Economics of Networks

Valerio Leone Sciabolazza[1]   Eleonora Patacchini[2]

[1]University of Naples Parthenope

[2]Cornell University

2021

A growing literature has sparked on various aspects of networks and their role in explaining various social and economic phenomena among economic theorists, empirical researchers and, more recently, econometricians.

The integration of models of social interactions within economic theory is an active area of research (Benhabib et al., 2011, Jackson et al. 2017, de Paula et al. 2021, Hsieh et al. 2021).

Observation that many individual outcomes vary much more *between* social groups than *within* them.

Application contexts:

- *Social interactions*: e.g., consumption, academic achievement, unemployment, crime, . . . ;
- *Strategic interactions*: e.g., finance, production, IO, international trade, . . . ;
- *Spatial externalities*: e.g., agglomeration effects, agriculture, . . . ;
- *Neighborhood effects*: e.g., new economic geography, housing market, ...

Theory:
Models of social interactions are widely used

Empirics:
Convincing tests of such models are still quite limited. This is because:

- Appropriate data sets difficult to find;
- Identification and measure of such peer effects is a quite difficult exercise;
- Econometrics of networks is lagging behind: the use of panel data econometrics and spatial econometrics is not trivial.

Methodological tools:

- Graph theory tools and spatial statistics techniques
- Spectral analysis of matrices

There is substantial empirical evidence showing that peer effects matter in (among other outcomes):

- alcohol use (Kremer and Levy, 2008),
- crime (Ludwig et al., 2001; Patacchini and Zenou, 2012; Damm and Dustmann, 2014),
- education (Calvo-Armengol et al., 2009; Epple and Romano, 2011; Sacerdote, 2011),
- environmentally friendly behavior (Brekke et al., 2010; Czajkowski et al., 2017),
- obesity (Christakis and Fowler, 2007),
- participation in extracurricular activities (Boucher, 2016),
- performance in the workplace (Herbst and Mas, 2015),
- political networks (Cohen et al., 2014, Battaglini et al. 2018, 2019, 2020)
- risky behavior (Clark and Loheac, 2007; Hsieh and Lin, 2017),
- smoking (Powell, Tauras, and Ross, 2005),
- substance use (Lundborg, 2006), and
- tax compliance and tax evasion (Fortin et al., 2007; Alm et al., 2017).

This course provides a selective overview on:

- Centrality Measures
- Modeling Peer Effects
- Modeling Network formation

- Alm, J., Bloomquist, K.M., McKee, M., 2017. When you know your neighbour pays taxes: information, peer effects, and tax compliance. Fisc. Stud.38, 587–613.
- Battaglini, M., E. Patacchini (2018), Influencing connected legislators, Journal of Political Economy, 126, 6: 2277-2322.
- Battaglini, M., E. Patacchini (2018), Social Networks in Policy Making, Annual Reviews of Economics, 11: 473-494.
- Battaglini, M., V. Leone Sciabolazza, E. Patacchini (2020), "Effectiveness of connected legislators," American Journal of Political Science 64(4): 739-756.
- Benhabib J., A., Bisin, M. O., Jackson (eds.) (2011), Handbook of Social Economics, Volume 1 (A-B), North Holland.
- Boucher, V., 2016. Conformism and self-selection in social networks. J. Public Econ.136, 30–44.
- Brekke, K.A., Kipperberg, G., Nyborg, K., 2010. Social interaction in responsibility ascription: the case of household recycling. Land Econ.86, 766–784.
- Calvo-Armengol, A., Patacchini, E., Zenou, Y., 2009. Peer effects and social networks in education. Rev. Econ. Stud.76, 1239–1267.
- Hsieh C.-S., Lin X., Patacchini E. (2011), Social Interaction Methods, Zimmerman K. (ed.), Handbook of Labor, Human Resources and Population, Springer Nature, CEPR wp n. 14141.
- Clark, A.E., Loheac, Y., 2007. It wasn't me, it was them! Social influence in risky behaviour by adolescents. J. Health Econ.26, 763–784.
- Cohen L., Malloy C. J. (2014), Friends in high places, American Economic Journal: Economic Policy, 6 (3): 63-91.
- Christakis, N.A., Fowler, J.H., 2007. The spread of obesity in a large social network over 32 years. N. Engl. J. Med.357, 370–379.
- Czajkowski, M., Hanley, N., Nyborg, K., 2017. Social norms, morals and self-interest as determinants of pro-environment behaviours: the case of household recycling. Environ. Resour. Econ.66, 647–670

- Damm, A.P., Dustmann, C., 2014. Does growing up in a high crime neighborhood affect youth criminal behavior? Am. Econ. Rev.104, 1806–1832.
- de Paula A, Tamer E., Voia M.-C. 2021. The Econometrics of Networks (Advances in Econometrics). Elsevier, Amsterdam.
- Epple, D., Romano, R., 2011. Peer effects in education: a survey of the theory and evidence. In: Benhabib, J., Bisin, A., Jackson, M.O. (Eds.), Handbook of Social Economics, vol. 1B. Elsevier, Amsterdam, pp.1053–1163.
- Fortin, B., Lacroix, G., Villeval, M.-C., 2007. Tax evasion and social interactions. J. Public Econ.91, 2089–2112.
- Herbst, D., Mas, A., 2015. Peer effects in worker output in the laboratory generalize to the field. Science350, 6260.
- Hsieh, C.-S., Lin, X., 2017. Gender and racial peer effects with endogenous network formation. Reg. Sci. Urban Econ.67, 135–147.
- Jackson M. O., Rogers B., Zenou Y. (2017), The Economic Consequences of Social-Network Structure, Journal of Economic Literature, 55(1), 49-95.
- Kremer, M., and D. Levy (2008): "Peer effects and alcohol use among college students," The Journal of Economic Perspectives, 22(3), 189–189.
- Ludwig, J., Duncan, G.J., Hirschfield, P., 2001. Urban poverty and juvenile crime: evidence from a randomized housing-mobility experiment. Q. J. Econ.116, 655–679.
- Lundborg, P. (2006): "Having the wrong friends? Peer effects in adolescent substance use," Journal of Health Economics, 25(2), 214–233.
- Patacchini, E., Zenou, Y., 2012. Juvenile delinquency and conformism. J. Law Econ. Organ.28, 1–31.
- Powell, L. M., J. A. Tauras, and H. Ross (2005): "The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior," Journal of Health Economics, 24(5), 950–968.
- Sacerdote, B.I., 2011. Peer effects in education: how might they work, how big are they and how much do we know thus far? In: Hanushek, E., Machin, S., Woessmann, L. (Eds.), Handbook of the Economics of Education. Elsevier, Dordrecht, pp.249–277.

- Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2011), "Identification of Social Interactions," in Handbook of Social Economics, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, vol. 1, pp. 853 – 964. North-Holland, Amsterdam.
- De Paula, A. (2017), "Econometrics of Network Models," in B. Honoré, A. Pakes, M. Piazzesi, L. Samuelson (Eds.), *Advances in Economics and Econometrics: Eleventh World Congress* Cambridge University Press, 268-323.
- Durlauf, S. N., and Y. M. Ioannides (2010): "Social interactions," Annual Review of Economics, 2, 451–478.
- Glaeser, E., and J. Scheinkman (2001): "Measuring social interactions," in Social Dynamics, ed. by S. N. Durlauf, and H. P. Young, pp. 83–132. MIT Press, Cambridge, MA.
- Kline, B., Tamer, E., (2021). Econometric analysis of models with social interactions. In: Graham, B., De Paula, A. (Eds.), The Econometric Analysis of Network Data. Elsevier, Amsterdam.
- Manski, C. F. (2000): "Economic analysis of social interactions," The Journal of Economic Perspectives, 14(3), 115–136.
- Scheinkman, J. A. (2008): "Social interactions," in The New Palgrave Dictionary of Economics, ed. by S. N. Durlauf, and L. E. Blume, vol. 2. Palgrave Macmillan, London.

Outline

Outline

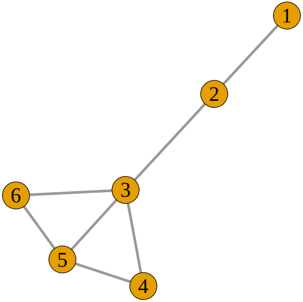
Centrality is a key concept in network studies.

There are many different ways in which a node can be important to a structure.

- **Absolute measures**: a node's influence is evaluated on the sole basis of the node's structural characteristics in the network. E.g.: degree, betweenness, closeness centrality (Freeman, 1979).
- **Relative measures**: a node's status is a function of the statuses of the nodes to which it is connected. E.g.: eigenvector, Bonacich (1987), Page Rank (Brin and Page, 1998) centrality.

Adjacency Matrix

Each network can be represented as an adjacency matrix G, where $g_{i,j} = 1$ if i is connected to j ($i \neq j$) and $g_{i,j} = 0$, otherwise.

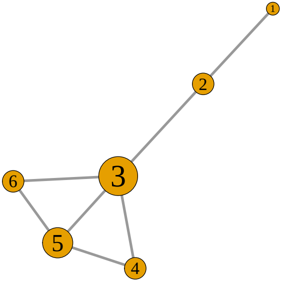


$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

This formalization can be used to compute centrality measures.

Absolute measures: degree centrality

Degree centrality is the row/column sum of the adjacency matrix.

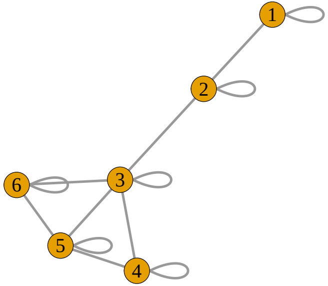


$$d(G) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 2 \\ 3 \\ 2 \end{bmatrix} \quad ‘$$

Relative measures: eigenvector centrality

Gould (1967): The connectedness of a vertex is not just how many other vertices it is connected too, but also how connected those vertices are.

Eigenvector centrality requires to consider loops.

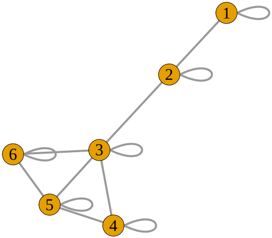


$$I+G = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} ,$$

Relative measures: eigenvector centrality

Remember that:

- $I + G$ is an n x n non-negative primitive matrix: e.g. all entries are positive and $(I+G)^k$ is strictly positive.
- by the Perron-Frobenius theorem, the largest eigenvalue of $I + G$ is positive and it has a unique eigenvector with all positive entries (this will come in hand in a few moments).
- the entry j in row i of $(I+G)^k$ gives all k length paths from i to j.

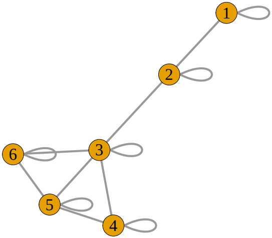


$$(I+G)^2 = \begin{bmatrix} 2 & 2 & 1 & 0 & 0 & 0 \\ 2 & 3 & 2 & 1 & 1 & 1 \\ 1 & 2 & 5 & 3 & 4 & 3 \\ 0 & 1 & 3 & 3 & 3 & 2 \\ 0 & 1 & 4 & 3 & 4 & 3 \\ 0 & 1 & 3 & 2 & 3 & 3 \end{bmatrix}$$

Note: values on the diagonal are also referred to as *subgraph centrality* (Estrada, Rodriguez-Velazquez, 2005).

Relative measures: eigenvector centrality

Let us rank agents: for each i multiply i's degree by the sum of the degrees of all agent i's connections (including i) at any path length k (Gould, 1967).

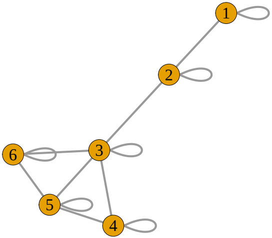


$$(I+G)^{k\to\infty}\cdot d(G) \cong \begin{bmatrix} 5.880630e+56 \\ 1.591056e+57 \\ 3.716676e+57 \\ 2.588712e+57 \\ 3.287309e+57 \\ 2.588712e+57 \end{bmatrix}$$

This is awkward to read and difficult to interpret. Gould (1967) suggests to normalize these values using the norm of the vector: $\frac{(I+G)^{k\to\infty}}{||(I+G)^{k\to\infty}||}$

Relative measures: eigenvector centrality

Following Gould (1967), we compute $\frac{(I+G)^{k\to\infty}}{||(I+G)^{k\to\infty}||}$.



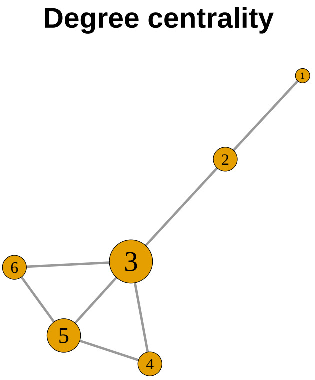
$$\frac{(I+G)^{k\to\infty}}{||(I+G)^{k\to\infty}||}\cdot d(G) \cong \begin{bmatrix} 0.09195198 \\ 0.24878406 \\ 0.58115487 \\ 0.40478177 \\ 0.51401731 \\ 0.40478177 \end{bmatrix}$$

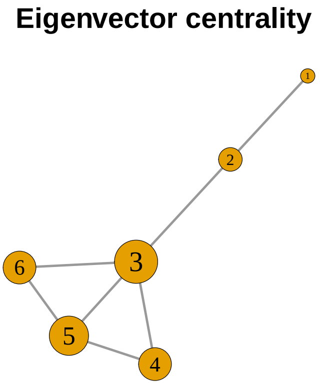
This value is called Gould index of accessibility. It measures how accessible is a vertex and how important are the vertices to which it is connected.

The Gould index is equal to the eigenvector associated to the largest eigenvalue of $I+G$ (Perron-Frobenius Theorem).

Relative measures: eigenvector centrality

The Gould index is equal to eigenvector centrality.





Note: 4 and 6 are more important than 2 even if they all have the same degree centrality (they are closer to nodes with higher degree centrality).

Relative measures: Bonacich centrality

Bonacich centrality counts the number of paths connecting i to j at any $k \in [0, \infty]$.

$$\sum_{k=0}^{\infty} G^k \cdot 1$$

It assigns higher importance to paths for which k is lower: nodes further away are less influential.

$$\sum_{k=0}^{\infty} \left(\frac{1}{\delta}\right)^k G^k \cdot 1 = \sum_{k=0}^{\infty} \phi^k G^k \cdot 1 \cong \begin{bmatrix} 1.359068 \\ 1.795340 \\ 2.617634 \\ 1.987209 \\ 2.318410 \\ 1.987209 \end{bmatrix}$$

Note 1: $\phi = 0.25$.
Note 2: We also refer to Katz-Bonacich centralities as parameter-dependent centrality measures.

Definition (Katz, 1953; Bonacich, 1987)

Given a vector $\boldsymbol{u} \in \mathbb{R}^n_+$, and $\phi \geq 0$ a small enough scalar, the vector of Bonacich centralities of parameter ϕ in network G is defined as:

$$b(g, \phi) = \sum_{k=0}^{\infty} \phi^k G^k u = (I - \phi G)^{-1} u$$

Relative measures: Bonacich centrality

Note that the definition relies on interesting properties of Taylor expansions:

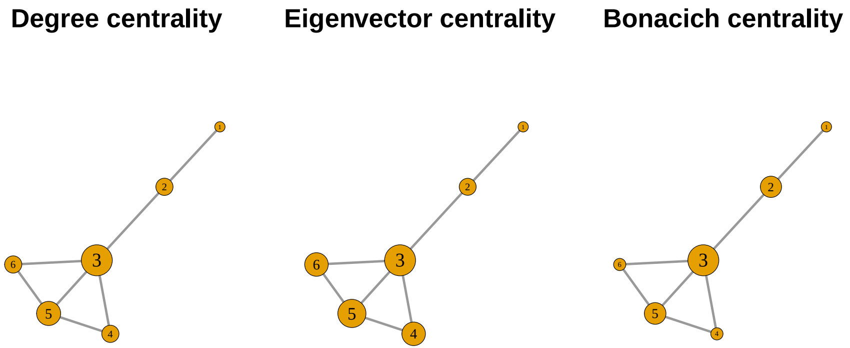
- When $\phi < 1$, then $\sum_{k=0}^{\infty} \phi^k G^k \cdot 1 \rightarrow (I - \phi G)^{-1} \cdot 1$.

Note also that:

- For $(I - \phi G)$ to be invertible, it is necessary to impose a condition on ϕ, that is $\frac{1}{\delta_{min}} < \phi < \frac{1}{\delta_{max}}$, where δ_{min} and δ_{max} are respectively the smallest and largest eigenvalue of G.

- If G is a stochastic matrix (row/column normalized), then $\delta_{min} = -1$ and $\delta_{max} = 1$, and $-1 < \phi < 1$

Relative measures: Bonacich centrality



Bonacich centrality assigns more importance to node 2 than nodes 4 and 6. This is because all nodes that want to reach 1, they must pass through 2.

Degree, eigenvector, and Bonacich centralities can be nested in a general formula (Banerjee et al., 2014):

$$C(G_{norm}, q, K) = \Big[\sum_{k=1}^{K} (\phi G_{norm})^k\Big]1 =$$

$$= \begin{cases} K = 1 \text{ and } \phi = 1 & \text{It is proportional to degree centrality.} \\ K = \infty \text{ and } \phi < \frac{1}{\delta_{max}} & \text{It coincides with Bonacich centrality.} \\ K \to \infty \text{ and } \phi \geq \frac{1}{\delta_{max}} & \text{It approaches eigenvector centrality.} \end{cases}$$

Different values of K and ϕ reflect different diffusion processes. If network externalities are local, then $K = \phi = 1$. If network externalities are global, then $K \to \infty$ and ϕ is alternatively higher or lower than $\frac{1}{\delta_{max}}$.

References:

- Banerjee, A., A. Chandrasekhar, E. Duflo, and M. Jackson (2014), "Gossip: Identifying Central Individuals in a Social Network," MIT Working Paper.
- Bonacich, P. (1987), "Power and centrality: A family of measures," *American Journal of Sociology*, 92(5): 1170-1182.
- Cliff, A.D. and Ord, J.K. (1981), *Spatial Processes*, Pion.
- Gould P. (1967), "The geographical interpretation of eigenvalues," *Transactions of the Institute of British Geographers*, 42: 53-85.
- Estrada E., Rodriguez-Velazquez J. A. (2005), "Subgraph centrality in Complex Networks", Physical Review E 71, 056103.
- Freeman, L. C. (1979), "Centrality in social networks conceptual clarification," *Social networks*, 1(3): 215-239.
- Haining, R. (1990): *Spatial Data Analysis in the Social and Environmental Science*, Cambridge University Press.
- Katz L., "A new index derived from sociometric data analysis," *Psychometrika*, 18: 39-43.
- Straffin, P., D. (1980), "Linear Algebra in Geography: Eigenvectors of Networks," *Mathematics Magazine*, 53(5).

Outline

Outline

A taxonomy of peer effects models
An emerging literature shows that measures of centrality, which were introduced as descriptive, have an interpretation within equilibrium models of behavior.

Most popular models are:

- Local Average Model: individuals have a preference to conform to the **average** action of their neighbors in a social network (Patacchini *et al.*, 2012).
- Local Aggregate Model: individuals have a preference to conform to the **sum** of actions of their neighbors in a social network (Ballester *et al.*, 2006, Calvo-Armengol *et al.*, 2009).

Note: Bramoullé et al.(2009) provide conditions for identification in the local-average model, Liu et al.(2014) derive conditions for identification in the local-aggregate model.

Local Aggregate Model:

- A node's outcome is influenced by the aggregate outcomes of their neighbours.
- Each individual i selects an effort $y_i \geq 0$ and obtains a payoff given by the following utility function:

$$u_i(y, g) = \underbrace{x_i y_i}_{\text{benefits from own effort}} - \underbrace{\frac{1}{2} y_i^2}_{\text{costs}} + \underbrace{\phi \sum_{j=1}^{n} g_{i,j} y_i y_j}_{\text{benefits from own and friends' effort}}$$

Local Aggregate Model

Interpretation:

- Individual outcomes results from both idiosyncratic characteristics (x) and peer effects ($\sum_{j=1}^{n} g_{i,j} y_i y_j$).
- Payoffs are interdependent and agents choose their levels of activity simultaneously.

Result:
If $\phi\ \mu_1(G) < 1$, where $\mu_1(G)$ is the spectral radius of G, there is a unique Nash equilibrium of the peer effect game y_i^*,where each individual provides effort proportional to that of her/his reference group and to her/his idiosyncratic characteristics.

$$y_i^* = \phi \sum_{j=1}^{N} g_{i,j} y_j^* + x_i \tag{1}$$

Local Average Model

- A node's outcome is influenced by the average outcome of its neighbours.
- Each individual i selects an effort $y_i \geq 0$ and obtains a payoff given by the following utility function:

$$u_i(y,g) = \underbrace{x_i y_i}_{\text{benefits from own effort}} - \underbrace{\frac{1}{2}y_i^2}_{\text{costs}} - \underbrace{\frac{d}{2}(y_i - \bar{y}_i)}_{\text{benefits from own and friends' effort}}$$

Where $d > 0$ and $\bar{y} = \frac{\sum_{j=1}^{n} g_{i,j} y_j}{g_{i,j}}$.

Local Average Model

Interpretation:

- Individual outcomes results from both idiosyncratic characteristics (x) and peer effects ($y_i - \bar{y}_i$).
- Payoffs are interdependent and agents choose their levels of activity simultaneously.

Result:
If $\delta\ \mu_1(G) < 1$, where $\mu_1(G)$ is the spectral radius of G, and $\delta = \frac{d}{1+d}$, there is a unique Nash equilibrium of the peer effect game y_i^*,where each individual provides effort proportional to that of her/his reference group and to her/his idiosyncratic characteristics.

$$y_i^* = \delta \sum_{j=1}^{N} g_{i,j} y_j^* + (1-\delta) x_i$$

The key difference between the Local Average and the Local Aggregate model is that (Ushchev *et al.*, 2020):

- The Local Aggregate model highlights the role of knowledge spillovers on outcomes (Ballester et al., 2006,2010; Bramoullé et al., 2014; De Marti and Zenou, 2015).
- The Local Average model aims to capture the role of social norms, such as conformist behavior or peer pressure, on outcomes (Patacchini and Zenou, 2012; Liu et al., 2014; Blume et al., 2015; Topa and Zenou, 2015; Boucher, 2016).

Each model entails a different approach to policy making. For instance:

how does adding a new link to the existing network affect the equilibrium efforts?

- In the Local Aggregate model, all agents will change the level of effort.

- In the Local Average model, the agents will change the level of effort only when the difference between the individual effort and the group average is changed.

If the planner wants to reduce total crime, which link should she remove from the network?

In this class of local aggregate models each agent can leverage global network effects as captured by **Bonacich centrality**: *complete information about the network topology*.

Alternative:

- A node's outcome is influenced by the number of neighbours, each agent knows how many neighbors she herself has (i.e. her **degree centrality**) and the degree distribution: *incomplete information about the network topology*.
- Each individual i selects an effort $y_i \geq 0$ and obtains a payoff given by the following utility function:

$$u_i(y) = \underbrace{f(y_i + \phi \sum_{j=1}^{k} y_j)}_{\text{benefits from own and friends' effort}} - \underbrace{c(y_i)}_{\text{costs}}$$

Note: This is a model of conformist society. All agents choose an optimal level of effort, which is equal to the weighted mean of individual effort, whereby the weights are proportional to the degree of the agents in the network. This is essentially the definition of best-shot public games (Hirshleifer, 1983): when ϕ increases, the impact of the social norm on own effort increases.

Interpretation:
The parameter $\phi \in \mathbb{R}$ determines the nature of the externality across players' actions:

- $\phi < 0$: information sharing as a local public good (Bramoulle' et al., 2007), co-ordination problems (Jackson and Wolinsky, 1996).
- $\phi > 0$: vaccination (Galeotti et al., 2010), collaboration among firms (Goyal and Moraga-Gonzalez, 2001).

Result:
There is a unique Nash equilibrium of the peer effect game y_i^*,where each individual provides effort proportional to that of her/his direct connection and to her/his idiosyncratic characteristics: $y_i^* = f(\phi, d(G))$.

Models of complete information: Reduced form

We re-write (1) in matrix form:

$$Y = \phi G Y + X$$

Where Y is the level of effort solving our local average/aggregate model.

By assuming that G is invertible, it is easy to show that Y is uniquely identified:

$$\begin{aligned} Y - \phi G Y &= X \\ (I - \phi G) Y &= X \\ Y &= (I - \phi G)^{-1} X \end{aligned}$$

Hence:

$$Y = (I - \phi G)^{-1} X = X + \phi G X + \phi^2 G^2 X + \phi^3 G^3 X + \dots \tag{2}$$

Models of complete information: Reduced form

$$Y = (I - \phi G)^{-1} X = X + \phi G X + \phi^2 G^2 X + \phi^3 G^3 X + \ldots$$

Interpretation:

This class of model corresponds to a perfect information game (with quadratic utility functions) in which Eq. (2) is the best-reply function of individual i choosing action (outcome) y_i, where:

- individuals do not care about others when $\phi = 0$. There are no social spillovers among agents. Eq. (2) tells us that agents' effort is determined exclusively by the individual characteristics of agents (X).
- individuals want to conform when $\phi > 0$. There are social spillovers among connected agents. The effort of any agent depends on the characteristics of all other agents, with each agent weighted using their distance in the network.

Note: this is essentially the definition of pure strategy Nash equilibrium play with complete information, where the observed outcome of each individual i is the response to the observed outcome of individuals $-i$. Caveat: it is not always reasonable to rule out mixed strategies (Camerer, 2003), e.g. non-best mutual responses, or complete information over the network topology (Kline *et al.*, 2019).

Three important results:

1. There is a unique Nash equilibrium in which agent i's effort is equal to the vector $(I - \phi G)^{-1}(X)$, which coincides with the vector of weighted Bonacich centralities, with weights X: i.e. $Y = (I - \phi G)^{-1} X$
2. There are magnifying or *social multiplying* effects due to network relationships, as captured by the Katz-Bonacich centrality: more central agents in the network will exert more effort.
3. From a theoretical point of view, a key novelty is the fact that agents choose the optimal effort by taking as given their rational expectations of the other agents' levels of effort.

References (1/2):

- C. Ballester, A. Calvo-Armengol, and Y. Zenou (2006), "Who's who in networks. Wanted: the key player," *Econometrica*, 74:1403-1417.
- C. Ballester, A. Calvo-Armengol, and Y. Zenou (2010), "Delinquent networks," *Journal of the European Economic Association*, 8:34-61.
- Y. Bramoulle, H. Djebbari, and B. Fortin (2009), "Identification of Peer Effects through Social Networks," *Journal of Econometrics*, 150:41-55.
- Y. Bramoulle, R. Kranton, and M. D'Amours (2014), "Strategic Interaction and Networks," *American Economic Review*, 104(3):898-930.
- Blume, L.E., Brock, W.A., Durlauf, S.N., Jayaraman, R. (2015), "Linear social interactions models," *Journal of Political Economy*, 123, 444–496.
- Boucher, V., (2016), "Conformism and self-selection in social networks," *Journal of Public Economics*, 136: 30–44.
- Calvo-Armengol, A., Patacchini, E., Zenou, Y. (2009), "Peer effects and social networks in education," *The Review of Economic Studies*, 76(4): 1239-1267.
- Calvo-Armengol, A., Y. Zenou (2004), "Social networks and crime decisions. The role of social structure in facilitating delinquent behavior," *International Economic Review*, 45: 939-958.
- Camerer, C. (2003), Behavioral Game Theory: Experiments in Strategic Interaction, Princeton University Press, Princeton, NJ.
- De Giorgi, G., Pellizzari, M. and S. Redaelli (2010), "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2: 241-275.

References (2/2):

- De Marti, J., Zenou, Y., (2015), "Networks games under incomplete information", *Journal of Mathematical Economics*, 61: 221–240.
- Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., and Yariv, L. (2010), "Network games," *The review of economic studies*, 77(1): 218-244.
- Goyal, S., Moraga-Gonzalez, J. L. (2001), "R&D networks," *Rand Journal of Economics*, 32, 4, 686-707.
- Hirshleifer, J. (1983), "From weakest-link to best-shot: The voluntary provision of public goods", Public Choice, 41, 371–386.
- Jackson, M. O., and Wolinsky, A. (1996), "A strategic model of social and economic networks," *Journal of economic theory*, 71(1): 44-74.
- Kline, B., Tamer, E., (2019), "Econometric analysis of models with social interactions," in Graham, B., De Paula, A. (Eds.), The Econometric Analysis of Network Data. Elsevier, Amsterdam. Forthcoming.
- Liu, X., Patacchini, E., Zenou, Y. (2014), Endogenous peer effects: local aggregate or local average? *Journal of Economic Behavior & Organization*, 103: 39–59.
- Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economics, and Organization*, 28: 1-31.
- Topa, G., Zenou, Y. (2015), "Neighborhood and network effects," in Duranton, G., Henderson, V., Strange, W. (Eds.), "Handbook of Regional and Urban Economics," vol. 5A. Elsevier, Amsterdam, pp.561–624.
- Ushchev P., Zenou Y. (2020), "Social norms in networks," *Journal of Economic Theory*: 185, 104969.

Outline

Consider again (1) in matrix form:

$$Y = \phi G Y + X\beta$$

The reduced form is:

$$Y = (I - \phi G)^{-1}(X\beta)$$

Interpretation:
The outcome of agent i will not only be affected by the changes in the exogenous characteristics of i (direct effects), but also by the changes in the characteristics of all other agents (indirect effects).

Average Total Impact:

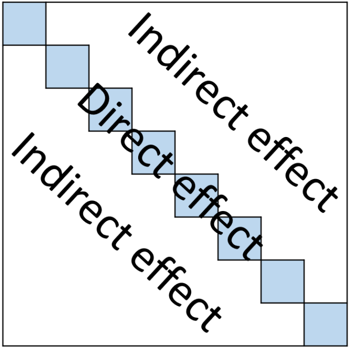
$$\bar{M}^k_{tot} = \frac{\hat{\beta}_k}{1 - \hat{\phi}}$$

Average Direct Impact:

$$\bar{M}^k_{dir} = n^{-1}\ tr[\frac{I\hat{\beta}_k}{(I - \hat{\phi}G)^{-1}_{ii}}]$$

Average Indirect Impact:

$$\bar{M}^k_{ind} = \bar{M}^k_{tot} - \bar{M}^k_{dir}$$



Example

Suppose that:

- $\hat{\beta}_k = 0.5$
- $\hat{\phi}_k = 0.397$
- $\bar{M}^k_{tot} = 0.829$
- $\bar{M}^k_{dir} = 0.586$
- $\bar{M}^k_{ind} = 0.243$

If the model is specified in levels, we say that:

- There is a positive effect of variable X_k equal to 0.5.
- There is a positive feedback effect of variable X_k equal to $\bar{M}^k_{dir} - \hat{\beta}_k = 0.086$ arising from the impact on other agents' outcome and getting back to the agent herself.
- There is a positive spillover effect arising from changes in the variable X_k of other agents equal to 0.243.

Example

Suppose that:

- $\hat{\beta}_k = 0.5$
- $\hat{\phi}_k = 0.397$
- $\bar{M}^k_{tot} = 0.829$
- $\bar{M}^k_{dir} = 0.586$
- $\bar{M}^k_{ind} = 0.243$

If the model is specified in logged levels, we can estimate impact estimates as elasticities, e.g.:

A 10% increase in X_k would result in a 8.29% ($\bar{M}^k_{tot}$) increase in y, of which:

- Around 70% comes from the direct effect with magnitude of 0.586.
- Around 30% comes from the social spillover impact estimate (e.g. 0.243).

References:

- Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Springer.
- Cliff, A.D. Ord, J.K. (1981), *Spatial Processes*, Pion.
- Gibbons S., Overman H., Patacchini E. (2015), *Spatial Methods*, Duranton G., Henderson V., Strange W. (eds), Handbook of Regional and Urban Economics, vol. 5, Elsevier.
- LeSage, J., Pace, R. K. (2009), *Introduction to spatial econometrics*, Chapman and Hall/CRC.

Outline

Outline

Assume that the utility of the agent i from effort profile y given the network structure (G) and agent attribute (X) is:

$$u_i(y, G, X) = v_i(G, X)y_i - \frac{1}{2}y_i + \phi G y_i \tag{3}$$

with $|\phi| < 1$ and $v_i(G, X)$ equal to:

$$\begin{aligned} v_i(G, X) &= A + X\beta + \bar{X}\gamma + U_i = \\ &= A + X\beta + GX\gamma + U_i \end{aligned}$$

Assume that:

- the observed effort Y corresponds to a Nash equilibrium where no agent can increase her utility by changing her effort given the efforts of all other agents in the network.
- The econometrician observes the triple (Y, X, G).
- Agents *do* observe A and U.

The utility function (3) posits the existence of two types of peer effects:

1. The marginal utility associated with changes in the average effort of one's peers (Gy):
$$\frac{\partial^2 u_i(y, G, X)}{\partial y_i \partial Gy} = \phi$$
ϕ is referred to as **endogenous peer effects**.[1]
The propensity of an individual to behave in some way varies with the behavior of the reference group.

2. The marginal utility associated with an increase in y_i is increasing with peer attributes:
$$\frac{\partial^2 u_i(y, G, X)}{\partial y_i \partial GX} = \gamma$$
γ is referred to as **contextual effects**.
The propensity of an individual to behave in some way varies with the exogenous characteristics of the reference group.

[1] If $\phi > 0$, then own- and peer-effort are complements.

The utility function (3) includes also **correlated effects**:

Agents located in an area of the network with high values of A will choose higher efforts:

$$\frac{\partial^2 u_i(y, G, X)}{\partial y_i \partial A} = 1$$

A is referred to as **correlated effects**:
Agents in the same group behave similarly because they face similar institutional environments or common shocks.

Examples:

Education

1. **Endogenous Effects**: Classmates' decisions directly affects her own decision (interdependent decisions).
2. **Contextual Effects**: Distribution of background characteristics (similar parental education, family income) lead similar behavior.
3. **Correlated Effects**: Similar environment/characteristics (teacher quality, similar activities) lead similar behavior.

Crime

1. **Endogenous Effects:** Neighbors' decisions directly affects own decisions (interdependent decisions).
2. **Contextual Effects:** Distribution of background characteristics lead to similar behavior (similar social structure of families in the neighborhood, e.g. single-family households).
3. **Correlated Effects:** Similar environment/characteristics (neighborhood quality, employment opportunities, police) lead to similar behavior, e.g. individuals in the same neighborhood choose to commit crime because they face low punishment.

The first order condition for optimal behavior associated with (3) generates the following best response function:

$$Y = A\iota + X\beta + GX\gamma + \phi GY + U \tag{4}$$

Equation (4) is called the **linear-in-means model** (Manski, 1993): agent's best reply varies with

- the average effort of those to whom she is directly connected (GY);
- her own observed attributes (X);
- the average attributes of her direct peers (GX);
- the unobserved network effect (A);
- the unobserved own attributes (U_i).

The identification problem is to recover (ϕ, β, γ).

Outline

Identification

The reduced form of the model can be written as:

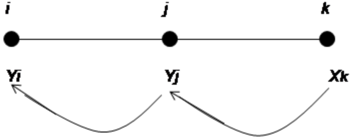
$$Y = A\iota + X\beta + GX\gamma + \phi GY + U = A\iota + X\beta + \bar{X}\gamma + \phi\bar{Y} + U \qquad (5)$$

Where $\bar{Y} = GY$ is equal to:

$$\bar{Y} = \frac{A}{1-\phi}\iota + \bar{X}\beta + [\sum_{k=0}^{\infty}\phi^k G^k]G\bar{X}(\beta\phi + \gamma) + [\sum_{k=0}^{\infty}\phi^k G^k]\bar{U} \qquad (6)$$

▸ Demonstration

We can use (6) as a vector of "first-stage" equations to instrument $\bar{Y}$ in (5).
Note: $G\bar{X} = G^2X$ (friends' of friends' characteristics) is the exclusion restriction (remember time series).



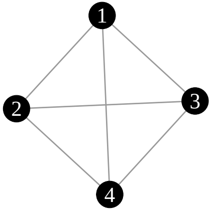
Note that (6) is identified when:

- $\beta\phi + \gamma \neq 0$. This condition is satisfied when β (contextual effects) and γ (exogenous effects) have the same sign, endogenous effects are positive ($\phi > 0$) and $\beta \neq 0$.
- Observe that (6) is equal to:

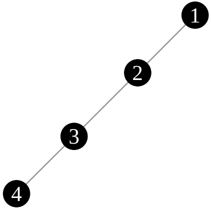
$$\bar{Y} = \frac{A(I\iota)}{(I - \phi G)} + \frac{GX\beta + G^2X\gamma}{(I - \phi G)} + \frac{\bar{U}}{(I - \phi G)}$$

  This equation is identified if (I, G, G^2) are linearly independent. Moreover, if this is true, (I, G, G^2) is a valid instrument for $\bar{Y}$.

The identifying power of intransitive triads:
(I, G, G^2) are linearly dependent if, e.g., G^2 is a linear combination of G, which is verified if the diameter of G is minor than 3: i.e. there are no intransitive triads.



$$G = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \qquad G^2 = 3*I + 2G = \begin{bmatrix} 3 & 2 & 2 & 2 \\ 2 & 3 & 2 & 2 \\ 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 3 \end{bmatrix}$$



$$G = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad G^2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Cases:

- Reflection problem (Manski, 1993): G^2 is a linear combination of G.
- Identification (Bramoulle *et al.*, 2009; Calvo-Armengol *et al.*, 2009; De Giorgi *et al.*, 2010): agents have some (but not all) common friends.

Theorem (Bramoulle *et al.*, 2009; Calvo-Armengol *et al.*, 2009)

Suppose that $|I - \phi G| \neq 0$ *and* $\beta\phi + \gamma \neq 0$. *Then:*
- Without group fixed effects, the model is identified iff I, G *and* G^2 *are linearly independent.*
- With group fixed effects, the model is identified iff I, G, G^2, *and* G^3 *are linearly independent.*

Manski (1993): it is important to *separately* identify peer or endogenous effects from contextual or exogenous effects.

Example: The Framingham study (*Connected*, by Christakis and Fowler): When a Framingham resident became obese, his or her friends were 57 % more likely to become obese, too. A Framingham resident was roughly 20 % more likely to become obese if the friend of a friend became obese . . . even if the connecting friend didn't put on a single pound.

Claim: If you're at the center of a network, you are going to be more susceptible to anything that spreads through it. However there are at least two possible explanations:

1. *Homophily*: the tendency of people to gravitate toward others who are like them: People who are gaining weight prefer to hang out with others who are also gaining weight.
2. The shared environment, and not social contagion, might be causing the people of Framingham to change in groups: McDonald's opens up causing a cluster of people living nearby to gain weight. The cluster of people appears as they are sharing a contagious form of behavior, but it is an illusion.

Note that desirable policy may be totally different, depending on the source of seemingly related behavior:

- If a social behavior is subject to *endogenous social effects*, a policy that decreases that behavior of an individual or a group of individuals will affect other individuals who were not directly targeted by the policy: i.e., the effect of the policy is *multiplied* through social interactions.
- *Contextual* changes do not imply the same multiplier effect responses to an exogenous shock.
- Failure to adequately control for *correlated effects* can lead to spurious conclusions about the importance of social influences on individual choices.

References:

- Bramoulle, Y., H. Djebbari, B. Fortin (2009), "Identification of peer effects through social networks," *Journal of econometrics* 150(1): pp. 41 – 55.
- Calvo-Armengol, A., Patacchini, E., Zenou, Y. (2009), "Peer effects and social networks in education," *The Review of Economic Studies*, 76(4): 1239-1267.
- De Giorgi, G., M. Pellizzari, S. Redaelli (2010), "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2(2): 241 - 275.
- Graham B. (2015), "Methods of identification in social networks," *Annual Review of Economics*, 7: 465 - 485.
- Manski, C. F. (1993), "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 60(3): 531 – 542.

Outline

The LIM can be correctly estimated using a **generalized spatial two-stage least square (GS-2SLS)** (Kelejian *et al.*, 1998; Lee, 2007). This will yield consistent estimates and asymptotically valid standard error estimates (see for more details De Giorgi *et al.*, 2010).

Why OLS are biased?

Consider pure first order spatial autoregressive model for simplicity:

$$Y = \phi GY + \epsilon$$

To ease notations, let $Z = GY$. The OLS estimator for ϕ is

$$\hat{\phi} = (Z'Z)^{-1}Z'Y$$

Substituting the expression for Y:

$$\hat{\phi} = \phi(Z'Z)^{-1}Z'\epsilon$$

The convergence of $\hat{\phi}$ towards the true value ϕ requires both:

$$\operatorname*{plim}_{n\to\infty} n^{-1}(Z'Z) = P \tag{7}$$

$$\operatorname*{plim}_{n\to\infty} n^{-1}(Z'\epsilon) = 0 \tag{8}$$

where P is a non zero scalar.

- Condition (7) can be satisfied with suitable restrictions on the value of ϕ and on the structure of G (see Anselin, 1988).
- Condition (8) does not hold because of the feedback effects generated by the lagged endogenous variable (Z is not an exogenous regressor here).

Demonstration: OLS inconsistency

Demonstration: OLS simultaneity bias

Outline

- H_0: There is no social spillover ($\phi = 0$). The data generation process is $Y = X\beta + \epsilon$
- H_1: There are social spillovers ($\phi \neq 0$). The data generation process is $Y = \phi GYX\beta + \epsilon$

In a ML framework, standard specification tests may be used (Anselin, 1988):

- Wald Test (W);
- Lagrange Test (LM);
- Likelihood Ratio Test (LR).

In addition, one case use the Moran's I test (Moran, 1950):

- H_0: y is randomly distributed in the network.
- H_1: y tends to cluster in the network (e.g. high values cluster near other high values; low values cluster near other low values). Warning: The data generation process can be either $Y = \phi GYX\beta + \epsilon$ or $Y = X\beta + G\epsilon$.

References:

- Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Springer.
- Cliff, A.D. Ord, J.K. (1981), *Spatial Processes*, Pion.
- Haining, R. (1990): *Spatial Data Analysis in the Social and Environmental Science*, Cambridge University Press.
- Kelejian, H. H., and Prucha, I. R. (2010), "Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances." *Journal of Econometrics*, 157(1): 53-67.
- Lee, L.-F. (2007). "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," *Journal of Econometrics*, 140 (2): 307 - 335.
- LeSage, J., Pace, R. K. (2009), *Introduction to spatial econometrics*, Chapman and Hall/CRC.
- Moran, P. A. P. (1950), "Notes on Continuous Stochastic Phenomena," *Biometrika*, 37 (1): 17-23.

Outline

In most applications, the network is endogenous: this threatens the identification.

This might happen if unobservables affect both outcomes and links:

- Outcomes are determined by a linear-in-means model.
- Links are determined by a dyadic model.
- Some unobservable η_i enters linearly into y_i and via $\eta_{i,j}$ into $g_{i,j}$.

Two different approaches:

- Joint model of outcomes and links (Battaglini et al. 2021, Goldsmith-Pinkham *et al.*, 2013).
- Two-stage correction as in Heckman (1979) (Arduini *et al.*, 2015; Battaglini *et al.*, 2020; Johnnson *et al.*, 2017; Qu *et al.* 2015).

Consider the linear-in-means model

$$Y = A\iota + X\beta + \bar{X}\gamma + \phi\bar{Y} + U$$

Assume that:

- Agents may assortatively match on some attribute and cluster together (homophily);
- The surplus generated by connections may vary with agent attributes, generating degree heterogeneity (prefential attachment)

$$g_{i,j} = \delta_0 + \sum_{l} \delta_{l+1}|x_i^l - x_j^l| + \eta_{i,j} \tag{9}$$

Note: Equation (9) assumes dyadic independence[2]: i.e., agents' choices are not influenced by others decisions - each link occurs with the same probability.

[2] see, e.g., Lai *et al.*, 2000; Fafchamps *et al.*, 2007; Mayer *et al.*, 2008; Apicella et al., 2012; Attanasio *et al.*, 2012, Graham B., 2017.

In addition to dyadic independence, we assume that:
$U = (u_1, u_2, u_3, \ldots, u_n)'$ and $\eta_i = (\eta_{i,1}, \eta_{i,2}, \ldots, \eta_{i,n})'$ are jointly normal with:

- $E(u_i^2) = \sigma_u^2$;
- $E(u_i \eta_{i,j}) = \sigma_{u\eta}^2$, $\forall i \neq j$;
- $E(\eta_{i,j}\eta_{i,k}) = \sigma_\eta^2$, $\forall j = k$;
- $E(\eta_{i,j}\eta_{i,k}) = 0$, $\forall j \neq k$.

The selection effect (i.e. the correlation between unobservable characteristics determining link formation and unobservable characteristics driving the outcome, as measured by $\sigma_{u\eta}$) is the same for all agents.

Under this assumptions, it can be shown that the expected value of the error term (u) on the link formation is

$$E(u_i|\eta_{i,1},\ \ldots,\ \eta_{i,n}) = \psi \sum_{j \neq i} \eta_{i,j}$$

Where $\psi = \frac{\sigma_{u\eta}}{\sigma_{\eta}^2}$.

It follows that our model can be re-written as:

$$Y = A\iota + X\beta + \bar{X}\gamma + \phi\bar{Y} + \psi\zeta + U$$

Where $\psi\zeta_i = \sum_{j \neq i} \eta_{i,j}$ captures the selection bias.

References:

- Apicella, C. L., F. W. Marlowe, J. H. Fowler, N. A. Christakis (2012), "Social networks and cooperation in hunter-gatherers," *Nature*, 481(7382): 497 - 501.
- Arduini T., Patacchini E., Rainone E. (2015), "Parametric and Semiparametric IV Estimation of Network Models with Selectivity," *EIEF Working Papers Series 1509*
- Attanasio, O., A. Barr, J. C. Cardenas, G. Genicot, C. Meghir (2012), "Risk pooling, risk preferences, and social networks," *American Economic Journal: Applied Economics*, 4(2): 134-167.
- Battaglini, M., V. Leone Sciabolazza, E. Patacchini (2020), "Effectiveness of connected legislators," *American Journal of Political Science* 64(4): 739-756.
- Battaglini, M., Patacchini E., Rainone E. (2021), "Endogenous Social Interactions with Unobserved Networks," *Review of Economic Studies* forthcoming.
- Fafchamps, M., F. Gubert (2007), "Risk sharing and network formation," *American Economic Review*, 97(2): 75-79.
- Goldsmith-Pinkham, P., G. W. Imbens (2013), "Social networks and the identification of peer effects," *Journal of Business & Economic Statistics*, 31(3): 253 - 264.
- Graham B. (2017), "An econometric model of network formation with degree heterogeneity," *Econometrica* 85(4): 1033 - 1063.
- Heckman, J. (1979), "Sample selection bias as a specification error," *Econometrica*, 47 (1): 153 – 61.
- Johnnson, I., H. R. Moon (2017), "Estimation of peer effects in endogenous social networks: control function approach," *arXiv:1709.10024 working paper*.
- Lai, B., D. Reiter (2000), "Democracy, politic al similarity, and international alliances, 1816-1992," Journal of Conflict Resolution, 44(2): 203-227.
- Mayer, A., S. L. Puller (2008), "The old boy (and girl) network: social network formation on university campuses," *Journal of Public Economics*, 92(1): 329 - 347.
- Qu, X., L. Lee (2015), "Estimating a spatial autoregressive model with an endogenous spatial weight matrix," *Journal of Econometrics* 184(2): 209 – 232.

Outline

Consider the equation:

$$Y = A\iota + X\beta + \bar{X}\gamma + \phi\bar{Y} + \psi\zeta + U$$

In reduced form we have:

$$Y = (I - \phi G)^{-1}(A\iota + X\beta + \bar{X}\gamma + \psi\zeta + U) \tag{10}$$

Note that ζ is a generated regressor, hence standard errors in (10) are biased.

However, no closed form solution is available.
Use the residual bootstrap procedure in spatial econometrics (e.g., Anselin, 1990), where resampling is performed on the structural errors U.[3]

[3] This requires errors to be i.i.d.

Bootstrap procedure:

1. Remove autocorrelations in the residuals:

$$(I - \hat{\phi}G) * U = U^*$$

2. Resample U^* (i.e., extraction with re-insertion) and obtain U^*_{res}.
3. Obtain predicted values:

$$\hat{Y} = (I - \hat{\phi}G)^{-1}(A\iota + X\hat{\beta} + \bar{X}\hat{\gamma} + \hat{\psi}\zeta + U^*_{res})$$

4. Estimate equation (10) using $\hat{Y}$ as a new dependent variable and store point estimates.
5. Replicate points 2 - 4 a sufficient number of times (e.g., 1,000).
6. Use the standard deviation of point estimates obtained in point 5 as standard error in equation (10).

Other popular models of network formation are:

$$g_{i,j} = \delta_0 + \sum_l \delta_{l+1}|x_i^l - x_j^l| + \delta_2\mu_i + \delta_3\mu_j + \eta_{i,j}$$

Where μ_i and μ_j are individual fixed effects. This formulation has the benefit to perfectly control for the **degree distributio**n if we are working with a cross-section dataset (Graham, 2017).

$$g_{i,j,t} = \delta_0 + \sum_l \delta_{l+1}|x_i^l - x_j^l| + \delta_2\ g_{i,j,t-1} + \delta_3 k_{i,j} + \eta_{i,j}$$

Where $g_{i,j,t-1}$ controls for connections at time $t-1$, and $k_{i,j}$ adds to the model a "network effect flavor" (Fafchamps et al., 2007) and it measures either:

- the distance between i and j when the link between them is removed.
- the number of shared connections between i and j.

Other network formation models

References:

- Anselin, L. (1990), "Some robust approach to testing and estimation in spatial econometrics," *Regional Science and Urban Economics*, 20, 141 - 163.
- Graham B. (2017), "An econometric model of network formation with degree heterogeneity," *Econometrica* 85(4): 1033 - 1063.

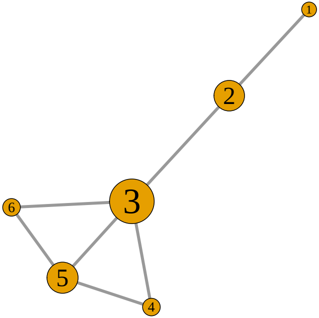
Outline

Outline

The linear-in-means model allow to estimate agent's performance as a function of individual characteristics weighted by the importance of the individual in the network.

$$Y = (I - \phi G)^{-1}(X\beta + \epsilon)$$

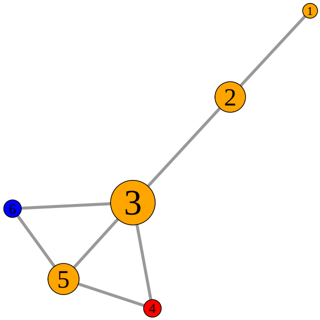
Agent's importance in the network is given by its **relative position** in the network (e.g. Bonacich centrality), rather than its structural characteristics (e.g. degree centrality).



All else being equal, the different performance of agent i and j is determined by the statuses of the nodes to which they are connected.

The LIM posits the existence of homogenous network spillovers:
Network spillover benefits in the same way (ϕ) agents with equal Bonacich centrality, but different individual characteristics.

Assumption: individual characteristics do not affect how agents "use" their connections



We can modify the LIM to account for differences in agent's characteristics.

Suppose that:

- agent 4 has characteristic z_1 (e.g., male) and
- agent 6 has characteristic z_2 (e.g., female)

We can estimate the impact of

- The connections of agent 4 interacted with x_1: e.g., $\theta_0 + \theta_1 z_1 \sum_j g_{4,j}$
- The connections of agent 6 interacted with x_2: e.g., $\theta_0 + \theta_1 z_2 \sum_j g_{6,j}$

Where θ_0 is a simple rescaling factor, and θ_1 quantifies the effect of the interaction between the adjacency matrix G and the characteristic z of the agent.

Observe that the higher is θ_1, the more agent i is able to use her/his characteristic to benefit from her/his connections: θ_1 is a measure of the extent to which agent i is able to influence her/his peers, by using characteristic z_i, and obtain higher values of y_i.

For agent i, the model is:

$$y_i = \left(1 - \sum_{j=1, j\neq 1}^{n-1} \theta_0 g_{i,j} + \sum_{j=1, j\neq 1}^{n-1} \theta_1 g_{i,j} z_i\right)\left(\beta X + \epsilon\right)$$

In matrix form, we get:

$$Y = (I - \Theta \Lambda G)^{-1}(X\beta + \epsilon)$$

Where:

- Λ is a diagonal matrix of ones.
- $\Theta = \theta_0 + \theta_1 z$ measures the extent to which the peers of an agent with a given characteristic are susceptible to her/his influence.

The same exercise can be performed to account for differences in the characteristics of agents' connections:

$$y_i = \left(1 - \sum_{j=1, j\neq 1}^{n-1} \eta_0 g_{i,j} + \sum_{j=1, j\neq 1}^{n-1} \eta_1 g_{i,j} z_j\right)\left(\beta X + \epsilon\right)$$

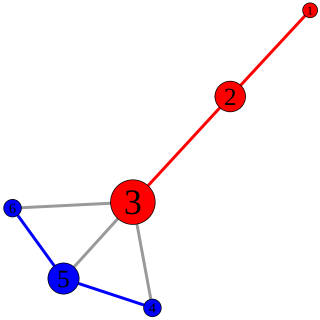
$$Y = (I - EG\Lambda)^{-1}(X\beta + \epsilon)$$

Where:

- Λ is a diagonal matrix of ones.
- $E = \eta_0 + \eta_1 z$ measures the extent to which an agent is more susceptible to the influence of her/his own peers with a given characteristic: e.g., all else being equal, the extent to which being connected to males/females matters in your performance.

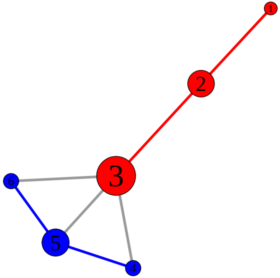
We can also consider the possibility of heterogeneous links, rather than nodes.

Specifically, we consider the case where agents belong to two different groups and interactions are different between and within groups.



In order to estimate this effect, we need to rearrange the adjacency matrix, so that:

- The block on the upper-left side registers the connections within group 1;
- The block on the lower-right side registers the connections within group 2;
- The block on the upper-right side registers the connections between group 1 and group 2;
- The block on the lower-left side registers the connections between group 2 and group 1;



$$\left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right]$$

Now define:

- G_{within} the matrix registering connections on the diagonal of G (e.g. blocks on the upper-left and lower right side).
- $G_{between}$ the matrix registering the connections out the diagonal of G (e.g. blocks on the upper-right and lower left side).

$$G_{within} = \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right] \qquad G_{between} = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right] \text{‘}$$

Note that that $G = G_{within} + G_{between}$

We can estimate:

$$Y = (I - \phi_1 G_{within} - \phi_2 G_{between})^{-1}(X\beta + \epsilon)$$

Where:

- ϕ_1 captures within groups spillovers, and
- ϕ_2 captures between groups spillovers.

Wrapping up:

We measure the extent to which the peers of an agent with a given characteristic are susceptible to her/his influence, with:

$$Y = (I - \Theta\Lambda G)^{-1}(X\beta + \epsilon) \tag{11}$$

We measures the extent to which an agent with a given characteristic is more susceptible to her/his peers influence, with:

$$Y = (I - EG\Lambda)^{-1}(X\beta + \epsilon) \tag{12}$$

We measure the effect of belonging to two different groups, when interactions are different between and within groups, with:

$$Y = (I - \phi_1 G_{within} - \phi_2 G_{between})^{-1}(X\beta + \epsilon) \tag{13}$$

Outline

We can estimate the modified LIM model, by using:

- Non-Linear Least Squares (NLLS): the basis of the method is to approximate the model by a linear one and to refine the parameters by successive iterations.
- Maximum Likelihood Estimation (MLE).

CAVEAT: in the latter case, we need to write a MLE function. This is easier than one would think.

Consider the traditional ML function of the LIM:

$$\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\Omega| - \frac{1}{2}\left[y - (I - \phi G)^{-1}X\beta\right]' \Omega^{-1}\left[y - (I - \phi G)^{-1}X\beta\right]$$

Where n is the total sample size and $\Omega = \sigma^2(I - \phi G)^{-1}(I - \phi G')^{-1}$.

In order to estimate the effect of:

- Agent i's characteristics in leveraging his/her connections, we substitute ϕG with $\Theta\Lambda G$;
- the characteristics of Agent i's connections in leveraging the connection with i, we substitute ϕG with $EG\Lambda$;
- the impact of different types of connections, we substitute ϕG with $\phi_1 G_{within} - \phi_2 G_{between}$;

References:

- Battaglini, M., V. Leone Sciabolazza, E. Patacchini (2020), "Effectiveness of connected legislators," *American Journal of Political Science* 64(4): 739-756.
- Battaglini, M., V. Leone Sciabolazza, E. Patacchini, S. Peng (2018), "econet: An R package for parameter-dependent network centrality measures," Mimeo (available at https://goo.gl/Lt1ueL).
- Battaglini, M., E. Patacchini (2018), "Influencing connected legislators," *Journal of Political Economy*, 126(6): 2277-2322.
- Dieye, R., B. Fortin (2017), "Gender Peer Effects Heterogeneity in Obesity," *Center for Interuniversity Research and Analysis on Organizations*.

Outline

The R package econet provides methods for estimating parameter-dependent network centrality measures with linear-in-means models.

Both nonlinear least squares and maximum likelihood estimators are implemented.

The methods allow for both link and node heterogeneity in network effects, endogenous network formation and the presence of unconnected nodes.

The routines also compare the explanatory power of parameter-dependent network centrality measures with those of standard measures of network centrality.

econet allows to implement the reduced form of two model behavior.

Model A: Battaglini et al. (2018)

$$y = \alpha \cdot (I - \phi G)^{-1} + X_r^\top \beta + \epsilon \tag{14}$$

Model B: Battaglini et al. (2020)

$$y = (I - \phi G)^{-1} \left(\alpha + X_r^\top \beta\right) + \epsilon \tag{15}$$

The main function of the econet package is net_dep().

Field specification in net_dep()

Model	Hypothesis	Equation	Centrality measure $\boldsymbol{b}(g,\phi)$	
Model A	lim	14	ϕ: homogeneous	$\boldsymbol{b}(g,\phi) = (I - \phi G)^{-1} 1$
	het	16	ϕ: heterogeneous by node type[4]	$\boldsymbol{b}(g,\phi) = [I - G(\phi I + \gamma\Lambda)]^{-1} 1$
Model B	lim	15	ϕ: homogeneous	$\boldsymbol{b}(g,\phi) = (I - \phi G)^{-1} 1$
	het_l	11	ϕ: heterogenous outgoing influence by node type	$\boldsymbol{b}(g,\phi) = (I - \theta\Lambda G)^{-1} 1$
	het_r	12	ϕ: heterogenous ingoing influence by node type	$\boldsymbol{b}(g,\phi) = (I - \eta G\Lambda)^{-1} 1$
	par	13	ϕ: heterogenous by link type	$\boldsymbol{b}(g,\phi) = (I - \phi_1 G_{wit} \phi_2 G_{btw})^{-1} 1$

Addional fields are detailed in the R help: i.e., ?net_dep().

[4] The extension of *model A* to the heterogenous case is:

$$y = \alpha \cdot [I - G(\phi I + \gamma\Lambda)]^{-1} + X_r^{\top} \beta + \epsilon \tag{16}$$

Example: Model A.

```
R> library("econet")
R> data("a_db_alumni")
R> data("G_alumni_111")
R> db_model_A <- subset(a_db_alumni, time == 3)
R> G_model_A <- a_G_alumni_111
R> are_factors <- c("party", "gender", "nchair", "isolate")
R> db_model_A[are_factors] <- lapply(db_model_A[are_factors],
+    factor)
R> db_model_A$PAC <- db_model_A$PAC/1e+06
R> f_model_A <- formula("PAC ~ gender + party + nchair + isolate")
R> starting <- c(alpha = 0.47325, beta_gender1 = -0.26991,
+    beta_party1 = 0.55883, beta_nchair1 = -0.17409,
+    beta_isolate1 = 0.18813, phi = 0.21440)
R> lim_model_A <- net_dep(formula = f_model_A, data = db_model_A,
+    G = G_model_A, model = "model_A", estimation = "NLLS",
+    hypothesis = "lim", start.val = starting)
R> summary(lim_model_A)
R> lim_model_A$centrality
```

Example: Model B, endogeneity.

```
R> data("db_cosponsor")
R> data("G_alumni_111")
R> db_model_B <- subset(db_cosponsor, time == 3)
R> G_model_B <- G_cosponsor_111
R> G_exclusion_restriction <- G_alumni_111
R> are_factors <- c("gender", "party", "nchair")
R> db_model_B[are_factors] <- lapply(db_model_B[are_factors] ,
+    factor)
R> f_model_B <- formula("les ~ gender + party + nchair")
R> starting <- c(alpha = 0.23952, beta_gender1 = -0.22024,
+    beta_party1 = 0.42947, beta_nchair1 = 3.09615,
+    phi = 0.40038, unobservables = 0.07714)
R> lim_model_B <- net_dep(formula = f_model_B, data = db_model_B,
+    G = G_model_B, model = "model_B", estimation = "NLLS",
+    hypothesis = "lim", endogeneity = TRUE, correction = "heckman",
+    first_step = "standard",
+    exclusion_restriction = G_exclusion_restriction,
+    start.val = starting)
R> summary(lim_model_B)
R> summary(lim_model_B, print = "first.step")
```

Example: Model A, heterogeneity by node type.

```
R> z <- as.numeric(as.character(db_model_A[, "gender"]))
R> f_het_model_A <- formula("PAC ~ party + nchair + isolate")
R> starting <- c(alpha = 0.44835, beta_party1 = 0.56004,
+    beta_nchair1 = -0.16349, beta_isolate1 = 0.21011,
+    beta_z = -0.26015, phi = 0.34212, gamma = -0.49960)
R> het_model_A <- net_dep(formula = f_het_model_A, data = db_model_A,
+    G = G_model_A, model = "model_A", estimation = "NLLS",
+    hypothesis = "het", z = z, start.val = starting)
R > summary(het_model_A)
```

Example: Model B, heterogeneity by link type.

```
R> z <- as.numeric(as.character(db_model_B[, "party"]))
R> starting <- c(alpha = 0.242486, beta_gender1 = -0.229895,
+    beta_party1 = 0.42848, beta_nchair1 = 3.0959,
+    phi_within = 0.396371, phi_between = 0.414135)
R> party_model_B <- net_dep(formula = f_model_B,
+    data = db_model_B,
+    G = G_model_B, model = "model_B", estimation = "NLLS",
+    hypothesis = "par", z = z, start.val = starting)
R> summary(party_model_B)
```

quantify(): it allows to estimate direct and indirect peer effects when "model = model_b".[5]

```
R>  quantify(fit = lim_estimate_model_B, Gn = G_model_B)
```

boot(): it provides robust standard errors when "endogeneity = TRUE"

```
R>  boot_lim_estimate <- boot(object = lim_model_B,
+    hypothesis = "lim", group = NULL, niter = 2, weights = FALSE)
R>  boot_lim_estimate
```

horse_race(): it runs a horse race across different centrality measures.

```
R> horse_model_B <- horse_race(formula = f_model_B,
+    centralities = "betweenness", directed = TRUE, weighted = TRUE,
+    normalization = NULL, data = db_model_B, G = G_model_B,
+    model = "model_B", estimation = "NLLS")
R> summary(horse_model_B)
R> summary(horse_model_B, centrality = "betweenness")
```

[5] Note that the estimated effect of network centrality when "model = model_a" is captured by the parameter α.

References:

- Battaglini, M., V. Leone Sciabolazza, E. Patacchini (2020), "Effectiveness of connected legislators," *American Journal of Political Science*, 64(4): 739-756.
- Battaglini, M., V. Leone Sciabolazza, E. Patacchini, S. Peng (2021), "econet: An R package for parameter-dependent network centrality measures," Journal of Statistical Software, forthcoming.
- Battaglini, M., E. Patacchini (2018), "Influencing connected legislators," *Journal of Political Economy*, 126, 6: 2277-2322.

Outline

Outline

Each network can be classified in terms of:

- **Structure**: the link generation process connecting the considered population (i.e. the realization of a probability distribution).
- **Composition**: the distribution of a set of attributes in the considered population.

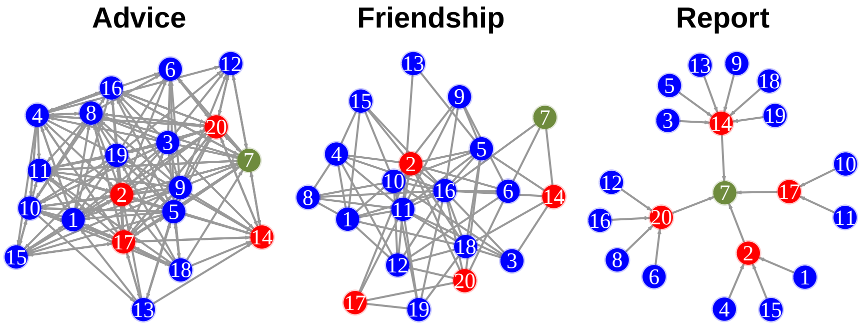
The structure of a network can be characterized:

- at the **local level**: by looking at the centrality of each agent.
- at the **global level**: by looking at the network as a whole, studying e.g. the degree distribution, the clustering coefficient, the average path length.

The composition of a network can be studied as a function of:

- **composition**: e.g. the propensity for individuals of the same sex to form partnerships (homophily).
- **topology**: e.g., the propensity for individuals to form triangles of partnerships (transitivity).
- **structure of a different network**, when agents have multiple affiliations.

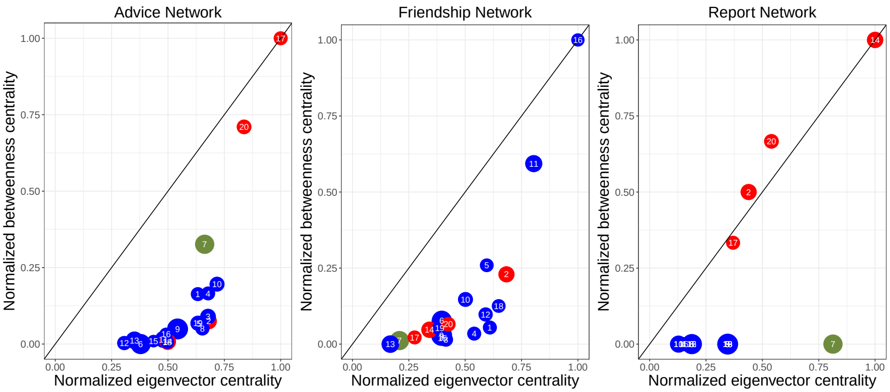
Case study: The managers of a high-tech company.



Reference:
Krackhardt, D. (1987), "Cognitive social structures," *Social networks* 9(2): pp. 109 - 134.

Network composition and local measures of connectivity

Vice-presidents are gatekeepers, the president is a diffuser leader.



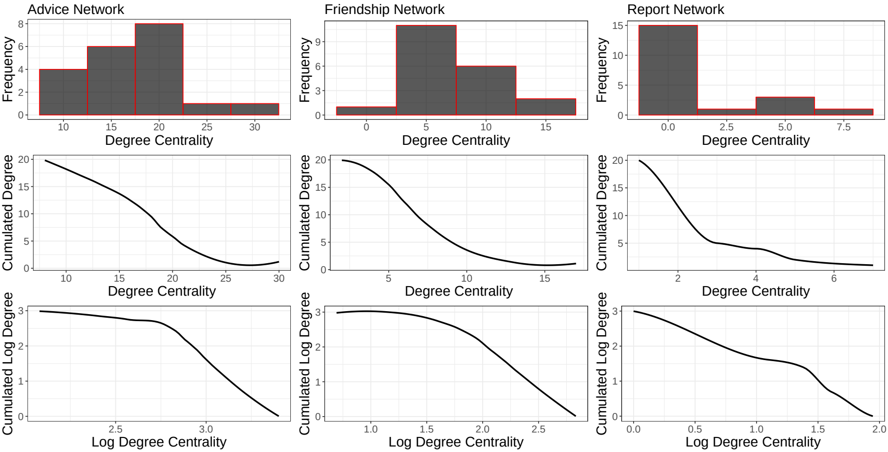
Tyschuk et al. (2015):

- gatekeeper: high betweenness, and low eigenvector. Control of information flow.
- diffuser leader: low betweenness, and high eigenvector. Unique access to gatekeepers.

Global measures of connectivity: degree centrality

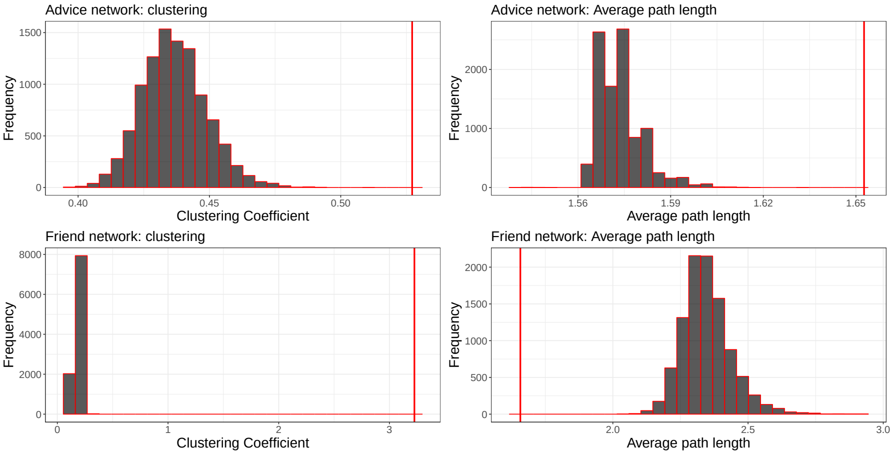
A common hypothesis is preferential attachment (power law distribution):

Mechanism of cumulative advantage where the probability of being attached to another node is proportional to the degree of that node (Barabasi et al., 1999).



Global measures of connectivity: Average Path Length (APL) and Clustering Coefficient (CC).

Another common hypothesis is that the network is a random graph:
i.e. it is generated by a Bernoulli distribution (Erdős and Rényi 1959, 1960, 1961).



Outline

The researcher has limited information to model the joint distribution of agents' linking decision: i.e. the decisions to connect.

First approach: dyadic dependence

Suppose the researcher knows that every two nodes are linked with probability p, where p is the parameter to be estimated. Two extreme cases are possible:

- any two pairs of nodes are independent of each other.
- all pairs are perfectly correlated.

We need to make assumptions on the extent to which an event, e.g. i is linked to j but k is not linked to i, influences the likelihood of some other nodes u and v being linked.

Second approach: agent heterogeneity [Previous lection]

Agents may assortatively match on some attribute and cluster together (homophily). Moreover, the surplus generated by connections may vary with agent attributes, generating degree heterogeneity (prefential attachment).

Outline

Hypothesis: any two pairs of nodes are independent of each other.

Model: the network is a Simple Random Graph (SRG).
The network is one of the many realizations of a distribution of networks consisting of: i) n nodes, ii) with each tie in the network occurring with the same specified probability (θ) (Bernoulli distribution).

Case study: asking advice to one colleague is independent of all other counseling relationships created in the network, including the ones that the two colleagues maintain with other network members.

Let the set of all possible socio-matrices be:

$$\Upsilon = \{Y : y_{i,j} \in \{0,1\}, y_{i,i} = 0\}$$

Note: Υ is the distribution from which our network is drawn.

Assumption:

The entries of Y (i.e. the connections between i and j) are independent and identically distributed: i.e. any two pairs of nodes are independent of each other: $y_{1,2}, ..., y_{n-1,n} \sim i.i.d\ binary\ (\theta)$.

Implication:

$Y = [y_{1,2}, ..., y_{n-1,n}]$ is generated by a **Bernoulli distribution**.

A SRG probability model (P) over Υ assigns a number $P(Y)$ to each $Y \in \Upsilon$:

$$0 \leq P(Y) \leq 1 \; \forall Y \in \Upsilon \; ; \; \sum_{Y \in \Upsilon} P(Y) = 1; \; Y \in \Upsilon \sim i.i.d.$$

The probability of a graph $Y \in \Upsilon$ is given by a binomial distribution (i.e., the sum of i.i.d. Bernoulli realizations):

$$P_\theta(Y) = \prod_{i \neq j} \theta^{y_{i,j}} (1-\theta)^{1-y_{i,j}} = \theta^{\sum y_{i,j}} (1-\theta)^{\sum (1-y_{i,j})}$$

Put differently, the probability of a graph $Y \in \Upsilon$ is given by:

- The probability of observing a certain number of links ($\sum y_{i,j}$) in Y: θ; and
- The probability of not observing a certain number of links ($1 - \sum y_{i,j}$) in Y: $1-\theta$.

Depending on the value of θ, the SRG model will produce a distribution of networks Υ, and we can test what is the probability that Y is a realization of Υ.

To this purpose, we define all the network distributions Υ generated by the SRG model as:

$$P = \{p(Y|\theta) : \theta \in \Theta\}$$

where:

- θ is an unknown parameter;
- Θ is the space parameter of θ (i.e. $[0, 1]$);
- $p(Y|\theta) = \mathit{binomial}(\theta)$ is one of the many networks generated by Υ: i.e. an i.i.d. realization of the binary model $P_\theta(Y) = \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})}$.

Define the observed network (e.g. the advice network) as $\mathbf{Y}$.

How can we test the hypothesis that $\mathbf{Y}$ is generated by a SRG model: i.e. $\mathbf{Y}$ is the realization of a distribution Υ?

Step 1:

- Comparing the observed data ($\mathbf{Y}$) with every SRG distribution in P [This is computationally unfeasible]
- Finding a value of θ such that $P_\theta(Y)$ generates a network similar to $\mathbf{Y}$ (ML approach)

Step 2:

- Test the extent to which observed data resembles the networks generated by the chosen SRG distribution: test of goodness of fit (GOF).

Step 1

Define the probability of occurrence of one link in **Y** as $m\bar{y}$, where m is the number of pairs and $\bar{y}$ is density of **Y**, respectively.

$$p(Y|\theta) = \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})} = \theta^{m\bar{y}}(1-\theta)^{m(1-\bar{y})}$$

In logs, we have:

$$log\ \ p(Y|\theta) = m\bar{y}\ log\ \theta + m(1-\bar{y})\ log\ (1-\theta)\ = \ m[\bar{y}\ log\ \theta + (1-\bar{y})\ log\ (1-\theta)]$$

Using a maximum likelihood approach, we now need to estimate the parameter θ such that the probability to observe $log\ \ p(Y|\theta)$ under this relation is maximized. Recalling that the maximum occurs where the derivative (slope) is zero, we have:

$$\frac{d}{d\theta} log\ \ p(Y|\theta)\ =\ m[\frac{\bar{y}}{\theta} - \frac{1-\bar{y}}{1-\theta}]\ = 0$$

This condition occurs if:

$$\frac{\bar{y}}{1-\bar{y}} = \frac{\theta}{1-\theta} \tag{17}$$

Equation 17 is verified when $\theta = \bar{y}$.

Hence, by the maximum likelihood criterion, the distribution of networks Υ with realizations

$$P = p(Y|\theta) = binomial(\theta) : \theta \in [0, 1]$$

that is closest to $\mathbf{Y}$ is the distribution:

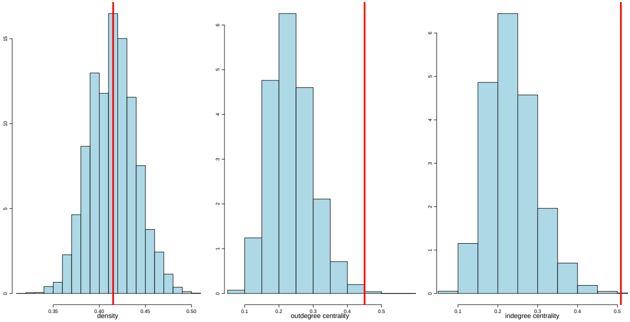
$$p(Y|\theta) = binomial(\bar{y})$$

Among all the possible distributions (graphs) generated by the SRG model ($Y \in \Upsilon$), we choose the one generating links with probability $\bar{y}$.

Intuition: $\bar{y}$ is the rate of occurence of links in $\mathbf{Y}$

Step 2 (GOF)

Monte Carlo methods are used to implement a t-test comparing the graphs generated by $p(Y|\theta) = \textit{binomial}(\bar{y})$ with **Y**.



H0: The advice network (**Y**) is an SRG.

Outline

The Row-Column Effect (RCE) model

It is a class of tie-focused logit models (which is also called two-way ANOVA).

$$Pr(Y_{i,j} = 1|\mu, \alpha_i, \beta_j) = \frac{e^{\mu+\alpha_i+\beta_j}}{1 + e^{\mu+\alpha_i+\beta_j}}$$

Where α measures agent **sociability**, and β measures agent **attractiveness**.

Odds ratios for tie preferences

Hypothesis testing:
How does the "best" RCE model compare to the data?

<u>Step 1</u>

The RCE model $Pr(Y|\mu, \alpha, \beta)$ that maximizes the fit with **Y** is

$$\hat{\theta} = \hat{\mu} + \hat{\alpha} + \hat{\beta}$$

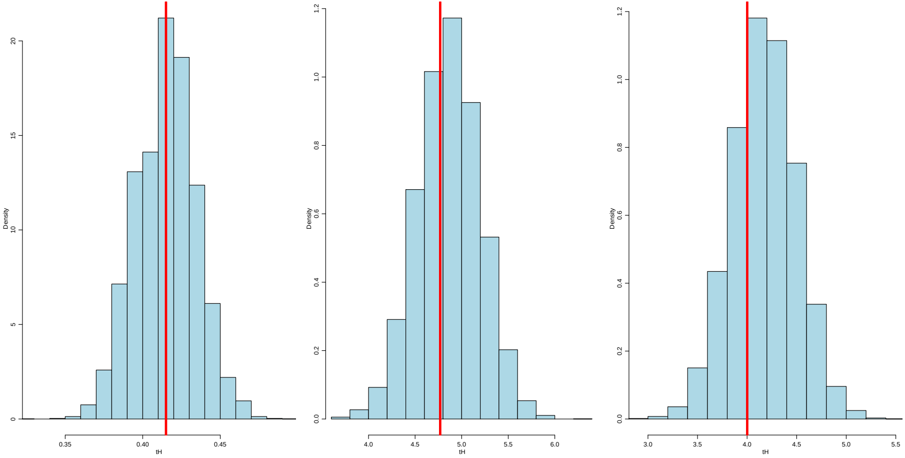
Where $\hat{\mu}, \hat{\alpha}, \hat{\beta}$ are the values estimated by the logistic regression.

<u>Note</u>:

Logistic regression estimates will be unbiased **only if** the assumption **dyadic independence** holds: i.e.,
Any two pairs of nodes are independent of each other (a test is provided by Pan et al. 2015).

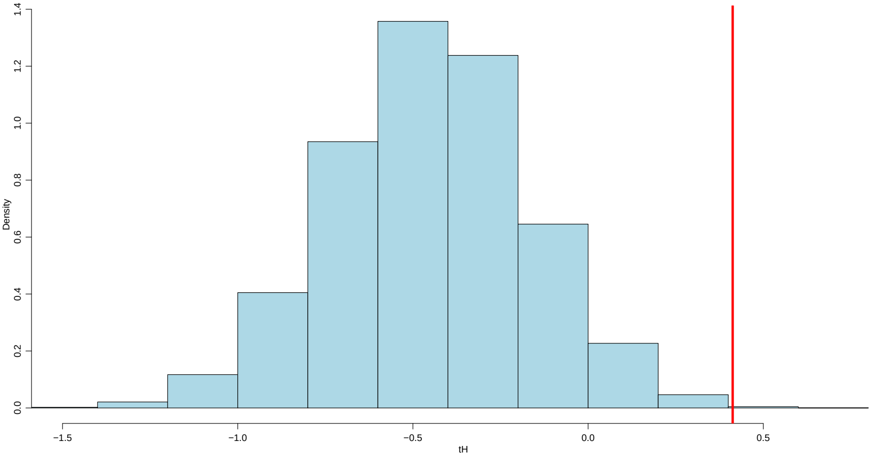
Step 2 (GOF)

Metrics: Density, In-degree, Out-degree.



Step 2 (GOF)

Metrics: Mutuality.



Why RCE doesn't recover mutuality effects?
Mutuality is not a parameter of our model.

Let's rearrange our equation.

$$Pr(Y|\mu,\alpha,\beta) = \prod_{i\neq j} \frac{e^{(\mu+\alpha_i+\beta_j)y_{i,j}}}{1+e^{(\mu+\alpha_i+\beta_j)y_{i,j}}} = exp(\mu y_{..} + \sum_i \alpha_i y_{i,.} + \sum_j \beta_j y_{.,j}) = exp[t(y)\cdot\theta]g(\theta)$$

where:

- $y_{..}$ represents all dyads,
- $y_{i,.}$ represents out degrees (sociability);
- $y_{.,j}$ represents in degrees (attractivity);
- $t(y) = (y_{.,.}, y_{i,.}, y_{.,j})$ is a vector of statistics;
- $\theta = (\mu,\ \alpha_{1,.},\ldots,\ \alpha_{n,.},\ \beta_{.,1},\ldots,\ \beta_{.,n})$ is a vector of parameters.

Important result: The RCE model is nested in the exponential family:

▸ Hammersley and Clifford theorem

Outline

So far, we have assumed that any two pairs of nodes are independent of each other.

Now, we hypothesize that there is a higher probability for j to be connected with i if i is connected to j (mutuality):

$$Pr(Y_{i,j} = y_{i,j} \bigcap Y_{j,i} = y_{i,j}) = Pr(Y_{i,j} = y_{i,j} | Y_{j,i} = y_{i,j}) Pr(Y_{j,i} = y_{i,j})$$

This can be tested using a **p1 model** (Holland et al., 1981).

A p1 model can be thought as a RCE model that includes a term for mutuality, even though the two models are non nested.

Note: also p1 is nested in the exponential family model.

The P1 model

$$p(y_{i,j}, y_{j,i}|\mu, \alpha_i, \beta_j, \gamma) = \frac{e^{\mu_{i,j}y_{i,j}+\mu_{j,i}y_{j,i}+\gamma y_{i,j}y_{j,i}}}{1 + e^{\mu_{i,j}} + e^{\mu_{j,i}} + e^{\mu_{i,j}+\mu_{j,i}+\gamma}}$$

Where $\mu_{i,j} = \mu + \alpha_i + \beta_j$.

Sufficient statistics for p1 are:

- the total number of edges (μ);
- outdegrees ($y_{j,i}$) and indegrees ($y_{i,j}$);
- total mutual dyads ($y_{i,j}y_{j,i}$).

Computational issues:

- logistic regression estimates are not reliable anymore: observations are not i.i.d.

The details for estimating a p1 model are in Robins et al. (2007).

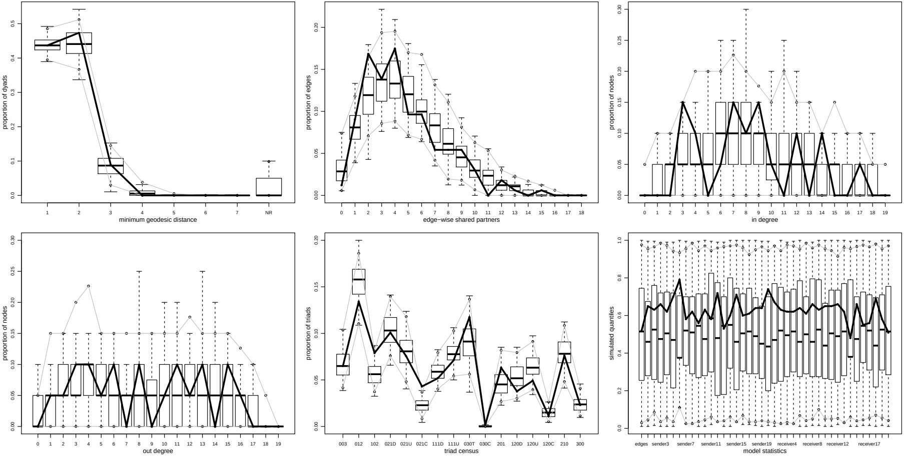
Intuition

The P1 estimation method is similar to the ad hoc approach: i.e.,

- The observed graph is compared against a distribution of random graphs generated by a stochastic process parameterized according to pre-defined parameter values.
- These values are chosen in each iteration of the estimation algorithm according to a maximum likelihood criterion, so to ensure that the generated networks are each time closer to the observed data.
- The quality of the estimates are once again evaluated by t-statistics.

A MCMC algorithm repeatedly performs these steps until t-statistics are close to 0.

Step 2 (GOF)



Triad Census

Outline

We can make hypotheses beyond the dyadic-level.

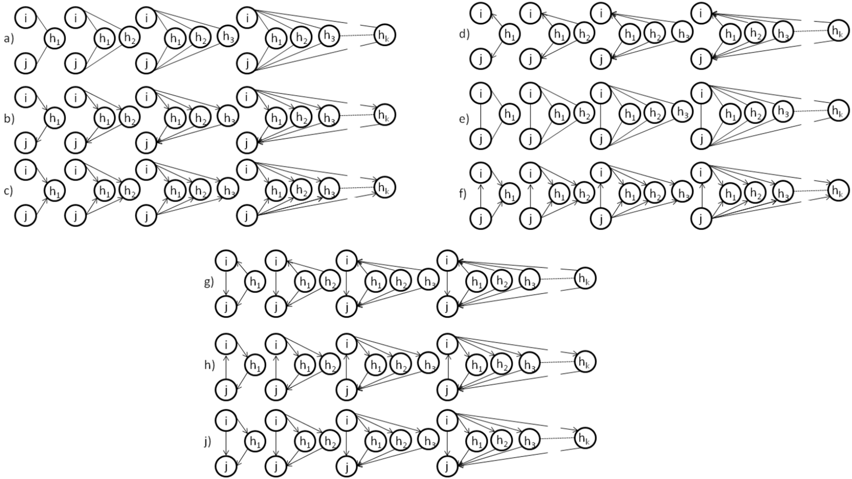
Markov Dependence Assumption (Frank and Strauss, 1986)

- nodes are seen as tie variables.
- each tie-variables is assumed to be independent conditional on the rest of the graph.

Example:
if node i connects the possible edges $(i;j)$ and $(i;k)$, we say that the tie-variables corresponding to $(i;j)$ and $(i;k)$ are independent conditional on the rest of the graph: i.e.

- if agent 1 refers to agent 2 may depend on whether agent 1 refers to 3.
- the probability of agent 2 to refer to agent 3, may depend on whether 2 and 3 both refer to agent 1.

Realization-Dependent Assumption (Baddeley et al. 1989, Pattison et al. 2002)
It generalizes the notion of conditional independence: i.e.
Two variables are statistically independent given the state of a third variable, even when do not share a node.



Exponential Random Graph Models (**ERGMs**) can be used to model these assumptions.

Each network tie is regarded as a random variable: i.e., the model is not going to make perfect deterministic predictions.

A dependence hypothesis has to be proposed to define contingencies among the network variables (e.g. mutuality, transitivity, etc.).

The dependence hypothesis implies a particular form to the model, which is described by different count statistics (e.g. number of mutual dyads, triangles, k-stars).

The general form for an ERGM can be written as:

$$P(Y = y) = \frac{\exp(\theta' g(y))}{k(\theta)}$$

where:

- Y is the random variable for the state of the network (with realization y);
- $g(y)$ is a vector of model statistics for network y;
- θ is the vector of coefficients for those statistics; and
- $k(\theta)$ represents the quantity in the numerator summed over all possible networks (typically constrained to be all networks with the same node set as y).

ERGMs relies on the **Hammersley & Clifford theorem**:
Every network can be expressed in the exponential family using count of graph statistics:

Note:
Parameters are simplified through homogeneity or other constraints: e.g., we assume the same mutual effect across the entire network.

The general form for an ERGM can be re-expressed in terms of the conditional log-odds of a single tie between two actors:

$$\text{logit}\left(Y_{ij} = 1 | y_{ij}^c\right) = \theta' \delta(y_{ij})$$

Where:

- Y_{ij} is the random variable for the state of the actor pair i, j (with realization y_{ij});
- y_{ij}^c signifies the complement of y_{ij}, i.e. all dyads in the network other than y_{ij};
- The vector $\delta(y_{ij})$ contains the change statistic for each model term.

The change statistic records how $g(y)$ term changes if the y_{ij} tie is toggled on or off. So:

$$\delta(y_{ij}) = g(y_{ij}^{+}) - g(y_{ij}^{-})$$

Where

- y_{ij}^{+} is defined as y_{ij}^{c} along with y_{ij} set to 1; and
- y_{ij}^{-} is defined as y_{ij}^{c} along with y_{ij} set to 0.

That is,

- $\delta(y_{ij})$ equals the value of $g(y)$ when $y_{ij} = 1$ minus the value of $g(y)$ when $y_{ij} = 0$;
- but all other dyads are as in $g(y)$.

The model terms $g(y)$ in an ERGM are network statistics, functions of network configurations that we hypothesize may be more or less common than what would be expected in a simple random graph (where all ties have the same probability).

The coefficient θ can be interpreted as the log-odds of an individual tie conditional on all others.[6]

Note 1:
Unlike traditional covariates in a linear model, network statistics are not exogenous measures, they are functions of the network itself, typically sums of dyad states or products of dyad states that represent specific configurations.

Note 2:
Also in this case, a test of GOF is required.

[6] Probabilities are obtained from log odds computing $exp(\theta)/(1 + exp(\theta))$.

Goodness of fit:
ERGMs can be seen as generative models where

- local processes govern the formation of micro-configurations (e.g. a triangle);
- Locally generated processes aggregate up to produce characteristic global network properties, even though these global properties are not explicit terms in the model.

GOF tests how well local processes reproduce the observed global network properties that are not in the model (Hunter et al., 2008).

Standard global network properties checked by GOF are:

- Degree distribution;
- Minimum geodesic distance;
- Edgewise shared partner distributions (Hunter et al., 2008);
- Triad census distribution.

References:

- Baddeley, A. and J. Møller (1989). "Nearest-neighbour Markov point processes and random sets," *International Statistical Review/Revue Internationale de Statistique*, pp. 89–121.
- Barabasi, A.-L., R. Albert (1999), "Emergence of scaling in random networks," *Science* 286(5439): pp. 509 – 512.
- Frank, O., and D. Strauss (1986), "Markov graphs," *Journal of the American Statistical Association*, 81(395), pp. 832–842.
- Holland, P. W. and S. Leinhardt (1981), "An exponential family of probability distributions for directed graphs," *Journal of the American Statistical association*, 76(373), pp. 33–50. Hunter, D. R., S. M. Goodreau, M. S. Handcock (2008), "Goodness of fit of social network models," *Journal of the American Statistical Association*, 103(481).
- Lehmann, E. L. and J. P. Romano (2006), *Testing statistical hypotheses*, Springer.
- Lusher, D., J. Koskinen, and G. Robins (2012), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press.
- Newman, Mark E J (2003). "Ego-centered networks and the ripple effect," *Social Networks* 25(1): pp. 83 – 95.
- Pan, R. and H. Wang (2015), "A note on testing conditional independence for social network analysis," *Science China Mathematics*, 58 (6), pp. 1179 – 1190.
- Pattison, P. and G. Robins (2002), "Neighborhood–based models for social networks," *Sociological Methodology*, 32(1), pp. 301–337.
- Robins, G., P. Pattison, Y. Kalish, D. Lusher (2007), "An introduction to exponential random graph (p*) models for social networks," Social networks, 29(2), pp. 173–191.
- Robins, G., T. Snijders, P. Wang, M. Handcock, and P. Pattison (2007), "Recent developments in exponential random graph (p*) models for social networks," Social networks, 29(2), pp. 192–215
- Valente, T. W., K. Coronges, C. Lakon, and E. Costenbader (2008), "How correlated are network centrality measures?," *Connections* 28(1).

Valerio Leone Sciabolazza
valerio.leonesciabolazza@uniparthenope.it
www.valerioleonesciabolazza.com

Outline

Define:

- G a row-normalized matrix $n * n$;
- ι a vector of ones $1 * n$;
- ϕ a scalar $\in (-1, 1)$.

Rememeber that:

- $(I - \phi G)^{-1} = \sum_{k=0}^{\infty} \phi^k G^k$.
- $G\iota = \iota$, and $G^k \iota = \iota$.
- $A(I - \phi G)^{-1}\iota = A[\sum_{k=0}^{\infty} \phi^k G^k]\iota = A(1 + \phi + \phi^2 + \phi^3 + \dots) = \frac{A}{1-\phi}\iota$

Re-arrange the structural equation:

$$
\begin{aligned}
Y &= A\iota + X\beta + GX\gamma + \phi GY + U = \\
&= A(I - \phi G)^{-1}\iota + (I - \phi G)^{-1}(X\beta + GX\gamma) + (I - \beta G)^{-1} U = \\
&= \frac{A}{1-\phi}\iota + \sum_{k=0}^{\infty} \phi^k G^k X\beta + GX\gamma + [\sum_{k=0}^{\infty} \phi^k G^k] U = \\
&= \frac{A}{1-\phi}\iota + \sum_{k=0}^{\infty} \phi^k G^k X\beta + \sum_{k=0}^{\infty} \phi^k G^{k+1} X\gamma + [\sum_{k=0}^{\infty} \phi^k G^k] U = \\
&= \frac{A}{1-\phi}\iota + X\beta + \sum_{k=1}^{\infty} \phi^k G^k X\beta + \sum_{k=0}^{\infty} \phi^k G^{k+1} X\gamma + [\sum_{k=0}^{\infty} \phi^k G^k] U = \\
&= \frac{A}{1-\phi}\iota + X\beta + \sum_{k=0}^{\infty} \phi^k G^{k+1} X\beta\phi + \sum_{k=0}^{\infty} \phi^k G^{k+1} X\gamma + [\sum_{k=0}^{\infty} \phi^k G^k] U = \\
&= \frac{A}{1-\phi}\iota + X\beta + \sum_{k=0}^{\infty} \phi^k G^{k+1} X(\beta\phi + \gamma) + [\sum_{k=0}^{\infty} \phi^k G^k] U
\end{aligned}
$$

Define:

- $GX = \bar{X}$ and $G^2X = G\bar{X}$.

Observe that:

- $\sum_{k=0}^{\infty} G^{k+1}X = \bar{X} + \sum_{k=1}^{\infty} G^k\bar{X}$

Then:

$$Y = \frac{A}{1-\phi}\iota + X\beta + [\sum_{k=0}^{\infty}\phi^k G^{k+1}X](\beta\phi+\gamma) + [\sum_{k=0}^{\infty}\phi^k G^k]U$$

Multiplying by G:

$$\bar{Y} = \frac{A}{1-\phi}\iota + \bar{X}\beta + [\sum_{k=0}^{\infty}\phi^k G^{k+1}\bar{X}](\beta\phi+\gamma) + [\sum_{k=0}^{\infty}\phi^k G^k]\bar{U} =$$

$$= \frac{A}{1-\phi}\iota + \bar{X}\beta + [\sum_{k=0}^{\infty}\phi^k G^k]G\bar{X}(\beta\phi+\gamma) + [\sum_{k=0}^{\infty}\phi^k G^k]\bar{U}$$

Back to slides

Outline

Let us show that condition (8) does not hold.

Consider the reduced form:

$$Y = (I - \phi G)^{-1}\epsilon$$
$$Z = GY = G(I - \phi G)^{-1}\epsilon$$

Condition (8) can be written as:

$$\operatorname*{plim}_{n \to \infty} n^{-1}(Z^{'}\epsilon) = \operatorname*{plim}_{n \to \infty} n^{-1}\epsilon^{'} G(I - \phi G)^{-1}\epsilon \neq 0$$

The presence of the spatial matrix G results in a quadratic form in the error term ϵ. Therefore, except in the trivial case in which $\phi = 0$, the plim expression will not equal to zero.

Back to slides

Outline

Because of the correlation in the error term ($E(\epsilon_i, \epsilon_j) \neq 0$, $var(\epsilon_i, \epsilon_j) \neq 0$), the OLS will suffer from the simultaneity bias: i's outcome (y) is a linear function of j-s' error term (ϵ).

Let us calculate the simultaneity bias in a simple case.
Consider three individuals, we can represents the matrix form as a system of simultaneous equations as follows:

$$\begin{cases} y_1 = \phi_1 \frac{y_2+y_3}{2} + \epsilon_1 \\ y_2 = \phi_2 \frac{y_1+y_3}{2} + \epsilon_2 \\ y_3 = \epsilon_3 \end{cases}$$

Now, substitute the first and third equation into the second one:

$$y_2 = \phi_2 \frac{\phi_1 \frac{y_2+y_3}{2} + \epsilon_1 + y_3}{2} + \epsilon_2$$

$$y_2 = \phi_2 \frac{\phi_1 \frac{y_2+\epsilon_3}{2} + \epsilon_1 + \epsilon_3}{2} + \epsilon_2$$

This yields to the reduced form:

$$y_2(1 - \phi_1\phi_2) = \phi_2 \frac{\phi_1 \frac{\epsilon_3}{2} + \epsilon_1 + \epsilon_3}{2} + \epsilon_2$$

$$y_2 = \frac{\phi_2 \frac{\phi_1 \frac{\epsilon_3}{2} + \epsilon_1 + \epsilon_3}{2} + \epsilon_2}{(1 - \phi_2\phi_1)}$$

If we add y_3 and divide all by 2, we have:

$$\frac{y_2 + y_3}{2} = y_3 + 2\,\frac{\phi_2\phi_1\frac{\epsilon_3}{2} + \epsilon_1 + \epsilon_2 + \epsilon_3}{(1 - \phi_1\phi_2)} + \epsilon_2$$
$$= \epsilon_3 + 2\,\frac{\phi_2\phi_1\frac{\epsilon_3}{2} + \epsilon_1 + \epsilon_2 + \epsilon_3}{(1 - \phi_1\phi_2)} + \epsilon_2$$

Now, note that $E(\frac{y_2+y_3}{2}, \epsilon_1)$ is equal to:

$$E\left(\epsilon_3 + 2\,\frac{\phi_2\phi_1\frac{\epsilon_3}{2} + \epsilon_1 + \epsilon_2 + \epsilon_3}{(1 - \phi_1\phi_2)} + \epsilon_2, \epsilon_1\right) = \frac{1}{(1 - \phi_1\phi_2)}\sigma_\epsilon^2$$

Since $E(\epsilon_i, \epsilon_j) \neq 0$.[7]

Back to slides

[7] Remember that $\sigma_\epsilon^2 = Var(\epsilon) > 0$

Outline

We use a frequentist approach to infer whether agents' characteristics determined the characteristics of whom they choose as friends.

$$
\begin{aligned}
Pr(x_j = 1|y_{i,j} = 1, x_i = 1) &= \frac{Pr(y_{i,j} = 1|x_j = 1, x_i = 1)Pr(x_j = 1|x_i = 1)}{Pr(y_{i,j} = 1|x_i = 1)} \\
&= \frac{Pr(y_{i,j} = 1|x_j = 1, x_i = 1)Pr(x_j = 1)}{Pr(y_{i,j} = 1|x_i = 1)} \\
&= p_{11}\frac{Pr(x_j = 1)}{Pr(y_{i,j} = 1|x_i = 1)}
\end{aligned}
$$

Where $p_{1,1}$ can be interpreted as, e.g., "what is the probability that two nodes of level 3 (managers) are friends?".

In terms of odds, we have:

$$odds(x_j = 1|y_{i,j} = 1, x_i = 1, y_{i,j} = 1, x_i = 1) = \frac{Pr(x_j = 1|y_{i,j} = 1, x_i = 1)Pr(x_j = 0|y_{i,j} = 1, x_i = 0)}{Pr(x_j = 0|y_{i,j} = 1, x_i = 1)Pr(x_j = 1|y_{i,j} = 1, x_i = 0)}$$

Now recalling that:

- $\mathrm{Pr}(x_j = 1|y_{i,j} = 1, x_i = 1) = p_{11}[Pr(x_j = 1)/Pr(y_{i,j} = 1|x_i = 1)]$
- $\mathrm{Pr}(x_j = 0|y_{i,j} = 1, x_i = 1) = p_{10}[Pr(x_j = 0)/Pr(y_{i,j} = 1|x_i = 1)]$

We have that:

$$odds(x_j = 1|\{y_{i,j} = 1, x_i = 1\}, \{y_{i,j} = 1, x_i = 1\}) = \frac{p_{11}}{p_{10}} \frac{Pr(x_j = 1)}{Pr(x_j = 0)}$$

Where $\frac{p_{11}}{p_{10}}$ is an absolute ratio revealing tie preferences.

Odd ratios can be used to obtain network characteristics:

- Attractiveness (indegree) of agents with attribute $x = 1$:

  $\frac{p_{01}}{p_{00}} = \frac{odds(x_j=1|y_{i,j}=1, x_i=0)}{odds(x_j=1)}$

- Sociability (outdegree) of agents with attribute $x = 1$:

  $\frac{p_{10}}{p_{00}} = \frac{odds(x_j=0|y_{i,j}=1, x_i=1)}{odds(x_j=1)}$

- Preference for homophily:

  $\frac{p_{11}p_{00}}{p_{10}p_{01}} = Odds\ ratio(x_j = 1|\{y_{i,j} = 1, x_i = 1\}, \{y_{i,j} = 1, x_i = 0\}) = \gamma$

  This ratio represents the relative preference of node j with $x = 1$ versus $x = 0$ to tie to nodes with $x = 1$.

Now consider a logistic regression model, where the probability for i and j to be connected is a function of their characteristics:

$$Pr(y = 1|x_1, x_2) = \frac{e^{\beta_0+\beta_1 x1+\beta_2 x_2+\beta_{12}x_1x_2}}{1 + e^{\beta_0+\beta_1 x1+\beta_2 x_2+\beta_{12}x_1x_2}}$$

In terms of odds, we get:

$$odds(y = 1|x_1, x_2) = exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12}x_1x_2)$$

This model summarizes the statistics previously mentioned, since e.g.:

- $odds(y = 1|0, 0) = exp(\beta_0)$;
- $odds(y = 1|1, 0) = exp(\beta_0 + \beta_1)$.

Then:

- $odds\ ratio(y = 1|(1, 0), (0, 0)) = \frac{exp(\beta_0+\beta_1)}{exp(\beta_0)} = \exp(\beta_1)$: i.e. $\frac{p_{10}}{p_{00}}$ (sociability);

In log odds:

$$log\ odds(y = 1|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

Given that $\exp(\beta_1) = ratio(y = 1|(1,0),(0,0))$, in logistic regression:

- β_1 is the "effect" of x_1, it represents the log odds ratio

  $(y = 1|(1,0),(0,0))$ (sociability)
- β_2 is the "effect" of x_2, it represents the log odds ratio

  $(y = 1|(0,1),(0,0))$ (attractiveness)
- β_{12} is the "effect" of x_2, it represents the log odds ratio

  $\frac{(y=1|(1,1),(0,1))}{(y=1|(1,0),(0,0))}$ (homophily)

A tie-focused approach that exploit this feature of logit models is the RCE model (which is also called two-way ANOVA), where x_1 and x_2 are coded as fixed effects (the reference category is excluded), and they are constrained so that $\sum x_1 = \sum x_2 = 0$. Back to slides

Outline

Every network can be expressed in the exponential family using count of graph statistics.

We have already seen using our ad hoc approach for random graphs, the probability of occurrence of G can be expressed as:

$$\begin{aligned} P(G) &= p^{L(g)}(1-p)^{\frac{n(n-1)}{2-L(G)}} \\ &= (\frac{p}{1-p})^{L(G)}(1-p)^{\frac{n(n-1)}{2}} \\ &= exp\Big[log\Big(\frac{p}{1-p}\Big)L(G) - log\Big(\frac{1}{1-p}\Big)\frac{n(n-1)}{2}\Big] \end{aligned}$$

Note that:

- $\left(\frac{1}{1-p}\right)\frac{n(n-1)}{2}$ is a constant not involving the number of links;
- $L(G)$ is exactly the number of links of G, that is the summary statistics of the graph.

The probability of G can be expressed as a function of its summary statistics.

Assume that:

- $L(G) = s_1(G)$;
- $\left(\frac{1}{1-p}\right)^{\frac{n(n-1)}{2}} = c$; and
- $log\,\frac{p}{1-p} = \beta_1$

Then, we can rewrite the function in this way:

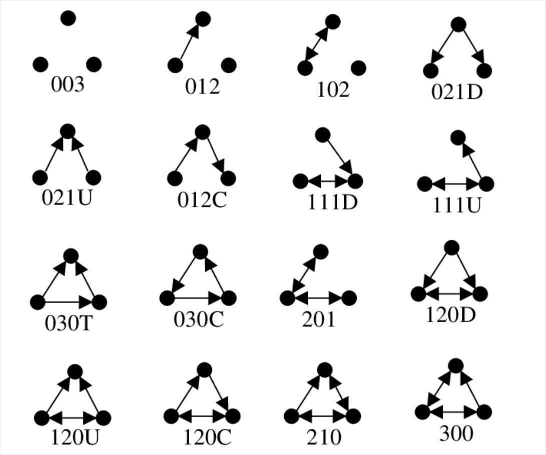
$$P(G) = exp\left[\beta_1 s_1(G) - c\right]$$

As a result, we have expressed our model in terms of an exponential distribution.

Back to slides

Outline

Triad census



Back to slides