Persuasion on Networks[*]

Georgy Egorov   Konstantin Sonin

Northwestern University   University of Chicago

July 2020

Abstract

We analyze persuasion in a model, in which each receiver might buy a direct access to the sender's signal or to rely on her network connections to get it. For the sender, a higher slant increases the impact per direct receiver, yet diminishes the willingness of agents to receive information. Contrary to naive intuition, the optimal propaganda might target peripheral, rather than centrally-located agents, and is at its maximum levels when the probability that information flows between agents is close to zero or nearly one, but not in-between. The impact of the network density depends on this probability as well.

Keywords: Propaganda, networks, Bayesian persuasion, percolation.
JEL Classification: P16, D85, L82.

1 Introduction

Autocratic governments use propaganda to maintain their grip on power and achieve other goals (Guriev and Treisman, 2019). 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, reduce alcohol consumption and embrace a healthy lifestyle, increase and, in other time periods, decrease the birth rate, and many others (Service, 2003). In the communist China, mass campaigns were used, in addition to maintaining support for the party and its leaders, to promote development of individual industrial plants, encourage extermination of rats, 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. 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). In Afghanistan, the US government has successfully organized information operations to encourage local population to report road-side bombs and increase support for the reintegration of former Taliban fighters (Sonin and Wright, 2019).

In this paper, we analyze the limits on propaganda that stems from the existence of information networks that connect dictator's subjects. We do not focus on the spread of opposition information: the only information that is transmitted through the network is the dictator's own signal. The limit on propaganda comes from the incentive constraints of the receivers. If getting information is costly – e.g., 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 subscribe. Thus, there is a trade off for any producer of propaganda: when the signal is more slanted, the impact on those who get information from them is higher, yet the willingness of agents to become direct receivers is lower.

To analyze the impact of propaganda, we use the basic Bayesian persuasion model (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019; Kamenica, 2019). 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
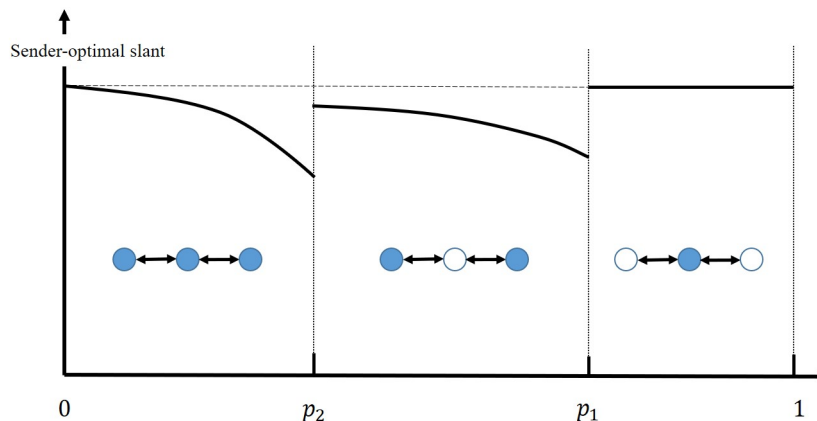
Figure 1: Sender-optimal equilibria as a function of p. Subscribers are solid blue. The optimal slant is decreasing on $[0, p_2)$ and $[p_2, p_1)$, and constant, at the maximum level, on $[p_1, 1]$.

or a propaganda campaign. Dictators do not edit news in the real time: instead, they appoint biased editors or establish institutions of censorship to generate the slanted signal (King, Pan and Roberts, 2013, 2017; Lorentzen, 2014; Rozenas and Stukal, 2019).

There is a single sender who maximizes the expected amount of action by agents connected in a network. In a political economy setting, this will be, e.g., mobilization in support of a leader or abstention from the anti-government protest. The sender chooses the information design without knowing the state of the world – this is a critical element of Bayesian persuasion. After the sender chooses the information design, receivers compare the cost of getting this information with the value that information provides to them; less slant makes information more valuable. Example 1 shows the consequences of the optimal choice by receivers (whether to buy subscription themselves or to get information for free from their neighbors) and the sender who optimizes the impact of propaganda in a very simple 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 slant are optimal. If p is high enough $(p > p_1)$, the sender-optimal structure is such that the slant 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 $(p < p_2)$, then the impact is maximized if everyone is a subscriber. Finally, for intermediate values of p, $p_2 < p < p_1$, the sender-preferred equilibrium is such that the slant is below the maximum level, the two peripheral agents buy subscription, and the central one receives information from them.
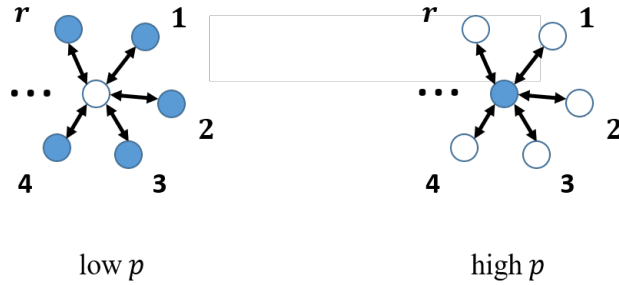
Figure 2: The sender-optimal equilibrium structure for different levels of p. Subscribers are solid.

In Example 1, the sender has to reduce the slant 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 slant that induces the desired action to have a higher probability that the signal reaches every agent. Propositions 1 and 2 demonstrate this logic in the case of one-way (directed) networks, and Propositions 3 and 4 describe equilibria in two-way (undirected) networks.

Example 1 show that the sender of propaganda would not necessarily target the centrally connected individual. In general, when an agent with a high degree of centrality (an "influencer") receives information directly, this crowds out the incentives of those who can receive information from this agent. In Example 2, when the edge probability p is sufficiently low, the sender would prefer that every peripheral agent becomes a subscriber, yet the central agent, for whom the incentive compatibility constraint is tighter, does not.[1] Proposition 8 provides this result for any star network.

Example 2. Consider a star network with r peripheral nodes depicted in Figure 2. When the edge probability p is high, then the naive logic applies: the impact is maximized when the central agent is a subscriber (solid blue), and other receive information from her. 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.

---

[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 et al. (2010), as an example of network games with externalities. In a model of information diffusion in a development context, Akbarpour and Jackson (2018) have effect similar to that in Example 2: the optimal seeding structure might target peripheral, rather than the centrally connected agents.
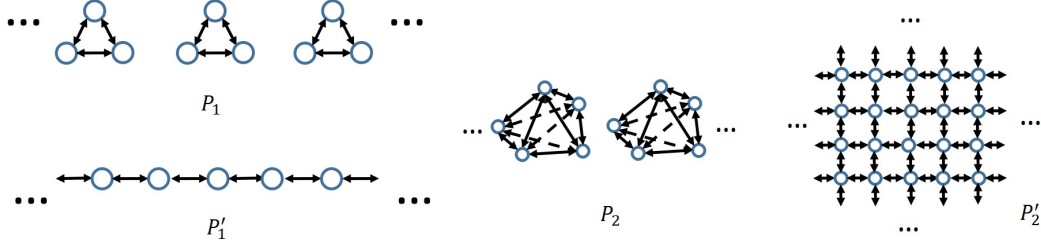
Figure 3: In networks P_1 and P_1', as well as in networks P_2 and P_2', each agent has the same number of links (two and four, resp.). When p is above a certain threshold, the government propaganda is more effective in P_1' and P_2'. However, in P_1, the sender-optimal subscription set is a singleton if and only if $p = 1$, while in P_2' is a singleton whenever p is larger than some $\bar{p} < 1$.

In our model, the immediate consequence of the presence of a network 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 slant reduces the citizens' incentives to subscribe. As a result, a 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, many totalitarian regimes discouraged all kind of networking, even strictly non-political, save for directly state-supervised one (Service, 2003; Roberts, 2018). Proposition 6 illustrates the impact of network density on the optimal slant: a more dense network with the same frequency of subscribers features lower optimal slant.

Example 3 demonstrates another mechanism through which network structure limits persuasion. We compare two networks, with each agent having the same number of links in either of them, yet networks having different topology. As a result, the sender's ability to persuade receivers connected via such networks is different.

Example 3. Consider two pairs of networks in Figure 3: P_1 and P_1', and P_2 and P_2'. In each pair, agents have the same number of connections (2 and 4, respectively). 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 $\bar{p}$ such that for any $p > \bar{p}$, the government would survive for a wider range of parameters with polity P_1 than with polity P_1' and with P_2 rather than P_2'. When p is high enough (not necessarily close to 1), the sender-optimal slant in P_1 and P_2 is at the maximum level and each connected component has a unique subscriber. For P_1', it is never optimal to have slant at its maximum unless $p = 1$; for P_2', again, there exists another threshold $\tilde{p} < 1$, such that the maximum slant is optimal for any $p > \tilde{p}$.

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 equilibrium is sender-optimal when the distance that is consistent with the subscribers' incentive compatibility constraint is at the minimum. This distance approaches infinity as p tends to 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 (Proposition 7). Every other agent does not subscribe, expecting information to reach him with a sufficient probability – regardless of the distance from the unique subscriber. The reason for this is percolation (Broadbent and Hammersley, 1957): when p is large enough, there exists, almost surely, a unique infinite cluster of connected nodes. 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 $\mathbb{Z}^2$.

Finally, the discussion in Section 7 highlights another important feature of authoritarian propaganda. 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 slant is needed to satisfy each subscriber's incentive compatibility constraint. When p is close to 1, there is a single subscriber, so the slant is at the maximum level.

In our 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, 2016, 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. 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 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 directed networks. Section 5 analyzes undirected networks, while Section 6 relates our results to percolation on infinite lattices. Section 7 applies the model's logic to propaganda in non-democratic regimes. Section 8 concludes, and the Appendix contains technical proofs.

2 Related Literature

Bayesian Persuasion. Kamenica and Gentzkow (2011) introduced the general notion of Bayesian persuasion that encompassed many models, in which commitment to information disclosure serves as a persuasion mechanism (e.g., Aumann and Maschler, 1966; Brocas and Carrillo, 2007; Ostrovsky and Schwarz, 2010; Rayo and Segal, 2010). Dewatripont and Tirole (2005) introduced the moral hazard aspect of communication with a conflict of interest between sender and receiver. Calvó-Armengol, De Martí and Prat (2015) extended their model to network communication, in a setting very different from what we consider in this paper.

Bayesian persuasion with multiple receivers was considered in Bergemann and Morris (2013, 2019), Mathevet, Perego and Taneva (2020), and Taneva (2019). Alonso and Câmara (2016) and Kosterina (2018) assume that signals are public, in Wang (2015) signals are independent, and Arieli and Babichenko (2016), Bardhi and Guo (2018), and Chan et al. (2019) allow for arbitrary private signals. In Kolotilin (2015, 2018), the persuasion mechanism that the sender designs relies on private information reported by the receivers. In our model, the persuasion part is elementary compared to that in any of these papers. Still, it allows us to describe the limits that the network structure puts on the sender's persuasion power.

Gentzkow and Kamenica (2014) provide an important connection between the literatures on Bayesian persuasion and rational inattention (e.g., Sims, 2003). In our model, the receivers weight the benefits and the costs of having a direct access to information, which makes the underlying intuition somewhat similar to the rational inattention mechanism. In Bloedel and Segal (2018), a rationally inattentive receiver bears an entropic cost to process the sender's signals; the special case of their model with a binary state and state-independent sender preferences is a deep generalization of our single-receiver version. (Also, see Lipnowski, Mathevet and Wei 2020, and the follow-up paper Wei 2018). Finally, Matyskova (2018) considers, in a single-receiver setting, Bayesian persuasion in the case which is "opposite" to ours in the sense that learning from the sender is free but acquiring additional information about the state is privately costly. Again, the contribution of this paper is in identifying the impact of the network structure on persuasion power even if the persuasion problem is very basic.

Information in Networks. We rely on Jackson (2008) as an exhaustive review of the early literature and an 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. (See also Acemoglu et al., 2011.) 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 (2020) considers information contagions in a discrete network with Bayesian players that do not necessarily know the network structure. In Lipnowski and Sadler (2019), if two players are linked in a network, they get information about each others' strategies. (In our model, information flows through a network link with some probability.) 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 have 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. 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 et al. (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, Malladi 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. (See Example 1 and Proposition 8.)

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 agent. In our 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 models 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 Treisman (2019) propose a theory of the "informational autocracy", in which the dictator chooses between repression and propaganda. (See also Chen, Lu and Suen, 2016 and Huang, 2018). 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 Strömberg (2013) are two excellent surveys on the role of persuasion in politics, with the emphasis on a democratic context. Adena et al. (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 networks in the rise of Nazis, but do not analyze the impact of the network structure. Yanagizawa-Drott (2014), King, Pan and Roberts (2017, 2014), and Chen and Yang (2019) use experimental technique to analyze empirically propaganda tactics of an authoritarian government. Chen and Yang (2019) 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 Hammersley (1957). Kesten (1980) completed characterization of critical probability in $\mathbb{Z}^m$, which we use in Section 6. Van Der Hofstad (2010) and Duminil-Copin (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 consider an environment with a sender and multiple agents, potential receivers of the information provided by the sender, connected via a network. The agents might acquire information not only from the source controlled by the sender, e.g., a newspaper or a TV channel, directly, but indirectly from other agents in the network. The network X consists of nodes $x \in X$ and edges $(x, y) \in E(X)$. We will assume that X is 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 agent's individual action:

$$u_x(a_x = 1, s = 1) = 1 - q,$$

$$u_x(a_x = 0, s = 0) = q,$$

$$u_x(a_x = 1, s = 0) = u_x(a_x = 0, s = 1) = 0,$$

where $q \in (0, 1)$. The common prior is $P(s = 1) = \mu < q$, 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 signal $\widehat{s}$. The signal is structured so that

$$P(s = 1 | \widehat{s} = 1) = \frac{\mu}{\mu + (1 - \mu)\beta},$$

$$P(s = 0 | \widehat{s} = 0) = 1,$$

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, $0 < c < \mu$. If an agent is indifferent, she subscribes and acts.

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. The original Erdős-Rényi random graph is a particular case of this construction when X is a *complete* graph, i.e., each node is connected to every other node.[2] 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 subscription set* 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. Conditional on information, each agent x's action $a_x \in \{0, 1\}$ maximizes her expected utility.

The sender wants to have the expected amount of action 1 per agent as high as possible. We will specifically focus on sender-optimal pure strategy Bayes-Nash equilibria, the outcome of the game with following timing. First, the sender commits to the level of propaganda β, then

---

[2] Sonin (2020) considers the setup when X is a complete network. In such a network, every subscriber and every non-subscriber face the same incentive constraints, which grossly simplifies analysis.

everything is determined in an equilibrium given β.

We say that network X is *linear* if it has two vertices of degree 1 and every other vertex has degree 2, and *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 there is at most one node that belongs to more than one link. (The network in Example 1 is both a linear and a star network.)

Finally, $|M|$ denotes the cardinality of set M.

4   Directed Networks

We start by analyzing economics of propaganda on directed networks, which are relatively straightforward to study. We use this analysis to develop tools for the analysis of a more general case, and also to demonstrate the important features of persuasion on networks.

One-way Linear Network.   We start our analysis with the case of an infinite linear network, in which information flows in one direction, $X = \mathbb{Z}_+ = \{0, 1, 2, 3, ...\}$.

In the absence of information, agent $x \in X$ chooses $a_x = 0$ and gets the expected payoff of $(1 - \mu) q$. With access to information, the agent's expected utility is $\mu (1 - q) + (1 - \mu)(1 - \beta) q$. Thus, the value of information is

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

Agent $x = 0$ subscribes if and only if $V(\beta) \geq c$. Consider the decision by agent $k \in \{1, 2, 3, ...\}$ when agent 0 is a subscriber and agents $1, ..., k - 1$ do not subscribe. Buying subscription gives the expected value of $V(\beta) - c$. Not buying gives $V(\beta)$ with probability p^k and 0 with probability $1 - p^k$. That is, agent k buys subscription if and only if $V(\beta) - c \geq V(\beta)p^k$, or, equivalently

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

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

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

Conditions 1 define K such that agent $x = K - 1$ does not want to be a subscriber, yet agent $x = K$ does (Figure 4). Then K defines the equilibrium size of the "subscription cell". The set of subscribers is $S(X) = \{nK | n = 0, 1, 2, ...\}$. Naturally, the size of the cell $K = K(\beta, \mu, q, c, p)$
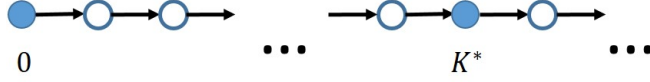
Figure 4: A symmetric equilibrium on a one-way linear network.

is increasing in c and β and is decreasing in μ and p.

The expected action per agent in network $X = \mathbb{Z}_+$ is

$$\frac{1}{K} \left(\mu + (1-\mu)\beta\right) \frac{1-p^K}{1-p}. \tag{2}$$

If a network consists of a single agent, the sender-optimal slant can be derived from the incentive compatibility constraint $(V(\beta) - c \geq 0)$:

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

What is the optimal slant on $X = \mathbb{Z}_+$? It has to maximize (2) and satisfy $\left(1 - p^{K^*}\right) V(\beta^*) = c$, so

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

Note that the optimal level of slant is strictly lower in the case of the linear network than in the case of a single receiver: $\beta^*\left(\mathbb{Z}_+\right) < \beta^{\max}$ for any p such that $0 < p < 1$. It is intuitive: the presence of other would-be subscribers forces the sender to scale down the slant. This result is also very general: it extends to any network 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. In such a network, setting the slant at $\beta^{\max}$ results in no other subscribers as each agent compares zero informational rent if she subscribes to a positive cost of subscription. One-way tree networks and any connected two-way networks are examples of such networks.

The following proposition summarizes the above discussion about persuasion on one-way linear network.

Proposition 1. *Let $X = \mathbb{Z}_+$ be a one-way (directed) linear network.*

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

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

11

The optimal slant $\beta^* = \beta^*(\mathbb{Z}_+)$ 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 optimal slant $\beta^*(\mathbb{Z}_+, p)$ 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 slant that the sender chooses.

(iii) The optimal size of the subscription cell $K^*(c)$ and the level of slant $\beta^*(\mathbb{Z}_+, c)$ 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 slant.

Technically, the comparative statics of parts (ii) and (iii) of Proposition 1 rely on Lemma A2 in the Appendix that demonstrates that (2) satisfies increasing differences condition in (K, p) and (K, c) and provides the basis for comparative statics with respect to p and c (Milgrom and Shannon, 1994). Substantively, Proposition 1 features the main trade-off that a sender of propaganda faces: a higher edge probability results in a higher optimal slant. A higher p does discourage subscription: the subscription cell becomes longer in equilibrium.

Tree Networks. Consider the network $X = \mathbb{T}_r$, a regular tree; that is, a directed graph with each node having $r > 1$ outgoing links. In a study of persuasion, a tree network is interesting for a number of reasons. First, it is a natural extension of directed path networks: each path through a tree is isomorphic to a directed linear network. Second, it is a far more realistic approximation of spread of information through a society than the directed linear network. Finally, this is the simplest possible network in which, in contrast with the linear network case, there might be infinitely many agents choosing $a_x = 1$ even when $p < 1$ and there is a unique subscriber.

What is the sender-optimal level of slant β in a tree network? Each directed path through a tree is isomorphic to the directed linear network. Define $K(\beta)$ as in (1). Now, suppose that β^* is the optimal level of β. Then, the incentive compatibility constraint should be binding for $K(\beta^*)$:

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

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 $\frac{r^K - 1}{r - 1}$. So, the problem is to maximize the average actions:

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

The next Proposition 2 extends the logic of Proposition 1 to tree networks. Still, there is

an important new result that relates the optimal slant to the number of tree branches, i.e., the agents that have a chance to receive information from a single agent (part (ii)).

Proposition 2. *Let $X = \mathbb{T}_r$. Suppose that $rp > 1$.*

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

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

(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.

Technically, the proof of Proposition 2 is similar to that of Proposition 1. If $rp > 1$, Lemma A3 guarantees that (4) is single-peaked with respect to K. As in the linear network case, this guarantees existence of an optimal K^*. The comparative statics rely on Lemma A4, which asserts that (4) satisfies increasing differences in (K, c), $(K, -\mu)$, and (K, r). While Lemma A4 deals with a continuous function, the monotone comparative statics applies to the discrete maximization as well.

In addition to extending the results of Proposition 1, Proposition 2 establishes 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 slant 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 chooses a higher slant, putting more weight on persuading those who are reached rather than on having a higher rate of subscription.

Persuasion and Percolation on Regular Trees. There is a critical difference between spread of information on a directed linear network and a tree with $r \geq 2$. In a linear network, the probability that a single subscriber results in an infinite number of agents choosing the sender-preferred action is 0 for any edge probability $p < 1$. In an r-arnary tree network, this probability depends on the relationship between p and r. Specifically, if $p > \frac{1}{r-1}$, then a single subscriber at the root of $X = \mathbb{T}_r$ results in infinitely many actions with a non-zero (and non-vanishing!) probability. Still, despite that there will be infinitely many agents receiving information, the expected action per agent resulting from a single subscriber at the root of the tree will be zero.

We will see a qualitatively different result, the infinitely many agents with a positive action per agent resulting from a unique subscriber when the network is $X = \mathbb{Z}^2$ (see Section 6).

5   Two-Way Networks

In this Section, we start with a two-way flow infinite linear network $X = \mathbb{Z}$, the simplest possible two-way network. Then, we compare the cases $X = \mathbb{Z}$ and $X = \mathbb{Z}^2$, which are dramatically different in terms of sender-optimal persuasion.

Infinite Linear Network.   Fix any β, $0 \leq \beta \leq \beta^{max}$. We will focus on symmetric equilibria on two-way network $X = \mathbb{Z}$; that is, equilibria that have the same number of non-subscribers between each neighboring pair of subscribers. Proposition 3 establishes existence and provides an estimate for the difference between the equilibrium with the minimum distance between subscribers (which is sender-optimal with fixed β) and the maximum.

Proposition 3. *Let $X = \mathbb{Z}$ be a two-way network, and consider any slant β.*

(i) There exist two integer thresholds $a = a\,(p, \beta, c)$ and $b = b\,(p, \beta, c)$ such that for any integer d, $a \leq d \leq b$, $S^d = \{dn \mid n \in \mathbb{Z}\}$ is a symmetric 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 X satisfies $a \leq \left\lfloor \frac{b}{2} \right\rfloor$, where $\lfloor t \rfloor$ denotes the integer part of a real number t.

Part (i) of Proposition 3 is proved in the Appendix. Here we will prove part (ii) as the proof provides insights into the source of multiplicity of symmetric equilibria in the two-way case.

As in part (i), let $a = a\,(p, \beta, c)$ and $b = b\,(p, \beta, c)$ be the minimum and the maximum number of non-subscribers between two subscribers in an equilibrium, respectively. Denote S^b the corresponding set of subscribers, $S^b = \{(b+1)\,n \mid n \in \mathbb{Z}\}$. 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. 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 $\left\lfloor \frac{b+1}{2} \right\rfloor^{\text{th}}$ agent in S^b, and the distance between any two other subscribers is closer in $S^{\frac{b}{2}}$ than in S^b, $S^{\frac{b}{2}}$ defines a symmetric equilibrium. So, we just demonstrated that $a \leq \left\lfloor \frac{b}{2} \right\rfloor$.

Naturally, for each β, the sender-optimal equilibrium set of subscribers is S^a as this equilibrium features the highest frequency of subscribers. As every non-subscriber may receive information from her neighbouring two subscribers only, the minimum distance between subscribers

14

guarantees that the probability that information reaches non-subscribers is at the maximum. The optimal slant on $X = \mathbb{Z}$ satisfies $c = V(\beta^*) \left(1 - p^{a^*}\right)^2$, where $a^* = \min_\beta \{a(\beta)\}$. In higher dimensions, the story will be very different as the specifics of one-dimensional case is that nobody could receive any information but from their immediate neighbors (see Section 6).

Finite Linear Network. Part (ii) of Proposition 3 can be used to prove existence of an equilibrium on a finite linear network, which is a less straightforward exercise. We will use the same thresholds $a = a(p, \beta, c)$, the minimum number of non-subscribers between two subscribers in a symmetric equilibrium on $\mathbb{Z}$, and $b = b(p, \beta, c)$, the maximum number. Proposition 3 established that $a \leq \left\lfloor \frac{b}{2} \right\rfloor$.

Consider a finite two-way linear 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. the number of agents in X, $|X|$, does not exceed $b + 1$. 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 $a \leq \left\lfloor \frac{b}{2} \right\rfloor$, the distance between $l(b + 1) + 1$ and $|X| - l(b + 1) - 1$ exceeds $\left\lfloor \frac{b}{2} \right\rfloor$. Thus, we ended up with an equilibrium subscription set.[3] Formally, one can state the following result.

Proposition 4. *Let X be a finite two-way linear network. For any slant β, there exists a symmetric equilibrium on network X.*

6 Persuasion and Information Percolation

In Sections 4 and 5, we built a theory of persuasion on linear networks, extended it to directed tree networks, in which every path is isomorphic to the linear network, and then to two-way linear

---

[3]A similar argument establishes that in any circular network, an equilibrium exists for any β.

network. However, it appears that a small increase in network dimensionality leads to drastically different results with respect to persuasion. This is a result of "percolation" (Broadbent and Hammersley, 1957; Kesten, 1980), the effect that is not present on linear networks, yet plays an important role whenever the network is similar (locally) to two-dimensional lattice.

We will need some basic definitions and results from percolation theory (Duminil-Copin, 2018). For any (connected) network X and one subscriber $x \in X$, we define the connected component C_x, a random variable with the values in all subnetworks of X that contain x. The observed value of C_x is the set of all agents who received information from agent x. Now we can define the *percolation function* $\theta(p) = P(|C_x| = \infty)$, the probability that the connected cluster that originated in x contains infinitely many agents.

There is the following central result in the percolation theory: in many networks X, 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 $p > p_c$, $\theta(p) > 0$ (Kesten, 1980). When the edge probability is lower than the critical probability, the cluster that starts with a single subscriber is finite with probability 1. Yet when the edge probability p exceeds the critical threshold $p_c(X)$, the infinite cluster forms with a positive probability $\theta(p) > 0$ which increases with p. For r-regular trees and multidimensional integer lattices, the function $\theta(p)$ is continuous on $[0, 1]$ and is increasing on $(p_c, 1]$.

Persuasion in Higher Dimensions. 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 environment, 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. In this case, construction of a symmetric equilibrium is relatively straightforward.

Proposition 5. *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,

16

$h \in \mathbb{N}$. For each h, define the set

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

(S_2^2 is the right display in Figure 5.) 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. $\qquad\square$

Let us connect the frequency of subscribers in the network with the sender-optimal slant. To do this, we will look at networks $\mathbb{Z}^m$ for different ms. Consider a sequence of lattices $\mathbb{Z}^m$, $m = 1, 2, ...$, with links $\{((x_1, ...x_l, ..., x_m), (x_1, ...x_l + 1, ..., x_m)) \,|\, (x_1, ...x_l, ..., x_m) \in \mathbb{Z}^m\}$, and the set of subscribers

$$S^m = \{(2l_1, 2l_2, ..., 2l_m), (2l_1 + 1, 2l_2 + 1, ..., 2l_m + 1) \,|\, l_1, l_2, ..., l_m \in \mathbb{Z}\}. \qquad (3)$$

(See Figure 5 for $m = 1$ and $m = 2$.)

In the larger dimensions, each non-subscriber is connected to more subscribers. Thus, if the frequency of subscribers is the same, the slant compatible with the subscription structure should be lower in higher dimensions. In other words, a network of a higher density puts stricter limits on the persuasion power of the sender, than a less dense one.

Proposition 6. *Let $X = \mathbb{Z}^m$, let the set S^m be defined by (3), 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.*[4]

While Proposition 6 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 incentive compatibility constraint is satisfied for a lower level of propaganda.

---

[4]The incentive compatibility constraint for the non-subscriber is $\left(1 - (1 - p)^{2m}\right) V(\beta) < V(\beta) - c$, 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_k is a strictly decreasing function for $k \in [0, \overline{m}]$.
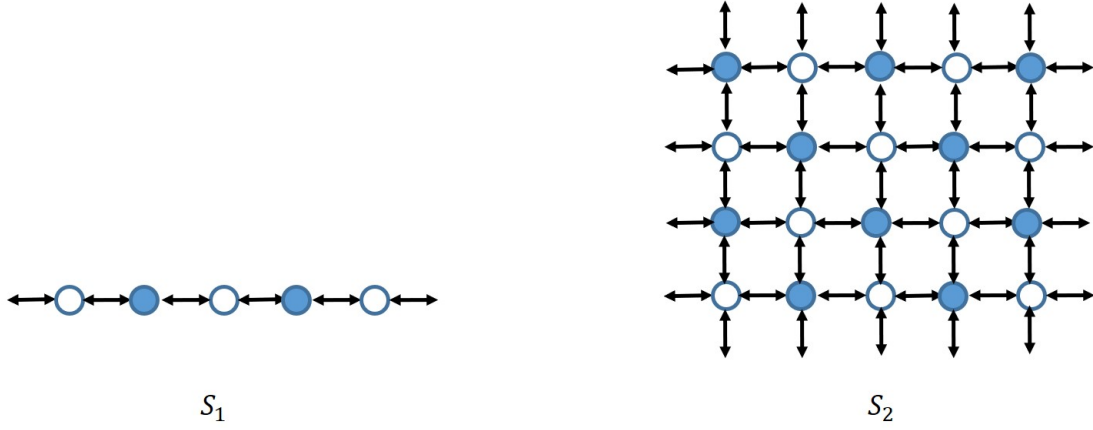
S_1

S_2

Figure 5: Sets of subscribers S_1 and S_2 (solid blue) in $\mathbb{Z}$ and $\mathbb{Z}^2$, respectively, result in the same frequency of subscribers, $\frac{1}{2}$, but the maximum level of slant consistent with S_2 is lower.

The Impact of Percolation. Consider again $X = \mathbb{Z}^m$. 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 5 and 6 results in a higher expected average action than any one-subscriber structure. For $p > p_c(X)$, the situation is different: 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.

Take $m = 2$, and consider the subscription structure S_2 (see Figure 5). 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 incentive compatibility constraint of the subscriber is satisfied if $f(p)V(\beta) \le V(\beta) - c$, and the expected action per agent is $\frac{1}{2}(1+p)\frac{1}{q}\left(\mu - \frac{c}{1-f(p)}\right)$, which is non-negative. So, whenever $p < p_c(\mathbb{Z}^2) = \frac{1}{2}$, this structure is preferable to that of a finite number of subscribers with $\beta^{\max}$.

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

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

Inequality (4) cannot be true for all pairs (c, p), $c > 0$, $p > 0$, yet for any $p < 1$, there exists a

18

threshold $c(p) > 0$ such that (4) 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, almost surely, a cluster of second-hand receivers of a finite size. 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

$$\theta \left(p \right) V \left(\beta \right) > V \left(\beta \right) - c, \tag{5}$$

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 (Duminil-Copin, 2018). That is, the probability that information reaches x is at least $\theta \left(p \right)$, which makes her unwilling to subscribe. Therefore, we can state the following formal result.

Proposition 7. *Let $X = \mathbb{Z}^d$, $d \geq 2$, and suppose that $p > p_c \left(\mathbb{Z}^d \right)$. Then, for any β such that $V \left(\beta \right) < \frac{c}{1 - \theta(p)}$, any equilibrium has a unique subscriber. Vice versa, for any $x \in X$, $S(X) = \{x\}$ is an equilibrium set of subscribers.*

Proposition 7 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 (5) is not fulfilled, then finite equilibrium groups of subscribers with $\beta < \beta^{\max}$ might 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. Still, this argument is correct if and only if subscribers cannot "block" each other; that is, there exist infinite connected paths that consist of non-subscribers only. If this assumption is violated (as in S_2 on Picture 5), there might exist an equilibrium with infinitely many subscribers.

19

7 Optimal Propaganda

In this Section, we apply the results discussed in the previous sections to the issue of optimal propaganda. Specifically, we will discuss the role of "influencers", the agents that have a high degree of centrality in a network.

The Dictator's Preference over Network Characteristics. Our first observation about propaganda is almost trivial. If the sender takes the probability p as given, his optimal choice is a function of p. As Example 1 demonstrated, this choice is non-monotonic, resulting in the maximum slant for very low or very high values of p and p. Thus, if the sender has a choice of the edge probability p, atop of the optimal choice of slant after choosing p, then the optimal choice of p is either $p^* = 0$ or $p^* = 1$, and the propaganda level is at its maximum:

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

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 $\mu + \beta^*(1-\mu) = \mu + \frac{c - \mu(1-q)}{q} = \frac{1}{q}(\mu - c)$. This result is true regardless of whether or not the network X is directed or undirected.

Alternatively, $p^* = 1$, and in a two-way network there will necessary be a single subscriber. Again, choosing $\beta^* = \beta^{\max}$ maximizes the expected average amount of individual action, which is the same as in the case of $p^* = 0$. In a directed (one-way) network, the result is more subtle: it is still optimal to have $p^* = 1$, yet there should be an additional condition on the network X: there must be a "root" for any infinite sequence of one-way links. (Otherwise, the one way network $X = \mathbb{Z}$ is a counterexample: if $p = 1$, there is no pure-strategy equilibrium for any β.) In a one-way network, there might be more than one subscriber even if $p = 1$.

Both of the two extreme edge 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.

The Role of a Centrally Connected Agent. 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 always prefers 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 8. *Let X be a star network that connects $r + 1$ agents.*

(i) There exist thresholds $0 < \bar{p}_r \leq ... \leq \bar{p}_1 < 1$ such that for any $p \in \left[\bar{p}_k, \bar{p}_{k-1}\right]$, there are exactly k agents who subscribe in a sender-optimal equilibrium.

(ii) The sender-optimal slant $\beta^* (p)$ is strictly decreasing over each interval $\left[\bar{p}_k, \bar{p}_{k-1}\right]$, $k = 1, ..., r + 1$. It is constant, $\beta^* (p) = \beta^{\max}$, on $\left[\bar{p}_1, 1\right]$.

(iii) The central agent is a subscriber in a sender-optimal equilibrium if and only if either $p \leq \bar{p}_r$, or $p \geq \bar{p}_1$

Part (i) implies that the set of subscribers in the sender-optimal equilibrium (weakly) shrinks with p. When p is close to zero, the incentive compatibility constraint is not binding for every subscriber. The first subscriber to have the constraint binding is the agent at the center of the star X : so, the first threshold $\bar{p}_r$ is determined by this constraint. Indeed, when all peripheral agents is the set of subscribers, the probability that one peripheral agent receives information from other $r - 1$ peripheral agents 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 these agents. So, the incentive compatibility constraint for the subscriber 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_r is determined by the

21

equation

$$(r+1)\left(\mu - \frac{c}{(1-p)^r}\right) = (r + (1-(1-p)^r))\left(\mu - \frac{c}{1 - p\left(1 - (1-p)^{r-1}\right)}\right). \quad (6)$$

The left-hand side is the total expected action when all $r+1$ agents are subscribers; the optimal slant comes from the central agent's incentive compatibility constraint.

Proposition 8 extends to any two-way tree. When p is close to 1, there is no chance to satisfy the incentive constraint for more than one subscriber. In this situation, it is better to have the central agent (the root) 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.

Proposition 8 challenges the conventional wisdom about targeting "influencers", agents with a high degree of centrality. When p is low, they are the *worst* possible target for influence. This 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.

Discussion. 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'. 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

22

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 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 et al., 2010 and Sadler, 2020.) Unlike in the local public goods games (Bramoullé and Kranton, 2007; Galeotti et al., 2010), agent's decision to subscribe 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 degree of centrality are not sender-preferred targets of propaganda.

8   Conclusion

We consider an environment in which agents that are influenced by a sender who has 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. In equilibrium, a subscriber crowds out the incentives to subscribe of those who could get information from her. A higher slant chosen by the sender increases the probability that those who receive signal would do what the sender wants, but further reduces agents' willingness to subscriber. 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 the possibility that agents could strategically share their signals or exchange information that is harmful to the sender would be an interesting extension, yet will leave the main effects identified in our paper in place.

References

Acemoglu, Daron, Munther A Dahleh, Ilan Lobel and Asuman Ozdaglar. 2011. "Bayesian learning in social networks." *The Review of Economic Studies* 78(4):1201–1236.

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

Akbarpour, Mohammad and Matthew O 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." *Available at SSRN 3062830* .

Alonso, Ricardo and Odilon Câmara. 2016. "Persuading voters." *American Economic Review* 106(11):3590–3605.

Arieli, Itai and Yakov Babichenko. 2016. "Random extensive form games." *Journal of Economic Theory* 166:517–535.

Aumann, Robert J and M Maschler. 1966. "Game theoretic aspects of gradual disarmament." *Report of the US Arms Control and Disarmament Agency* 80:1–55.

Bala, Venkatesh and Sanjeev Goyal. 1998. "Learning from neighbours." *The Review of Economic Studies* 65(3):595–621.

Bala, Venkatesh and Sanjeev Goyal. 2000. "A noncooperative model of network formation." *Econometrica* 68(5):1181–1229.

Bardhi, Arjada and Yingni Guo. 2018. "Modes of persuasion toward unanimous consent." *Theoretical Economics* 13(3):1111–1149.

Bergemann, Dirk and Stephen Morris. 2013. "Robust predictions in games with incomplete information." *Econometrica* 81(4):1251–1308.

Bergemann, Dirk and Stephen Morris. 2019. "Information design: A unified perspective." *Journal of Economic Literature* 57(1):44–95.

Bloedel, Alexander W and Ilya R Segal. 2018. "Persuasion with Rational Inattention." *Available at SSRN 3164033* .

Bramoullé, Yann and Rachel Kranton. 2007. "Public goods in networks." *Journal of Economic Theory* 135(1):478–494.

Broadbent, Simon R and John M Hammersley. 1957. Percolation processes: I. Crystals and mazes. In *Mathematical Proceedings of the Cambridge Philosophical Society.* Vol. 53 Cambridge University Press pp. 629–641.

Brocas, Isabelle and Juan D Carrillo. 2007. "Influence through ignorance." *The RAND Journal of Economics* 38(4):931–947.

Calvó-Armengol, Antoni, Joan De Martí and Andrea Prat. 2015. "Communication and influence." *Theoretical Economics* 10(2):649–690.

Candogan, Ozan. 2019. Persuasion in networks: Public signals and k-cores. In *Proceedings of the 2019 ACM Conference on Economics and Computation.* pp. 133–134.

Chan, Jimmy, Seher Gupta, Fei Li and Yun Wang. 2019. "Pivotal persuasion." *Journal of Economic Theory* 180:178–202.

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 Y Yang. 2019. "The impact of media censorship: 1984 or brave new world?" *American Economic Review* 109(6):2294–2332.

Chwe, Michael Suk-Young. 2000. "Communication and coordination in social networks." *The Review of Economic Studies* 67(1):1–16.

Crawford, Vincent P and Joel Sobel. 1982. "Strategic information transmission." *Econometrica: Journal of the Econometric Society* pp. 1431–1451.

DellaVigna, Stefano and Matthew Gentzkow. 2010. "Persuasion: empirical evidence." *Annual Review of Economics* 2(1):643–669.

Dewatripont, Mathias and Jean Tirole. 2005. "Modes of communication." *Journal of Political Economy* 113(6):1217–1238.

Duminil-Copin, Hugo. 2018. "Sixty years of percolation." *Proceeding of the International Congress of Mathematicians* 4:2847–2874.

Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. "Why resource-poor dictators allow freer media: A theory and evidence from panel data." *American Political Science Review* 103(4):645–668.

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

Galperti, Simone and Jacopo Perego. 2019. "Belief Meddling in Social Networks: An Information-Design Approach." *Available at SSRN 3340090* .

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* 104(5):457–62.

Golub, Benjamin and Matthew O Jackson. 2010. "Naive learning in social networks and the wisdom of crowds." *American Economic Journal: Microeconomics* 2(1):112–49.

Guriev, Sergei and Daniel Treisman. 2019. "Informational autocrats." *Journal of Economic Perspectives* 33(4):100–127.

Huang, Haifeng. 2018. "The Pathology of Hard Propaganda." *The Journal of Politics* 80(3):1034–1038.

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

Jackson, Matthew O, Brian Rogers and Yves Zenou. 2016. Networks: An Economic Perspective. In *Oxford Handbook of Social Network Analysis*. Oxford University Press.

Jackson, Matthew O and Leeat Yariv. 2011. Diffusion, strategic interaction, and social structure. In *Handbook of Social Economics*. Vol. 1 Elsevier pp. 645–678.

Kamenica, Emir. 2019. "Bayesian persuasion and information design." *Annual Review of Economics* 11:249–272.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101(6):2590–2615.

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 E Roberts. 2013. "How censorship in China allows government criticism but silences collective expression." *American Political Science Review* 107(2):326–343.

King, Gary, Jennifer Pan and Margaret E Roberts. 2014. "Reverse-engineering censorship in China: Randomized experimentation and participant observation." *Science* 345(6199):1251722.

King, Gary, Jennifer Pan and Margaret E Roberts. 2017. "How the Chinese government fabricates social media posts for strategic distraction, not engaged argument." *American Political Science Review* 111(3):484–501.

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(2):607–635.

Kosterina, Svetlana. 2018. "Persuasion with unknown beliefs." *Working Paper* .

Lipnowski, Elliot and Evan Sadler. 2019. "Peer-Confirming Equilibrium." *Econometrica* 87(2):567–591.

Lipnowski, Elliot, Laurent Mathevet and Dong Wei. 2020. "Attention management." *American Economic Review: Insights* 2(1):17–32.

Lorentzen, Peter. 2013. "Regularizing rioting: Permitting public protest in an authoritarian regime." *Quarterly Journal of Political Science* 8(2):127–158.

Lorentzen, Peter. 2014. "China's strategic censorship." *American Journal of Political Science* 58(2):402–414.

Mathevet, Laurent, Jacopo Perego and Ina Taneva. 2020. "On information design in games." *Journal of Political Economy* 128(4):1370–1404.

Matyskova, Ludmila. 2018. "Bayesian persuasion with costly information acquisition." *CERGE-EI Working Paper Series* (614).

Milgrom, Paul and Chris Shannon. 1994. "Monotone comparative statics." *Econometrica: Journal of the Econometric Society* pp. 157–180.

Milgrom, Paul R. 1981. "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics* pp. 380–391.

Ostrovsky, Michael and Michael Schwarz. 2010. "Information disclosure and unraveling in matching markets." *American Economic Journal: Microeconomics* 2(2):34–63.

Prat, Andrea and David Strömberg. 2013. "The political economy of mass media." *Advances in Economics and Econometrics* 2:135.

Rayo, Luis and Ilya Segal. 2010. "Optimal information disclosure." *Journal of Political Economy* 118(5):949–987.

Roberts, Margaret E. 2018. *Censored: distraction and diversion inside China's Great Firewall.* Princeton University Press.

Rozenas, Arturas and Denis Stukal. 2019. "How Autocrats Manipulate Economic News: Evidence from Russias State-Controlled Television." *Journal of Politics* 81(3):982–96.

Sadler, Evan. 2020. "Diffusion games." *American Economic Review* 110(1):225–70.

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. 2003. *A history of modern Russia: From Nicholas II to Putin.* Penguin.

Shadmehr, Mehdi and Dan Bernhardt. 2015. "State censorship." *American Economic Journal: Microeconomics* 7(2):280–307.

Shapiro, Judith. 2001. *Mao's war against nature: Politics and the environment in revolutionary China.* Cambridge University Press.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of Monetary Economics* 50(3):665–690.

Sonin, Konstantin. 2020. "Networks and Limits on Propaganda." *mimeo* .

Sonin, Konstantin and Austin L Wright. 2019. "Information Operations Increase Civilian Security Cooperation." *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2019-130).

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

Taneva, Ina. 2019. "Information design." *American Economic Journal: Microeconomics* 11(4):151–85.

Taylor, Philip M. 1999. *British propaganda in the 20th century: Selling democracy.* Edinburgh University Press Edinburgh.

Thaler, Richard H and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness.* Penguin.

Van Der Hofstad, Remco. 2010. "Percolation and random graphs." *New Perspectives in Stochastic Geometry* pp. 173–247.

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

Wang, Yun. 2015. "Bayesian persuasion with multiple receivers." *Available at SSRN 2625399* .

Wei, Dong. 2018. "Persuasion under costly learning." *Available at SSRN 3188302* .

Yanagizawa-Drott, David. 2014. "Propaganda and conflict: Evidence from the Rwandan genocide." *The 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 $(2 - p)pV(\beta) \le V(\beta) - c$ 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 $a \in \left(\frac{3}{100}, \frac{9}{200}\right)$, $f_2\left(\frac{1}{2}\right) > \max\left\{f_1\left(\frac{1}{2}\right), f_3\left(\frac{1}{2}\right)\right\}$. ∎

Algebra for One-Way Linear Network

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

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

Now, the expected average action is

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

(So, we effectively replaced $\lim_{m\to\infty}\frac{1}{m}$ by $\lim_{l\to\infty}\frac{1}{lK}$.)

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

Proof. Take a derivative:
$$\varphi'(t) = \frac{-p^t t \ln p - \alpha + p^t}{t^2}.$$

Let $\phi(t) = \frac{-p^t t \ln p - \alpha + p^t}{p^t} = -t\ln p - \alpha\left(\frac{1}{p}\right)^t + 1$. 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 $p < 1$, $1 - \alpha\left(\frac{1}{p}\right)^t$ is a negative of an exponential function; it crosses y-axis at $1 - \alpha > 0$. The linear function $-t\ln p$ crosses y-axis at 0 (below $1 - \alpha$) and then have a single crossing with $1 - \alpha\left(\frac{1}{p}\right)^t$. □

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).*

A-2

Proof.

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

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

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

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

$\square$

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

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

A-3

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

$$(1 - p^t)^2 - (c/\mu)(1 - p^t) - (c/\mu)p^t < 0$$
$$(1 - p^t)^2 < c/\mu.$$

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

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

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

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

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

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

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

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

A-4

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 $\frac{d^2 \varphi(t, r)}{dt dr} > 0$. 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). $\qquad\square$

Proof of Proposition 3

(i) 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 $A > \arg\max_{m \geq 0} \{2p^m - p^{2m}\}$. Then S^0 is an equilibrium subscription set; everyone is a subscriber.

Now, suppose that $A \leq \arg\max_{m \geq 0} \{2p^m - p^{2m}\}$. The fact $\varphi(x)$ is converging to zero guarantees that the set $\{m | 2p^m - p^{2m} \leq A\}$ is non-empty. Take $m^* = \min\{m | 2p^m - p^{2m} \leq A\}$. 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).

(ii) is demonstrated in the text. ∎