Economics 696, Topics in Econometrics

Lecture Note 13: Social Interactions (updated May 3, 2010)

Based on Manski (1993), Graham and Hahn (2005), and Bramoulle, Djebbari, and Fortin (2009).

1 Introduction

As a concrete example, suppose that we want to determine how a child i's achievement in school depends on his or her own characteristics, the characteristics of the child's classmates, and the *performance* of other classmates. We could imagine various stories why all of these "effects" might be present.

Following Manski (1993), we could start by writing down an equation for performance of child i in classroom g as:

$$y_i = \alpha + E_g[y]\beta + E_g[x]'\gamma + x_i'\delta + u_i,$$

Here, y_i is an outcome (e.g. test score), and x_i is a measured characteristic of the child (e.g. parent's income). The notation E_g means average within classroom g. So $E_g[y]$ is the average of outcome y among all children in the child's class g (including child i), and $E_g[x]$ is the average of the regressor x among the child's class.

To complete the model we need to say something about u_i. Let's assume that $E[u_i|X_g] = 0$, where X_g contains all the x values for children in classroom g (not just i).

This is a "linear-in-means" model for social interactions. The parameter β measures how the average outcome in i's classroom affects i's individual outcome. This effect is called an *endogenous* social effect. The γ measures how the average characteristics of children in i's classroom affects i's outcome. This is called an *exogenous* social effect. There may also be a *correlated* effect, in that u_i and u_j may be correlated for children i and j in the same classroom.

Then the identification question is: from knowledge of the joint distribution of $\{y_i, x_i\}$, can we recover β and γ (and possibly α)? This does not follow immediately from standard OLS arguments, because of the presence of terms like $E_g[y]$ on the right hand side of the main equation of interest.

2 Sampling Framework

The model in §1 is somewhat heuristic. Let's try to put it in a formal probability framework. Intuitively, we want to have many classes (or *groups*), drawn from some joint distribution, and some number of individuals within each group. This puts the social interaction model into the framework of panel data, as pointed out by Graham and Hahn (2005).

Let g denote a given group. Group g has individuals $i = 1, \dots, N_g$. For each individual i in group g, their outcome y_{gi} satisfies:

$$y_{gi} = \left[\frac{1}{N_g} \sum_{j=1}^{N_g} y_{gj} \right] \beta + \left[\frac{1}{N_g} \sum_{j=1}^{N_g} x_{gj} \right]' \gamma + x'_{gi}\delta + \alpha_g + \epsilon_{gi}. \tag{1}$$

We assume that the ϵ_{gi} are mutually uncorrelated and have mean zero conditional on $X_g := (x_{g1}, \dots, x_{gN_g})$ and α_g.[1]

Comparing this to the original equation in §1, we have explicitly spelled out $E_g[y]$ and $E_g[x]$, and we have rewritten $\alpha + u_i = \alpha_g + \epsilon_{gi}$. The term α_g captures both an overall mean α and the part of u_{gi} that is correlated between individuals. In general, we allow α_g to be correlated with X_g.

So for each group g, there is a set of variables (not all observed):

$$(\alpha_g, N_g, X_g, Y_g, E_g), \text{ where } Y_g := (y_{g1}, \dots, y_{gN_g}), E_g := (\epsilon_{g1}, \dots, \epsilon_{gN_g}).$$

In turn, each group g is a draw from the population distribution of groups:

$$(\alpha_g, N_g, X_g, Y_g, E_g) \overset{\text{i.i.d.}}{\sim} F_g,$$

where F_g is such that Equation (1) and the conditions on ϵ_{gi} are satisfied with probability one.

We could relax the i.i.d. assumption on groups somewhat, but it's easy to see now that random sampling over groups and observability of y_{gi}, x_{gi} will identify the joint distribution of (N_g, Y_g, X_g).

3 Model with Endogenous Interactions

Let's first consider a simple version of Equation (1). Suppose that there are no exogenous effects. Then, using $\overline{y}_g$ for the group average of y_{gi}, we have

$$y_{gi} = \alpha_g + \overline{y}_g \beta + x'_{gi}\delta + \epsilon_{gi}. \tag{2}$$

Assume this equation holds for all $i = 1, \dots, N_g$, so we can take the average over i to get:

$$\overline{y}_g = \alpha_g + \overline{y}_g \beta + \overline{x}'_g \delta + \overline{\epsilon}_g,$$

with $\overline{x}_g$ and $\overline{\epsilon}_g$ defined in analogy with $\overline{y}_g$. Let's assume $\beta \neq 1$, so we can solve for $\overline{y}_g$:

$$\overline{y}_g = \frac{\alpha_g}{1-\beta} + \overline{x}'_g \frac{\delta}{(1-\beta)} + \frac{1}{1-\beta} \overline{\epsilon}_g.$$

[1] We could go further and say that the ϵ_{gi} are i.i.d., but we will not need this for the later arguments.

Then substitute the solution for $\overline{y}_g$ back into (2) to get the reduced form equations

$$\begin{aligned} y_{gi} &= \alpha_g + \overline{y}_g\beta + x'_{gi}\delta + \epsilon_{gi} \\ &= \alpha_g + \left[\frac{\alpha_g}{1-\beta} + \overline{x}'_g\frac{\delta}{(1-\beta)} + \frac{1}{1-\beta}\overline{\epsilon}_g\right]\beta + x'_{gi}\delta + \epsilon_{gi} \\ &= \left[\alpha_g + \frac{\alpha_g\beta}{1-\beta}\right] + \overline{x}'_g\frac{\delta\beta}{1-\beta} + x'_{gi}\delta + v_{gi} \\ &= \frac{\alpha_g}{1-\beta} + \overline{x}'_g\frac{\delta\beta}{1-\beta} + x'_{gi}\delta + v_{gi}, \end{aligned}$$

where

$$v_{gi} = \frac{\beta}{1-\beta}\overline{\epsilon}_g + \epsilon_{gi}.$$

Hence (using $\alpha_g^* = \alpha_g/(1-\beta)$),

$$E[y_{gi}|\alpha_g, X_g] = \alpha_g^* + \overline{x}'_g\frac{\delta\beta}{1-\beta} + x'_{gi}\delta.$$

Clearly, if $\delta = 0$ then β is not identified. But even if $\delta \neq 0$, there is an issue because α_g^* is expected to be correlated with the x_{gi}, and hence $\overline{x}_g$. For example, suppose we try to estimate this equation using a fixed effects estimator. The FE estimator sweeps out regressors that do not vary within the group, hence we cannot estimate the compound coefficient $\delta\beta/(1-\beta)$ using the FE estimator. A possible solution is to try to find and instrumental variable for $\overline{x}_g$. See Graham and Hahn (2005).

4 Interactions in Other Network Structures

The setup above imposes a particular structure to the interactions between individuals: individuals within a group interact with all of each other (through the group means), and not at all with other individuals outside the group. Bramoulle, Djebbari, and Fortin consider more general types of social networks and show that in some cases, one can identify interaction effects without additional instrumental variables.

As an example, suppose that each group g consists of three individuals ($g1, g2, g3$), but that the influence of individuals within a group runs in a particular direction:

$$g1 \to g2 \to g3 \cdots$$

Thus each group is an *intransitive directed network*.

Then a model (following our previous linear specification) could have, for $i > 1$,

$$y_{gi} = \alpha_g + \beta y_{g,i-1} + \delta x_{gi} + \gamma x_{g,i-1} + \epsilon_{gi},$$

with $E[\epsilon_{gi}|\alpha_g, X_g] = 0$ for all i. Then

$$\Delta y_{gi} = \beta\Delta y_{g,i-1} + \delta\Delta x_{gi} + \gamma\Delta x_{g,i-1} + \Delta\epsilon_{gi},$$

and

$$E[\Delta\epsilon_{gi}|X_g] = 0.$$

Hence the $\Delta x_{gi}, \Delta x_{g,i-1}, \Delta x_{g,i-2}, \ldots$ are valid instruments for the model.

More generally, we can think of each individual i having a group of other individuals P_i who influence him or her. Then the linear in means model can be written

$$y_i = \alpha + \beta \left[\frac{1}{|P_i|} \sum_{j \in P_i} y_j \right] + \delta x_i + \gamma \left[\frac{1}{|P_i|} \sum_{j \in P_i} x_j \right] + u_i,$$

where $E[u_i|X] = 0$ and the u_i are allowed to be correlated across individuals. Here X refers to all x_j.

For $i = 1, \ldots, N$ and $j = 1, \ldots, N$, let $G_{ij} = 1/P_i$ if $j \in P_i$, and 0 otherwise. Then let G_N be the $N \times N$ matrix with elements G_{ij}. Then we can write the model in stacked form as

$$Y_N = \alpha\iota + \beta G_N Y_N + \delta X_N + \gamma G_N X_N + U_N,$$

where

$$Y_N = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad X_N = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}, \quad U_N = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix},$$

and ι is the $N \times 1$ vector of ones. Note that $E[U_N|X_N] = 0$.

At this level of generality, it's a bit tricky to embed this model into a rigorous probability structure that will hold as $N \to \infty$ and permit use of laws of large numbers. But we'll ignore this for now and plunge ahead.

Assume that $|\beta| < 1$. Then $I_N - \beta G_N$ is invertible and we can write

$$Y_N = \alpha(I_N - \beta G_N)^{-1}\iota + (I_N - \beta G_N)^{-1}(\delta I_N + \gamma G_N)X_N + (I_N - \beta G_N)^{-1}U_N.$$

Note that

$$(I_N - \beta G_N)^{-1} = \sum_{k=0}^{\infty} \beta^k G_N^k.$$

Also, if no individual i is isolated, then the intercept for i is simply $\alpha/(1-\beta)$ as before. Then (assuming no isolated individuals) we have

$$Y_N = \alpha/(1-\beta)\iota + \delta X_N + (\delta\beta + \gamma) \sum_{k=0}^{\infty} \beta^k G_N^{k+1} X_N + \sum_{k=0}^{\infty} \beta^k G_N^k U_N.$$

Clearly if $\delta\beta + \gamma = 0$ then we cannot hope to estimate β, since there is no overall interaction effect. If $\delta\beta + \gamma \neq 0$, then identification will depend on the structure of G.

Multiplying the previous expression by G and taking conditional expectation we have

$$E[GY_N|X_N] = \alpha/(1-\beta)\iota + \delta G_N X_N + (\delta\beta + \gamma) \sum_{k=0}^{\infty} \beta^k G^{k+2} X_N.$$

Now return to the equation

$$Y_N = \alpha\iota + \beta G_N Y_N + \delta X_N + \gamma G_N X_N + U_N.$$

The idea is to use terms like $G_N^k X_N$ for $k \geq 2$ as instruments for $G_N Y_N$ in the equation of interest. This corresponds to using higher lags of x as instruments. However, this will not work, for example, if $G_N^2 = G_N^1$, because then $G_N^2 X_N = G_N X_N$ and the potential instrument $G_N^2 X_N$ is collinear with the regressor $G_N X_N$. Bramoulle et al show that if I, G_N, and G_N^2 are linearly independent, then this identification strategy works.

5 Nonparametric Models

See (inter alia) Graham, Imbens, and Ridder, 2009, "Complementarity and Aggregrate Implications of Assortative Matching: A Nonparametric Analysis," Manski, 2010, "Identification of Treatment Response with Social Interactions," and Lazzati, 2010, "Treatment Response with Social Interactions: A Nonparametric Approach to Partial Identification."