# Persuasion in Networks:

## a Model with Heterogenous Agents

Melika Liporace*

Bocconi University

September 2021

Latest Version Here

---

**Abstract:**

This paper studies a Bayesian persuasion problem in a connected world. A sender wants to induce receivers to take some actions by committing to a signal structure about a payoff-relevant state. I wonder about the role of a network on information provision when signals are shared among neighbors. Receivers differ in their prior beliefs; the sender wants to persuade some receivers without dissuading the others. I present and characterize novel strategies through which the network is exploited. Were receivers' priors homogenous, such strategies would underperform with respect to a public signal. However, when priors are heterogenous, these strategies can prove useful to the sender. In particular, if the average degree of the nodes who should not be dissuaded is sufficiently low, strategies exploiting the network convince more receivers than public signals, conditional on the adverse state realizing. Furthermore, I show how connectivity can be beneficial to the sender, in particular in segregated networks; and how strategies exploiting the network perform better when one group is especially hard to persuade.

---

**JEL codes:** D83, D85, M30, L82

**Keywords:** Bayesian Persuasion, Network, Heterogenous Priors, Social Media

*I am grateful to Fernando Vega-Redondo for his continued support. I thank Margaret Meyer, Francesco Massazza and Andrea Pasqualini for their insightful comments. I gratefully acknowledge financial support from *Fondazione Invernizzi*. Any mistake is my own.

# 1 Introduction

Information is transmitted through social networks constantly. Social media, for instance, are an ever more prevalent source of information for many. The structure of the network is essential to understand how information fares. But, not only does it constrain the *spread* of information; it also shapes the *content* available. How different is communication on a network?

A sender who wants to communicate a payoff-relevant state to multiple receivers might want to communicate privately to different receivers. This would allow him to tailor information differently to different receivers. This would be particularly relevant if receivers differ substantially, for instance in their priors. Yet, receivers can communicate among themselves. Hence, the information delivered in the network can spread to unintended receivers and lead to suboptimal outcomes. Can the sender exploit a communication network in any way? When should she prefer it to public communication? What role do the differences between receivers and the network structure play?

To answer these questions, I propose a model of persuasion in networks with multiple heterogenous receivers. As is standard with persuasion, the sender is assumed to commit to a signal structure in order to induce as many receivers as possible to take some given action. Receivers are Bayesian and want to take the right action for the right state. Differently from the classical persuasion problem, receivers can hold different priors, which is common knowledge. Priors are such that they would induce different actions. The sender thus needs to persuade one group of receivers without dissuading the other one. Furthermore, receivers are arranged on a communication network. Each receiver does not only observe their own signal realization, but also that of their neighbors.

The sender knows the distribution of connectivity within each group and between groups. She can design the signal structure to allow for different probability distribution for each group; she can correlate signals between groups too. As she does not know the exact network structure, she cannot target nodes individually; she can only design the information conditional on receivers' priors. She wants to maximize the persuasion value, that is, the probability for a receiver taken at random to take her preferred action.

This framework corresponds to many contexts. For instance, a firm might use social media to credibly advertise a new technology. Some users are enthusiastic enough to adopt the technology without further information; others might require proof in order to switch to the firm. I consider the following motivating example: a pharmaceutical company claims that one of their previously approved products is an effective preventive treatment against COVID-19. The company encourages individuals to skip the vaccine and to purchase their treatment instead. A spokesperson can decide to go on television to announce the launch of a clinical trial; and later on present the evidence. Alternatively, the company can reach the same people through Facebook. One or several community managers can announce clinical trials to different users, and follow up with results. The community managers cannot withhold or falsify evidence. However, the company can decide how to hold these clinical trials. In particular, they can decide whether to advertise the same trials for all Facebook users; whether to hire the same scientific team to hold the trials; and how precisely to design each individual trials. Users frequent different groups: some are active on anti-vax groups, and would buy the treatment against COVID-19 without further evidence; the other users frequent pro-vax groups, and need to be convinced about the effectiveness of the treatment.

What would the company do? First, the community managers can do as good as the spokesperson by advertising the same trial to every Facebook user: the persuasion value of a public signal is reproducible in any network. Furthermore, if both types of users have more connection within themselves than between them, the company has no reason to advertise clinical trials to anti-vaxers, as these users would buy the treatment in any case. However, anti-vaxers might have pro-vaxers friends that will display trials result on their wall. How can the company utilize the difference in pattern connections in order to persuade pro-vaxers without dissuading anti-vaxers?

I focus on strategies that ensure for the state favoring the sender to be signaled when it realizes. I propose two novel strategies that exploit the network and show that there does not exist any other. Both strategies rely on a similar mechanism. The sender splits information among different receivers; in particular, she designs the signal structure so that each bit of information is not very informative. Therefore, among agents who need to be persuaded, only those who observe enough bits of information can be persuaded; but as each bit of information is less informative, nodes who should not be dissuaded are less likely to actually detect the adverse state. This relies on the assumption that the nodes who should not be dissuaded observe less signals from the other group than the other group does within itself.

For instance, say that the pharmaceutical company focuses on ensuring that users buy the treatment when the treatment is actually effective. Then, the proposed strategies would be equivalent to the company designing a lot of small scale, not well-controlled trials. The result of any of these trials taken individually would not be sufficient to convince pro-vaxers; but the accumulation of congruent signals would eventually persuade them. As each trial is less likely to give adverse results, anti-vaxers are less likely to come across a result disproving the effectiveness of the treatment, as long as pro-vaxers are friends with more pro-vaxers than anti-vaxers.

The two strategies rely on the same mechanism but play on different parameters of the distribution. While, with what I call a *multiple-message* (MM) strategy, each bit of information is delivered by the actual signal realization; the *network-specific* (NS) strategy exploits correlation of signals. It is then the similarity of signals that is informative about the state. With a MM strategy, nodes must only observe successes in order to take the sender's preferred action; with a NS strategy, receivers must observe all identical signals, i.e. that all or none are successful.

How well do these strategies do? With a public signal, as everybody observes the result from the same trial, they all take the same action; in particular, they buy the treatment if the treatment is effective; if it is not, the results are still sometimes encouraging and they buy it. With a strategy exploiting the network, some pro-vaxers might not be connected enough to ever be receive enough information to buy the treatment. The other pro-vaxers, those who are connected enough, buy the treatment less often than with a public signal; however most anti-vaxers, those who are poorly-connected to pro-vaxers, buy the treatment more often than with a public signal.

Now, if anti-vaxers were as skeptical on the treatment as pro-vaxers, introducing this heterogeneity of informativeness would be suboptimal: the company should go on television! However, when the company does need to convince these two groups differently, strategies exploiting the network can be very useful. When the treatment happens not to be effective, and if the average degree of users susceptible to buy the treatment is small enough compared to the minimum connectivity required to persuade pro-vaxers, the strategies exploiting the network are doing better than the public signal. Furthermore, the NS strategy is better at exploiting degree differences than the MM strategy, so that under the same condition on average degree, the company is better off by relying on correlation rather than on frequency of signal realizations.

I further study these strategies in simplified networks. Take a regular network, that is, pro-vaxers all have the same number of pro-vaxers friends; and anti-vaxers all have the same number of pro-vaxers friends. Then, all pro-vaxers are connected well-enough to be persuaded as often as with a public signal; anti-vaxers, however, are less likely to be dissuaded as they observe the outcome of only a few, poorly informative trials. Therefore, the strategies exploiting a regular network are always better than public communication.

This simplified case underlines how connectivity is not harmful *per se*. The company is happy to divide informativeness into bits that is spread within a well connected group of pro-vaxers; it allows the true effectiveness of the treatment to be hidden to anti-vaxers more often, while still being well disclosed to pro-vaxers. The type of connectivity detrimental to the company is the between group connectivity: the company is better off if anti-vaxers do not talk to too many pro-vaxers. This result has an intuitive link with homophily and segregation on social networks. It underlines how those social habits can be detrimental to the provision of information.

Furthermore, it also emphasize the importance of pro-vaxers' skepticism. The harder are pro-vaxers to convince, the more the strategies exploiting the network outperform public signals. Indeed, when pro-vaxers are so skeptical, a public signal would need to be extremely informative; hence, it is very likely for the anti-vaxers to be dissuaded from purchasing the treatment. It is thus all the more useful to design a strategy that exploits the network and allows to dissuade anti-vaxers less often.

These two intuitions are expected to carry on to general networks. Future versions of this work should contain more precise results in this respect. Furthermore, further research will explore another category of strategies: instead of ensuring for the treatment to be bought when it is effective, the sender could focus on never letting some receivers detect that it does not work. Strategies exploiting the networks can accommodate this objective. Therefore, the same comparison to public communication should be carried out with this new category of strategy. This analysis should eventually allow to study better the role of polarization and its interaction with the network.

## 1.1 Related literature

I mainly contribute to the information design literature, in particular the strands interested in Bayesian persuasion and the endogenous provision of information in networks. To the best of my knowledge, this paper is the first to combine multiple receivers with heterogenous priors and a communication network between receivers.

The seminal work Kamenica and Gentzkow [2011] exposes a persuasion problem with one sender and one receiver. The particularity of this approach is that the sender can commit to a signal distribution about a payoff-relevant state. The authors characterize the sender-optimal signal structures: the sender optimally designs a distribution which is minimally informative to still induce receivers to take the sender's preferred action upon receiving the relevant signal realization. Taneva [2019] generalizes this result by proposing a systematic methodological approach to finding the optimal information structure for the sender in static finite environments. While, in Kamenica and Gentzkow [2011], the emphasis is put on the case of a unique receiver, the authors underline how their analysis can directly be applied to multiple receivers, but only if the latter have homogenous priors and the influence of their actions on the sender's payoff is separable.

The relaxation of either of these two conditions has been explored in the literature. About the latter, a few papers have considered the importance of persuasion in a voting context. For instance, Wang [2013] asks whether private or public communication is preferable for a sender facing strategic voters with heterogenous preferences. Alonso and Câmara [2016] also consider receivers with homogenous priors and heterogenous preferences in a voting context. Kerman et al. [2020] wonder about correlation of private signals between homogenous strategic voters.

The role of heterogenous priors among receivers has seldom been explored. Some authors have studied cases in which priors are unknown to the sender, which could be read as an interpretation of heterogenous priors. Kosterina [2018] studies the problem of a sender who wants to optimize the signal distribution when the worst-case-scenario priors realize. Castiglioni et al. [2021] study an online persuasion problem, in which the sender repeatedly faces receivers with homogenous but unknown priors. Finally, Guo and Shmaya [2019] consider a unique receiver with unknown priors and proposes a nested-interval structure as optimal.

This is reminiscent of Innocenti [2021] who explicitly assumes multiple receivers with heterogenous but known priors. Receivers can hold two different priors, each inducing different actions. He shows how, with a public signal, the tradeoff between persuading some receivers and avoiding to dissuade others lead to two type of strategies, which he coins hard and soft news. These echo, respectively, fully pooling and separating mechanisms in Guo and Shmaya [2019]. I adopt a similar environment with receivers who hold either of two priors inducing different actions. I also borrow Innocenti's terminology. However, rather than assuming a public signal, I allow for partially private signals, by introducing a network structure and the possibility for the sender to target different groups with different signal structures, possibly correlated.

A persuasion problem involving receivers communicating over signals in a network is the object of Kerman and Tenev [2021]'s analysis. However, the authors assume homogenous receivers. Furthermore, in their model, the network structure is known to the sender who can thus target individual nodes. Finally, their analysis takes place in a voting context; by contrast I consider a payoff for the sender that is linear in the number of receivers taking the preferred action. Reminiscent of some of the insights from my model, Kerman and Tenev [2021] find that a higher network density does not necessarily translate into a decrease in the sender's gain from persuasion.

More generally, I contribute to the information design literature. Bergemann and Morris [2019] provides an overview of the literature by unifying the strand of the literature interested in Bayesian persuasion with the rest of the literature. The authors underline how the presence of multiple receivers naturally raises the question of the sender's preference for public or private information. While they consider the possibility for the sender to have preferences or strategical interests in correlated actions, they ignore the possibility for agents to communicate among themselves. My paper contributes to endogenize how privately signals are communicated, subject to the constraints of the communication network.

Finally, a few recent papers have introduced networks in information design. Egorov and Sonin [2020] study the receivers choice to subscribe to a signal structure designed by a sender. Candogan [2019] considers the problem of a sender who wants to persuade receivers arranged on a network by communicating publicly. The receivers payoffs are subject to strategic complementarities. Finally, Galperti and Perego [2019] provide a general framework to study the maximal impact that information revealed to some seeds in a network can have. Evocative of the strategies I propose to exploit the network in the specific context considered, Galperti and Perego underline how messages can be coded to be understood only be nodes who receive all individual signals composing the message.

The remainder of the paper is organized as follows. The model is presented in Section 2. In Section 3, the private and public information cases are explained and their relation to networks are underlined. Section 4 proposes novel strategies exploiting the network, studies their performance and explores future developments. Section 5 concludes.

# 2 Model

## 2.1 Environment

Consider a persuasion problem with one sender (she) and $N$ receivers (they/he). They communicate over a payoff-relevant binary state $\omega \in \{0, 1\}$. I allow for agents to hold different priors over the state. The sender's prior is denoted $\mu := \Pr(\omega = 1)$. The receivers are partitioned in two groups $A$ and $B$. Let $a := A/N$ and $b := B/N$ so that $a + b = 1$.[1] The groups are characterized by their members' prior beliefs, with $\mu_A > \mu_B$. The group to which a node belongs is common

[1] To ease notation, $A$ and $B$ refer to both the sets of nodes and its cardinality.

knowledge among all players.

Receivers are arranged on an undirected network.[2] Any node $i$ observes his own signal realization $s_i \in \{0,1\}$. In addition, he observes the signal realization of all of his neighbors. Let $\mathcal{N}_i$ be $i$'s neighborhood including himself; and $\mathcal{N}_i^I := \mathcal{N}_i \cap I$ for $I \in \{A, B\}$ be $i$'s $I$-neighborhood. Finally, $i$'s $I$-degree, denoted $d_i^I$ is defined as $|\mathcal{N}_i^I|$.

Both the network structure and the communication patterns are exogenous. Furthermore, the network's topology is unknown to the sender; only the degree distribution is accessible to her.[3] Denote $\delta_I(d_i^I)$ the portion of node in group $I \in \{A, B\}$ who have $d_i$ neighbors belonging to the same group; and $\delta_I(d_i^{-I})$ the portion of node in $I$ who have $d_i$ neighbors belonging to the other group.

Receivers update their beliefs about the state according to Bayes' rule. Let $\beta_I(m_i)$ be the belief of agent $i$ from group $I$ after observing $m_i$, where $m_i$ is the sum of relevant signal realizations, possibly accounting for group belonging.[4] Given the binary nature of the state, $m_i$ contains all information accessible to $i$. The index $i$ is kept in $m_i$ to account for the $i$'s neighborhood's size, that is, the number of signals observed.

Finally, as the sender can positively correlate signals, the probability for the joint realization of signals must be defined. I assume that the number of success for dependent and simultaneous Bernoulli trials is characterized by the following distribution:

$$\tilde{p}_k^n = \Pr(X = k|r, n, \alpha) = (1-\alpha)\binom{n}{k} r^k (1-r)^{n-k} + \alpha\big[(1-r)\mathbb{1}_{k=0} + r\mathbb{1}_{k=n}\big]$$

where $X$ is the number of successes in a sequence of $n$ dependent trials, $r$ is the success probability of any trial and $\alpha$ is the pair-wise correlation of any two trials. Appendix A details how this distribution is consistent with the parameters $r$ and $\alpha$. For the remainder of this work, let $p_k^n$ be the probability of $k$ successes among $n$ independent trials, that is, the probability prescribed by the standard Binomial variable.

## 2.2 Objectives, Timing and Equilibrium Concept

All receivers have the same preferences. Their payoff is such that there is a unique preferred action for any realized state. The optimal action given a state is denoted by the state. For instance, if receivers were to know that the state is 0, they would take action 0. However, because they cannot observe the state, their action is conditional on their belief $\beta(m_i)$. Let $t$ be the belief that makes them indifferent between action 0 and 1. They are assumed to take action 1 if and only if they believe with probability greater or equal to $t$ that the state is 1, i.e. iff $\beta(m_i) \geq t$. Furthermore, $t$ is such that the receivers' different priors induce different actions, i.e. $\mu_A > t > \mu_B$. I denote $\alpha_A := \frac{\mu_A(1-t)}{(1-\mu_A)t}$ and $\alpha_B := \frac{\mu_B(1-t)}{(1-\mu_B)t}$, so that $\alpha_A > 1 > \alpha_B$. $\alpha_I$ indicates how

[2] The analysis in a directed network would be identical. Instead of degrees, the sender would focus on in-degrees. The rest of the analysis would not be affected.

[3] This can be interpreted as a random network, or as limited information on the side of the sender.

[4] Relevant signals are those which are informative. For now, we can define $m_i := (m_i^A, m_i^B)$ where $m_i^I := \sum_{j \in \mathcal{N}_i^I} s_j$. The definition is provided in Section 3.2.

much information is needed to make them change action.

The sender's objective is to induce receivers to take her preferred action, regardless of the state. Without loss of generality, I assume her preferred action is 1. I refer to 1 as the *favorable* action or state. The sender only cares about the expected number of receivers taking the favorable action.[5] I denote $V$ the *value of persuasion*, i.e. the sender's expected payoff from the persuasion problem; it is defined as the probability that a receiver takes action 1. In order to persuade receivers, she commits to a signal structure $\pi$. Conditional on the state, the distribution specifies a probability of success for each group $I$, $\Pr(s_I = 1|\omega)$, as well as the signal correlation within groups $\mathrm{Corr}(s_I, s_I|\omega) \geq 0$ and between groups $\mathrm{Corr}(s_A, s_B|\omega) \geq 0$. The notation is as follows: $p_I := \Pr(s_I = 1|\omega = 0)$, $q_I := \Pr(s_I = 1|\omega = 1)$, $\rho_{IJ} := \mathrm{Corr}(s_I, s_J|\omega = 0)$ and $\varphi_{IJ} := \mathrm{Corr}(s_I, s_J|\omega = 1)$.

The game is played sequentially. The timing is the following:

1. The sender commits to a signal structure $\pi$.
2. All uncertainty realizes, in particular $\omega$ and $m_i \forall i \in N$.
   Receivers observe $\pi$ and $m$. They update their beliefs about $\omega$ using Bayes rule.
   Given their updated beliefs $\beta(m_i)$, they take action $a_i \in \{0, 1\}$.

Note that receivers observe $\pi$ fully, including correlations between signals. I look at Subgame Perfect Equilibria (SPE), using backward induction.

# 3 Benchmarks: Private and Public Information

## 3.1 Unconnected World

For now, let us abstract from the existence of a network. A sender facing multiple receivers might be able to communicate with each of them in isolation or, alternatively, might need to commit to signals that are publicly observed. Those are referred to as the private information and public information cases respectively.

In the private information case, the sender can optimize over each receiver separately. Indeed, because each signal realization is only observed by the given receiver, the sender does not need to worry about information spreading to unintended receivers. Furthermore, because her payoff is separable in each receiver's decision, she can optimize the signal structure for each receiver separately. Her strategy therefore corresponds to the one prescribed in a standard persuasion problem, i.e. in Kamenica and Gentzkow [2011].

[5] Implicitly, this means that she is risk neutral, as is standard with Bayesian persuasion. However, in the context of multiple receivers, this assumption is stronger. Each receivers' action needs to enter the sender's payoff linearly. Therefore, I abstract from global social effects. For instance, I ignore cases in which the sender might benefit increasingly or decreasingly from each marginal adopter; or in which she needs a quota of receivers to take the action, such as voting context.

The sender optimally designs a signal structure which is minimally informative but informative enough to persuade receivers upon receiving some given signal realization. It translates into $q_I = 1$ and $p_I = \max\{\alpha_I, 1\}$. I refer to it as the *standard* strategy. Any receiver belonging to group A would take action 1 without any information; so $p_A = 1 = q_A$.[6] For any receiver belonging to $B$, the sender sets $q_B = 1$ and $p_B = \alpha_B$.

As for correlations, since receivers only observe their own signal, any correlation structure would be optimal as long as it is consistent with the signal described above, i.e. such that $\rho_{AB} = 0$. In particular, $\rho_{II} = \varphi_{II} = 1$ would correspond to one common signal realization for all members of each group $I$, and would lead to the same outcome, in expectation, as letting all agents get a different signal realizations $\rho_{II} = \varphi_{II} = 0$. The persuasion value is:

$$V_{PI} = \mu + (1-\mu)[a + \alpha_B b]$$

When signal is public, there exists a tradeoff between persuading agents in $B$ and avoiding members of $A$ to be dissuaded. Innocenti [2021] studies the case of a unique public signal and shows that there are two potential optimal strategies: persuade both groups as often as possible with the favorable signal, but sometimes dissuading group $A$ – a *hard news* strategy –; or never dissuade group $A$ at the price of persuading group $B$ less often – a *soft news* strategy. The former implies $q = 1, p = \alpha_B$;[7] the latter prescribes $q = \frac{\alpha_A - 1}{\alpha_A - \alpha_B}, p = \alpha_B q$. The persuasion value are respectively:

$$V_{HN} = \mu + (1-\mu)\alpha_B \quad \text{and} \quad V_{SN} = a + b\Big[\mu q + (1-\mu)\alpha_B q\Big]$$

The determination of the optimal strategy between these two candidates depends on: the polarization of prior beliefs, defined as $\alpha_A - \alpha_B$; the probability to be in an unfavorable state according to the sender, $\mu$; and the groups' relative sizes, $a, b$.

Because all signal realizations are observable, the sender can replicate the public information case in an unconnected world by sending the same signal realization to every receiver – $\rho_{IJ} = \varphi_{IJ} = 1$; or by sending any amount of informative signals, as long as the set of signals inducing agents to take the favorable action realizes with the probabilities specified above.

### 3.2 Implementation in a Network

These two benchmarks can relate to a connected world. An empty network – i.e. $\delta_A(d^A = 1) = \delta_B(d^B = 1) = 1$ – corresponds to the private information case: the sender can persuade each group in isolation. Actually, any network in which there are no connections between group $A$ and group $B$ – i.e. $\delta_A(d^B = 0) = 1$ – can deliver the same value of persuasion. However, the correlation structure now matters. Setting $\rho_{II} = \varphi_{II} = 1$ is optimal. [8]

[6]Any signal $\pi$ with $p_A = q_A$ would be uninformative, so the strategy specified in the main text is not unique.

[7]Because the signal is public, $q_A = q_B$ and $p_A = p_B$. I omit the index in this case.

[8]If the degree distribution within $B$ was homogenous, the correlation structure would be irrelevant as long as the set of signal realizations inducing nodes in $B$ to take the favorable action occurs with the optimal probability. If the degree distribution within $B$ is not homogenous, however, not perfectly correlated signals introduces a suboptimal heterogeneity in the effective informativeness of the signal structure. This intuition will be made

On the other hand, a complete network – i.e. $\delta_I(d^A = A) = \delta_I(d^B = B) = 1$ – corresponds to the public information case: every agent can observe the signal realization of every other agent. However, the public information case is reproducible in any network.

**Lemma 1.** *The persuasion value of the public information is reproducible in any network by setting* $\rho_{IJ} = \varphi_{IJ} = 1$.

*Proof.* If $\rho_{IJ} = \varphi_{IJ} = 1$, then $\forall i, j \in N, s_i = s_j$. Because receivers are fully Bayesian and observe correlations in $\pi$, their effective information set is composed of a unique signal realization, common to every receiver. This is equivalent to the sender designing a unique signal observed by everyone. □

In other words, the presence of a network can never hurt the sender. If anything, she might prefer sending a common message rather than exploiting the network. Therefore, if the network is exploited, it is because it improves upon the persuasion value with respect to a public signal. On the other hand, a private message is the best the sender can hope for. Therefore, the value persuasion in a network will never exceed that of private communication.

**Remark 1.** The value of persuasion in a network with heterogenous priors is bounded between the value of persuasion of the public information case and that of the private information case.

Intuitively, the sender would like to provide information on the state to $B$ without it spreading to $A$. Therefore, it can never be optimal to provide information to $A$, even if the sender relies on this to inform $B$. Indeed, in the latter case, providing that information to $B$ would be (at least weakly) preferable.

**Lemma 2.** *The sender designs* $\pi$ *so that* $(s_i)_{i \in I}$ *is uninformative for some group* $I \in \{A, B\}$, *that is* $p_I = q_I$, $\rho_{II} = \varphi_{II}$ *and* $\rho_{IJ} = \varphi_{IJ} = 0$.

*Proof.* Let us assume that $\pi$ is designed to deliver informative signals to $B$. As members of group $A$ do not need to be persuaded, any informativeness spent on them would be wasted. The sender does not need to exploit signals to group $A$ in order to inform members of group $B$, as this can be achieved through targeting nodes in $B$ directly. Therefore, provide group $A$ with informativeness would reduce the probability for agents belonging to $A$ to take the favorable action without allowing to increase the number of nodes in $B$ taking the actions beyond what is achievable through the informativeness to $B$. The same argument applies if informative signals are delivered to $A$, in order to convince $B$. □

**Remark 2.** Because the sender designs uninformative signals for some group $I$, the distribution parameters related to $s_I$ are irrelevant as long as they are consistent with Lemma 2. The group to which to provide informativeness is the group with higher connectivity with $B$ members than with $A$ members.[9] For the remainder of the analysis, $s_A$ will be assumed to be uninformative.

more formal in Lemma 4

[9]See Theorem 2.

To ease notation, denote $p_B = p$, $q_B = q$, $d_i^B = d_i$ $\rho_{BB} = \rho$ and $\varphi_{BB} = \varphi$. Note that all results would carry on if only $A$ was delivered informative news, reinterpreting in particular $d_i$ as nodes' $A$-degree. Furthermore, without loss of generality I focus on cases with $p \geq 1/2$.[10]

This section closes with a few formal concepts that should allow for more efficient analysis.

**Definition 1.** The set of *messages* a node $i \in I$ can receive is $\mathcal{M}_i := \{m_i \in \mathbb{N} : m_i = \sum_{j \in \mathcal{N}_i^B} s_j\}$.

**Definition 1a.** The set of *favorable messages* for node $i \in I$ is $\mathcal{FM}_i := \{m_i \in \mathcal{M}_i : \beta_I(m_i) \geq t\}$.

**Definition 1b.** The set of *persuading messages* for node $i \in I$ is $\mathcal{PM}_i := \{m_i \in \mathcal{M}_i : \beta_I(m_i) = t\}$.

A message is simply a realization of signals in $i$'s $B$-neighborhood, that is aggregated but contains all relevant information. A favorable message corresponds to a realization of signals in $i$'s neighborhood that would induce $i$ to take the favorable action. The set of *unfavorable* messages is similarly defined as the set of signal realization that would induce $i$ to take the unfavorable action, i.e. $\mathcal{M}_i - \mathcal{FM}_i$. The set of persuading messages corresponds to the messages that would make node $i$ indifferent between action 0 and 1. This set can be empty.

**Definition 2a.** The *targeted nodes* is the subset of nodes for whom the set of persuading messages is non-empty, i.e. $\{i \in N : \mathcal{PM}_i \neq \varnothing\}$.

**Definition 2b.** The *susceptible nodes* is the subset of nodes for whom the set of favorable messages is non-empty, i.e. $\{i \in N : \mathcal{FM}_i \neq \varnothing\}$.

The targeted nodes are the nodes that the sender constraints herself to make indifferent between actions. In standard Bayesian persuasion, the sender uses minimal informativeness to persuade nodes as often as possible. In that case, such persuaded nodes are targeted, because they are the ones made *just* indifferent between actions. If nodes differ in terms of priors, persuading some nodes as often as possible might result in dissuading the others too often. Therefore, the sender might want to avoid ever dissuading some other nodes. Then, such non-dissuaded nodes are also targeted: sender still makes them indifferent; she wants to send a favorable message to the former nodes as often as possible given that it does not cause the latter nodes to ever be dissuaded. In the networks considered, in addition to their priors, nodes might differ in their $B$-degree; a same realization of signals in the network could then lead to different posteriors. Therefore, the sender might be able to only target a subset of nodes, defined not only by their priors but also by their degree.

The susceptible nodes are simply those susceptible to receive favorable message. In other words, for each of them, there exists a signal realization that would induce them to take the favorable action. In standard Bayesian persuasion, whether or not agents hold heterogenous priors, all nodes are susceptible under an optimal information structure. However, in a network, posteriors might depend on the degree in addition to the group belonging; this means that $B$

[10]This relates to the general irrelevance of the content of the signal; only the distribution of signals matter to interpret them. To simplify exposition, I do not expose strategies that mirror the ones discussed with inverted signal realizations.

nodes with some certain degrees might never reach a posterior high enough to be persuaded, irrespective of the realization of signals in the network, such nodes would not be susceptible. They would be referred to as *non-susceptible*.

# 4 The Role of the Network

## 4.1 Strategies Exploiting the Network

**Definition 3.** A hard news strategy is a signal structure $\pi$ such that for any targeted node $i$, $\Pr(m_i \in \mathcal{PM}_i | \omega = 1) = 1$

**Definition 3a.** The *unconnected* (U) hard news strategy is a signal structure $\pi$ that is independent of any of the targeted nodes' degree and in which all nodes $i \in B$ are targeted.

**Definition 3b.** A *multiple-message* (MM) hard news strategy is a signal structure $\pi$ that is dependent of the targeted nodes' degree and in which for any targeted node $i \in B$, $\mathcal{PM}_i = \{d_i\}$.

**Definition 3c.** A *network-specific* (NS) hard news strategy is a signal structure $\pi$ that is dependent of the targeted nodes' degree and in which for any targeted node $i \in B$, $\mathcal{PM}_i = \{0, d_i\}$.

As introduced by Innocenti [2021], a hard news strategy convinces targeted nodes in $B$ as often as possible. The unconnected hard news strategy generally replicates the persuasion value of the public information case.[11]

The multiple-message strategy relies on a message composed of multiple successful signals in order to convince targeted nodes. Hence, such strategy is irrelevant if agents do not observe multiple signals. The sender designs $\pi$ such that each individual signal realization is less informative than under an unconnected hard news strategy. Therefore, only the accumulation of successful signals can persuade agents in $B$; however, it is also less likely for nodes in $A$ to observe an unfavorable message as the probability for any signal realization to be a success is greater under this strategy.

The network-specific strategy relies on the information contained in the similarity of signal realizations rather than on individual realizations. Hence, it cannot exist outside a network. The sender informs receivers about the state through correlation rather than a probability of success. Therefore, individual signal realizations are not informative *per se*; it is a consensus among messages that should inform receivers on the state. In particular, the sender renders the realization of identical signals more likely in the favorable state than in the unfavorable one, so that agents are persuaded by observing identical signals in their neighborhood. As the MM strategy, the NS strategy capitalizes on the $B$-degree difference among nodes of different groups: it is more likely for nodes in $A$ than in $B$ to observe identical but uncorrelated signal realizations in their $B$-neighborhood because their $B$-neighborhood is smaller.

[11] It does in networks such that $\delta_A(d_i = 0) = 0$. The persuasion value of this strategy is provided in Proposition 1.

I consider multiple-messages and network-specific strategies that target members of $B$ with a given $B$-degree. The $B$-degree of the targeted nodes is denoted $\hat{d}$. Note that these strategies could be declined to never dissuade targeted agents in $A$; this would correspond to a soft-news strategy. These strategies are omitted in the current version of this work.

However, a natural question arises: are there other hard news strategy that should be considered? I show that there does not exist any such strategies.

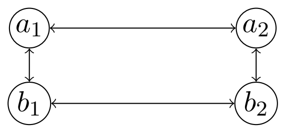
**Theorem 1.** *If a signal structure is a hard news strategy, then it is either an unconnected strategy, or a multiple-message strategy, or a network-specific strategy.*

*Proof.* See Appendix A. □

The proof relies on a very intuitive argument: there are only limited tools to insure that all targeted nodes receive a persuading message in the favorable state. In particular, with a hard news strategy, nodes who observe an unfavorable message perfectly infer that the unfavorable state of the world realized. Therefore, any a persuading message different from 0 or $d_i$ would require for the set of all messages to be persuading; but this is impossible, as the signal structure must be informative in order to persuade nodes.

Example 1 below illustrates these strategies in a simple network and shows how they can outperform public information.

**Example 1.** Assume $\mu = t = 0.5, \mu_A = 2/3, \mu_B = 1/3$ so that $\alpha_A = 2, \alpha_B = 1/2$. Furthermore, consider the network depicted below.




The hard news strategy with $\rho = \varphi = 1$ would lead to a value of persuasion $V_{HN} = \mu + (1-\mu)\alpha_B = 3/4$; the soft news strategy with $\rho = \varphi = 1$ would mean $q = 2/3$ and $p = 1/3$ so that $V_{SN} = 1/2 + 1/2[\mu \cdot q + (1-\mu) \cdot p] = 3/4$. If the sender was able to communicate privately within each group, the value of persuasion would become $V_{PI} = 1/2 + 1/2[\mu + (1-\mu) \cdot \alpha_B] = 7/8$.

Now, the two alternative strategies MM and NS would entail the following specification:

1. For MM, the sender does not need to use correlations. Hence she sets $\rho = \varphi = 0$. One can study the sender's strategy using the following table:

| | $\omega = 0$ | $\omega = 1$ |
|---|---|---|
| $m_B = 0$ | $(1-p)^2$ | $(1-q)^2$ |
| $m_B = 1$ | $2p(1-p)$ | $2q(1-q)$ |
| $m_B = 2$ | $p^2$ | $q^2$ |

Because the sender is designing a hard news strategy, $q = 1$. Agents in $B$ are persuaded iff they observe $s_{b_1} = s_{b_2} = 1$. Therefore, the sender would like to maximize the probability of this realization in the unfavorable state. She would thus set: $p^2 = \alpha_B$. It means that $m_B < 2$ is an unfavorable message for all members of $B$. However, if $m_B = 1$ was to realize, only half of $A$'s member would be dissuaded because only half would observe $s = 0$. This is an advantage compared to the unconnected hard news strategy, that dissuades all nodes in $A$ upon the realization of an unfavorable message. In other words, because $p_{MM} = \sqrt{2}^{-1} > 2^{-1} = p_{HN}$, it is less likely for nodes in A to observe an unfavorable message signal realizations of their neighbors in $B$. Hence the persuasion value under MM is: $V_{MM} = \mu + (1-\mu)p_B^2 + (1-\mu)\frac{a}{2}2p(1-p) = 3/4 + 1/2 \cdot 1/4 \cdot 2\frac{\sqrt{2}-1}{2} \approx 13/16$

2. In this case, the sender mainly uses correlations. Let her set $\rho = 0$ and $\varphi = 1$. The sender's strategy is now represented by :

| | $\omega = 0$ | $\omega = 1$ |
|---|---|---|
| $m_B$=0 | $(1-p)^2$ | $1-q$ |
| $m_B$=1 | $2p(1-p)$ | 0 |
| $m_B$=2 | $p^2$ | $q$ |

The set of persuading messages contains both $m_B = 0$ and $m_B = 2$. For this, she needs $(1-q)\alpha \geq (1-p)^2$ and $q\alpha \geq p^2$ so $p^2 + (1-p)^2 \leq \alpha$. Because $V$ is increasing in the probability that $m_B = 0$ or $m_B = 2$ realizes, the constraint is binding. The sender sets $p_B = 1/2 = q_B$. This replicates the persuasion value of private information. Indeed, $a_1$ and $a_2$ would never be dissuaded, as observing one signal realization of $B$ is fully uninformative. Hence, members of group $A$ are always persuaded. On the other hand, $b_1$ and $b_2$ are persuaded upon seeing $s_{b_1} = s_{b_2}$. In the unfavorable state, this persuasion occurs with probability $Pr(s_1 = s_2|\omega = 0) = p^2 + (1-p)^2 = 1/2$. Therefore, the persuasion value under NS is: $V_{NS} = \mu + (1-\mu)(a + b\alpha_B) = 7/8$.

Let us generalize the insights from Example 1 to any network. Recall that $\hat{d}$ is the $B$-degree of the targeted nodes. $\hat{d}$ corresponds to the minimum number of signal realizations agents in $B$ need to observe in order to potentially be persuaded, with either MM or NS. For MM, these realizations all need to be a success. For NS, these realizations all need to be the same.

**Lemma 3.** *Consider any MM or NS strategy.*

*(i)* $\forall i \in A, \mathcal{FM}_i \neq \varnothing$ *and* $\forall i \in B, d_i \geq \hat{d} \Leftrightarrow \mathcal{FM}_i \neq \varnothing$

*(ii)* $\forall i \in N : \mathcal{FM}_i \neq \varnothing,\ \Pr(m_i \in \mathcal{FM}_i|\omega = 1) = 1$

*(iii)* $\forall i \in N : \mathcal{FM}_i \neq \varnothing,\ \Pr(m_i \in \mathcal{FM}_i)$ *is decreasing in* $d_i$.

*Proof.* See Appendix A. □

The first result characterizes the susceptible nodes in strategies exploiting the network. Because the multiple-message and network-specific strategies exploit the differences of connectivity within $B$ and between $A$ and $B$, they are meant to convey different level of informativeness to nodes with different $B$-degree. However, this comes at the price of introducing heterogeneity within groups of otherwise identical nodes. In particular, if members of $B$ are too scarcely-connected in $B$, they do not observe enough signals and thus perceive messages to be too little informative for them to ever be induced to take the favorable action.

Furthermore, the definition of hard news strategy insures that all susceptible nodes are induced to take the favorable action in the favorable state of the world. The only realization of signals permitted by a hard news strategy are easily characterized: because targeted nodes must observe a persuading message with probability 1, either all signals to $B$ are successes, or, all signals to $B$ are identical. In either case, this insures that all susceptible nodes receive a favorable message.

Finally, the probability that a susceptible node receives a favorable signal is decreasing in the node's $B$-degree. Indeed, the more signals a node observes, the more informative his message. Therefore, if nodes are *too* connected to members of $B$, they are induced to take the favorable action less often than with a public signal. In such a case, they are hit by messages that are too informative. This applies to both groups.

Proposition 1 below characterizes the persuasion value of such strategies in a general network.

**Proposition 1.**

*(i) The unconnected strategy targeting all agents in $B$ is such that $\varphi = \rho = q = 1$ and $p = \alpha_B$. The persuasion value associated with such strategy $V_U = \mu + (1-\mu)\left[\alpha_B + a\ \delta_A(0)(1-\alpha_B)\right]$.*

*(ii) The multiple-message strategy targeting agents in $B$ whose $B$-degree is $\hat{d}$ is such that $q = 1$ and $\rho p + (1-\rho)p^{\hat{d}} = \alpha_B$. The persuasion value associated, $V_{MM}$, can be written as:*

$$\mu\left[a + b\sum_{d_i=\hat{d}}^{B}\delta_B(d_i)\right] + (1-\mu)\left[a\delta_A(0) + \sum_{d_i=1}^{B}\left[a\delta_A(d_i) + \mathbb{1}_{d_i\geq\hat{d}}\ b\delta_B(d_i)\right]\left[\rho\ p + (1-\rho)p^{d_i}\right]\right]$$

*(iii) The network-specific strategy targeting agents in $B$ whose $B$-degree is $\hat{d}$ is such that $\varphi = 1$, $\rho + (1-\rho)\left(p^{\hat{d}} + (1-p)^{\hat{d}}\right) = \alpha_B$, and $q = \alpha_B^{-1}\left(\rho p + (1-\rho)p^{\hat{d}}\right)$. The persuasion value associated, $V_{NS}$, can be written as:*

$$\mu\left[a+b\sum_{d_i=\hat{d}}^{B}\delta_B(d_i)\right]+(1-\mu)\left[a\delta_A(0)+\sum_{d_i=1}^{B}\left[a\delta_A(d_i)+\mathbb{1}_{(d_i\geq\hat{d})}\ b\delta_B(d_i)\right]\left[\rho+(1-\rho)\left(p^{d_i}+(1-p)^{d_i}\right)\right]\right]$$

*This strategy is implementable if and only if $\alpha_B > 2^{-(\hat{d}-1)}$.*

*Proof.* See Appendix A □

While the technical details of the proofs are reported in the appendix, the expressions for $V_{MM}$ and $V_{NS}$ are intuitive. When the favorable state realizes, $s_i = 1$ for all $i \in B$ if the MM strategy is adopted; and $s_i = s_j$ for all $i, j \in B$ if the NS strategy is adopted. This cannot dissuade any agent in $A$. Furthermore, the realization of signals persuades all agents with $d_i \geq \hat{d}$. Indeed, members of group $B$ require at least $\hat{d}$ sucesses or identical signals to be persuaded. When the unfavorable state realizes, nodes of either group are taking the unfavorable action as soon as they observe: an unsuccessful signal realization for MM; or a signal realization for some neighbor in $B$ that is different from the others for NS. For any node with $B$-degree $d_i$, the probability to observe a favorable message is $\sum_{s=0}^{B-d_i} \tilde{p}^B_{s+d_i}$. In case of NS, the probability to observe a favorable message is $\sum_{s=0}^{B-d_i} \tilde{p}^B_{s+d_i} + \sum_{s=0}^{B-d_i} \tilde{p}^B_s$.

The general characterization of the optimal $\hat{d}$ is left for a future version of this work. However the next sections, I provide sufficient conditions to fully characterize the signal structure of each strategy and the ranking of persuasion value for different strategies.

### 4.2 Homogenous Priors

First, let us compare the different strategies when agents do not differ in their prior. In such a context, the standard strategy and the unconnected hard news strategy are the same. It is however interesting to study how well the hard news MM and NS strategies do.

**Lemma 4.** *For any random network populated by receivers with homogenous priors, the multiple-message and network-specific strategies weakly underperform with respect to the standard strategy.*

*Proof.* See Appendix A. □

Intuitively, the MM and NS strategies level on the potential differences in connectivity patterns between groups with different priors. These strategies allow group $A$, which is expected to be connected to $B$ less than $B$ is to itself, to observe a less informative signal exactly because of their lack of connectivity. But this comes at a price: the strategy will perfectly target nodes with some degree, but is too informative for nodes that are more connected, and not informative enough for agents that are less connected.

Now, when the priors are homogenous, it is suboptimal to let group $A$ with less informative signals. In other words, when groups have the same priors, exploiting the network introduces suboptimal heterogeneity. In particular, the sender loses all nodes with $d_i < \hat{d}$ in any state, and she wastes informativeness on nodes with $d_i > \hat{d}$ in the unfavorable state.

### 4.3 Heterogenous Priors

When receivers hold different priors, the strategies exploiting the network allow the sender to design a signal structure that is less informative for some nodes. This is a double edge-sword: on the one hand, the sender can design $\pi$ to be less informative to some nodes in $A$; on the other hand, this might cause some nodes to receive too informative messages.

Theorem 2 provides sufficient conditions to characterize the optimal signal structure with such strategies and for such strategies to be beneficial to the sender.

**Theorem 2.** *Denote* $\tilde{\delta}(d_i) := a\delta_A(d_i) + \mathbb{1}_{d_i \geq \hat{d}}\, b\delta_B(d_i)$.

*(i) If* $\mathbb{E}_{\tilde{\delta}}(d_i) \leq \hat{d}$, $\rho_{MM} = 0$ *and* $V_{MM} < V_{NS}$. *If furthermore* $\mathbb{E}_{\tilde{\delta}}({d_i}^2) \leq \hat{d} \cdot \mathbb{E}_{\tilde{\delta}}(d_i)$, $\rho_{NS} = 0$.

*(ii) If* $\mathbb{E}_{\tilde{\delta}}(d_i) \leq \hat{d}$, *when the unfavorable state realizes, the multiple-message strategy outperforms the unconnected strategy.*

The sender designing MM or NS strategies has to tradeoff $\rho$ for $p$, as both are decreasing the informativeness of favorable messages. Why does the sender prefer increasing $p$ than $\rho$?

For the MM strategy, if the sender was decreasing $p$ in order to increase $\rho$, it would increase the probability for a node in $A$ to observe an unfavorable message, and hence to be dissuaded. Intuitively, a MM signal structure is designed to diversify the type of unfavorable messages that realize for a node in $B$. While all unfavorable messages have the same consequence on $B$, the ones containing more successes have a lower chance to dissuade members of $A$ that have a lower $B$-degree. Therefore, because she is constraint, the sender prefers to allow for different signal realizations.

This intuition applies if the gain from exploiting the lower connectivity of $A$ members overcompensates for the nodes who are hit with *too much* informativeness due to their higher connectivity. Indeed, when the sender relies on a MM strategy, it is less likely for a poorly-connected member of $A$ to see an unfavorable message, but it is impossible for a poorly-connected member of $B$ to observe a favorable message. Because of the rate at which the probability varies with connectivity, it is sufficient – but not necessary – that the average degree of susceptible nodes is lower than the degree of targeted nodes.

To sum up, the strategy relies on the difference in connection patterns between groups. Correlating messages renders this difference less salient. Hence, it is suboptimal for the sender to correlate signals when this difference compensates the excess of informativeness for well-connected nodes and the loss of non-susceptible members of $B$.

A similar reasoning applies to the NS strategy. Only successes in a node's neighborhood can be due to highly correlated signals, or to a very high probability for any of them to be a success. As before, agents with bigger $B$-neighborhood are better at distinguishing whether signals are identical out of luck – that is, because of a high (or low) probability of success – or because of correlation. For instance, the probability for two uncorrelated signals to be identical, even for outcomes as random as $p = 1/2$, is relatively high, 50% in this case; but it vanishes fast when more signals are considered, to slightly more than 6% in the example considered if five uncorrelated signals are observed. If the signals are correlated, the number of signals observed becomes less and less relevant in determining the informativeness of the message. Because this effect is more extreme with changes in $d_i$, the sufficient condition for the NS strategy, compared to the MM one, is more demanding, as it does not only consider average degree, but also include a notion of the variance of degrees of susceptible nodes.

The second result in Theorem 2 underlines that the NS strategy is better at capitalizing on degree differences than the MM strategy when the unfavorable state of the world realizes. The effective informativeness of messages issued from a strategy designed for a given $B$-degree varies more for different $d_i$ with NS than with MM. As long as the average degree of susceptible nodes is lower than the degree of targeted nodes, the average probability for the susceptible nodes to actually receive a favorable message is also higher with NS than with MM. Because both strategies exploit the network, both might equally lose non-susceptible nodes when the favorable state of the world realizes.

A similar intuition applies regarding the comparison between MM and U strategies. When the unfavorable state of the world realizes, MM might, on average, deliver more favorable signals than U, as the higher probability of favorable signals for poorly-connected signals in $A$ more than compensates the lower probability of favorable signals for highly connected nodes. This is indeed the case if the probability of a favorable signal is higher for a node with the average degree of susceptible node than the minimum probability required for agents in $B$ to be persuade, which is the probability of persuasion upon the unfavorable state realizing. However, in the favorable state, the unconnected strategy convinces all nodes, while the multiple-message strategy loses the non-susceptible nodes.

The relative size of these effects depends on the parameters. In particular, one would expect that losing the non-susceptible is compensated by the higher average probability of persuading nodes with the MM strategy when the favorable state of the world is rather unlikely or when few nodes are non-susceptible.

**Corollary 1.** *For $\mathbb{E}_{\tilde{\delta}}(d_i) < \hat{d}$, $V_{NS} > V_{MM} > V_U$ for relatively small $\mu$, $b$ or $\sum_{d_i=1}^{\hat{d}} \delta_B(d_i)$.*

*Proof.* See Appendix A □

To compare the strategies exploiting the network to the unconnected strategy, one needs to compare the benefits of MM against U in the unfavorable state of the world with the loss from MM against U in the favorable state of the world; as well as the relative probability that each state will occur.

When $\mu$ is small, the unfavorable state of the world is much more likely than the favorable one, so it is much more likely for benefits to occur than losses, making MM better than U. Likewise, when $b$ or $\sum_{d_i=1}^{\hat{d}} \delta_B(d_i)$ are small, the magnitude of the cost is rather small, so that the benefits from the unfavorable state are more likely to compensate for the small costs from the favorable state.

To better understand the performance of strategies exploiting the network when receivers differ in their prior, I consider specific or simplified network structures below.

### 4.3.1 Regular Networks

Generalizing example 1, consider a network in which all nodes within one group have the same number of $B$-neighbors. In such a context, the targeted nodes are straightforward to define.

This approach also has the advantage of a reduced set of relevant parameters, which still allows powerful predictions.

Let $d_B$ be the $B$-degree of $B$ nodes, while $d_A$ is the number of neighbors of $A$ members who belong to $B$.

**Corollary 2.** *Take a network in which $\delta_B(d_i^B = d_B) = 1$ and $\delta_A(d_i^B = d_A) = 1$ with $d_B > d_A$.*

*(i) A sender designing a multiple-message or a network-specific strategy would set $\rho = 0$.*

*(ii) The network-specific strategy outperforms the multiple-message strategy, which outperforms the unconnected hard news strategy: $V_{NS} > V_{MM} > V_U$*

*(iii) $V_{NS}$, $V_{MM}$ and $V_{NS} - V_{MM}$ are increasing in $d_B$ but decreasing in $d_A$.*

*(iv) $V_{MM}$, $V_U$ and $V_U - V_{MM}$ are increasing in $\alpha_B$.*

*Proof.* See Appendix A □

Following Theorem 2, it is optimal for the sender to design the information structure such that $\rho = 0$. Indeed, in such a regular network there is no variance of degree: all nodes in $B$ have the same degree, so they are all targeted and hence all delivered optimal informativeness. Likewise, all nodes in $A$ have a lower degree than the targeted ones, so they all observe effectively less informative messages than $B$ members. This implies that it is optimal for the designer to exploit degree differences by setting $\rho = 0$; and that such strategy exploiting degree differences are performing better than the ones ignoring it.

As explained above, the NS strategy is better at capitalizing on degree differences than MM. The effective informativeness of messages issued from a strategy designed for a given $B$-degree varies more for different $d_i$ with NS than with MM. Furthermore, because all nodes are susceptible, there is no loss from exploiting the network. As a result, any strategy relying on the connectivity differences does better than the standard unconnected strategy.

The third result of the proposition reinforces the intuitions previously underlined. The strategies relying on the difference of $B$-degree between groups perform better, the larger are these differences. Furthermore, as the NS strategy is better than the MM strategy at exploiting such differences, the former gets increasingly better compared to the latter as the difference increases. This echoes the condition $\mathbb{E}_{\tilde{\delta}}(d_i) \leq \hat{d}$ from Theorem 2. The following equivalent for a general network is expected to appear in future version of this work:

**Conjecture 1.** *$V_{NS}$, $V_{MM}$ and $V_{NS} - V_{MM}$ are increasing in $\mathbb{E}_{\delta_B}(d_i)$ but decreasing in $\mathbb{E}_{\delta_A}(d_i)$.*

This result also shows how connectivity *per se* is not detrimental to the sender. Actually, it can benefit her, if this connectivity allows a group to be more interconnected. The problem is the number of connections between the groups. One can informally link this to homophily. While it could be that group $A$ is simply less connected overall, and therefore not necessarily display homophily, it is interesting to note that, were all nodes having the same overall degree, the pattern $d_A < d_B$ would indicate homophily. In this context, homophily would serve the sender, as it would increase the difference between $d_A$ and $d_B$.

Finally, the last result underlines how the magnitude of skepticism from the members of $B$ influences the persuasion value, but also which type of strategy to use. Unsurprisingly, as very skeptical receivers are harder to convince, the sender needs to make the signal structure more informative, lower the probability for any node to take the favorable action. This applies given any strategy. However, when members of $B$ are very hard to convince, exploiting the network becomes more beneficial. When $\alpha_B$ is low, the differences in connection patterns between groups matter more, as decreasing the effective informativeness of messages observed by nodes in $A$ is more important. I expect this intuition to carry through in general networks:

**Conjecture 2.** *For $\mathbb{E}_{\tilde{\delta}}(d_i) < \hat{d}$, $V_{NS} > V_{MM} > V_U$ for relatively small $\alpha_B$.*

This conjecture echoes Corollary 1. It posits that since the benefits from exploiting the networks in the unfavorable state are higher for more skeptical members of $B$, these benefits are more likely to compensate the loss from never persuading non-susceptible nodes.

#### 4.3.2 High, medium and low connectivity within B

So far, only the persuasion value of the different strategies exploiting the network have been analyzed for any arbitrary $\hat{d}$. To gain insights into this problem, I consider the following simple network: all nodes in $A$ have a medium number of connection to $B$, denoted $d_M$; while nodes in $B$ can have a low, a medium or high number of connections, denoted $d_L$, $d_M$ and $d_H$ respectively. Let $d_L < d_M < d_H$.

**Lemma 5.** *Take a network in which $\delta_B(d_i^B = d_L) + \delta_B(d_i^B = d_M) + \delta_B(d_i^B = d_H) = 1$ and $\delta_A(d_i^B = d_M) = 1$ with $d_L < d_M < d_H$. The NS and MM strategies targeting $\hat{d} \in \{d_L, d_M\}$ underperforms the U strategy.*

*Proof.* See Appendix A □

Targeting nodes with higher degree have two effects: increase the number of non-susceptible nodes in $B$; and decrease the informativeness of messages to $A$. The former reduces the persuasion value, the latter raises it. For the persuasion value of MM or NS strategies surpass that of the unconnected strategy, the informativeness of messages to $A$ has to be decreased below its value for targeted node; hence the targeted nodes have to have a higher degree than nodes in $A$.

Intuitively, targeting $d_L$ would underperform with respect to the unconnected strategy because in the unfavorable state of the world, the effective informativeness of messages would be weakly higher than the one needed to persuade nodes in $B$ *for everyone*. Targeting $d_M$ would underperform with respect to the unconnected strategy because in the favorable state of the world, all nodes in $B$ with degree $d_L$ would be lost. This cannot be compensated by the outcome when the unfavorable state of the world realizes, as the effective informativeness of messages for $A$ would still be the same as the one in the unconnected strategy.

While exploiting the network has non-linear effects on the persuasion value, one would expect that this reasoning still applies when nodes in $A$ have heterogenous degrees. This leads to the following conjecture.

**Conjecture 3.** *If the targeted nodes have a $B$-degree $\hat{d} < \mathbb{E}_{\delta_A}(d_i)$, then $V_{NS} < V_U$ and $V_{MM} < V_U$*

Further research is needed to understand the conditions characterizing the optimal targeted nodes.

### 4.4 Further Research

The first element that needs to be addressed is which degree in $B$ to target. Using Proposition 1, one can study the different persuasion value under the same strategy with different targeted nodes. This should allow to derive some necessary conditions for a certain degree to be optimal to target.

Similarly, the conditions provided in Theorem 2 are sufficient but not necessary for the ranking of the considered strategies. It will be important to determine necessary conditions for strategies exploiting the network to be beneficial to the sender.

A second step in this work will be to consider soft news strategies exploiting the network. Such a strategy would be defined as a signal structure such that for any targeted node $i$ in $A$, $\Pr(m_i \in \mathcal{PM}_i) = 1$; in other words, the sender choses some nodes in $A$ and makes sure that they can never be dissuaded. This requires for the favorable state not to be signaled for sure to members of $B$. The proposed strategies would then be adapted. In particular, as before, any targeted nodes in $B$ should be made indifferent, i.e. $\Pr(m_i \in \mathcal{PM}_i|\omega = 0, i \in B) = \alpha_B \Pr(m_i \in \mathcal{PM}_i|\omega = 1, i \in B)$. But now, targeted nodes in $A$ should not detect $\omega = 0$, which requires $\Pr(m_i \in \mathcal{PM}_i|\omega = 1, i \in B) \neq 1$. Actually, for the targeted nodes in $A$ to be indifferent upon seeing any message, we need $1 - \Pr(m_i \in \mathcal{PM}_i|\omega = o, i \in B) = \alpha_A\Big(1 - \Pr(m_i \in \mathcal{PM}_i|\omega = 1, i \in B)\Big)$. This forms a system that implicitly determines $q$, $\varphi$, $p$ and $\rho$ for the MM and NS strategies.

Proposition 1 would then have a soft news strategy equivalent, in which the persuasion value of such strategies would be characterized for any degree of targeted nodes in $A$ and in $B$. The persuasion value would thus be determined by two degree thresholds. These strategies could then be compared between themselves, but also with equivalent hard news strategies. The role of polarization and its relation to the network could thus be underlined.

## 5 Conclusion

This paper explores the role of a network in shaping the provision of information in a persuasion problem. How and when should the sender exploit a network in order to induce communicating agents with heterogenous priors to act a certain way?

I find that there exists strategies which exploit the network in order to send messages with different informativeness to different nodes. In particular, the information necessary to persuade one type of receiver can be split among many of them. This applies particularly well when the network is rather segregated, that is, when nodes with different priors are connected less than nodes with the same prior. It allows receivers who need to be persuaded to observe an informative message by combining each bit of information that has been disseminated among themselves; while preventing the other receivers from accessing too informative messages, as they cannot observe many bits of information, each of which are barely informative.

I present two ways in which the sender can implement such scheme. She can make each bit of information informative by itself. Then, each signal is informative, but poorly so; it thus takes many positive signals to persuade the skeptical receivers. Alternatively, the sender can code information in the *similarity* of signal realizations. In such a case, the signal is not informative by itself; but, in association with other messages, it is. Again, the more signals one observes, the more informative of a message they deliver when combined. However, the informativeness varies relatively more with degree when the sender utilizes the similarity of signals rather than their individual realization.

Which of these strategies is more valuable to the sender? Well, it depends on the context. If the average degree of nodes who need to be persuaded is different enough from those who should not be dissuaded, the sender would prefer to exploit the network through correlated signals than through low informativeness of individual realization. This is due to the greater capacity for the strategy relying on correlation to utilize degree differences. Whether the sender wants to exploit the network in the first place also depends on the relative degree averages. If the receivers who should not be dissuaded are connected poorly enough, then exploiting the network is better when the adverse state realizes, as the sender would, on average, persuade more nodes. However, when the favorable state realizes, the standard strategy that does not exploit the network persuades all receivers; whereas some nodes who need to be persuaded and are too poorly connected are left out if the network is exploited. Therefore, when the favorable state is unlikely to begin with, or when only very few nodes are impossible to persuade, exploiting the network is better.

These results are put in perspective in a regular network. I consider the case in which all nodes who should not be dissuaded have the same number of connections to nodes who should be persuaded, who are themselves equally connected to each other. In this case, the strategies exploiting the network are unambiguously better; and the one who uses correlation even more so.

In such a context, the sender benefits from greater cohesion within the group who she needs to persuade. This shows how connectivity is not necessarily detrimental to the sender. Even though agents can communicate, the sender can easily adapt the informativeness of each signal so that the messages observed are not too informative. This scheme is of course more efficient if each group's connectivity is rather homogenous. However, higher connectivity between groups *is* detrimental to the sender. This puts into perspective the role of homophily in the network. The sender is not necessarily hurt by connectivity; but by cross-group connectivity, she is. The reverse applies to the amount of information provided to the receivers: it does not necessarily decrease with connectivity, but it surely does if the network is more segregated. It underlines how the network structure, and in particular homophily, is detrimental to the provision of information.

Using the same regular network, I show that strategies exploiting the network benefit the sender even more when agents are harder to persuade. Indeed, relying on the network to decrease the effective informativeness of messages observed by agents who should not be dissuaded is all the more valuable when a lot of information is delivered to other agents. Therefore, one would expect that the sender would always prefer to communicate about polarizing topics on the network.

To better asses the role of polarization on the value of strategies exploiting the network, other strategies should be considered. The strategies presented above are the unique strategies exploiting the network and ensuring that the favorable state is signaled when it occurs. However, the sender might prefer to never dissuade some nodes. The previous strategies, adapted for this objective, would allow to explicitly contrast beliefs of each group.

Yet, the results presented already offer many interesting insights. The way existing information spreads in a network depends on its structure; this has been largely documented. However, the current work offers insights into the role of the network structure on the provision of information in the first place. It addresses topical concerns about polarization and segregation in social media, and shows how these weaknesses can be exploited even when agents are fully rational.

## References


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

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

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

Matteo Castiglioni, Alberto Marchesi, Andrea Celli, and Nicola Gatti. Multi-receiver online bayesian persuasion. *Working Paper*, 2021.

Georgy Egorov and Konstantin Sonin. Persuasion on networks. Technical report, National Bureau of Economic Research, 2020.

Simone Galperti and Jacopo Perego. Belief meddling in social networks: An information-design approach. *Working Paper*, 2019.

Yingni Guo and Eran Shmaya. The interval structure of optimal disclosure. *Econometrica*, 87 (2):653–675, 2019.

Federico Innocenti. Can media pluralism be harmful to news quality? Technical report, University of Mannheim, 2021.

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

Toygar Kerman and Anastas P Tenev. Persuading communicating voters. *Working Paper*, 2021.

Toygar Kerman, P Herings, and Dominik Karos. Persuading strategic voters. *Working Paper*, 2020.

Jerome Klotz. Statistical inference in bernoulli trials with dependence. *The Annals of statistics*, pages 373–379, 1973.

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

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

Yun Wang. Bayesian persuasion with multiple receivers. *Working Paper*, 2013.

# A Computations and Proofs

***Distribution of $n$ dependent Bernoulli trials***

Recall that the probability of $k$ successes among $n$ dependent trials is denoted $\tilde{p}_k^n$. Furthermore, denote any individual trial $x_i$.

The probability for any single trial $x_i$ to be a success is $p$, hence $p = \mathbb{E}(x_i) = \sum_{k=0}^{n-1} \tilde{p}_{k+1}^n$. Likewise, conditioning on $x_i = 0$, we find $1 - p = \sum_{k=0}^{n-1} \tilde{p}_k^n$

By the definition, $\text{Corr}(x_i, x_j)[\mathbb{E}(x_i^2) - \mathbb{E}(x_i)^2] + \mathbb{E}(x_i)\mathbb{E}(x_j) = \mathbb{E}(x_i x_j)$. Therefore $\rho[p - p^2] + p^2 = \sum_{k=0}^{n-2} \tilde{p}_{k+2}^n$ as the expectation of $\mathbb{E}(x_i x_j)$ is the probability for both $x_i, x_j$ to be successes, regardless of what happens with all other trials.

Therefore, $\tilde{p}_k^n$ must satisfy: (i) $p = \sum_{k=0}^{n-1} \tilde{p}_{k+1}^n$; (ii) $1-p = \sum_{k=0}^{n-1} \tilde{p}_k^n$; (iii) $\rho(1-p)p + p^2 = \sum_{k=0}^{n-2} \tilde{p}_{k+2}^n$.

In particular, for $\rho = 0$, $\tilde{p}_k^n$ is the standard Binomial PMF, while for $\rho = 1$, $\tilde{p}_k^n = \begin{cases} 1-p & \text{if } k = 0 \\ p & \text{if } k = n \\ 0 & \text{otherwise} \end{cases}$

It is easy to verify that the PML proposed fulfills (i), (ii) and (iii). Among the PMLs of the form: $(1-\rho)^x \binom{n}{k} p^k (1-p)^{n-k} + \rho^y \big[(1-p)\mathbb{1}_{k=0} + p\mathbb{1}_{k=n}\big]$, only $x = y = 1$ fulfills all three conditions. If $x \neq 1$ or $y \neq 1$, then $\sum_{k=0}^{n} \tilde{p}_k^n \neq 1$.

***Proof of Theorem 1***

Let us first derive $\Pr(m_i = x|\omega = 0)$ and $\Pr(m_i = x|\omega = 1)$.

For any given node $i$ with $B$-degree $d_i$, all realizations of $m_B$ with at least $x$ successes could allow for $m_i$ to equate $x$. For any number of successes $s \geq x$ within $B$-nodes, there are $\binom{B-d_i}{s-x}\binom{d_i}{x}$ out of the $\binom{B}{s}$ possible realizations of signals that allow for $m_i$ to be $x$. Indeed, there are $\binom{d_i}{x}$ ways to arrange the $x$ successes within $i$'s neighborhood; and $\binom{B-d_i}{s-x}$ ways to arrange the other $(s-x)$ successes outside of $i$'s neighborhood. Therefore, the probability for this given node $i$ to observe $x$ successes in his $B$-neighborhood when $s$ successes occurred among the $B$ nodes is $\frac{\binom{B-d_i}{s-x}\binom{d_i}{x}}{\binom{B}{s}} \tilde{p}_s^B$. Fixing $x$ successes within $i$'s neighborhood, anything can happen in the rest of the network, i.e. from $s - x = 0$ to $B - d_i$ other successes. Hence:

$$\Pr(m_i = x|\omega = 0) = \sum_{s-x=0}^{B-d_i} \frac{\binom{B-d_i}{s-x}\binom{d_i}{x}}{\binom{B}{s}} \tilde{p}_s^B = (1-\rho) \sum_{s-x=0}^{B-d_i} \binom{B-d_i}{s-x}\binom{d_i}{x} p^s (1-p)^{B-s} + \rho\big[(1-p)\mathbb{1}_{x=0} + p\mathbb{1}_{x=n}\big]$$

$$= (1-\rho)\binom{d_i}{x} \underbrace{\sum_{s-x=0}^{B-d_i} \binom{B-d_i}{s-x} p^{s-x}(1-p)^{B-d_i-(s-x)}}_{=1} p^x (1-p)^{d_i-x} + \rho\big[(1-p)\mathbb{1}_{x=0} + p\mathbb{1}_{x=n}\big]$$

Likewise, $\Pr(m_i = x|\omega = 1) = (1-\varphi)\binom{d_i}{x} q^x (1-q)^{d_i-x} + \varphi\big[(1-q)\mathbb{1}_{x=0} + q\mathbb{1}_{x=n}\big]$.

By definition of a hard news strategy, we need $\Pr(m_i \in \mathcal{PM}_i|\omega = 1) = 1$. Therefore, for any hard news strategy, it must be that $\sum_{m_i \in \mathcal{PM}_i} \Pr(m_i = x|\omega = 1) = 1$. We show that the only sets of messages that fulfill this condition are associated to one of the three strategies U, MM or NS. Assume by contradiction that $0 < m_i < d_i$ is a persuading message for the targeted nodes. Then $q \neq 1 \neq \varphi$, hence:

$$\sum_{x \in X} (1-\varphi)\binom{d_i}{x} q^x (1-q)^{d_i - x} + \varphi\left[q \mathbb{1}_{x=d_i} + (1-q)\mathbb{1}_{x=0}\right] = 1 \Leftrightarrow X = \{0, 1..., d_i\}$$

This would thus require $\mathcal{PM}_i = \mathcal{M}_i$, which is impossible, as messages have to be informative in order to be persuading.

***Proof of Lemma 3***

*Proof.* Recall that $\Pr(m_i = x|\omega = 1) = \varphi\binom{d_i}{x} q^x (1-q)^{d_i - x} + (1-\varphi)\left[q \mathbb{1}_{x=d_i} + (1-q)\mathbb{1}_{x=0}\right]$.

(i) By definition, $\mathcal{FM}_i \neq \varnothing$ requires that there exists a $m_i$ such that $\beta(m_i) \geq t$. First, because MM and NS are hard news strategies, upon observing any signal realization that is not a success for MM, or any signal realization different to other signals for NS, nodes perfectly detect that the unfavorable state realizes. Hence, $\mathcal{FM}_i \in \{\varnothing, \{d_i\}\}$ for MM and $\mathcal{FM}_i \in \{\varnothing, \{0, d_i\}\}$ for NS. Now, $\forall i \in A$, $\beta(m_i = d_i) \geq t$. Indeed, because they induce informative signals, both strategies insure that $\Pr(m_i = d_i|\omega = 0) \leq \Pr(m_i = d_i|\omega = 1)$; hence, $\forall i \in A, \Pr(m_i = d_i|\omega = 0) \leq \alpha_A \Pr(m_i = d_i|\omega = 1)$. Furthermore, $\forall i \in B : d_i \geq \hat{d}$, $\Pr(m_i = d_i|\omega = 0) \leq \Pr(m_i = \hat{d}|\omega = 0) = \alpha_B \Pr(m_i = \hat{d}|\omega = 1) = \alpha_B \Pr(m_i = d_i|\omega = 1)$ where the first inequality follows from the expression for $\Pr(m_i = d_i|\omega = 0)$ being decreasing in $d_i$; and the last equality follows from the definition of hard news strategy. Finally $\forall i \in B : d_i < \hat{d}$, $\beta(m_i = 0) < t$ and $\beta(m_i = d_i) < t$. Indeed, $d_i < \hat{d}$ implies $\Pr(m_i = d_i|\omega = 0) > \Pr(m_i = \hat{d}|\omega = 0)$. Therefore, $\Pr(m_i = d_i|\omega = 0) > \alpha_B \Pr(m_i = d_i|\omega = 1)$. Likewise, $\forall i \in B : d_i < \hat{d}$, $\Pr(m_i = 0|\omega = 0) > \alpha_B \Pr(m_i = 0|\omega = 1)$.

(ii) For U and MM, $q = 1$ insures $\forall i \in B s_i = 1$, so that $\forall i \in N, m_i = d_i$; for NS, $\varphi = 1$ implies $\forall i, j \in B, s_i = s_j$, so that $\forall i \in N, m_i \{0, d_i\}$.

(iii) $\Pr(m_i = x|\omega = 0)$ is decreasing in $d_i$.

□

***Proof of Proposition 1***

*Proof.* By definition, the beliefs of targeted nodes upon observing a persuading signal must equate $t$. Because $\beta(m) = \frac{\Pr(m|\omega=1)}{\Pr(m|\omega=0)+\Pr(m|\omega=1)}$, the sender needs to set $\Pr(m|\omega = 0) = \alpha_B \Pr(m|\omega = 1)$. Because we consider hard news strategies, $\Pr(m|\omega = 1) = 1$.

(i) For the strategy to be independent of the targeted nodes' degree, the posterior has to be independent of the number of signals a node observes. Therefore, it requires $\varphi = \rho = 1$.

Because it is a hard news strategy, $q = 1$. Because nodes in $B$ are targeted, $p = \alpha_B q = \alpha_B$. Following Lemma 2, all nodes in $A$ whose $B$-degree is zero, observe no informative signals and thus always take the favorable action. Therefore, if the unfavorable state realizes, the probability of a node taking the favorable action is the probability for the favorable message to realize, plus the probability that the favorable message not to realize times the probability of the node being member of $A$ without connection to $B$. Hence, the persuasion value of the unconnected strategy is $V_U = \mu + (1-\mu)\left[\alpha_B + a\ \delta_A(0)(1-\alpha_B)\right]$.

(ii) Under a $MM$ strategy, for the targeted nodes $i$, $\mathcal{PM}_i = \{d_i\}$. From Lemma 3's proof, $\Pr(m_i = d_i|\omega = 1) = (1-\rho)p^{d_i} + \rho\ p$ and $\Pr(m_i = d_i|\omega = 0)(1-\rho)q^{d_i} + \rho\ q$. Because $B$-nodes with $B$-degree $\hat{d}$ are targeted, we have $\Pr(m_i = \hat{d}|\omega = 1) = \alpha_B \Pr(m_i = \hat{d}|\omega = 0)$. Furthermore, $\Pr(m = \hat{d}|\omega = 0) = \varphi q + (1-\varphi)q^{\hat{d}} \overset{!}{=} 1$ requires $q = 1$. Therefore, the sender sets $q = 1$ and $\rho p + (1-\rho)p^{\hat{d}} = \alpha_B$.

Now, by Lemma 3, $m_i = d_i$ induces node $i$ to take the favorable action iff $i \in A$ or ($i \in B$ and $d_i \geq \hat{d}$). Therefore, the probability that any random node takes the favorable action, i.e. the value of persuasion, under this strategy is:

$$\sum_{\omega} \Pr(\omega)\left[\Pr(i \in A)\sum_{d=0}^{B}\Pr(d_i = d)\Pr(m_i = d_i|\omega) + \Pr(i \in B)\sum_{d=\hat{d}}^{B}\Pr(d_i = d)\Pr(m_i = d_i|\omega)\right]$$

Which translates into:

$$\mu\left[a + b\sum_{d_i=\hat{d}}^{B}\delta_B(d_i)\right] + (1-\mu)\left[a\delta_A(0) + \sum_{d_i=1}^{B}\left[a\delta_A(d_i) + \mathbb{1}_{d_i\geq\hat{d}}\ b\delta_B(d_i)\right]\left[\rho\ p + (1-\rho)p^{d_i}\right]\right]$$

(iii) Under a $NS$ strategy, for the targeted nodes $i$, $\mathcal{PM}_i = \{0, d_i\}$. Therefore, $\Pr(m_i \in \mathcal{PM}_i|\omega = 0) = \rho + (1-\rho)\left[p^{d_i} + (1-p)^{d_i}\right]$ and $\Pr(m_i \in \mathcal{PM}_i|\omega = 1) = \varphi + (1-\varphi)\left[q_i^d + (1-q)_i^d\right]$.
Because $B$-nodes with $B$-degree are targeted, we have $\Pr(m_i \in \{0, d_i\}|\omega = 1) = \alpha_B \Pr(m_i \in \{0, d_i\}|\omega = 0)$. Furthermore, $\Pr(m \in \{0, \hat{d}\}|\omega = 0) = \varphi + (1-\varphi)\left[q^{\hat{d}} + (1-q)^{\hat{d}}\right] \overset{!}{=} 1$ requires $\varphi = 1$. Therefore, the sender sets $\varphi = 1$ and $\rho + (1-\rho)\left[p^{\hat{d}} + (1-p)^{\hat{d}}\right] = \alpha_B$. Finally, for $\Pr(m_i = \hat{d}|\omega = 1) = \alpha_B \Pr(m_i = \hat{d}|\omega = 0)$, we need $q = \alpha_B^{-1}\left(\rho p + (1-\rho)p^{\hat{d}}\right)$ Similarly to above, $m_i \in \{0, d_i\}$ induces node $i$ to take the favorable action iff $i \in A$ or ($i \in B$ and $d_i \geq \hat{d}$). Therefore, the probability that any random node takes the favorable action, i.e. the value of persuasion, under this strategy is:

$$\sum_{\omega} \Pr(\omega)\left[\Pr(i \in A)\sum_{d=0}^{B}\Pr(d_i = d)\Pr(m_i \in \{0, d_i\}|\omega) + \Pr(i \in B)\sum_{d=\hat{d}}^{B}\Pr(d_i = d)\Pr(m_i \in \{0, d_i\}|\omega)\right]$$

Which translates into:

$$\mu\left[a + b\sum_{d_i=\hat{d}}^{B}\delta_B(d_i)\right] + (1-\mu)\left[a\delta_A(0) + \sum_{d_i=1}^{B}\left[a\delta_A(d_i) + \mathbb{1}_{(d_i\geq\hat{d})}\ b\delta_B(d_i)\right]\left[\rho + (1-\rho)\left(p^{d_i} + (1-p)^{d_i}\right)\right]\right]$$

$\square$

***Proof of Lemma 4***

*Proof.* • Consider the MM strategy. From Proposition **??** the persuasion value of this strategy without difference in group priors is:

$$V_{MM} = \mu \sum_{d_i=\hat{d}}^{N} \delta(d_i) + (1-\mu) \sum_{d_i=\hat{d}}^{N} \delta(d_i)\big[\rho\ p + (1-\rho)p^{d_i}\big]$$

This is to be compared with the value of persuasion with a standard strategy which is:

$$V_{std} = \mu + (1-\mu)\alpha$$

Because $\hat{d} \geq \text{argmin}_i d_i$, $\sum_{d_i=\hat{d}}^{N} \delta(d_i) \leq 1$. Furthermore, $\sum_{d_i=\hat{d}}^{N} \delta(d_i)\big[\rho\ p + (1-\rho)p^{d_i}\big] \leq \alpha$. Indeed, by definition, $\rho\ p + (1-\rho)p^{\hat{d}} = \alpha$. Therefore, $\forall d_i > \hat{d}, \rho\ p + (1-\rho)p^{\hat{d}} < \alpha$. We conclude $V_{MM} \leq V_{std}$. The equation holds with equality iff $\forall i \in N, d_i = \hat{d}$.

- Consider the NS strategy. From Proposition **??** the persuasion value of this strategy without difference in group priors is:

$$\mu \sum_{d_i=\hat{d}}^{N} \delta_B(d_i) + (1-\mu) \sum_{d_i=\hat{d}}^{N} \delta_B(d_i)\Big[\rho + (1-\rho)\big(p^{d_i} + (1-p)^{d_i}\big)\Big]$$

This is to be compared with the value of persuasion with a standard strategy which is:

$$V_{std} = \mu + (1-\mu)\alpha$$

Because $\hat{d} \geq \text{argmin}_i d_i$, $\sum_{d_i=\hat{d}}^{N} \delta(d_i) \leq 1$. Furthermore, $\sum_{d_i=\hat{d}}^{N} \delta(d_i)\big[\rho + (1-\rho)(1-\rho)\big(p^{d_i} + (1-p)^{d_i}\big)\big] \leq \alpha$. Indeed, by definition, $\rho + (1-\rho)(1-\rho)\big(p^{\hat{d}} + (1-p)^{\hat{d}}\big) = \alpha$. Therefore, $\forall d_i > \hat{d}, \big(p^{d_i} + (1-p)^{d_i}\big) < \alpha$. We conclude $V_{NS} \leq V_{std}$. The equation holds with equality iff $\forall i \in N, d_i = \hat{d}$.

□

***Proof of Theorem 2***

(i) I consider each strategy separately. I first show that the two conditions imply $\rho_{MM} = 0$ and $\rho_{NS}$ respectively. I then proceed to show that the first condition is sufficient for $V_{MM} < V_{NS}$.

   – For the multiple-message strategy, the sender sets $(p, \rho)$ subject to the constraint $\rho\ p + (1-\rho)p^{\hat{d}} = \alpha_B$ to maximize:

$$(1-\mu) \sum_{d_i=1}^{B} \tilde{\delta}(d_i)\Big[\rho\ p + (1-\rho)p^{d_i}\Big]$$

One can rewrite this objective as:

$$\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[\rho\ p+(1-\rho)p^{d_i}\Big]=\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[\alpha_B+(1-\rho)(p^{d_i}-p^{\hat{d}})\Big]$$
$$=\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\left[\alpha_B+\left(1-\frac{\alpha_B-p^{\hat{d}}}{p-p^{\hat{d}}}\right)(p^{d_i}-p^{\hat{d}})\right]$$
$$=\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\left[\alpha_B+\left((p-\alpha_B)\frac{p^{d_i}-p^{\hat{d}}}{p-p^{\hat{d}}}\right)\right]$$

Which is increasing in $p$ if $\mathbb{E}_{\tilde{\delta}}(d_i)<\hat{d}$. Indeed, both $p-\alpha_B$ and $\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\frac{p^{d_i}-p^{\hat{d}}}{p-p^{\hat{d}}}$ are increasing in $p$. While it is trivial to show this for $p-\alpha_B$, it is less straightforward for $\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\frac{p^{d_i}-p^{\hat{d}}}{p-p^{\hat{d}}}$. We have:

$$\frac{\partial\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\frac{p^{d_i-1}-p^{\hat{d}-1}}{1-p^{\hat{d}-1}}}{\partial p}$$
$$=\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\left[\frac{\big[(d_i-1)p^{d_i-2}-(\hat{d}-1)p^{\hat{d}-2}\big](1-p^{\hat{d}-1})+(\hat{d}-1)p^{\hat{d}-2}(p^{\hat{d}-1}-p^{\hat{d}-1})}{(1-p^{\hat{d}-1})^2}\right]$$
$$=\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\left[\frac{(d_i-1)p^{d_i-2}-(\hat{d}-1)p^{\hat{d}-2}+(\hat{d}-d_i)p^{\hat{d}+d_i-3}}{(1-p^{\hat{d}-1})^2}\right]$$

Whose sign is proportional to numerator, which itself is proportional to

$$\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[(d_i-1)-(\hat{d}-1)p^{\hat{d}-d_i}+(\hat{d}-d_i)p^{\hat{d}-1}\Big]$$

This expression is decreasing in $p$:

$$\frac{\partial\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[(d_i-1)-(\hat{d}-1)p^{\hat{d}-d_i}+(\hat{d}-d_i)p^{\hat{d}-1}\Big]}{\partial p}$$
$$=\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[-(\hat{d}-1)(\hat{d}-d_i)p^{\hat{d}-d_i}+(\hat{d}-d_i)(\hat{d}-1)p^{\hat{d}-1}\Big]\geq 0$$

since $0<\hat{d}-\mathbb{E}_{\tilde{\delta}}(d_i)<\sum_{d_i=1}^{B}\tilde{\delta}(d_i)(\hat{d}-d_i)p^{-d_i+1}$. The inequality holds strictly for $p<1$. Therefore,

$$\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[(d_i-1)-(\hat{d}-1)p^{\hat{d}-d_i}+(\hat{d}-d_i)p^{\hat{d}-1}\Big]$$
$$\geq\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[(d_i-1)-(\hat{d}-1)p^{\hat{d}-d_i}+(\hat{d}-d_i)p^{\hat{d}-1}\Big]\Bigg|_{p=1}=0$$

Again, the inequality holds strictly for $p<1$. We conclude that to maximize $\sum_{d_i=1}^{B}\tilde{\delta}(d_i)\Big[\rho\ p+(1-\rho)p^{d_i}\Big]$, the sender sets $p$ as high as possible, which requires $\rho_{MM}=0$.

– For the network-specific strategy, the sender sets $(p,\rho)$ subject to the constraint $\rho + (1-\rho)\big(p^{\hat{d}} + (1-p)^{\hat{d}}\big) = \alpha_B$ to maximize:

$$(1-\mu)\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\Big[\rho + (1-\rho)\big(p^{d_i} + (1-p)^{d_i}\big)\Big]$$

One can rewrite this objective as:

$$\begin{aligned}
&\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\Big[\rho + (1-\rho)\big(p^{d_i} + (1-p)^{d_i}\big)\Big] \\
&\qquad = \textstyle\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\left[\frac{\alpha - p^{d_B} - (1-p)^{\hat{d}}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}} + \left(1 - \frac{\alpha - p^{\hat{d}} - (1-p)^{\hat{d}}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}\right)\big(p^{d_i} + (1-p)^{d_i}\big)\right] \\
&\qquad = \textstyle\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\left[\frac{\alpha_B - p^{\hat{d}} - (1-p)^{\hat{d}} - (1-\alpha_B)\left(p^{d_i}+(1-p)^{d_i}\right)}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}\right] \\
&\qquad = \textstyle\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\left[1 - \frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}\right]
\end{aligned}$$

Which is increasing in $p$ for $p \geq 1/2$. Indeed, $\sum_{d_i=2}^{B} \tilde{\delta}(d_i)\frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}$ is decreasing in $p \geq 1/2$.[12]

$$\frac{\partial \sum_{d_i=2}^{B} \tilde{\delta}(d_i)\frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}}{\partial p} = \sum_{d_i=2}^{B} \tilde{\delta}(d_i)\frac{-d_i\left[p^{d_i-1}-(1-p)^{d_i-1}\right](1-p^{\hat{d}}-(1-p)^{\hat{d}}) + \hat{d}\left[p^{\hat{d}-1}-(1-p)^{\hat{d}-1}\right](1-p^{d_i}-(1-p)^{d_i})}{(1-p^{\hat{d}}-(1-p)^{\hat{d}})^2}$$

Is equal to 0 in $p = 1/2$, but it is null for no other $p \in [0,1]$. Hence, the function is monotone between $[0,1/2)$ and $(1/2,1]$, and $p = 1/2$ is a stationary point. Using Hospital rule twice, we have:

$$\left.\frac{\partial \sum_{d_i=2}^{B} \tilde{\delta}(d_i)\frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}}{\partial p}\right|_{p=0} = \frac{-\sum_{d_i=2}^{B} \tilde{\delta}(d_i)d_i(d_i-1)\hat{d} + \hat{d}(\hat{d}-1)\sum_{d_i=1}^{B} \tilde{\delta}(d_i)d_i}{2\hat{d}^2} \geq 0$$

for $\mathbb{E}_{\tilde{\delta}}(d_i{}^2) \leq \hat{d} \cdot \mathbb{E}_{\tilde{\delta}}(d_i)$; and likewise:

$$\left.\frac{\partial \sum_{d_i=2}^{B} \tilde{\delta}(d_i)\frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}}{\partial p}\right|_{p=1} = \frac{\sum_{d_i=2}^{B} \tilde{\delta}(d_i)d_i(d_i-1)\hat{d} - \hat{d}(\hat{d}-1)\sum_{d_i=1}^{B} \tilde{\delta}(d_i)d_i}{2\hat{d}^2} \leq 0$$

Therefore, for $\mathbb{E}_{\tilde{\delta}}(d_i{}^2) \leq \hat{d} \cdot \mathbb{E}_{\tilde{\delta}}(d_i)$, $p = 1/2$ is a maximum. We conclude that to maximize $\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{\hat{d}}-(1-p)^{\hat{d}}}$ the sender sets $p$ as high as possible, which requires $\rho_{NS} = 0$.

Furthermore, consider again $\hat{d} \leq \mathbb{E}_{\tilde{\delta}}(d_i)$, so that $\rho_{MM} = 0$. I will show that $V_{MM} < V_{NS}|_{\rho_{NS}=0} \leq V_{NS}$ where the last inequality follows from the sender's optimization. Hence, assume $\rho_{NS} = 0$. The same reasoning applies for picking $\hat{d}$, so I can compare

[12] We can sum from $d_i = 2$ as in $d_i$, the expression is 0

them assuming the targeted nodes within each strategy are the same.

As $p_{MM} = \alpha_B^{\frac{1}{\hat{d}}}$ and $\alpha_B = p_{NS}^{\hat{d}} + (1-p_{NS})^{\hat{d}}$, $p_{MM} = \left(p_{NS}^{\hat{d}} + (1-p)_{NS}^{\hat{d}}\right)^{\frac{1}{\hat{d}}}$. To ease notation, we denote $p_{NS}$ with $p$ and $p_{MM}$ with $\tilde{p}$ in this proof. Note that $\tilde{p} \geq p$ as $\tilde{p}$ is decreasing in $\hat{d}$ with $\lim_{d\to\infty} \tilde{p} = p$. We have:

$$V_{NS} - V_{MM} = (1-\mu)\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\left[p^{d_i} + (1-p)^{d_i} - \tilde{p}^{d_i}\right]$$

For the remainder of this proof, I denote $f(d) = \left(p^{\hat{d}} + (1-p)^{\hat{d}}\right)^{\frac{1}{\hat{d}}} - p^{d_i} - (1-p)^{d_i}$. I derive the shape of this function. In particular, it is increasing, then decreasing. I show that as long as it is increasing, it is concave; and that if it is decreasing for some $d_i$, then it is decreasing for all subsequent $d_i$. I consider the discrete variations in $f(d)$ as $d \in \mathbb{N}$. $\Delta f(d) \coloneqq f(d+1) - f(d) = (1-p)p^d + p(1-p)^d - (1-\tilde{p})\tilde{p}^d$; and $\Delta\Delta f(d) \coloneqq \Delta f(d+1) - \Delta f(d) = (1-p)^2 p^d + p^2(1-p)^d - (1-\tilde{p})^2\tilde{p}^d$. It is easy to verify that for any $d < \hat{d}$, $f(d) < 0$ and $\Delta f(d) > 0$. Indeed $a^x$ is concave for any $x < 1$ so that $(a+b)^x > a^x + b^x$; i.e. $\left(p^{\hat{d}} + (1-p)^{\hat{d}}\right)^{\frac{d}{\hat{d}}} > p^d + (1-p)^d$.

Now, if $f(d)$ is increasing, $(1-p)p^d + p(1-p)^d > (1-\tilde{p})\tilde{p}^d$; which implies $(1-p)^2p^d + p^2(1-p)^d > (1-\tilde{p})(1-p)p^d + (1-\tilde{p})p(1-p)^d > (1-\tilde{p})^2\tilde{p}^d$ where the first inequality follows from $p > 1-p > 1-\tilde{p}$. Therefore, $\Delta f(d) > 0 \Rightarrow \Delta\Delta f(d) < 0$.

Furthermore, as soon as $f(d)$ is decreasing, it stays decreasing. Indeed, $f(d) < 0$ means $(1-p)p^d + p(1-p)^d < (1-\tilde{p})\tilde{p}^d$ so that $(1-p)p^{d+1} + p(1-p)^{d+1} < (1-p)p^d\tilde{p} + p(1-p)^d\tilde{p} < (1-\tilde{p})\tilde{p}^{d+1}$ i.e. $f(d+1) < 0$, using again $\tilde{p} > p > 1-p$.

We want to show that $\sum_{d_i=1}^{B} \tilde{\delta}(d_i)f(d_i) < 0$. To show this, consider:

$\tilde{f}(d) \coloneqq \begin{cases} f(d) & \text{if } \Delta f(d) > 0 \\ \max f(d) & \text{if } \Delta f(d) \leq 0 \end{cases}$. $\tilde{f}(d)$ function is concave; and $f(d) \leq \tilde{f}(d)$. We have:

$$\sum_{d_i=1}^{B} \tilde{\delta}(d_i)f(d_i) \leq \sum_{d_i=1}^{B} \tilde{\delta}(d_i)\tilde{f}(d_i) < \tilde{f}\left(\sum_{d_i=1}^{B} d_i\right) = f\left(\sum_{d_i=1}^{B} d_i\right) < 0$$

where the second inequality follows from $\tilde{f}(d)$ being concave;[13] and the equality because $\mathbb{E}_{\tilde{\delta}}(d_i) < \hat{d}$ such that $f\left(\mathbb{E}_{\tilde{\delta}}(d_i)\right) < 0$.

(ii) For $\rho_{MM} = 0$, $p_{MM} = \alpha_B^{\frac{1}{\hat{d}}}$, so that conditional on $\omega = 0$, the difference of probability for a node to take the favorable action between MM and U is:

$$\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\alpha_B^{\frac{d_i}{\hat{d}}} - \alpha_B > 0$$

as $f(x) = \alpha_B^{\frac{x}{\hat{d}}}$ is a decreasing convex function for $\alpha_B < 1$, so that $\sum_{d_i=1}^{B} \tilde{\delta}(d_i)f(d_i) > f\left(\sum_{d_i=1}^{B} \tilde{\delta}(d_i)d_i\right)$, i.e. $\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\alpha_B^{\frac{d_i}{\hat{d}}} > \alpha_B^{\frac{\mathbb{E}_{\tilde{\delta}}(d_i)}{\hat{d}}} \geq \alpha_B$ for $\mathbb{E}_{\tilde{\delta}}(d_i) \leq \hat{d}$.

[13]It is strict because $\hat{d} \leq \mathbb{E}_{\tilde{\delta}}(d_i)$ implies $\exists i : d_i < \hat{d}$, and that $\tilde{f}(d)$ is strictly concave for $d_i < \hat{d}$.

***Proof of Corollary 1***

When the favorable state of the world realizes, the difference of probability for a node to take the favorable action between MM and U is: $-b\sum_{d_i=1}^{\hat{d}}$. Therefore, $V_{MM} > V_U$ for $\mu b\sum_{d_i=1}^{\hat{d}} \delta_B(d_i) < (1-\mu)\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\alpha_B^{\frac{d_i}{\hat{d}}} - \alpha_B$. Now $\mu b\sum_{d_i=1}^{\hat{d}}$ is increasing in $\mu$, $b$ and $\sum_{d_i=1}^{\hat{d}}$. Likewise $(1-\mu)\sum_{d_i=1}^{B} \tilde{\delta}(d_i)\alpha_B^{\frac{d_i}{\hat{d}}} - \alpha_B$ is decreasing in $\mu$ and $b$ – through $\tilde{\delta}$. Therefore, the inequality is more likely to hold for small $\mu$, $b$ and $\sum_{d_i=1}^{\hat{d}}$.

***Proof of Corollary 2***

*Proof.* (i) Using Theorem 2, as $\mathbb{E}_{\tilde{\delta}}(d_i) = ad_a + bd_B < d_B = \hat{d}$ and $\mathbb{E}_{\tilde{\delta}}(d_i{}^2) = ad_a^2 + bd_B^2 < d_B(ad_A + bd_B) = \hat{d}\cdot\mathbb{E}_{\tilde{\delta}}(d_i)$, we know $\rho = 0$. Let us consider each strategy separately:

(ii) As $\sum_{d_i=1}^{\hat{d}-1} \delta(d_i) = 0$, the value of persuasion of MM and U are the same in the favorable state. Therefore, using Theorem 2, $V_{MM} > V_U$.

(iii) The persuasion value when using a MM strategy is:

$$V_{MM} = \mu + (1-\mu)\Big[a\ \alpha_B{}^{\frac{d_A}{d_B}} + b\ \alpha_B\Big]$$

Because $\alpha_B < 1$, $V_{MM}$ is decreasing in $\frac{d_A}{d_B}$, i.e. increasing in $d_B$ but increasing in $d_A$.
The persuasion value when using a NS strategy is:

$$V_{NS} = \mu + (1-\mu)\Big[a\ \big(p_{NS}{}^{d_A} + (1-p_{NS})^{d_A}\big) + b\ \alpha_B\Big] \text{ s.t } p_{NS}{}^{d_B} + (1-p_{NS})^{d_B} = \alpha_B$$

For any given $p$, $p_{NS}{}^{d_A} + (1-p_{NS})^{d_A}$ is decreasing in $d_A$, making $V_{NS}$ decreasing in $d_A$. For any given $d_A$, $p_{NS}{}^{d_A} + (1-p_{NS})^{d_A}$ is increasing in $p_{NS}$, which is increasing in $d_B$ due to the constraint.
Finally, $V_{NS} - V_{MM} = \big(p_{NS}{}^{d_A} + (1-p_{NS})^{d_A}\big) - \big(p_{NS}{}^{d_B} + (1-p_{NS})^{d_B}\big)^{\frac{d_A}{d_B}}$. Because $f(x) = a^x$ gets less concave as $x \in (0,1]$ approaches 1, $\big(p_{NS}{}^{d_B} + (1-p_{NS})^{d_B}\big)^{\frac{d_A}{d_B}}$ approaches $p_{NS}{}^{d_A} + (1-p_{NS})^{d_A}$ as $\frac{d_A}{d_B}$ approaches 1, hence when $d_A$ increases or $d_B$ decreases.

(iv) The derivative of the MM persuasion value is:

$$\frac{\partial V_{MM}}{\partial \alpha_B} = (1-\mu)\left[a\frac{d_A}{d_B}\alpha_B{}^{-1+\frac{d_a}{d_B}} + b\right] > 0$$

Furthermore,

$$\frac{\partial V_U - V_{MM}}{\partial \alpha_B} = (1-\mu)\left[1 - b - a\frac{d_A}{d_B}\alpha_B{}^{-1+\frac{d_a}{d_B}}\right] > 0$$

since $\frac{d_A}{d_B}\alpha_B{}^{-1+\frac{d_a}{d_B}} < 1$. Indeed, $\ln\left(\frac{d_A}{d_B}\right) + \ln(\alpha_B)\left(-1+\frac{d_a}{d_B}\right) < \left(\frac{d_A}{d_B} - 1\right) + \ln(\alpha_B)\left(-1+\frac{d_A}{d_B}\right) = \left(\frac{d_A}{d_B} - 1\right)(1 - \ln(\alpha_B)) < 0$

□

***Proof Lemma 5***

*Proof.* With $\hat{d} = d_L$, $\frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{d_L}-(1-p)^{d_L}} \leq 1$ for any $d_i$, so the probability for a node to take the favorable action with NS is smaller than $\alpha_B$. Likewise, $\frac{p^{d_i}-p^{\hat{d}}}{p-p^{\hat{d}}} \leq 0$ for any $d_i$, so the probability for a node to take the favorable action with NS is smaller than $\alpha_B$.

With $\hat{d} = d_M$, there is a probability $b\delta_B(d_L)$ for a random node to be non-susceptible.Furthermore $\frac{1-p^{d_i}-(1-p)^{d_i}}{1-p^{d_L}-(1-p)^{d_L}} \leq 1$ for $d_i$ of any susceptible node; hence, $V_{NS} < V_U$. Likewise, $\frac{p^{d_i}-p^{\hat{d}}}{p-p^{\hat{d}}} \leq 0$ for $d_i$ of any susceptible node; hence $V_{MM} < V_U$. $\square$