

Using Network Structure to Identify Peer Effects

Econometric Methods for Networks,
GCEP, May 8th & 9th, 2017

Bryan S. Graham

University of California - Berkeley

Overview

- Large empirical literature on peer group effects based on the “linear-in-means” model of social interactions (Manski, 1993).
- Quality of research in this area is uneven and has been heavily criticized (e.g., Angrist, 2013).
 - over 20 years since Manski (1993) conditions for identification (and their interpretation) evidently not fully understood by some practitioners.

Overview (continued)

- Recent work on network games with linear best reply functions (e.g., Jackson and Zenou, 2015; Bramouille, Kranton and D'Amours, 2014).
 - provides micro-foundations for linear-in-means model.
 - facilitates intuitive assessment of conditions for identification.
 - also connected to older literature on input-output models recently explores by Acemoglu and co-authors.

Key References

- Manski (1993, *Review of Economic Studies*)
- Brock and Durlauf (2001, *Handbook of Econometrics*)
- Bramouille, Djebbari and Fortin (2009, *Journal of Econometrics*)

Notation

- Let $G = \text{diag}(\mathbf{D}\iota_N)^{-1} \mathbf{D}$ be the **row-normalized network adjacency matrix**.
 - Note that all rows of this matrix sum to 1 by construction.
 - The matrix is row-stochastic (when graph is connected).
 - Focus on undirected links today, but extension to directed case follows.

Notation (continued)

- Let G_i denotes the i^{th} row of G and define

$$G_i y = \sum_{j \neq i} G_{ij} y_j \stackrel{def}{=} \bar{y}_{n(i)}$$

$$G_i X = \sum_{j \neq i} G_{ij} X_j \stackrel{def}{=} \bar{X}_{n(i)}.$$

These equal

- the average action of player i 's peers.
- the average of her peers' attribute vector.

Utility

- Assume that the utility agent i receives from action profile \mathbf{y} , given network structure (\mathbf{D}) and agent attributes (\mathbf{X}), is

$$\begin{aligned} u_i(\mathbf{y}; \mathbf{D}, \mathbf{X}) &= v_i(\mathbf{D}, \mathbf{X}) y_i - \frac{1}{2} y_i^2 + \beta \bar{y}_{n(i)} y_i \\ &= v_i(\mathbf{D}, \mathbf{X}) y_i - \frac{1}{2} y_i^2 + \beta \mathbf{G}_{i\mathbf{y}} y_i. \end{aligned} \quad (1)$$

Utility (continued)

- Assume that $|\beta| < 1$ and define $v_i(\mathbf{D}, \mathbf{X})$ as

$$\begin{aligned} v_i(\mathbf{D}, \mathbf{X}) &= X_i' \gamma + \bar{X}_{n(i)}' \delta + A + U_i \\ &= X_i' \gamma + (\mathbf{G}_i \mathbf{X})' \delta + A + U_i. \end{aligned}$$

- Comment: alternative is provided by quadratic “conformist” preferences (e.g., Akerlof, 1997).
- Comment: recall our discussion of social multiplier centrality earlier

Equilibrium

- The observed action \mathbf{Y} corresponds to a Nash equilibrium.
 - No agent can increase her utility by changing her action given the actions of all other agents in the network.
- The econometrician observes the triple $(\mathbf{Y}, \mathbf{X}, \mathbf{D})$.
 - she does not observe A , nor does she observe \mathbf{U} , the $N \times 1$ vector of individual-level heterogeneity terms.
 - agents *do* observe (A, \mathbf{U}) .

Endogenous and Exogenous Social Effects

- **endogenous:** the marginal utility associated with an increase in y_i is increasing in the average action of one's peers, $\bar{y}_{n(i)}$:

$$\frac{\partial^2 u_i(y, D, X)}{\partial y_i \partial \bar{y}_{n(i)}} = \beta.$$

- **exogenous or contextual:** the marginal utility associated with an increase in y_i varies with peer attributes:

$$\frac{\partial^2 u_i(y, D, X)}{\partial y_i \partial \bar{X}'_{n(i)}} = \delta.$$

Endogenous and Exogenous Social Effects (continued)

- Endogenous and exogenous effects have different policy implications (except under special network structures)
 - effects of a “local” intervention may spread across the entire network in the presence of endogenous effects
 - effects are localized if only exogenous effects are present

Correlated Effects

- **correlated effects:** agents located in networks with high values of A will choose higher actions.

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

- Endogenous, contextual and correlated effects all cause outcomes across members of a common network to covary.
- Attributing this covariance to true spillovers, whether endogenous or contextual, versus group-level heterogeneity is difficult.

Policy Implications

- Spillovers raise the possibility that
 - rewirings of the network – the addition or subtraction of links – could improve the distribution of outcomes.
 - intervening at different locations of the network will have different effects on the distribution of outcomes.
- These claims will become clear shortly.

Linear Best Replies

- F.O.C for optimal behavior generates best response functions of the form

$$Y_i = A + \beta \bar{Y}_{n(i)} + X_i' \gamma + \bar{X}_{n(i)}' \delta + U_i$$

for $i = 1, \dots, N$.

- Called the **linear-in-means** model of social interactions (e.g., Brock and Durlauf, 2001).
- Basis of most empirical work on peer effects.

Linear Best Replies (continued)

An agent's best reply varies with

(i) the average action of those to whom she is directly connected $\bar{Y}_{n(i)}$,

(ii) her own observed attributes X_i ,

(iii) the average attributes of her direct peers $\bar{X}_{n(i)}$,

(iv) the unobserved network effect, A , and

(v) unobserved own attributes, U_i .

A System of Simultaneous Equations

- The N best reply functions define an $N \times 1$ system of (linear) simultaneous equations.
- A least squares fit of Y_i onto a constant, $\bar{Y}_{n(i)}$, X and $\bar{X}_{n(i)}$ will not provide consistent estimates of $\theta_0 = (A_0, \beta_0, \gamma'_0, \delta'_0)'$.
- Manski (1993) calls this feature of the linear-in-means model the **reflection problem**.

Anatomy of the Reflection Problems

- Define the index set

$$\mathcal{N}(i) = \{j : D_{ij} = 1\}$$

with cardinality N_i .

- Y_i is a component of the best response functions of $j \in \{j : j \in \mathcal{N}(i)\}$.
- U_i will be correlated with all $Y_j \in \{Y_j : j \in \mathcal{N}(i)\}$.
- $\Rightarrow U_i$ will covary with $\bar{Y}_{n(i)}$!

Reduced Form

- Write the system of best replies in matrix form:

$$\mathbf{Y} = A\iota_N + \mathbf{X}\gamma + \mathbf{G}\mathbf{X}\delta + \beta\mathbf{G}\mathbf{Y} + \mathbf{U}. \quad (2)$$

- If $|\beta| < 1$, then $I_N - \beta\mathbf{G}$ is strictly (row) diagonally dominant & hence non-singular.
- Solving for the equilibrium action vector as a function of \mathbf{D} , \mathbf{X} , A and \mathbf{U} alone yields

$$\begin{aligned} \mathbf{Y} = & A(I_N - \beta\mathbf{G})^{-1}\iota_N + (I_N - \beta\mathbf{G})^{-1}(\mathbf{X}\gamma + \mathbf{G}\mathbf{X}\delta) \\ & + (I_N - \beta\mathbf{G})^{-1}\mathbf{U}. \end{aligned}$$

Reduced Form

It is helpful to simplify the reduced form in a number of ways. First, using the series expansion

$$(I_N - \beta \mathbf{G})^{-1} = \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k,$$

as well as the fact that $\mathbf{G}\iota_N = \iota_N$ (and hence that $\mathbf{G}^k\iota_N = \iota_N$ for $k \geq 1$) we get the simplification:

$$\begin{aligned} A(I_N - \beta \mathbf{G})^{-1} \iota_N &= A \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}^k \right] \iota_N \\ &= A (1 + \beta + \beta^2 + \beta^3 + \dots) \iota_N \\ &= \frac{A}{1 - \beta} \iota_N. \end{aligned}$$

The Social Multiplier REDUX!

- Further manipulation yields a reduced form of

$$\mathbf{Y} = \frac{A}{1-\beta} \iota_N + \mathbf{X}\gamma + \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}^{k+1} \mathbf{X} \right] (\gamma\beta + \delta) + \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}^k \right] \mathbf{U}.$$

- Consider a policy which changes the value of X_i by Δ .
- What is the effect of this intervention on the distribution of outcomes?

The Social Multiplier (continued)

- We can conceptualize the effect of the intervention as spreading out in a series of “waves”.
- Let \mathbf{c}_i be a vector with a 1 in the i^{th} element and zeros elsewhere. For simplicity assume $\delta = 0$ (i.e., no exogenous effects).
- In the first “wave” the intervention changes agent i 's action alone. The effect on the distribution of outcomes is

$$\Delta' \gamma \mathbf{c}_i$$

The Social Multiplier (continued)

- In the second “wave” agent i 's friends revise their best response in reaction to i 's initial change in action. The effect on the distribution of outcomes is

$$\Delta' \gamma \beta \mathbf{G} \mathbf{c}_i$$

- In the third “wave” agent i 's friends' friends revise their best response in reaction to i 's friends' wave two changes in action. The effect on the distribution of outcomes is

$$\Delta \gamma \beta^2 \mathbf{G}^2 \mathbf{c}_i.$$

The Social Multiplier (continued)

- In the k^{th} wave we have a change in the action vector of

$$\Delta\gamma\beta^{k-1}\mathbf{G}^{k-1}\mathbf{c}_i.$$

- The “long-run” or full effect of the change in X_i on the entire distribution of outcomes is

$$\Delta\gamma(I_N - \beta\mathbf{G})^{-1}\mathbf{c}_i. \tag{3}$$

- The planner can use the form of \mathbf{G} to efficiently target interventions.

Reduced Form (continued)

- $\mathbf{GX} = \bar{\mathbf{X}}$ is a matrix consisting of the average of friends' characteristics (with i^{th} row $\bar{X}_{n(i)}$).
- $\mathbf{G}^2\mathbf{X} = \mathbf{G}\bar{\mathbf{X}}$ is a matrix consisting of an average of your friends' friends' average attributes (with i^{th} row $\bar{X}_{n(i)}^{ff}$).
- $\mathbf{G}^3\bar{\mathbf{X}}$ is an average of your friends' friends' average of their friends' average attributes (with i^{th} row $\bar{X}_{n(i)}^{fff}$)

Reduced Form (continued)

- Extra credit: describe $\mathbf{G}^4 \bar{\mathbf{X}}$ in words.
- Use this notation we get

$$\mathbf{Y} = \frac{A}{1-\beta} \iota_N + \mathbf{X}\gamma + \bar{\mathbf{X}}(\gamma\beta + \delta) + \left[\sum_{k=1}^{\infty} \beta^k \mathbf{G}^k \bar{\mathbf{X}} \right] (\gamma\beta + \delta) + \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}^k \right] \mathbf{U}$$

- In equilibrium, an agent's action will vary with own attributes, her peers', her peers' peers' and so on.

Connection to Dynamic Panel Data

- The endogenous effect induces a distributed lag in \mathbf{X} in the reduced form expression for \mathbf{Y} .
- In dynamic linear panel data models with strictly exogenous regressors, state dependence induces an analogous structure (Chamberlain, 1984; Arellano, 2003).

Formulation as an IV Problem

- Bramoulle, Djebbari and Fortin's (2009) propose a linear IV procedure.
- Our **structural equations** are

$$\mathbf{Y} = A\iota_N + \beta\bar{\mathbf{Y}} + \mathbf{X}\gamma + \bar{\mathbf{X}}\delta + \mathbf{U}$$

Formulation as an IV Problem (continued)

- Let $\bar{\mathbf{Y}} = \mathbf{G}\mathbf{Y}$ to be the $N \times 1$ of peer average actions. Multiplying the reduced form by \mathbf{G} yields the **first stage equations**

$$\begin{aligned}\bar{\mathbf{Y}} = & \frac{A}{1-\beta} \iota_M + \bar{\mathbf{X}}\gamma + \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}^{k+1} \bar{\mathbf{X}} \right] (\gamma\beta + \delta) \\ & + \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}^k \right] \bar{\mathbf{U}}\end{aligned}$$

Estimation

- The dataset consists of a random sample of networks indexed by c
 - with the size of network c equal to N_c and
 - with action profile \mathbf{Y}_c , adjacency matrix \mathbf{D}_c and attribute matrix \mathbf{X}_c

Estimation (continued)

- Assume that $\mathbb{E} [\mathbf{U}_c | \mathbf{D}_c, \mathbf{X}_c, N_c] = 0$.
- This (effectively) restricts the network formation process (in many cases unrealistically).

Estimation (continued)

- The following moment restriction holds at the population vector θ_0

$$\mathbb{E} \left[\begin{pmatrix} \iota_{N_c} & \mathbf{G}_c \bar{\mathbf{X}}_c & \mathbf{X}_c & \bar{\mathbf{X}}_c \end{pmatrix}' \right. \\ \left. \times \left(\mathbf{Y}_c - A_0 \iota_{N_c} - \beta_0 \bar{\mathbf{Y}}_c - \mathbf{X}_c \gamma_0 - \bar{\mathbf{X}}_c \delta_0 \right) \right] = 0$$

- If I_{N_c} , \mathbf{G}_c and \mathbf{G}_c^2 are linearly independent and $\gamma\beta + \delta \neq 0$, then a GMM estimator will be consistent (Bramouille, Djebbari and Fortin (2009, Proposition 1)).

Friends-of-Friends Instrument

- Linear IV fit of Y_{ci} onto a constant, $\bar{Y}_{cn(i)}$, X_{ci} and $\bar{X}_{cn(i)}$ with $\bar{X}_{cn(i)}^{\text{ff}}$ serving as an excluded instrument for $\bar{Y}_{cn(i)}$.
 - consistent estimates of β , γ , and δ
 - see Di Giorgi, Pellizzari and Redaelli (2010, *AEJ*) for an illustrative application

Non-identification Result of Manski (1993)

- Consider the case where \mathbf{G}_c equals

$$\mathbf{G}_c = \left(\iota_{N_c} \iota'_{N_c} - I_{N_c} \right) \frac{1}{N_c - 1}.$$

- Often used in economics of education applications.
- Under this network structure we have

$$\mathbf{G}_c^2 = \frac{1}{N_c - 1} I_{N_c} + \frac{N_c - 2}{N_c - 1} \mathbf{G}_c$$

Non-identification Result of Manski (1993)

- If groups/networks vary in size, then I_{N_c} , G_c and G_c^2 will be linearly independent (cf., Lee, 2007).
- If groups are equal in size identification fails.
- $N_c \rightarrow \infty$, which is (essentially) Manski's (1993) case, gives $G_c^2 = G_c$.

Identification via Non-Transitivity

- Bramoulle, Djebbari and Fortin (2009) note that if the pair (i, j) are not connected then $D_{ij} = 0$.
- If they share some friends in common, then $(i, j)^{th}$ element of \mathbf{D}^2 , which equals $\sum_k D_{ik}D_{kj}$, will be non-zero.
- The presence of intransitive triads (i.e., two-stars), in at least some networks, guarantees linear independence of I_{N_c} , \mathbf{G}_c and \mathbf{G}_c^2 .

Network Effects

- One generalization of the model allows the intercept to vary across sampled networks.
- If A_c varies across networks we get a reduced form of

$$\begin{aligned} \mathbf{Y}_c = & \frac{A_c}{1-\beta} \iota_{N_c} + \mathbf{X}_c \gamma + \bar{\mathbf{X}}_c (\gamma \beta + \delta) \\ & + \left[\sum_{k=1}^{\infty} \beta^k \mathbf{G}_c^k \bar{\mathbf{X}}_c \right] (\gamma \beta + \delta) + \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}_c^k \right] \mathbf{U}_c \end{aligned}$$

Network Effects (continued)

- Subtracting “first stage” from this equation eliminates the “network effect”, yielding

$$\begin{aligned} \mathbf{Y}_c - \bar{\mathbf{Y}}_c &= (\mathbf{X}_c - \bar{\mathbf{X}}_c) \gamma + (I_{N_c} - \mathbf{G}_c) \bar{\mathbf{X}} (\gamma\beta + \delta) \\ &\quad + \left[\sum_{k=1}^{\infty} \beta^k \mathbf{G}^k (I_{N_c} - \mathbf{G}_c) \bar{\mathbf{X}} \right] (\gamma\beta + \delta) \\ &\quad + \left[\sum_{k=0}^{\infty} \beta^k \mathbf{G}_c^k \right] (\mathbf{U}_c - \bar{\mathbf{U}}_c). \end{aligned}$$

- If I_{N_c} , \mathbf{G}_c , \mathbf{G}_c^2 and \mathbf{G}_c^3 are linearly independent θ_0 is identified (need networks with diameter of at least three).

Estimation with Network Effects

- Let $\bar{Y}_{cn(i)}^{ff}$ equal the i^{th} element of $\mathbf{G}_c^2 \mathbf{Y}_c$.
 - equals the average of my friends' averages of their friends behavior.
- Recall that $\bar{X}_{cn(i)}^{fff}$ is the i^{th} row of $\mathbf{G}_c^3 \mathbf{X}$.
 - equals a (weighted) average of agent characteristics up to three degrees away from i .

Estimation with Network Effects (continued)

- A linear IV fit of $Y_{ci} - \bar{Y}_{cn(i)}$ onto $\bar{Y}_{cn(i)} - \bar{Y}_{cn(i)}^{\text{ff}}$, $X_{ci} - \bar{X}_{cn(i)}$ and $\bar{X}_{cn(i)} - \bar{X}_{cn(i)}^{\text{ff}}$ with
 - $\bar{X}_{cn(i)}^{\text{ff}} - \bar{X}_{cn(i)}^{\text{fff}}$ serving as an excluded instrument for $\bar{Y}_{cn(i)} - \bar{Y}_{cn(i)}^{\text{ff}}$ and
 - standard errors “clustered” at the network level
 - yields consistent estimates of θ_0 and asymptotically valid standard error estimates.

Empirical Work

- Identification of θ_0 requires maintaining fairly strong assumptions about the network formation process.
- Condition $\mathbb{E}[\mathbf{U}_c | \mathbf{D}_c, \mathbf{X}_c, N_c, A_c] = 0$ provides a useful way for assessing the plausibility of empirical work.
- Can I predict the idiosyncratic component of behavior using network structure, agent characteristics and/or network size?