NCRM
National Centre for Research Methods

PEPA Programme Evaluation for Policy Analysis

Identifying social effects from policy experiments

Arun Advani (UCL & IFS) and Bansi Malde (IFS & UCL)

PEPA is based at the IFS and CEMMAP

Institute for Fiscal Studies

cemmap centre for microdata methods and practice

Introduction

- Individuals, households, firms, countries, etc are **linked** with one another through kinship, social and transactional ties
- Network = Map of these interactions between these units
- Such networks play an important role in many types of interactions:
  - Information transmission
  - Trade and exchange
  - Influence preferences
- Consequently shape the beliefs, preferences and constraints of economic agents → Affect socioeconomic outcomes

Introduction

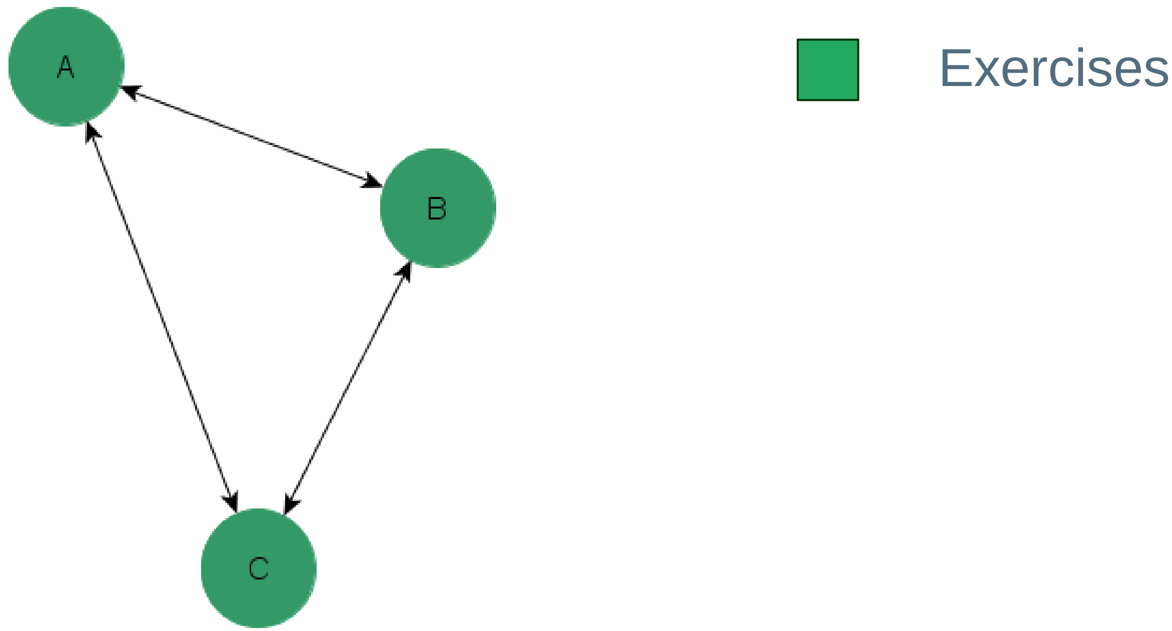
- Understanding and quantifying the effects of networks is of great interest, to academics and policymakers
- Of pasrticular interest is the effect of **social networks** on socio-economic outcomes
  - **Social network** = Links between individuals or households
- Refer to the effect of the social network on outcomes as a **social effect**
  1. For example, influence of the **average** behaviour of an individual's friends on the individual's own behaviour
  2. Or effect of the **total** behaviour of an individual's friends on the individual
  3. Or effect of individual's proximity to central individuals in the network on his outcomes
- Focus on social effects of type (1) above in this talk

Introduction

- This type of social effect is relevant when research question of interest is of the type:
  - Are teenagers more likely to smoke if their friends smoke?
  - Is an individual more likely to exercise if her friends exercise?

- Not very straightforward to obtain causal estimates of these effects

An Example



Did A and B’s exercising influence C to exercise?

Challenges in answering this question

- Did A and B influence C, or did C influence A and B?

- There could be some unobserved factor influencing A, B and C to exercise (e.g. they live in the same neighbourhood and a gym has just opened up)

- Or A, B and C all like exercising and became friends because they like to exercise; OR they are all very social, which influenced them to be become friends AND to exercise
  - There have similar unobserved preferences

Policy Experiments

- Policy experiments offer one way of resolving some of these issues

- Policy experiments offer a policy or programme to some units in a manner that is **random** or **close to random (quasi-random)**

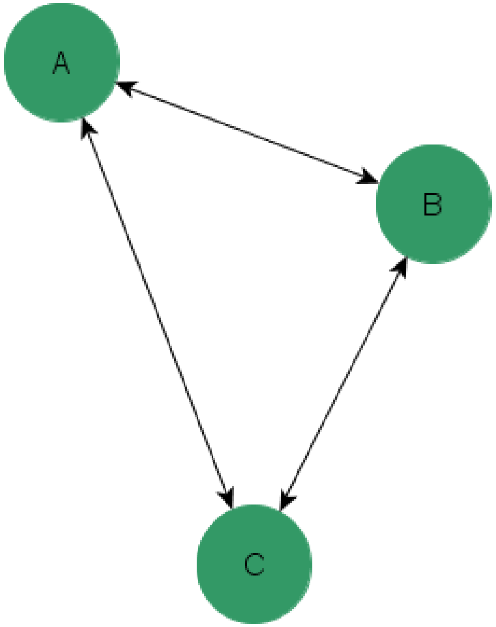
- Examples:
  - Random = Like allocating policy via the toss of a (fair) coin
  - Quasi-random = Policy allocated based on a cut-off, where units just below and just above the cut-off may be very similar in other respects

This talk...

- Illustrates how random or quasi-random variation from policy experiments can be used to identify social effects
- Economics focused approach
- Reduced form effects
  - Policy experiments within a network
- Discuss data requirements

Some definitions

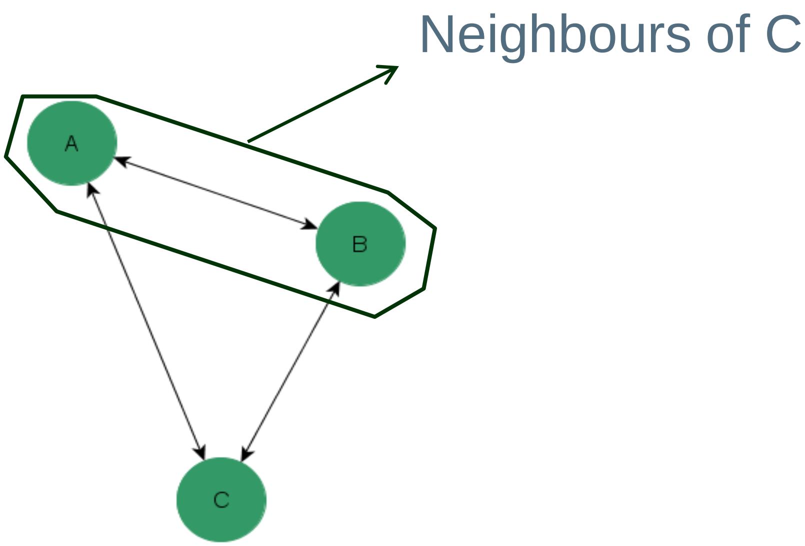
- Network = A set of **nodes** connected to each other by **links**
  - **Nodes** = Individuals, households, firms, countries
  - **Links** = Friendship, kinship, transactions, employment or other commercial relationship



Some Definitions

- We refer to data with detailed information on nodes and edges as **network data**

- **Neighbours** = the set of other nodes that a node is directly linked with



Policy experiments and social effects

- Consider a programme or policy that is allocated **(quasi)-randomly** to a **subset** of a network
  - E.g. Providing free gym membership to some individuals in a network to encourage exercise
- Note that it is important that the policy shifts the behaviour of some of the directly treated
  - Giving free gym memberships must induce some of those receiving them to exercise

Assumptions on network

- It has **well-defined and known boundaries**
  - For example, a village or classroom
- It is **fixed**
  - Policy doesn't change the network
- Policy should be uncorrelated with underlying network
  - Network should not have been formed to withstand the type of shock brought about by the policy
    - Separate out effect of neighbours' actions on own action from factors influencing formation of network

Treatment Status vs. Treatment Exposure

- To proceed, distinguish between **treatment status** and **treatment exposure** (Manski, 2013; and Aronow and Samii, 2013)

- **Treatment status** = whether node directly receives the policy or not

- If there are social effects, then an individual's receipt of the policy will also influence indirectly the outcomes of his neighbours

- **Treatment exposure** = Includes all direct and indirect influences (through the network) of the treatment allocation on the individual
  - Depends on the underlying network structure, and treatment allocation

Defining Treatment Exposure

- Many different ways of defining treatment exposure
  - Proportion of an individual's neighbours that receive policy
  - Number of an individual's neighbours receiving policy
  - Position in network relative to those receiving policy
    - Direct neighbour or indirect neighbour

- Choice depends on what one believes to be the mechanism through which providing the policy to an individual influences the outcomes of his neighbours

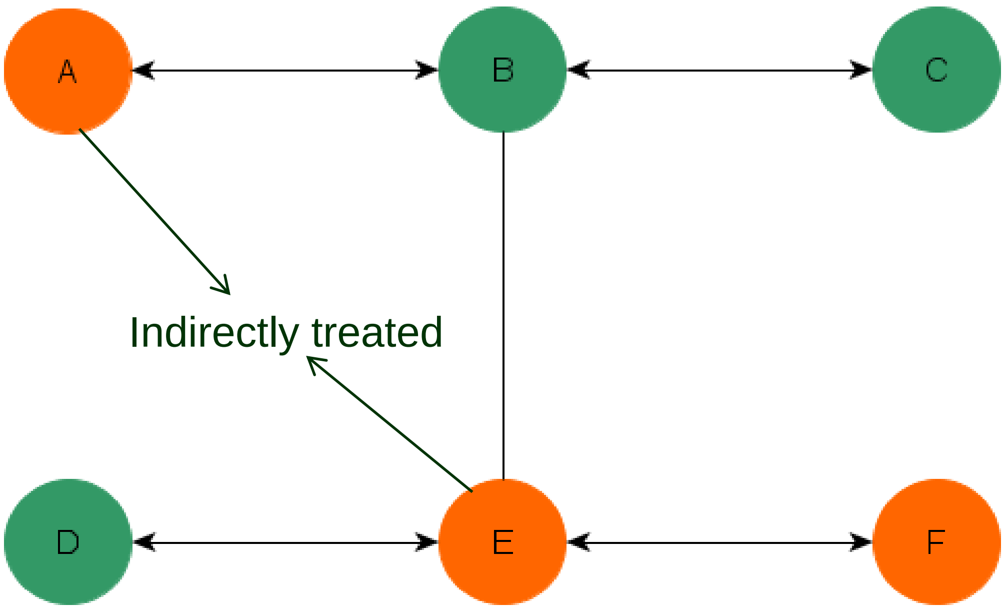
- In this talk, we assume that the policy affects the treated individual and his direct neighbours only

More definitions

- Subset of individuals receiving the policy (or treatment) = **Directly Treated**

- Individuals who don’t receive the policy themselves, but whose neighbours do = **Indirectly Treated**

Example



Receive policy, i.e. 'Directly Treated'

Don't receive policy

Some intuition

- Treatment exposure of indirectly treated individuals varies with
  - Position of individual in the network
  - The treatment allocation
- Policy evaluation literature: If a policy is randomly or quasi-randomly allocated, then comparing the average outcome of the treated and the untreated provides a credible estimate of policy effect

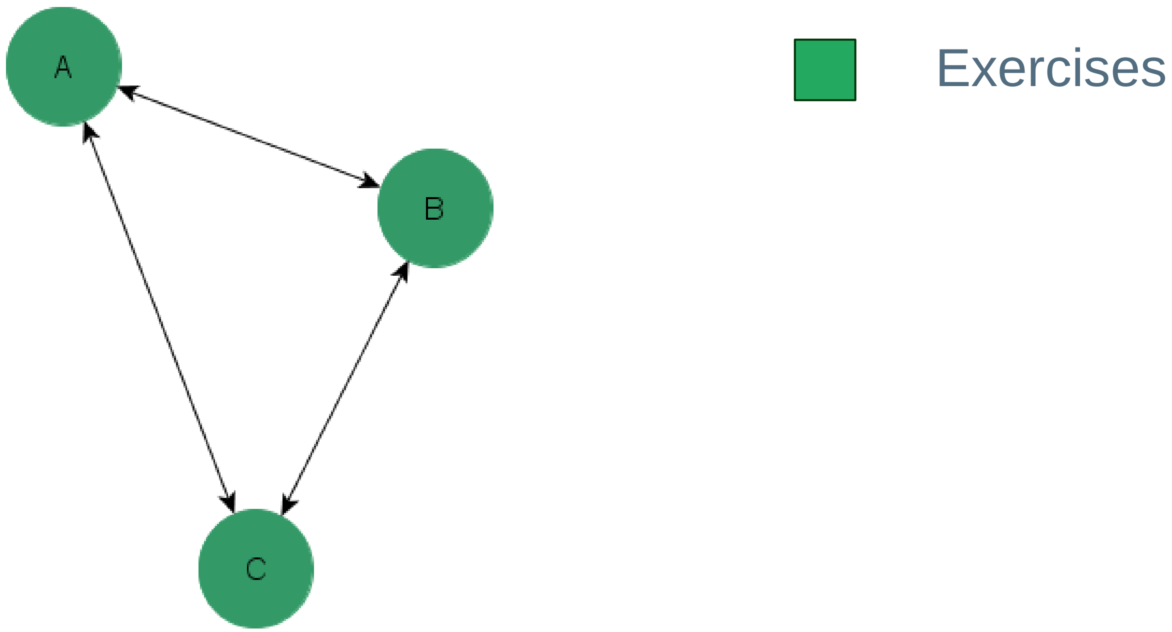
Within Network Variation – Reduced Form Effect

- A natural comparison to make is to compare average outcomes of the indirectly treated, across different levels of treatment exposure

$$\beta(s', s'') = E(y \mid s = s') - E(y \mid s = s'')$$

- For this to be computed, one must observe nodes with treatment exposure levels s' and s'' in the data (Manski 2013)
  - Support condition
- What does the experimental variation get us?

An Example



Did A and B's exercising influence C to exercise?

What challenges does experimental variation resolve?

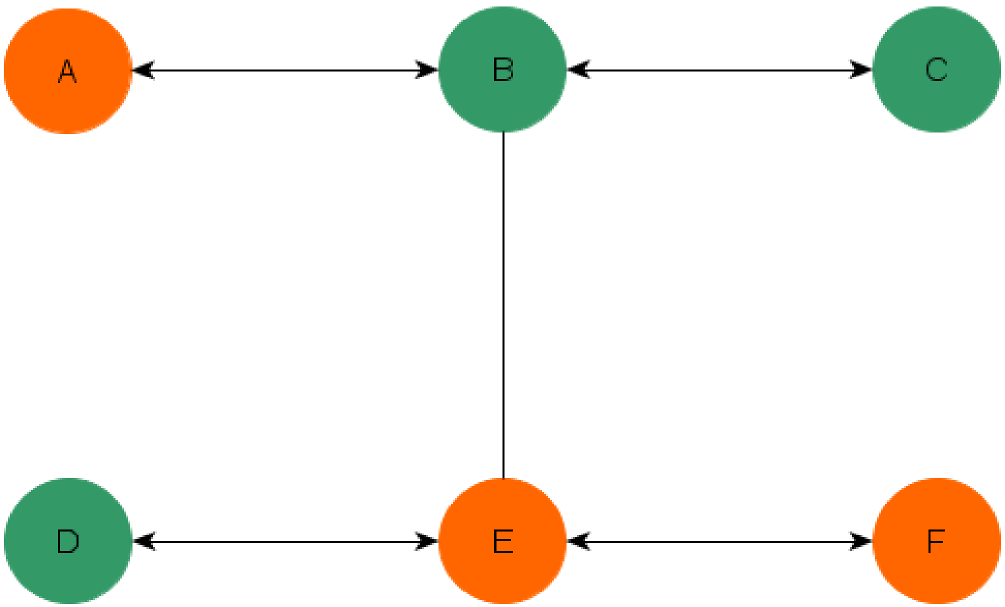
1. Did A and B influence C, or did C influence A and B?
   – We know who received the policy and who didn't
2. There could be some unobserved factor influencing A, B and C to exercise (e.g. they live in the same neighbourhood where a gym has just opened up)
   – Random or quasi-random allocation of the free gym memberships ensures that the policy is uncorrelated with these (unobserved) background factors
3. Or A, B and C all like exercising and became friends because they like to exercise; Or A, B and C are all very social which makes them likely to have these friends and also to exercise
   – Since treatment exposure depends on the network structure, the support condition and (quasi-)random allocation insufficient to overcome this

Within Network Variation

- So, endogenous formation of networks is still an issue
  - Further assumptions are needed

- Manski (2013) suggests partitioning individuals in the network into **‘types’**, based on:
  - Observed characteristics of the nodes (e.g. gender, age, etc)
  - Measures of their network position (e.g. number of links)

Example



- Types of untreated (orange) based on number of edges:
  - Type 1 = 1 edge = {A, F}
  - Type 2 = 3 edges = {E}

Identifying Reduced Form Social Effect

- Assume that individuals of the same 'type' have similar values of unobserved variables that affect the outcome and their linking decisions

- Then reduced form social effect can be calculated by:
  - Calculating $\beta(s', s'')$ for each 'type'
  - Take a weighted average of the 'type' specific social effects

- For this to be possible, require that nodes with treatment exposure levels s' and s'' are observed **for every type** for which the exposure levels are feasible
  - Stronger support condition

Will this support condition hold in practice?

- Not always
- Going back to our simple example, we see that it fails (albeit with an artificially small network)
- More generally, it is likely to fail with networks data, since the treatment of any node constrains the **realised** treatment exposure for its neighbours (Manski 2013).
- Note also that for **inference**, need a sufficient proportion of nodes with each exposure level of interest within each ‘type’
- Parametric assumptions could be made to get around support condition
  - E.g. Assume linear relationship between outcome and treatment exposure

What data is needed for this?

- Outcomes for the untreated individuals
- Treatment status of all individuals in the network
- Know enough about a network to be able to calculate the treatment exposure from the treatment allocation
  - If treatment exposure is measured as proportion of an individual's friends who are treated, need to know the friends of the individual
- Data from a sample can be used, but need to know the treatment status of **all** individuals that influence treatment exposure.

Summary

- Outline when (quasi)-random allocation of policies and programmes can be used to uncover social effects
- Highlight the key assumptions required for recovering reduced form social effects
- Briefly outline data requirements for such analysis

Further Reading

- A. Advani and B. Malde (2014), “Empirical Methods for Networks Data: Social Effects, Network Formation and Measurement Error”, IFS *mimeo* (Survey article).

- C. Manski. ”Identification of Treatment Response with Social Interactions”. *Econometrics Journal,* 16:S1–S23, 2013.

- P. Aronow and C. Samii. ”Estimating Average Causal Effects Under General Interference”. *Unpublished Manuscript, 2013.*

Thanks!

arun_a@ifs.org.uk

bansi_m@ifs.org.uk