M [G \mid N_1] < \mathbb{P}[G \mid N_1] < \mathbb{P}_M \left[Y_2 \mid Y_1^0\right]$$

So if c is such that

$$\mathbb{P}_M [G \mid N_1] < c \leq \mathbb{P}[G \mid N_1]$$

the conditions of both claims above are satisfied. From (12) and (13)

$$\pi_M - \pi = \rho(1-\varepsilon)^2 (1+2\varepsilon)(1-c) + \rho(1-\varepsilon)\varepsilon^2(-c) - (\rho - c)$$

and it may be verified that when $c \approx \mathbb{P}_M [G \mid N_1]$, the difference is negative.

B.3 Random seeds

Denote by $\mathbb{P}^1$ the probability distribution over Ω when 1 is the seed, by $\mathbb{P}^2$ the probability distribution over Ω when 2 is the seed, and by $\widetilde{\mathbb{P}}$ the probability distribution over Ω when the seeds are randomly chosen

Consider agent 1 when uninformed. The probability that this agent assigns to G is

$$\begin{aligned}\widetilde{\mathbb{P}}\left[G \mid N_1\right] &= \frac{\rho\left(\frac{1}{2}\varepsilon + \frac{1}{2}\left(\varepsilon + \left(1-\varepsilon\right)\varepsilon\right)\right)}{\rho\left(\frac{1}{2}\varepsilon + \frac{1}{2}\left(\varepsilon + \left(1-\varepsilon\right)\varepsilon\right)\right) + 1 - \rho} \\ &= \frac{\frac{1}{2}\rho\varepsilon\left(3-\varepsilon\right)}{\frac{1}{2}\rho\varepsilon\left(3-\varepsilon\right) + 1 - \rho}\end{aligned}$$

To see why, note that the numerator, the probability of $G \cap N_1$, involves three possibilities in state G: (i) 1 is the seed (probability $\frac{1}{2}$) and the message from the planner to 1 was lost (probability ε); (ii) 2 is the seed and the message from the planner to 2 was lost; and (iii) 2 is the seed, the message from the planner to 2 was received but then lost when sent to 1. The denominator takes into account the fact that when $\theta = b$, no messages are sent.

Also, recall from (1) that

$$\mathbb{P}^1\left[G \mid N_1\right] = \frac{\rho\varepsilon}{\rho\varepsilon + 1 - \rho}$$

and note that $\mathbb{P}^1\left[G \mid N_1\right] < \widetilde{\mathbb{P}}\left[G \mid N_1\right]$.

Let ε be small enough so that $\widetilde{\mathbb{P}}\left[G \mid N_1\right] < \mathbb{P}^1\left[Y^* \mid Y_1\right]$.

Claim B.4 *If*

$$\mathbb{P}^1\left[G \mid N_1\right] < c < \widetilde{\mathbb{P}}\left[G \mid N_1\right]$$

then with random seeding, there exists an equilibrium in which both agents adopt regardless of whether they are informed or not.

Proof. Suppose 2 always adopts. Then the only uncertainty that N_1 faces is regarding G and since $c < \widetilde{\mathbb{P}}\left[G \mid N_1\right]$, it is optimal for N_1 to also adopt. Then it is certainly optimal for Y_1 to adopt as well.

Agents 1 and 2 are symmetrically placed and so the claim is established. ■

When 1 is the only seed, Proposition 3.1 shows that when $\mathbb{P}^1\left[G \mid N_1\right] \leq c < \mathbb{P}^1\left[Y^* \mid Y_1\right]$, it is *not* an equilibrium for both agents to adopt regardless of whether they are informed or not.

Thus, with random seeding, the probability that everyone adopts is greater than with a single seed. The payoffs of the two agents are also greater with random seeding.

References

[1] Akbarpour, M., S. Malladi and A. Saberi (2023): "Just a Few Seeds More: The Inflated Value of Network Data for Diffusion," https://web.stanford.edu/~mohamwad/

[2] Banerjee, A., E. Breza, A. Chandrasekhar and B. Golub (2023): "When Less is More: Experimental Evidence on Information Delivery During India's Demonetization," *Review of Economic Studies,* forthcoming.

[3] Banerjee, A., A. Chandrasekhar, E. Duflo and M. Jackson (2013): "The Diffusion of Microfinance," *Science*, 341, 364–370.

[4] Banerjee, A., A. Chandrasekhar, E. Duflo and M. Jackson (2019): "Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials," *Review of Economic Studies*, 86, 2453–2490.

[5] Beaman, L., A. B. Yishay, J. Magruder and M. Mobarak (2021): "Can Network Theory-Based Targeting Increase Technology Adoption?," *American Economic Review*, 111, 1918–1943.

[6] Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, University of California Press, 93–102.

[7] Chetty, R., J. Friedman and E. Saez (2013): "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings," *American Economic Review*, 103, 2683–2721.

[8] Coles, P. and R. Shorrer (2012): "Correlation in the Multiplayer Electronic Mail Game," *BE Journal of Theoretical Economics*, 12.

[9] Crouzet, N., A. Gupta and F. Mezzanotti (2023): "Shocks and Technology Adoption: Evidence from Electronic Payment Systems," *Journal of Political Economy*, 131, 3003–3065.

[10] De Jaegher, K. (2015): "Beneficial Long Communication in the Multiplayer Electronic Mail Game," *American Economic Journal: Microeconomics*, 7, 233–251.

[11] George, S., S. Gupta and Y. Neggers (2019): "Coordinating Voters Against Criminal Politicians: Experimental Evidence from India," https://sites.google.com/view/siddharthgeorge/research

[12] Gottlieb, J. (2016): "Common Knowledge and Voter Coordination: Experimental Evidence from Mali." In *Voting Experiments*, edited by A. Blais, J-F. Laslier and K. Van der Straeten, New York: Springer Press.

[13] Jackson, M. (2008): *Social and Economic Networks*, Princeton: Princeton University Press.

[14] Kajii, A. and S. Morris (1997): "The Robustness of Equilibria to Incomplete Information," *Econometrica*, 65, 1283–1309.

[15] Monderer, D., and D. Samet (1989): "Approximating Common Knowledge with Common Beliefs," *Games and Economic Behavior*, 1, 170–190.

[16] Monderer, D., and L. Shapley (1996): "Potential Games," *Games and Economic Behavior*, 124–143.

[17] Oyama, D. and S. Takahashi (2020): "Generalized Belief Operator and Robustness in Binary-Action Supermodular Games," *Econometrica*, 88, 693–726.

[18] Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *American Economic Review*, 79, 385–391.

[19] Sadler, E. (2020): "Diffusion Games," *American Economic Review*, 110, 225–270.