Learning Dynamics in Social Networks

Simon Board and Moritz Meyer-ter-Vehn

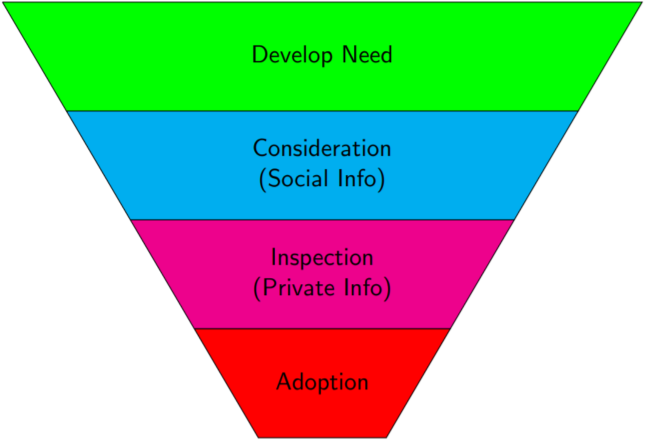
Department of Economics, UCLA

October 29, 2021

Introduction

- How do the entire societies learn about innovations?
  - consumers learning a new brand of electric car from friends
  - farmers learning about a novel crop from neighbors
  - entrepreneurs learning about a source of finance from nearby business
- Two sources of information:
  - social information acquired from neighbors
  - private information if inspect innovation
- How does diffusion depend on the network?
  - is diffusion faster in more interconnected societies?
  - is diffusion faster in more centralized societies?

The Social Purchasing Funnel



Model

- Players and products:
  - I players i on exogenous, directed network G.
  - Product quality $\theta \in \{L, H\}$, where the prior $\Pr(H) = \pi_0$.
- Timing:
  - Player i observes which of her neighbors N_i adopt product by t_i.
  - Player i can inspect product at iid cost $\kappa_i \sim F[\underline{\kappa}, \overline{\kappa}]$.
  - Player i adopts product iff inspected and $\theta = H$.
- Payoffs:
  - Player gets 1 if adopts high quality product net of inspection cost κ_i.
  - Player gets $-M$ if adopts low quality product net of inspection cost κ_i.
  - 0, if no adoption.

Examples: Directed Trees

Example (1. Directed Pair $i \to j$)

Suppose two agents, Iris and John. John has no social information, while Iris observes John.

- Let $x_{j,t}$ be the probability that John adopts product H by time t.
- The time-derivative $\dot{x}_{j,t}$ equals the probability he adopts conditional on entering at time t.
- Since he inspects iff $\kappa_j \leq \pi_0$ and then always adopts product H, we have

$$\dot{x}_j = \Pr(j \text{ adopt}) = \Pr(j \text{ inspect}) = F(\pi_0),$$

where we drop the time subscript.

Examples: Directed Trees

- Iris can learn from John's adoption.
- We can interpret John's adoption curve x_j as Iris's social learning curve.
- If John has adopted, Iris infers that quality is high and also adopts.
- If John has not adopted, Iris's posterior that quality is high is given by Bayes' rule

$$\pi\left(x_j\right) := \frac{\left(1 - x_j\right)\pi_0}{\left(1 - x_j\right)\pi_0 + \left(1 - \pi_0\right)}.$$

- Iris inspects if $\kappa_i \leq \pi(x_j)$. $\pi(x_j)$ is decreasing in x_j.
- Iris's adoption rate equals

$$\begin{aligned}\dot{x}_i &= 1 - \Pr(i \text{ not adopt }) \\ &= 1 - \Pr(j \text{ not adopt }) \times \Pr(i \text{ not inspect } \mid j \text{ not adopt }) \\ &= 1 - \left(1 - x_j\right)\left(1 - F\left(\pi\left(x_j\right)\right)\right) =: \Phi\left(x_j\right)\end{aligned}$$

Examples: Directed Chain

- Suppose there is an infinite chain of agents, so Kata observes Lili, who observes Moritz, so on.
- The adoption in the symmetric equilibrium in governed by ODE

$$\dot{x} = \Phi(x).$$

- This captures the idea that Kata's decision takes into account Lili's decision, which takes into account Moritz's decision, and so on.

General Networks

- Let x_i denotes agent i's probability of adopting product H by x_i.
- Let $x_{i,G,\xi}$ be agent i's realized adoption curve given (G, ξ) after taking expectation over others' entry time t_j and cost draws κ_j.
- Taking expectation over (G, ξ_{-i}), let

$$x_{i,\xi_i} := \sum_{G,\xi_{-i}} \mu(G, \xi_{-i} \mid \xi_i) \cdot x_{i,G,\xi}$$

be i's interim adoption curve given her signal ξ_i.

- Let $y_{i,G,\xi}$ be the probability that at least one of i's neighbor adopts product H by time $t \leq t_i$ in network G given signals ξ.
- Let

$$y_{i,\xi_i} := \sum_{G,\xi_{-i}} \mu(G, \xi_{-i} \mid \xi_i) \cdot y_{i,G,\xi}$$

be the expectation conditional on ξ_i.

General Networks

- To solve for i 's realized adoption curve $x_{i,G,\xi}$, consider two cases.
- If she sees one of her neighbors adopt, she updates her belief to $\pi = 1$ and adopts blindly.
- If she sees no adoption, she updates her belief to $\pi_i = \pi(y_{i,\xi_i}) \leq \pi_0$ and inspects iff her inspection cost is below this cutoff, that is $\kappa_i \leq c_{i,\xi_i} := \pi_i$.
- Hence, i's realized adoption curve follows

$$\dot{x}_{i,G,\xi} = 1 - \left(1 - y_{i,G,\xi}\right)\left(1 - F\left(\pi\left(y_{i,\xi_i}\right)\right)\right) =: \phi\left(y_{i,G,\xi}, y_{i,\xi_i}\right).$$

- Taking expectation over (G, ξ_{-i}), agent i's interim adoption curve is then

$$\dot{x}_{i,\xi_i} = 1 - \left(1 - y_{i,\xi_i}\right)\left(1 - F\left(\pi\left(y_{i,\xi_i}\right)\right)\right) = \phi\left(y_{i,\xi_i}, y_{i,\xi_i}\right) = \Phi\left(y_{i,\xi_i}\right).$$

General Networks

Assumption (BHR)

The distribution of costs has a bounded hazard rate (BHR) *if*

$$\frac{f(\kappa)}{1-F(\kappa)} \leq \frac{1}{\kappa(1-\kappa)} \quad \text{for } \kappa \in [0, \pi_0]$$

Lemma (1)

If F has a bounded hazard rate, then i's interim adoption probability x_{i,ξ_i} increases in her information y_{i,ξ_i}.

- Recall

$$\Phi(y) = 1 - \underbrace{(1-y)}_{\downarrow \text{ in } y} \cdot \underbrace{(1 - F(\pi(y)))}_{\uparrow \text{ in } y}.$$

$$\Phi'(y) = 1 - F(\pi(y)) - \pi(y) \cdot (1 - \pi(y)) \cdot f(\pi(y))$$

General Networks

Proposition (1)

In any random network $\mathcal{G}$ there exists a unique equilibrium.

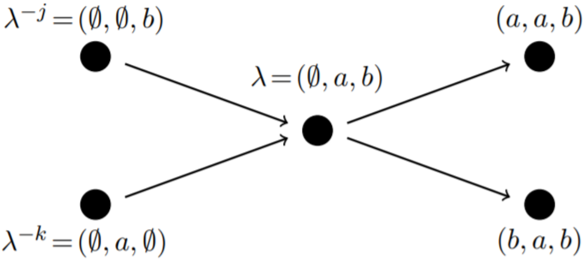
Idea of proof:

- State of network $\lambda \in \{\emptyset, a, b\}^I$
  - $\lambda_i = \emptyset$: i hasn't moved yet, $t \leq t_i$
  - $\lambda_i = a$: i has moved, tried, and adopted the product.
  - $\lambda_i = b$: i has moved, but not adopted the product.
- Agent i's knowledge in state λ

$$\Lambda(i, \lambda) := \left\{\lambda' : \lambda'_i = \lambda_i, \lambda_j = a \text{ iff } \lambda'_j = a \text{ for all } j \in N_i\right\}$$

- Additional notation
  - Distribution $z = \left(z_\lambda^\theta\right)$, and $z_\Lambda^\theta := \sum_{\lambda \in \Lambda} z_\lambda^\theta$ for sets Λ
  - For λ with $\lambda_i = a, b$, write λ^{-i} for "same state with $\lambda_i = \emptyset$ ".

State Transition



ODE for General Networks

$$
\begin{aligned}
(1-t)\dot{z}_{\lambda,G,\xi} = & -\underbrace{\sum_{i:\lambda_i=\emptyset} z_{\lambda,G,\xi}}_{\text{out-flow}} \\
& + \underbrace{\sum_{\substack{i:\lambda_i=a,\\ \exists j\in N_i(G):\lambda_j=a}} z_{\lambda-i}, G, \xi}_{\text{in-flows, if } \lambda_i=a \text{ and neighbors adopt}} \\
& + \underbrace{\sum_{\substack{i:\lambda_i=a,\\ \forall j\in N_i(G):\lambda_j\neq a}} z_{\lambda^{-i},G,\xi} F\left(\pi\left(y_{i,\xi_i}\right)\right)}_{\text{in-flows, if } \lambda_i=a \text{ and no neighbor adopt}} \\
& + \underbrace{\sum_{\substack{i:\lambda_i=b,\\ \forall j\in N_i(G):\lambda_j\neq a}} z_{\lambda^{-i},G,\xi}\left(1-F\left(\pi\left(y_{i,\xi_i}\right)\right)\right)}_{\text{in-flows, if entered and not adopted}}
\end{aligned}
$$

ODE for General Networks

- The existence and uniqueness reduce to check the existence and uniqueness of solution to ODE.
- The existence of a unique equilibrium follows from the Picard-Lindelof theorem since the boundedness of f implies the system is Lipschitz.

Examples: Undirected Networks

Example (3. Undirected Pair $i \leftrightarrow j$)

Agent i 's social learning curve equals i 's expectation of j 's adoption curve at $t \leq t_i$; for convenience we denote this by $\bar{x}_j$.

- Two probability assessment of the event that j observes i adopt.
- From i's objective perspective, this probability equals 0 (i knows she has not entered at $t < t_i$).
- From j's subjective perspective, this probability equals $\bar{x}_i$ (j thinks he is learning from agent i given $t \leq t_j$).

$$\dot{\bar{x}}_j = \phi\left(0, \bar{x}_i\right) = F\left(\pi\left(\bar{x}_i\right)\right).$$

- By symmetry, $\bar{x}_i = \bar{x}_j =: \bar{x}$, we can reduce it to a one-dimensional ODE. The actual (unconditional) adoption probability follows $\dot{x} = \Phi(\bar{x})$.

Examples: Complete Networks

Example (4. Complete Networks)

Consider the complete network of $I+1$ agents. When $I=1$, it reduces to the undirected pair (Example 3). With more agents, agent j's adoption before i enters depends on agent k's adoption before both i and j enter.

- We can think about the game from the first mover's perspective, before anyone else has adopted.
- Let the first adopter probability $\hat{x}$ be the probability an agent adopts given that no one else has yet adopted.
- Since everyone is symmetric, intuition suggests that the first adopter attaches subjective probability $(1-\hat{x})^I$ to the event that none of the other potential first adopters has adopted.
- The first adopter observes no adoption herself, we define $\hat{x}$ as the solution of

$$\dot{\hat{x}} = \phi\left(0, 1-(1-\hat{x})^I\right) = F\left(\pi\left(1-(1-\hat{x})^I\right)\right).$$

Examples: Complete Networks

Lemma (2)

In the complete network with $I+1$ agents, any agent's social learning curve is $1-(1-\hat{x})^I$; the adoption probability follows $\dot{x} = \Phi\left(1-(1-\hat{x})^I\right)$.

Directed Networks with Multiple Types

- Let us generate the random network $\mathcal{G}_I$ via the configuration model.
- For any agent i, independently draw a finite type $\theta \in \Theta$.
- For any agent with type θ, independently draw a vector of labeled outlinks $d = (d_{\theta'})_{\theta'} \in \mathbb{N}^{\Theta}$, random vector $D_\theta = \left(D_{\theta,\theta'}\right)_{\theta'}$.
- The agent i's signal ξ_i consists of her degree $d \in \mathbb{N}^{\Theta}$ after the pruning.
- Let $x_d, y_d, c_d = \pi(y_d)$ as the adoption probability, learning curves and cost thresholds of a degree-d agent.
- Taking expectation of the degree a type θ agent, we write it as $x_\theta = E[x_{D_\theta}]$.

Directed Networks with Multiple Types

- For large I, the random network locally resembles a tree where the adoption probabilities of an agent's neighbors are approximately independent.
- The probability that an agent with degree $d = (d_{\theta'})$ observes an adoption is then approximated by

$$y_d \approx 1 - \prod_{\theta'} (1 - x_{\theta'})^{d_{\theta'}} .$$

- Substituting this approximation into ODE, we define (x_θ^*) to be the solution of

$$\dot{x}_\theta = E\left[\Phi\left(1 - \prod_{\theta'} (1 - x_{\theta'})^{D_{\theta,\theta'}}\right)\right]$$

Directed Networks with Multiple Types

Definition (ϵ_I equilibrium)

We say that a vector of cutoff costs (c_d) is a limit equilibrium of the large directed random network with degree distribution D if it is an ϵ_I-equilibrium in $\mathcal{G}_I$ for some sequence (ϵ_I) with $\lim_{I\to\infty} \epsilon_I = 0$.

Proposition (2.)

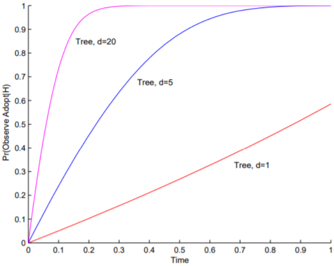
The cutoffs (c_d^*) are the unique limit equilibrium of the large directed random network with degree distribution D.

Directed Networks with Multiple Types

Theorem (1.)

Theorem 1. Assume F has a bounded hazard rate. Social learning and welfare improve with links: If $\tilde{D} \succeq_{FOSD} D$,

1. *For any degree $d, \tilde{y}^*_d \geq y^*_d$.*
2. *For any type $\theta, E\left[\tilde{y}^*_{\tilde{D}_\theta}\right] \geq E\left[y^*_{D_\theta}\right]$.*



Regular tree with d neighbors, $c \sim U[0,1]$, $\pi_0 = 1/2$.

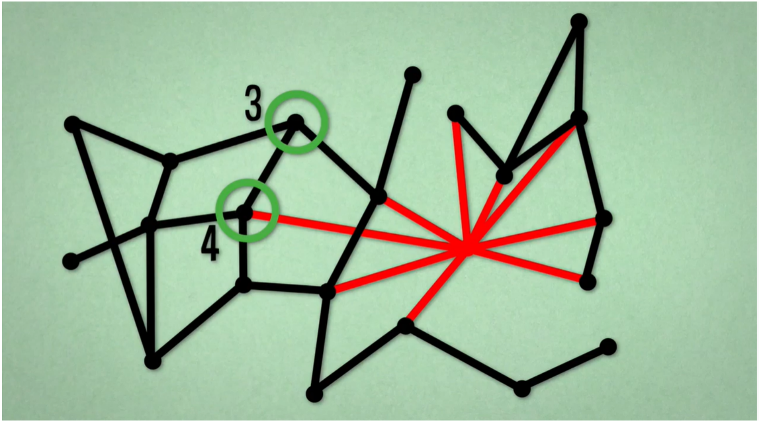
Undirected Networks

1. Consider a single type undirected configuration model.
2. Each agent independently draws $d \in \mathbb{N}$ link-stubs generated by a random variable D.
3. An important feature of random undirected networks is the friendship paradox. Namely, i's neighbors typically have more neighbors than i herself.
4. Formally, we define the neighbor's degree distribution D' by

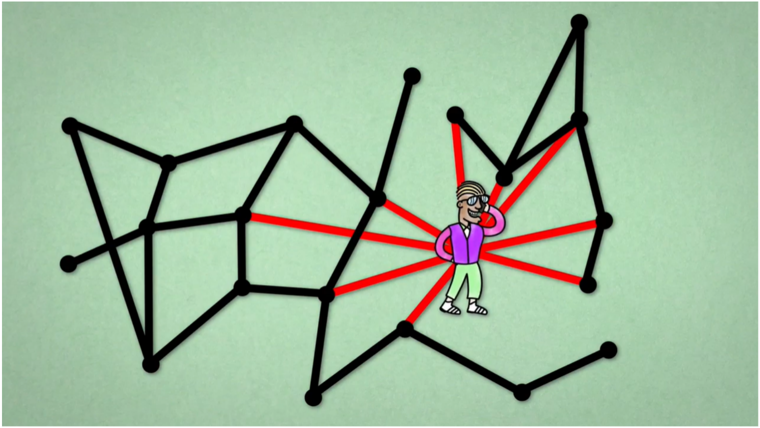
$$\Pr\left(D' = d\right) := \frac{d}{E[D]} \Pr(D = d).$$

5. For example, in an Erdős-Rényi network, D is Poisson and $D' = D + 1$, whereas in a regular network, $D' = D \equiv d$.

Friendship Paradox



Friendship Paradox



Undirected Networks

1. We write $\bar{x}$ for the probability that i 's neighbor j has adopted at $t \leq t_i$.
2. With general degree distribution D, neighbor j additionally learns from another $D' - 1$ independent links, from which he observes no adoption with probability $(1 - \bar{x})^{(D'-1)}$.
3. Agent i expects j to observe an adoption with objective probability $1 - (1 - \bar{x})^{(D'-1)}$, while j expects to observe an adoption with the higher, subjective probability $1 - (1 - \bar{x})^{D'}$.
4. So motivated, define $\bar{x}^*$ as the solution of

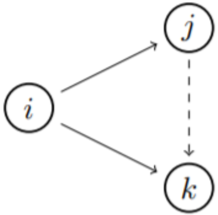
$$\dot{\bar{x}} = E\left[\phi\left(1 - (1 - \bar{x})^{D'-1}, 1 - (1 - \bar{x})^{D'}\right)\right].$$

5. Agent i 's actual, unconditional adoption rate then equals

$$E\left[\Phi\left(1 - (1 - \bar{x}^*)^{D}\right)\right].$$

Are All Links Beneficial?

1. Indirect links induce neighbors to inspect.
2. Learn from neighbors' inspections and adoptions.
3. How about correlating and backward links?



Adding a Backward Link

1. With backward link (undirected $i \leftrightarrow j$)

$$\dot{\bar{x}}_j = F(\pi(\bar{x}_i)) \Rightarrow \dot{\bar{x}}_j \leq F(\pi_0) \Rightarrow \bar{x}_j \leq F(\pi_0)\, t.$$

2. Without backward link (directed $i \to j$)

$$\dot{x}_j = F(\pi_0) \Rightarrow F(\pi_0)\, t.$$

3. Clearly, we have $\bar{x}_j \leq x_j$. $\bar{x}_j \leq F(\pi_0)\, t$.
4. Thus, the link $j \to i$ lowers i 's social information and her utility.

Welfare Implication

1. We compare a network where agents have D directed links to one with D undirected links.
2. We write $y_d^* = 1 - (1 - x^*)^d$ and $\bar{y}_d = 1 - (1 - \bar{x}^*)^d$ for the respective social learning curves.

Theorem (2)

Assume $D' - 1 \preceq D$ in the FOSD-order. Social learning and welfare are higher when the network is directed rather than undirected: For any degree $d, y_d^* > \bar{y}_d^*$.

1. Intuitively, fixing the degree distribution, directed networks generate better information than undirected networks.

Clustering

1. To capture clustering, we consider the following variant of the configuration model.
2. Each agent independently draws D pairs of link-stubs, which are then randomly connected to two other pairs of link-stubs to form a triangle.
3. In Example 4, where $D = 1$, $I = 2$, the learning curve is determined by the adoption probability $\hat{x}$ of the first adopter,

$$\dot{\hat{x}} = \phi\left(0, 1 - (1 - \hat{x})^2\right).$$

4. For general D, agent i 's neighbors additionally learn from another $D' - 1$ independent triangles, from which they observe no adoption with probability $(1 - \hat{x})^{2\left(D'-1\right)}$.
5. Define $\hat{x}^*$ as the solution of

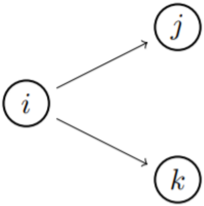
$$\dot{\hat{x}} = E\left[\phi\left(1 - (1 - \hat{x})^{2\left(D'-1\right)}, 1 - (1 - \hat{x})^{2D'}\right)\right].$$

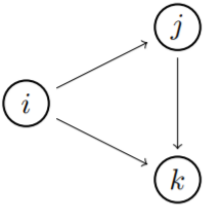
6. Agent i 's actual, unconditional adoption rate then equals

$$E\left[\Phi\left(1 - (1 - \hat{x})^{2D}\right)\right].$$

Correlating Link

1. Assume agent i initially observes two uninformed agents j and k.
2. The probability that neither adopts is $(1 - F(\pi_0)\, t)^2$.
3. Now, suppose we add a link from j to k.
4. Agent k's behavior is unchanged (social information does not change).
5. But the probability that agent i sees an adoption decreases.
6. This is because the probability $x_{j|\neg k}$ that j adopts conditional on k not adopting follows $\dot{x}_{j|\neg k} = F(\pi(x_k)) < F(\pi_0)$.





Clustering

1. To address the overall welfare effect, we compare an undirected network with D pairs of link-stubs to one with $2D$ bilateral link-stubs, as illustrated in Figure 5 .
2. The social learning curve equals $\hat{y}^*_{2d} = 1 - (1 - \hat{x}^*)^{2d}$ in the former network, and $\bar{y}^*_{2d} = 1 - (1 - \bar{x}^*)^{2d}$ in the latter, where $\hat{x}^*$ and $\bar{x}^*$ solve

$$\dot{\hat{x}} = E\left[\phi\left(1 - (1 - \hat{x})^{2(D'-1)}, 1 - (1 - \hat{x})^{2D'}\right)\right]$$

$$\dot{\bar{x}} = E\left[\phi\left(1 - (1 - \bar{x})^{D'-1}, 1 - (1 - \bar{x})^{D'}\right)\right]$$

Theorem (3.)

Clustering reduces social learning and welfare: For any degree d, $\hat{y}^*_{2d} < \bar{y}^*_{2d}$

Correlation Neglect

1. To model correlation neglect, we consider a configuration model where agents draw D pairs of undirected triangular stubs, but agents believe all their information is independent.
2. That is, i believes that her neighbors are not connected, believes her neighbors think their neighbors are not connected, and so on.
3. Consider the limit as I grows large. Since agent i believes that links are generated bilaterally, her subjective probability assessment that any of her neighbors has adopted, $\bar{x}^*$, solves

$$\dot{\bar{x}} = E\left[\phi\left(1-(1-\bar{x})^{2D'-1}, 1-(1-\bar{x})^{2D'}\right)\right].$$

4. An agent with $2d$ links thus uses cutoff $\pi\left(1-(1-\bar{x}^*)^{2d}\right)$ when choosing whether to inspect

Correlation Neglect

1. In reality, agent i 's neighbors form triangles (i, j, k), and so the objective probability $\check{x}^*$ that the first adopter in a triangle adopts follows a variant of the usual first adopter triangle formula

$$\dot{\check{x}} = E\left[\phi\left(1-(1-\check{x})^{2(D'-1)}, 1-(1-\bar{x}^*)^{2D'}\right)\right]$$

2. Intuitively, the first adopter in (i, j, k) expects to see an adoption with probability $\bar{y}^*_{2d} = 1-(1-\bar{x}^*)^{2D'}$, but the objective adoption probability is $\check{y}^*_{2d} = 1-(1-\check{x}^*)^{2(D'-1)}$.

Theorem (4.)

Correlation neglect reduces social learning: For any degree d, $\check{y}^*_{2d} < \hat{y}^*_{2d}$.

General Undirected Networks

1. Let us consider random networks that encompass the undirected links and cliques.
2. To define these networks $\hat{\mathcal{G}}_I$, suppose every agent independently draws $\bar{D}$ bilateral stubs and $\hat{D}$ pairs of triangle stubs with finite expectations.
3. We connect pairs of bilateral stubs and triples of triangular stubs at random, and then prune self-links.

General Undirected Networks

1. Let $\bar{x}$ is the probability that i 's bilateral neighbor j adopts before t_i.
2. Let $\hat{x}$ is the probability that the first adopter j in one of i 's triangles adopts before t_i.
3. Define $(\bar{x}^*, \hat{x}^*)$ as the solution to the two-dimensional ODE

$$\dot{\bar{x}} = E\left[\phi\left(1-(1-\bar{x})^{\bar{D}'-1}(1-\hat{x})^{2\hat{D}}, 1-(1-\bar{x})^{\bar{D}'}(1-\hat{x})^{2\hat{D}}\right)\right]$$
$$\dot{\hat{x}} = E\left[\phi\left(1-(1-\bar{x})^{\bar{D}}(1-\hat{x})^{2\left(\hat{D}'-1\right)}, 1-(1-\bar{x})^{\bar{D}}(1-\hat{x})^{2\hat{D}'}\right)\right].$$

Proposition (2')

The cutoffs $\left(c^*_{\bar{d},\hat{d}}\right)$ are the unique limit equilibrium of $\hat{\mathcal{G}}_I$.

Conclusion

- Social learning plays a crucial role in the diffusion of new products, financial innovations, and new production techniques.
- This paper proposes a tractable model of social learning on large random networks, characterizes equilibrium in terms of simple differential equations, and studies the effect of network structure on learning dynamics.
- They showed that clustering slows learning by correlating neighbors' adoption decisions.
- Future work: pricing, advertising and seeding the network.

Thanks!