NETWORK DATA

(M2 ETE Econometrics II)

Koen Jochmans

Toulouse School of Economics

Last revised on February 17, 2023

Outline

Social interactions

Agents do not make decisions in isolation.

- Market-entry decisions of firms.
- Effort-level decisions by workers in a team/students in a classroom.

Agents can incur indirect benefits/costs from actions of others

- Individual health risk as a function of vaccine uptake.
- Immigration propensity as a function of migrant stock.

Interactions are at the heart of economic theory. But econometric practice has been slow to follow.

Data structure: Peers/Reference groups

We have n agents that interact in a network.

We observe (y_i, x_i) for each individual.

Interactions can be summarized through the $n \times n$ adjacency matrix

$$(A)_{i,j} := \begin{cases} 1 & \text{if } j \text{ is connected to } i \\ 0 & \text{if not} \end{cases} .$$

The ith row of A indicates the peer group/reference group of individual i.

The notation allows for directed interaction, where $(A)_{i,j} \neq (A)_{j,i}$.

Typical sampling framework consists of seeing many independent networks, $g = 1, \ldots, G$, and network g has n_g units. Think, e.g., many classrooms. For now, the notation leaves this implicit.

Peer effects

Several types of possible effects of interaction with peers:

Correlated effects.
Individuals behave similarly because they are in the same environment.

Exogenous effects, contextual effects, or spillover effects.
The propensity of an individual to behave in a certain way varies with the distribution of characteristics in the reference group.

Endogenous effects.
The propensity of an individual to behave in a certain way varies with the prevalence of that behavior in the reference group.

Example: Vaccine efficacy

Let

$$y_i = \begin{cases} 1 & \text{if } i \text{ is infected} \\ 0 & \text{if not} \end{cases}, \qquad x_i = \begin{cases} 1 & \text{if } i \text{ is vaccinated} \\ 0 & \text{if not} \end{cases}.$$

The regression

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

ignores any effect of peers being vaccinated or being infected.

A sensible way forward is to include the vaccination propensity among peers:

$$y_i = \alpha + \beta x_i + \gamma \tilde{x}_i + \varepsilon_i, \qquad \tilde{x}_i := \sum_{j=1}^{n} \left(\frac{(A)_{i,j}}{\sum_{j'=1}^{n} (A)_{i,j'}} \right) x_j.$$

Here β is the direct effect of the vaccine, and γ captures the indirect spillover effect.

In a treatment-effect context the spillover effect is an example of interference.

If vaccination decisions are correlated among peers the simple regression will not yield the direct effect (unless $\gamma = 0$).

The extended specification still ignores the infection rate among peers.

We further generalize the specification to

$$y_i = \alpha + \delta \tilde{y}_i + \beta x_i + \gamma \tilde{x}_i + \varepsilon_i,$$

where $\tilde{y}_i$ is defined in the same manner as $\tilde{x}_i$.

This is an example of the so-called linear-in-means model.

A simple special case has

$$(A)_{i,j} = 1 \text{ for all } j \neq i,$$

so that everyone interacts will everyone else, but the above allows general structures of reference groups.

We can further extend the model by bringing in further characteristics and by letting unobservables be correlated between peers.

Linear-in-means model

Let

$$(H)_{i,j} = \begin{cases} \frac{1}{\sum_{j'=1}^{n}(A)_{i,j'}} & \text{if } (A)_{i,j} = 1 \\ 0 & \text{otherwise} \end{cases}$$

be the row-normalized adjacency matrix.

This is a transition matrix on the network induced by A.

Then

$$\tilde{y}_i = \sum_{j=1}^{n}(H)_{i,j}y_j, \qquad \tilde{x}_i = \sum_{j=1}^{n}(H)_{i,j}x_j.$$

and can write the model in matrix notation:

$$y = I\alpha + \delta Hy + \beta x + \gamma Hx + \varepsilon.$$

Complete the model with an exogeneity assumption:

$$\mathbb{E}(\varepsilon|A, x) = 0;$$

the covariates and the network structure are (strictly) exogenous.

Reflection problem

The model has a simultaneity issue embedded in it.

To make progress we work out the reduced form for Hy.

First re-arrange the equation to get

$$(I - \delta H)y = I\alpha + \beta x + \gamma Hx + \varepsilon.$$

Next (provided that $\delta \neq 1$ and each unit has at least one peer) we can invert to obtain

$$\begin{aligned} y &= (I - \delta H)^{-1} (I\alpha + \beta x + \gamma Hx + \varepsilon) \\ &= \left(\sum_{k=0}^{\infty} \beta^k H^k \right) (I\alpha + \beta x + \gamma Hx + \varepsilon) \\ &= \frac{\alpha}{1-\delta} I + \beta x + (\delta\beta + \gamma) \left(Hx + H^2 x + \cdots \right) + (\varepsilon + H\varepsilon + \cdots) \end{aligned}$$

Reduced form and instrumental variables

Then pre-multiplying by H and taking expectations yields

$$\mathbb{E}(Hy|H,x) = \frac{\alpha}{1-\delta} I + \beta Hx + (\delta\beta + \gamma)\left(H^2x + H^3x + \cdots\right)$$

If $\gamma = -\delta\beta$ endogenous and exogenous effects cancel each other and this is just a linear function of Hx.

If not, then $H^2x, H^3x, \cdots$ provide exogenous variation in Hy and can be used as instrumental variables.

The simplest case is just using H^2x.

Informally, this amounts to using peers of peers to instrument peers. Indeed,

$$(H^2)_{i,j} = \sum_{j'=1}^{n} (H)_{i,j'} (H)_{j',j} = \Pr(\text{Going from } i \text{ to } j \text{ in 2 steps}).$$

We do need I, H, and H^2 to be linearly independent, otherwise the added moment condition is a linear combination of the ones already included, and we do not get identification.

Optimal instrument

The information obtained from included further $H^q y$ as instruments will normally be decreasing in q; as $q \to \infty$, H^q converges to the steady state of the Markov chain on the network.

A just-identified MM problem that uses all information can be obtained by working with the optimal instrument.

Given that the reduced-form is linear, under homoskedasticity, the (feasible) optimal instrument is (an estimated version of)

$$\mathbb{E}(Hy|H,x) = H(I - \delta H)^{-1}(I\alpha + \beta x + \gamma Hx)\,.$$

Intuitively, this exploits peers of peers at all distances.

With heteroskedasticity the optimal instrument depends on the conditional variance of the shocks; this leads to a more complicated form, and more difficult feasible instrument.

Variation in peer groups

The relevance condition requires that units are exposed to different peers.

The absence of links in A yields exclusion restrictions in the n-dimensional simultaneous-equation model for y.

A relevant case where this condition fails is in the 'complete' network, where

$$(A)_{i,j} = 1$$

for all $i \neq j$. Think of students (units) in a classroom (networks), where all pupils are peers of one another.
For $n = 3$,

$$H = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

and

$$H^2 = \frac{1}{4}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \frac{1}{2}I + \frac{1}{2}H$$

so identification fails.

Variation in peer group sizes

However, in the complete network we can leverage variation in peer group size.

We have

$$\sum_{j=1}^{n}(H)_{i,j}\,x_j = \frac{1}{n-1}\sum_{j\neq i}x_j = \frac{1}{n-1}(n\,\bar{x} - x_i),$$

and

$$\sum_{j=1}^{n}(H)_{i,j}\,y_j = \frac{1}{n-1}\sum_{j\neq i}y_j = \frac{1}{n-1}(n\,\bar{y} - y_i),$$

and so

$$\left(1+\delta\frac{n}{n-1}\right)y_i = \alpha + \left(\beta - \gamma\frac{n}{n-1}\right)x_i + \gamma\frac{n}{n-1}\bar{x} + \delta\frac{n}{n-1}\bar{y} + \varepsilon_i$$

Averaging over i yields

$$\left(1+\delta\frac{n}{n-1}\right)\bar{y} = \alpha + \left(\beta - \gamma\frac{n}{n-1}\right)\bar{x} + \gamma\frac{n}{n-1}\bar{x} + \delta\frac{n}{n-1}\bar{y} + \bar{\varepsilon}.$$

By taking the difference between both these equations we find

$$\left(1+\delta\frac{n}{n-1}\right)(y_i-\bar{y}) = \left(\beta-\gamma\frac{n}{n-1}\right)(x_i-\bar{x}) + (\varepsilon_i-\bar{\varepsilon}).$$

The coefficient from the regression of $(y_i - \bar{y})$ on $(x_i - \bar{x})$ is identified. It equals

$$\frac{\left(\beta-\gamma\frac{n}{n-1}\right)}{\left(1+\delta\frac{n}{n-1}\right)}$$

and varies with n.

To learn the three parameters β, γ, δ it suffices to recover this coefficient for three (or more) different group sizes.

This then gives three equations from which the parameters can be uniquely recovered, provided that $\gamma \neq -\delta\beta$.

Note that this is the same condition as before

Can do estimation by minimum distance after least-squares.

Exogeneity

Note that, in the presence of endogenous effects, identification requires:

The presence of exogenous regressors.

Identification fails when $\gamma = \beta = 0$.

Peer-group formation to be exogenous.

Randomized assignment to peer groups has been exploited.

The identifying power of exogenous regressors is similar in spirit to in the general dynamic panel data problem.

If the network is endogenously formed both $\tilde{x}_i$ and $\tilde{y}_i$ become endogenous in the linear-in-means model.

Accounting for correlated effects

The model can be augmented with network-level fixed effects (i,e, network specific intercepts).

Then first-demean the data and largely proceed as before.

Will not do more detail here.

The setting of many small networks brings us close to an autoregressive panel data problem where we get identification through the presence of strictly exogenous regressors.

Testing for endogenous effects

In the absence of endogenous effects identification is straightforward, and estimation can be done by least squares.

Can be useful to first test whether such effects are present.

First, ignore regressors. Baseline model has

$$y_{g,i} = \alpha_g + \delta \tilde{y}_{g,i} + \varepsilon_{g,i}.$$

In matrix form

$$y_g = I_{n_g} \alpha_g + \delta H_g y_g + \varepsilon_g$$

The normal-equation for within-network least-squares is

$$\sum_{g=1}^{G} y_g' H_g' M_g \left(y_g - \delta H_g y_g \right) = 0$$

and is biased.

However, under the null that $\delta = 0$,

$$\mathbb{E}\left(\sum_{g=1}^{G} y_g' H_g' M_g \left(y_g - \delta H_g y_g \right) \right) = \sum_{g=1}^{G} \mathbb{E}\left(\varepsilon_g' H_g' M_g \varepsilon_g \right) = \sum_{g=1}^{G} \text{trace}\left(H_g' M_g \Sigma_g \right).$$

Take $\Sigma_g = \sigma_g^2 I_{n_g}$ for simplicity. Then

$$\text{trace}\left(H_g' M_g \Sigma_g \right) = \sigma_g^2 \text{trace}\left(M_g H_g \right) = -\sigma_g^2 \frac{\iota_{n_g}' H_g \iota_{n_g}}{n_g} = -\sigma_g^2.$$

Under the null, an unbiased estimator of σ_g^2 is

$$\frac{y_g' M_g y_g}{n_g - 1}$$

so that

$$\sum_{g=1}^{G} y_g' H_g' M_g y_g + (n_g - 1)^{-1} y_g' M_g y_g$$

has mean zero.

This is just a sample average of mean-zero random variables across the number of networks G.

Can build a t-test based on this to test for the presence of endogenous peer effects.

In the presence of regressors, we first partial them out by least squares and then proceed as before.

Nonlinear model

Market entry decision of two firms:

Produce (enter) if profit is positive:

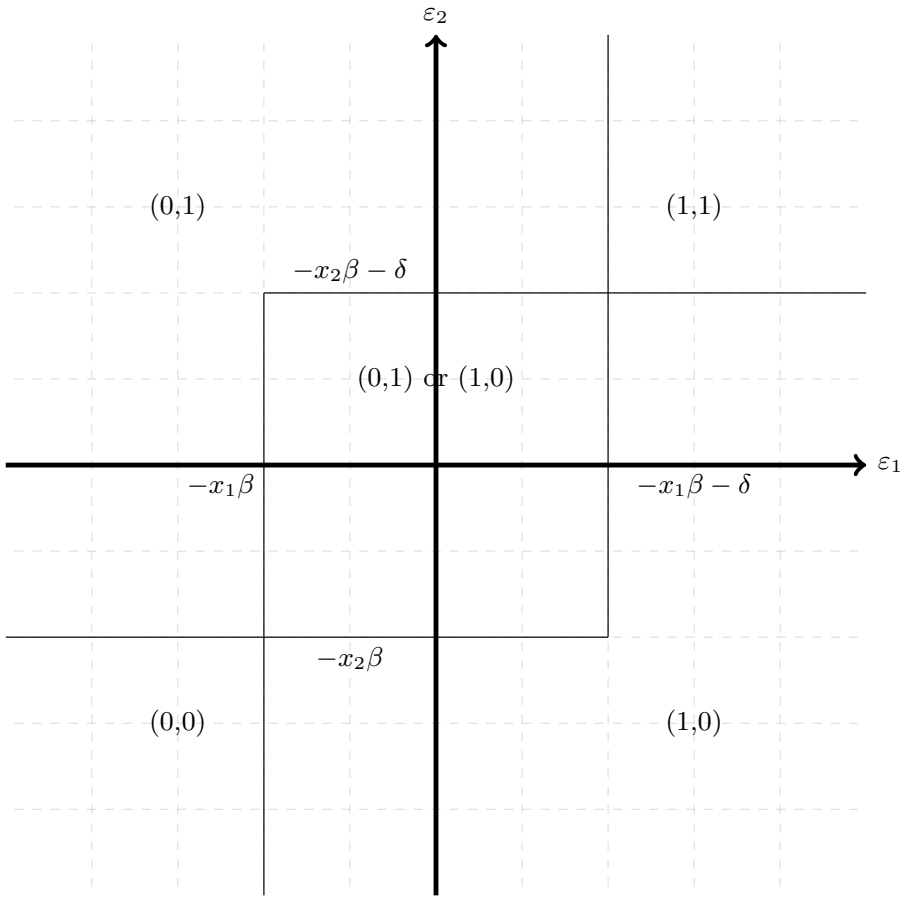
$$y_i = \begin{cases} 1 & \text{if } x_i\beta + \delta y_j + \varepsilon_i > 0 \\ 0 & \text{if not} \end{cases} .$$

Production decision of other firm affects own profits.

		Firm 2	
		0	1
Firm 1	0	$(0, 0)$	$(0, x_2\beta + \varepsilon_2)$
	1	$(x_1\beta + \varepsilon_1, 0)$	$(x_1\beta + \delta + \varepsilon_1, x_2\beta + \delta + \varepsilon_2)$

(Here take $\delta < 0$)

Outcome depends on where $(\varepsilon_1, \varepsilon_2)$ lies in the plot



This is an incomplete model.

Outline

Dyadic interaction

Units (i, j) with $i < j$ in the network pair up to jointly produce an output y_{ij}.

We could have $y_{ij} \neq y_{ji}$ but do not do so here.

So, if all units interact we have

$$\frac{n(n-1)}{2}$$

outcomes.

A linear model

A simple linear-specification is

$$y_{ij} = \alpha_i + \alpha_j + x_{ij}\beta + \varepsilon_{ij}.$$

together with

$$\mathbb{E}(\varepsilon_{ij}|x_{12}, \ldots, x_{(n-1)n}, \alpha_1, \ldots, \alpha_n) = 0$$

(strict exogeneity).

With all units i interacting with all $j \neq i$ this is very similar to a two-way panel data problem.

Estimation of β here is not difficult; can use least-squares dummy variable regression.

This model is often used in its linear probability form to consider network formation.

So, $y_{ij} \in \{0, 1\}$ codes whether units are connected ('are friends').

The degree of unit i is $\sum_{i<j} y_{ij} + \sum_{i>j} y_{ji}$ (i.e., the number of friends).

There, x_{ij} would be a distance measure, capturing whether units are similar in some sense.

Then $x_{ij}\beta > 0$ represents a taste for homophily

Here, the α_i would capture degree heterogeneity across units, i.e., a taste for having friends.

Least squares

We solve

$$\min_{\alpha_1,\ldots,\alpha_n} \sum_{i=1}^{n} \sum_{i<j} (y_{ij} - \alpha_i - \alpha_j)^2.$$

The first-order condition for a given α_i is

$$\left(\sum_{i<j} y_{ij} + \sum_{i>j} y_{ji} \right) - (n-1)\alpha_i - \sum_{j \neq i} \alpha_j = 0.$$

Re-arranging gives

$$\alpha_i = \frac{1}{n-2} \left(\sum_{i<j} y_{ij} + \sum_{i>j} y_{ji} - \sum_{j=1}^{n} \alpha_j \right).$$

Summing over α_i yields

$$\frac{2}{n} \sum_{i=1}^{n} = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{i<j} y_{ij} =: \bar{y}.$$

Then

$$\hat{\alpha}_i = \frac{1}{n-2}\left((n-1)\bar{y}_i - \frac{n}{2}\bar{y}\right)$$

with

$$\bar{y}_i := \frac{1}{n-1}\left(\sum_{i<j} y_{ij} + \sum_{i>j} y_{ji}\right)$$

A small calculation verifies that

$$\mathbb{E}(\hat{\alpha}_i | \alpha_1, \ldots, \alpha_n) = \alpha_i.$$

The variance of $\hat{\alpha}_i$ is of the order n^{-1}.

Functionals

We typically think of α_i as i.i.d. draws of a random variable from some distribution.

We might care about functionals of this distribution.

For example,

$$\mathbb{E}(\alpha_i), \qquad \mathbb{E}(\alpha_i^2) - \mathbb{E}(\alpha_i)^2.$$

The plug-in estimators of the mean and variance are

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i, \qquad \frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i\right)^2.$$

Note that

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i = \frac{1}{n}\sum_{i=1}^{n}\alpha_i + \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha}_i - \alpha_i)$$

has mean $\mathbb{E}(\alpha_i)$ and variance

$$\frac{\mathrm{var}(\alpha_i)}{n} + o(n^{-1}).$$

Estimation noise is irrelevant and we have

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i - \mathbb{E}(\alpha_i)\right) \xrightarrow{d} \mathbf{N}(0, \mathrm{var}(\alpha_i)).$$

For nonlinear functionals such as the variance we have a bias.

As

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i^2 = \frac{1}{n}\sum_{i=1}^{n}\alpha_i^2 + \frac{2}{n}\sum_{i=1}^{n}\alpha_i\left(\hat{\alpha}_i - \alpha_i\right) + \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha}_i - \alpha_i)^2$$

we have

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i^2\right) = \mathbb{E}(\alpha_i^2) + \mathbb{E}((\hat{\alpha}_i - \alpha_i)^2) = \mathbb{E}(\alpha_i^2) + \mathbb{E}(\mathrm{var}(\hat{\alpha}_i|\alpha_1, \ldots, \alpha_n)).$$

The bias is of order n^{-1}, which is the same order as the variance.

Still obtain

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i^2 - \mathbb{E}(\alpha_i^2)\right) \xrightarrow{d} \mathbf{N}(0, \mathrm{var}(\alpha_i^2)).$$

May wish to consider bias correction in practice.

Limited interaction

The situation is more complicated when not all units in the network interact.

In practice, many units have a small degree.

The difficulty is that $\text{var}(\hat{\alpha}_i|\alpha_1, \ldots, \alpha_n)$ does not vanish as the network grows.

The bias in nonlinear functionals can become important, and even dominate the standard deviation.

Then inference then breaks down.

Sometimes possible to remove bias. But small degrees imply high leverage and can lead to non-standard asymptotic behavior.

We discuss this in some more detail later.

Nonlinear models

A binary-choice version would be, e.g.,

$$y_{ij} = \alpha_i + \alpha_j + x'_{ij}\beta \geq \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim \text{i.i.d. } F$$

In dense networks we can do maximum likelihood estimation.

If interest lies in, say, β, we need to correct for asymptotic bias in the limit distribution.

Can adapt the type of bias correction that was discussed for long panel data.

Sparse networks again more complicated.

Logit model

Let

$$z := {}^1\!/_2(y_{13} - y_{14}) - {}^1\!/_2(y_{23} - y_{24}).$$

Then

$z = {}^1\!/_2(1-0) - {}^1\!/_2(0-1) = \quad 1$ in subgraph 1, and

$z = {}^1\!/_2(0-1) - {}^1\!/_2(1-0) = -1$ in subgraph 2.

Conditional on $z \in \{-1, 1\}$, z (and regressors and fixed effects) is Bernoulli with probability

$$\frac{\mathbb{P}(z=1)}{\mathbb{P}(z=1)+\mathbb{P}(z=-1)} = \frac{1}{1+\frac{\mathbb{P}(z=-1)}{\mathbb{P}(z=\ \ 1)}}.$$

Here,

$$\frac{\mathbb{P}(z=-1)}{\mathbb{P}(z=\ \ 1)} = \frac{\mathbb{P}(y_{13}=1, y_{14}=0, y_{23}=0, y_{24}=1)}{\mathbb{P}(y_{13}=0, y_{14}=1, y_{23}=1, y_{24}=0)} = e^{-((x_{13}-x_{14})-(x_{23}-x_{24}))'\beta}$$

which does not depend on fixed effects.

The conditional success probability has a logit structure with a 'difference in differences' form.

Outline

Network formation

Say that

$$(A)_{i,j} = G(\alpha_i, \alpha_j)$$

and go back to the specification

$$y_i = \alpha + \beta x_i + \gamma \tilde{x}_i + \varepsilon_i, \qquad \tilde{x}_i := \sum_{j=1}^{n} \left(\frac{(A)_{i,j}}{\sum_{j'=1}^{n} (A)_{i,j'}} \right) x_j.$$

Network exogeneity requires that α_i and ε_i are independent.

Reference groups are often self selected.

Then

$$y_i = \alpha + \beta x_i + \gamma \left(\sum_{j=1}^{n} \left(\frac{G(\alpha_i, \alpha_j)}{\sum_{j'=1}^{n} G(\alpha_i, \alpha_{j'})} \right) x_j \right) + \varepsilon_i,$$

implies an endogeneity problem even if the regressor is strictly exogenous.

Instrumental variables

In our example where $(A)_{i,j} = G(\alpha_i, \alpha_j)$,

$$y_i = \alpha + \beta x_i + \gamma \left(\sum_{j=1}^{n} \left(\frac{G(\alpha_i, \alpha_j)}{\sum_{j'=1}^{n} G(\alpha_i, \alpha_{j'})} \right) x_j \right) + \varepsilon_i.$$

The characteristic α_i of individual i causes endogeneity.

Suppose that the $\alpha_1, \ldots, \alpha_n$ are independently distributed and independent of the regressor.

Then the characteristics of the other individuals (α_j for all $j \neq i$) satisfy the properties of an instrumental variable.

Of course, α_j is not observed.

But $(A)_{i,j}$ and $(A)_{j,i}$ are functions of α_j for all j.

We can use link decisions of other individuals in the network to construct instruments.

Leave-own-out subnetworks

A simple and effective way of creating instruments is by using

$$(A_{-i'})_{i,j} = \begin{cases} (A)_{i,j} & \text{if } i = i' \\ (A)_{i,j} & \text{if } j = i' \\ (A)_{i,j} & \text{if } i \neq i' \text{ and } j \neq i' \end{cases}$$

to construct

$$(Q)_{i,j} = \frac{1}{n-1} \sum_{i'=1}^{n} (H_{-i})_{i',j}$$

and set up instrument

$$z_i = \sum_{j=1}^{n} (Q)_{i,j} x_j$$

for $\tilde{x}_i$.

The entries of the matrix Q give marginal probabilities of walking to j in the leave-i-out network.

Overidentification is possible by constructing higher-order versions of Q, which would consist of probabilities of arriving at j in two, three, or more steps.

Control function

An alternative is to use a control function.

Specify, e.g., that

$$(A)_{i,j} = \alpha_i + \alpha_j \geq v_{ij}, \qquad v_{ij}|\alpha_i, \alpha_j \sim G$$

with v_{ij} independent of everything.

Then

$$\mathbb{E}(\varepsilon_i|A, x) = \mathbb{E}(\varepsilon_i|\alpha_i) = \int \varepsilon \, f(\varepsilon|\alpha_i) \, d\varepsilon =: h(\alpha_i).$$

so that we can write

$$\varepsilon_i = h(\alpha_i) + v_i$$

with $\mathbb{E}(v_i|A, x) = 0$.

Can then proceed by first estimating α_i (as a fixed effect) and then estimate

$$y_i = \alpha + \beta x_i + \gamma \tilde{x}_i + h(\alpha_i) + v_i$$

replacing α_i by $\hat{\alpha}_i$. We can parametrize h or estimate it non-parametrically.

This approach requires a good estimator of α_i and so a large and dense network.

Outline

Bipartite problems

Consider a network with two types of nodes.

Say n_1 workers and n_2 firms (or students and teachers), observed over T periods.

Simple model: At time t worker i works at firm $j = j(i,t)$ and produces a wage

$$y_{it} = \alpha_i + \gamma_{j(i,t)} + \varepsilon_{it}.$$

This gives a bipartite network.

Key parameters of interest here are the variance and covariance of the worker effects and firm effects.

'Do high-quality workers work for high-quality firms?'

Matched data sets are typically very large. However, mobility of workers across firms is what matters.

Bipartite network

The adjacency matrix A of the bipartite network is the $(n_1+n_2)\times(n_1+n_2)$ matrix

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

The diagonal $n_1 \times n_1$ and $n_2 \times n_2$ blocks A_{11} and A_{22} are all zeros.

The upper-right $n_1 \times n_2$ block A_{12} gives the number of wage observations of a worker in a given firm.

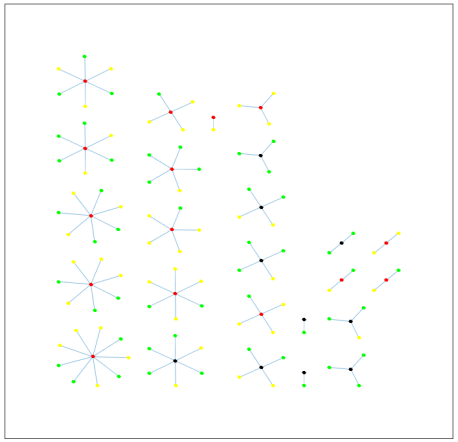
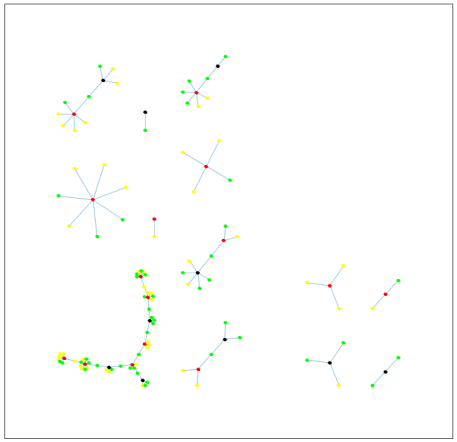
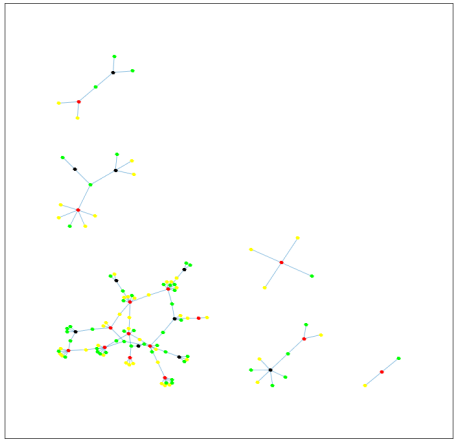
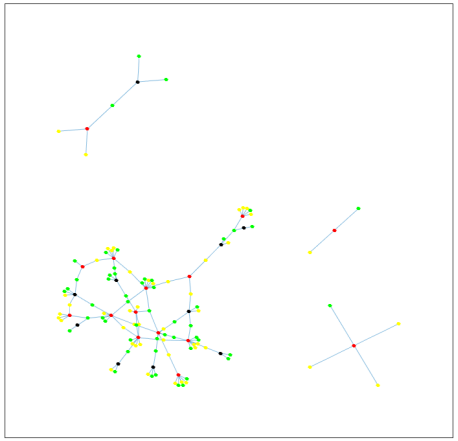
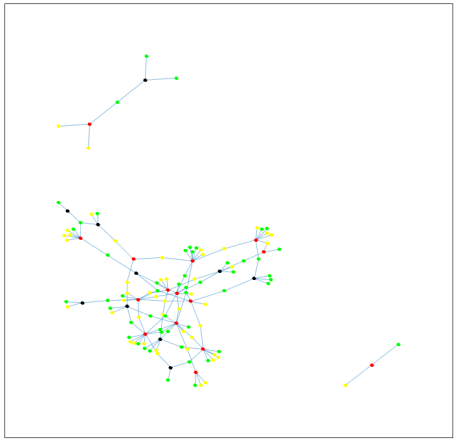
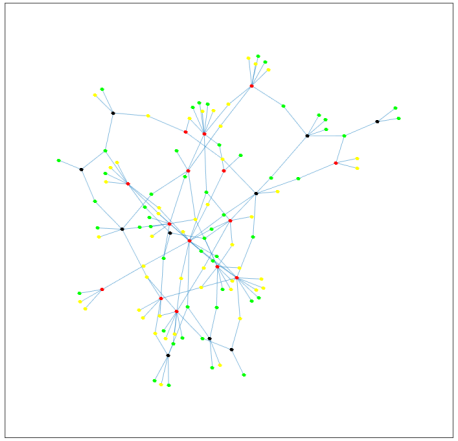
The diagonal matrix $D = \mathrm{diag}(D_1, D_2)$ contains the number of observations involving a given worker/firm:

$$(D_1)_{i,i} = \sum_{j=1}^{n_2} (A_{12})_{i,j}, \quad (D_2)_{j,j} = \sum_{i=1}^{n_1} (A_{21})_{j,i}.$$

The Laplacian matrix of the network is

$$L := D - A.$$

(Dis-)connected graphs

The above plot is for a stationary model where workers and firms are of binary types (high and low quality), and mobility of workers is introduced through exogenous lay-offs.

A one-period network has as many components as firms.

Cannot disentangle worker effects from firm effects.

Adding time periods makes workers switch to different firms.

Makes the graph more connected.

With enough mobility, the graph becomes connected.

Connectivity allows for the least-squares estimator to be well defined (subject to one normalization on the fixed effects) but is not enough for it to have good properties.

Connectivity

The smallest eigenvalue of L is always zero.

The number of zero eigenvalues equals the number of connected components.

The second smallest eigenvalue λ_2 (non-zero iff the graph is connected) is a measure of global connectivity.

Best possible variance rate for a given (worker) effect is d_i^{-1}. But achieving this demands certain conditions on connectivity:

$$\lambda_2 h_i \to \infty,$$

for

$$h_i = \left(\frac{1}{d_i} \sum_{j=1}^{n_2} \frac{(A_{12})_{i,j}^2}{d_j} \right)^{-1},$$

which is a local connectivity measure.

Consequently, bias in plug-in estimator of functionals will not vanish, and certain parts of the network can remain influential in the limit.

Outline

Absence of complementarities in the linear model

Suppose that workers and firms are either high or low type (binary zero-one).

Then expected wage given types, $w_{\alpha,\gamma}$ can take four values.

This leads to the saturated specification

$$w_{\alpha\gamma} = w_{00} + \alpha\left(w_{10} - w_{00}\right) + \gamma\left(w_{01} - w_{00}\right) + \alpha\gamma\left(\left(w_{11} - w_{10}\right) - \left(w_{01} - w_{00}\right)\right)$$

which has an interactive term.

Linearity requires that

$$\left(w_{11} - w_{10}\right) = \left(w_{01} - w_{00}\right).$$

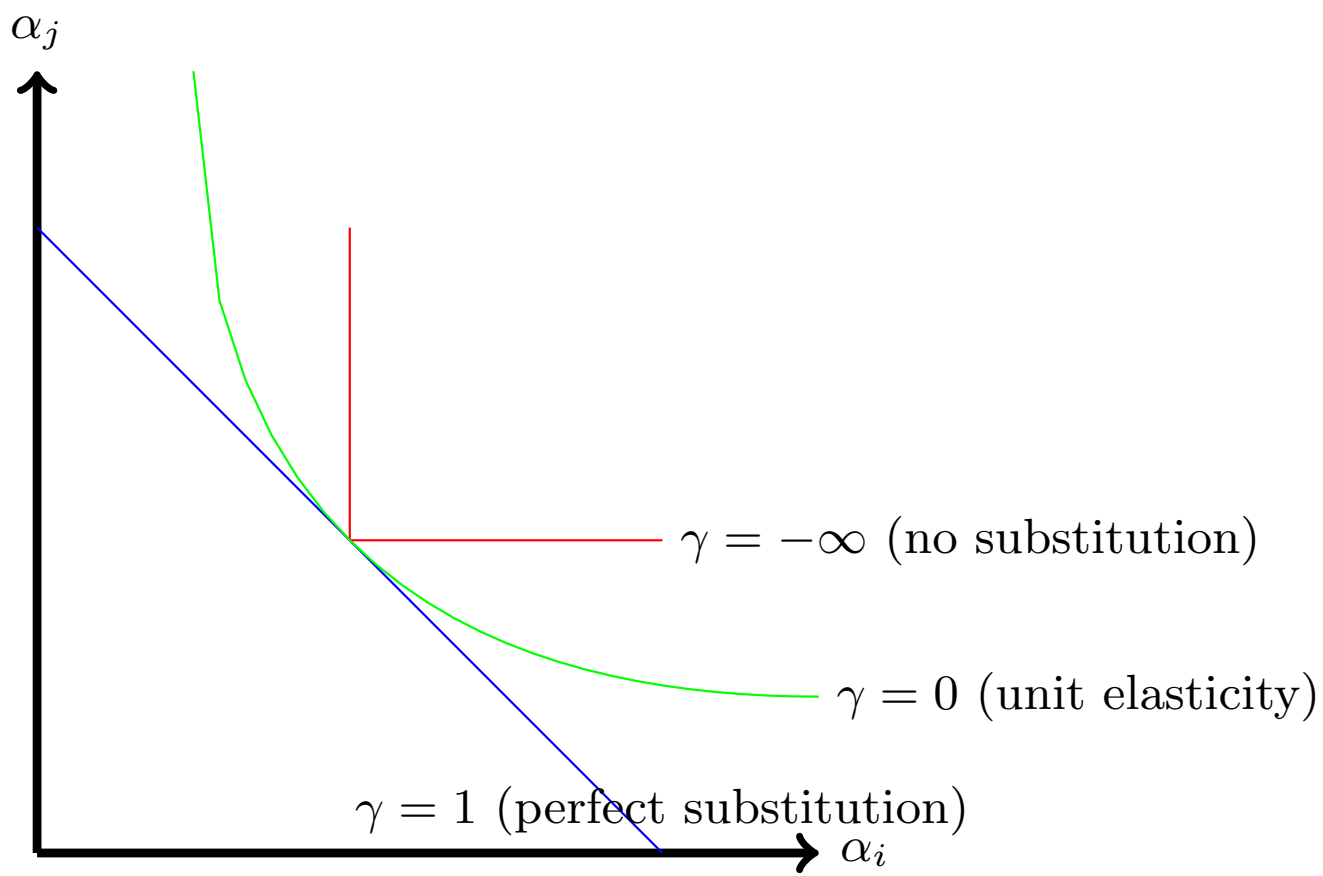
High-quality workers and low-quality workers benefit in the same way from switching from a low-quality to a high-quality firm.

Dyadic production

A standard (stochastic) CES production function would be

$$y_{ij} = \delta \left({}^1\!/_2 \alpha_i^\gamma + {}^1\!/_2 \alpha_j^\gamma \right)^{1/\gamma} \varepsilon_{ij}$$

with ε_{ij} independent of α_i, α_j and mean one. This permits complementarity.



Stochastic block model

Type heterogeneity is again a useful paradigm.

Units are of one of m latent types, according to distribution

$$p_z := \mathbb{P}(\alpha_i = z) > 0$$

for $z = 1, \ldots, m$.

Output has a distribution that depends on the type of both units:

$$F(y|z_i, z_j) := \mathbb{P}(y_{ij} \leq y|z_i, z_j).$$

The marginal distribution of outcomes is

$$\mathbb{P}(y_{ij} \leq y) = \sum_{z_1=1}^{m} \sum_{z_2=1}^{m} F(y|z_1, z_2)\, p_{z_1}\, p_{z_2}$$

which has a mixture form.

Dependence in outcomes is due to the presence of types.

We can re-write this model as

$$y_{ij} = \mu(\alpha_i, \alpha_j; \varepsilon_{ij})$$

where the function μ is symmetric in α_i, α_j but can depend arbitrarily on them, and $u_{ij}|\alpha_i, \alpha_j$ are independent but have a distribution that can depend arbitrarily on α_i, α_j.

We can use this to study complementarity.

See how, for example, $\mathbb{E}(y_{ij}|\alpha_i, \alpha_j)$ changes with (α_i, α_j).

We can apply this to network formation to study sorting.

See how $\mathbb{P}(\alpha_i, \alpha_j|y_{ij})$ depends on y_{ij}.

Note: Types labelling is arbitrary. Can 'fix' an ordering by conditions such as, e.g., $\mathbb{E}(y_{ij}|\alpha_i)$ is increasing in α_i.

Identification

Note that $y_{i1}, \ldots, y_{in}$ are independent conditional on α_i.

This is like a finite mixture in a panel data problem.

Can recover

$$F(y|z) := \mathbb{P}(y_{ij} \leq y | \alpha_i = z)$$

and p_z for each z provided that the $F(\cdot|z)$ functions are linearly independent.

Then look at interaction between four units.

Say you wish to recover

$$\varphi(z, z') := \mathbb{E}(\varphi(y_{ij})|\alpha_i = z, \alpha_j = z')$$

for some chosen function φ.

Let

$$(M)_{p,q} = \mathbb{E}(\{y_{12} \leq y_p\}\, \varphi(y_{23})\, \{y_{34} \leq y_q\})$$

for a collection of points $y_1, \ldots, y_{m'}$ for $m' \geq m$.

Also collect

$$(G)_{p,z} = F(y_p|z), \qquad (P)_{z,z} = p_z.$$

and write

$$(H)_{z,z'} = \varphi(z, z').$$

Then

$$M = G\, P\, H\, P\, G'$$

by conditional independence.

Thus,

$$H = P^{-1}(G'G)^{-1}G'MG(G'G)^{-1}P^{-1}$$