Revealed Social Network

Chris Chambers Yusufcan Masatlioglu Chris Turansick

Motivation

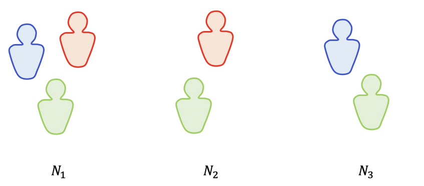
- Our goal is to develop a revealed preference style test and identification results for a specific model of peer effects in networks.
    - The linear in means model (LiM)

Preliminaries

- A person, i, makes choices in the context of a group of agents, N
- Choice is made from a fixed set of alternatives X
- Choice is probabilistic
  - $p_i^N(x)$ is how often agent i chooses x in the context of group N
  - p_i^N denotes the entire vector of agent i's choice frequencies in group N
  - p_{-i}^N denotes the choices of every agent besides i in group N

Primitive

- $|X| > 1$ finite set of alternatives (fixed)
- No menu variation
- group variation



- $\mathcal{A}$: the grand set of agents, and N: a typical group of agents
- Data: $p_i^N(x)$ for $i \in N$

$$p_i^N(x) > 0 \text{ for all } x \in X \quad \text{and} \quad \sum_{x \in X} p_i^N(x) = 1$$

The Linear in Means Model

$$p_i^N(x) = \pi_i^N(i)v_i(x) + \sum_{j \in N \setminus i} \pi_i^N(j)p_j^N(x) \quad \text{for all } x$$

- An agent's choice is a convex combination of their bliss point and the other agents' choices
  - v_i is agent i's bliss point
  - $\pi_i^N(j)$ is the weighting agent i puts on agent j in group N
- As many have noted, this choice function arises from quadratic loss utility

$$u(p_i^N, p_{-i}^N) = -\pi_i^N(i)\sum_{x \in X}(p_i^N(x) - v_i(x))^2 - \sum_{j \in N \setminus i} \pi_i^N(j) \sum_{x \in X}(p_i^N(x) - p_j^N(x))^2$$

The Linear in Means Model

$$p_i^N(x) = \pi_i^N(i)v_i(x) + \sum_{j \in N \setminus i} \pi_i^N(j) p_j^N(x) \quad \text{for all } x$$

- An agent's choice is a convex combination of their bliss point and the other agents' choices
    - v_i is agent i's bliss point
    - $\pi_i^N(j)$ is the weighting agent i puts on agent j in group N
- As many have noted, this choice function arises from quadratic loss utility

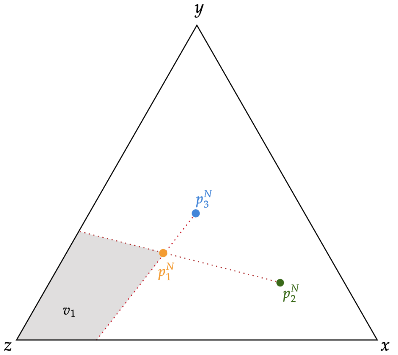
$$u(p_i^N, p_{-i}^N) = -\pi_i^N(i) \sum_{x \in X} (p_i^N(x) - v_i(x))^2 - \sum_{j \in N \setminus i} \pi_i^N(j) \sum_{x \in X} (p_i^N(x) - p_j^N(x))^2$$

Group Variation

- In the paper we consider three models of group variation
- General (GLM):
  - $\pi_i^N(i) > 0$ for all N
  - No other restrictions on $\pi_i^N(j)$ across groups
- Luce (LLM)
  - $w_i(j) > 0$ is a weighting function
  - $\pi_i^N(j) = \frac{w_i(j)}{\sum_{k \in N} w_i(k)}$
- Uniform (ULM)
  - $\pi_i^N(j) = \frac{1}{|N|}$

Feasible Bliss Points

- Given observed choices in a group, we can directly recover the set of feasible bliss points for each agent
- It is the set of points for which p_i^N is a convex combination of that point at p_{-i}^N
- This turns out to be a cone (intersected with the simplex)
- For each agent, there must be some bliss point which is feasible in each of their groups



$$co^{-1}(\Delta(p^N), p_i^N) = \{v \in \Delta(X) | v = \sum_{j \in N} \gamma_j p_j^N, \gamma_j \leq 0 \; \forall j \in N \setminus i, \sum_{j \in N} \gamma_j = 1\}$$

First Test

Theorem

The following are equivalent.

- A data set $\{p^N\}_{N \in \mathcal{N}}$ is consistent with GLM.
- For every i, the collection of sets $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N \in \mathcal{N}_i}$ has a point of mutual intersection.

- The only testable content of GLM is the existence of a feasible bliss point for each agent
- We can test this agent by agent

Further Developing the Test

- Our first test is fully geometric
- Can we develop a linear programming approach?
- Yes!
- We will interpret this test as the existence of a profitable bet on agent i's behavior

Definition

A set of vectors $\{b^N\}_{N \in \mathcal{N}_i}$ with $b^N \in \mathbb{R}^X$ for each $N \in \mathcal{N}_i$ is called a **bet on agent i**.

Feasibility

Definition

A bet on agent i is **strictly feasible** if $\sum_{N \in \mathcal{N}_i} b^N \ll 0$.

- Feasibility is a statement about ex-ante profitability of the bet
- A bank is willing to front the bet if, for each $x \in X$, the bank makes a profit

Individual Rationality

Definition

A bet on agent i is **individually rational** if $b^N \cdot p_i^N > 0$ for each $N \in \mathcal{N}_i$.

- Individual rationality is a statement about profitability of the bet to the bet maker
- An outside observer is willing to make the bet if, in each group, they make a profit (in expectation)

Definition

A bet on agent i is **incentive compatible** if $b^N \cdot (p_i^N - p_j^N) \geq 0$ for each $N \in \mathcal{N}_i$ and each $j \in N \setminus i$.

- Incentive compatibility is a statement about the better's incentives to deviate from their bet
- A bet is a bet on agent i
- An outside observer is willing to stick with their bet if, in each group, there is no other agent j in that group that gives higher (expected) profit

Second Test

Definition

We say that a dataset $\{p^N\}_{N\in\mathcal{N}}$ satisfies **no incentive compatible money pump** if for each i there are no strictly feasible, individually rational, and incentive compatible bets on agent i.

Theorem

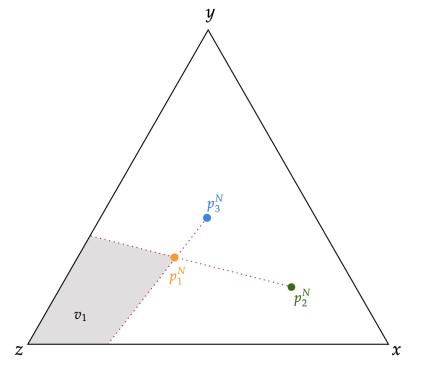
The following are equivalent.

- A data set $\{p^N\}_{N\in\mathcal{N}}$ is consistent with GLM.
- For every i, the collection of sets $\{co^{-1}(\Delta(p^N), p_i^N)\}_{N\in\mathcal{N}_i}$ has a point of mutual intersection.
- $\{p^N\}_{N\in\mathcal{N}}$ satisfies no incentive compatible money pump.

- Suppose $|X| = 2$ (i.e. labor vs leisure or test scores)
- $co^{-1}(\Delta(p^N), p_i^N)$ is $[0,1]$ unless $p_i^N(x)$ is either the min or max among $\{p_j^N(x)\}$
- Let p_i^- be the min value of $p_i^N(x)$ over N where $p_i^N(x)$ is the min among $\{p_j^N(x)\}$
  - Define p_i^+ similarly
- Ignoring a lower dimensional case for today, we have consistency with GLM iff for each i we have $p_i^- \geq p_i^+$

Identification

- So far we have focused on testing
- Our goal now is to show how we can get identification
  - Of v_i
  - Of π_i^N
- We start with partial identification bounds



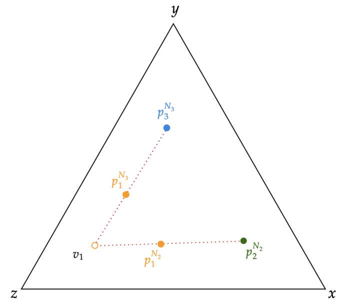
Proposition

In GLM, the tight identified set for v_i is given by $\bigcap_{N \in \mathcal{N}_i} co^{-1}(\Delta(p^N), p_i^N)$.

Corollary

Let $N_j = \{i, j\}$ and $N_k = \{i, k\}$. Suppose that $\{p^N\}_{N \in \mathcal{N}}$ is consistent with GLM, $N_j, N_k \in \mathcal{N}_i^{ext}$, and that the vectors $(p_i^{N_j} - p_j^{N_j})$ and $(p_i^{N_k} - p_k^{N_k})$ are linearly independent. Then v_i is point identified.

- With two binary groups, we are able to identify v_i
- The logic extends to larger group sizes
- We effectively need $|N|$ groups of size $|N|$ (plus linear independence) to identify v_i
  - These have to be groups where $co^{-1}(\Delta(p^N), p_i^N)$ is not the entire simplex



Identifying Network Structure

Proposition

Suppose that v_i is point identified. Then π_i^N is point identified if the set of vectors including v_i and $\{p_j^N\}_{j \in N \setminus i}$ is linearly independent.

- When we have more alternatives than people ($|X| \geq |N|$), we can generically recover each π_i^N
- Recovering π_i^N for each i in a group recovers the entire network structure of that group

Extensions

- Varying network structure instead of group variation
  - Suppose we observe some instrument Z which is independent of v_i but is correlated with network structure
  - All of our tests and identification results go through
  - Each realization of R corresponds to a different group
- Exogenous vs Endogenous components of choice
  - Suppose v_i varies with some set of observed characteristics
  - Our procedure lets us recover endogenous effects (π_i^N) and exogenous effects (v_i) from choices
  - Once we do so, we can treat v_i as our outcome variable
- Product Attribute Variation
  - Suppose we keep X fixed but now vary the attributes of some of the choice alternatives
  - Since we can identify v_i in our setup, we can treat v_i as an outcome variable and estimate how it responds to attribute variation

Conclusion

- We develop revealed preference style tests and identification results for the linear in means model of peer effects
- A key takeaway from our analysis is that identification can be recovered when data is more granular
    - Instead of data on labor vs leisure time, we want data on how labor and leisure time is used
- All of our analysis carries through with either group variation or network variation

Thanks!