Influence Campaigns

Evan Sadler

Columbia University

GraphEx Conference
April 24, 2019

Motivation

Huge investments to manipulate opinions and beliefs

Motivation

Huge investments to manipulate opinions and beliefs

- 2016 US elections (ongoing Russian interference...)

Motivation

Huge investments to manipulate opinions and beliefs

- 2016 US elections (ongoing Russian interference...)
- Political advocacy (climate change)

Motivation

Huge investments to manipulate opinions and beliefs

- 2016 US elections (ongoing Russian interference...)
- Political advocacy (climate change)
- Public health (vaccines)

Motivation

Huge investments to manipulate opinions and beliefs

- 2016 US elections (ongoing Russian interference...)
- Political advocacy (climate change)
- Public health (vaccines)
- Advertising

Motivation

Huge investments to manipulate opinions and beliefs

- 2016 US elections (ongoing Russian interference...)
- Political advocacy (climate change)
- Public health (vaccines)
- Advertising

Some common features

- Expressed opinions and beliefs are *coarse* and *diverse*
- People are influenced by their friends
- Heterogeneous susceptibility to influence

Motivation

Huge investments to manipulate opinions and beliefs

- 2016 US elections (ongoing Russian interference...)
- Political advocacy (climate change)
- Public health (vaccines)
- Advertising

Some common features

- Expressed opinions and beliefs are *coarse* and *diverse*
- People are influenced by their friends
- Heterogeneous susceptibility to influence

High-level question: how shoud social networks affect evaluation and design of influence campaigns?

Opinion Dynamics

Introduce model with three features:

- Discrete opinions
- Random opportunities to update
- Stubborn agents

Opinion Dynamics

Introduce model with three features:

- Discrete opinions
- Random opportunities to update
- Stubborn agents

No consensus, non-degenerate steady-state distribution

- Opinions fluctuate over time

Opinion Dynamics

Introduce model with three features:

- Discrete opinions
- Random opportunities to update
- Stubborn agents

No consensus, non-degenerate steady-state distribution

- Opinions fluctuate over time

New questions:

- What drives opinion swings?
- Risk-reward tradeoffs in influence campaigns

Contribution

Tractable model of opinion dynamics with non-trivial steady state

Contribution

Tractable model of opinion dynamics with non-trivial steady state

Characterize influence, convergence rate, and variance of public opinion

Contribution

Tractable model of opinion dynamics with non-trivial steady state

Characterize influence, convergence rate, and variance of public opinion

Random graph limit theorems

- Dimensionality reduction

Contribution

Tractable model of opinion dynamics with non-trivial steady state

Characterize influence, convergence rate, and variance of public opinion

Random graph limit theorems

- Dimensionality reduction

Comparative statics:

- Homophily and polarization lead to smaller fluctuations
- View of minority group benefits from integration

Contribution

Tractable model of opinion dynamics with non-trivial steady state

Characterize influence, convergence rate, and variance of public opinion

Random graph limit theorems

- Dimensionality reduction

Comparative statics:

- Homophily and polarization lead to smaller fluctuations
- View of minority group benefits from integration

Framework to study targeted propaganda

Road Map

Model and interpretation

Influence, convergence rates, and variance in arbitary graphs

Types and the stochastic block model

Examples

A framework for propaganda

Related Work

Voter models (Yildiz et al., 2013; Mossel and Tamuz, 2014)

Related Work

Voter models (Yildiz et al., 2013; Mossel and Tamuz, 2014)

DeGroot models (Golub and Jackson, 2012)

Related Work

Voter models (Yildiz et al., 2013; Mossel and Tamuz, 2014)

DeGroot models (Golub and Jackson, 2012)

Social Learning literature

- Bayesian (Acemoglu et al., 2011; Lobel and Sadler, 2015, 2016)
- Non-Bayesian (Molavi et al., 2018)
- With misspecified models (Bohren and Hauser, 2016, 2018)

Communicating Views in a Network

n agents, two possible views $v \in \{0, 1\}$

Communicating Views in a Network

n agents, two possible views $v \in \{0, 1\}$

Undirected graph G describes connections among agents

- Assume connected

Communicating Views in a Network

n agents, two possible views $v \in \{0, 1\}$

Undirected graph G describes connections among agents

- Assume connected

Each period, nature selects $ij \in G$ uniformly at random

- i talks to j
- j adopts i's view with probability $\pi_j \in [0, 1]$

Communicating Views in a Network

n agents, two possible views $v \in \{0, 1\}$

Undirected graph G describes connections among agents

- Assume connected

Each period, nature selects $ij \in G$ uniformly at random

- i talks to j
- j adopts i's view with probability $\pi_j \in [0, 1]$

π_j is j's *persuadability*

- If $\pi_j = 0$, call j stubborn, otherwise j is open
- Assume at least one stubborn agent

Interpretation

What is a view?

- Opinion, political position, religion, belief about some fact
- Easy to generalize to arbitrary number of views

Interpretation

What is a view?

- Opinion, political position, religion, belief about some fact
- Easy to generalize to arbitrary number of views

Who are these stubborn agents?

- Depending on setting, partisans, zealots, subject matter experts

Interpretation

What is a view?

- Opinion, political position, religion, belief about some fact
- Easy to generalize to arbitrary number of views

Who are these stubborn agents?

- Depending on setting, partisans, zealots, subject matter experts

Alternatively, stubbornness could represent tendency to revert to initial view

- Formally, model multiple selves

Steady-State Views

Model is a finite Markov chain, 2^n states

Steady-State Views

Model is a finite Markov chain, 2^n states

Simplification: look at first and second moments

- $x_i = \mathbb{P}(v_i = 1)$
- $y_{ij} = \mathbb{P}(v_i = v_j = 1)$

Steady-State Views

Model is a finite Markov chain, 2^n states

Simplification: look at first and second moments

- $x_i = \mathbb{P}(v_i = 1)$
- $y_{ij} = \mathbb{P}(v_i = v_j = 1)$

Define $\omega = \frac{1}{n}\sum_{i=1}^{n} v_i$,

$$\mathbb{E}[\omega] = \frac{\mathbf{1}\cdot\mathbf{x}}{n}, \quad \mathrm{Var}[\omega] = \frac{\mathbf{1}^T(Y - \mathbf{x}\mathbf{x}^T)\mathbf{1}}{n^2}$$

Expected Views and Influence

Let d_i denote i's degree, $m = \sum_{i=1}^{n} d_i$, set of open agents O, set of stubborn agents S

Expected Views and Influence

Let d_i denote i's degree, $m = \sum_{i=1}^{n} d_i$, set of open agents O, set of stubborn agents S

Define row-stochastic matrix $\tilde{G}$ via

$$\tilde{g}_{ij} = \begin{cases} \chi(i = j) & \text{if } i \in S \\ \frac{\pi_i}{m} & \text{if } i \in O,\ ij \in G \\ 1 - \frac{\pi_i d_i}{m} & \text{if } i \in O,\ j = i. \end{cases}$$

Expected Views and Influence

Let d_i denote i's degree, $m = \sum_{i=1}^{n} d_i$, set of open agents O, set of stubborn agents S

Define row-stochastic matrix $\tilde{G}$ via

$$\tilde{g}_{ij} = \begin{cases} \chi(i = j) & \text{if } i \in S \\ \frac{\pi_i}{m} & \text{if } i \in O,\ ij \in G \\ 1 - \frac{\pi_i d_i}{m} & \text{if } i \in O,\ j = i. \end{cases}$$

$\tilde{g}_{ij}$ is probability that i adopts j's view in any period

$$\mathbf{x}(t+1) = \tilde{G}\mathbf{x}(t)$$

Expected Views and Influence

Theorem

$\mathbf{x} = \lim_{t \to \infty} \tilde{G}^t \mathbf{v}(0) = M\mathbf{x}(0)$

Expected Views and Influence

Theorem

$\mathbf{x} = \lim_{t\to\infty} \tilde{G}^t \mathbf{v}(0) = M\mathbf{x}(0)$

Steady state views must satisfy $\mathbf{x} = \tilde{G}\mathbf{x}$, write $\tilde{G}$ in block form

$$\tilde{G} = \begin{bmatrix} I & 0 \\ \tilde{G}_{OS} & \tilde{G}_{OO} \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} \mathbf{x}_S \\ \mathbf{x}_O \end{bmatrix}$$

Expected Views and Influence

Theorem

$\mathbf{x} = \lim_{t\to\infty} \tilde{G}^t \mathbf{v}(0) = M\mathbf{x}(0)$

Steady state views must satisfy $\mathbf{x} = \tilde{G}\mathbf{x}$, write $\tilde{G}$ in block form

$$\tilde{G} = \begin{bmatrix} I & 0 \\ \tilde{G}_{OS} & \tilde{G}_{OO} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \mathbf{x}_S \\ \mathbf{x}_O \end{bmatrix}$$

$\mathbf{x}_O = (I - \tilde{G}_{OO})^{-1}\tilde{G}_{OS}\mathbf{x}_S$

Convergence Rate

Using block form representation of $\tilde{G}$, can compute

$$\tilde{G}^t = \begin{bmatrix} I & 0 \\ \left(\sum_{k=0}^{t-1} \tilde{G}_{OO}^k\right) \tilde{G}_{OS} & \tilde{G}_{OO}^t \end{bmatrix}$$

Convergence Rate

Using block form representation of $\tilde{G}$, can compute

$$\tilde{G}^t = \begin{bmatrix} I & 0 \\ \left(\sum_{k=0}^{t-1} \tilde{G}_{OO}^k\right) \tilde{G}_{OS} & \tilde{G}_{OO}^t \end{bmatrix}$$

Implies

$$\mathbf{x}_O - \mathbf{x}_O(t) = \left(\sum_{k=t}^{\infty} \tilde{G}_{OO}^k\right) \tilde{G}_{OS}\mathbf{x}_S - \tilde{G}_{OO}^t \mathbf{x}_O(0) = \tilde{G}_{OO}^t \left(\mathbf{x}_O - \mathbf{x}_O(0)\right)$$

Convergence Rate

Using block form representation of $\tilde{G}$, can compute

$$\tilde{G}^t = \begin{bmatrix} I & 0 \\ \left(\sum_{k=0}^{t-1} \tilde{G}_{OO}^k\right) \tilde{G}_{OS} & \tilde{G}_{OO}^t \end{bmatrix}$$

Implies

$$\mathbf{x}_O - \mathbf{x}_O(t) = \left(\sum_{k=t}^{\infty} \tilde{G}_{OO}^k\right) \tilde{G}_{OS}\mathbf{x}_S - \tilde{G}_{OO}^t \mathbf{x}_O(0) = \tilde{G}_{OO}^t \left(\mathbf{x}_O - \mathbf{x}_O(0)\right)$$

Spectral radius of $\tilde{G}_{OO}$ gives a measure of convergence rate

Variance of Public Opinion

Recall matrix of second moments, $y_{ij} = \mathbb{P}(v_i = v_j = 1)$

Variance of Public Opinion

Recall matrix of second moments, $y_{ij} = \mathbb{P}(v_i = v_j = 1)$

If k persuades j at time t, then $v_i(t) = v_j(t) = 1$ iff $v_i(t-1) = v_k(t-1) = 1$

Variance of Public Opinion

Recall matrix of second moments, $y_{ij} = \mathbb{P}(v_i = v_j = 1)$

If k persuades j at time t, then $v_i(t) = v_j(t) = 1$ iff $v_i(t-1) = v_k(t-1) = 1$

Steady state condition for distinct i and j:

$$y_{ij} = \tilde{\mathbf{g}}_j \cdot \mathbf{y}_i + \tilde{\mathbf{g}}_i \cdot \mathbf{y}_j - y_{ij} \implies y_{ij} = \frac{1}{2}\left(\tilde{\mathbf{g}}_j \cdot \mathbf{y}_i + \tilde{\mathbf{g}}_i \cdot \mathbf{y}_j\right)$$

Variance of Public Opinion

Recall matrix of second moments, $y_{ij} = \mathbb{P}(v_i = v_j = 1)$

If k persuades j at time t, then $v_i(t) = v_j(t) = 1$ iff $v_i(t-1) = v_k(t-1) = 1$

Steady state condition for distinct i and j:

$$y_{ij} = \tilde{\mathbf{g}}_j \cdot \mathbf{y}_i + \tilde{\mathbf{g}}_i \cdot \mathbf{y}_j - y_{ij} \implies y_{ij} = \frac{1}{2}\left(\tilde{\mathbf{g}}_j \cdot \mathbf{y}_i + \tilde{\mathbf{g}}_i \cdot \mathbf{y}_j\right)$$

Diagonal entry $y_{ii} = x_i$

Variance of Public Opinion

Theorem

$$Y = \frac{1}{2}\left(Y\tilde{G}^T + \tilde{G}Y\right) - \frac{1}{2}\,\mathrm{diag}\left(Y\tilde{G}^T + \tilde{G}Y\right) + \Lambda_{\mathbf{x}}$$

Variance of Public Opinion

Theorem

$$Y = \frac{1}{2}\left(Y\tilde{G}^T + \tilde{G}Y\right) - \frac{1}{2}\,\mathrm{diag}\left(Y\tilde{G}^T + \tilde{G}Y\right) + \Lambda_{\mathbf{x}}$$

Can write Y in block form

$$Y = \begin{bmatrix} Y_{SS} & Y_{SO} \\ Y_{OS} & Y_{OO} \end{bmatrix}$$

Three blocks are easy: $Y_{SS} = \mathbf{x}_S\mathbf{x}_S^T$, $Y_{SO} = \mathbf{x}_S\mathbf{x}_O^T$, $Y_{OS} = \mathbf{x}_O\mathbf{x}_S^T$

Variance of Public Opinion

Theorem

$$Y = \frac{1}{2}\left(Y\tilde{G}^T + \tilde{G}Y\right) - \frac{1}{2}\operatorname{diag}\left(Y\tilde{G}^T + \tilde{G}Y\right) + \Lambda_{\mathbf{x}}$$

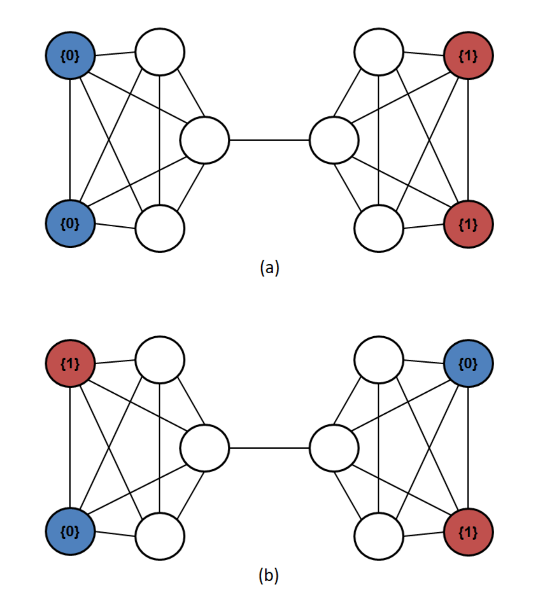
Can write Y in block form

$$Y = \begin{bmatrix} Y_{SS} & Y_{SO} \\ Y_{OS} & Y_{OO} \end{bmatrix}$$

Three blocks are easy: $Y_{SS} = \mathbf{x}_S\mathbf{x}_S^T$, $Y_{SO} = \mathbf{x}_S\mathbf{x}_O^T$, $Y_{OS} = \mathbf{x}_O\mathbf{x}_S^T$

Last block Y_{OO} is only block in which off diagonal entries contribute to the variance

An Example



The Stochastic Block Model

Finite collection of types Θ, $|\Theta| \times |\Theta|$ matrix P

- $p_{\theta\theta'}$ is probability a link exists between types θ and θ'

The Stochastic Block Model

Finite collection of types Θ, $|\Theta| \times |\Theta|$ matrix P

- $p_{\theta\theta'}$ is probability a link exists between types θ and θ'

Sequence of graphs $G(n)$ with $n \to \infty$

- Fraction of type θ converges to $q_\theta > 0$
- Fraction of these that are stubborn converges to $s_\theta \in (0, 1)$
- Fraction of these with view 1 converges to $z_{S\theta}$

The Stochastic Block Model

Finite collection of types Θ, $|\Theta| \times |\Theta|$ matrix P

- $p_{\theta\theta'}$ is probability a link exists between types θ and θ'

Sequence of graphs $G(n)$ with $n \to \infty$

- Fraction of type θ converges to $q_\theta > 0$
- Fraction of these that are stubborn converges to $s_\theta \in (0, 1)$
- Fraction of these with view 1 converges to $z_{S\theta}$

Define

$$\tilde{d}_\theta = \sum_{\theta' \in \Theta} q_{\theta'} p_{\theta\theta'}, \quad \tilde{m} = \sum_{\theta \in \Theta} q_\theta \tilde{d}_\theta$$

Expected Views and Influence

Block matrix

$$H = \begin{bmatrix} I & 0 \\ H_{OS} & H_{OO} \end{bmatrix}$$

$$h_{\theta\theta'}^{(OS)} = \frac{s_{\theta'} q_{\theta'} p_{\theta\theta'}}{\tilde{m}}, \quad h_{\theta\theta'}^{(OO)} = \begin{cases} \frac{(1-s_{\theta'}) q_{\theta'} p_{\theta\theta'}}{\tilde{m}} & \text{if } \theta' \neq \theta \\ 1 - \frac{\tilde{d}_\theta}{\tilde{m}} + \frac{(1-s_\theta) q_\theta p_{\theta\theta}}{\tilde{m}} & \text{if } \theta' = \theta \end{cases}$$

Expected Views and Influence

Block matrix

$$H = \begin{bmatrix} I & 0 \\ H_{OS} & H_{OO} \end{bmatrix}$$

$$h^{(OS)}_{\theta\theta'} = \frac{s_{\theta'} q_{\theta'} p_{\theta\theta'}}{\tilde{m}}, \quad h^{(OO)}_{\theta\theta'} = \begin{cases} \frac{(1-s_{\theta'})q_{\theta'}p_{\theta\theta'}}{\tilde{m}} & \text{if } \theta' \neq \theta \\ 1 - \frac{\tilde{d}_\theta}{\tilde{m}} + \frac{(1-s_\theta)q_\theta p_{\theta\theta}}{\tilde{m}} & \text{if } \theta' = \theta \end{cases}$$

Theorem

Define $\mathbf{z}_O = (I - H_{OO})^{-1} H_{OS} \mathbf{z}_S$. *Given any* $\epsilon > 0$, *the probability that there exists* i *such that* $|x_i - z_{O\theta_i}| > \epsilon$ *converges to zero as* $n \to \infty$.

Expected Views and Influence

Expected view of type θ agents concentrates around $z_{O\theta}$

Expected Views and Influence

Expected view of type θ agents concentrates around $z_{O\theta}$

Entry $\theta\theta'$ of matrix $(I - H_{OO})^{-1} H_{OS}$ describes influence of stubborn type θ' agents on open type θ agents

Expected Views and Influence

Expected view of type θ agents concentrates around $z_{O\theta}$

Entry $\theta\theta'$ of matrix $(I - H_{OO})^{-1}H_{OS}$ describes influence of stubborn type θ' agents on open type θ agents

Obtain measure of individual influence through renormalization

- Multiply entry $\theta\theta'$ by $\frac{(1-s_\theta)q_\theta}{s_{\theta'}q_{\theta'}}$

Convergence Rates

Theorem

$$\lim_{n\to\infty} \|\tilde{G}_{OO}(n)\|^n = \|e^{H_{OO}-I}\|$$

almost surely.

Convergence Rates

Theorem

$$\lim_{n\to\infty} \|\tilde{G}_{OO}(n)\|^n = \|e^{H_{OO}-I}\|$$

almost surely.

Sharp result on convergence rate

Convergence Rates

Theorem

$$\lim_{n\to\infty} \|\tilde{G}_{OO}(n)\|^n = \|e^{H_{OO}-I}\|$$

almost surely.

Sharp result on convergence rate

Why does the exponential appear?

- All agents update at once vs. sequentially at random
- Somewhat slower mixing

Asymptotic Variance

Need a different matrix $\mathcal{H}_O$, dimension $|\Theta|^2 \times |\Theta|^2$, matrix $\mathcal{H}'_O$, dimension $|\Theta|^2 \times |\Theta|$

- Entries derived from those of H

Asymptotic Variance

Need a different matrix $\mathcal{H}_O$, dimension $|\Theta|^2 \times |\Theta|^2$, matrix $\mathcal{H}'_O$, dimension $|\Theta|^2 \times |\Theta|$

- Entries derived from those of H

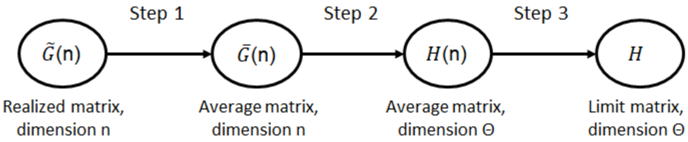
Define $q_\theta^{(O)} = (1 - s_\theta)q_\theta$, $r_{\theta\theta'}^{(O)} = q_\theta^{(O)} q_{\theta'}^{(O)}$

Theorem

$$\lim_{n\to\infty} n \,\mathrm{Var}[\omega] = \mathbf{q}^{(O)} \cdot (I - \Lambda_{\mathbf{z}_O})\mathbf{z}_O + \mathbf{r}^{(O)} \cdot (I - \mathcal{H}_O)^{-1} \mathcal{H}'_O (I - \Lambda_{\mathbf{z}_O})\mathbf{z}_O$$

almost surely.

A Proof Template



The Key Lemma

Suppose A is a $n \times n$ matrix with random entries that are independent with mean zero and magnitude uniformly bounded by 1

The Key Lemma

Suppose A is a $n \times n$ matrix with random entries that are independent with mean zero and magnitude uniformly bounded by 1

Lemma

There exist constants $C, c > 0$ such that

$$\mathbb{P}\left(\|A\| > k\sqrt{n}\right) \leq Ce^{-ckn}$$

for all k.

The Key Lemma

Suppose A is a $n \times n$ matrix with random entries that are independent with mean zero and magnitude uniformly bounded by 1

Lemma

There exist constants $C, c > 0$ such that

$$\mathbb{P}\left(\|A\| > k\sqrt{n}\right) \leq Ce^{-ckn}$$

for all k.

Proof is an application of Hoeffding's inequality, allows us to bound difference between $\tilde{G}$ and its expectation

The Islands Model

Two types, link to own type w.p. p_s, link to other type w.p. p_d, assume $p_s \geq p_d$

The Islands Model

Two types, link to own type w.p. p_s, link to other type w.p. p_d, assume $p_s \geq p_d$

Vary homophily by changing ratio p_s/p_d

The Islands Model

Two types, link to own type w.p. p_s, link to other type w.p. p_d, assume $p_s \geq p_d$

Vary homophily by changing ratio p_s/p_d

Vary polarization by varying $|z_{S1} - z_{S2}|$

The Islands Model

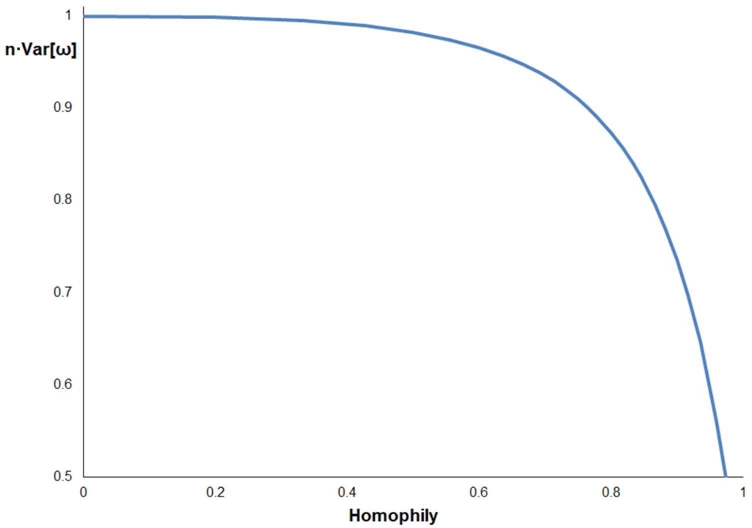
Two types, link to own type w.p. p_s, link to other type w.p. p_d, assume $p_s \geq p_d$

Vary homophily by changing ratio p_s/p_d

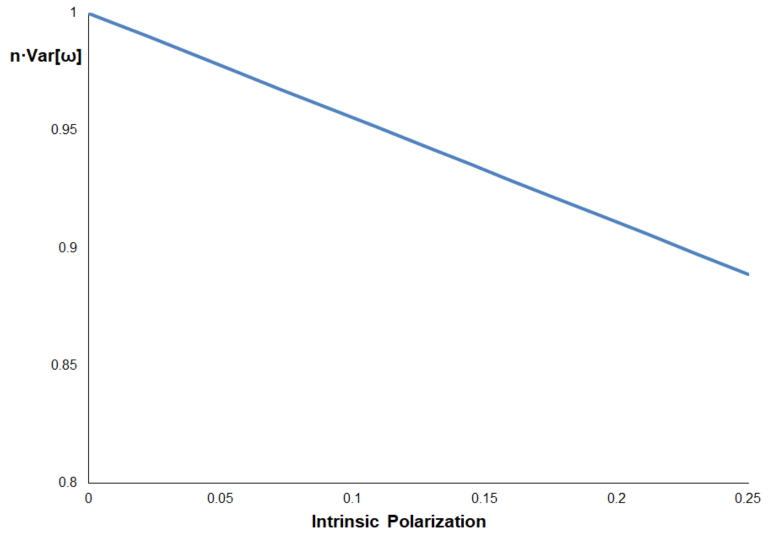
Vary polarization by varying $|z_{S1} - z_{S2}|$

Homophily and polarization work together to reduce variance

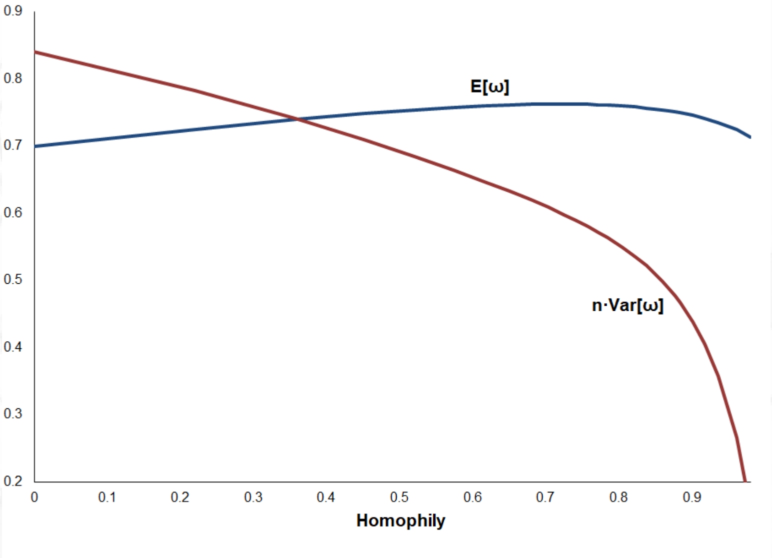
Homophily and Variance



Polarization and Variance



Different Group Sizes and Integration



A Framework to Study Propaganda

Suppose third party can broadcast propaganda to the agents

A Framework to Study Propaganda

Suppose third party can broadcast propaganda to the agents

Natural representation: add a type, fully stubborn

- Links with other types capture intensity of targeting

A Framework to Study Propaganda

Suppose third party can broadcast propaganda to the agents

Natural representation: add a type, fully stubborn

- Links with other types capture intensity of targeting

Augmented influence matrix

$$\hat{H} = \begin{bmatrix} 1 & \mathbf{0}^T & \mathbf{0}^T \\ \mathbf{0} & I & 0 \\ \boldsymbol{\alpha} & H_{OS}^{(\alpha)} & H_{OO}^{(\alpha)} \end{bmatrix}$$

$\boldsymbol{\alpha}$ tells us where propaganda is targeted, H_{OS} and H_{OO} rescaled so $\hat{H}$ is row stochastic

The Effect of Propaganda

Expected views:

$$\hat{\mathbf{z}}_O = \left(I - H^{(\alpha)}\right)^{-1} \begin{bmatrix} \boldsymbol{\alpha} & H_{OS}^{(\alpha)} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{z}_S \end{bmatrix}$$

The Effect of Propaganda

Expected views:

$$\hat{\mathbf{z}}_O = \left(I - H^{(\alpha)}\right)^{-1} \begin{bmatrix} \boldsymbol{\alpha} & H_{OS}^{(\alpha)} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{z}_S \end{bmatrix}$$

First column of

$$W^{(\alpha)} := \left(I - H^{(\alpha)}\right)^{-1} \begin{bmatrix} \boldsymbol{\alpha} & H_{OS}^{(\alpha)} \end{bmatrix}$$

captures influence of propaganda

Remarks

Tractable model of opinion dynamics

- Non-degenerate steady state distribution

Comparative statics on influence and variance of public opinion

Framework to evaluate propaganda

Further work:

- Competition for influence
- Model estimation