Identification of Social Interaction Effects with Social Networks

Yves Zenou
Stockholm University

February 2013

Linear-in-means model

Agents interact in groups, i.e., the social network is partitioned in groups and individuals are affected by all others in their group but by none outside of it.

Assessing social interaction effects: the standard framework

The basic linear-in-means model can be written as:

$$y_{i,r} = \alpha + \beta\,\mathbb{E}(y_r) + \delta\,\mathbb{E}(x_r) + \gamma x_{i,r} + \varepsilon_{i,r} \tag{1}$$

$y_{i,r}$ is the outcome (education, crime, etc.) of individual i belonging to group r,

$x_{i,r}$ is the set of covariates of individual i's (i.e. i's characteristics) in group r,

$\mathbb{E}(y_r)$ denotes the average of outcomes in the peer group r of individual i,

$\mathbb{E}(x_r)$ denotes the average of the characteristics (or characteristics specific to group r) in the peer group r of individual i.

$\beta > 0 \Rightarrow$ endogenous peer effects, $\delta > 0 \Rightarrow$ exogenous contextual effects.

Example: Education, Peer group: Classroom (or school)

$$y_{i,r} = \alpha + \beta\,\mathbb{E}(y_r) + \delta\,\mathbb{E}(x_r) + \gamma x_{i,r} + \varepsilon_{i,r}$$

$y_{i,r}$ is the grade of individual i belonging to class r,

$x_{i,r}$ is the set of individual i's characteristics (parents' education, race, gender, etc.) in class r,

$\mathbb{E}(y_r)$ denotes the average of grades of all students belonging to class r,

$\mathbb{E}(x_r)$ denotes the average of the characteristics of all students belonging to class r.

$\beta > 0 \Rightarrow$ endogenous effect (effect of grades' students in the same class), $\delta > 0 \Rightarrow$ exogenous (contextual) effects (effect of parents' friends education or teacher quality).

Example: Crime, Peer group: Neighborhood

$$y_{i,r} = \alpha + \beta\,\mathbb{E}(y_r) + \delta\,\mathbb{E}(x_r) + \gamma x_{i,r} + \varepsilon_{i,r}$$

$y_{i,r}$ is the crime effort (how often he/she is committing crime) of individual i belonging to neighborhood r,

$x_{i,r}$ is the set of individual i's characteristics (parents' education, race, gender, etc.) living in neighborhood r,

$\mathbb{E}(y_r)$ denotes the crime rate average of individuals living in neighborhood r,

$\mathbb{E}(x_r)$ denotes the average of the characteristics of all individuals living in neighborhood r.

$\beta > 0 \Rightarrow$ endogenous effect (influence of criminal neighbors), $\delta > 0 \Rightarrow$ exogenous effects (influence of neighborhood characteristics or friends' characteristics).

What is the reflection problem?

$$y_{i,r} = \alpha + \beta\,\mathbb{E}(y_r) + \delta\,\mathbb{E}(x_r) + \gamma x_{i,r} + \varepsilon_{i,r}$$

Assume $\mathbb{E}(\varepsilon_{i,r}|y_r, x_r) = 0$ and take the average over peer group r:

$$\mathbb{E}(y_r) = \alpha + \beta\,\mathbb{E}(y_r) + \delta\,\mathbb{E}(x_r) + \gamma\,\mathbb{E}(x_r)$$

Solve this equation:

$$\mathbb{E}(y_r) = \frac{\alpha}{1-\beta} + \left(\frac{\delta+\gamma}{1-\beta}\right) \mathbb{E}(x_r)$$

Plugging this value in the initial equation: $y_{i,r} = \alpha + \beta\,\mathbb{E}(y_r) + \delta\,\mathbb{E}(x_r) + \gamma x_{i,r} + \varepsilon_{i,r}$, we obtain:

$$y_{i,r} = \frac{\alpha}{1-\beta} + \left[\frac{\gamma\beta+\delta}{1-\beta}\right] \mathbb{E}(x_r) + \gamma x_{i,r} + \varepsilon_{i,r}$$

If one estimates this equation, there is an identification problem since β (endogenous peer effects) and δ (exogenous or contextual effects) cannot be separated identified.

3 estimated coefficients, 4 structural parameters: identification fails

This is the **reflection problem** (Manski, 1993).

Manski (1993, 2000) and Moffitt (2001): it is important to *separately* identify peer or endogenous effects from contextual or exogenous effects (policy implications).

Endogenous effects: when the propensity of an individual to behave in some way varies with the behavior of the reference group.

Contextual effects: when the propensity of an individual to behave in some way varies with the exogenous characteristics of the reference group.

Correlated effects: when individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments.

Social networks can solve the reflection problem: Some intuition

So far the reference group was the same for all individuals.

Peer effects: an average intra-group externality that affects identically all the members of a given group.

Group boundaries: arbitrary and at a quite aggregate level

Peer effects in crime: neighborhood level using local crime rates

Peer effects in school: classroom or school level using average school achievements

Social networks: smallest unit of analysis for cross influences: the dyad (two-person group)

Reference group of individual i is his/her best friends.

Reference group of individual j, who is a best friend of i, is not the same as i because individual j may have some best friends that are not i's best friends.

A simple example

Consider a network of students in infinite number arrayed on a line with each student being influenced **only by his left-hand friend** in his/her choice of (educational) activities.

$$\mathbf{G} = \begin{cases} 1 & j = i-1 \\ 0 & j \neq i-1 \end{cases}$$

Structural model:

$$\begin{aligned} y_{i,r} &= \alpha + \beta\, y_{i-1,r} + \delta x_{i-1,r} + \gamma x_{i,r} + \eta_r + \varepsilon_{i,r} \\ \mathbb{E}(\varepsilon_{i,r} &\mid x_{-\infty,r}, ..., x_{+\infty,r}) = 0 \end{aligned}$$

Using the panel data model terminology, lags of $x_{i,r}$ may be used as instruments for $y_{i-1,r}$.

This captures the intuition that the **characteristics of the friends' friends** of a student who are not his friends may serve as instruments for the **actions of his own friends**.

This example illustrates the case of a network in which we can find **intransitive triads**.

These are sets of three students i, j, k such that i is affected by j and j is affected by k (that is, a triad), but i is not affected by k.

Here, i, j, k forms an **intransitive triad** for any i when $j = i - 1$ and $k = i - 2$, since i is not directly affected by $i - 2$.

We will show below that the presence of intransitive triads is a sufficient (but not necessary) condition for the identification of social effects in the absence of correlated effects.

Now consider the case where there may be **network fixed effects** potentially correlated with the family background of students.

Assuming that the $x_{i,r}$'s are strictly exogenous **conditional** on α_r, and maintaining our other assumptions, we have

$$\begin{aligned} y_{i,r} &= \alpha + \beta\, y_{i-1,r} + \delta\, x_{i-1,r} + \gamma x_{i,r} + \eta_r + \varepsilon_{i,r} \\ \mathbb{E}(\varepsilon_{i,r} &\mid x_{-\infty,r}, ..., x_{+\infty,r}, \eta_r) = 0 \end{aligned}$$

Define

$$\Delta z_{i,r} = z_{i,r} - z_{i-1,r} \text{ for } z = y, x, \varepsilon$$

Differencing this equation gives

$$
\begin{aligned}
y_{i,r} - y_{i-1,r} &= \alpha - \alpha + \beta\, y_{i-1,r} - \beta\, y_{i-1,r} + \delta\, x_{i-1,r} - \delta\, x_{i-2,r} \\
&\quad + \gamma x_{i,r} - \gamma x_{i-1,r} + \eta_r - \eta_r + \varepsilon_{i,r} - \varepsilon_{i-1,r}
\end{aligned}
$$

$$
\Delta y_{i,r} = \beta\, \Delta y_{i,r} + \delta\, \Delta x_{i-1,r} + \gamma\, \Delta x_{i,r} + \Delta\varepsilon_{i,r}
$$

and

$$
\mathbb{E}(\Delta\varepsilon_{i,r} \mid \Delta x_{-\infty,r}, ..., \Delta x_{+\infty,r}) = 0
$$

Therefore, lags $(\Delta x_{-\infty,r}, ..., \Delta x_{+\infty,r})$ can be used as valid identifying instruments.

Go back to the **general model with network effects**

Equation is now:

$$y_{i,r_i} = \alpha + \beta\,\mathbb{E}(y_{r_i}) + \delta\,\mathbb{E}(x_{r_i}) + \gamma x_{i,r_i} + \varepsilon_{i,r_i}$$

where r_i is **now** the group of best friends of individual i and thus

$$\mathbb{E}(y_{r_i}) = \frac{1}{g_{i,r_i}} \sum_{j=1}^{n_{r_i}} g_{ij,r_i}\, y_{j,r_j}$$

n_{r_i} is the number of friends of i,

$g_{i,r_i} = \sum_{j=1}^{n_{r_i}} g_{ij,r_i}$ is the total number of friends of i,

$g_{ij,r_i} = 1$ means that individuals i and j are best friends.

We have $g_{ij,r_i} = 1$ implies that $g_{ji,r_j} = 1$ (undirected network).

Write the same equation for individual j for whom $g_{ij,r_j} = 1$ we have

$$y_{j,r_j} = \alpha + \beta\,\mathbb{E}(y_{r_j}) + \delta\,\mathbb{E}(x_{r_j}) + \gamma x_{i,r_j} + \varepsilon_{i,r_j}$$

where

$$\mathbb{E}(y_{r_j}) = \frac{1}{g_{j,r_j}} \sum_{k=1}^{n_{r_j}} g_{jk,r_j}\, y_{k,r_k}$$

with

$$\mathbb{E}(y_{r_i}) \neq \mathbb{E}(y_{r_j})$$

Consider our model in matrix form:

$$\mathbf{Y}_r = \alpha \mathbf{1} + \beta \mathbf{G}_r^* \mathbf{Y}_r + \gamma \mathbf{X}_r + \delta \mathbf{G}_r^* \mathbf{X}_r + \epsilon_r,$$

$\bar{r}$: total number of networks in the sample, n_r: number of individuals in the rth network

$n = \sum_{r=1}^{\bar{r}} n_r$ total number of sample observations.

$\mathbf{Y}_r$ is a $n \times 1$ vector of observations on the dependent (decision) variable,

$\mathbf{G}_r$ is the $n \times n$ adjacency matrix in network r,

$\mathbf{X}_r$ is a $n \times 1$ vector of observations on the exogenous variables,

$\mathbf{G}_r^*$ is the $n \times n$ row normalized matrix of $\mathbf{G}_r$.

$\epsilon_{i,r}$'s are i.i.d. innovations with zero mean and variance σ^2 for all i and r.

Assume $E\left[\boldsymbol{\epsilon} | \mathbf{G}_r, \mathbf{X}_r\right] = 0$.

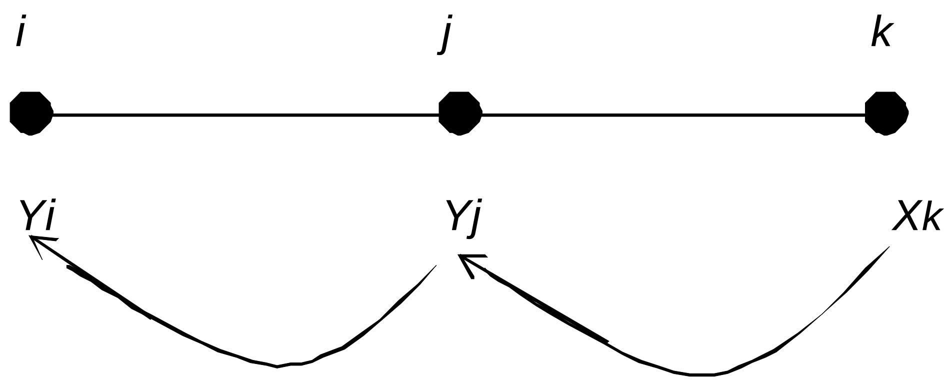
Similar to that of a spatial autoregressive (SAR) model

This model is identified if and only if $E\left(\mathbf{G}_r^{*}\mathbf{Y}_r \mid \mathbf{X}_r\right)$ is not perfectly collinear with the regressors ($\mathbf{X}_r$, $\mathbf{G}_r^{*}\mathbf{X}_r$) so that instruments can be found for the endogenous vector $\mathbf{G}_r\mathbf{Y}_r$.

Bramoulle et al (J of Econometrics 2009): This condition is equivalent to $\mathbf{I}_{n_r}$, $\mathbf{G}_r$ and $\mathbf{G}_r^2$ are linearly independent. Only for *row-normalized* matrix of peer effects, i.e. $\mathbf{G}_r = \mathbf{G}_r^{*}$ so that $\beta\mathbf{G}_r^{*}\mathbf{Y}_r$.

This is true as long as the networks are partially overlapping: some individuals may not be friends with his/her friends' friends (i is friend to j and j is friend to k but k is not friend with i).

For individual i, the characteristics of peers of peers $\mathbf{G}_r^2\mathbf{X}_r$ (i.e. $x_{k,r}$) is a valid instrument for peers' behavior $\mathbf{G}_r^2\mathbf{Y}_r$ (i.e. $y_{j,r}$) since $x_{k,r}$ affects $y_{i,r}$ only indirectly through its effect on $y_{j,r}$ (distance 2)



The natural exclusion restrictions induced by the network structure (existence of an intransitive triad) guarantee identification of the model.

Model

- Structural model in matrix notation:

$$\mathbf{y} = \alpha\boldsymbol{\iota} + \beta\mathbf{G}\mathbf{y} + \gamma\mathbf{x} + \delta\mathbf{G}\mathbf{x} + \boldsymbol{\varepsilon} \quad (2)$$

  - $\mathbf{y}$ is an $n \times 1$ vector of recreational activities for the l network ($l = 1, ..., L$ and stands for the fixed network index)
  - $\mathbf{G}$ is an $n \times n$ interaction matrix, with G_{ij} capturing the strength of interaction between i and j
  - $\boldsymbol{\iota}$ is an $n \times 1$ vector of ones
  - assume: $E[\boldsymbol{\iota} \otimes \mathbf{x}]$ has full rank

Reduced form and identification

(2) can be re-written as

$$\mathbf{y} = \alpha(\mathbf{I} - \beta\mathbf{G})^{-1}\boldsymbol{\iota} + (\mathbf{I} - \beta\mathbf{G})^{-1}(\gamma\mathbf{I} + \delta\mathbf{G})\mathbf{x} + (\mathbf{I} - \beta\mathbf{G})^{-1}\boldsymbol{\varepsilon} \qquad (3)$$

because $|\beta| < 1$, hence $\mathbf{I} - \beta\mathbf{G}$ is invertible.

- Individual i is said to be isolated if his friends' group is empty. Then, the intercept is α, otherwise $\frac{\alpha}{1-\beta}$.
- Social effects can be identified if and only if $\boldsymbol{\theta}$ can be uniquely recovered from the unrestricted reduced-form parameters in (3)

Reduced form and identification

- Since $(\mathbf{I} - \beta\mathbf{G})^{-1} = \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k$ and assuming no isolated students (3) can be expanded as

$$\mathbf{y} = \frac{\alpha}{1-\beta}\boldsymbol{\iota} + \gamma\mathbf{x} + (\gamma\beta + \delta)\sum_{k=0}^{\infty} \beta^k \mathbf{G}^{k+1}\mathbf{x} + \sum_{k=0}^{\infty} \beta^k \mathbf{G}^k \boldsymbol{\varepsilon} \qquad (4)$$

Results

Theorem 1

Suppose that $(\gamma\beta + \delta) \neq 0$. If the matrices $\mathbf{I}, \mathbf{G}$ and $\mathbf{G}^2$ are linearly independent social effects are identified. If the matrices $\mathbf{I}, \mathbf{G}$ and $\mathbf{G}^2$ are linearly dependent and no individual is isolated, social effects are not identified.[a]

[a] Proof provided in the report

- Condition $(\gamma\beta + \delta) \neq 0$ directly refers to (4), i.e. student's expected recreational activities depend upon the family background of his friends. This negated equality holds for $\beta > 0$, $\gamma \neq 0$, and when γ and δ have the same sign
- When $(\gamma\beta + \delta) = 0$, then endogenous and exogenous effects are zero or cancel each other out. Thus, social effects are unidentified.

Implications of Theorem 1

- In Manski (1993) $\mathbf{G}^2 = \mathbf{G}$ resulting in perfect collinearity between the expected mean recreational activities of the friends' groups and the mean family background of the group. Hence, second part of Theorem 1 holds.
- In Moffitt (2001) $\mathbf{G}^2 = \frac{1}{s-1}\mathbf{I} + \frac{s-2}{s-1}\mathbf{G}$ for group size $s \geq 2$. $\mathbf{G}^2$ is linearly dependent on $\mathbf{I}$ and $\mathbf{G}$, which satisfies the second part of Theorem 1.
- In Lee (2007) groups have different sizes ($s_1 \neq s_2$), hence $\mathbf{I}$, $\mathbf{G}$ and $\mathbf{G}^2$ are linearly independent- social effects are identified.

Results cont.

Theorem 2

Suppose that individuals interact in groups. If all groups have the same size, social effects are not identified. If (at least) two groups have different sizes, and if $(\gamma\beta + \delta) \neq 0$, social effects are identified.

- Variations in group sizes create exogenous variations in the reduced form coefficients (3), which leads to identification of social effects

Results with network interactions

- In networks with intransitive triads (a friend of a friend of mine is not my friend), social effects can be identified ($\mathbf{I}$, $\mathbf{G}$ and $\mathbf{G}^2$ are linearly independent)
- In transitive networks (a friend of a friend of mine is my friend) identification generally holds and relies on the nature of the links ($\mathbf{G}^2 \neq 0$)

Theorem 3

Suppose that individuals do not interact in groups. Suppose that $(\gamma\beta + \delta) \neq 0$. If $\mathbf{G}^2 \neq 0$, social effects are identified. If $\mathbf{G}^2 = 0$, social effects are identified when $\alpha \neq 0$, but not when $\alpha = 0$.

Correlated effects/selection

Manski (1993): Can we disentangle "endogenous effects" from "correlated effects", i.e. those due to the fact that individuals in the same group tend to behave similarly because they face a common environment?

The formation of peer group is not random and individuals do select into groups of friends.

It is therefore important to separate the endogenous peer effects from the correlated effects (Manski, 1993), i.e. the same criminal activities may be due to common unobservable variables (such as, for example, the fact that individuals from the same network like bowling together) faced by individuals belonging to the same network rather than peer effects.

This is also very important for crime policies since, for example, if the high-crime rates are due to the fact that teenagers like to bowling together, then obviously the implications are very different than if it is due to peer effects.

Correlated effects might originate from the possible sorting of agents into "groups"

If the variables that drive this process of selection are not fully observable, potential correlations between (unobserved) group-specific factors and the target regressors are major sources of bias.

Selection on unobservables

Assume agents self-select into different networks in a first step, and that link formation takes place within groups in a second step.

Bramoullé *et al.* (2009): if link formation is uncorrelated with the observable variables, this two-step model of link formation generates network fixed effects.

Assuming additively separable network heterogeneity, a within group specification is able to control for selection issues

Bramoullé et al. (2009): by subtracting from the individual-level variables the network average, social effects are again identified and one can disentangle endogenous effects from correlated effects

Consider our model with **network fixed effects**:

$$\mathbf{Y}_r = \alpha \mathbf{1} + \beta \mathbf{G}_r^* \mathbf{Y}_r + \gamma \mathbf{X}_r + \delta \mathbf{G}_r^* \mathbf{X}_r + \eta_r \mathbf{1}_{n_r} + \epsilon_r,$$

We can eliminate the network fixed effect by the network-mean transformation (**global differences**), that is by multiplying this equation by the matrix: $\mathbf{J}_r = \mathbf{I}_{n_r} - \frac{1}{n_r}\mathbf{1}_r\mathbf{1}_r^\top$ ($\mathbf{I}_{n_r}$ identity matrix, $\mathbf{1}_r$ vector of 1).

$$\mathbf{J}_r = \mathbf{I}_{n_r} - \frac{1}{n_r}\mathbf{1}_r\mathbf{1}_r^\top = \begin{pmatrix} 1-\frac{1}{n_r} & -\frac{1}{n_r} & \dots & -\frac{1}{n_r} \\ \dots & \dots & \dots & \dots \\ -\frac{1}{n_r} & \dots & 1-\frac{1}{n_r} & -\frac{1}{n_r} \\ -\frac{1}{n_r} & \dots & \dots & 1-\frac{1}{n_r} \end{pmatrix}$$

with $\mathbf{J}_r\mathbf{1}_r = \mathbf{0}$.

Observe that **global differences**: eliminate the network fixed effect by the network-mean transformation, that is by multiplying this equation by the matrix: $\mathbf{J}_r = \mathbf{I}_{n_r} - \frac{1}{n_r}\mathbf{1}_r\mathbf{1}_r^\top$ ($\mathbf{I}_{n_r}$ identity matrix, $\mathbf{1}_r$ vector of 1).

Average

$$\mathbf{Y}_r = \alpha\mathbf{1} + \beta\mathbf{G}_r^*\mathbf{Y}_r + \gamma\mathbf{X}_r + \delta\mathbf{G}_r^*\mathbf{X}_r + \eta_r\mathbf{1}_{n_r} + \epsilon_r,$$

over all students in i's network, and subtract from i's equation.

The equation being subtracted is identical for all students in the same network

Model becomes:

$$\mathbf{J}_r\mathbf{Y}_r = \beta\mathbf{J}_r\mathbf{G}_r^*\mathbf{Y}_r + \gamma\mathbf{J}_r\mathbf{X}_r + \delta\mathbf{J}_r\mathbf{G}_r^*\mathbf{X}_r + \mathbf{J}_r\epsilon_r$$

Model can be written as:

$$\widehat{\mathbf{Y}_r} = \beta \mathbf{G}_r^* \widehat{\mathbf{Y}}_r + \gamma \widehat{\mathbf{X}}_r + \delta \mathbf{G}_r^* \widehat{\mathbf{X}}_r + \widehat{\epsilon}_r$$

where $\widehat{\mathbf{Y}_r} = \mathbf{J}_r \mathbf{Y}_r$, $\widehat{\mathbf{X}}_r = \mathbf{J}_r \mathbf{X}_r$, $\widehat{\epsilon}_r = \mathbf{J}_r \epsilon_r$.

The model can be identified if and only if $\mathbb{E}\left(\mathbf{G}_r^* \widehat{\mathbf{Y}}_r \mid \widehat{\mathbf{X}}_r\right)$ is not perfectly collinear with the regressors $\left(\widehat{\mathbf{X}}_r, \mathbf{G}_r^* \widehat{\mathbf{X}}_r\right)$.

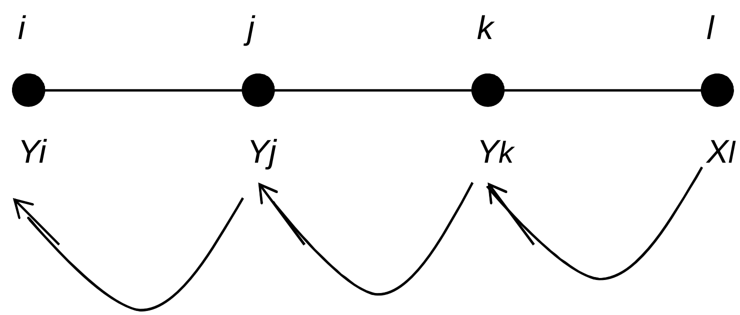
This condition is equivalent to $\mathbf{I}_{n_r}$, $\mathbf{G}_r$, $\mathbf{G}_r^2$ and $\mathbf{G}_r^3$ are linearly independent.

The condition is more demanding because some information has been used to deal with the fixed effects.

Bramoulle et al (2009) show that if two agents i and j in a network are separated by a link of distance 3, then $\mathbf{I}_r$, $\mathbf{G}_r$, $\mathbf{G}_r^2$ and $\mathbf{G}_r^3$ are linearly independent. Model is identified.

Consider four individuals: ij, jk, kl, but l is not friend with i.

$x_{l,r}$ can serve as an instrument for $y_{j,r}$ in individual i's equation since $x_{l,r}$ affects $y_{i,r}$ but only indirectly through its effect on $y_{k,r}$.



Example with star-shaped network with $n = 3$ and 1 in the center

Consider our model with **network fixed effects**:

$$\mathbf{Y}_r = \alpha \mathbf{1} + \beta \mathbf{G}_r^* \mathbf{Y}_r + \gamma \mathbf{X}_r + \delta \mathbf{G}_r^* \mathbf{X}_r + \eta_r \mathbf{1}_{n_r} + \epsilon_r,$$

$$y_{i,r} = \alpha + \beta \frac{\sum_{j \in P_i} y_{j,r}}{n_i, r} + \gamma x_{i,r} + \delta \frac{\sum_{j \in P_i} x_{j,r}}{n_{i,r}} + \eta_r + \varepsilon_{i,r}$$

$$\mathbf{G}_S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; \mathbf{G}_S^* = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; \mathbf{1}_S = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \; \mathbf{1}_S^\mathsf{T} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

Initial model

$$\mathbf{Y}_r = \alpha\mathbf{1} + \beta\mathbf{G}_r^*\mathbf{Y}_r + \gamma\mathbf{X}_r + \delta\mathbf{G}_r^*\mathbf{X}_r + \eta_r\mathbf{1}_{n_r} + \epsilon_r,$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix} + \beta \begin{pmatrix} y_2 + y_3 \\ y_1 \\ y_1 \end{pmatrix} + \gamma \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \delta \begin{pmatrix} 0.5\,(x_2 + x_3) \\ x_1 \\ x_1 \end{pmatrix} + \begin{pmatrix} \eta_r \\ \eta_r \\ \eta_r \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}$$

Local differences

We average

$$y_{i,r} = \alpha + \beta \frac{\sum\limits_{j \in P_i} y_{j,r}}{n_i, r} + \gamma x_{i,r} + \delta \frac{\sum\limits_{j \in P_i} x_{j,r}}{n_{i,r}} + \eta_r + \varepsilon_{i,r}$$

over all student i's friends, and subtract it from i's equation.

This approach is local since it does not fully exploit the fact that the fixed effect is not only the same for all i's friends but also for all students of his network.

$$\mathbf{Y}_r = \alpha \mathbf{1} + \beta \mathbf{G}_r^* \mathbf{Y}_r + \gamma \mathbf{X}_r + \delta \mathbf{G}_r^* \mathbf{X}_r + \eta_r \mathbf{1}_{n_r} + \epsilon_r,$$

We obtain:

$$\begin{aligned} \left(\mathbf{I}_{n_r} - \mathbf{G}_r^*\right) \mathbf{Y}_r &= \beta \left(\mathbf{I}_{n_r} - \mathbf{G}_r^*\right) \mathbf{G}_r^* \mathbf{Y}_r + \gamma \left(\mathbf{I}_{n_r} - \mathbf{G}_r^*\right) \mathbf{X}_r \\ &\quad + \delta \left(\mathbf{I}_{n_r} - \mathbf{G}_r^*\right) \mathbf{G}_r^* \mathbf{X}_r + \left(\mathbf{I}_{n_r} - \mathbf{G}_r^*\right) \epsilon_r \end{aligned}$$

Note $\left(\mathbf{I}_{n_r} - \mathbf{G}_r^*\right) \mathbf{1}_{n_r} = \mathbf{0}$.

$$
\begin{aligned}
(\mathbf{I}_3 - \mathbf{G}_S^*) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1.0 & -0.5 & -0.5 \\ -1.0 & 1.0 & 0 \\ -1.0 & 0 & 1.0 \end{pmatrix}
\end{aligned}
$$

and

$$
(\mathbf{I}_3 - \mathbf{G}_S^*)\,\mathbf{G}_S^* = \begin{pmatrix} -1 & 0.5 & 0.5 \\ 1 & -0.5 & -0.5 \\ 1 & -0.5 & -0.5 \end{pmatrix}
$$

Thus, transformed model

$$\begin{pmatrix} y_1 - 0.5y_2 - 0.5y_3 \\ y_2 - y_1 \\ y_3 - y_1 \end{pmatrix} = \beta \begin{pmatrix} -y_1 + 0.5y_2 + 0.5y_3 \\ y_1 - 0.5y_2 - 0.5y_3 \\ y_1 - 0.5y_2 - 0.5y_3 \end{pmatrix}$$

$$+\gamma \begin{pmatrix} x_1 - 0.5x_2 - 0.5x_3 \\ x_2 - x_1 \\ x_3 - x_1 \end{pmatrix} + \delta \begin{pmatrix} -x_1 + 0.5x_2 + 0.5x_3 \\ x_1 - 0.5x_2 - 0.5x_3 \\ x_1 - 0.5x_2 - 0.5x_3 \end{pmatrix}$$

$$+ \begin{pmatrix} \varepsilon_1 - 0.5\varepsilon_2 - 0.5\varepsilon_3 \\ \varepsilon_2 - \varepsilon_1 \\ \varepsilon_3 - \varepsilon_1 \end{pmatrix}$$

Global differences

We can eliminate the network fixed effect by the network-mean transformation, that is by multiplying this equation by the matrix: $\mathbf{J}_r = \mathbf{I}_{n_r} - \mathbf{H}_r$ (where $\mathbf{H}_r = \frac{1}{n_r}\mathbf{1}_r\mathbf{1}_r^{\mathsf{T}}$).

$$\begin{aligned}
\mathbf{J}_S &= \mathbf{I}_3 - \frac{1}{3}\mathbf{1}_S\mathbf{1}_S^{\mathsf{T}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}
\end{aligned}$$

and

$$\mathbf{J}_S\mathbf{1}_S = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We obtain:

$$\mathbf{J}_r\mathbf{Y}_r = \beta\mathbf{J}_r\mathbf{G}_r\mathbf{Y}_r + \gamma\mathbf{J}_r\mathbf{X}_r + \delta\mathbf{J}_r\mathbf{G}_r^*\mathbf{X}_r + \mathbf{J}_r\epsilon_r$$

In contrast to the local difference approach, the equation being subtracted is now identical for all students in the same network.

$$
\begin{pmatrix} \frac{2}{3}y_1 - \frac{1}{3}y_2 - \frac{1}{3}y_3 \\ -\frac{1}{3}y_1 + \frac{2}{3}y_2 - \frac{1}{3}y_3 \\ -\frac{1}{3}y_1 - \frac{1}{3}y_2 + \frac{2}{3}y_3 \end{pmatrix} = \beta \begin{pmatrix} \frac{2}{3}\left(-y_1 + y_2 + y_3\right) \\ \frac{1}{3}\left(y_1 - y_2 - y_3\right) \\ \frac{1}{3}\left(y_1 - y_2 - y_3\right) \end{pmatrix}
$$

$$
+\gamma \begin{pmatrix} \frac{2}{3}x_1 - \frac{1}{3}x_2 - \frac{1}{3}x_3 \\ -\frac{1}{3}x_1 + \frac{2}{3}x_2 - \frac{1}{3}x_3 \\ -\frac{1}{3}x_1 - \frac{1}{3}x_2 + \frac{2}{3}x_3 \end{pmatrix}
$$

$$
+\delta \begin{pmatrix} \frac{1}{3}x_2 - \frac{2}{3}x_1 + \frac{1}{3}x_3 \\ \frac{1}{3}x_1 - \frac{1}{6}x_2 - \frac{1}{6}x_3 \\ \frac{1}{3}x_1 - \frac{1}{6}x_2 - \frac{1}{6}x_3 \end{pmatrix}
$$

$$
+ \begin{pmatrix} \frac{2}{3}\epsilon_1 - \frac{1}{3}\epsilon_2 - \frac{1}{3}\epsilon_3 \\ -\frac{1}{3}\epsilon_1 + \frac{2}{3}\epsilon_2 - \frac{1}{3}\epsilon_3 \\ -\frac{1}{3}\epsilon_1 - \frac{1}{3}\epsilon_2 + \frac{2}{3}\epsilon_3 \end{pmatrix}
$$

Correlated Effects- Model

- Identification problem: correlation between family background of students and unobserved variables common to students within particular network
- For any network l and any student i belonging to l

$$y_{li} = \alpha_l + \beta \frac{\sum_{j \in P_i} y_{lj}}{n_i} + \gamma x_{li} + \delta \frac{\sum_{j \in P_i} x_{lj}}{n_i} + \varepsilon_{li} \tag{5}$$

  - α_l captures unobserved variables that have universal impact on all individuals within the network (network fixed effect)
  - $E[\varepsilon_{li} | \mathbf{x}_l, \alpha_l] = 0$- strict exogeneity of $\mathbf{x}_l$ conditional on α_l

Correlated Effects- Model

- Remedy: elimination of network specific unobservables by differencing structural equations:

- **Local differences**

Average eq. (5) over all student *i*'s friends and subract it from eq. (5)

- **Global differences**

Average eq. (5) over all students in *i*'s network and subract it from eq. (5)

- Within local transformation model

$$(\mathbf{I}-\mathbf{G})\mathbf{y}_l = \beta(\mathbf{I}-\mathbf{G})\mathbf{G}\mathbf{y}_l + \gamma(\mathbf{I}-\mathbf{G})\mathbf{x}_l + \delta(\mathbf{I}-\mathbf{G})\mathbf{G}\mathbf{x}_l + (\mathbf{I}-\mathbf{G})\varepsilon_l \quad (6)$$

- Within global transformation model

$$(\mathbf{I}-\mathbf{H})\mathbf{y}_l = \beta(\mathbf{I}-\mathbf{H})\mathbf{G}\mathbf{y}_l + \gamma(\mathbf{I}-\mathbf{H})\mathbf{x}_l + \delta(\mathbf{I}-\mathbf{H})\mathbf{G}\mathbf{x}_l + (\mathbf{I}-\mathbf{H})\varepsilon_l \quad (7)$$

Results (local transformation)

Theorem 4

Consider model (6). Suppose that $(\gamma\beta + \delta) \neq 0$*. Social effects are identified if and only if the matrices* $\mathbf{I}$, $\mathbf{G}$, $\mathbf{G}^2$ *and* $\mathbf{G}^3$ *are linearly independent.*

Note: in the presence of correlated effects, identification condition is stronger than in previous theorems

Corollary 5

Consider model (6) and suppose that $(\gamma\beta + \delta) \neq 0$*. If the diameter of the network is greater than or equal to 3, social effects are identified.*

Note: diameter is the maximal friendship distance between any two students in the network

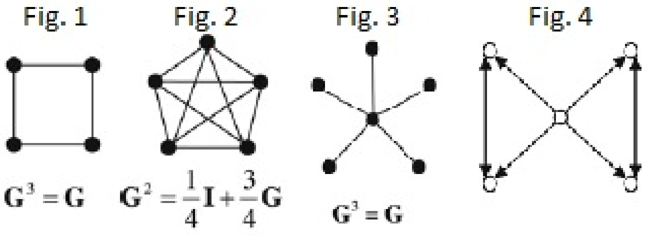
Results (global transformation)

Theorem 6

Consider model (7). Suppose that $(\gamma\beta + \delta) \neq 0$*. If the matrices* $\mathbf{I}$, $\mathbf{G}$, $\mathbf{G}^2$ *and* $\mathbf{G}^3$ *are linearly independent, social effects are identified. Next, suppose that* $\mathbf{G}^3 = \lambda_0\mathbf{I} + \lambda_1\mathbf{G} + \lambda_2\mathbf{G}^2$*. If* $rank(\mathbf{I} - \mathbf{G}) < n - 1$ *and* $2\lambda_0 + \lambda_1 + 1 \neq 0$*, social effects are identified. In contrast, if* $rank(\mathbf{I} - \mathbf{G}) = n - 1$*, social effects are not identified.*

Note: if social effects are identified when using local transformation, they are also identified when taking global differences. If social effects are not identified when using global transformation, they are never identified.

Lack of identification: examples



- In figures 1-3, the matrices **I**, **G**, $\mathbf{G}^2$ and $\mathbf{G}^3$ are linearly dependent, hence social effects are not identified
- Figure 4 is an example of the network, for which identification holds under global, but not local, differences.

Estimation

- US in-school Add Health data (September 1994 - April 1995), sample of 80 high schools and 52 middle schools, 55208 observations
- Correlated effects eliminated by appropriate local transformation:

$$(\mathbf{I}-\mathbf{G})\mathbf{y} = \beta(\mathbf{I}-\mathbf{G})\mathbf{G}\mathbf{y} + \gamma(\mathbf{I}-\mathbf{G})\mathbf{X} + \delta(\mathbf{I}-\mathbf{G})\mathbf{G}\mathbf{X} + \mathbf{v} \quad (8)$$

- Generalized 2SLS procedure implemented:
  - first step: estimate a 2SLS using instruments $\mathbf{S} = \left[(\mathbf{I}-\mathbf{G})\mathbf{X} \quad (\mathbf{I}-\mathbf{G})\mathbf{G}\mathbf{X} \quad (\mathbf{I}-\mathbf{G})\mathbf{G}^2\mathbf{X}\right]$ and obtain $\widehat{\theta}^{2SLS}$
  - second step: estimate a 2SLS with instruments $\widehat{\mathbf{Z}} = \mathbf{Z}\left(\widehat{\theta}^{2SLS}\right)^2$

[2] Please refer to the report

Inferences

- Statistically significant exogenous social effects: age, presence of parents, parents' education, parents' participation in the labour market
  - e.g. student's recreational activities index depends negatively on the mean age of his friends and positively on friends' mean of parents' labour participation
- Statistically significant own characteristics: age, gender, race, parents' education, parents' participation in the labour market
  - e.g. index decreases with age and with being white, increases with being female and parents' labour participation
- Endogenous social effect is significant at 10% level
  - the mean of student's friends' actions has a positive impact on his recreational activity

Do we have a theoretical model backing up the econometrics model?

Local average ($\mathbf{G}_r^*$) or local aggregate ($\mathbf{G}_r$) model?

Microfoundations?

Local-aggregate model

Finite set of agents $N = \{1, \ldots, n\}$ is partitioned into r^{max} networks

Adjacency matrix $G_r = [g_{ij,r}]$

Utility

$$u_{i,r}(y_{i,r}; Y_r, G_r) = \left(a_{i,r} + \eta_r + \epsilon_{i,r}\right) y_{i,r} - \frac{1}{2} y_{i,r}^2 + \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{i,r} y_{j,r}$$

where $\phi_1 \geq 0$.

Unique Nash equilibrium if $0 \leq \phi_1 g_r^{\max} < 1$ ($g_r^{\max} = \max_i g_{i,r}$ the highest degree of network r) given by (best-reply function for individual i)

$$y_{i,r} = \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + a_{i,r} + \eta_r + \epsilon_{i,r}.$$

Matrix form ($\pi_{i,r} = a_{i,r} + \eta_r + \epsilon_{i,r}$, $\Pi_r = (\pi_{1,r}, \cdots, \pi_{n_r,r})'$)

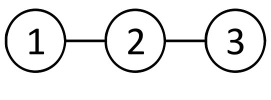
$$Y_r = (I_{n_r} - \phi_1 G_r)^{-1} \Pi_r$$

Local-average model

$G_r^* = [g_{ij,r}^*]$, where $g_{ij,r}^* = g_{ij,r}/g_{i,r}$: *Row-normalized* adjacency matrix of network r.

By construction, $0 \leq g_{ij,r}^* \leq 1$.

Examples



$$G_r = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \text{ and } G_r^* = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}.$$

Figure 1: an example network with corresponding adjacency matrices

Denote: $y_{i,r}^{av} = \sum_{j \in N_{i,r}} g_{ij,r}^{*} y_{j,r}$ the average effort of individual i's friends.

Payoff:

$$u_{i,r}(y_{i,r}; Y_r, G_r) = \left(a_{i,r}^{*} + \eta_r^{*} + \epsilon_{i,r}^{*}\right) y_{i,r} - \frac{1}{2} y_{i,r}^{2} - \frac{d}{2}(y_{i,r} - y_{i,r}^{av})^{2}$$

with $d \geq 0$.

If $0 \leq \phi_2 < 1$, unique Nash equilibrium given by:

$$y_{i,r} = \phi_2 \sum_{j=1}^{n_r} g^*_{ij,r} y_{j,r} + a_{i,r} + \eta_r + \epsilon_{i,r},$$

where $\phi_2 = d/(1+d)$, $a_{i,r} = (1-\phi_2)a^*_{i,r}$, $\eta_r = (1-\phi_2)\eta^*_r$, and $\epsilon_{i,r} = (1-\phi_2)\epsilon^*_{i,r}$.

Matrix form

$$Y_r = (I_{n_r} - \phi_2 G^*_r)^{-1} \Pi_r$$

Hybrid network model

Integrating *local-aggregate* and *local-average* effects into the same model.

Utility function

$$
\begin{aligned}
u_{i,r}(y_{i,r}; Y_r, G_r) \;=\; & \left(a^*_{i,r} + \eta^*_r + \epsilon^*_{i,r}\right) y_{i,r} - \frac{1}{2} y^2_{i,r} + d_1 \sum_{j=1}^{n_r} g_{ij,r} y_{i,r} y_{j,r} \\
& - \frac{d_2}{2} (y_{i,r} - y^{av}_{i,r})^2
\end{aligned}
$$

where $d_1 \geq 0$ and $d_2 \geq 0$

Unique NE if $\phi_1 \geq 0$, $\phi_2 \geq 0$ and $\phi_1 g_r^{\text{max}} + \phi_2 < 1$,

$$y_{i,r} = \phi_1 \sum_{j=1}^{n_r} g_{ij,r} y_{j,r} + \phi_2 \sum_{j=1}^{n_r} g^*_{ij,r} y_{j,r} + a_{i,r} + \eta_r + \epsilon_{i,r}$$

where $\phi_1 = d_1/(1 + d_2)$, $\phi_2 = d_2/(1 + d_2)$, $a_{i,r} = (1 - \phi_2)a^*_{i,r}$, $\eta_r = (1 - \phi_2)\eta^*_r$, and $\epsilon_{i,r} = (1 - \phi_2)\epsilon^*_{i,r}$.

Comparisons: Local aggregate vs local average model

Consider the case where all individuals are ex ante identical apart from their positions in the network such that $\pi_{i,r} = \pi_r$ for $i = 1, \cdots, n_r$.

For the local-aggregate model, if $0 \leq \phi_1 g_r^{\max} < 1$, the unique Nash equilibrium is:

$$Y_r = \pi_r \left(I_{n_r} - \phi_1 G_r\right)^{-1} l_{n_r},$$

where l_{n_r} is an n_r-dimensional vector of ones.

$\left(I_{n_r} - \phi_1 G_r\right)^{-1} l_{n_r}$ represents the Bonacich centrality of a network. The more central an individual's position is, the higher is her equilibrium effort and equilibrium utility.

Local-average model: if $0 \leq \phi_2 < 1$, the the unique interior Nash equilibrium is

$$Y_r = \pi_r (1 - \phi_2)^{-1} l_{n_r}.$$

The position in the network plays no role and all individuals provide the same equilibrium effort level $\pi_r/(1-\phi_2)$ in network r.

Fundamental differences with the local-aggregate model where, even if agents are ex ante identical, because of social multiplier effects, the position in the network determines their effort activity so that more central persons exert more effort than less central individuals

Econometrics models

Let $Y_r = (y_{1,r}, \cdots, y_{n_r,r})'$, $X_r = (x_{1,r}, \cdots, x_{n_r,r})'$,

and $\epsilon_r = (\epsilon_{1,r}, \cdots, \epsilon_{n_r,r})'$.

$$
\begin{aligned}
Y_r &= \phi_1 G_r Y_r + X_r\beta + G_r^* X_r \gamma + \eta_r l_{n_r} + \epsilon_r, \\
Y_r &= \phi_2 G_r^* Y_r + X_r\beta + G_r^* X_r \gamma + \eta_r l_{n_r} + \epsilon_r, \\
Y_r &= \phi_1 G_r Y_r + \phi_2 G_r^* Y_r + X_r\beta + G_r^* X_r \gamma + \eta_r l_{n_r} + \epsilon_r.
\end{aligned}
$$

Identification for the **local aggregate model**

$$Y = \phi_1 GY + X\beta + G^* X\gamma + L\eta + \epsilon \tag{2}$$

Proposition 0.1

- *When G_r has non-constant row sums for some network r,* $\mathsf{E}(JZ_1)$ *of the local-aggregate network model (2) has full column rank if:* (i) $I_{n_r}, G_r, G_r^*, G_r G_r^*$ *are linearly independent and* $|\beta| + |\gamma| + |\eta_r| \neq 0$*; or* (ii) $G_r G_r^* = c_1 I_{n_r} + c_2 G_r + c_3 G_r^*$ *and* Λ_1 *has full rank.*

- *When G_r has constant row sums such that $g_{i,r} = g_r$ for all r,* $\mathsf{E}(JZ_1)$ *has full column rank if:* (iii) $I, G, G^*, GG^*, G^{*2}, GG^{*2}$ *are linearly independent and* $|\beta| + |\gamma| \neq 0$*;* (iv) I, G, G^*, GG^*, G^{*2} *are linearly independent,* $GG^{*2} = c_1 I + c_2 G + c_3 G^* + c_4 GG^* + c_5 G^{*2}$*, and* Λ_2 *has full rank; or* (v) $g_r = g$ *for all* r*,* I, G^*, G^{*2}, G^{*3} *are linearly independent, and* $\phi_1 \beta g + \gamma \neq 0$*.*

The identification conditions for the *local-aggregate* model are weaker than the conditions for the *local-average* model in Bramoullé et al. (2009).

Here is an example (star-shaped network) where identification is possible for the local-aggregate model but fails for the local-average model.

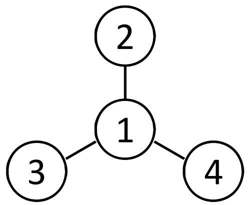
The adjacency matrix G is a block-diagonal matrix with diagonal blocks being G_r.

For the row-normalized adjacency matrix G^*, it is easy to see that $G^{*3} = G^*$. Thus local-average model is not identified.

G_r has non-constant row sums and $I_{n_r}, G_r, G_r^*, G_r G_r^*$ are linearly independent: local-aggregate model can be identified.

Identification easier with the local-aggregate model

Figure 2 gives an example where identification is possible for the local-aggregate model but fails for the local-average model.

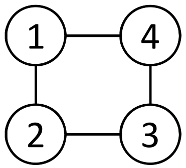


$$G_r = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad G_r^* = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

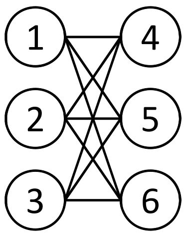
Here is another example where the local-average model cannot be identified while the local-aggregate model can.

Top of Figure 3 (regular network or circle).

Bottom of Figure 3 (bi-partite network).



$$G_1 = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad G_1^* = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix},$$



$$G_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$G_2^* = \begin{bmatrix} 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \end{bmatrix}.$$

For these two networks, the adjacency matrix G is a block-diagonal matrix with diagonal blocks being either G_1 or G_2 given in Figure 3.

For the row normalized adjacency matrix G^*, it is easy to see that $G^{*3} = G^*$. Local-average model not identified.

The two different types of networks have different row sums,

I, G, G^*, GG^*, G^{*2} are linearly independent and $GG^{*2} = G$.

Local-aggregate model identified.

Which model is better?

J test for model selection

The local-aggregate and local-average models can be written more compactly as:

$$
\begin{aligned}
H_1 &: Y = \phi_1 GY + X^* \delta_1 + L\eta_1 + \epsilon_1, \\
H_2 &: Y = \phi_2 G^* Y + X^* \delta_2 + L\eta_2 + \epsilon_2,
\end{aligned}
$$

where $X^* = (X, G^* X)$, and δ_1, δ_2 are corresponding vector of coefficients.

The test of model H_1 against model H_2

To test against the model specification H_2, one can estimate the following augmented model of H_1,

$$Y = \alpha_1 Y_{H_2} + \phi_1 GY + X^* \delta_1 + L\eta_1 + \epsilon_1,$$

where Y_{H_2} is a predictor of Y under H_2 such that $Y_{H_2} = \phi_2 G^* Y + X^* \delta_2 + L\eta_2$.

Thus, a test of the null model against the alternative one would be in terms of the hypotheses: $H_0 : \alpha_1 = 0$ against $H_a : \alpha_1 \neq 0$.

If the estimated α_1 is insignificant, then this is evidence against model H_2.

The test of model H_2 against model H_1

The test of model H_2 against model H_1 can be carried out in a similar manner.

Consider the following augmented model of H_2,

$$H_2 : Y = \alpha_2 Y_{H_1} + \phi_2 G^* Y + X^* \delta_2 + L\eta_2 + \epsilon_2,$$

where Y_{H_1} is a predictor of Y under H_1 such that $Y_{H_1} = \phi_1 GY + X^* \delta_1 + L\eta_1$.

Thus, the test of the null model against the alternative would be in terms of the hypotheses $H_0 : \alpha_2 = 0$ against $H_a : \alpha_2 \neq 0$.

If the estimated α_2 is significant, then that is evidence against model H_2.

Empirical Application

Data

Unique database on friendship networks from the National Longitudinal Survey of Adolescent Health (AddHealth).

Adolescents' behavior in the United States by collecting data on students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in years 1994-95.

Every pupil attending the sampled schools on the interview day is asked to compile a questionnaire (*in-school data*) containing questions on respondents' demographic and behavioral characteristics, education, family background and friendship.

This sample contains information on roughly 90,000 students.

A subset of adolescents selected from the rosters of the sampled schools, about 20,000 individuals, is then asked to compile a longer questionnaire containing more sensitive individual and household information (*in-home and parental data*).

Those subjects of the subset are interviewed again in 1995–96 (wave II), in 2001–2 (wave III), and again in 2007-2008 (wave IV).

Here only wave I because the network information is only available in the first wave.

Friendships: Pupils were asked to identify their best friends from a school roster (up to five males and five females).

We denote a link from i to j as $g_{ij,r} = 1$ if i has nominated j as her friend in network r, and $g_{ij,r} = 0$, otherwise.

Four different outcomes: (i) school performance; (ii) sport activities, such as playing baseball, softball, basketball, soccer or football; (iii) screen activities, such as playing video or computer games, (iv) criminal activities.

Empirical results

Tests of ϕ_1 and ϕ_2 of the hybrid network model

Tests of the augmented models where the null hypothesis is $\alpha_1 = 0$, i.e., the local average model does not matter for the first model, and $\alpha_2 = 0$, i.e., the local aggregate model does not matter for the second model.

Screen activities: peer effects are not important in explaining own screen activity.

The latter appears to be explained by own characteristics and contextual effects.

For example, male, black and lower grade students are more likely to participate in screen activities than other students.

Sport activities: it is the *sum of the effort of the friends* (i.e. the local aggregate model) and not *their average effort* that matters for explaining own sport activity.

For **education** (i.e. GPA index), both social norms (local average) and social multiplier (local aggregate) matter.

However, the magnitude of the effects is *higher* for the local-average model compared to the local-aggregate one.

A one standard deviation increase in the average activity of individual i's reference group translates roughly into a 0.29 increase of a standard deviation of individual i's GPA score while it is only 0.10 for the sum of activity of friends.

For **crime** (i.e. crime index), only the local-aggregate model matters, i.e. sum of friends.

Policy implications

An effective policy for the *local-average* model would be to change people's perceptions of "normal" behavior (i.e. their social norm) so that a *group-based policy* should be implemented

For the *local-aggregate* model, this would not be necessary and an *individual-based policy* should instead be implemented.

Crime

Local-aggregate model: key-player policy is the most effective policy since the effort of each criminal and thus the sum of one's friends crime efforts will be reduced.

The removal of the key player can have large effects on crime because of the feedback effects or "social multipliers" at work.

Education

Teachers:

Debate in the United States of giving incentives to teachers to improve teacher quality.

If the local aggregate model is at work among teachers, then we would need to have a teacher-based incentive policy since teachers will influence each other.

If the local average model, then one should implement a school-based incentive policy because this will be the only way to change the social norm of working hard among teachers.

Students

Local-aggregate model: any individual-based policy (for example, *vouchers*) would be efficient.

Here local-average model: should change the social norm in the school or the classroom and try to implement the idea that it is "cool" to work hard at school (acting white literature).

Example of a policy that has tried to change the social norm of students in terms of education is the *charter-school* policy.

The charter schools are very good in screening teachers and at selecting the best ones.

In particular, the "No Excuses policy" (Angrist et al., 2010, 2012) is a highly standardized and widely replicated charter model that features a long school day, an extended school year, selective teacher hiring, strict behavior norms, and emphasizes traditional reading and math skills.

The main objective is to change the social norms of disadvantage kids by being very strict on discipline.

The local-average model can also help us design an adequate policy in terms of tracking at school (Betts, 2011).

Should we “track” students in a way that separates high achievers from low achievers or should we mix them?

Local-average: Should separate high achievers from low achievers but then have an exogenous intervention (charter school) on the low achievers in order to change their social norms.

Local aggregate: Classes should be heterogenous with respect to students’ test scores, with the high performing students distributed among the classes.

Under this scenario, high achievers will have a positive impact on low achievers but will not be able to change the social norm of the low achievers.

To sum-up:

An effective policy for the *local-average* model would be to change people's perceptions of "normal" behavior (i.e. their social norm) so that a *group-based policy* should be implemented

For the *local-aggregate* model, this would not be necessary and an *individual-based policy* should instead be implemented.

CONSUMPTION NETWORK EFFECTS

Giacomo De Giorgi[1]   Anders Frederiksen[2]   Luigi Pistaferri[3]

[1]Stanford, BREAD, CEPR and NBER
[2]Aarhus University
[3]Stanford, EIEF, CEPR and NBER

The Economics of Coordination and Communication
Ravello, 1-3 June 2012

Motivation

- Consumption may be influenced by socially proximate agents, i.e., friends, neighbors, co-workers, etc. (Veblen, 1899; Duesenberry, 1948)
- It is hard to find clean evidence of "peer effects"
    - peers definition
    - the "reflection" problem - endogeneity vs. correlated shocks
- The existence of network effects may be relevant from a policy point of view, as meaningful multiplier effects magnify intended effects of a policy
- Network effects may also create intertemporal and/or intratemporal distortions that may be policy relevant

Our goals

- Test if and how much *co-worker* networks affect consumption
- Investigate the possible mechanisms behind our findings
  - Conspicuous Consumption or *Status*
  - Keeping Up with the *Jensens* (KUJ)
  - Risk Sharing
- Understand aggregate implications of consumption network effects

Our paper

- Previous work:
  - individuals sharing similar characteristics (Maurer and Meier, 2008)
  - racial group within a US state (Charles et al., 2009)
  - neighbors within a zip code (Kuhn et al., 2011), or city (Ravina, 2010)
- Co-workers as a more reasonable group to investigate peer effects
  - Adult-life equivalent of class-mates
  - Spend more time with co-workers than with family members
  - Evidence from sociology - friendship as determinant of co-workership (Granovetter, 1995)
- Data
  - Administrative tax records on Danish households (including data on income and assets), coupled with information on place of work
  - Use network structure to tackle identification problem - spouses add *nodes* to otherwise isolated networks
  - Use a household consumption survey to investigate mechanisms

Preview of the Results

- Peers' consumption affects household consumption
- In particular husbands' co-workers' consumption has a significant, non-negligible, positive effect
- Estimate a multiplier effect of about 1.4
- Preliminary evidence
  - Mechanisms: suggestive of "Keeping-up-with-the-Jensens"
  - Effects on aggregate consumption

An empirical model: Linear-in-means (1)

Standard Manski's problem (1993)

$$c_{h,t} = \alpha + \beta_1 \overline{c_t^w} + \beta_2 \overline{c_t^m} + \gamma_1 \overline{X_t^w} + \gamma_2 \overline{X_t^m} + \delta_1 X_{h,t}^w + \delta_2 X_{h,t}^m + \epsilon_{h,t}$$

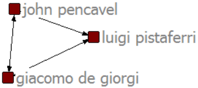
- Where h, t, w, m indicate household, time, wife and husband respectively.
  - $\overline{c_{h,t}^w}, \overline{c_{h,t}^m}$: (average) logged (per adult equivalent) consumption levels of the wife's and husband's co-workers;
  - $\overline{X_{h,t}^w}, \overline{X_{h,t}^m}$: (average) characteristics of the wife's and husband's co-workers;
  - $X_{h,t}^w, X_{h,t}^m$: the wife's and husband's observable characteristics (permanent income determinants)

- Main parameters of interest:
  - $\beta' s$: *endogenous* effect
  - $\gamma' s$: *contextual* effects
  - $\delta' s$: ancillary parameters of interest
  - *Correlated* effects (through $\epsilon_{h,t}$) also allowed for (i.e., common firm shock).

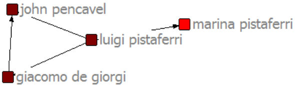
An empirical model: Linear-in-means (2)

- Identification of the parameters of interest in this model is notoriously problematic:
  - Reflection problem $\longrightarrow$ Simultaneity
  - Self-selection/Correlated effects $\longrightarrow E\left(\varepsilon|X,\overline{X}\right) \neq 0$
- Our solution:
  - Assume networks do not perfectly overlap
  - Construct the network of co-workers (distance-1 peers), co-workers' spouses (distance-2), co-workers of co-workers' spouses (distance-3), and so on.
  - "Valid" instruments for co-workers' ("distance-1") consumption: The X's of distance-3 peers.

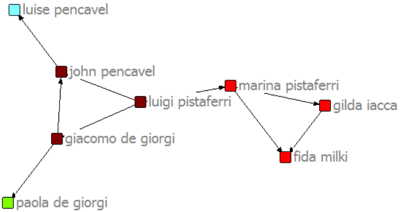
Identification: the Stanford Example – Luigi's coworkers



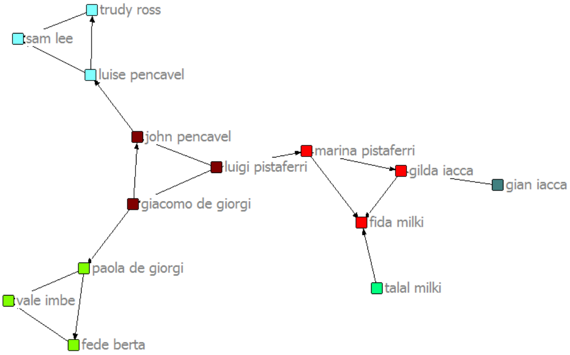
Identification: the Stanford Example – Luigi (d-1)



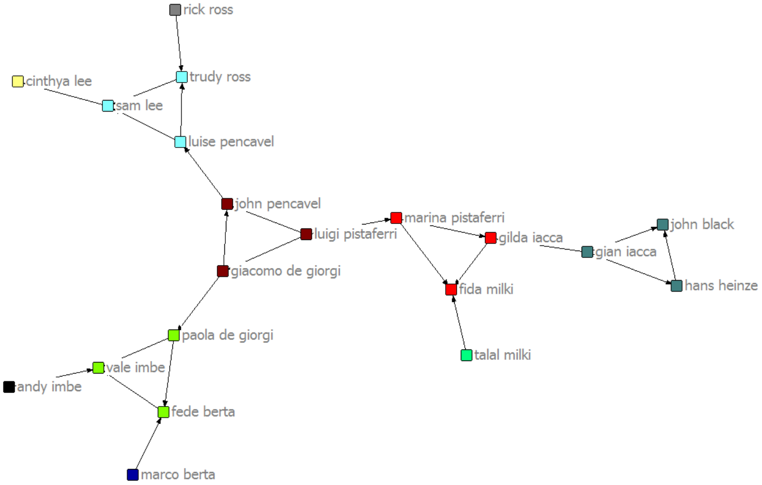
Identification: the Stanford Example – Luigi (d-2)



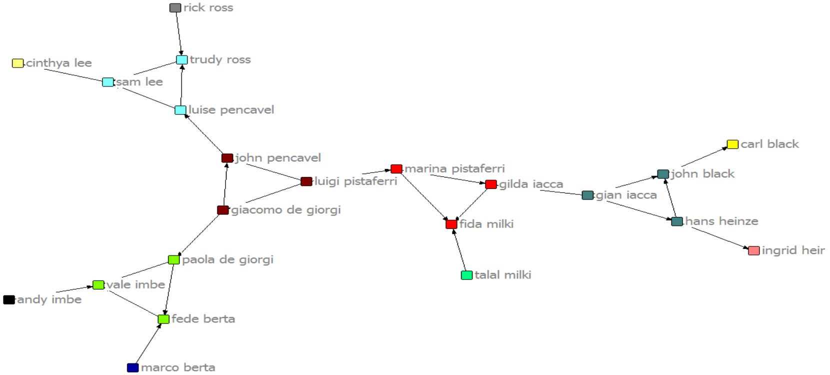
Identification: the Stanford Example – Luigi (d-3)



Identification: the Stanford Example – Luigi (d-4)

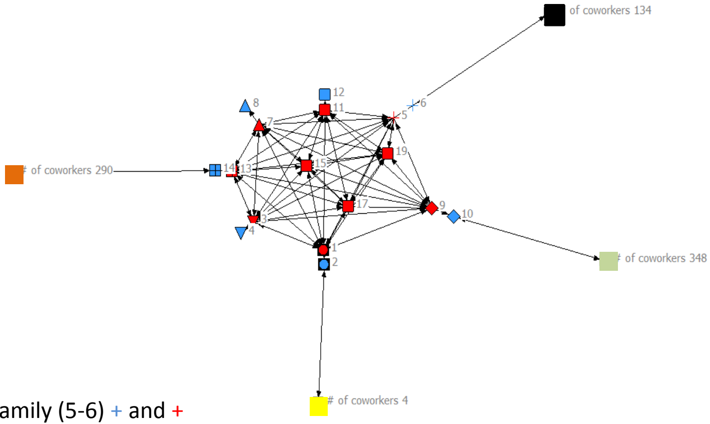


Identification: the Stanford Example – All



- Pistaferri's (distance-1) peers. Luigi (G. De Giorgi+J. Pencavel), Marina (G. Iacca+ F.Milki)
- Instruments (distance-3 peers): (F. Berta+V. Imbe) , (S. Lee+T. Ross), and (J. Black +H. Heinze)

An Actual Network from our data



Consider family (5-6) + and +

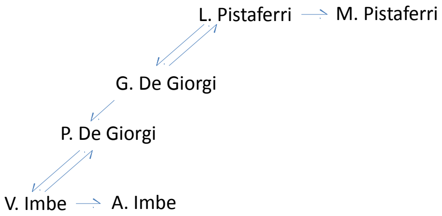
His peers: Red symbols

Her peers: ■

Instruments: ■ , ■ , and ■

Identification: An Example (1)

- Take simplified "Stanford" network:



$$
\begin{aligned}
c_P &= \alpha + \beta c_D + \gamma x_D + \delta x_P + \varepsilon_P \\
c_D &= \alpha + \beta \left(\frac{c_P + c_I}{2} \right) + \gamma \left(\frac{x_P + x_I}{2} \right) + \delta x_D + \varepsilon_D \\
c_I &= \alpha + \beta c_D + \gamma x_D + \delta x_I + \varepsilon_I
\end{aligned}
$$

Identification: An Example (2)

- c_D is endogenous in c_P equation, i.e., $E(\varepsilon_P | c_D) \neq 0$:

$$
\begin{aligned}
c_P &= \alpha + \beta c_D + \gamma x_D + \delta x_P + \varepsilon_P \\
c_D &= \alpha + \beta \left(\frac{c_P + c_I}{2} \right) + \gamma \left(\frac{x_P + x_I}{2} \right) + \delta x_D + \varepsilon_D \\
c_I &= \alpha + \beta c_D + \gamma x_D + \delta x_I + \varepsilon_I
\end{aligned}
$$

- In fact: Write reduced form for c_D:

$$c_D = f(x_P, x_D, x_I, \varepsilon_D, \varepsilon_P, \varepsilon_I)$$

  so c_D moves with ε_P.

- Also, reduced form for c_I is:

$$c_I = g(x_P, x_D, x_I, \varepsilon_D, \varepsilon_P, \varepsilon_I)$$

Identification: An Example (3)

- Need an instrument for c_D in c_P's equation
- What are the variables that explain c_D (condition I for an IV to be valid) and do not belong to c_P (condition II for an IV to be valid)?

$$
\begin{aligned}
c_P &= \alpha + \beta c_D + \gamma x_D + \delta x_P + \varepsilon_P \\
c_D &= \alpha + \beta \left(\frac{c_P + c_I}{2} \right) + \gamma \left(\frac{x_P + x_I}{2} \right) + \delta x_D + \varepsilon_D
\end{aligned}
$$

- List includes:
  1. $c_P \longrightarrow$ of course not
  2. $c_I \longrightarrow$ No, because in the reduced form $c_I = g\left(x_P, x_D, x_I, \varepsilon_D, \underline{\varepsilon_P}, \varepsilon_I\right)$
  3. $x_P, x_D \longrightarrow$ No, because they are "included" variables
  4. $x_I \longrightarrow$ Yes! Distance-3 peer's exogenous characteristics.

Identification: A General Model (1)

- Identification strategy similar to Bramoullé et al. (2009), Calvó-Armengol et al. (2009), De Giorgi et al. (2010): It relies on the network structure
- In matrix notation, household level analysis, and with a single covariate $\mathbf{X}$:

$$\mathbf{c} = \alpha \mathbf{i} + \beta \mathbf{G} \mathbf{c} + \gamma \mathbf{G} \mathbf{x} + \delta \mathbf{x} + \epsilon$$

  where $\mathbf{G}$ is a $n \times n$ (weighting) adjacency matrix with generic element G_{hk}
- $G_{hk} = \mathbf{1}\{1/(n_h - 1)$ if h, k co-workers$\}$, with n_h the size of h's firm, and $h, k = 1, ..., n$
- Assume ψ a network level shock, so that $E(\epsilon|\mathbf{x}) \neq 0$, $E(\epsilon|\psi, \mathbf{x}) = 0$. Premultiply by $\mathbf{G}$:

$$\mathbf{G}\mathbf{c} = \alpha \mathbf{i} + \beta \mathbf{G}^2 \mathbf{c} + \gamma \mathbf{G}^2 \mathbf{x} + \delta \mathbf{G} \mathbf{x} + \mathbf{G}\epsilon$$

- Take within network transformation (which eliminates correlated effect):

$$(\mathbf{I} - \mathbf{G})\,\mathbf{c} = \beta \mathbf{G}(\mathbf{I} - \mathbf{G})\,\mathbf{c} + \gamma \mathbf{G}(\mathbf{I} - \mathbf{G})\,\mathbf{x} + \delta(\mathbf{I} - \mathbf{G})\,\mathbf{x} + (\mathbf{I} - \mathbf{G})\,\epsilon$$

Identification: A General Model (2)

- Reduced form (no worker is isolated) of above model is:

$$(\mathbf{I}-\mathbf{G})\,\mathbf{c} = (\mathbf{I}-\beta\mathbf{G})^{-1}(\gamma\mathbf{G}+\delta)(\mathbf{I}-\mathbf{G})\,\mathbf{x} + (\mathbf{I}-\beta\mathbf{G})^{-1}(\mathbf{I}-\mathbf{G})\,\epsilon$$

- Recall that $(\mathbf{I}-\beta\mathbf{G})^{-1} = \lim_{l\to\infty}\sum_{k=0}^{l}\beta^k\mathbf{G}^k$, for $|\beta| \in (0,1)$. Expand (2^{nd} term) and rearrange to yield:

$$(\mathbf{I}-\mathbf{G})\,\mathbf{c} = \delta(\mathbf{I}-\mathbf{G})\,\mathbf{x} + (\gamma+\beta\delta)(\mathbf{I}-\mathbf{G})\,\mathbf{G}\mathbf{x} + \beta(\gamma+\beta\delta)(\mathbf{I}-\mathbf{G})\,\mathbf{G}^2\mathbf{x} + \beta^2\gamma(\mathbf{I}-\mathbf{G})\,\mathbf{G}^3\mathbf{x} + \mathbf{v}$$

- As long as $(\gamma+\beta\delta) \neq 0$, and $\mathbf{I}, \mathbf{G}, \mathbf{G}^2, \mathbf{G}^3$ linearly independent — then the 3 parameters of interest are identified
- Note: Identification requires availability of co-workers of co-workers' spouses (or distance-3 nodes, as we consider husband and wife to be distance-1 peers).

Data (1)

- We use administrative longitudinal tax records for the Danish population (1980-1996)
  - Wealth tax abolished in 1996
- We match these data with the IDA, an employer-employee data set , which includes demographics and firm ID (plant level) $\Rightarrow$ Co-workers
- We also match the tax records with the DES, a CEX-type cross-sectional survey (1994-96)

Data (2)

- Household heads aged 18-65 (Descriptives)
  - Focus on couples where both spouses working
  - Drop couples where spouses are working in the same plant
- Consumption is not measured in administrative tax data.
  - We use the dynamic budget constraint to calculate total consumption
  - Between 1980 and 1996, households paid a tax on assets
  - Consumption is calculated as the difference between after-tax annual income and asset changes

$$C_{ht} = Y_{ht} - \Delta A_{ht}$$

  - Similar to Browning and Leth-Petersen (2005) and Koijen, Van Nieuwerburgh and Vestman (2012) (Details)
- We use the DES primarily to investigate the "mechanisms" behind our findings

Peers

- In the baseline results peers are co-workers within the same plant *and* occupation (blue collar, white collar, manager)
- We also consider alternative definitions of co-workers
- Future drafts: Neighbors (multiplexity), and alternative weighting scheme

Results: Main Table (1)

Table 4: Baseline Results

	OLS (1)	OLSX (2)	IVX (3)	IVFEX (4)
Wife's peers ln C	0.41***	0.16***	0.20***	0.06
	(0.001)	(0.002)	(0.027)	(0.050)
Husband's peers ln C	0.46***	0.20***	0.44***	0.29***
	(0.001)	(0.002)	(0.030)	(0.050)
p-value $[\beta_h = \beta_w]$	0.0000	0.0000	0.0000	0.0000
F-stat first stage				
Wife			976.64	
Husband			767.98	
Endogeneity Wu-Hausman F [p-value]			0.00	
Observations		4,514,496		

Dependent variable: Log of adult equivalent consumption. **Controls**: Age, Age^2, Years of schooling, Occupation dummies, Firm size. **Contextual controls**: Age, Age^2 , Years of schooling, share of female peers. **Household controls**: Year dummies, Municipality dummies, Sector dummies, Sector$\times$ Year dummies. **IV's**: age, years of schooling, share of female peers, occupation dummies

Table 5: Imposing the restriction $\beta_h = \beta_w$

	OLS	OLSX	IVX	IVFEX
	(1)	(2)	(3)	(4)
Household peers ln C	0.44***	0.18***	0.31***	0.17***
	(0.001)	(0.001)	(0.019)	(0.034)
First stage F	–	–	907.44	
Endogeneity Wu-Hausman F [p-value]			0.00	
N	4,514,496			

Dependent variable: Log of adult equivalent consumption. **Controls**: Age, Age^2, Years of schooling, Occupation dummies, Firm size. **Contextual controls**: Age, Age^2 , Years of schooling, share of female peers. **Household controls**: Year dummies, Municipality dummies, Sector dummies, Sector$\times$ Year dummies. **IV's**: age, years of schooling, share of female peers, occupation dummies

Results: Summary

- OLS is biased because of usual reasons: Self-selection, endogeneity, reflection, measurement error
- IV's and FE address all these issues at once
- Our preferred estimator is IVFEX, where a clear difference emerges between husband's and wife's network effects
- We estimate elasticities of 0.29 and 0.06 (insignificant) for husband and wife, respectively. Pooled elasticity: 0.17
- These estimates imply a multiplier effect of about 1.2-1.4 (pooled/husbands)
  - But aggregate effects will depend also on "degree of connectedness"