Simultaneous Equations Models

Semester 2/2013

Part 2

Chayanee Chawanote

Identifying and estimating a structural equation

- To solve the problem of endogeneity, we apply 2SLS to SEMs.
- We first need to know how to identify these equations.
- How we observe demand/suppy shifts?
  $q = \alpha_1 p + \beta_1 z_1 + u_1$ (1)
  $q = \alpha_2 p + u_2$ (2)
- Which is an identified equation?

Identification in a two-equation system

- A general two-equation model:

  $y_1 = \beta_{10} + \alpha_1 y_2 + \mathbf{z}_1\beta_1 + u_1$ (3)

  $y_2 = \beta_{20} + \alpha_2 y_1 + \mathbf{z}_2\beta_2 + u_2$ (4)

  - $z_1 = (z_{11}, z_{12}, ..., z_{1k_1})$; $z_2 = (z_{21}, z_{22}, ..., z_{2k_2})$: sets of different exogenous variables (if we impose exclusion restriction).
  - Solve (3) and (4) for $y_1$and y_2 as linear functions of all exogenous variables and the structural errors to get the reduced forms.
  - Under what conditions can we estimate the parameters in (3), or (4)?

Rank condition for identification of a structural equation

- The 1st equation in a 2-equation SEM is identified if, and only if, the 2nd equation contains at least one exogenous variable (with nonzero coefficient) that is excluded from the the 1st equation.
- For the 1st equation to be identified, we need
  - Order condition: at least one exogenous variable is excluded from the 1st equation
  - Rank condition: at least one of the exogenous variables excluded from the first equation must have a nonzero population coeffecient in the 2nd equation

Rank condition for identification of a structural equation

- Example: labor supply and demand for married women
  $hours = \beta_{10} + \alpha_1 log(wage) + \beta_{11} educ + \beta_{12} age + \beta_{13} kids6 + \beta_{14} nwifeinc + u_1$ (5)
  $log(wage) = \beta_{20} + \alpha_2 hours + \beta_{21} educ + \beta_{22} exper + \beta_{23} exper^2 + u_2$ (6)
  - kids6 - #children less than 6 years, nwifeinc - woman's nonwage income, including husband's earnings
  - Others than *log(wage)* and *hours* are exogenous variables
- Which equation is the labor supply for married women?
- If we want to estimate the labor supply,
  - order condition:
  - rank condition:

Estimation by 2SLS

- Once we determined that an equation is identified, we can estimate it by 2SLS with IVs consist of the exogenous variables appearing in either equation.

Systems with more than two equations

- Suppose we have three equations, intercept suppressed for simplicity:
  $y_1 = \alpha_{12} y_2 + \alpha_{13} y_3 + \beta_{11} z_1 + u_1$ (7)
  $y_2 = \alpha_{21} y_1 + \beta_{21} z_1 + \beta_{22} z_2 + \beta_{23} z_3 + u_2$ (8)
  $y_3 = \alpha_{32} y_2 + \beta_{31} z_1 + \beta_{32} z_2 + \beta_{33} z_3 + \beta_{34} z_4 + u_3$ (9)
- Which of these equations can be estimated?
- Order condition for identification: it satisfies the order condition if the number of excluded exogenous variables from the equation is at least as large as the number of RHS endogenous variables
- Rank condition: need matrix algebra

General linear restrictions and structural equations

- Suppose we have G equations:
  $\mathbf{y}\gamma_{\mathbf{1}} + \mathbf{z}\beta_{\mathbf{1}} + \mathbf{u_1} = 0$
  $\vdots$
  $\mathbf{y}\gamma_{\mathbf{G}} + \mathbf{z}\beta_{\mathbf{G}} + \mathbf{u_G} = 0$
  - $\mathbf{y} \equiv (y_1, y_2, ..., y_G)$ is $1 \times G$ vector of all endogenous variables
  - $\mathbf{z} \equiv (z_1, z_2, ..., z_M)$ is $1 \times M$ vector of all exogenous variables, containing unity(intercept)
- That is, we have $\mathbf{y}` + \mathbf{z}\beta + \mathbf{u} = 0$ (10)
- Hence, the reduced form (11) will be:

General linear restrictions and structural equations

restrictions and rank condition

- Suppose we consider identification of the first equation: $\mathbf{y}\gamma_\mathbf{1} + \mathbf{z}\beta_\mathbf{1} + \mathbf{u_1} = 0$ (12)
- The normalization restriction in equation (12): one element of γ_1 is -1
  - one variable is taken to be the LHS explained variable
- Let $B_1 \equiv (\gamma_1', \delta_1')$ be the (G+M) $\times$ 1 vector of structural parameters in the first equation.
- With a normalization restriction, there are (G+M) - 1 unknown elements in B_1
- Assume that prior knowledge about B_1 can be expressed as $R_1 B_1 = 0$ (13)
  - R_1 is a $J_1 \times (G + M)$ matrix of known constants, J_1 is the number of restrictions on B_1
  - We assume that rank$(R_1) = J_1$

General linear restrictions and structural equations

restrictions and rank condition

- Example (how to define R_1): consider the first equation in a system with G = 3 and M = 4:
  $y_1 = \gamma_{12} y_2 + \gamma_{13} y_3 + \beta_{11} z_1 + \beta_{12} z_2 + \beta_{13} z_3 + \beta_{14} z_4 + u_1$ (14)
- We have $\gamma_{\mathbf{1}} = (-1, \gamma_{12}, \gamma_{13})', \beta_{\mathbf{1}} = (\beta_{11}, \beta_{12}, \beta_{13}, \beta_{14})'$, and $B_1 = (-1, \gamma_{12}, \gamma_{13}, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{14})'$
- Suppose the restrictions on the structural parameters are $\gamma_{12} = 0$ and $\beta_{13} + \beta_{14} = 3$. Then, $J_1 = 2$ and
  $$R_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

General linear restrictions and structural equations

restrictions and rank condition

- *Rank condition for identification*: Let B_1 be the (G+M) $\times$ 1 vector of structural parameters in the first equation, with the normalization restriction that one of the coefficients on an endogenous variable is -1. Let the additional information on B_1 be given by restriction $R_1B_1 = 0$. Then, B_1 is identified if and only if the rank condition, rank $(R_1B) =$ G -1, holds.
- $R_1B = [R_1B_1, R_1B_2, ..., R_1B_G]$, where $B_g$is the (G+M) $\times$ 1 vector of structural parameters in equation g.
- Since the first column of R_1B is the zero vector (by $R_1B_1 = 0$), R_1B cannot have rank larger than G-1.
- *Order condition for identification*: under assumption $R_1B_1 = 0$, a necessary condition for the first equation to be identified is $J_1 \geq G - 1$

Check identification

Unidentified, Just identified, and overidentified equations

- Steps for checking whether the first equation in the system is identified.

1. Set one element of γ_1 to -1 as a normalization
2. Define the $J_1 \times (G + M)$ matrix R_1 such that eq(13) captures all restrictions on B_1
3. If $J_1 < G - 1$, the first equation is not identified
4. If $J_1 \geq G - 1$, the equation might be identified. Compute $R_1 B$ and check the rank condition.
    4.1 If rank condition fails, we say that the equation is unidentified.
    4.2 If $J_1 = G - 1$, the equation is just identified.
    4.3 If $J_1 > G - 1$, dropping one or more restrictions on parameters still achieve identification, the equation is overidentified.