Moshe Buchinsky
Department of Economics
UCLA

Economics 203C
Spring 2003

LECTURE NOTE 3
SIMULTANEOUS-EQUATION MODEL (SEM)

I. THE MODEL

We have a system of m equations with some (or all) of the endogenous variables being explanatory variables as well.

ENDOGENOUS VARIABLES:

$$y_i \equiv (y_{1i}, \ldots, y_{mi})', \quad m \times 1 \text{ vector.}$$

EXOGENOUS VARIABLES:

$$z_i \equiv (z_{1i}, \ldots, z_{ki})', \quad k \times 1 \text{ vector.}$$

ERROR TERM:

$$\epsilon_i \equiv (\epsilon_{1i}, \ldots, \epsilon_{mi})', \quad m \times 1 \text{ vector.}$$

SIMULTANEOUS-EQUATION MODEL (SEM):

$$y_i'\Gamma = z_i'B + \epsilon_i',$$

where Γ is an $m \times m$ matrix of coefficients and B is a $k \times m$ matrix of coefficients.

Let

$$Y = \begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix}, \quad Z = \begin{pmatrix} z_1' \\ \vdots \\ z_n' \end{pmatrix}, \quad E = \begin{pmatrix} \epsilon_1' \\ \vdots \\ \epsilon_n' \end{pmatrix},$$

where Y is an $n \times m$ matrix, Z is an $n \times k$ matrix, and E is an $n \times m$ matrix. Hence, we can write the model in matrix form as

$$Y\Gamma = ZB + E$$

or

$$\begin{aligned} Y &= Y(I - \Gamma + ZB + E \\ &= \begin{pmatrix} Y & Z \end{pmatrix} \begin{pmatrix} I - \Gamma \\ B \end{pmatrix} + E \\ &= \begin{pmatrix} Y & Z \end{pmatrix} A + E, \end{aligned}$$

where

$$A = \begin{pmatrix} I - \Gamma \\ B \end{pmatrix}$$

an $(m + k) \times m$ matrix of coefficients. This is the *structural model.*

REMARK:

(1) We need to present the model in *reduced form*, because as it is now we cannot estimate it.

(2) We will need to restrict some of the elements of A, in order to be able to identify the rest of the coefficients. As it is right now, we have altogether $m(k+m)$ coefficients in Γ and B. Restriction can be based on prior knowledge about symmetry, zero coefficients, etc.

II. INDIVIDUAL EQUATION

Write Y and E in the following way

$$Y = (Y_1, \ldots, Y_j, \ldots, Y_m),$$

$$E = (E_1, \ldots, E_j, \ldots, E_m),$$

where Y_j and E_j $(j = 1, \ldots, m)$ are $n \times 1$ vectors. That is, each column represents one equation.

Without loss of generality, assume that $\{\Gamma\}_{jj} = \gamma_{jj} = 1$. Then the j^{th} column of SEM is of the form:

$$Y_j = X_j\beta_j + \epsilon_j,$$

or

$$y_{ji} = x'_{ji}\beta_j + \epsilon_{ji} \qquad (i = 1, \ldots, n).$$

The vector β_j is the vector containing the non-zero components of A, while X_j contains the corresponding components of the columns of (Y, Z).

REMARKS:

1. Note that the j^{th} equation is written in a form so that all the other endogenous variables are on the right hand side of the equation.

2. Since X_j contains some columns of Y, it follows that in general $E[x_{ji}\epsilon_{ji}] \neq 0$, so that the NeoCR model does not hold.

III. REDUCED FORM OF THE MODEL

Solves for y_i given z_i and ϵ_i:

$$y'_i\Gamma = z'_iB + \epsilon'_i.$$

So,

$$\begin{aligned} y'_i &= z'_iB\Gamma^{-1} + \epsilon'_i\Gamma^{-1} \\ &= z'_i\Pi + v'_i, \end{aligned}$$

where $\Pi \equiv B\Gamma^{-1}$ a $k \times m$ matrix of reduced form coefficients, and $v'_i = \epsilon'_i\Gamma^{-1}$. We assume that $\det(\Gamma) \neq 0$, so that Γ^{-1} exists.

In matrix form we have for the SEM:

$$\begin{aligned} Y &= ZB\Gamma^{-1} + E\Gamma^{-1} \\ &= Z\Pi + V, \end{aligned}$$

where $V \equiv E\Gamma^{-1}$.

REMARKS:

1. Note that $E[v_i' \mid z_i] = E[\epsilon_i'\Gamma^{-1} \mid z_i] = 0$, by assumption. Hence, each column of the reduced form follows the NeoCR model.

2. If the data are drawn from a random sample, then the reduced form is simply a SUR model with identical X's. Therefore, the best unrestricted estimator of Π is a LS equation-by-equation (which is identical to the GLS estimator). That is, if we write $\Pi = (\pi_1, \ldots, \pi_m)$, then

$$\hat{\pi}_j = (Z'Z)^{-1} Z'Y_j.$$

IV. RELATION BETWEEN STRUCTURAL AND REDUCED FORM COEFFICIENTS

The relation between the structural parameters of Γ and B and the reduced form parameters Π are quite important, since it governs the uniqueness of Γ and B.

$$\Pi = B\Gamma^{-1} \quad \Longleftrightarrow \quad \Pi\Gamma = B \quad \Longleftrightarrow \quad B - \Pi\Gamma = 0.$$

That is, B and Γ solve the linear equation system:

$$\begin{pmatrix} I & -\Pi \end{pmatrix} \begin{pmatrix} B \\ \Gamma \end{pmatrix} = 0.$$

V. IDENTIFICATION

Given the restrictions on Γ and B, when does $B - \Pi\Gamma = 0$ have unique solution for B and Γ given Π?

FACTS TO BE CONSIDERED:

1. Π can be consistently estimated by LS. Hence, we can treat Π as "known in large samples," i.e., identified.

2. There are

   (i) km elements in Π;

   (ii) m^2 elements in Γ; and

   (iii) km elements in B.

   Hence, if Γ and B are unrestricted, $B - \Pi\Gamma = 0$ gives km equations with $m(k+m)$ unknowns. Therefore, no unique solution can be obtained. We must have at least m^2 restrictions (m restriction for each structural equation) to be able to obtain a unique solution.

TYPICAL RESTRICTIONS:

1. Non-zero restriction: $\gamma_{jj} = 1$ $(j = 1, \ldots, m)$ gives m restrictions.

2. Zero restrictions $\{B\}_{jk} = 0$ and/or $\{\gamma\}_{lp} = 0$. That is, other components of A beside the diagonal elements of Γ equal zero.

VI. ORDER CONDITION FOR IDENTIFICATION (necessary conditions):

In order to identify β_j, in the j^{th} equation, where

$$Y_j = X_j\beta_j + \epsilon_j,$$

we need *at least* m zeros in the j^{th} column of A for β_j to be identified.

INSTRUMENTAL VARIABLE INTERPRETATION TO THE ORDER CONDITION

Recall that in the structural model $y_{ji} = x'_{ji}\beta_j + \epsilon_{ji}$.

Let

$$z_i = (z_{1i}, \ldots, z_{ki})$$

be the instrumental variables for x_{ji}.

Let R_j be the number of zero restrictions in the j^{th} column of A.

Then,

$$\dim(\beta_j) = \dim(x_{ji}) = k + m - R_j,$$

and

$$\dim(z_i) = k.$$

Thus, $R_j \geq m$ implies that

$$\dim(x_{ji}) = k + m - R_j \leq k = \dim(z_i).$$

That is, the number of IV's in z_i exceeds the number of regressors in x_{ji}. So, can solve the problem for the j^{th} equation.

Intuitively, we need to have at least as many excluded exogenous variables as included endogenous variables in the j^{th} equation for it to be identified.

VII. RANK CONDITION FOR IDENTIFICATION (necessary and sufficient condition):

Let A_j be the sub-matrix of rows of A with zeros in the j^{th} column. Then β_j is identified *if and only if*

$$\text{rank}(A_j) = m - 1.$$

INTUITION: What does the requirement $\text{rank}(A_j) < m-1$ mean (when $R_j = m$)? It means that $m-1$ columns (or the rows) of A_j, without the column of zeros, are linearly dependent. So, we can normalize A so that one additional row in A_j will be zero, and the variable corresponds to that row is not in the SEM.

REMARKS:

1. A_j is an $R_j \times m$ matrix with the j^{th} column equals to zero by construction. Hence,

   $$\text{rank}(A_j) \leq m - 1$$

   Note that $\text{rank}(A_j) < m - 1$ only if $R_j < m - 1$. But for the order condition to hold we need that $R_j \geq m$, because the order condition is a necessary condition!.

2. The rank condition is rarely checked, since it depends on unknown parameters, except in simple cases.

   EXAMPLE: Identification fails if $R_j = m$ and the l^{th} column of A_j $(l \neq j)$ equals zero, because in that case $\text{rank}(A_j) = m - 2$.

   That is, this particular variable does not enter into SEM.

3. If the rank condition holds and

   (a) if $R_j = m$, then the j^{th} equation is said to be *just identified*.

   (b) if $R_j > m$, then the j^{th} equation is said to be *over identified*.

   (c) if $R_j < m$, then the j^{th} equation is said to be *not identified*, or *under identified*.