

# Learning and Manipulation in Social Networks

James Siderius<sup>1</sup>, Mohamed Mostagir<sup>2</sup>, and Asu Ozdaglar<sup>1</sup>

Network Science and Economics Conference

April 28, 2019

---

<sup>1</sup>MIT EECS LIDS

<sup>2</sup>University of Michigan Ross School of Business

# Motivation

"Vaccine hesitancy - the reluctance or refusal to vaccinate despite the availability of vaccines - threatens to reverse progress made in tackling vaccine-preventable diseases. Measles, for example, has seen a 30% increase in cases globally."

*World Health Organization (WHO)*  
February 2019



Health Topics ▾

Countries ▾

News ▾

Emergencies ▾

About Us ▾

## Vaccine hesitancy



Vaccine hesitancy – the reluctance or refusal to vaccinate despite the availability of vaccines – threatens to reverse progress made in tackling vaccine-preventable diseases. Vaccination is one of the most cost-effective ways of avoiding disease – it currently prevents 2-3 million deaths a year, and a further 1.5 million could be avoided if global coverage of vaccinations improved.

Measles, for example, has seen a 30% increase in cases globally. The reasons for this rise are complex, and not all of these cases are due to vaccine hesitancy. However, some countries that were close to eliminating the disease have seen a resurgence.

The reasons why people choose not to vaccinate are complex: a vaccine's safety group is WHO identified complacency, inequities in accessing vaccines, and lack of confidence are key reasons underlying

# Motivation

"Vaccine hesitancy - the reluctance or refusal to vaccinate despite the availability of vaccines - threatens to reverse progress made in tackling vaccine-preventable diseases. Measles, for example, has seen a 30% increase in cases globally."

*World Health Organization (WHO)*  
February 2019



## Vaccine hesitancy



Vaccine hesitancy – the reluctance or refusal to vaccinate despite the availability of vaccines – threatens to reverse progress made in tackling vaccine-preventable diseases. Vaccination is one of the most cost-effective ways of avoiding disease – it currently prevents 2.3 million deaths a year, and a further 1.5 million could be avoided if global coverage of vaccinations improved.

Measles, for example, has seen a 30% increase in cases globally. The reasons for this are complex, and not all of these cases are due to vaccine hesitancy. However, some countries that were close to eliminating the disease have seen a resurgence.

The reasons why people choose not to vaccinate are complex: a vaccine advisory group to WHO identified complacency, inequities in accessing vaccines, and lack of confidence as key reasons underlying

"Whereas bots that spread malware and unsolicited content disseminated antivaccine messages, Russian trolls promoted discord. Accounts masquerading as legitimate users create false equivalency, eroding public consensus on vaccination."

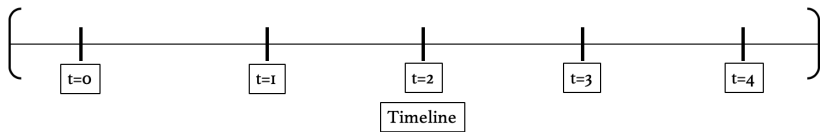
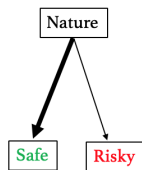
*American Journal of Public Health*  
October 2018

## WORLD

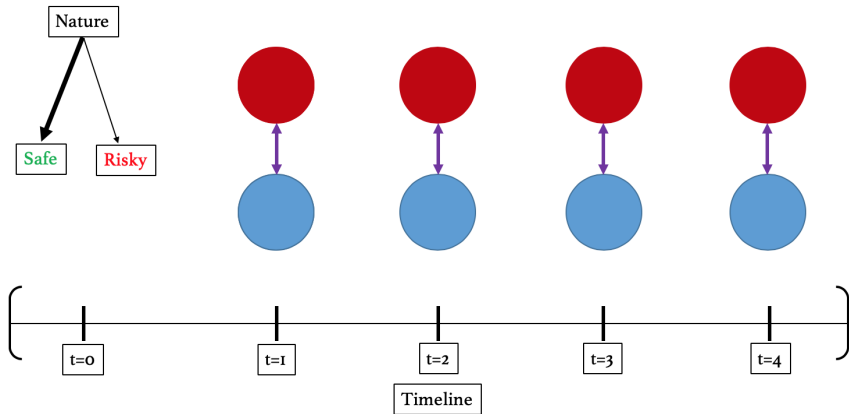
# RUSSIAN TROLLS PROMOTED ANTI-VACCINATION PROPAGANDA THAT MAY HAVE CAUSED MEASLES OUTBREAK, RESEARCHER CLAIMS

BY CRISTINA MAZA ON 2/14/19 AT 3:52 PM EST

# This Talk



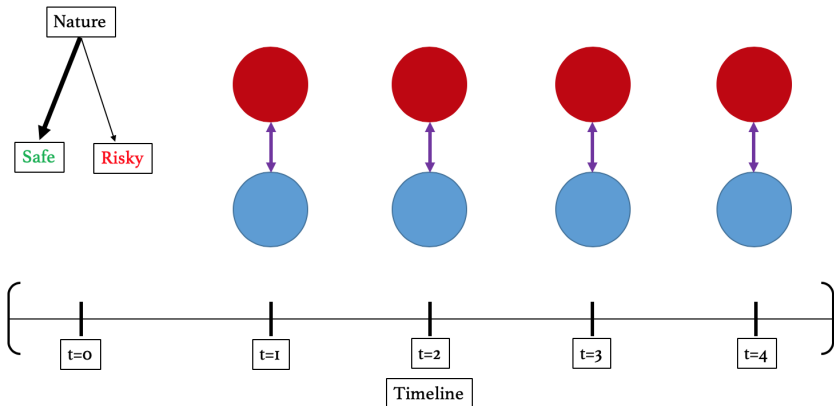
# This Talk



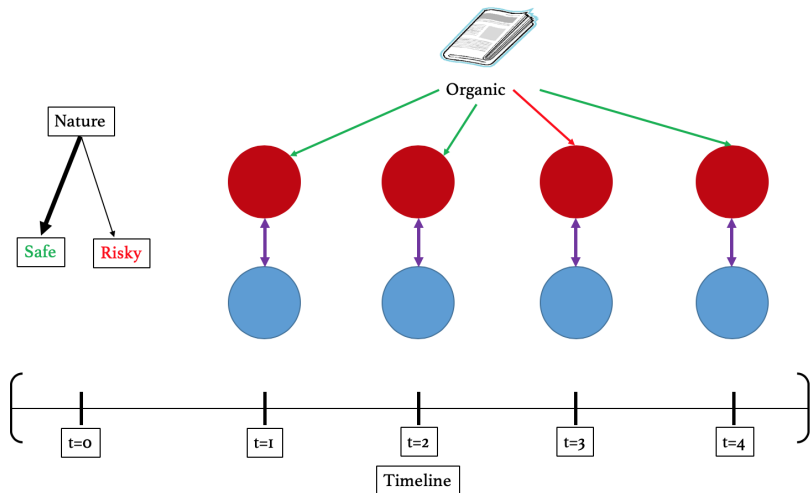
# This Talk



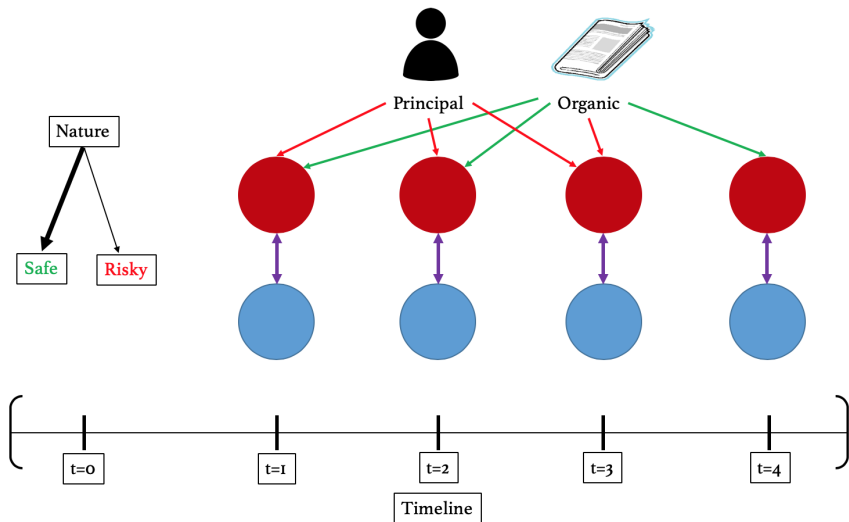
Organic



# This Talk

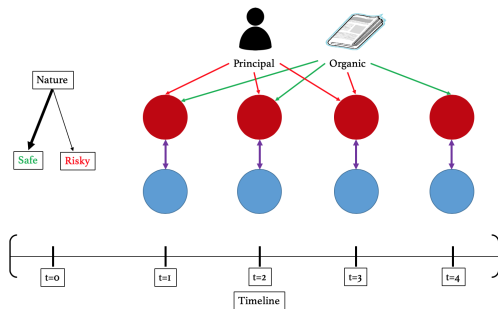


# This Talk



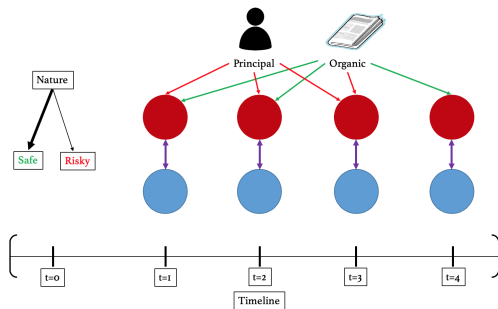


# This Talk



- We study the classical **learning problem** with unknown state of the world  $y \in \{S, R\}$ . Is vaccination safe ( $S$ ) or risky ( $R$ )?

# This Talk



- We study the classical **learning problem** with unknown state of the world  $y \in \{S, R\}$ . Is vaccination safe ( $S$ ) or risky ( $R$ )?
  - **Organic news**: News generation process for each agent, receive signals  $s_{i,t} \in \{S, R\}$  correlated with the underlying state at arrival  $t$ .
  - **Fake news**: Spurious process with signals unrelated to the state.

## Related Literature

- Bayesian learning
  - Acemoglu et al (2011), Bikhchandani et al (1992)
- DeGroot-style learning
  - Golub and Jackson (2010), Jadbabaie et al (2012), Molavi et al (2018)
- Mixed-learning environments
  - Mueller-Frank (2014), Chandrasekhar et al (2015), Bohren and Hauser (2017), Pennycook and Rand (2018)
- Propagation of fake news and misinformation
  - Candogan and Drakopoulos (2017), Papanastasiou (2018), Acemoglu et al (2010)
- Reputation effects
  - Kreps and Wilson (1982), Milgrom and Roberts (1982), Fudenberg-Levine (1989), Gossner (2011)

# Main Contribution

- **Question:** How does learning depend on...

# Main Contribution

- **Question:** How does learning depend on...
  - ...the sophistication and reasoning abilities of the agents?

# Main Contribution

- **Question:** How does learning depend on...
  - ...the sophistication and reasoning abilities of the agents?
  - ...the signals received by the agents?

# Main Contribution

- **Question:** How does learning depend on...
  - ...the sophistication and reasoning abilities of the agents?
  - ...the signals received by the agents?
  - ...the underlying communication structure?

# Main Contribution

- **Question:** How does learning depend on...
  - ...the sophistication and reasoning abilities of the agents?
  - ...the signals received by the agents?
  - ...the underlying communication structure?
- Assume there is a **principal** who interacts with the agents and may have a specific agenda (i.e., wants them to mislearn).
  - Examples: Russian propaganda, oil interests, marketing campaigns, etc.
- When the signal structure is **endogenous**, can agents learn the truth?



# Basic Model

- $n$  agents interact in continuous-time over a finite horizon  $T$ .

# Basic Model

- $n$  agents interact in continuous-time over a finite horizon  $T$ .
- Unknown state of the world  $y \in \{S, R\}$  drawn once at  $t = 0$  with common prior  $q$  of state  $S$ .

# Basic Model

- $n$  agents interact in continuous-time over a finite horizon  $T$ .
- Unknown state of the world  $y \in \{S, R\}$  drawn once at  $t = 0$  with common prior  $q$  of state  $S$ .
- Agent  $i$ 's **organic news** generation process is Poisson with unknown intensity  $\lambda_i \geq 0$ .
- Each Poisson arrival at time  $t$  provides a signal  $s_{i,t}$  with distribution  $\mathbb{P}[s_{i,t} = y] = p_i > 1/2$ , independent over time and across agents.

# Basic Model

- $n$  agents interact in continuous-time over a finite horizon  $T$ .
- Unknown state of the world  $y \in \{S, R\}$  drawn once at  $t = 0$  with common prior  $q$  of state  $S$ .
- Agent  $i$ 's **organic news** generation process is Poisson with unknown intensity  $\lambda_i \geq 0$ .
- Each Poisson arrival at time  $t$  provides a signal  $s_{i,t}$  with distribution  $\mathbb{P}[s_{i,t} = y] = p_i > 1/2$ , independent over time and across agents.
- If  $\lambda_i = 0$  agent  $i$  is **oblivious** - she cannot distinguish the state on her own.
- If  $\lambda_i > 0$  agent  $i$  is **informed** - she can distinguish the true state when only organic news is received.

# Principal Setup

- Principal has a type  $\omega$  either truthful ( $\mathcal{T}$ ) or strategic ( $\mathcal{S}$ ).

# Principal Setup

- Principal has a type  $\omega$  either truthful ( $\mathcal{T}$ ) or strategic ( $\mathcal{S}$ ).
- Principal chooses a **message**  $\hat{y} \in \{S, R\}$  and then chooses a **network strategy**  $\mathbf{x} \in \{0, 1\}^n$ .

# Principal Setup

- Principal has a type  $\omega$  either truthful ( $\mathcal{T}$ ) or strategic ( $\mathcal{S}$ ).
- Principal chooses a **message**  $\hat{y} \in \{S, R\}$  and then chooses a **network strategy**  $\mathbf{x} \in \{0, 1\}^n$ .
- If the principal is truthful, he is committed to playing  $\mathbf{x} = 0$ . If the principal is strategic, he may choose any strategy he wants.

# Principal Setup

- Principal has a type  $\omega$  either truthful ( $\mathcal{T}$ ) or strategic ( $\mathcal{S}$ ).
- Principal chooses a **message**  $\hat{y} \in \{S, R\}$  and then chooses a **network strategy**  $\mathbf{x} \in \{0, 1\}^n$ .
- If the principal is truthful, he is committed to playing  $\mathbf{x} = 0$ . If the principal is strategic, he may choose any strategy he wants.
- Principal pays an investment cost of  $\epsilon > 0$  for every agent  $i$  influenced (i.e.,  $c(\mathbf{x}) = \sum_{i=1}^n \epsilon \mathbf{1}_{x_i=1}$ ).



# Principal Setup

- Principal has a type  $\omega$  either truthful ( $\mathcal{T}$ ) or strategic ( $\mathcal{S}$ ).
- Principal chooses a **message**  $\hat{y} \in \{S, R\}$  and then chooses a **network strategy**  $\mathbf{x} \in \{0, 1\}^n$ .
- If the principal is truthful, he is committed to playing  $\mathbf{x} = 0$ . If the principal is strategic, he may choose any strategy he wants.
- Principal pays an investment cost of  $\epsilon > 0$  for every agent  $i$  influenced (i.e.,  $c(\mathbf{x}) = \sum_{i=1}^n \epsilon \mathbf{1}_{x_i=1}$ ).
- Investment to influence agent  $i$ : send an independent Poisson process of fixed intensity  $\lambda^*$  with signal  $s = \hat{y}$ .

# Principal Setup

- Principal has a type  $\omega$  either truthful ( $\mathcal{T}$ ) or strategic ( $\mathcal{S}$ ).
- Principal chooses a **message**  $\hat{y} \in \{S, R\}$  and then chooses a **network strategy**  $\mathbf{x} \in \{0, 1\}^n$ .
- If the principal is truthful, he is committed to playing  $\mathbf{x} = 0$ . If the principal is strategic, he may choose any strategy he wants.
- Principal pays an investment cost of  $\epsilon > 0$  for every agent  $i$  influenced (i.e.,  $c(\mathbf{x}) = \sum_{i=1}^n \epsilon \mathbf{1}_{x_i=1}$ ).
- Investment to influence agent  $i$ : send an independent Poisson process of fixed intensity  $\lambda^*$  with signal  $s = \hat{y}$ .
  - **Key assumption**: Agents cannot distinguish between organic news and principal's fake news.

# Agent Setup

- Agents have different reasoning abilities: either **Bayesian** ( $\mathcal{B}$ ) or **DeGroot** ( $\mathcal{D}$ ).

# Agent Setup

- Agents have different reasoning abilities: either **Bayesian** ( $\mathcal{B}$ ) or **DeGroot** ( $\mathcal{D}$ ).
- Each Bayesian agent  $i$  has a neighborhood  $\mathcal{N}_i \subset \{1, \dots, n\}$  of friends it observes.

# Agent Setup

- Agents have different reasoning abilities: either **Bayesian** ( $\mathcal{B}$ ) or **DeGroot** ( $\mathcal{D}$ ).
- Each Bayesian agent  $i$  has a neighborhood  $\mathcal{N}_i \subset \{1, \dots, n\}$  of friends it observes.
- Each DeGroot agent  $i$  is influenced by agent  $j$ 's beliefs proportionally to  $\alpha_{ij}$  and by its own news signals proportionally to  $\theta_i$ , with  $\theta_i + \sum_{j=1}^n \alpha_{ij} = 1$ .

# Agent Setup

- Agents have different reasoning abilities: either **Bayesian** ( $\mathcal{B}$ ) or **DeGroot** ( $\mathcal{D}$ ).
- Each Bayesian agent  $i$  has a neighborhood  $\mathcal{N}_i \subset \{1, \dots, n\}$  of friends it observes.
- Each DeGroot agent  $i$  is influenced by agent  $j$ 's beliefs proportionally to  $\alpha_{ij}$  and by its own news signals proportionally to  $\theta_i$ , with  $\theta_i + \sum_{j=1}^n \alpha_{ij} = 1$ .
- For every interval of length  $\Delta > 0$ , agents exchange beliefs  $\pi_{i,t}$  about the state of the world  $y$ :

# Agent Setup

- Agents have different reasoning abilities: either **Bayesian** ( $\mathcal{B}$ ) or **DeGroot** ( $\mathcal{D}$ ).
- Each Bayesian agent  $i$  has a neighborhood  $\mathcal{N}_i \subset \{1, \dots, n\}$  of friends it observes.
- Each DeGroot agent  $i$  is influenced by agent  $j$ 's beliefs proportionally to  $\alpha_{ij}$  and by its own news signals proportionally to  $\theta_i$ , with  $\theta_i + \sum_{j=1}^n \alpha_{ij} = 1$ .
- For every interval of length  $\Delta > 0$ , agents exchange beliefs  $\pi_{i,t}$  about the state of the world  $y$ :
  - Bayesian agents play a **perfect Bayesian equilibrium** against the principal observing beliefs  $\pi_{j,t}$  in their neighborhood  $\mathcal{N}_i$ .

# Agent Setup

- Agents have different reasoning abilities: either **Bayesian** ( $\mathcal{B}$ ) or **DeGroot** ( $\mathcal{D}$ ).
- Each Bayesian agent  $i$  has a neighborhood  $\mathcal{N}_i \subset \{1, \dots, n\}$  of friends it observes.
- Each DeGroot agent  $i$  is influenced by agent  $j$ 's beliefs proportionally to  $\alpha_{ij}$  and by its own news signals proportionally to  $\theta_i$ , with  $\theta_i + \sum_{j=1}^n \alpha_{ij} = 1$ .
- For every interval of length  $\Delta > 0$ , agents exchange beliefs  $\pi_{i,t}$  about the state of the world  $y$ :
  - Bayesian agents play a **perfect Bayesian equilibrium** against the principal observing beliefs  $\pi_{j,t}$  in their neighborhood  $\mathcal{N}_i$ .
  - DeGroot agents update beliefs taking news at **face value**:

$$\pi_{i,t+1} = \theta_i \cdot \text{BU}(\text{news}_t) + \sum_{j=1}^n \alpha_{ij} \pi_{j,t}$$



# Terminal Payoffs

- At time  $T$ , agent  $i$  is forced to commit to one action  $a_i \in \{S, R\}$ .

# Terminal Payoffs

- At time  $T$ , agent  $i$  is forced to commit to one action  $a_i \in \{S, R\}$ .
  - Agents want to **match** their action  $a_i$  with the unknown state  $y$ .

# Terminal Payoffs

- At time  $T$ , agent  $i$  is forced to commit to one action  $a_i \in \{S, R\}$ .
  - Agents want to **match** their action  $a_i$  with the unknown state  $y$ .
  - Principal always wants agents to play  $a_i = R$ , regardless of  $y$ .

# Terminal Payoffs

- At time  $T$ , agent  $i$  is forced to commit to one action  $a_i \in \{S, R\}$ .
  - Agents want to **match** their action  $a_i$  with the unknown state  $y$ .
  - Principal always wants agents to play  $a_i = R$ , regardless of  $y$ .

		Agent	
		R	S
State $y$	R	1, $1 + b$	0, 0
	S	1, $b$	0, 1

Table: Terminal Game.

- Agent's payoff  $u_i^a(y, a_i)$  and assume  $b \in (-1, 1)$ .
- Principal's benefit given additively  $u^p(\mathbf{a}) = \sum_{i=1}^n u_i^p(a_i)$ .
- Total payoff of principal is **benefit less investment**:  $u^p(\mathbf{a}) - c(\mathbf{x})$ .

# Manipulation

- We say that agent  $i$  is **manipulated** if in equilibrium:

# Manipulation

- We say that agent  $i$  is **manipulated** if in equilibrium:
  - (a) Her terminal action  $a_i$  *does not* match the underlying state  $y$  when the principal's type is **strategic**.

# Manipulation

- We say that agent  $i$  is **manipulated** if in equilibrium:
  - (a) Her terminal action  $a_i$  *does not* match the underlying state  $y$  when the principal's type is **strategic**.
  - (b) Her terminal action  $a_i$  *does* match the underlying state  $y$  when the principal's type is **truthful**.

# Manipulation

- We say that agent  $i$  is **manipulated** if in equilibrium:
  - (a) Her terminal action  $a_i$  *does not* match the underlying state  $y$  when the principal's type is **strategic**.
  - (b) Her terminal action  $a_i$  *does* match the underlying state  $y$  when the principal's type is **truthful**.
- We refer to the collection of neighborhoods  $\{\mathcal{N}_i\}_{i \in \mathcal{B}}$  and weights  $(\theta, \mathbf{A})$  as the *network structure*.
- We say the network is **impervious** to manipulation if no agent is manipulated, and otherwise **susceptible**.



# Manipulation

- We say that agent  $i$  is **manipulated** if in equilibrium:
  - (a) Her terminal action  $a_i$  *does not* match the underlying state  $y$  when the principal's type is **strategic**.
  - (b) Her terminal action  $a_i$  *does* match the underlying state  $y$  when the principal's type is **truthful**.
- We refer to the collection of neighborhoods  $\{\mathcal{N}_i\}_{i \in \mathcal{B}}$  and weights  $(\theta, \mathbf{A})$  as the *network structure*.
- We say the network is **impervious** to manipulation if no agent is manipulated, and otherwise **susceptible**.
- By **Jadbabaie et al (2012)**: under standard network connectedness assumptions, all agents learn the true state  $y$  for large  $T$ .

# Manipulation

- We say that agent  $i$  is **manipulated** if in equilibrium:
  - (a) Her terminal action  $a_i$  *does not* match the underlying state  $y$  when the principal's type is **strategic**.
  - (b) Her terminal action  $a_i$  *does* match the underlying state  $y$  when the principal's type is **truthful**.
- We refer to the collection of neighborhoods  $\{\mathcal{N}_i\}_{i \in \mathcal{B}}$  and weights  $(\theta, \mathbf{A})$  as the *network structure*.
- We say the network is **impervious** to manipulation if no agent is manipulated, and otherwise **susceptible**.
- By **Jadbabaie et al (2012)**: under standard network connectedness assumptions, all agents learn the true state  $y$  for large  $T$ .
- What if the principal is strategic and  $T \rightarrow \infty$ ?

# Bayesian Learning

## Theorem. (Bayesian Learning)

Under standard network connectedness assumptions, there exists  $p^* < 1$  such that if there is one informed agent with  $p_i > p^*$ , no Bayesian agent is manipulated.

# Bayesian Learning

## Theorem. (Bayesian Learning)

Under standard network connectedness assumptions, there exists  $p^* < 1$  such that if there is one informed agent with  $p_i > p^*$ , no Bayesian agent is manipulated.

- Need for “experts” as in [Acemoglu et al \(2011\)](#); otherwise may be profitable to invest  $x = 1$  when  $\epsilon$  is small and fool the entire community.

# Bayesian Learning

## Theorem. (Bayesian Learning)

Under standard network connectedness assumptions, there exists  $p^* < 1$  such that if there is one informed agent with  $p_i > p^*$ , no Bayesian agent is manipulated.

- Need for “experts” as in [Acemoglu et al \(2011\)](#); otherwise may be profitable to invest  $x = 1$  when  $\epsilon$  is small and fool the entire community.
- Note the “expert” need not be a Bayesian, and it is even possible all Bayesian agents be oblivious (i.e., receive no news themselves)!

# DeGroot Learning

- Bayesian agents are eventually **stubborn agents** with accurate beliefs (i.e.,  $\pi_i(y) = 1$  for all  $i \in \mathcal{B}$ ).

# DeGroot Learning

- Bayesian agents are eventually **stubborn agents** with accurate beliefs (i.e.,  $\pi_i(y) = 1$  for all  $i \in \mathcal{B}$ ).
- If  $\lambda_i > \lambda^*/(2p_i - 1)$ , DeGroot agent  $i$  is a *specialist* and otherwise she is *amenable*.

# DeGroot Learning

- Bayesian agents are eventually **stubborn agents** with accurate beliefs (i.e.,  $\pi_i(y) = 1$  for all  $i \in \mathcal{B}$ ).
- If  $\lambda_i > \lambda^*/(2p_i - 1)$ , DeGroot agent  $i$  is a *specialist* and otherwise she is *amenable*.
  - **Specialist** agents are discerning of the true state from their own news regardless of the presence of fake news.
  - **Amenable** agents have little precision in their signal distribution, so if  $x_i = 0$  the news appears as state  $y$ , but if  $x_i = 1$  it appears as the opposite state.



# DeGroot Learning

- Bayesian agents are eventually **stubborn agents** with accurate beliefs (i.e.,  $\pi_i(y) = 1$  for all  $i \in \mathcal{B}$ ).
- If  $\lambda_i > \lambda^*/(2p_i - 1)$ , DeGroot agent  $i$  is a *specialist* and otherwise she is *amenable*.
  - **Specialist** agents are discerning of the true state from their own news regardless of the presence of fake news.
  - **Amenable** agents have little precision in their signal distribution, so if  $x_i = 0$  the news appears as state  $y$ , but if  $x_i = 1$  it appears as the opposite state.
- Limit beliefs of DeGroots can be characterized by:

$$\pi(y) = (\mathbf{I} - \tilde{\mathbf{A}})^{-1}(\boldsymbol{\theta} \otimes \boldsymbol{\gamma})$$

where  $\gamma_i = 1$  if agent  $i$  is either Bayesian, a specialist, or  $x_i = 0$ .

## Sufficiency for Imperviousness

- Let  $\mathcal{P}_{ij}$  be the set of paths between agents  $i$  and  $j$  in the network.
- We can define the **log-diameter** of the network  $\mathbf{G}$  to be:

$$d_{\mathbf{G}} \equiv \max_{i,j \in \mathcal{D}} \min_{P_{ij} \in \mathcal{P}_{ij}} \sum_{(k \rightarrow \ell) \in P_{ij}} -\log_n(\alpha_{k\ell})$$

where we “skip” over Bayesian agents in the path summation.

## Sufficiency for Imperviousness

- Let  $\mathcal{P}_{ij}$  be the set of paths between agents  $i$  and  $j$  in the network.
- We can define the **log-diameter** of the network  $\mathbf{G}$  to be:

$$d_{\mathbf{G}} \equiv \max_{i,j \in \mathcal{D}} \min_{P_{ij} \in \mathcal{P}_{ij}} \sum_{(k \rightarrow \ell) \in P_{ij}} -\log_n(\alpha_{k\ell})$$

where we “skip” over Bayesian agents in the path summation.

### Theorem. (Small Log-Diameter)

For any  $\delta$ , there exists  $m$  such that if network  $\mathbf{G}$  has log-diameter at most  $\delta$ , then it is **impervious** to manipulation if there are at least  $m$  Bayesian agents.

## Sufficiency for Imperviousness

- Let  $\mathcal{P}_{ij}$  be the set of paths between agents  $i$  and  $j$  in the network.
- We can define the **log-diameter** of the network  $\mathbf{G}$  to be:

$$d_{\mathbf{G}} \equiv \max_{i,j \in \mathcal{D}} \min_{P_{ij} \in \mathcal{P}_{ij}} \sum_{(k \rightarrow \ell) \in P_{ij}} -\log_n(\alpha_{k\ell})$$

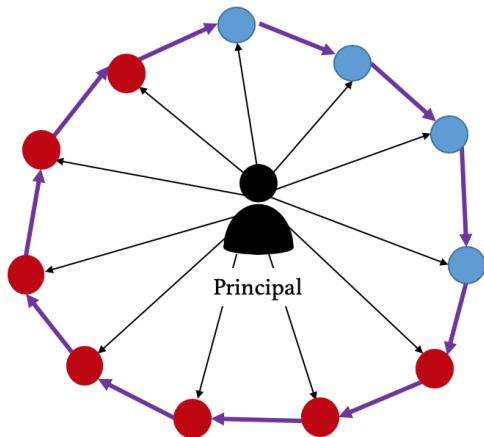
where we “skip” over Bayesian agents in the path summation.

### Theorem. (Small Log-Diameter)

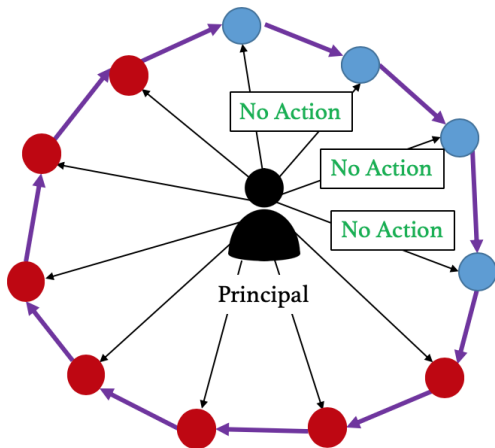
For any  $\delta$ , there exists  $m$  such that if network  $\mathbf{G}$  has log-diameter at most  $\delta$ , then it is **impervious** to manipulation if there are at least  $m$  Bayesian agents.

- Implies complete network is impervious; star network is also impervious *even if* Bayesians are on the periphery.

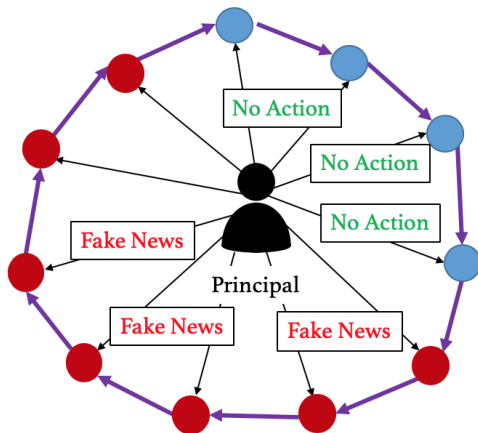
# The Ring Network



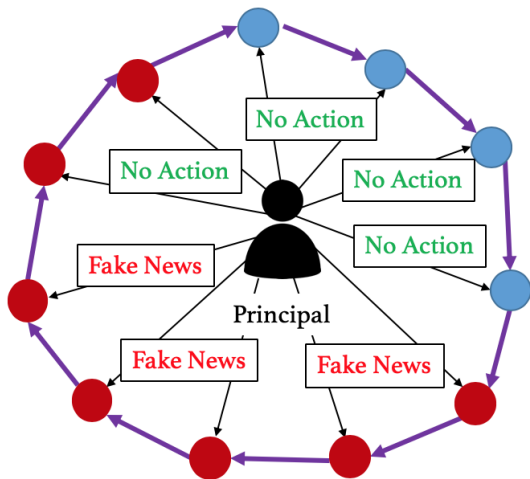
# The Ring Network



# The Ring Network

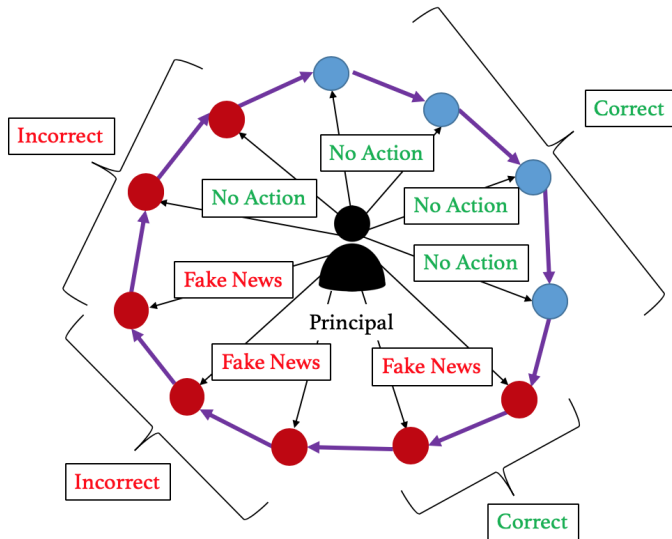


# The Ring Network

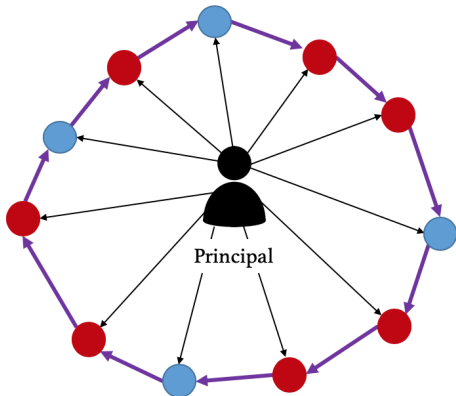




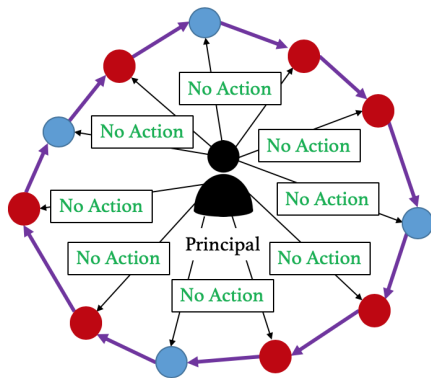
# The Ring Network



# The Ring Network, cont.

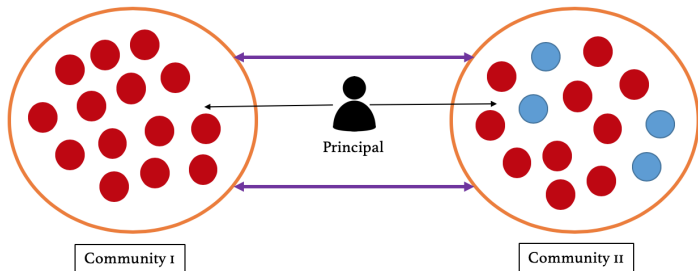


# The Ring Network, cont.

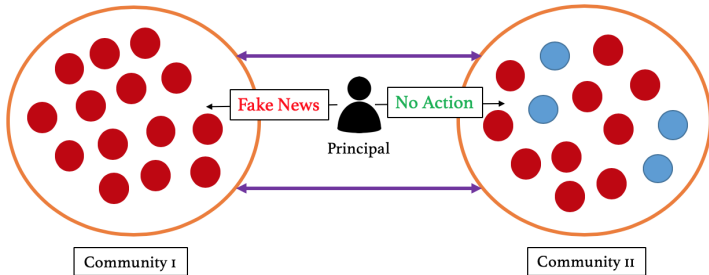


- Need a linear number of Bayesian agents **sprinkled** throughout!

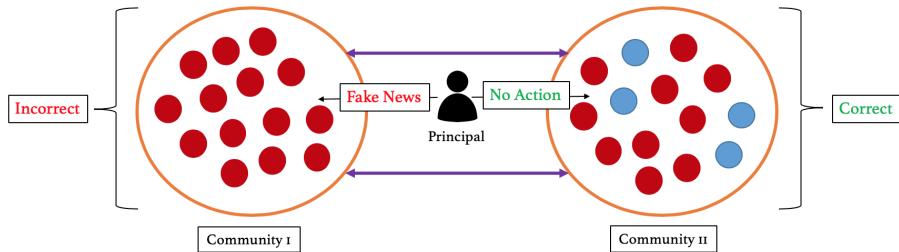
# Weak Homophily



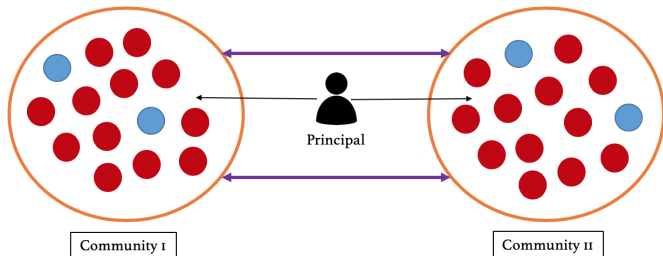
# Weak Homophily



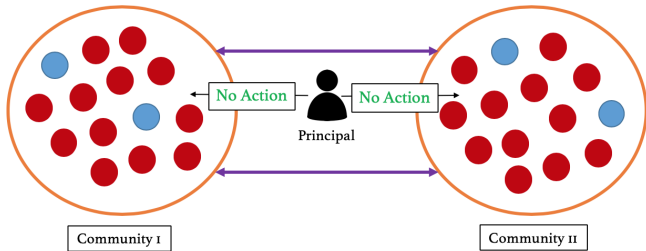
# Weak Homophily



# Weak Homophily

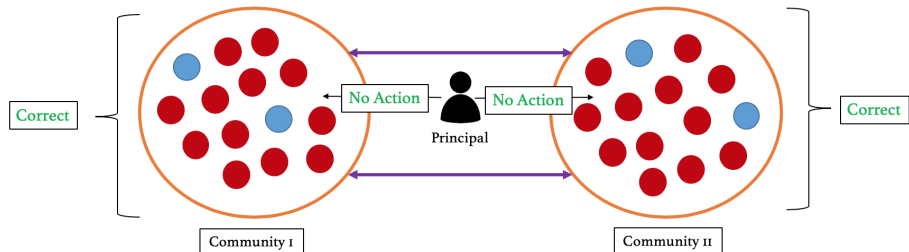


# Weak Homophily

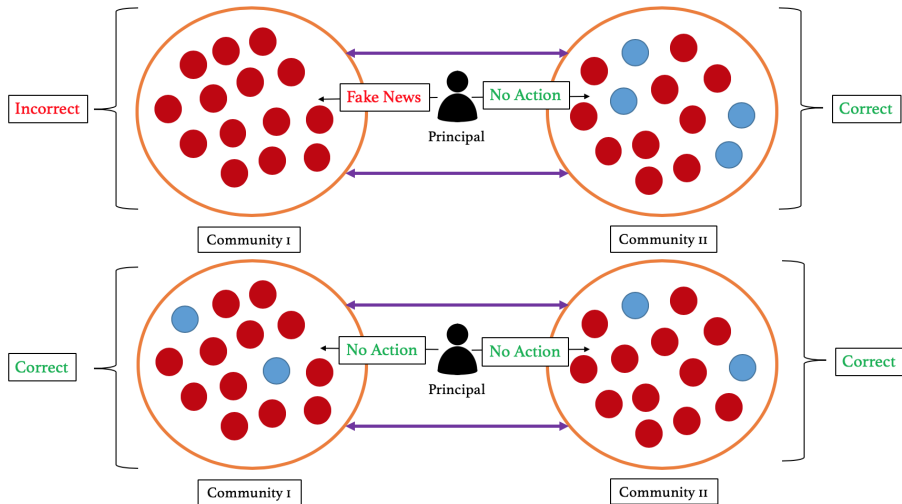




# Weak Homophily



# Weak Homophily



# Weak Homophily

- **Weak homophily model:**  $k$  groups with proportion of the population  $(s_1, \dots, s_k)$ . Within-group link probability is  $p_s$  and between-group link probability is  $p_d$ .

# Weak Homophily

- **Weak homophily model:**  $k$  groups with proportion of the population  $(s_1, \dots, s_k)$ . Within-group link probability is  $p_s$  and between-group link probability is  $p_d$ .
- Under certain regularity conditions, manipulation in random network *reduces* to manipulation in expected network (see paper).

# Weak Homophily

- **Weak homophily model:**  $k$  groups with proportion of the population  $(s_1, \dots, s_k)$ . Within-group link probability is  $p_s$  and between-group link probability is  $p_d$ .
- Under certain regularity conditions, manipulation in random network *reduces* to manipulation in expected network (see paper).
- With communities of the **same size** ( $s_1 = \dots = s_k = 1/k$ ) the network cannot go from impervious to susceptible if we (i) decrease  $p_s$ , (ii) increase  $p_d$ , or (iii) more evenly distribute the Bayesian agents. However:

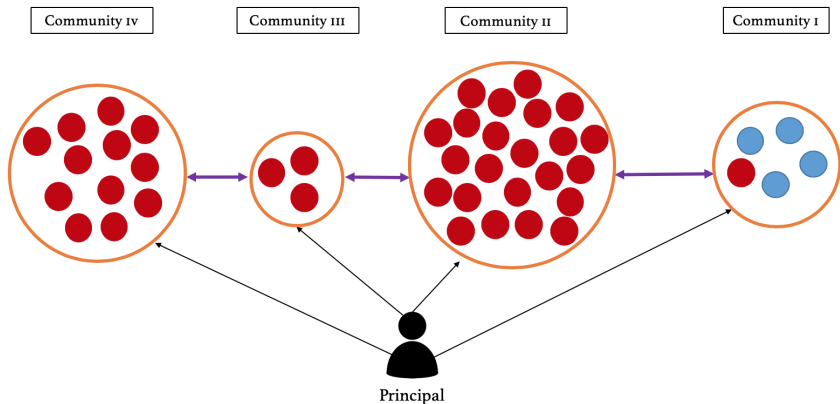
# Weak Homophily

- **Weak homophily model:**  $k$  groups with proportion of the population  $(s_1, \dots, s_k)$ . Within-group link probability is  $p_s$  and between-group link probability is  $p_d$ .
- Under certain regularity conditions, manipulation in random network *reduces* to manipulation in expected network (see paper).
- With communities of the **same size** ( $s_1 = \dots = s_k = 1/k$ ) the network cannot go from impervious to susceptible if we (i) decrease  $p_s$ , (ii) increase  $p_d$ , or (iii) more evenly distribute the Bayesian agents. However:
  - All of these can cause manipulation to **increase** if manipulation already exists.

# Weak Homophily

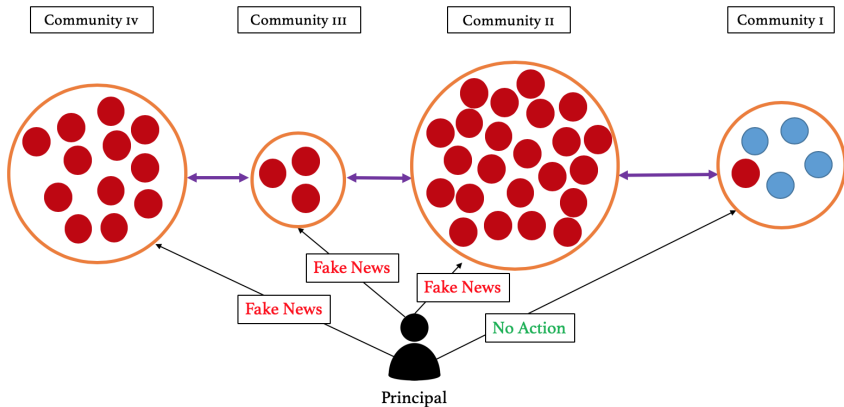
- **Weak homophily model:**  $k$  groups with proportion of the population  $(s_1, \dots, s_k)$ . Within-group link probability is  $p_s$  and between-group link probability is  $p_d$ .
- Under certain regularity conditions, manipulation in random network *reduces* to manipulation in expected network (see paper).
- With communities of the **same size** ( $s_1 = \dots = s_k = 1/k$ ) the network cannot go from impervious to susceptible if we (i) decrease  $p_s$ , (ii) increase  $p_d$ , or (iii) more evenly distribute the Bayesian agents. However:
  - All of these can cause manipulation to **increase** if manipulation already exists.
  - Does not hold for communities of **different sizes**.

# Strong Homophily

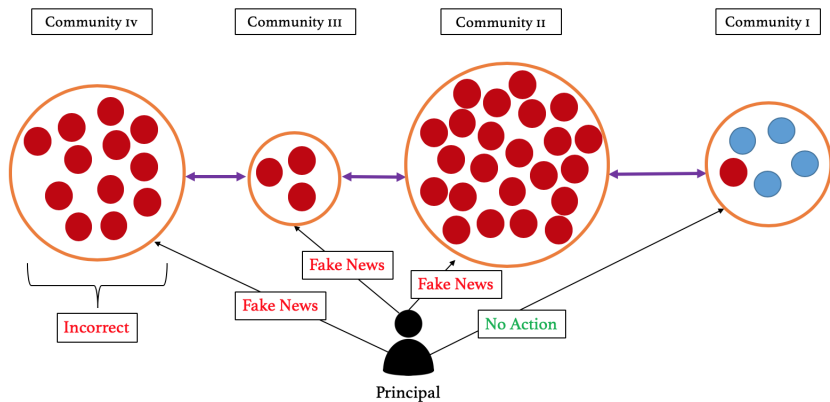




# Strong Homophily



# Strong Homophily



# Strong Homophily

- **Strong homophily:** Groups are totally ordered. Within-group link probability is still  $p_s$ , but between-group link probability is  $p_d$  only for the two nearest communities (and otherwise 0).
  - For simplicity, assume one island contains all of the Bayesian agents.

# Strong Homophily

- **Strong homophily:** Groups are totally ordered. Within-group link probability is still  $p_s$ , but between-group link probability is  $p_d$  only for the two nearest communities (and otherwise 0).
  - For simplicity, assume one island contains all of the Bayesian agents.
- For large  $n$ , the strong homophily model with sufficiently many communities  $\{s_\ell\}_{\ell=1}^k$  is **susceptible** to manipulation even with  $O(n)$  Bayesians.

# Strong Homophily

- **Strong homophily:** Groups are totally ordered. Within-group link probability is still  $p_s$ , but between-group link probability is  $p_d$  only for the two nearest communities (and otherwise 0).
  - For simplicity, assume one island contains all of the Bayesian agents.
- For large  $n$ , the strong homophily model with sufficiently many communities  $\{s_\ell\}_{\ell=1}^k$  is **susceptible** to manipulation even with  $O(n)$  Bayesians.
  - Can be shown the principal manipulates all but  $O(\sqrt{n})$  DeGroot agents.

# Strong Homophily

- **Strong homophily:** Groups are totally ordered. Within-group link probability is still  $p_s$ , but between-group link probability is  $p_d$  only for the two nearest communities (and otherwise 0).
  - For simplicity, assume one island contains all of the Bayesian agents.
- For large  $n$ , the strong homophily model with sufficiently many communities  $\{s_\ell\}_{\ell=1}^k$  is **susceptible** to manipulation even with  $O(n)$  Bayesians.
  - Can be shown the principal manipulates all but  $O(\sqrt{n})$  DeGroot agents.
  - The weak homophily model with the same community structure  $\{s_\ell\}_{\ell=1}^k$  is impervious for  $O(1)$  Bayesians because of small log-diameter.

# Extensions

- Endogenize the choice of fake news intensity  $\lambda^*$ .
  - Reduces to **idiosyncratic costs** of investment  $\epsilon_i$  for each agent.

# Extensions

- Endogenize the choice of fake news intensity  $\lambda^*$ .
  - Reduces to **idiosyncratic costs** of investment  $\epsilon_i$  for each agent.
- Allow the strategic principal to send **mixed messages**.
  - Does not change equilibrium analysis for Bayesians, can only improve learning for mechanical DeGroots.



# Extensions

- Endogenize the choice of fake news intensity  $\lambda^*$ .
  - Reduces to **idiosyncratic costs** of investment  $\epsilon_i$  for each agent.
- Allow the strategic principal to send **mixed messages**.
  - Does not change equilibrium analysis for Bayesians, can only improve learning for mechanical DeGroots.
- Allow the strategic principal to vary the intensity over time.
  - For large  $T$ , equivalent to the fixed intensity equal to average.

# Extensions

- Endogenize the choice of fake news intensity  $\lambda^*$ .
  - Reduces to **idiosyncratic costs** of investment  $\epsilon_i$  for each agent.
- Allow the strategic principal to send **mixed messages**.
  - Does not change equilibrium analysis for Bayesians, can only improve learning for mechanical DeGroots.
- Allow the strategic principal to vary the intensity over time.
  - For large  $T$ , equivalent to the fixed intensity equal to average.
- More than two states, **uncertainty** about strategic principal's agenda.
  - Can be accommodated under a few more network knowledge assumptions.

# Extensions

- Endogenize the choice of fake news intensity  $\lambda^*$ .
  - Reduces to **idiosyncratic costs** of investment  $\epsilon_i$  for each agent.
- Allow the strategic principal to send **mixed messages**.
  - Does not change equilibrium analysis for Bayesians, can only improve learning for mechanical DeGroots.
- Allow the strategic principal to vary the intensity over time.
  - For large  $T$ , equivalent to the fixed intensity equal to average.
- More than two states, **uncertainty** about strategic principal's agenda.
  - Can be accommodated under a few more network knowledge assumptions.
- Characterization of equilibrium for **small  $T$** .
  - Principal may worry about reputation effects.