Network Formation Models for Dynamic Network Data

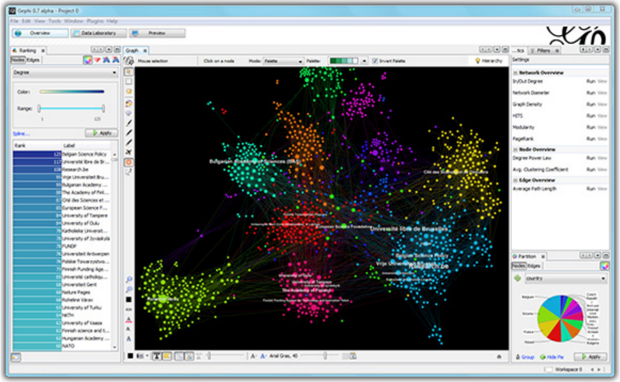
Chih-Sheng Hsieh

Department of Economics
National Taiwan University

April 15, 2024

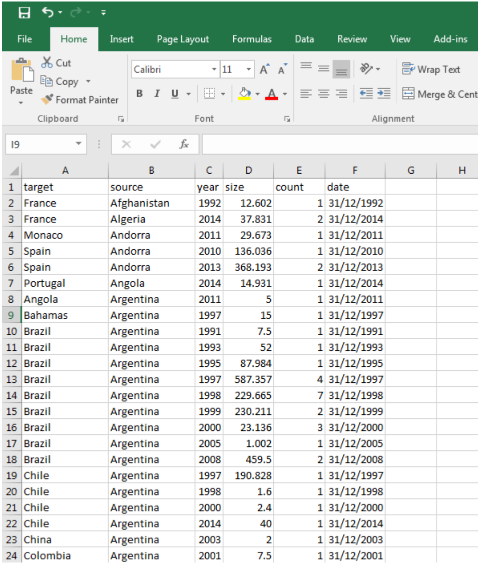
Visualizing Network Evolution by Gephi

- Gephi https://gephi.org/ (need Java 7 or 8 during installation: also pay attention to whether your OS is 32-bit or 64-bit).
- tutorial: https://gephi.org/users/tutorial-visualization/



Visualizing Network Evolution by Gephi

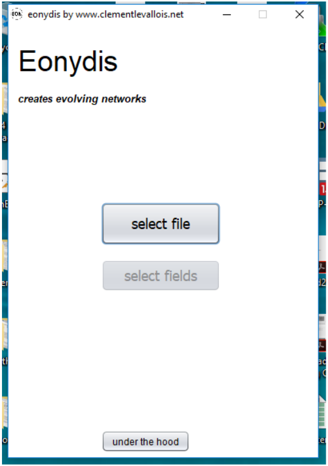
- We use the cross-border Merger & acquisition (M&A) data from 1984 to 2014 as an illustrating example.
- We are interested in observing the evolving pattern in the cross border M&A network. See which countries are central and the growing importance of certain countries such as China, India, Brazil, Russia, etc.
- Potentially the similar idea can be applied to internal trade data, international migration data, or smuggle data.

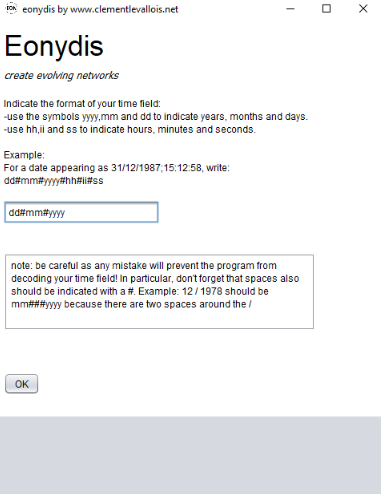


	A	B	C	D	E	F	G	H
1	target	source	year	size	count	date		
2	France	Afghanistan	1992	12.602	1	31/12/1992		
3	France	Algeria	2014	37.831	2	31/12/2014		
4	Monaco	Andorra	2011	29.673	1	31/12/2011		
5	Spain	Andorra	2010	136.036	1	31/12/2010		
6	Spain	Andorra	2013	368.193	2	31/12/2013		
7	Portugal	Angola	2014	14.931	1	31/12/2014		
8	Angola	Argentina	2011	5	1	31/12/2011		
9	Bahamas	Argentina	1997	15	1	31/12/1997		
10	Brazil	Argentina	1991	7.5	1	31/12/1991		
11	Brazil	Argentina	1993	52	1	31/12/1993		
12	Brazil	Argentina	1995	87.984	1	31/12/1995		
13	Brazil	Argentina	1997	587.357	4	31/12/1997		
14	Brazil	Argentina	1998	229.665	7	31/12/1998		
15	Brazil	Argentina	1999	230.211	2	31/12/1999		
16	Brazil	Argentina	2000	23.136	3	31/12/2000		
17	Brazil	Argentina	2005	1.002	1	31/12/2005		
18	Brazil	Argentina	2008	459.5	2	31/12/2008		
19	Chile	Argentina	1997	190.828	1	31/12/1997		
20	Chile	Argentina	1998	1.6	1	31/12/1998		
21	Chile	Argentina	2000	2.4	1	31/12/2000		
22	Chile	Argentina	2014	40	1	31/12/2014		
23	China	Argentina	2003	2	1	31/12/2003		
24	Colombia	Argentina	2001	7.5	1	31/12/2001		

Visualizing Network Evolution by Gephi

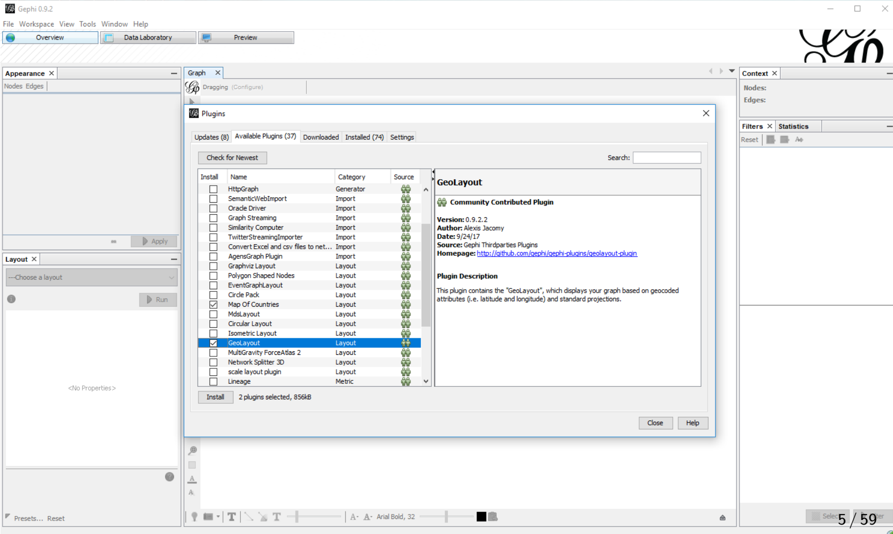
- First step: transforming the excel file into GEXF file by eonydis
- Be sure to set the time format as dd#mm#yyyy





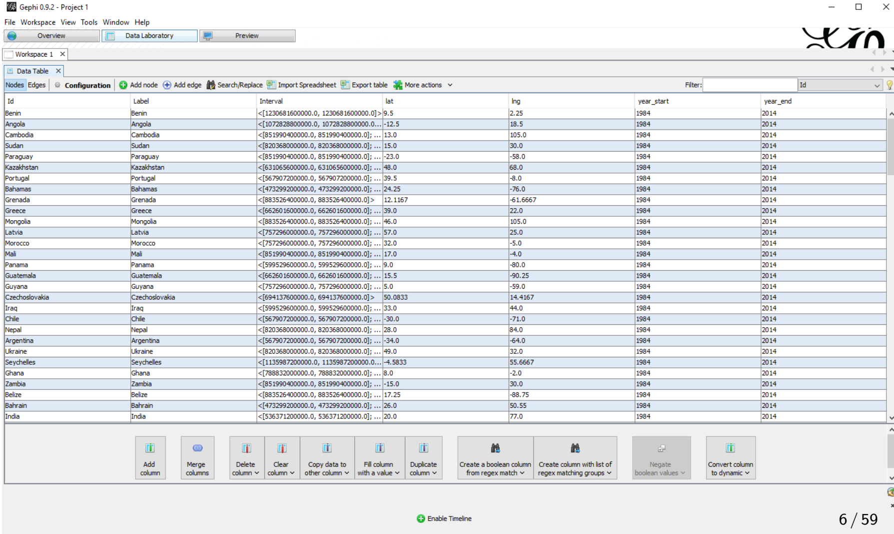
Visualizing Network Evolution by Gephi

- Second step: update Gephi and add two plug-ins: Map of countries, GeoLayout



Visualizing Network Evolution by Gephi

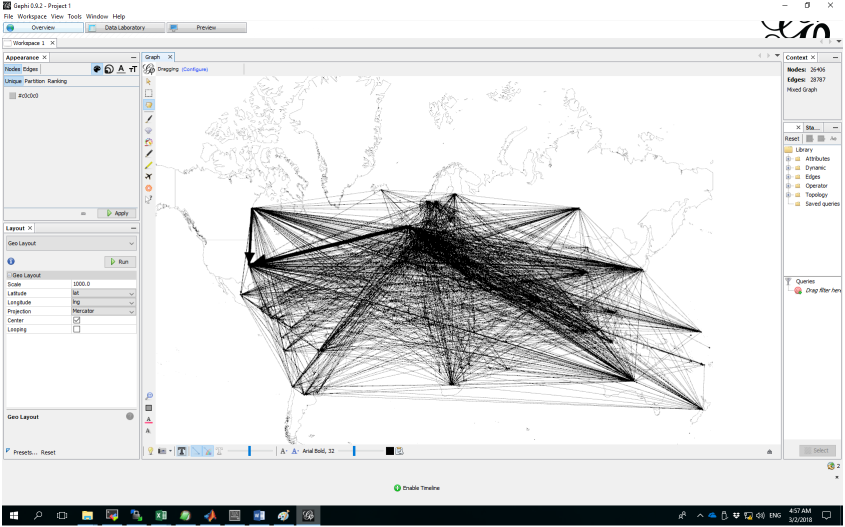
- Third step: load the GEXF file into Gephi and add latitude and longitude information of each country from a separated excel file



Id	Label	Interval	lat	lng	year_start	year_end
Benin	Benin	<[1230681600000.0, 1230681600000.0]>	9.5	2.25	1984	2014
Angola	Angola	<[1072828800000.0, 1072828800000.0…	-12.5	18.5	1984	2014
Cambodia	Cambodia	<[851990400000.0, 851990400000.0]; …	13.0	105.0	1984	2014
Sudan	Sudan	<[820368000000.0, 820368000000.0]; …	15.0	30.0	1984	2014
Paraguay	Paraguay	<[851990400000.0, 851990400000.0]; …	-23.0	-58.0	1984	2014
Kazakhstan	Kazakhstan	<[631065600000.0, 631065600000.0]; …	48.0	68.0	1984	2014
Portugal	Portugal	<[567907200000.0, 567907200000.0]; …	39.5	-8.0	1984	2014
Bahamas	Bahamas	<[473299200000.0, 473299200000.0]; …	24.25	-76.0	1984	2014
Grenada	Grenada	<[883526400000.0, 883526400000.0]>	12.1167	-61.6667	1984	2014
Greece	Greece	<[662601600000.0, 662601600000.0]; …	39.0	22.0	1984	2014
Mongolia	Mongolia	<[883526400000.0, 883526400000.0]; …	46.0	105.0	1984	2014
Latvia	Latvia	<[757296000000.0, 757296000000.0]; …	57.0	25.0	1984	2014
Morocco	Morocco	<[757296000000.0, 757296000000.0]; …	32.0	-5.0	1984	2014
Mali	Mali	<[851990400000.0, 851990400000.0]; …	17.0	-4.0	1984	2014
Panama	Panama	<[599529600000.0, 599529600000.0]; …	9.0	-80.0	1984	2014
Guatemala	Guatemala	<[662601600000.0, 662601600000.0]; …	15.5	-90.25	1984	2014
Guyana	Guyana	<[757296000000.0, 757296000000.0]; …	5.0	-59.0	1984	2014
Czechoslovakia	Czechoslovakia	<[694137600000.0, 694137600000.0]>	50.0833	14.4167	1984	2014
Iraq	Iraq	<[599529600000.0, 599529600000.0]; …	33.0	44.0	1984	2014
Chile	Chile	<[567907200000.0, 567907200000.0]; …	-30.0	-71.0	1984	2014
Nepal	Nepal	<[820368000000.0, 820368000000.0]; …	28.0	84.0	1984	2014
Argentina	Argentina	<[567907200000.0, 567907200000.0]; …	-34.0	-64.0	1984	2014
Ukraine	Ukraine	<[820368000000.0, 820368000000.0]; …	49.0	32.0	1984	2014
Seychelles	Seychelles	<[1135987200000.0, 1135987200000.0…	-4.5833	55.6667	1984	2014
Ghana	Ghana	<[788832000000.0, 788832000000.0]; …	8.0	-2.0	1984	2014
Zambia	Zambia	<[851990400000.0, 851990400000.0]; …	-15.0	30.0	1984	2014
Belize	Belize	<[883526400000.0, 883526400000.0]; …	17.25	-88.75	1984	2014
Bahrain	Bahrain	<[473299200000.0, 473299200000.0]; …	26.0	50.55	1984	2014
India	India	<[536371200000.0, 536371200000.0]; …	20.0	77.0	1984	2014

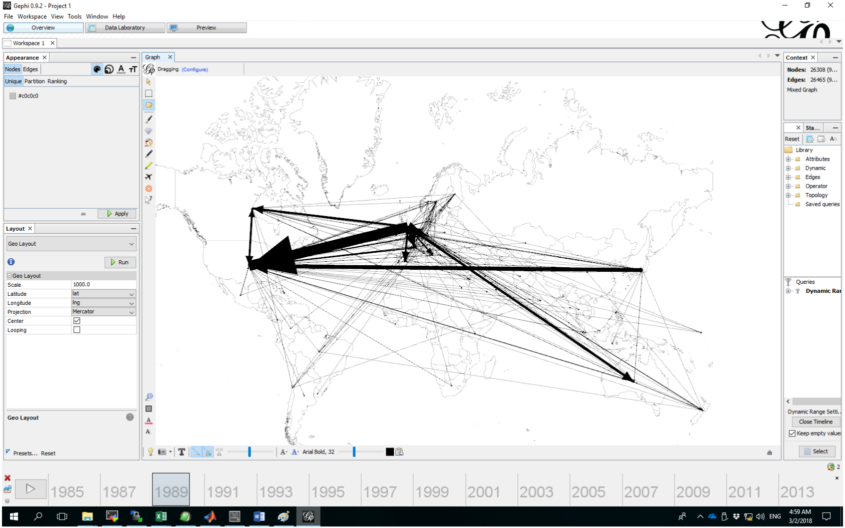
Visualizing Network Evolution by Gephi

- Fourth step: choose Layout to show the network



Visualizing Network Evolution by Gephi

- Fifth step: enable time line



Spatial dynamic panel data (SDPD) model

Lee and Yu (2010) specify the SDPD model as follows:

$$\begin{aligned} Y_{gt} =& \lambda \overline{W}_{gt} Y_{gt} + \rho Y_{g,t-1} + \mu \overline{W}_{g,t-1} Y_{g,t-1} + X_{gt}\beta_1 + \overline{W}_{gt} X_{gt}\beta_2 \\ & + \tau_g + l_g \alpha_{gt} + U_{gt}, \quad g = 1, 2, \cdots, G, \ t = 1, 2, \cdots, T. \end{aligned}$$

- λ captures the contemporary peer effect.
- ρ captures the time persistence effect on outcome.
- μ captures the temporal peer effect.
- $\tau_g = (\tau_1, \cdots, \tau_{n_g})'$ is the $n_g \times 1$ individual effect
- α_{gt} is the group-time effect for group g at time t
- When W_{gt} is endogenously formed, estimates of λ, μ and other coefficients will be biased.

A general dynamic network formation model

$$
P(W_{gt}|Z_{gt}, M_{gt}, W_{g,t-1}, Y_{g,t-1}, \Gamma) = \prod_{i=1}^{n_g}\prod_{j\neq i}^{n_g}\left(\frac{\exp(\psi_{ijgt})}{1+\exp(\psi_{ijgt})}\right)^{I(w_{ijgt}=1)}\left(\frac{1}{1+\exp(\psi_{ijgt})}\right)^{I(w_{ijgt}=0)},
$$

where

$$
\begin{aligned}
\psi_{ijgt} = {} & \underbrace{\gamma_0 + c_{igt}'\gamma_1 + c_{jgt}'\gamma_2 + c_{ijgt}'\gamma_3}_{\text{direct and homophily effects from obs. chara.}} + \underbrace{\gamma_4 w_{ijg,t-1} + \gamma_5 r_{ijg,t-1}}_{\text{persistence and transitivity}} \\
& + \underbrace{\gamma_6 y_{ig,t-1} + \gamma_7 y_{jg,t-1} + \gamma_8 |y_{ig,t-1} - y_{jg,t-1}|}_{\text{direct and homophily effects from activity outcomes}} + \underbrace{\sum_{p_1=1}^{\bar{p}_1} \delta_{p_1} |z_{ip_1g,t} - z_{jp_1g,t}|}_{\text{unobserved homophily}} \\
& + \underbrace{\sum_{p_2=1}^{\bar{p}_2} \xi_{p_2} m_{ip_2gt} + \sum_{p_2=1}^{\bar{p}_2} \zeta_{p_2} m_{jp_2gt}}_{\text{unobserved degree heterogeneity}}.
\end{aligned}
$$

Combining the SDPD model and the dynamic network formation model

- The SDPD model and the dynamic network formation model are connected by specifying the correlation structure between U_{gt} and Z_{gt} and M_{gt}:

$$\begin{aligned} U_{gt} &= Z_{nt}\kappa_1 + M_{gt}\kappa_2 + V_{gt}, \\ &= H_{gt}\kappa + \ V_{gt} \sim \mathcal{N}(0, \sigma_v^2 I_{n_g}). \end{aligned}$$

- The Selection-Corrected (SC)-SDPD model is

$$\begin{aligned} Y_{gt} =& \lambda \overline{W}_{gt} Y_{gt} + \rho Y_{g,t-1} + \mu \overline{W}_{g,t-1} Y_{g,t-1} + X_{gt}\beta_1 + \overline{W}_{gt} X_{gt}\beta_2 \\ &+ H_{gt}\kappa + \tau_g + l_g \alpha_{gt} + V_{gt}. \end{aligned}$$

- $H_{gt}\kappa$: linear control function for the endogenous W_{gt}, or time-varying individual effects in the SDPD model.

The likelihood function

- Let $\theta = (\lambda, \rho, \mu, \beta', \kappa', \sigma_v^2)'$. The likelihood function of Y_{gt}, conditional on $(W_{gt}, Y_{g,t-1}, W_{g,t-1}, H_{gt})$, is,

$$P(Y_{gt}|W_{gt}, Y_{g,t-1}, W_{g,t-1}, H_{gt}, \theta, \tau_g, \alpha_{gt})$$
$$= (2\pi)^{-\frac{n_g}{2}} \cdot (\sigma_v^2)^{-\frac{n_g}{2}} \cdot |S_{gt}(\lambda)| \cdot \exp\left(-\frac{V'_{gt}V_{gt}}{2\sigma_v^2}\right),$$

for $t = 1, 2, \cdots, T$ and $g = 1, 2, \cdots, G$.

- Assume the initial period Y_{g0} and W_{g0} are exogenously given, the joint likelihood function of $\{Y_{gt}\}$ and $\{W_{gt}\}$, conditional upon $\{H_{gt}\}$, can be written as

$$P(\{Y_{gt}\}, \{W_{gt}\}|\{H_{gt}\}, \{\alpha_{gt}\}, \tau_g, \theta, \Gamma) =$$
$$\prod_{g=1}^{G}\prod_{t=1}^{T} P(Y_{gt}|W_{gt}, Y_{g,t-1}, W_{g,t-1}, H_{gt}, \theta, \tau_g, \alpha_{gt})P(W_{gt}|H_{gt}, Y_{g,t-1}, \Gamma).$$

Bayesian estimation – the prior and posterior distributions

- We impose the following priors $\pi(\cdot)$ for θ, Γ, τ_g and $\{\alpha_{gt}\}$:

$$\gamma \sim \mathcal{N}_{6+2\bar{l}_1+\bar{l}_2}(\gamma_o, G_o),\ \ \Phi = (\delta', \xi', \zeta')' \sim \mathcal{N}_{\bar{p}}(\Phi_0, P_0),\ \ \lambda \sim U(-1, 1)$$
$$\rho|\lambda \sim U(-1 + |\lambda|, 1 - |\lambda|), \mu|\rho, \lambda \sim U(-1 + |\lambda| + |\rho|, 1 - |\lambda| - |\rho|),$$
$$\beta \sim \mathcal{N}_{2k}(\beta_0, B_0),\ \ \kappa \sim \mathcal{N}_{\bar{p}}(\kappa_0, K_0),\ \ \sigma_v^2 \sim \mathcal{IG}\left(\frac{a}{2}, \frac{b}{2}\right).$$
$$\tau_g \sim \mathcal{N}_{n_g}(\tau_0, E_0), \alpha_{gt} \sim \mathcal{N}(\alpha_0, F_0),\ \ g = 1, 2, \cdots, G,\ \ t = 1, 2, \cdots, T.$$

- By Bayes' theorem, the posterior distribution of parameters is,

$$P(\theta, \Gamma, \{\tau_g\}, \{\alpha_{gt}\}, \{H_{gt}\}|\{Y_{gt}\}, \{W_{gt}\}) \propto \pi(\theta) \times \pi(\Gamma) \times \pi(\{\tau_g\})$$
$$\times \pi(\{\alpha_{gt}\}) \times P(\{H_{gt}\}|\Gamma) \times P(\{Y_{gt}\}, \{W_{gt}\}|\{H_{gt}\}, \{\alpha_{gt}\}, \tau_g, \theta, \Gamma).$$

An overview of the MCMC algorithm

- Step 1: sample h_{ig1} from $P(h_{ig1}|Y_{g1}, W_{g1}, h_{-i,g1}, h_{ig2}, \Gamma, \Psi, \beta, \kappa, \sigma_v^2, \tau_g)$
- Step 2: sample h_{igt} from $P(h_{igt}|Y_{gt}, W_{gt}, h_{ig,t+1}, h_{-i,gt}, h_{ig,t-1}, \Gamma, \Psi, \beta, \kappa, \sigma_v^2, \tau_g, \alpha_{gt})$
- Step 3: sample h_{igT} from $P(h_{igT}|Y_{gT}, W_{gT}, h_{ig,T-1}, \Gamma, \Psi, \beta, \kappa, \sigma_v^2, \tau_g, \alpha_{gT})$
- Step 4: sample Γ from $P(\Gamma|\{W_{gt}\}, \{H_{gt}\})$
- Step 5: sample κ from $P(\kappa|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \Psi, \beta, \{\tau_g\}, \{\alpha_{gt}\}, \sigma_v^2)$
- Step 6: sample $\Psi = (\lambda, \rho, \mu)$ from $P(\Psi|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \beta, \kappa, \sigma_v^2, \{\tau_g\}, \{\alpha_{gt}\})$.
- Step 7: sample β from $P(\beta|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \Psi, \kappa, \sigma_v^2, \{\tau_g\}, \{\alpha_{gt}\})$
- Step 8: sample σ_v^2 from $P(\sigma_v^2|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \Psi, \beta, \kappa, \{\tau_g\}, \{\alpha_{gt}\})$
- Step 9: sample τ_g from $P(\tau_g|\{Y_{gt}\}, \{W_{gt}\}, \{H_{gt}\}, \Psi, \beta, \kappa, \sigma_v^2, \{\alpha_{gt}\})$
- Step 10: sample α_{gt} from $P(\alpha_{gt}|Y_{gt}, W_{gt}, H_{gt}, \Psi, \beta, \kappa, \sigma_v^2, \tau_g)$

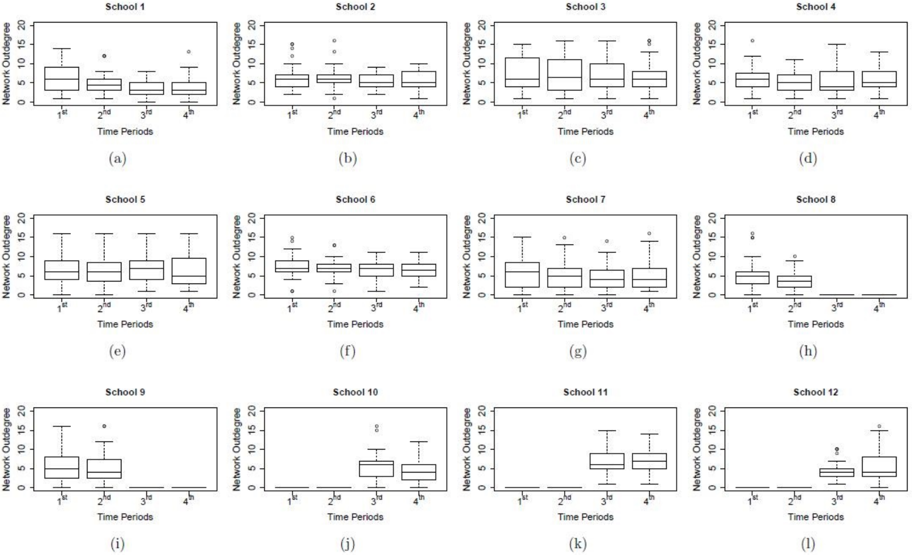
Table: Simulation Result (I)

		Sample I: m=50,G=20,T=5									
		Full-D2		Full-D1		Unobs. homo.		Unobs. Deg.		SDPD	
Para.	DGP	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
λ	0.3000	0.3027	0.0083	0.3712	0.0284	0.3446	0.0206	0.2184	0.0375	0.4288	0.0264
ρ	0.2000	0.1959	0.0094	0.2733	0.0272	0.2877	0.0149	0.1396	0.0199	0.4515	0.0163
μ	-0.1000	-0.0969	0.0087	-0.1128	0.0168	-0.1270	0.0216	-0.1091	0.0429	-0.1142	0.0172
β_1	1.0000	1.0029	0.0102	1.0329	0.0143	1.0358	0.0155	0.9905	0.0192	1.0483	0.0152
β_2	1.0000	0.9940	0.0209	0.9263	0.0423	0.9725	0.0393	1.0695	0.0537	0.8749	0.0443
κ_{11}	1.2000	1.1271	0.1182	0.5782	0.3727	1.0767	0.0985	-	-	-	-
κ_{12}	0.4000	0.4072	0.2294	-	-	0.7030	0.2131	-	-	-	-
κ_{21}	1.0000	0.9958	0.0628	1.1321	0.0659	-	-	0.5647	0.0590	-	-
κ_{22}	0.5000	0.5310	0.0822	-	-	-	-	0.5610	0.0813	-	-
γ_0	-1.0000	-0.9532	0.1005	-1.6506	0.1725	0.4960	0.1445	-3.5724	0.1638	-	-
γ_1	0.5000	0.5072	0.0689	0.4284	0.1099	0.4357	0.0959	0.3935	0.0780	-	-
γ_2	0.5000	0.5072	0.0166	0.4485	0.0184	0.4132	0.0205	0.3932	0.0191	-	-
γ_3	0.5000	0.5096	0.0172	0.4477	0.0169	0.4171	0.0196	0.3952	0.0180	-	-
γ_4	0.5000	0.5106	0.0701	0.6677	0.1234	1.0132	0.0943	0.7445	0.0823	-	-
γ_6	0.3000	0.3127	0.0115	0.2615	0.0132	0.3714	0.0178	0.2513	0.0180	-	-
γ_7	0.3000	0.3033	0.0147	0.2372	0.0121	0.3349	0.0208	0.2236	0.0150	-	-
γ_8	-1.0000	-1.0143	0.0250	-0.4629	0.0180	-0.7176	0.0338	-0.7836	0.0279	-	-
δ_1	-2.0000	-1.9337	0.0989	-2.1208	0.1738	-2.0375	0.1164	-	-	-	-
δ_2	-1.0000	-1.1998	0.1261	-	-	-1.8668	0.0814	-	-	-	-
ξ_1	1.0000	1.0026	0.0614	1.0551	0.0538	-	-	0.9015	0.0488	-	-
ζ_1	0.5000	0.4923	0.0780	0.9321	0.0505	-	-	0.5850	0.0915	-	-
ξ_2	0.7000	0.7017	0.0758	-	-	-	-	0.8514	0.0894	-	-
ζ_2	1.0000	1.0079	0.0460	-	-	-	-	0.9729	0.1110	-	-
σ_v^2	1.0000	0.9754	0.0537	1.7936	0.2763	1.8903	0.1024	3.2030	0.1026	3.9850	0.1381
AICM		3968.8	303.30	6320.5	636.10	5185.1	478.90	5420.2	599.10	-	-

Note: This Monte Carlo study contains 100 repetitions. The mean and standard deviation of the point estimates across repetitions are reported. The column "Two dimensions" refers to the true model that generates the artificial data, which has latent variables in two dimensions. The column "One dimension" refers to the model that has only one-dimensional latent variable. The column "Unobs. Deg." refers to the alternative network formation model for unobserved degree heterogeneity. The column "SDPD" refers to the SDPD model which neglects endogenous network formation.

Empirical Study

- The data set comes from the project *Luminous Shine and Dark Shadow: The Duality of Late Adolescents to Early Adults' Friendship Networks* conducted by the Institute of Sociology, Academia Sinica, Taiwan between 2008 and 2011.
- The project conducted surveys in 12 high school classes in southern Taiwan, and consists of two questionnaires:
  - **Network questionnaire** collects information on students' friend and foe nominations; collected 3 times each semester; A total of 18 waves.
  - **General questionnaire** collects information about students' family background, academic performance, experience of school life, and so on; Collected twice each year.
- After matching the two, we obtain a panel data of 12 classes (networks) in 4 time snapshots.



(a) (b) (c) (d)
(e) (f) (g) (h)
(i) (j) (k) (l)

Empirical Study

- The main dependent variable in outcome equation: student's academic performance in terms of percentile ranking.
- We investigate 4 types of networks: friendship, foe, study mate, and cram-school mate, all of which are time-varying.
- We control an array of individual characteristics in the SDPD outcome equation, including gender, age, height, weight, and family background, such as number of siblings, parents' education levels, parents' current job status, and whether parents were fighting recently or whether family was in financial hardship or not.
- In the network formation model, we control many dynamic features, such as the previous friendship link, previous academic ranking, and whether individuals share common friends or enemies previously.

Table 3: Variable Definition and Summary Statistics

Variable	Definition	1^{st} period		2^{nd} period		3^{rd} period		4^{th} period			
		Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.	Min	Max
Ranking	Student's percentile ranking	52.8438	27.6137	54.1561	27.1148	54.1868	26.6620	56.1251	26.3881	1.2195	98.7805
Male	Dummy of male student	0.5485	0.4983	0.5485	0.4983	0.5860	0.4932	0.5860	0.4932	0.0000	1.0000
Sibling	Number of siblings	1.5180	0.7604	1.5180	0.7604	1.4839	0.7786	1.4839	0.7786	0.0000	1.0000
Fehl	Dummy of father's education lower than high school	0.1967	0.3980	0.1967	0.3980	0.1882	0.3914	0.1882	0.3914	0.0000	1.0000
Fehh	Dummy of father's education higher than high school	0.3019	0.4597	0.3019	0.4597	0.2903	0.4545	0.2903	0.4545	0.0000	1.0000
Mehl	Dummy of mother's education lower than high school	0.2327	0.4231	0.2327	0.4231	0.2285	0.4204	0.2285	0.4204	0.0000	1.0000
Mehh	Dummy of mother's education higher than high school	0.2105	0.4082	0.2105	0.4082	0.1989	0.3997	0.1989	0.3997	0.0000	1.0000
Funemp	Dummy of father being unemployed	0.0582	0.2344	0.0693	0.2542	0.0645	0.2460	0.0538	0.2259	0.0000	1.0000
Fretired	Dummy of father being retired	0.0222	0.1474	0.0305	0.1721	0.0215	0.1453	0.0296	0.1696	0.0000	1.0000
Mumemp	Dummy of mother being umemployed	0.0332	0.1795	0.0305	0.1721	0.0215	0.1453	0.0134	0.1153	0.0000	1.0000
Mretired	Dummy of monther being retired	0.0111	0.1048	0.0111	0.1048	0.0054	0.0732	0.0054	0.0732	0.0000	1.0000
Housewife	Dummy of monther being housewife	0.2050	0.4043	0.2133	0.4102	0.2608	0.4396	0.2661	0.4425	0.0000	1.0000
Height	Student's height, measured by centimeter	166.5208	8.0853	166.7036	8.3042	167.7312	8.3510	167.8145	8.4681	146.0000	193.0000
Weight	Student's weight, measured by kilogram	58.0928	10.9713	58.3416	10.7433	59.4991	11.7241	59.9556	11.7543	35.0000	98.0000
Divorce	Dummy of patents being divorced	0.0471	0.2248	0.0471	0.2248	0.0349	0.1839	0.0457	0.2216	0.0000	1.0000
Age	Student's age	16.3352	0.4900	17.3130	0.5154	17.0860	0.6469	18.0887	0.6444	15.0000	20.0000
Lessmoney	Dummy of family in financial hardship	0.1801	0.3848	0.1357	0.3430	0.1586	0.3658	0.1344	0.3416	0.0000	1.0000
Parentfight	Dummy of parents have more conflicts or fights	0.0443	0.2061	0.0388	0.1933	0.0269	0.1620	0.0188	0.1361	0.0000	1.0000
Network density	A measure of network connectedness	0.3009	0.1389	0.3019	0.0974	0.3265	0.1077	0.3201	0.0796	0.0070	0.5172
Clustering coefficient	A measure of network transitivity	0.0955	0.0571	0.1154	0.0483	0.1304	0.0661	0.1136	0.0368	0.0000	0.2930
Number of schools		9		9		10		10			
Number of students		361		361		372		372			

Our empirical specification

- With a short panel, we only include group-time effect.
- We compare the following empirical specifications:
  - static spatial model with exogenous networks
  - SDPD model with exogenous networks
  - selection corrected SDPD model with dynamic and endogenous networks, with $\bar{p}_1$ and $\bar{p}_2$ of different dimensions.
- We adopt the Akaike information criterion Monte Carlo (AICM) in Raftery et al. (2007) to select the dimension of z_{igt}'s and m_{igt}'s: the model with $\bar{p}_1 = 2$ and $\bar{p}_2 = 2$ is selected.

Table: Estimation Results of Peer Effects on Academic Performance Under SAR and SDPD Models with Exogenous Friendship Networks

	(I)	(II)	(III)	(IV)
λ	0.299***	0.255***	0.121***	0.107***
	(0.043)	(0.047)	(0.029)	(0.035)
ρ	-	-	0.794***	0.793***
			(0.017)	(0.018)
μ	-	-	0.008	0.005
			(0.028)	(0.031)
Own and Contextual Effects	No	Yes	No	Yes
Group-Time Effect	Yes	Yes	Yes	Yes
σ^2	646.167	641.468	210.606	209.470
	(29.382)	(29.154)	(9.583)	(9.519)

Note: Column (I): cross-sectional SAR model without contextual effects. Column (II): cross-sectional SAR model with contextual effects. Column (III): SDPD model without contextual effects. Column (IV): SDPD model with contextual effects. The asterisks ***(**,*) indicates that its 99% (95%, 90%) highest posterior density range does not cover zero.

Table: Estimation Results of Peer Effects on Academic Performance from Other Relationship Networks

	Enemy	Study mate	Cram school mate
λ	-0.060	0.115***	0.065*
	(0.074)	(0.032)	(0.038)
ρ	0.810***	0.798***	0.809***
	(0.018)	(0.018)	(0.019)
μ	-0.017	-0.006	-0.020
	(0.024)	0.017	(0.020)
Own and Contextual Effects	Yes	Yes	Yes
Group-Time Effect	Yes	Yes	Yes
σ^2	213.594	210.3244	212.523
	(9.800)	(9.605)	(9.705)

Note: The asterisks $***$ ($**$,$*$) indicates that its 99%(95%, 90%) highest posterior density range does not cover zero.

Table: Estimation Results of Peer Effects on Academic Performance Under SDPD Model with Dynamic Friendship Network Formation

	(I) One dimension	(II) Two dimensions	(III) Three dimensions
λ	0.103***	0.094***	0.097***
	(0.033)	(0.035)	(0.033)
ρ	0.791***	0.789***	0.789***
	(0.019)	(0.018)	(0.018)
μ	0.006	0.003	0.011
	(0.031)	(0.032)	(0.032)
Own and Contextual Effects	Yes	Yes	Yes
Group-Time Effect	Yes	Yes	Yes
Endog. Network Formation	Yes	Yes	Yes
σ^2	209.888	208.712	210.789
	(9.657)	(9.667)	(9.692)
AICM	21344	17729	18696
se(AICM)	39.759	56.832	79.092

Table: Estimation Results of Network Formation

	(I) One dimension	(II) Two dimensions	(III) Three dimensions
Constant	-0.748^{***}	1.078^{***}	3.045^{***}
	(0.183)	(0.124)	(0.179)
Same gender	1.503^{***}	1.877^{***}	2.476^{***}
	(0.076)	(0.083)	(0.099)
Lessmoney (i,t)	-0.223^{***}	-0.309^{***}	-0.050
	(0.100)	(0.094)	(0.085)
Lessmoney (j,t)	-0.070	-0.108	0.062
	(0.058)	(0.089)	(0.08)
Parentfight (i,t)	-0.271	-0.036	-0.380
	(0.166)	(0.129)	(0.160)
Partenfight (j,t)	-0.111	-0.666^{***}	-1.142^{***}
	(0.120)	(0.167)	(0.153)
Friend (t-1)	2.968^{***}	3.367^{***}	3.934^{***}
	(0.046)	(0.057)	(0.072)
Common friend (t-1)	0.204^{***}	0.295^{***}	0.320^{***}
	(0.015)	(0.019)	(0.016)
Common enemy (t-1)	0.344	0.770^{***}	1.275^{***}
	(0.215)	(0.221)	(0.254)
Acad. ranking (i,t-1)	0.000	0.000	-0.006^{***}
	(0.001)	(0.001)	(0.001)
Acad. ranking (j,t-1)	0.003^{***}	0.004^{***}	0.002^{***}
	(0.001)	(0.001)	(0.001)
Diff. of ranking (t-1)	-0.007^{***}	-0.005^{***}	-0.010^{***}
	(0.001)	(0.001)	(0.001)

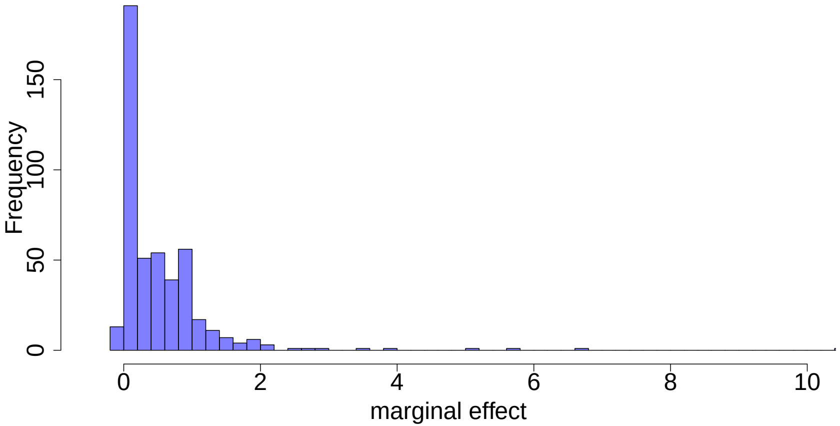
A summary of empirical results

- We get positive and significant contemporary peer effect from the static spatial model, the SDPD model, and the full model.
- The static model tends to over-estimate λ.
- There is some bias correction on λ when the dimension of z_{igt}'s and m_{igt}'s increase to 2.
- For network formation
  - Homophily effect strongly exists in gender dimension
  - Students make less friends when family is in financial hardship
  - Positive and significant effect from previous friendship link
  - Transitivity from common friends and enemies matter
  - Direct and homophily effect from the lagged academic ranking matters
  - Unobserved homophily and degree heterogeneity matter

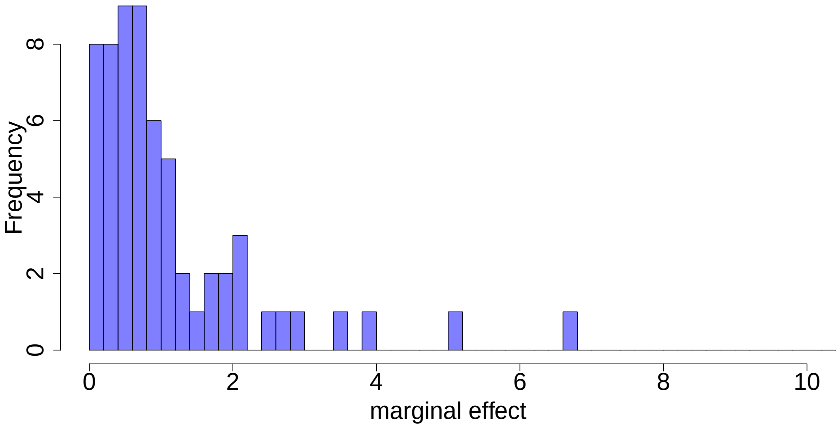
Counter-factual simulation

- Following the empirical framework of this paper, we study a policy scenario that government (or school) agencies provide financial assistance to students' families which are in financial hardships.
- From empirical results, we find the dummy variable "family in financial hardship" has significant negative effects on both student's friendship formation (-0.309 for sender & -0.108 for receiver) and academic performance (-2.504).
- We expect that the financial releasing policy (program) will help students to improve their social networking and school academic performance.
- We pay attention to the multiplier effect generated by this policy through network interactions.

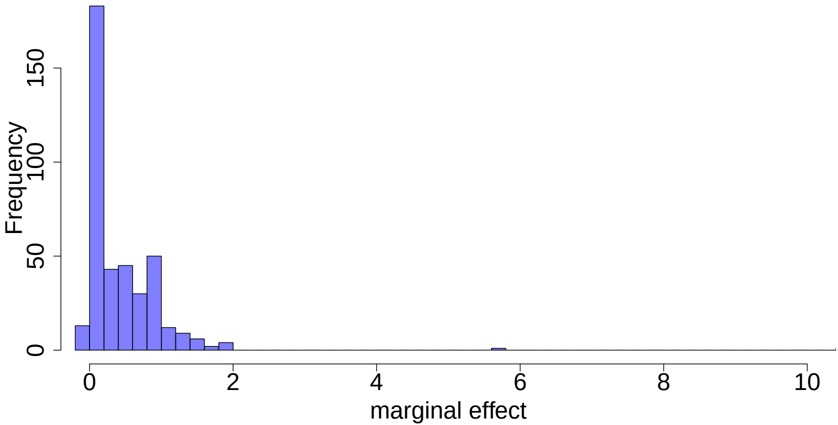
Counter-factual simulation of removing family financial hardship on students' friendship formation – the whole sample



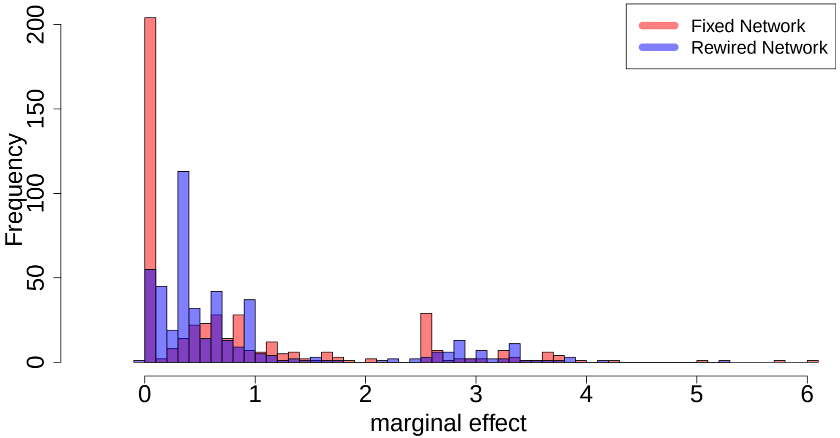
Counter-factual simulation of removing family financial hardship on students' friendship formation – the treated group



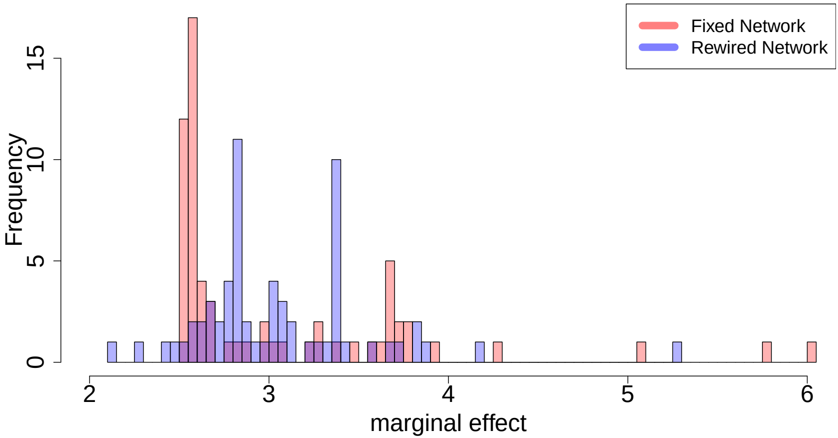
Counter-factual simulation (continue) – the untreated group



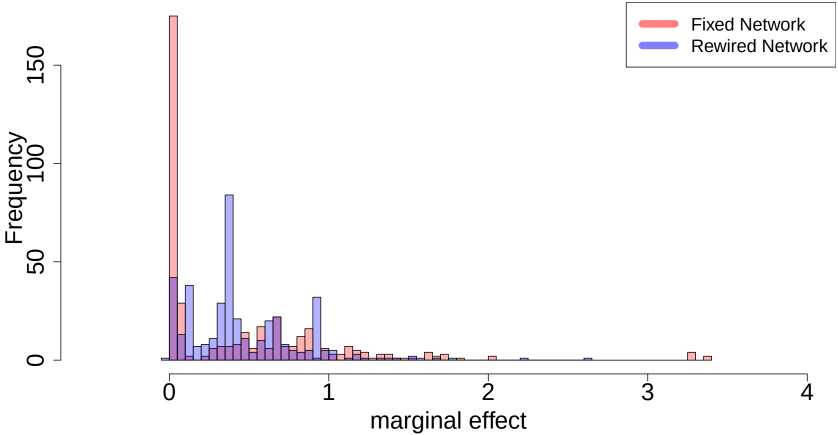
Counter-factual simulation of removing family financial hardship on child's academic performance – the whole sample



Counter-factual simulation of removing family financial hardship on child's academic performance – the treated group



Counter-factual simulation (continue) – the untreated group



The Stochastic Actor-Based Model for Network and Behavior Dynamics

- Applications of stochastic actor-based model (Snijders, 1996, 2001; Snijders et al., 2007) can be found in
  1. Statistics (e.g., Greenan, 2014),
  2. Psychology (e.g., Giletta et al., 2012),
  3. Medicine (e.g., Gesell et al., 2012),
  4. Public health (e.g., Valente et al., 2010),
  5. Sociology (e.g., Knecht et al., 2010),
  6. Marketing (e.g., Godinho et al., 2014) study the social influence on iPhone 3G adoption using mobile data from Europe.
  7. Multidisciplinary study (e.g., Lewis et al. (2012)): analyze the coevolution of friendship formation and shared tastes among friends using Facebook data.
- Snijders manages a website SIENA to promote the stochastic actor-based model

The Stochastic Actor-Based Model for Network and Behavior Dynamics

- We observe network g and several behavioral outcomes $(y_1, y_2, \cdots, y_H)$ at two or more discrete points in time (say $t_1 < t_2 < \cdots < t_M$), and they are treated as *snapshots* from a continuous process.
- The model assumes that between any two discrete time points, t_m and t_{m+1}, there are "micro" steps at stochastically determined moments that individuals have chances to change their network ties or behaviors.
- Changes of state variables, $z = (g, y_1, \cdots, y_H)$, are assumed to follow a continuous Markov process.
- The changes of network tie and behavior made by an individual are conditionally independent of each other, given the current state of the process.
- Each individual can make one and only one change at a time. Also, one and only one individual can make decision at a time.

The Stochastic Actor-Based Model for Network and Behavior Dynamics

- How frequent an individual i has an opportunity to make a change in a time period $t_m \leq t \leq t_{m+1}$ is determined by a Possion process with the parameter given by a rate function, λ_i. (for simplicity, we may assume $\lambda_i = \lambda$ in practice.)
- The core of the model is a micro-mechanism which depends on individual utility functions.
  For network $f_i^{[g]}(\beta^{[g]}, z) = \sum_k \beta_k^{[g]} s_{ik}^{[g]}(z)$
  For behavior $f_i^{[y_h]}(\beta^{[y_h]}, z) = \sum_k \beta_k^{[y_h]} s_{ik}^{[y_h]}(z)$ for $h = 1, \cdots, H$
- Proper $s_{ik}^{[g]}(z)$ and $s_{ik}^{[y_h]}(z)$ are chosen to capture homophily and social influence.

The Stochastic Actor-Based Model for Network and Behavior Dynamics

- The objective function measures individual utility with a given configuration of network and behaviors, which is a sum of a evaluation function $f_i(\beta, z)$ and a stochastic error term from the extreme type-I distribution.
- The probability of individual i going to some new state z is given by

$$\frac{\exp(f_i(\beta, z))}{\sum_{z\in\mathcal{C}} \exp(f_i(\beta, z'))}$$

- To capture the effect of social influence, one can use the "average similarity effect" (López-Pintado, 2008; Young, 2009)

$$s_i^{[y_h]}(z) = g_{i+,t-1}^{-1} \sum_{j\neq i} g_{ij,t-1}(sim_{ij}^{[y_{h,t}]} - \widehat{sim}^{[y_{h,t}]}),$$

where $sim_{ij}^{[y_{h,t}]} - \widehat{sim}^{[y_{h,t}]}$ is the centralized similarity score with $sim_{ij}^{[y_{h,t}]} = \frac{\max_{ik}|y_{hi,t}-y_{hk,t}|-|y_{hi,t}-y_{hj,t}|}{\max_{ij}|y_{hi,t}-y_{hj,t}|}$ and $\widehat{sim}^{[y_{h,t}]}$ is the mean of similarity scores.

The Stochastic Actor-Based Model for Network and Behavior Dynamics

- Similarly, we capture homophily from certain behavior variable y_h by the "similarity effect"

$$s_i^{[g]}(z) = \sum_{j \neq i} g_{ij,t}\left(sim_{ij}^{[y_{h,t-1}]} - \widehat{sim}^{[y_{h,t-1}]}\right)$$

- To prevent the identification problem due to simultaneity, we follow Snijders et al. (2007) and exploit the time order of the variables based on the concept of causality.
- Homophily is reflected by a "later" change in network tie, which follows an "earlier" configuration of behavior; whereas peer influence is reflected by a "later" change in behavior following an "earlier" configuration of network ties.

The Stochastic Actor-Based Model for Network and Behavior Dynamics

- Estimation of the stochastic actor-based model can be done by maximum likelihood, method of Moments, or Bayesian MCMC.
- For the method of moments, estimation works as follows:
  1. Starts from a given set of initial parameter values in the rate and objective functions.
  2. Between any two network observations, we simulate individual changes in network ties and behaviors.
  3. The parameter estimates are determined by searching for values under which simulated and observed sample moments resemble each other most closely.
- Model relies on substantial variations of network links and behaviors between observations to identify parameters in the rate and utility functions.

Empirical Example (I)

Teenage Friends and Lifestyle Study (incorporated in RSiena as an illustrating example)

- R script file can be downloaded from https://www.stats.ox.ac.uk/~snijders/siena/
- data description (http://www.stats.ox.ac.uk/~snijders/siena/s50_data.htm)
- two study objects: friendship network and drinking behavior.
- we like to study how network evolves over time and how it affects (and being affected) by drinking behavior.
- 50 girls and 3 sample waves.

Empirical Example (II): Hu, Hsieh, and Jia (2014)

Introduction

This empirical example is based on my working paper: Network Based Targeting: The Effectiveness of Peer Influence within Social Networks

- Mobile phones have been widely adopted worldwide. In 2014, approximately 4 billion people, more than half of the world's population, were using mobile phones.
- People use mobile phones to connect to the world and, more importantly, to others. These changes present wireless carriers and researchers with a great opportunity to study consumer behaviors, their social networks, and ultimately the interplay between the two.
- Previous marketing studies have used mobile data; however, most studies have only used the information about consumer characteristics and their customerships with the carriers, such as choices on plans and services. Information on customers' connections (networks) is not used.

Empirical Example (II): Hu, Hsieh, and Jia (2014)

Introduction

- In this paper, we take full advantage of a rich mobile database from a carrier in China. We simultaneously consider how consumers' networks evolve and how their behaviors change.
- The main challenge of analyzing such a dataset is to distinguish the effect of peer influence from the effect of homophily.
- Peer influence is an expression of an individual's conformity (Young, 2009). If peer influence is strong, an individual is inclined to adopt a behavior similar to that of his/her peers, particularly when the behavior is exhibited by many of those peers.
- In this study, we focus on how peer adoption affects individuals' decisions; specifically, we examine how peer influence affects the likelihood of an individual adopting a Samsung Note II mobile phone, Samsung high-end mobile phone, and Samsung brand mobile phone.

Empirical Example (II): Hu, Hsieh, and Jia (2014)

Introduction

- Homophily (Lazarsfeld et al., 1954) is based on the observation that individuals who behave similarly tend to associate with each other.
- In this study, connections (associations) between customers are built based on their call detailed record (CDR).
- Our longitudinal data gives us chances to distinguish these two effects.
- we apply the stochastic actor-based dynamic network model (Snijders, 1996; 2001; Snijders et al., 2007) to analyze such a dataset. The model explains the co-evolution of networks and behaviors based on observations at several time snapshots.

Empirical Example (II): Hu, Hsieh, and Jia (2014)

Introduction

- As peer influence may not exist for every networks, we further ask the following question: under what circumstances (i.e., network structures) is peer influence effective?
- There are numbers of theoretical predictions in the sociological and economic literature, e.g., Coleman (1988) stressed that presence of network closures, i.e., cohesive ties, promote peer influences. On the other hand, Burt (1992) argued that more brokerage opportunities, i.e., less network closures, are beneficial for peer influences to take place.
- In this paper, we provide an empirical investigation on various network structures to see which structures can facilitate peer influences on phone adoptions.

Empirical Example (II): Hu, Hsieh, and Jia (2014)

Roadmap of our investigation

Network Clustering
Using T- CLAP Algorism
(Godinho de Matos, et al 2014)

↓

Identify Peer Effects in Each Network
Using the Stochastic Actor-Based Dynamic
Network Model

↓

Adoption of Samsung Note II	Adoption of Samsung High-end Phones	Adoption of Samsung Phones
		

↓

Explore Relationships b/w
Network Structure Measures
and Peer Effects
(Meta Analysis)

↓

Policy Simulation

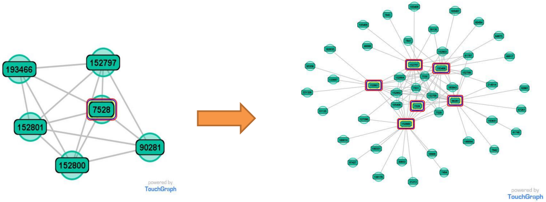
Compare individual based
seeding strategy and network
based seeding

Empirical Example (II): Hu, Hsieh, and Jia (2014)

What we are doing in this study - Stage 1

Network Construction -

- All Samsung Note II adopters as initial "seeds."
- Two layer snowball sampling from initial seeds (Chen et al., 2013)
- The resulting snowball networks might be overlapped.
- We use the modified T-CLAP algorithm (Godinho de Matos et al., 2014) to partition the network sample so that the resulting network samples maintain high relevance but with a controlled overlapping rate.



Empirical Example (II): Hu, Hsieh, and Jia (2014)

What we are doing in this study - Stage 2

Peer influence estimation -

- It is hard to distinguish between *social influence* and *homophily* (Bramoulle et al., 2009; Shalizi and Thomas 2011; Hsieh and Lee 2015) using cross-sectional data.
- Aral et al. (2009) proposed an approach using propensity score matching. But it is designed for one single large online network and doesn't fit our purpose.
- We propose to use the *Stochastic Actor-Based Model for Network and Behavior Dynamics* (Snijders 1996, 2001, 2007).
- The model estimates the co-evolution of network formation and individual behaviors using longitudinal network information.

Empirical Example (II): Hu, Hsieh, and Jia (2014)

What we are doing in this study - Stage 3

Meta regression of network structure measures on peer influences -

- Average customer characteristics are controlled. In addition, we also include the following network structure characteristics:
    - Clustering coefficient
    - Network structure entropy based on degree and eigenvector centralities.
    - Standard deviation of edge numbers across time
    - Minimum and maximum eigenvalue of adjacency matrix
    - Epidemic threshold and assortativity
- Moreover, to prevent high correlations between network structure measures affect meta regression outcomes, we conduct a factor analysis based on all of the network measures of interest and identify three orthogonal factors that separately capture the major variations in the network measures – they can be categorized into network cohesiveness, network heterogeneity, and spectrum of network matrices.

Empirical Example (II): Hu, Hsieh, and Jia (2014)
Summary of Findings

- About peer influence and homophily
  - From 313 individual social networks in the data, we identify 34.6% of the networks exhibit social influence for Samsung Note II adoption, 34.4% for Samsung high-end phone adoption and 26.7% for Samsung brand adoption.
  - Overall, we didn't find homophily driven by the adoption of the products. People don't get connected because they use the same phone.
- About network structure and peer influence
  - Two characteristics of network structure can be used to predict the effectiveness of peer influence within a network: 1. network cohesiveness (i.e., clustering coefficient and assortativity) and 2. the heterogeneity (i.e., entropy and epidemic threshold) and dynamics of the network.

Empirical Example (II): Hu, Hsieh, and Jia (2014)

Results - Peer Influence

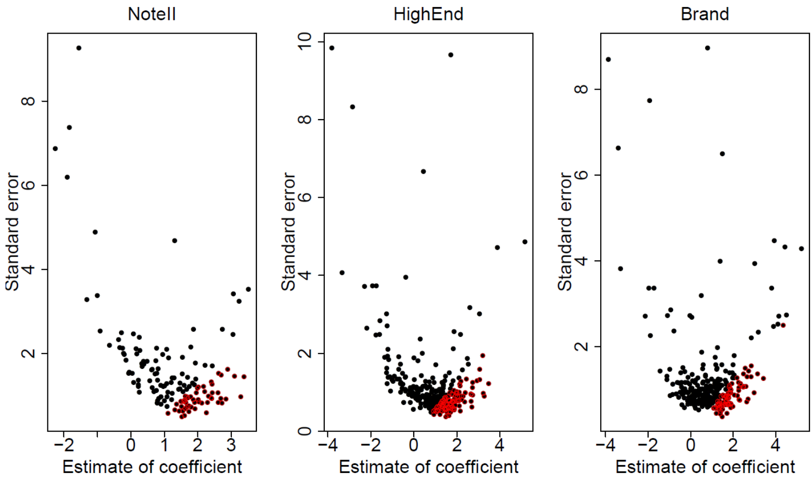


Figure: Scatter Plots of Estimated Social Influence

Empirical Example (II): Hu, Hsieh, and Jia (2014)
Results - Homophily

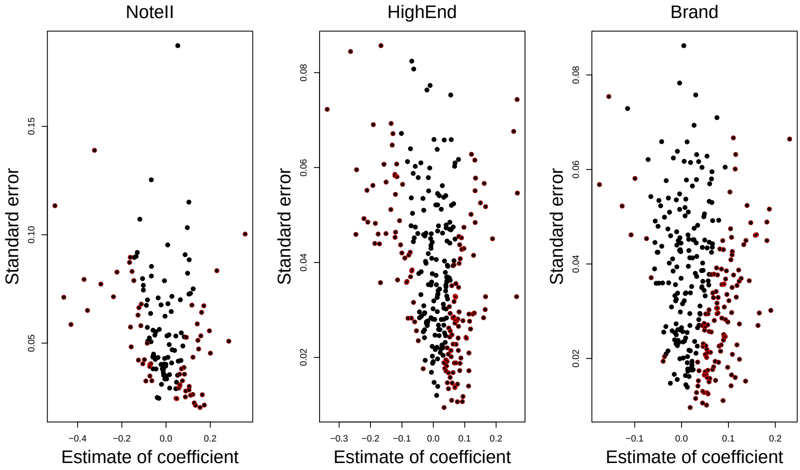


Figure: Scatter Plots of Estimated Homophily

Table: Summary Statistics of the Variables Used in the Meta-Regressions

	Note II		High End		Brand	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Peer Influence						
Significance	0.346	0.477	0.344	0.476	0.267	0.442
Magnitude	1.110	1.530	0.838	1.434	0.895	1.366
Demography						
Sex ratio	0.654	0.097	0.643	0.084	0.644	0.082
Age	38.250	2.156	38.139	2.178	38.208	2.000
Tenure	38.518	6.531	33.213	6.633	37.405	6.375
Network Measures						
Clustering coefficient	0.442	0.100	0.422	0.105	0.423	0.104
Assortativity	0.062	0.133	0.025	0.143	0.021	0.144
Degc. entropy	4.768	0.383	4.742	0.378	4.727	0.376
Eigc. entropy	4.486	0.502	4.444	0.495	4.423	0.490
Sd of edge number	162.255	212.506	130.358	196.678	115.455	178.009
Minimum eigenvalue	6.253	4.229	5.935	3.902	5.862	3.878
Maximum eigenvalue	22.722	26.638	20.969	24.667	20.450	24.406
Epidemic threshold	0.042	0.025	0.047	0.026	0.048	0.026
Sample size	147		296		300	

Note: Degc. stands for degree centrality. Eigc. stands for eigenvector centrality.

Table: Correlation of Network Statistics

Clustering coefficient	1							
Assortativity	0.578***	1						
Degc. entropy	-0.102	0.365***	1					
Eigc. entropy	-0.220***	0.194***	0.833***	1				
S.d. of edge number	-0.219***	0.0968	0.421***	0.529***	1			
Minimum eigenvalue	0.0191	-0.0650	-0.0927	-0.0311	0.0562	1		
Maximum eigenvalue	0.00493	-0.0717	-0.0796	-0.0270	0.0388	0.947***	1	
Epidemic threshold	-0.347***	-0.438***	-0.570***	-0.692***	-0.416***	-0.0236	-0.0202	1

$* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

Table: Factor Analysis of Network Statistics

	Factor-1	Factor-2	Factor-3	Uniqueness
Clustering coefficient			0.9288	0.1091
Assortativity			0.8144	0.2416
Degc. entropy	0.8756			0.2179
Eigc. entropy	0.9439			0.1041
S.D. of edge number	0.7100			0.4523
Minimum eigenvalue		0.9854		0.0287
Maximum eigenvalue		0.9837		0.0321
Epidemic threshold	-0.7496			0.2179

Empirical Example (II): Hu, Hsieh, and Jia (2014)

Meta Regression Analysis

- Binary probit model for the indicator of significance
- Random effect meta regression model for the magnitude of estimated social influence

$$y_i|\theta_i \sim N(\theta_i, \sigma_i^2), \quad \text{where } \theta_i \sim N(\mathbf{x}_i\beta, \tau_i^2),$$

so equivalently the model can be rewritten as

$$y_i = \mathbf{x}_i\beta + u_i + \epsilon_i, \quad \text{where } u_i \sim N(0, \tau^2) \text{ and } \epsilon_i \sim N(0, \sigma_i^2).$$

To estimate, we use the weighted least square (WLS) method with the unknown between-network variance τ^2 estimated by the restricted maximum likelihood (REML) method.

- We include averages of customers' age, sex, and tenure in regressions as controls of demographic backgrounds.

Table: Meta-regression of Peer Influence – Separated Regressions

	Samsung NoteII		Samsung HighEnd		Samsung Brand	
	significance	magnitude	significance	magnitude	significance	magnitude
Clustering coefficient	0.566 (0.48, 0.010)	0.405 (0.58, 0.004)	2.978** (3.69, 0.076)	2.578** (5.29, 0.120)	4.493*** (4.91, 0.113)	3.657*** (8.23, 0.212)
Assortativity	1.155 (1.40, 0.020)	-0.303 (-0.62, 0.004)	2.098*** (3.75, 0.077)	1.168** (3.45, 0.073)	3.748*** (5.90, 0.148)	2.006*** (6.49, 0.152)
Degc. entropy	1.061*** (3.60, 0.081)	-0.020 (-0.14, 0.002)	1.141*** (5.28, 0.116)	0.293*** (2.62, 0.058)	1.515*** (6.40, 0.168)	0.387*** (3.43, 0.068)
Eigc. entropy	0.995** (4.14, 0.108)	0.216^{+} (1.78, 0.023)	1.007** (5.81, 0.136)	0.418** (4.75, 0.105)	1.089*** (5.70, 0.142)	0.376*** (4.21, 0.086)
S.d. of edge number	0.00236*** (3.70, 0.087)	0.00197*** (5.79, 0.193)	0.00104* (2.49, 0.055)	0.00166*** (5.93, 0.139)	0.00126** (2.72, 0.057)	0.00186*** (5.37, 0.117)
Minimum eigenvalue	-0.0129 (-0.50, 0.011)	-0.000583 (-0.04, 0.002)	0.0105 (0.54, 0.039)	0.0105 (0.85, 0.038)	-0.0275 (-1.20, 0.040)	-0.00362 (-0.26, 0.031)
Maximum eigenvalue	-0.00228 (-0.55, 0.011)	-0.000686 (-0.28, 0.002)	0.00195 (0.63, 0.040)	0.001 (0.50, 0.036)	-0.00545 (-1.44, 0.042)	-0.00178 (-0.78, 0.033)
Epidemic threshold	-22.13** (-3.89, 0.109)	–10.87*** (-3.90, 0.098)	-32.15*** (-6.71, 0.201)	-15.64*** (-8.87, 0.241)	-36.62*** (-6.65, 0.221)	-14.08*** (-8.47, 0.221)
Observations	147		296		300	

Note: The results reported in this table are collected from several separated regressions; each one focuses on a specific network statistic. In all of the separated regressions, we control for customer demographics. The values reported in parentheses are t statistics of the coefficients (left) and R^2's of the regressions (right). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table: Meta-regression of Peer Influence – Three Factors

	Samsung NoteII		Samsung HighEnd		Samsung Brand	
	significance	magnitude	significance	magnitude	significance	magnitude
Sex_ratio	-1.874	0.698	-3.504***	-0.246	-1.410	0.128
	(-1.48)	(0.86)	(-3.43)	(-0.40)	(-1.24)	(0.21)
Age	0.0414	0.00437	-0.0950*	-0.0535*	-0.156**	-0.0674**
	(0.73)	(0.14)	(-2.43)	(-2.26)	(-3.20)	(-2.97)
Tenure	0.0144	-0.0198+	0.00886	-0.0189**	0.00215	-0.00259
	(0.77)	(-1.94)	(0.70)	(-2.62)	(0.15)	(-0.34)
Factor-1	0.535***	0.155*	0.530***	0.229***	0.664***	0.170***
	(4.41)	(2.47)	(5.82)	(5.20)	(6.11)	(3.88)
Factor-2	-0.00723	0.0328	0.0808	0.0698	-0.0842	0.0143
	(-0.07)	(0.56)	(1.02)	(1.58)	(-0.86)	(0.30)
Factor-3	0.121	0.00830	0.421***	0.235***	0.683***	0.347***
	(0.90)	(0.11)	(4.39)	(4.70)	(5.86)	(7.51)
Constant	-1.462	1.600	5.002**	3.892***	5.871**	3.489***
	(-0.71)	(1.47)	(3.28)	(4.49)	(2.99)	(4.10)
Observations	147	147	296	296	300	300
τ^2	-	0.00	-	0.00	-	0.00
R^2	0.126	0.030	0.188	0.187	0.279	0.251

Note: t statistics in parentheses. The variable significance is examined using a binary Probit model. The variable magnitude is examined using the (random) mixed-effect meta-regression model. The estimates of between-network variance, τ^2, are zero for all three cases. The pseudo (adjusted) R^2 values are reported for the dependent variable of significance (magnitude). $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Reference

Gesell, Sabina B, Eric Tesdahl, and Eileen Ruchman (2012) "The distribution of physical activity in an after-school friendship network," *Pediatrics*, Vol. 129, pp. 1064–1071.

Giletta, Matteo, Ron HJ Scholte, Mitchell J Prinstein, Rutger CME Engels, Emanuela Rabaglietti, and William J Burk (2012) "Friendship context matters: Examining the domain specificity of alcohol and depression socialization among adolescents," *Journal of Abnormal Child Psychology*, Vol. 40, pp. 1027–1043.

Greenan, Charlotte C (2014) "Diffusion of innovations in dynamic networks," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Knecht, Andrea, Tom AB Snijders, Chris Baerveldt, Christian EG Steglich, and Werner Raub (2010) "Friendship and delinquency: Selection and influence processes in early adolescence," *Social Development*, Vol. 19, pp. 494–514.

Lee, Lung-fei and Jihai Yu (2010) "Estimation of spatial autoregressive panel data models with fixed effects," *Journal of Econometrics*, Vol. 154, pp. 165–185.

Lewis, Kevin, Marco Gonzalez, and Jason Kaufman (2012) "Social selection and peer influence in an online social network," *Proceedings of the National Academy of Sciences*, Vol. 109, pp. 68–72.

López-Pintado, Dunia (2008) "Diffusion in complex social networks," *Games and Economic Behavior*, Vol. 62, pp. 573–590.

Raftery, Adrian E, Michael A Newton, Jaya M Satagopan, and Pavel N Krivitsky (2007) "Estimating the integrated likelihood via posterior simulation using the harmonic mean identity," *Bayesian Statistics*, Vol. 8, pp. 1–45.

Snijders, Tom AB (1996) "Stochastic actor-oriented models for network change," *Journal of Mathematical Sociology*, Vol. 21, pp. 149–172.

——— (2001) "The statistical evaluation of social network dynamics," *Sociological Methodology*, Vol. 31, pp. 361–395.

Snijders, Tom, Christian Steglich, and Michael Schweinberger (2007) "Modeling the coevolution of networks and behavior," in Han Oud Kees van Montfort and Albert Satorra eds. *Longitudinal models in the behavioral and related sciences*: Lawrence Erlbaum, pp. 41–71.

Valente, Thomas W, Kayo Fujimoto, Paula Palmer, and Sora Park Tanjasiri (2010) "A network assessment of community-based participatory research: linking communities and universities to reduce cancer disparities," *American journal of Public Health*, Vol. 100, p. 1319.

Young, H Peyton (2009) "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning," *The American economic review*, Vol. 99, pp. 1899–1924.