Persuasion on Networks[*]

Georgy Egorov
Northwestern University

Konstantin Sonin
University of Chicago

April 2019

Abstract

We analyze persuasion in a model, in which each receiver (of many) might buy a direct access to the sender's signal or to rely on her network connections to get the same information. For the sender, a more biased signal increases the impact per subscriber (direct receiver), yet diminishes the willingness of agents to become subscribers. Contrary to the naive intuition, the optimal propaganda might target peripheral, rather than centrally located agents, and is at its maximum level when the probability that information flows between agents is either zero, or nearly one, but not in-between. The density of the network has a non-monotonic effect on the optimal level of propaganda as well.

Keywords: Bayesian persuasion, networks, propaganda, percolation.
JEL Classification: D85, L82.

1 Introduction

Any economic model of persuasion should satisfy the incentive compatibility constraint for the target of persuasion: no one would consume information unless it is beneficial to do so. If receiving information is costly - even if the cost is the price of newspaper subscription or simply the alternative cost of switching on a radio - the signal should be informative enough so that the receiver would opt to bear the cost. In a network, an agent might receive information from a neighbor, which reduces her willingness to bear the cost. Thus, there is a trade off for any producer of propaganda: when the signal is more biased, the impact on direct receivers and those who get information from them is higher, yet the willingness of agents to become direct receivers is lower.

Recently, propaganda and censorship has become an active area of research both empirically (King, Pan, and Roberts, 2013, 2014), and theoretically (Lorentzen, 2013, 2014, Gehlbach and Sonin, 2014, Kolotilin, Li, Mylovanov, and Zapechelnyuk, 2017). Some studies combine propaganda models with basic network analysis. However, this work often relies on naive intuition that treats information as a fluid matter in continuum mechanics or infection in biomedicine, ignoring the basics of economics approach: any consumption of information is a strategic action of a rational individual. As our simple examples demonstrate, this naive intuition, while fruitful in some contexts (e.g., Banerjee, Chandrasekhar, Duflo, and Jackson, 2018), might be conceptually misleading in others.

To analyze the impact of propaganda when agents are linked via a network, we use the basic Bayesian persuasion model (Kamenica and Gentzkow, 2011, Bergemann and Morris, 2018, Kamenica, 2018). There is a single sender who maximizes the expected amount of action by agents. (In a political economy setting, our main application, this will be, e.g., mobilization in support of a leader or abstention from the anti-government protest.) Without information, each agent would not act. The sender does not know the state of the world, and chooses the information design to maximize her own payoff. With a single receiver, the optimal strategy is to commit to truthfully report the state of the world in which the receiver prefers to act, and bias, in the favor of the sender-preferred action, the report in the state when the receiver prefers not to act.

When the sender chooses the information design, receivers would compare the cost of getting this information with the value that information provides to them; less bias makes information more valuable. In Example 1, we look at the consequences of the optimal choice by receivers (whether to buy subscription themselves or to get information for free from their connections)
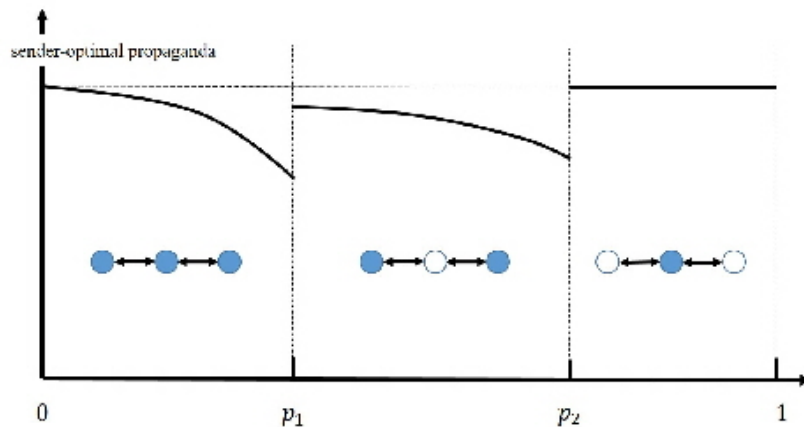
Figure 1: Sender-optimal equilibria as a function of p. Subscribers are solid. The optimal level of propaganda is decreasing on $[0, p_1)$ and $[p_1, p_2)$, and constant on $[p_2, 1]$.

and the sender who optimizes the impact of propaganda in a very simple linear network.

Example 1 Consider the network in Figure 1: the probability that information passes through any link connecting two agents is p; there is a non-zero cost of subscribing to the signal. For different values of p, different levels of propaganda (bias) are optimal. If p is high enough, the sender-optimal structure is such that the level of propaganda is at its maximum compatible with an individual agent buying subscription, the central agent chooses to subscribe, while each of peripheral agents opt to receive information from the central one. If p is low enough, then the impact is maximized if everyone is a subscriber. Finally, for intermediate values of p, the sender-preferred equilibrium is such that the level of propaganda is below the maximum level, the two peripheral agents buy subscription, and the central one receives information from them. (See the Appendix for full details on this example).

In Example 1, the sender has to reduce the bias in the message at the intermediate levels of p to make subscription incentive compatible for peripheral agents. If the central agent buys subscription, peripheral agents would get information with probability p. If the peripheral agents subscribe, then the central agent gets information with probability $1 - (1 - p)^2$, which exceeds p whenever $p \neq 0, 1$. Therefore, the sender trades off a lower level of bias that induces the desired action to have a higher probability that the signal reaches every agent. The sender's utility is maximized in two extreme cases. When p is very low, everyone is a subscriber and very little reduction of the maximum bias is needed to satisfy each subscriber's incentive compatibility constraint. When p is close to 1, there is a single subscriber, so the bias is at the maximum

3

level.

Example 1 is not an isolated phenomenon. Contrary to the naive intuition, the sender of propaganda would not necessarily target the centrally connected individual in a general star network. When the node with a high degree of centrality (Bloch, Jackson, and Tebaldi, 2019) receives information directly, this crowds out the incentives of those who can receive information from this node. In Example 2 (Figure 2), when the penetration probability p is sufficiently low, the sender would prefer that every peripheral agent becomes a subscriber.[1]
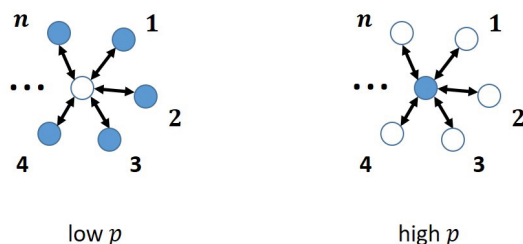

Figure 2: The sender-optimal equilibrium structure for different levels of p. Subcribers are solid.

Example 2 Consider a star network with r peripheral nodes depicted in Figure 2. When the penetration rate p is high, then the naive logic applies: the impact is maximized when the central agent is a subscriber, and other receive information from the central agent. However, when p is low, it might be optimal for the sender that everyone but the central agent opts to receive information directly, while the central agent receives information from r peripheral agents.

Proposition 1 describes the equilibrium information structure on a one-way infinite linear network and relates the size of the "subscription cell", the number of non-subscribers receiving information from a neighbor, to the models' fundamentals. While a part of our analysis is concerned with networks that do not allow cycles such as linear networks or star networks, the cyclic case is not necessarily different when information flows both ways on each link. At the same time, Example 4 in Subsection 6.1 demonstrates that sometimes there does not exist an equilibrium: a one-way network with three nodes provides such an example for a certain range

---

[1] Jackson and Yariv (2011) discuss the *best-shot public goods game*, a game of local public-good provision, first analyzed by Bramoullé and Kranton (2007) and Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2010), as an example of network games with externalities, which produce the effect similar to that in Figure 2. In a model of information diffusion in a development context, Akbarpour, Maladi, and Saberi (2018) have effect similar to that in Example 2: the optimal seeding structure might target peripheral, rather than centrally connected agents.

of parameters. Yet an equilibrium always exists if a two-way network is linear, finite or infinite, circular, or complete, or in many other situations.

Our main application is in political economics of authoritarian propaganda. The sender is the government that needs to maintain a certain level of support by the citizenry. The government commits to an information scheme that reports the state of the world: in practice, it might be an institutional structure that provides a certain level of censorship (King, Pan, and Roberts, 2013), a legal system that punishes dissemination of critical information in official media, or simply the policy of appointment of editors with a certain bias. In our simple model, the government commits to report the state of the world truthfully if it is favorable and bias the report in the unfavorable state. The problem for the sender is that once citizens rely on each other to receive information, there is less bias that the government can induce. An increase in the bias reduces the citizens' incentives to subscribe. As a result, more dense social network forces the government to report information more truthfully; this, in turn, makes authoritarian governments willing to deal with highly granular, atomized polities. Not surprisingly, totalitarian regimes discouraged all kind of networking, even strictly non-political, save for directly state-supervised one. Example 3 illustrates this intuition.

Example 3 Consider a polity in which the government (the sender) needs support from its citizens, i.e. action 1 in our model, to stay in power. There is an infinite countable set of citizens, whose support for the government depends on their estimate of the state of the world, which in turns depends on government's competence (or effort). In both pairs of polities, P_1 and P_1', and P_2 and P_2' in Figure 3, agents have the same number of connections (2 and 4, respectively) yet have markedly different network structures. In a sender-optimal equilibrium, it is easier to influence agents' actions and maintain government support in P_1 and P_2. In particular, there exists a threshold $\overline{p}$ such that for any $p > \overline{p}$, the government survives for a wider range of parameters with polity P_1 than with polity P_1' and with P_2 rather than P_2'. By Proposition 12, when p is high enough (not necessarily close to 1), the sender-optimal level of propaganda in P_1 and P_2 is $\beta^*(p) = \beta^{\max}$ and each connected component will have a unique subscriber. For P_1', it is never optimal to have propaganda at its maximum unless $p = 1$; for P_2', again, there exists another threshold $\widetilde{p} < 1$, such that $\beta^*(p) = \beta^{\max}$ for any $p > \widetilde{p}$.

As Example 3 illustrates, the equilibrium subscription structures are drastically different in P_1' and P_2'. In $P_1' = \mathbb{Z}$, for any p, $0 < p < 1$, the subscribers are allocated at a regular distance. The sender-optimal equilibrium is when the distance that is consistent with the subscribers' incentive compatibility constraint is minimal. This distance approaches infinity as p tends to 1,
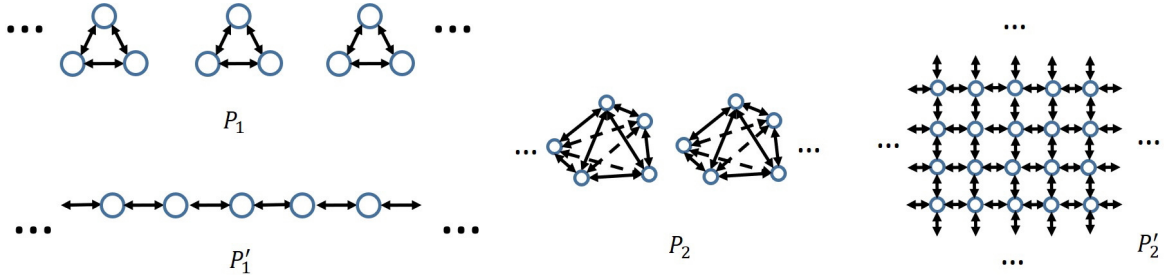
Figure 3: In networks P_1 and P_1', as well as in networks P_2 and P_2', each agent has the same number of links (two and four, resp.). When p is above a certain threshold, the government propaganda is more effective in P_1' and P_2'. However, in P_1, the sender-optimal subscription set is a singleton if and only if $p = 1$, while in P_2' is a singleton whenever p is larger than some $\overline{p} < 1$.

but it is never infinite unless $p = 1$. In contrast, in $P_2' = \mathbb{Z}^2$, for any level of propaganda and cost of subscription, there exists a threshold $\widetilde{p} < 1$ such that for any $p > \widetilde{p}$, there is exactly one subscriber. The reason for this is percolation (Broadbent and Hammersley, 1957, Duminil-Cupo, 2018): when p is large enough, there exists, almost surely, a unique infinite cluster of connected nodes. Therefore, for any agent in the network, the probability of getting informed is the probability of belonging to this infinite cluster. As a result, the spread of information, and the limits to propaganda, are very different in $\mathbb{Z}$ and $\mathbb{Z}^2$, and, more importantly, in networks which are locally similar to $\mathbb{Z}$ and $\mathbb{Z}^2$.

In our simple model of persuasion on networks, we ignore almost all of the peer effects discussed in the network literature (see, e.g., Jackson and Yariv, 2011, and Jackson, Rogers, and Zenou, 2019, on the central role of externalities in network analysis). The only externality that is present in our model is that a direct access to information by an agent discourages those who could get information from her to buy access themselves. Yet our model exhibits two critical economic features of information that differentiate it from a liquid matter flowing through a network. First, the higher is the informational content of a signal, the higher is the demand for it; if information is not valuable, it is not being consumed. Second, for each agent, different sources of information are gross substitutes: an agent does not have more information if it has reached her via two different paths. Still, this parsimonious setup allows to demonstrate the tension between information diffusion in a network and propaganda efforts by the sender, and challenge, at least theoretically, the naive wisdom that the agents with a high degree of network

6

centrality are the most obvious targets of the optimal propaganda.

The rest of the paper is organized as follows. Section 2 provides a brief literature review. Section 3 contains the setup. Section 4 starts the analysis with the case of a linear network; Section 5 extends the results to tree networks. Section 6 reports general existence results, including the analysis of complete and linear finite networks, and comparative statics, while Section 7 relates our results to percolation on infinite lattices. Then Section 8 applies the model's logic to propaganda in non-democratic regimes. Section 9 concludes, and the Appendix contains proofs of technical lemmas.

2 Related Literature

Bayesian Persuasion Kamenica and Gentzkow (2011) introduced the general concept of Bayesian persuasion and demonstrated that an optimal mechanism corresponds to a point on the concave closure of the sender's value function. The idea of commitment to information disclosure as an influence mechanism appeared in Aumann and Maschler (1966) and was developed, in different contexts, by, e.g., Brocas and Carrillo (2007), Rayo and Segal (2010), and Ostrovsky and Shwartz (2010). (See Kamenica, 2019, for the review of recent developments and Bergemann and Morris, 2018, for the general overview of information design.) Relative to other celebrated communication protocols such as cheap talk in Crawford and Sobel (1982), verifiable messaging in Milgrom (1981), and signaling in Spence (1973), Bayesian persuasion assumes more commitment power on behalf of the sender. This makes Bayesian persuasion a natural communication mechanism in a setup, where the sender's bias corresponds to censorship of a media or a propaganda campaign, which requires an institutional structure of a certain scale.

Bayesian persuasion with multiple receivers was considered in Bergemann and Morris (2013), Mathevet, Perego, and Taneva (2017), and Taneva (2016). Alonso and Camara (2016) and Kosterina (2018) assume public signals, in Wang (2015), signals are independent, and Arieli and Babichenko (2016), Bardhi and Guo (2018), and Chan, Gupta, Li, and Wang (2018) allow for arbitrary private signals. Kolotilin (2015, 2018) allow the sender to rely on private information about the receivers. While the persuasion part of our study is much simpler than in any of these papers, the contribution of this paper is on the relationship between the network structure and the persuasion mechanism.

Finally, Gentzkow and Kamenica (2014) provide a connection between the literatures on Bayesian persuasion and rational inattention (e.g., Sims 2003). As in our model subscribers

weight the benefits and the costs of having a direct access to information, the underlying intuition is somewhat similar to the rational inattention mechanism.

Information in Networks   We refer to Jackson (2008) for the exhaustive review of the early literature and excellent exposition of basic models of information diffusion in networks. Bala and Goyal (1998) pioneered a model of information diffusion in a network with Bayesian agents. Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) extend their results to a sequential setting with Bayesian learning. In Golub and Jackson (2010), agents receive independent signals about the true value of a variable and then update their beliefs by repeatedly taking weighted averages of neighbors' opinions. Sadler (2019) considers information contagions in a discrete network with Bayesian players that do not necessarily know the network structure. In Lipnowski and Sadler (2018), if two players are linked in a network, they get information about each others' strategies. Galperti and Perego (2019) consider a model, in which a designer sends signals to agents who then communicate in a network before playing the game. In our model, any two informed agents receive the same information, rather than two different signals, and therefore cannot learn from each other. A more general model would combine the two sets of assumptions: if, in addition to the signal from the media source, the agents receive private signals, our qualitative effects will stay the same.

Jackson and Yariv (2011) survey both the non-economics and economics literature on the role of externalities and diffusion in networks. (Akbarpour and Jackson, 2018, analyze the dynamics of such diffusion.) Chwe (2000) models people in a coordination game who use a communication network to tell each other their willingness to participate. Though the combinatorial patterns in our model when the network is linear infinite ($\mathbb{Z}$) and linear are resembling that of Chwe (2000), the incentives of agents are very different. Candogan (2019) considers a setting, in which agents linked a social network take binary actions, with the payoff of each agent depending, in addition to the underlying state of the world, on the number of her neighbors who take the sender-optimal action. Unlike Bramoullé and Kranton (2007) and Galeotti, Goyal, Fernando-Redondo, Jackson, and Yariv (2010), we do not assume any externalities in actions and payoffs: this allows us to concentrate on the impact of informational externalities that depend on the network structure.

In a model of information diffusion in a development context, Akbarpour, Maladi, and Saberi (2018) demonstrate that the optimal seeding structure might target peripheral, rather than centrally connected agents. In our model, information acquisition by a centrally connected agent crowds out more incentives to acquire information than a similar acquisition by a peripheral agent.

Propaganda In Gehlbach and Sonin (2014), citizens with heterogenous costs of accessing information, receive it from the censored media. Choosing the level of censorship, the government trades off the number of media customers and the impact on each receiver. In this paper, citizens might become direct receivers at a cost, yet there is an opportunity to receive information via their network connections which affects the government ability to influence them.

In Egorov, Guriev, and Sonin (2009), while information manipulation is an instrument of the dictator's survival, there is a cost in terms of the economic efficiency. Lorentzen considers a similar efficiency vs. stability trade off in the model of strategic protest restrictions (Lorentzen, 2013), and censorship (Lorentzen, 2014). In Shadmehr and Bernhardt (2015), the state does not censor modestly bad news to prevent citizens from making inferences from the absence of news that the news could have been far worse. Guriev and Triesman (2019) model the "informational dictatorship", in which the dictator chooses between repression and propaganda. (See also Chen, Lu, and Suen, 2016). Our model is the first, to the best of our knowledge, theoretical exercise to interact a propaganda mechanism with a social network structure.

DellaVigna and Gentzkow (2010) and Prat and Strőmberg (2013) are two excellent surveys on the role of persuasion in politics, with the emphasis on a democratic context. In an early attempt to connect political regimes with media freedom, Egorov, Guriev, and Sonin (2009) used panel data to confirm the relationship between oil wealth and information access: in dictatorships, more resources means less media freedom, whereas in democracies the effect disappears. Adena, Enikolopov, Petrova, Santarosa, and Zhuravskaya (2015) and Voigtländer and Voth (2015) analyze the role of propaganda in the rise of Nazis. Satyanath, Voigtländer, and Voth (2017) explore empirically the role of social capital in the rise of Nazis, but do not analyze the impact of the network structure. Yanigazawa-Drott (2014), King, Pan, and Roberts (2017a,b), and Chen and Yang (2018) use experimental technique to analyze empirically propaganda tactics of an authoritarian government. Chen and Yang (2018) use data on the social network of college dorm roommates for their estimates, but do not use specifics of the network structure.

Percolation The notion of (Bernoulli) percolation was introduced in Broadbent and Hammersle (1957). Kesten (1980) completed characterization of critical probability in $\mathbb{Z}^m$, which we use in Section 7. Van der Hofstadt (2010) and Duminil-Cupo (2018) are recent mathematical surveys that we refer to for basic definitions and results on percolation. Jackson (2008) describes applications of percolation insights to economic issues. We use results from percolation theory to demonstrate the limits that network structure puts on propaganda power of the government.

3 Setup

We will consider an environment, in which information is acquired not only from the source, e.g., a radio or a newspaper, directly, but indirectly from other members of the network. We will start with the simplest possible network, an infinite ordered sequence of agents numbered by nonnegative integers, but, in general, there is a network X that consists of nodes $x \in X$ and directed links (edges) $(x, y) \in E(X)$. We will assume that X is connected and locally finite (that is, each node belongs to a finite number of links).

There are two possible states of the world, $s \in \{0, 1\}$. Each agent $x \in X$ has to make, ultimately, an individual action, $a_x \in \{0, 1\}$, and payoff of each agent depends on the state of the world and the action taken as follows.

$$
\begin{aligned}
u_x(a_x &= 1, s = 1) = 1 - q, \\
u_x(a_x &= 0, s = 0) = q, \\
u_x(a_x &= 1, s = 0) = u_x(a_x = 0, s = 1) = 0.
\end{aligned}
$$

where $q \in (0, 1)$. The common prior is $P(s = 1) = \mu \leq q < \frac{1}{2}$, which guarantees that the default action in the absence of information is $a_x = 0$.

Prior to taking action, each agent x decides whether or not to subscribe to the media that provides an informative signal $\widehat{s}$. The signal is structured so that

$$
\begin{aligned}
P(s &= 1|\widehat{s} = 1) = \frac{\mu}{\mu + (1 - \mu)\beta}, \\
P(s &= 0|\widehat{s} = 0) = 1,
\end{aligned}
$$

where $\beta = P(\widehat{s} = 1|s = 0)$ is the control parameter of the sender. (Kamenica and Gentzkow, 2011, results guarantee that this is the optimal persuasion scheme in this setup.) There is a price of subscription, $c > 0$. We will assume that $c < \mu$. (In fact, we will assume c to be small enough when needed without explicitly mentioning this.) If the agent is indifferent, she subscribes. In Section 8, we briefly discuss the case of a random individual cost of subscription for each agent $x \in X$: qualitatively, most of results go through with $c_x \sim U[0, 2c]$.

With probability p, each agent x gets information from any agent y such that $(y, x) \in E(X)$. This is modeled using the Erdős-Rényi random graph approach (Jackson, 2008): given network X, each link disappears with probability $1 - p$ and information flows through the remaining links. We consider both one-way and two-way flow links (Bala and Goyal, 2000).

Definition 1 Given the sender's choice of β, an *equilibrium information structure* on network X is defined by the subset of subscribers $S(X)$ such that no agent would want to change

her subscription choice: any $x \in S(X)$ prefers subscription to no subscription, and any $y \in X \backslash S(X)$ prefers no subscription.

The sender maximizes the expected amount of action 1; in the case of an infinite network, the expected average amount of action 1.The timing is as follows. First, the sender commits to the level of propaganda β, then everything is determined in an equilibrium given β. If there is more than one such equilibrium, we will focus on sender-optimal equilibria.

Finally, we say that network X is *complete* if each node is connected to every other node, *(vertex)-transitive* if for any pair of nodes $x, y \in X$, there exists a graph isomorphism $X \to X$ (i.e., a bijective map on the set of nodes that maps links into links) that maps x on y, *star* if there exists a unique node x such that for any two distinct nodes y_1, y_2, all three nodes y_1, y_2, and x belong to a linear network, and *simple star* if the network is star and there is at most one node that belongs to more than one link.

4   Linear Networks

We start our analysis with the case of an infinite linear network, in which information flows in one direction, $X = \{0, 1, 2, 3, ...\}$ (Figure 4). The case of (directed) $X = \mathbb{Z}$ can be analyzed in a similar way; of course, in the case of an infinite linear networks, there is little qualitative difference between one-way and two-way flows.
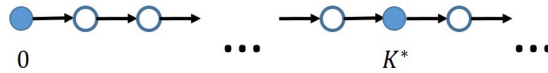
Figure 4: An equilibrium on a one-way linear network.

In the absence of information, any agent x chooses $a_x = 0$ and gets the expected pay-off of $(1 - \mu) q$. If there is access to information, the agent's expected utility is $\mu (1 - q) + (1 - \mu) (1 - \beta) q$. Thus, the value of information is

$$V = V(\beta) = \mu (1 - q) - (1 - \mu) \beta q.$$

Agent 0 subscribes if and only if $V \geq c$; we assume that this is always true. Consider the decision by agent $k \in \{1, 2, 3, ...\}$. Buying subscription gives the expected value of $V - c$.

Not buying gives V with probability p^k and 0 with probability $1 - p^k$. That is, agent k buys subscription if and only if $V - c \geq V p^k$, or, equivalently

$$V \geq c\frac{1}{1 - p^k}.$$

The right-hand side is increasing in k, so there exists K such that

$$c\frac{1}{1 - p^K} \leq V < c\frac{1}{1 - p^{K-1}}. \tag{1}$$

Note that function $K = K(\beta, \mu, q, c, p)$ is increasing in c and β and is decreasing in μ and p.

In equilibrium, the set of subscribers is $S(X) = \{nK | n = 0, 1, 2, ...\}$. The average amount of action by K people is

$$(\mu + (1 - \mu)\beta)\left(1 + p + ... + p^{K-1}\right) = (\mu + (1 - \mu)\beta)\frac{1 - p^K}{1 - p}.$$

(Recall that agents act if and only if they receive signal $s = 1$, which is translated by the media with probability $\mu + (1 - \mu)\beta$.)

To calculate the expected average action, take any m. Suppose, for now, that m is a multiple of K, $m = lK$. Then the amount of action among m agents is

$$\frac{m}{K}(\mu + (1 - \mu)\beta)\frac{1 - p^K}{1 - p}.$$

Now, the expected average action is

$$\lim_{m \longrightarrow \infty}\frac{1}{m}\left(\frac{m}{K}(\mu + (1 - \mu)\beta)\frac{1 - p^K}{1 - p}\right) = \frac{1}{K}(\mu + (1 - \mu)\beta)\frac{1 - p^K}{1 - p}.$$

(So, we effectively replaced $\lim_{m \longrightarrow \infty}\frac{1}{m}$ by $\lim_{l \longrightarrow \infty}\frac{1}{lK}$.) A straightforward observation is that the average expected amount of action is decreasing with K and, as a result, with p.

What is the optimal amount of propaganda β on such a network? The incentive compatibility constraint at the action stage ($a_x = 1$ when $s = 1$) for an individual subscriber comes from

$$\frac{\mu}{\mu + (1 - \mu)\beta}(1 - q) \geq \frac{(1 - \mu)\beta}{\mu + (1 - \mu)\beta}q,$$

which is equivalent to

$$\beta \leq \frac{\mu}{1 - \mu}\frac{1 - q}{q}. \tag{2}$$

The incentive compatibility constraint for agent 0 to subscribe, $V - c \geq 0$, is equivalent to $\mu(1 - q) - (1 - \mu)\beta q - c \geq 0$, i.e.

$$\beta \leq \frac{\mu(1 - q) - c}{(1 - \mu)q}. \tag{3}$$

12

As $c \geq 0$, if (3) is satisfied, then (2) is satisfied as well. Given the constraints, the optimal choice of the sender targeting one receiver is

$$\beta^{\max} = \frac{\mu (1 - q) - c}{(1 - \mu) q}.$$

As expected, $\beta^{\max}$, the maximum level of propaganda with one receiver, is lower when q, the intrinsic preference for action $a = 0$, is higher, or the cost of subscription, c, is higher. Naturally, when μ, the *ex ante* probability that action $a = 1$, is preferred, $\beta^{\max}$ is higher.

Now, consider

$$\beta^* = \arg \max_{\beta} \left\{ \frac{1}{K (\beta)} \left(\mu + (1 - \mu) \beta \right) \frac{1 - p^{K(\beta)}}{1 - p} \right\}.$$

Choosing optimal β^* can be done in two separate steps. First, chose $K = K(\beta^*)$, the number of the second (after agent 0) agent who buys subscription. Second, increase β^*, the bias that favors action, so that the IC constraint is binding for K, i.e.

$$V (\beta) = c \frac{1}{1 - p^K}.$$

which solves for

$$\beta(K) = \frac{1}{(1 - \mu) q} \left(\mu (1 - q) - c \frac{1}{1 - p^K} \right).$$

So, the maximization problem is

$$K^* = \arg \max_{K \in \mathbb{N}} \left\{ \frac{1}{K} \left(\mu - c \frac{1}{1 - p^K} \right) \left(1 - p^K \right) \right\}.$$

Let us make K real t, and analyze the continuous-variable optimization. (Multiplying by q and $1 - p$ does not affect maximization.)

$$t^* = \arg \max_{t \geq 0} \left\{ \frac{1}{t} \left(\mu \left(1 - p^t \right) - c \right) \right\}.$$

Technically, we prove Lemma A1 that shows that function $\frac{1}{t} \left(\mu \left(1 - p^t \right) - c \right)$ is single-peaked on $[0, +\infty)$ for any $a \in (0, 1)$. This does not necessarily imply that there always exists a unique K (and corresponding β) that maximizes the amount of action. However, it does prove that there is at maximum two such Ks, that such K is unique in any generic case, and also guarantees existence.

Proposition 1 *Let $X = \{0, 1, 2, ...\}$ define a one-way linear network.*

(i) Generically, there exists a unique optimal length of the "subscription cell" K^* such that

$$K^* = \arg \max_{K \in \mathbb{N}} \left\{ \frac{1}{K} \left(\mu \left(1 - p^K \right) - c \right) \right\}.$$

The corresponding propaganda level β^* satisfies $\left(1 - p^{K^*}\right) V(\beta^*) = c$, and maximizes the average amount of action in the network.

(ii) The subscription cell length $K^*(p)$ and the propaganda bias $\beta^*(p)$ are increasing functions of p : the higher is the probability that the signal passes, the less often the agents buy subscription, and the higher is the level of propaganda that the sender chooses.

(ii) $K^*(c)$ and $\beta^*(p)$ are increasing functions of c : the higher is the cost of subscription, the less often it is bought, and the higher is the sender-optimal level of propaganda.

One immediate corollary of Proposition 1 is that the optimal level of propaganda is lower in the case of the linear network than in the case of a single receiver. It is intuitive given that the sender has to trade off the impact on an individual receiver against the willingness of would be receivers to subscriber. This mechanism is very general: this result extends to any random graph in which for each node x the probability that x is reached from "the origin" is strictly positive and the set of subscribers is not a singleton. Indeed, setting the level of propaganda at the origin at β^{max} results in no other subscribers as each agent compares zero benefits of subscription to a positive cost of subscription. Technically, the comparative statics of parts (ii) and (iii) of Proposition 1 rely on Lemma A2 that demonstrates that the continuous function $\frac{1}{t}\left(\mu\left(1 - p^t\right) - c\right)$ satisfies increasing differences condition in (t, p) and (t, c) and provides the basis for comparative statics with respect to p and c. Then, the monotone comparative statics (Milgrom and Shannon, 1994) guarantees that the results extend to our main (discrete) case.

With the optimal β, every subscriber but agent 0 is on her incentive constraint. The subscriber #0 has a positive information. (If there are any non-subscribers at all; otherwise, $\beta = \beta^{\mathrm{max}}$, and there is no information rent left.). It is straightforward to extend this analysis to a two-way flow infinite linear network, $X = \mathbb{Z}$. (Lemma A5 formally establishes the existence of an equilibrium on $X = \mathbb{Z}$.) Of course, in the sender-optimal equilibrium in this case every subscriber will have zero surplus. In other dimensions, the story is very different as the specifics of one-dimensional case is that nobody could receive any information but from their immediate neighbors (see Section 7).

To describe equilibria on $X = \mathbb{Z}$ for any level of bias β, let $b = b(p, \beta, c)$ be the maximum number of non-subscribers between two subscribers in an equilibrium on $\mathbb{Z}$, and let S^b be the corresponding set of subscribers, $S^b = \left\{(b+1)\, n \mid n \in \mathbb{Z}\right\}$. Observe that $S^{\frac{b}{2}} = \left\{\left(\left\lfloor \frac{b}{2} \right\rfloor + 1\right) n \mid n \in \mathbb{Z}\right\}$ is an equilibrium set of subscribers as well. ($\lfloor t \rfloor$ denotes the integer part of t.) Indeed, if there were $b + 1$ agents between two subscribers, the agent $\left\lfloor \frac{b+1}{2} \right\rfloor$ would want to subscribe by construction of b. Since in $S^{\frac{b}{2}}$ every subscriber is in exactly the same position as this $\left\lfloor \frac{b+1}{2} \right\rfloor^{\mathrm{th}}$ agent
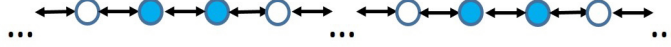
14

Figure 5: A special case of irregular equilibria on an infinite linear two-way network.

in S^b, and the distance between two subscribers is closer in $S^{\frac{b}{2}}$ than in S^b, $S^{\frac{b}{2}}$ defines an equilibrium. Define $a = a\left(p, \beta, c\right)$ to be the minimum distance such that $S^a = \left\{ \left(a+1\right) n \,|\, n \in \mathbb{Z} \right\}$ is an equilibrium set of subscribers. We just demonstrated that $a \leq \left\lfloor \frac{b}{2} \right\rfloor$. Trivially, for each β, the sender-optimal equilibrium set of subscribers is S^a. The optimal amount of propaganda on such a network satisfies $c = V(\beta^*) \left(1 - p^{a^*}\right)^2$, where $a^* = \min_\beta \left\{ a\left(\beta\right) \right\}$.

Proposition 2 summarizes the above discussion.

Proposition 2 *Let $X = \mathbb{Z}$, two-way, and consider any level of propoganda β.*

(i) There exist two integer thresholds $a = a\left(p, \beta, c\right)$ and $b = b\left(p, \beta, c\right)$ such that for any integer d, $a \leq d \leq b$, $S^d = \left\{ dn \,|\, n \in \mathbb{Z} \right\}$ is an equilibrium set of subscribers, and for any $d < a$ or $d > b$, S^d is not an equilibrium set.

(ii) The threshold a that defines the sender-optimal equilibrium on two-way $X = \mathbb{Z}$ satisfies $a \leq \left\lfloor \frac{b}{2} \right\rfloor$.

Part (ii) of Proposition Proposition 2 is an important element in our proof of existence of pure strategy equilibrium on a finite linear network for any β. Proposition 2 describes "symmetric" equilibria, in which subscribers are separated by the same number of subscribers. However, there might exist other equilibria, in which the gap between subscribers varies. Figure 5 demonstrates one such equilibrium. Still, any sender-optimal equilibrium is "symmetric".

5 Tree Networks

The analysis of infinite linear networks allows us to study persuasion on more general networks. Suppose the network $X = \mathbb{T}_r$ is a tree with each node having $r > 1$ outgoing links. (See Figure 6 illustrating the binary tree network.) The tree networks play a critical role in percolation on random graphs as in many networks any connected cluster that starts with an origin (a single subscriber in our model) is a tree. What is the optimal level of propaganda β in such a network?

Each directed path through a tree is isomorphic to the one-dimensional one-way case. Define $K(\beta)$ as in (1). Now, suppose that β^* is the optimal level of β. Then, the IC constraint should
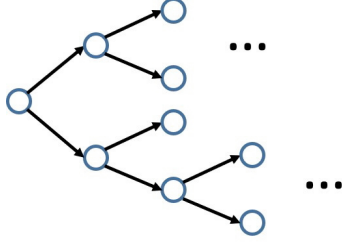
Figure 6: A binary tree network $(n = 2)$.

be binding for $K(\beta^*)$:

$$V(\beta^*) = c\frac{1}{1 - p^{K(\beta^*)}},$$

or, equivalently

$$\beta^* = \frac{1}{(1 - \mu)\, q}\left(\mu(1 - q) - c\frac{1}{1 - p^{K^*}}\right).$$

In the $r-$link network, the average action can be calculated as follows. For any K, the total number of agents at levels $0, ..., K - 1$ from the root is $\frac{r^K - 1}{r - 1}$. The expected number of agents who receive the propaganda signal at levels $0, ..., K - 1$ from the root is

$$1 + rp + (rp)^2 + ... + (rp)^{K-1} = \frac{(rp)^K - 1}{rp - 1}.$$

Each agent that received a signal acts with probability

$$\mu + (1 - \mu)\beta = \frac{1}{q}\left(\mu - \frac{c}{1 - p^K}\right).$$

So, the problem is to maximize the average action is as follows.

$$K^*(r) = \arg\max_{K \in \mathbb{N}}\left\{\frac{r - 1}{r^K - 1}\frac{1}{q}\left(\mu - \frac{c}{1 - p^K}\right)\frac{(rp)^K - 1}{rp - 1}\right\}.$$

If $rp > 1$, Lemma A3 guarantees that the continuous function $\varphi(t, r) = \frac{(rp)^t - 1}{(r^t - 1)}\left(1 - \frac{c/\mu}{1 - p^t}\right)$ is single-peaked with respect to t. As in the linear network case, this demonstrates that there exists a unique optimal point t^*, which guarantees existence (but not necessarily uniqueness) of the optimal K^*.

Proposition 3 *Let* $X = \mathbb{T}_r$. *Suppose that* $rp > 1$.

   (i) Generically, there exists a unique $K^*(r)$ *such that*

$$K^*(r) = \arg\max_{K \in \mathbb{N}}\left\{\frac{(rp)^K - 1}{r^K - 1}\left(1 - \frac{c/\mu}{1 - p^K}\right)\right\}.$$

16

(ii) The optimal length of the "subscription cell" $K^*(r)$ is an increasing function of r, the number of outgoing links in each node of regular-tree network X. Also, K^* is an increasing function of c, the cost of subscription, and a decreasing function of μ, the ex ante probability of the sender-desired state.

The comparative statics rely on Lemma A4, which asserts that function $\varphi(t) = \varphi(t|r, c, \mu)$ satisfies increasing differences in (t, c), $(t, -\mu)$, and (t, r). While Lemma A4 deals with the continuous function, the monotone comparative statics applies to the discrete maximization in $K^*(r)$. In particular, Proposition 3 yields that that the sender-optimal level of propaganda in an r-arnary tree network increases with r, the number of outdoing links in each node. As in the case of the linear network in Section 4, the sender trades off the frequency of subscribers, which requires less propaganda bias in the signal, against the impact of the signal. Now that each subscriber affects r groups similar to one group of the linear network, the sender puts in more propaganda bias, which results in a longer subscription cell.

6 General Networks

In this Section, we start with existence results for a few classes of networks. Then, we produce comparative statics results for the same network classes.

6.1 Existence

The first observation about existence of equilibrium in any two-way network is that if the sender moves first, then she could always choose a level of propaganda, for which the equilibrium division on subscribers and non-subscribers exists. Trivially, if $\beta^{\max}$ is the chosen level of propaganda, then the set of subscribers in any connected two-way network is a singleton. If the network is, in addition, complete, any agent can be this unique subscriber. Still, an equilibrium with $\beta^{\max}$ is not necessarily sender-optimal, even in a complete two-way network. On the other hand, with an arbitrary β and a general network X, existence is not guaranteed. In what follows, we identify certain classes of networks, in which an equilibrium exists for any level of propaganda. We start with the following simple observation.

Proposition 4 *Suppose that a finite information network X does not have any cycles. Then for any β, there exists an equilibrium.*

Proof. For βs such $V(\beta) < c$, the result is trivially true: no one buys subscription. Suppose that $V(\beta) \geq c$. Start with any node $x \in X$ that is not an end of any link. As $V(\beta) \geq c$, x buys

subscription. Then use induction to follow any path starting from x; the no-cycles assumption assures correctness of this procedure. Repeat until all nodes that are not ends of links are exhausted. ∎

While the absence of cycles guarantees the existence of an equilibrium, this need not be true when cycles are present. The simple Example 4 shows that an equilibrium in pure strategies might not exist.
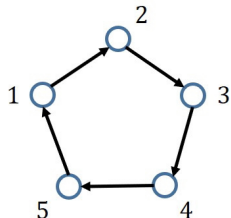

Figure 7: The cyclic network, for which there is no equilibrium for some open set of parameters.

Example 4 Consider the network in Figure 7 ("a wheel on five nodes", Jackson, 2008). Suppose that agent #1 is a subscriber. It is straightforward to verify that there exist parameters p, β, c, μ such that agent #3 prefers not to subscribe, yet agent #4 prefers subscription:

$$p^2 V\left(\beta\right) > V\left(\beta\right) - c > p^3 V\left(\beta\right).$$

Then, if #3 subscribes, #1 prefers to switch to non-subscription.

Note that all parameters in Example 4 are generic: for each of them, there exists an open neighborhood, in which the inequalities (and the non-existence result) stay true. Furthermore, any one-flow "wheel", with a suitable choice of β and p, will exhibit no equilibria. Indeed, for an "n-node wheel", choose p so that the optimal length of the subscription cell K is $\left\lfloor \frac{1}{2} n \right\rfloor + 1$, where $\lfloor t \rfloor$ denotes the integer part of t. Then, if this agent is a subscriber, the first subscriber will prefer not to be. At the same time, if a circular network is two-way, it has pure strategy equilibria for any β.

Remark 1 For a finite network, either directed or undirected, the existence of a mixed-strategy equilibrium is guaranteed by the Nash existence theorem. Another way to deal with the non-existence of pure-strategy equilibria on a "wheel with five nodes" might be to define a "dynamic equilibrium" on the network in Figure 7, in which the set of subscribers depends on time t. E.g., in period one $S_1(X) = \{\#1, \#4\}$, then $S_2(X) = \{\#2, \#5\}$, $S_3(X) = \{\#3, \#1\}$, etc. This concept would immediately extend to any one-way network on "wheels" and their direct

products (such as a one-way cell network on a two-dimensional torus). In the case of a "wheel", these equilibria will be periodic.

An important class of networks, in which existence of a pure strategy equilibrium is guaranteed for any β, is complete networks.

Proposition 5 *Suppose that a finite information network X is two-way and complete, i.e., each node is connected to every other node. Then for any β, there exists an equilibrium.*

Proof. Start with any node $x \in X$. As $V(\beta) \geq c$, x buys subscription. Then follow any path until some next agent buys subscription, and repeat the procedure until the network is exhausted. The critical reason why it is possible to do this in the case of a complete network is as follows. Suppose the procedure reached some agent $y \in X$, for which the value of subscription, $V(\beta) - c$, exceeds the value of non-subscription, $f_y(p)V(\beta)$. Then for any agent who has already subscribed, the value of subscription, once y is subscribed, is equal to $f_y(p)V(\beta)$. That is, none of the existing subscribers would want to change her strategy. ∎

Conjecture 1 *The result of Proposition 5 should extend to any two-way finite transitive network by, essentially, the same argument. However, the two-way assumption is necessary as the graph in Example 4 is transitive.*

Another important class of networks, in which existence is guaranteed for any β, is finite linear two-way networks. To prove this, we could use the parameters $b = b\left(p,\beta,c\right)$, the maximum number of non-subscribers between two subscribers in an equilibrium on $\mathbb{Z}$ (and so $S^b = \left\{\left(b+1\right)n \,|\, n \in \mathbb{Z}\right\}$ is the equilibrium set of subscribers) and $a = a\left(p,\beta,c\right)$, the minimum distance such that $S^a = \left\{\left(a+1\right)n \,|\, n \in \mathbb{Z}\right\}$ is an equilibrium set of subscribers. Recall from Proposition 2 in Section 4 that $a \leq \left\lfloor \frac{b}{2} \right\rfloor$.

Consider a finite two-way network X and make the left-most agent a subscriber. Choose $b + 2^{\text{th}}$ agent on X to be a subscriber, so that there is no agents who would want to subscribe between the first and second subscriber. Suppose that there is no such subscriber, i.e. $|X| \leq b+1$. If $|X|$ is between a and b, making the right-most agent a subscriber results in an equilibrium. If the length of the network X is less than a, and there exists d such that $d < a$, $p^d < 1 - \frac{c}{V(\beta)}$, make the agent $d + 1$ a subscriber. If there are non-subscribers to the right of $d + 1$ who want to subscribe, move the subscriber from $d + 1$ right-ward. Since $|X| < a$, moving a final number of steps will be sufficient, and there cannot be would-be subscribers in between the two by construction of a.

Now suppose that we made agents $1, b + 2, ..., l(b + 1) + 1$ subscribers and l is the largest such number satisfying $l(b + 1) + 1 \leq |X|$. If $|X| - l(b + 1) - 1 < d$, this set of subscribers is equilibrium. If $|X| - l(b + 1) - 1 \geq a$, there is no problem as well: making the right-most agent a subscriber completes the equilibrium set. If $|X| - l(b + 1) - 1 < a$, make agent with number $|X|$ a subscriber and move subscriber $l(b + 1) + 1$ to the left until $|X| - l(b + 1) - 1 = a + 1$. Since $a \leq \lfloor \frac{b}{2} \rfloor$, the distance between $l(b + 1) + 1$ and $|X| - l(b + 1) - 1$ exceeds $\lfloor \frac{b}{2} \rfloor$. Thus, we ended up with an equilibrium subscription set.

A similar argument establishes that in any circular network, an equilibrium exists for any β. An argument similar to one in Section 4 proves existence in the case of a star network, in which there are no terminal nodes.

Conjecture 2 *An equilibrium exists for any two-way star network and for any two-way tree.*

In linear networks, each subscriber's incentives are determined by the distance to at most two other subscribers. Now, let us consider the following important case in which this is not the case. Let X be an $m-$dimensional integer lattice, in which information flows in both direction on each link. For each m, there exists a threshold $p_c(\mathbb{Z}^m)$ such that for any $p < p_c(\mathbb{Z}^m)$, the probability that a cluster that originates with a subscriber is, in expectation, of a finite size (see more details in Section 7).

Proposition 6 *Let $X = \mathbb{Z}^m$ be an $m-$dimensional integer lattice, in which information flows in both direction on each link and suppose $p < p_c(\mathbb{Z}^m)$. For any β, there exists a symmetric equilibrium.*

Proof. Let $m = 2$. A helpful observation is that if the structure is "symmetric" (see below), than the value of non-subscription for each subscriber is the same.

Suppose that $V(\beta) \geq c$. We will consider the sequence of sets of potential subscribers S_h, $h \in \mathbb{N}$. For each h, define the set

$$S_h^2 = \{(2hl_1, 2hl_2), (2hl_1 + 1, 2hl_2 + 1) \,|\, l_1, l_2 \in \mathbb{Z}\}.$$

(See Figure 9 S_2^2.) First, for each $x \in S_h^2$, the alternative value of subscription (the amount of information that she receives from others), v_h, is the same. Second, it is decreasing in h, asymptotically approaching zero. (This is the place where the assumption $p < p_c(\mathbb{Z}^m)$ is used.) Thus, there exists h^* such that $v_{h^*-1} > V(\beta) - c$ and $v_{h^*} < V(\beta) - c$. To show that $S(X) = S_{h^*}^2$ is the equilibrium subscriber set, it is sufficient to demonstrate that any $x \in X \backslash S_{h^*}^2$ would not

want to subscribe. Observe that for any $x \in X \backslash S_{h^*}^2$, there exists at least one subscriber who is closer than the distance between two subscribers on $S_{h^*-1}^2$.

The proof for any $m > 2$ uses S_h^m in a similar way. ∎

6.2 Comparative Statics

Some comparative statics results are true for any network. For example, for any finite network, the number of subscribers in a sender-optimal equilibrium (weakly) decreases with p.[2] For infinite networks, this is true for the average, rather than total, number of subscribers. However, many other comparative statics results depend on both network structure and the penetration rate p.

Again, we start with networks that do not have any cycles. Such networks include directed trees, but also include networks in which different nodes can be connected by more than one path.

Proposition 7 *Suppose that two networks, X and Y, have no cycles, have the same set of nodes, and X has every link that Y has. Then, for any given p and β, the average amount of action is (weakly) lower in X than in Y.*

Proof. Take any path in Y that starts with a node that is not an end of any link. In the corresponding equilibrium in X, there cannot be a subscriber *earlier* (along the path) than in Y. (Because the amount of outside information that non-subscribers receive as a result of new links can only increase.) Therefore, X will have (weakly) less subscribers than Y. ∎

Proposition 7 suggests the next class of networks, to which the comparative statics might be extended. For example, consider an $r-$ regular (one-way) tree, in which nodes with the distance of l from the root are connected by additional two-way links. Consider two connected subnetworks, X and Y, that have the same set of nodes and "vertical" links as this tree, and suppose that X has every link that Y has. By induction, each "floor" of such network will contain less subscribers in X than in Y.

Let us now turn to complete graphs, i.e., two-flow networks, in which each node is connected to every other node. As it is shown in Proposition 5, any such network admits an equilibrium for any β. Example 5 describes the comparative statics with respect to probability p (see Figure 8) in a three node-network.

---

[2]This is not true for arbitrary equilibrium subscription structures. For example, for $X = \mathbb{Z}$, there exists a range of equilibria for any p as described in Proposition 2; they are payoff-ordered from the sender's standpoint.
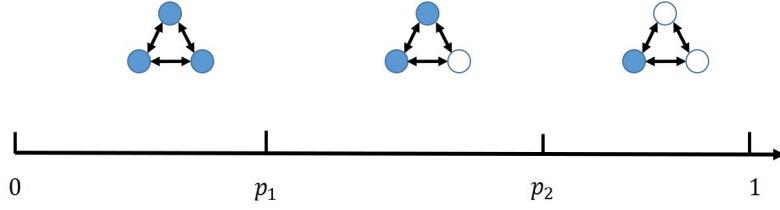
Figure 8: Sender-optimal equilibria in a full network with two-way information flow as a function of p. Subscribers are solid.

Example 5 Consider the complete network in Figure 8, in which each pair of agents is connected to each other. The probability that information passes through an link connecting two agents is p. Again, for different values of p, different levels of propaganda are optimal. If p is high enough, the sender's payoff is optimized at the level of propaganda such that only agent chooses to subscribe, while each of the two other agents opt to receive information from that one. If p is low enough, then the impact is maximized if everyone is a subscriber. Finally, for intermediate values of p, the sender-preferred equilibrium is such that two agents buy subscription, and the third one receives information from them.

The logic of Example 5 extends to any finite complete network. If such a network has n nodes, then for any m, $1 \leq m \leq n$, there exists a range of parameters $[p_{\min}(m), p_{\max}(m)]$ such that for any $p \in [p_{\min}(m), p_{\max}(m)]$, there is an equilibrium with exactly m subscribers. Also, Example 5 demonstrates that the comparative statics with respect to additional links depends critically on p. Indeed, consider a network X with nodes $\{1, 2, 3\}$ and two-way links $(1, 2)$ and $(2, 3)$. When p is low, the optimal structure is $S(X) = \{1, 3\}$ and adding link $(1, 3)$ forces the sender to reduce the bias; the impact of propaganda becomes smaller as there is more crowding-out effect. At the same time, when p is high, $S(X) = \{2\}$, $\beta^* = \beta^{\max}$ and adding link $(1, 3)$ increases the impact of propaganda as both agents 1 and 3 receive the same information with a strictly higher probability.

In a complete network, there is a recursive formula to calculate the probabilities with which agents get information. Denote $f(m, k)$ the probability that a non-subscriber in a complete network with m members and k subscribers receives information. Then the following recursive formula holds

$$f(m + 1, k) = 1 - (1 - p)^k (1 - pf(m, k))^{m-k}.$$

Indeed, the new $m + 1^{\text{th}}$ agent will not get any information from k subscribers with prob-

ability $(1-p)^k$ and will not get any information from $m-k$ non-subscribers with probability $(1-pf(m,k))^{m-k}$. (The probability of not getting information from one non-subscriber is $1-f(m,k)+(1-p)f(m,k)=1-pf(m,k)$.)

Consider two finite complete networks, X and Y, such that $Y \subset X$, $|X|=|Y|+1$. Take $\beta(Y)$, which is optimal for Y. Then, every subscriber $x \in S(Y)$ is on the IC constraint. (Recall that in a complete network every subscriber has the same expected pay off.) With $\beta(Y)$ in X, the $(|Y|+1)$th agent becomes a nonsubscriber as she receives more information than any nonsubscriber in Y before and every previous non-subscriber stays non-subscriber for the same reason (the only thing that changed is that there is an additional channel, via the $(|Y|+1)$th agent, to receive information). Some subscribers cannot be subscribers as they were on the IC constraint and benefits of non-subscription increased for the same reason. So, with $\beta(Y)$ in X the number of subscribers is strictly smaller than in Y. So, the optimal choice of $\beta(X)$ might be either higher than $\beta(Y)$ - that is, the sender would decide to put all those agents who are subscribers of X with $\beta(Y)$ on the IC constraint, or it might be lower, if the sender decides to increase the number of subscribers. While this explains why the comparative statics might be ambiguous locally, it is unambiguous when the number of network members is large. Indeed, the following result holds.

Proposition 8 *For any set of parameters μ, p, β, there exists $m^* = m^*(\mu, p, \beta)$ such that in any equilibrium in the complete two-way network X_m with m agents, the set of subscribers is a singleton, $|S(X_m)| = 1$.*

Proof. Suppose that there is more than one subscriber in an equilibrium in X_m, $m \geq 2$. Take any $x \in S(X_m)$. The probability that information reaches x cannot be lower than $g(m) = 1-(1-p)(1-p^2)^{m-2}$. Indeed, with probability $1-p$, x does not receive information from another subscriber (by assumption, there exists at least one) and with $(1-p^2)^{m-2}$ x does not receive information from any other non-subscriber (p^2 is that probability that x receives information from the other subscriber through a particular non-subscriber.) For $x \in S(X_m)$, the IC constraint is

$$V(\beta) < \frac{c}{1-g(m)} = \frac{c}{(1-p)(1-p^2)^{m-2}},$$

which cannot hold for large m. ∎

Proposition 8 implies that when m is large, the optimal level of propaganda in a complete network is $\beta^{\max}$ as the only constraint is the unique subscriber's IC constraint. However, this is not true for every m as Example 5 demonstrates: for low p, when $m=3$, the optimal level of propaganda might make two or even all three agents subscribers.

23

Not surprisingly, if m is constant, and p goes to one, the set of subscribers is, again, a singleton. This is a corollary to the celebrated Erdős-Rényi theorem: in any complete graph, the probability that the network is connected tends to 1 as long as the penetration probability p exceeds the threshold $\frac{\ln n}{n}$ (Erdős and Rényi, 1961, Jackson, 2008, Th. 4.1). This argument might be extended to any two-way finite network, also showing the existence of equilibria for any β, when p is close enough to 1.

Proposition 9 *Let X be any two-way (connected) finite network. Then, for any β, there exists $\overline{p} = \overline{p}(\beta) \in (0,1)$ such that for any $p > \overline{p}$, in any equilibrium the set of subscribers is a singleton.*

Proof. Suppose not, i.e. there are at least two subscribers, x and y. Let $f(p)$ be the probability that x's information reaches y. $f(p)$ is (obviously) an increasing function of p with $f(0) = 0$ and $f(1) = 1$. Thus, for any $c > 0$, the IC condition $f(p)V(\beta) < V(\beta) - c$ will be violated with p approaching 1. ∎

Note that Proposition 9 cannot be extended to an infinite network: e.g., it is not true in the case of an infinite linear network that we considered in Section 4. Still, with p approaching 1, the length of a subscription cell on $X = \mathbb{Z}$ or $X = \mathbb{T}_r$ tends to infinity for any β.

Consider a sequence of lattices $\mathbb{Z}^m$, $m = 1, 2, ...,$ with links $\{((x_1, ...x_l, ..., x_m), (x_1, ...x_l + 1, ..., x_m)) \,|\, (x_1, ...x_l, ..., x_m) \in \mathbb{Z}^m\}$. Consider the set of subscribers

$$S^m = \{(2l_1, 2l_2, ..., 2l_m), (2l_1 + 1, 2l_2 + 1, ..., 2l_m + 1) \,|\, l_1, l_2, ..., l_m \in \mathbb{Z}\}. \tag{4}$$

(Each line in $\mathbb{Z}^m$ is a linear network with $K = 2$. See Figure 9 for S_1 and S_2.)
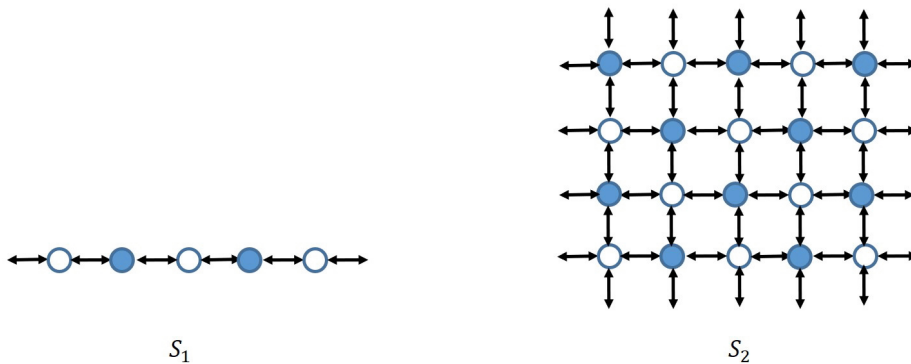

S_1 \qquad\qquad\qquad\qquad\qquad S_2

Figure 9: Sets of subscribers S_1 and S_2 in $\mathbb{Z}$ and $\mathbb{Z}^2$, respectively, result in the same frequency of subscribers, $\frac{1}{2}$, but the maximum level of propaganda consistent with S_2 is lower.

Proposition 10 *Let $X = \mathbb{Z}^m$, be two-way, and let the set S^m be defined by (4), and let β_m be the level of propaganda that makes S_m the equilibrium subscriber set for each m and guarantees the same average frequency of subscribers of $\frac{1}{2}$. Then β_m is a (strictly) decreasing function of m.*[3]

While Proposition 10 is formulated for the average frequency of subscribers of $\frac{1}{2}$, the results should be true for other frequencies as well. The reason is that whenever $n > m$, each agent has more links to receive information from in $\mathbb{Z}^n$ than in $\mathbb{Z}^m$, and so the IC constraint is satisfied for a lower level of propaganda.

7 Persuasion and Information Percolation

In Section 4, we built a theory of persuasion on linear networks, and then extended it to directed tree networks in Section **??**. However, it appears that a small change in network dimensionality leads to drastically different results with respect to persuasion. This is a result of "percolation" (Duminil-Copin, 2018), the effect that is not present on linear networks, yet play an important role whenever the network is a tree or is similar (locally) to two-dimensional integer lattice.

Take any (connected) network X. To analyze equilibrium subscription structures on X, let C denote the (random) connected component of the origin, the unique subscriber, and define the *percolation function* $\theta(p) = P(|C| = \infty)$, the probability that an infinite cluster forms. Then, on many networks, there exists a critical probability $p_c = p_c(X) \in (0, 1)$ such that for any $p < p_c$, $\theta(p) = 0$ and for any $p > p_c$, $\theta(p) > 0$. For r-regular trees and multidimensional integer lattices, the function $\theta(p)$ is continuous on $[0, 1]$ and is increasing on $(p_c, 1]$.[4] Another standard characteristics is the expected cluster size, given by the susceptibility function $\chi(p) = E_p(|C|)$. Of course, whenever $p > p_c$, $\chi(p) = \infty$.

In this Section, we use results from percolation theory to discuss the relationship between p and the sender-optimal subscription structures on regular trees and $\mathbb{Z}^m$.

---

[3]The IC constraint for the nonsubscriber is $\left(1 - (1 - p)^{2m}\right) V < V - c$, which is impossible to satisfy simultaneously for every $c > 0$ and every $m \in \mathbb{N}$. That is, the fully rigorous statement should be as follows. Let $\overline{m} = \overline{m}(\beta, c, p) = \arg\max \left\{ m | c < V(\beta)(1 - p)^{2m} \right\}$. Given c, β, and p, β_m is a strictly decreasing function on $[0, \overline{m}]$.

[4]Percolation in the critical case of $p = p_c$ is subject to active research (Duminil-Copin, 2018).

7.1 Persuasion on a Regular Tree

If $X = \mathbb{T}_r$, then $p_c(X) = \frac{1}{r-1}$ and the susceptibility

$$\chi(p) = E_p(|C|) = \frac{1+p}{1-(r-1)p}$$

if $p < \frac{1}{r-1}$ and $\chi(p) = \infty$ otherwise. This allows to obtain the following estimate for any graph for which the degree of every node is bounded by r (van der Hofstadt, 2010, Theorem 1.2) For $p < \frac{1}{r-1}$, for every subscriber x, the expected size of the subscription cluster is finite $E_p(|C(x)|) < \infty$.

So, if $p < \frac{1}{r-1}$, then the tree does not contain, in expectation, any infinite clusters. Thus, setting $\beta^* = \beta^{\max}$ and allowing the agent at the root of the tree to be a subscriber results in the zero expected average amount of action, while the optimal solution of Section 5 always results in a positive amount. If $p > \frac{1}{r-1}$, then X contains infinitely many clusters (except for the case when $p = 1$). When $p \to 1$, the expected average amount of action converges (from above) to the case of one subscriber at the root of the tree.

7.2 Persuasion on $\mathbb{Z}^m$

Consider $X = \mathbb{Z}^m$, two-way. For $p < p_c(X)$, the probability that one subscriber results in an infinite cluster is $\theta(p) = 0$, which means that the regular structure described in Propositions 6 and 10 results in a higher expected average action than any one-subscriber structure. For $p > p_c(X)$, this is less clear: with probability $\theta(p) > 0$, the cluster $C(0)$ is infinite; in this case, it is possible that choosing $\beta^{\max}$ results in a higher expected average action than in any regular structure.

Consider $X = \mathbb{Z}^2$, two-way, and the subscription structure S_2 (see Figure 9). Each neighbor y of any $x \in S_2$ receives information (not counting information from x) with probability $1 - (1-p)^3$. Agent x receives information from y with probability $p\left(1 - (1-p)^3\right)$. Thus, agent x receives information from someone with probability $f(p) = 1 - \left(1 - p\left(1 - (1-p)^3\right)\right)^4$, which is an increasing function of p with $f(0) = 0$ and $f(1) = 1$. The IC constraint of the subscriber is satisfied if $f(p)V(\beta) \leq V(\beta) - c$, and the expected average amount of action is $\frac{1}{2}(1+p)\frac{1}{q}\left(\mu - \frac{c}{1-f(p)}\right)$, which is non-negative. So, with $p < \frac{1}{2}$, this structure is preferable to that of a finite number of subscribers with $\beta^{\max}$ $(p_c(\mathbb{Z}^2) = \frac{1}{2}$, Kesten, 1980).

If $p > p_c(\mathbb{Z}^2) = \frac{1}{2}$, then S_2 is stable as long as

$$1 - \left(1 - p\left(1 - (1-p)^3\right)\right)^4 \leq 1 - \frac{c}{V(\beta)}. \tag{5}$$

Inequality (5) cannot be true for any pair (c, p), $c > 0$, $p > 0$, yet for any $p < 1$, there exists a $c(p) > 0$ such that (5) holds (and so S_2 is stable) for any $c < c(p)$.

When p is below the critical threshold $p_c\left(\mathbb{Z}^d\right)$, there cannot be an equilibrium with a finite number of subscribers for any $\beta < \beta^{\max}$. Indeed, each subscriber generates a cluster of second-hand receivers of a finite size (almost surely). Thus, there is always an agent far enough from the any finite group of subscribers that would prefer to subscribe.

When p is above the critical threshold $p_c\left(\mathbb{Z}^d\right)$, it is possible to have a finite group of subscribers when $\beta < \beta^{\max}$. Indeed, suppose that

$$\theta\left(p\right) V\left(\beta\right) > V\left(\beta\right) - c, \tag{6}$$

where $\theta\left(p\right)$ is the probability that an infinite cluster forms; $\theta\left(p\right)$ continuously maps $\left[\frac{1}{2}, 1\right]$ onto $[0, 1]$. As $p > p_c\left(\mathbb{Z}^d\right)$, the ergodicity implies that for any node $x \in \mathbb{Z}^d$, $\theta\left(p\right)$ is the probability that x belongs to the cluster that originates with the single subscriber. That is, the probability that information reaches x is at least $\theta\left(p\right)$, which makes her unwilling to subscribe. Therefore, in $X = \mathbb{Z}^d$, $d \geq 2$, we have an equilibrium with a unique subscriber for any $\beta < \beta^{\max}$ such that $\theta\left(p\right) V\left(\beta\right) > V\left(\beta\right) - c$. This stands in a sharp contrast with the case of $X = \mathbb{Z}$, in which a unique subscriber is possible only when $\beta = \beta^{\max}$. In any network, if the choice of the sender is over equilibrium structures with a unique subscriber, the sender-optimal choice is $\beta^{\max}$.

If $p > p_c\left(\mathbb{Z}^d\right)$, but (6) is not fulfilled, then finite equilibrium groups of subscribers with $\beta < \beta^{\max}$ do not exist for another reason. Indeed, suppose there is a finite number of subscribers. Each of them generates an infinite cluster with probability $\theta\left(p\right)$. Because of the Kolmogorov zero-one law, there exists, almost surely, exactly one infinite cluster. For any non-subscriber x the probability of getting information is close to $\theta\left(p\right)$, the probability to belong to the unique infinite cluster, which means that she will be willing to be another subscriber. Thus, for any finite number of subscribers, there always exists a non-subscriber who would want to subscribe. As a result, any equilibrium should include an infinite number of subscribers. For a certain combination of parameters, there exists an equilibrium (see S_2 on Picture 9). Whether or not there exists an equilibrium subscription structure for any set of fundamental parameters is an open question.

8 Optimal Propaganda

Propaganda and censorship is considered a major tool in any autocrat's disposal (Besley and Prat, 2005, Guriev, Egorov, and Sonin, 2009, Lorentzen, 2013, Guriev and Triesman, 2019). Still,

even in a totalitarian dictatorship, which shuts down all alternative channels of information, a major limit on the impact of propaganda stems from incentive compatibility. Not surprisingly, the states that have extensively used multiple propaganda campaigns have almost exclusively been states that consistently persecuted people for information exchange outside of the state purview. Over the years, the Soviet Union has engaged in mass campaigns to expand literacy, reduce Orthodox church participation, mobilize youth to participate in development programs in Siberia or Kazakhstan, reduce alcohol consumption and embrace a healthy lifestyle, increase and, in other time periods, decrease the birth rate, and many others (Service, 2005). In the communist China, mass campaigns were used, most notoriously, to promote development of individual industrial plants, encourage extermination of rats, flies, mosquitoes, and sparrows, or spearhead destruction of cultural artifacts (Shapiro, 2001).

In the democratic realm, governments have often used propaganda campaigns to mobilize population during wars. For example, during World War I, the British government has been using production of literature and films to increase the support for war both domestically and in the (still neutral) United States (Taylor, 1999). Currently, the US government finance efforts to nudge people to sign up for the health insurance provided by the Affordable Care Act (Thaler and Sunstein, 2009). Even more information operations are carried out outside the US. In Afghanistan, the US government has successfully organized information operations to encourage local population to report road-side bombs, the deadliest weapon used against government and international forces and increase support for the reintegration of former Taliban fighters (Sonin and Wright, 2019).

Our first obvservation about propaganda is almost trivial. If the sender takes the probability p as given, her optimal choice is a function of p. In the case of the linear network, an increase in p leads to a higher optimal level of propaganda. However, if the sender has a choice of p, then she will prefer, unambiguously, as low level of p as possible.

Proposition 11 *Suppose that the government chooses both p, the probability that information passes through the network, and the level of propaganda β, to maximize the expected amount of action. Then its optimal choice of p is either $p^* = 0$ or $p^* = 1$, and the propaganda level is at its maximum:*

$$\beta^* = \beta^{\max} = \frac{\mu(1-q) - c}{(1-\mu)q}.$$

Proof. By choosing $p^* = 0$, the government forces everyone to buy subscription. Then, choosing $\beta^* = \beta^{\max}$ maximizes the expected average amount of individual action, which is equal to $\mu + \beta^*(1-\mu) = \mu + \frac{c - \mu(1-q)}{q} = \frac{1}{q}(\mu - c)$. Alternatively, $p^* = 1$, and in a two-way network there

28

will be a single subscriber. In a directed network, there might be more than one subscriber, but with $p^* = 1$ information from one of them cannot reach another subscriber. Again, choosing $\beta^* = \beta^{\max}$ maximizes the expected average amount of individual action. ■

Both of the two extreme penetration probabilities, $p^* = 0$ or $p^* = 1$, result in a sender-optimal equilibrium. Still, the regime of $p^* = 0$, under which the government chooses to shut up all the channels of information exchange, forcing everyone to subscribe to propaganda, is different from that of $p^* = 1$, the regime of unrestricted networking. Let us focus on a two-way network. Suppose that the opportunity cost of subscription for each agent $x \in X$ is a random variable. Then, with $p^* = 0$, choosing any β^* would result in a different fraction of agents becoming subscribers. With $p^* = 1$, the impact of propaganda is higher: optimal β^* should maximize the probability that the minimum of agents' costs is such that the incentive constraint for the subscriber is satisfied.

Even if the government chooses p, the impact is still limited by agents intrinsic aversion to do the sender-preferred action $a = 1$, q, or by the *ex ante* probability μ that they attach to the state when action $a = 1$ is receiver-preferred. Trivially, if the government has means to lower c, the cost of subscription, the impact will always be higher.

What is the optimal structure of propaganda consumption when the government cannot change p? Consider a simple star network, i.e., a tree network with a single root and branches of length 1. Locally, any network in which there cannot be more than one link between nodes is a star network. The naive intuition would suggest that the sender would always prefer that the central agent is a subscriber and then information is "spread" to periphery. However, this argument does not take into account the negative externality: when the central agent subscribes, this crowds out the peripherals' incentives to subscribe (and thus diminishes the probability of action).

Proposition 12 *Let X be a simple star network with $r + 1$ nodes.*

(i) There exists a pair of thresholds $p_{\min,1}, p_{\min,2} > 0$ such that for any $p \in [p_{\min,1}, p_{\min,2}]$, in the sender-optimal equilibrium, the set of subscribers consists of all peripheral agents. The level of propaganda $\beta^* = \beta^*(p)$ is a strictly decreasing function on $[p_{\min,1}, p_{\min,2}]$; $\beta^*(0) = \beta^{\max}$.

(ii) There exists a $p_{\max} < 1$ such that for any $p \in [p_{\max}, 1]$, in the sender-optimal equilibrium, the set of subscribers is a singleton with the central agent being the only subscriber; $\beta^* = \beta^{\max}$ on $[p_{\max}, 1]$.

(iii) The set of subscribers in the sender-optimal equilibrium (weakly) shrinks with p.

Let us sketch the proof of Proposition 12. When all peripheral agents is the set of subscribers,

the probability that a peripheral agent receives information from others is $p\left(1 - (1-p)^{r-1}\right)$, where $1 - (1-p)^{r-1}$ is the probability that the central agent received a signal from the other $r-1$ peripheral agents. So, the IC constraint for the subscriber (as always, the IC constraint is an equality as we focus on sender-optimal equilibria) yields the probability that the signal is $s = 1$ of $\mu + (1 - \mu)\beta = \frac{1}{q}\left(\mu - \frac{c}{1 - p\left(1 - (1-p)^{r-1}\right)}\right)$, and the total expected amount of action is equal to $(r + (1 - (1-p)^r))\frac{1}{q}\left(\mu - \frac{c}{1 - p\left(1 - (1-p)^{r-1}\right)}\right)$. Then, the threshold $p_{\min,1}$ is determined by the equation

$$(r+1)\left(\mu - c\,(1-p)^{-r}\right) = (r + (1 - (1-p)^r))\left(\mu - \frac{c}{1 - p\left(1 - (1-p)^{r-1}\right)}\right), \qquad (7)$$

in which the left-hand side is derived for the situation when all $r+1$ agents are subscribers, which is sender-optimal when $p = 0$, and the relevant IC constraint is that of the central agent as she is the most likely to receive information.

It is a straightforward, though tedious, exercise to prove that there exists a solution $p_{\min,1}$ to this equation, and then to find $p_{\min,2}$ using the IC constraint in the situation of $r-1$ peripheral subscribers. The IC constraint in the left-hand side on (7) implies that

$$p_{\min,1} \leq 1 - (c/\mu)^{\frac{1}{r}},$$

which is decreasing in r.

To prove (ii), one may use Proposition 9 that guarantees a unique subscriber in a finite two-way network when p is close to 1. Choosing the central agent to be the subscriber is sender-optimal.

Analyzing the IC constraints in a star network, it is natural to formulate the following conjecture.

Conjecture 3 *The thresholds $p_{\min,1}$ and $p_{\max}(r)$ decrease with r.*

Proposition 12 extends to any finite two-way tree. When p is close to 1, there is no chance to satisfy the IC constraint for more than one subscriber. In this situation, it is better to have the central agent to subscribe. When $p = 0$, then everyone is a subscriber; with p increasing from zero, the nodes with high centrality are the first to drop from the subscriber set. With one-way (directed) tree, the results are essentially the same, but such a tree might have more than one "root", a node with a single outgoing link. In such a case, there might be more than

one subscriber even if $\beta^* = \beta^{\max}$.[5]

Proposition 12 challenges the conventional wisdom on the proper target of a government propaganda intervention. The effect is not confined to our particular setup: e.g., suppose that the government is choosing to allocate one free subscription in a star network, and the cost of subscription c_x is random, say $c_x \sim U[0, 2c]$. When p is high, the optimal allocation of the free subscription is the central agent. However, if p is relatively low, the government has to take into account the crowding-out effect. Allocating free subscription to a peripheral agent will have less of a crowding out effect: there will be a higher probability that peripheral agents become subscribers on their own. This result extends to any network: provided that an equilibrium exists, agents with high degree of centrality may not be the optimal direct receivers from the sender's standpoint. Instead, she would prefer to have agents with low degree of centrality to be the subscribers.

The above results allows to discuss the relationship between the network structure and the government's ability to maintain support. Consider networks in Figure 3. In both polities P_1 and P_1' each agent is connected to two others, yet the network architecture is different, and the government's ability to maintain support depends on both the architecture and p. In polities P_2 and P_2', each agent has four connections. In a sender-optimal equilibrium, it is typically, although not always, easier to influence agents' actions and maintain government support in P_1 and P_2. In particular, there exists a threshold $\bar{p}$ such that for any $p > \bar{p}$, the government survives for a wider range of parameters with polity P_1 than with polity P_1' and with P_2 rather than P_2'. By Proposition 9, when p is high enough (not necessarily close to 1), each connected component will have a unique subscriber and the sender-optimal level of propaganda will be $\beta^*(p) = \beta^{\max}$ in both in P_1 and P_2. In contrast, for P_1', it is never optimal to have propaganda at its maximum unless $p = 1$. For P_2', there exists another threshold $\widetilde{p} < 1$, such that $\beta^*(p) = \beta^{\max}$ for any $p > \widetilde{p}$.

Galperti and Perego (2019) demonstrate that a "deeper" network worsens the designer ability to manipulate beliefs of the networked receivers, with a complete network being the deepest possible network. In our random graph setup, the impact of network's "depth" depends, non-monotonically, on p. While for sufficiently high probabilities p, the sender prefers P_1 (a locally more deep network) to P_1' and P_2 to P_2', there is a range of ps for which the sender's preference is reversed. If the government has the ability to alter the network architecture (keeping the

---

[5]It is possible to have an equilibrium with more than one subscriber even if the level of propaganda is $\beta^{\max}$. Consider network X with three nodes $\{1, 2, 3\}$ and two one-way links $(1, 2)$ and $(3, 2)$. For any p exceeding some positive threshold $\bar{p}$, the sender-optimal equilibrim involves $S(X) = \{1, 2\}$ and the propaganda level of $\beta^{\max}$.

number of links constant), the following simple observation is true. For any p, there exists a division of network P_1 into m-nodes wheels (P_2 consists of 3-nodes wheels) so that the resulting structure (weakly) increases the amount of support that the government receives.

Finally, our main results are robust to the amount of information about the structure of network that the agents possess. In our model, the agents who make the subscription decision know the exact structure of the network. However, the main insights will stay intact if agents would know their own number of links and the distribution of links overall. (Such models are discussed in, e.g., Galeotti, Goyal, Fernando-Redondo, Jackson, and Yariv, 2010, and Sadler, 2019.) Unlike in the local public goods games (Bramoullé and Kranton, 2007, Galeotti, Goyal, Fernando-Redondo, Jackson, and Yariv, 2010), agent's decision to subscriber in equilibrium is not necessarily monotonic in the agent's degree. Still, when p is sufficiently low, it is monotonic in the sender-optimal equilibrium, which looks as follows: all agents with a node degree smaller than some threshold K^* will be subscribers, those agents with node degrees higher than K^* will be non-subscribers, and the agents with the degree equal to K^* will subscribe with a certain probability. As in our main model, agents with a high node centrality are not sender-preferred targets of propaganda.

9   Conclusion

We consider an environment in which agents that are influenced by a sender who committed to a certain information design face the following trade-off: subscribe, as a cost, to the news source and be sure to receive payoff relevant information or rely on network neighbors which might or might not pass the information. Our setup is the simplest possible model of Bayesian persuasion with multiple receivers who are linked in a network. When the network is dense, optimal propaganda needs to be less biased, because otherwise the subscription rate would fall. The sender might be interested in peripheral, rather than centrally connected agents subscribing to information. Adding or removing links in the network results in a higher or lower propaganda impact depending on the penetration rate for the network.

References

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

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya (2015) Radio and the Rise of The Nazis in Prewar Germany, *Quarterly Journal of Economics*, 130(4), 1885-1939.

Akbarpour, Mohammad and Matthew Jackson (2018) Diffusion in networks and the virtue of burstiness, *Proceedings of the National Academy of Sciences*, 115 (30), E6996-E7004.

Akbarpour, Mohammad, Suraj Malladi, and Amin Saberi (2018) Just a Few Seeds More: Value of Network Information for Diffusion, mimeo.

Alonso, Ricardo and Odilon Camara (2016) Persuading Voters, *American Economic Review*, 106, 3590-3605.

Aumann, Robert J., and Michael B. Maschler (1995) *Repeat ed Games with Incomplete Information*, Cambridge, MA: MIT Press

Bala, Venkatesh and Sanjeev Goyal (2000) A Noncooperative Model of Network Formation, *Econometrica*, 68(5), 1181-1230.

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

Banerjee, Abhijit, Arun Chandrasekhar, Esther Duflo, and Matthew Jackson (2018) Changes in Social Network Structure in Response to Exposure to Formal Credit Markets, mimeo.

Bloch, Francis, Matthew Jackson, and Pietro Tebaldi (2019) Centrality Measures in Networks, mimeo.

Bergemann, Dirk and Stephen Morris (2019) Information Design, *Journal of Economic Literature,* 57(1), 44-95.

Bergemann, Dirk and Stephen Morris (2013) Robust Predictions in Games With Incomplete Information, *Econometrica*, 81, 1251-1308.

Bramoullé, Yann and Rachelle Kranton (2007) Strategic Experimentation in Networks, *Journal of Economic Theory*, 135(1), 478-494.

Broadbent, S.R. and J. M. Hammersley (1957) Percolation processes. I. Crystals and mazes, Proceedings of the Cambridge Philosophical Society, 629–641.

Brocas, Isabelle and Juan Carrillo (2007) Influence through Ignorance. *RAND Journal of Economics.* 38(4):931-947.

Candogan, Ozan (2019) Persuasion in Networks: Public Signals and k-Cores, mimeo.

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang (2018) Pivotal Persuasion, mimeo.

Chen, Heng, Yang K. Lu, and Wing Suen (2016) The Power of Whispers: A Theory of Rumor, Communication and Revolution, *International Economic Review,* 57 (1): 89-116.

Chen, Yuyu, and David Yang (2018) The Impact of Media Censorship: Evidence from a Field Experiment in China, mimeo.

Chwe, Michael Suk-Young (2000) Communication and Coordination in Social Networks, *Review of Economic Studies*, 67, 1-16.

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

Duminil-Copin, Hugo (2018) Sixty Years of Percolation, *Proceeding of the International Congress of Mathematicians*, Rio de Janeiro, Vol. 3, 2817–2844.

Galeotti, Andrea, Sanjeev Goyal, Matthew Jackson, Fernando Vega-Redondo, and Leeat Yariv (2010), Network Games, *Review of Economic Studies*, 77(1), 218-244.

Galperti, Simone and Jacopo Perego (2019) Belief Meddling in Social Networks: An Information-Design Approach, mimeo.

Gehlbach, Scott and Konstantin Sonin (2014) Government Control of the Media, *Journal of Public Economics*, 118, 163-171.

Gentzkow, Matthew and Emir Kamenica (2014) Costly Persuasion, *American Economic Review Papers and Proceedings,* 104(5): 457-462.

Golub, Benjamin and Matthew Jackson (2010) Naïve Learning in Social Networks and the Wisdom of Crowds, *American Economic Journal: Microeconomics*, 2:1, 112–149.

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

Jackson, Matthew, Suraj Malladi, and David McAdams (2018) Learning Through the Grapevine: The Impact of Message Mutation, Transmission Failure, and Deliberate Bias, mimeo.

Jackson, Matthew, Brian Rogers, and Yves Zenou (2019) Networks: An Economic Perspective, in *Oxford Handbook of Social Network Analysis*, forthcoming.

Kamenica, Emir and Matthew Gentzkow (2011) Bayesian Persuasion, *American Economic Review*, 101, 2590-2615.

Kamenica, Emir (2018) Bayesian Persuasion and Information Design, *Annual Review of Economics,*

Kesten, Harry (1980). The Critical Probability of Bond Percolation on the Square Lattice Equals 1/2. *Communications in Mathematical Physics.* 74(1), 41-59.

King, Gary, Jennifer Pan, and Margaret Roberts (2013) "How Censorship in China Allows Government Criticism, but Silences Collective Expression." *American Political Science Review* 107 (2): 326-43.

King, Gary, Jennifer Pan, and Margaret Roberts (2014) "Reverse-engineering Censorship in China: Randomized Experimentation and Participant Observation." *Science* 345, 1251722.

Lorentzen, Peter. 2013. "Regularizing Rioting: Permitting Public Protest in an Authoritarian Regime." *Quarterly Journal of Political Science* 8: 127-58.

Lorentzen, Peter. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58 (2): 402-14.

Kolotilin, Anton (2015) Experimental Design to Persuade, *Games and Economic Behavior*, 90, 215–226.

Kolotilin, Anton (2018) Optimal Information Disclosure: A Linear Programming Approach, *Theoretical Economics,* 13: 607-635.

Kosterina, Svetlana (2018) Persuasion with unknown beliefs, mimeo.

Milgrom, Paul (1981) Good news and bad news: Representation theorems and applications, *Bell Journal of Economics.* 12(2): 380-291.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. (2017) On Information Design in Games, mimeo.

Ostrovsky, Michael and Michael Schwarz (2010) Information Disclosure and Unraveling in Matching Markets, *American Economic Journal: Microeconomics*, 2, 34–63.

Prat, Andrea and David Strömberg (2013) The Political Economy of Mass Media, in *Advances in Economics and Econometrics*, Cambridge Univ. Press, 135-187.

Rayo, Luis and Ilya Segal (2010) Optimal Information Disclosure, *Journal of Political Economy*, 118, 949-987.

Sadler, Evan (2019) Diffusion Games, mimeo.

Satyanath, Shanker, Nico Voigtländer, and Hans-Joachim Voth (2017) Bowling for Fascism: Social Capital and the Rise of the Nazi Party, *Journal of Political Economy,* 125:2, 478-526.

Service, Robert (2005) *A History of Modern Russia, from Nicholas II to Putin,* Harvard Univ. Press, Cambridge, MA.

Shapiro, Judith Rae (2001) *Mao's War Against Nature: Politics and the Environment in Revolutionary China,* Cambridge Univ. Press, UK.

Sims, Christopher (2003) Implications of Rational Inattention, *Journal of Monetary Economics.* 50: 665-690.

Sonin, Konstantin and Austin Wright (2019) Information Operations Increase Civilian Security Cooperation, mimeo.

Spence, Michael (1973) Job Market Signaling, *Quarterly Journal of Economics*, 87(3), 355-374.

Taneva, Ina (2016) Information Design, mimeo.

Taylor, Philip (1999) *British Propaganda in the 20th Century*, Edinburgh University Press, UK.

Thaler, Richard and Cass Sunstein (2009) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Penguin Books, NY.

Van Der Hofstad, Remco (2010) Percolation and Random Graphs, in *New Perspectives on Stochastic Geometry*, Elsevier, Amsterdam, 173-247.

Voigtlander, Nico and Hans-Joachim Voth (2015) "Nazi indoctrination and anti-Semitic beliefs in Germany, *Proceedings of the National Academy of Sciences of the United States of America* 112 (26):7931–7936.

Wang, Yun (2015) Bayesian Persuasion with Multiple Receivers, mimeo.

Yanagizawa-Drott, David (2014) Propaganda and Conflict: Evidence from the Rwandan Genocide, *Quarterly Journal of Economics*, 129 (4):1947–1994.

Appendix

Algebra for Example 1

There are 3 possible cases.

(1) If the central agent is a subscriber and the peripheral agents receive information from her, then for the optimal level of propaganda $\beta^* = \beta^{\max}$, one has $\mu + (1 - \mu)\beta^* = \frac{1}{q}(\mu - c)$, and the expected total action is

$$\frac{1}{q}(\mu - c)(1 + 2p). \tag{A1}$$

(2) If the peripheral agents are subscribers, and the central agent is the second-hand receiver, than the optimal amount of propaganda satisfies $p^2 V(\beta) = V(\beta) - c$ (otherwise, one of the two peripheral agents would drop subscription) so that $\mu + (1 - \mu)\beta^* = \frac{1}{q}\left(\mu - \frac{c}{1-p^2}\right)$. Then, the expected total action is

$$\frac{1}{q}\left(\mu - \frac{c}{1 - p^2}\right)(2 + (2 - p)p). \tag{A2}$$

(The central agent receives information with probability $1 - (1 - p)^2 = (2 - p)p > p^2$.).

It is straightforward, that (A2) exceeds (A1) when p is low and vice versa if p is relatively larger.

(3) All three agents are subscribers. Then, β should be such that $(2 - p)pV(\beta) \leq V(\beta) - c$ as the central agent is the least willing to subscribe. Then $\mu + (1 - \mu)\beta^* = \frac{1}{q}\left(\mu - \frac{c}{(1-p)^2}\right)$. Then the expected amount of action is

$$\frac{1}{q}\left(\mu - \frac{c}{(1 - p)^2}\right)3. \tag{A3}$$

Denote $a = c/\mu$ and compare three functions:

$$
\begin{aligned}
f_1(p) &= (1 + 2p)(1 - a), \\
f_2(p) &= (2 + (2 - p)p)\left(1 - \frac{a}{1 - p^2}\right), \\
f_3(p) &= 3\left(1 - \frac{a}{(1 - p)^2}\right).
\end{aligned}
$$

Observe that $f_1(0) < f_2(0) < f_3(0)$, $f_3'(p) < 0$ for any $p \in (0, 1)$, $f_2'(0) > 0$, and that $\lim_{p \to 1} f_2 = \lim_{p \to 1} f_3 = -\infty$ for any a [that satisfies the standard conditions in the model]. This implies that there exist thresholds p_1 and p_2 such that for any $p \in (0, p_1)$, (A3) provides the highest pay-off for the sender and for any $p \in (p_2, 1)$, the sender's payoff is (A1).

Let's consider $p = \frac{1}{2}$. The value of each of the three functions is a function of parameter a:

$$f_1\left(\frac{1}{2}\right) = 2 - 2a,$$

$$f_2\left(\frac{1}{2}\right) = \frac{11}{4} - \frac{11a}{3},$$

$$f_3\left(\frac{1}{2}\right) = 3 - 12a.$$

It is straightforward to check that if $a \in \left(\frac{3}{100}, \frac{9}{200}\right)$, $f_2\left(\frac{1}{2}\right) > \max\left\{f_1\left(\frac{1}{2}\right), f_3\left(\frac{1}{2}\right)\right\}$. $\blacksquare$

Lemma A1 *Suppose that $a \in (0,1)$. Then function $\varphi(t) = \frac{1}{t}\left(a - p^t\right)$ is single-peaked on $[0, +\infty)$.*

Proof. Take a derivative:

$$\varphi'(t) = \frac{-p^t t \ln p - a + p^t}{t^2}.$$

Let $\phi(t) = \frac{-p^t t \ln p - a + p^t}{p^t} = -t \ln p - a \left(\frac{1}{p}\right)^t + 1$. It is sufficient to show that there exists some t_0 such that $\phi(t) > 0$ for any $t \in [0, t_0)$ and $\phi(t) < 0$ for any $t \in (t_0, +\infty)$. Since $p < 1$, $1 - a\left(\frac{1}{p}\right)^t$ is a negative of an exponential function; it crosses y-axis at $1 - a > 0$. The linear function $-t \ln p$ crosses y-axis at 0 (below $1 - a$) and then have a single crossing with $1 - a\left(\frac{1}{p}\right)^t$. $\blacksquare$

Lemma A2 *Function $\frac{1}{t}\left(\mu\left(1 - p^t\right) - c\right)$ satisfies increasing differences condition in (t, p) and (t, c).*

Proof.

$$\frac{\partial}{\partial p}\frac{1}{t}\left(\mu\left(1 - p^t\right) - c\right) = \frac{1}{t}\left(-\mu p^{t-1}t\right) = -\mu p^{t-1},$$

$$\frac{\partial^2}{\partial p \partial t}\frac{1}{t}\left(\mu\left(1 - p^t\right) - c\right) = -p^{t-1}\mu \ln p > 0;$$

$$\frac{\partial}{\partial c}\frac{1}{t}\left(\mu\left(1 - p^t\right) - c\right) = -\frac{1}{t},$$

$$\frac{\partial^2}{\partial c \partial t}\frac{1}{t}\left(\mu\left(1 - p^t\right) - c\right) = \frac{1}{t^2} > 0.$$

$\blacksquare$

Lemma A3 *Suppose that $rp > 1$. Then there exists a threshold $\bar{t} \in (0,1)$ such that the function $\varphi(t, r) = \frac{(rp)^t - 1}{(r^t - 1)}\left(1 - \frac{c/\mu}{1 - p^t}\right)$ increases on $(0, \bar{t})$ and decreases on $(\bar{t}, 1)$.*

Proof.

$$\frac{d}{dt}\frac{(rp)^t - 1}{(r^t - 1)} = \frac{\left((rp)^t - 1\right)(\ln rp - \ln r)}{(r^t - 1)}$$

$$= \frac{\left((rp)^t - 1\right)\ln p}{(r^t - 1)}$$

$$\frac{d}{dt}\left(1 - \frac{c/\mu}{1 - p^t}\right) = -(c/\mu)\frac{p^t \ln p}{(1 - p^t)^2}$$

$$\frac{d}{dt}\frac{(rp)^t - 1}{(r^t - 1)}\left(1 - \frac{c/\mu}{1 - p^t}\right) = \frac{\left((rp)^t - 1\right)\ln p}{(r^t - 1)}\left(1 - \frac{c/\mu}{1 - p^t}\right) - \frac{(rp)^t - 1}{(r^t - 1)}(c/\mu)\frac{p^t \ln p}{(1 - p^t)^2}$$

The above derivative is positive if and only if

$$\left((rp)^t - 1\right)\ln p\left(1 - \frac{c/\mu}{1 - p^t}\right) - \left((rp)^t - 1\right)(c/\mu)\frac{p^t \ln p}{(1 - p^t)^2} > 0$$

as $r^t - 1 > 0$ for any $t > 0$. This is equivalent to

$$\left((rp)^t - 1\right)\left(1 - \frac{c/\mu}{1 - p^t}\right) - \left((rp)^t - 1\right)(c/\mu)\frac{p^t}{(1 - p^t)^2} < 0$$

as $\ln p < 0$. Finally, the above condition is equivalent to

$$\left(1 - \frac{c/\mu}{1 - p^t}\right) - (c/\mu)\frac{p^t}{(1 - p^t)^2} < 0$$

as long as $(rp)^t - 1 > 0$, which in turn is true *if and only if* $rp > 1$.

$$\left(1 - p^t\right)^2 - (c/\mu)\left(1 - p^t\right) - (c/\mu)\,p^t < 0$$

$$\left(1 - p^t\right)^2 < c/\mu.$$

The LHS is monotonically increasing in t (and asymptotically approach 1) on $(0, 1)$. ∎

Lemma A4 *Suppose that $rp > 1$. Then function $\varphi(t) = \varphi(t|r, c, \mu)$ satisfies increasing differences in (t, c), $(t, -\mu)$, and (t, r).*

Proof. Let us demonstrate that

$$\frac{d^2}{dt\,d(c/\mu)}\frac{(rp)^t - 1}{(r^t - 1)}\left(1 - \frac{c/\mu}{1 - p^t}\right) > 0.$$

$$\frac{d}{d(c/\mu)}\frac{(rp)^t - 1}{(r^t - 1)}\left(1 - \frac{c/\mu}{1 - p^t}\right) = -\frac{(rp)^t - 1}{(r^t - 1)}\frac{1}{1 - p^t}$$

$$\frac{d^2}{dt\,d(c/\mu)}\frac{(rp)^t - 1}{(r^t - 1)}\left(1 - \frac{c/\mu}{1 - p^t}\right) = \frac{d}{dt}\left(-\frac{(rp)^t - 1}{(r^t - 1)}\frac{1}{1 - p^t}\right)$$

$$= -\frac{\left((rp)^t - 1\right)(\ln rp - \ln r)}{(r^t - 1)}\frac{1}{1 - p^t} - \frac{(rp)^t - 1}{(r^t - 1)}\frac{p^t \ln p}{(1 - p^t)^2}.$$

A-3

As $rp > 1$, we can divide by $(rp)^t - 1$, multiply by $r^t - 1$ and $1 - p^t$, so the above expression is positive if and only if

$$-\ln p - \frac{p^t \ln p}{1 - p^t} > 0$$

if and only if

$$-\ln p(1 - p^t) - p^t \ln p > 0,$$

which is in turn equivalent to

$$-\ln p > 0,$$

which is always true.

Let us prove that

$$\frac{d^2}{dtdr} \frac{(rp)^t - 1}{(r^t - 1)} \left(1 - \frac{c/\mu}{1 - p^t}\right) > 0.$$

Calculate

$$
\begin{aligned}
\frac{d}{dr} \frac{(rp)^t - 1}{(r^t - 1)} \left(1 - \frac{c/\mu}{1 - p^t}\right) &= \left(1 - \frac{c/\mu}{1 - p^t}\right) \frac{d}{dr} \frac{(rp)^t - 1}{(r^t - 1)} \\
&= \left(1 - \frac{c/\mu}{1 - p^t}\right) \frac{d}{dr} \frac{tp(rp)^{t-1}}{tr^{t-1}} \\
&= \left(1 - \frac{c/\mu}{1 - p^t}\right) \frac{\left(tp(rp)^{t-1}\right)(r^t - 1) - tr^{t-1}\left((rp)^t - 1\right)}{(r^t - 1)^2} \\
&= \left(1 - p^t - c/\mu\right) \frac{tr^{t-1}}{(r^t - 1)^2}.
\end{aligned}
$$

Now

$$
\begin{aligned}
\frac{d}{dt} \left(1 - p^t - c/\mu\right) \frac{tr^{t-1}}{(r^t - 1)^2} &= -p^t \ln p \frac{tr^{t-1}}{(r^t - 1)^2} \\
&\quad + \left(1 - p^t - c/\mu\right) \frac{r^{t-1}(t \ln r + 1)(r^t - 1)^2 - 2tr^{t-1}(r^t - 1) r^t \ln r}{(r^t - 1)^4} \\
&= -p^t \ln p \frac{tr^{t-1}}{(r^t - 1)^2} + \left(1 - p^t - c/\mu\right) r^t \frac{r^t - 1 + (r^t + 1) t \ln r}{r(r^t - 1)^3}
\end{aligned}
$$

We will create another, sign-equivalent function:

$$\omega(r, t) = -p^t t \ln p + \left(1 - p^t - c/\mu\right) \frac{r^t - 1 + (r^t + 1) t \ln r}{r^t - 1}$$

If $\left(1 - p^t - c/\mu\right) > 0$, then $\omega(r, t) > 0$ which means that $\frac{d^2 \varphi(t,r)}{dtdr} > 0$. But $1 - p^t - \frac{c}{\mu} > 0$ is equivalent to $\mu - \frac{c}{1-p^t} > 0$, which is always true (our model is solved when this assumption is fulfilled). ∎

Lemma A5 *Let $X = \mathbb{Z}$, two-way. For any β, there exists an equilibrium on X.*

Proof. We will prove that there exists an $m \in \mathbb{Z}_+$ such that the set $S^m = \{mn|n \in \mathbb{Z}\}$ is an equilibrium subscription set. We need to find m such that, if S^m is a subscriber set, then, for any $x \in S^m$, the IC constraint looks as follows:

$$\left(1 - (1 - p^m)^2\right) V(\beta) \leq V(\beta) - c$$

or

$$2p^m - p^{2m} \leq A, \tag{A4}$$

if we denote $A = 1 - \frac{c}{V(\beta)}$ and note that $1 - (1 - p^m)^2 = 2p^m - p^{2m}$. If such m exists, we will need to check whether or not the IC constraint is satisfied for every non-subscriber as well.

Observe that the function $\varphi(x) = 2p^x - p^{2x}$ is non-negative, single-peaked on $[0, +\infty)$, and is converging to 0 when x approaches $+\infty$.

Start with the case when $A > \arg\max_{m \geq 0} \left\{2p^m - p^{2m}\right\}$. Then S^0 is an equilibrium subscription set; everyone is a subscriber.

Now, suppose that $A \leq \arg\max_{m \geq 0} \left\{2p^m - p^{2m}\right\}$. The fact $\varphi(x)$ is converging to zero guarantees that the set $\left\{m|2p^m - p^{2m} \leq A\right\}$ is non-empty. Take $m^* = \min\left\{m|2p^m - p^{2m} \leq A\right\}$. What remains to check is that if S^{m^*} is the subscriber set, than for any $x \notin S^{m^*}$, the incentive constraint is satisfied.

Suppose, on the contrary, that there exists an $x \in \mathbb{Z}$, $0 < x < m$ (so distance to the next left-ward subscriber, 0, is x, and to the next right-ward subscriber, m, is $m - x$) such that $1 - (1 - p)^x (1 - p)^{m-x} = p^x + p^{m-x} - p^m \leq A$. Without loss of generality $x \geq m - x$. Then $p^{m-x} \leq p^x$ as $p \leq 1$, and therefore $p^{m-x}(1 - p^{m-x}) \leq p^x(1 - p^{m-x}) = p^x - p^{m-x}$. This yields

$$2p^{m-x} - p^{2(m-x)} \leq p^x + p^{m-x} - p^m \leq A,$$

which contradics the choice of m^* as the minimal integer satisfying (A4). ∎