3. SIMULTANEOUS EQUATIONS MODELS (SEM)

Lecture Plan:

(1) Introduction
(2) Identification
(3) Single Equation Estimators (2SLS, LIML, etc.)
(4) Systems Estimators (3SLS, FIML, etc.)
(5) Comparisons of Estimators
(6) Specification Tests

[1] Introduction

- Example 1: Simple Keynesian Macro Model

  1) $C_t = \alpha + \beta INC_t + \varepsilon_{c,t.}$

  2) $INC_t = C_t + INV_t.$

  - Exogenous: INV; Endogenous: C and INC.
  - OLS estimators of equation 1) are not consistent:
    - If we solve the above equations for C_t and INC_t,

$$C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} INV_t + \frac{1}{1-\beta}\varepsilon_{c,t} \equiv \pi_{11} + \pi_{21} INV_t + v_t;$$

$$INC_t = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} INV_t + \frac{1}{1-\beta}\varepsilon_{c,t} \equiv \pi_{12} + \pi_{22} INV_t + v_t. \qquad \text{(RF)}$$

    - INC_t is correlated with ε_t.

- Comments:
  - The solution equations (RF) are called "reduced form" equations.
  - The parameters in the reduced form equations ($\pi_{11},\pi_{21},\pi_{12},\pi_{22}$) are called "reduced form" parameters.
  - Reduced form equations are essentially SUR with the same regressor s for different equations.
  - The parameters in the structural equations 1) and 2) are called "structural parameters.
  - Reduced form parameters are functions of structural parameters.
  - Reduced form equations indicate that the endogenous variables are correlated with the exogenous regressors. Thus, each equation can be estimated by 2SLS using all of the exogenous variables in the system.

- Funny notation:

  $C_t\bullet(-1) + INCt\bullet\beta + 1\bullet\alpha + INVt\bullet 0 + \varepsilon_{c,t} = 0;$

  $Ct\bullet 1 + INC_t\bullet(-1) + 1\bullet 0 + INVt\bullet 1 + 0 \quad = 0.$

$$\begin{pmatrix} C_t & INC_t \end{pmatrix}\begin{pmatrix} -1 & 1 \\ \beta & -1 \end{pmatrix} + \begin{pmatrix} 1 & INV_t \end{pmatrix}\begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \varepsilon_{c,t} & 0 \end{pmatrix} = (0,0).$$

$\rightarrow$ Let $Y_{t\bullet}' = (C_t, INC_t)$; $X_{t\bullet}' = (1, INV_t)$; $\varepsilon_{t\bullet}' = (\varepsilon_{c,t}, 0)$; and

$$\Gamma = \begin{pmatrix} -1 & 1 \\ \beta & -1 \end{pmatrix}; B = \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix}.$$

$\rightarrow Y_{t\bullet}'\Gamma + X_{t\bullet}'B + \varepsilon_{t\bullet}' = 0_{1\times 2}$ for each t.

$\rightarrow Y\Gamma + XB + E = 0_{T\times 2}$ (for all t), where Y endogenous and X exogenous.

- Example 2: Demand and Supply

1) $Q_t = a + bP_t + \varepsilon_{t1}$ (Demand)

2) $Q_t = \alpha + \beta P_t + \gamma W_t + \varepsilon_{t2}$ (Supply),

where W_t = dummy variable for weather.

$$\rightarrow (Q_t \quad P_t)\begin{pmatrix} -1 & -1 \\ b & \beta \end{pmatrix} + (1 \quad W_t)\begin{pmatrix} a & \alpha \\ 0 & \gamma \end{pmatrix} + (\varepsilon_{t1} \quad \varepsilon_{t2}) = (0,0).$$

- Example 3: Dynamic Keynesian Model

1) $C_t = \alpha + \beta INC_t + \gamma C_{t-1} + v_t$.

2) $INC_t = C_t + INV_t$.

$$(C_t \quad INC_t)\begin{pmatrix} -1 & 1 \\ \beta & -1 \end{pmatrix} + (1 \quad INV_t \quad C_{t-1})\begin{pmatrix} \alpha & 0 \\ 0 & 1 \\ \gamma & 0 \end{pmatrix} + (v_t \quad 0) = (0 \quad 0)$$

$\rightarrow Y_{t\cdot}'\Gamma + X_{t\cdot}'B + \varepsilon_{t\cdot}' = 0_{1\times 2}$, where $Y_{t\cdot}$ endogenous and $X_{t\cdot}$ predetermined (including exogenous).

- General Notation
  - M = # of equations (and # of endogenous variables).
  - K = # of exogenous or predetermined variables.
  - $Y_{t\bullet}' = (y_{t1}, \ldots , y_{tM})$.
  - $X_{t\bullet}'$ = 1×K vector of predetermined variables.
  - $\varepsilon_{t\bullet}' = (\varepsilon_{t1}, \ldots , \varepsilon_{tM})$.
  - Γ: M×M.
  - B: K×M.
  - Structural equations:

$$Y_{t\bullet}'\Gamma + X_{t\bullet}'B + \varepsilon_{t\bullet}' = 0_{1\times M}$$

$\rightarrow Y\Gamma + XB + E = 0_{T\times M}$, where Y, T×M; X, T×K; and E, T×M.

- Assumptions:

  1) $\varepsilon_{t\bullet}$ ~ iid $N(0_{M\times 1},\Sigma_{M\times M})$. [In fact, the $\varepsilon_{t\bullet}$ need not be normal.]

  2) Γ: nonsingular (for unique solution for endogenous variables).

  3) X is of full column and nonstochastic.

     - rank(X) = K (no multicollinearity).
     - "nonstochastic" implies $E(X_{t\bullet}\varepsilon_{tj}) = X_{t\bullet}E(\varepsilon_{tj}) = 0_{K\times 1}$ for all t and j.
     - Instead of the nonstochastic assumption, we can alternatively assume:

       $E(\varepsilon_{jt}|X_{1\bullet}, \ldots , X_{t\bullet}) = 0$ for all t and j.

       - When this assumption is satisfied, we say that the $X_{t\bullet}$ are weakly exogenous.
       - This alternative assumption allows "predetermined regressors.

  4) If there are lagged dependent variables, the model is dynamically stable (stationarity assumption).

  5) Some other regularity conditions (Schmidt, p. 120).

[2] Identification

- Important question: Can we estimate structural parameters?

- Continue Example 2: Demand and Supply

  1) $Q_t = a + bP_t + \varepsilon_{t1}$ (Demand)

  2) $Q_t = \alpha + \beta P_t + \gamma W_t + \varepsilon_{t2}$ (Supply), where W_t = weather.

$$\begin{pmatrix} Q_t & P_t \end{pmatrix} \begin{pmatrix} -1 & -1 \\ b & \beta \end{pmatrix} + \begin{pmatrix} 1 & W_t \end{pmatrix} \begin{pmatrix} a & \alpha \\ 0 & \gamma \end{pmatrix} + \begin{pmatrix} \varepsilon_{t1} & \varepsilon_{t2} \end{pmatrix} = (0,0).$$

$$\rightarrow \quad \Pi \equiv \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = -\mathrm{B}\Gamma^{-1} = \frac{1}{b-\beta} \begin{pmatrix} \alpha b - a\beta & \alpha - a \\ b\gamma & \gamma \end{pmatrix}$$

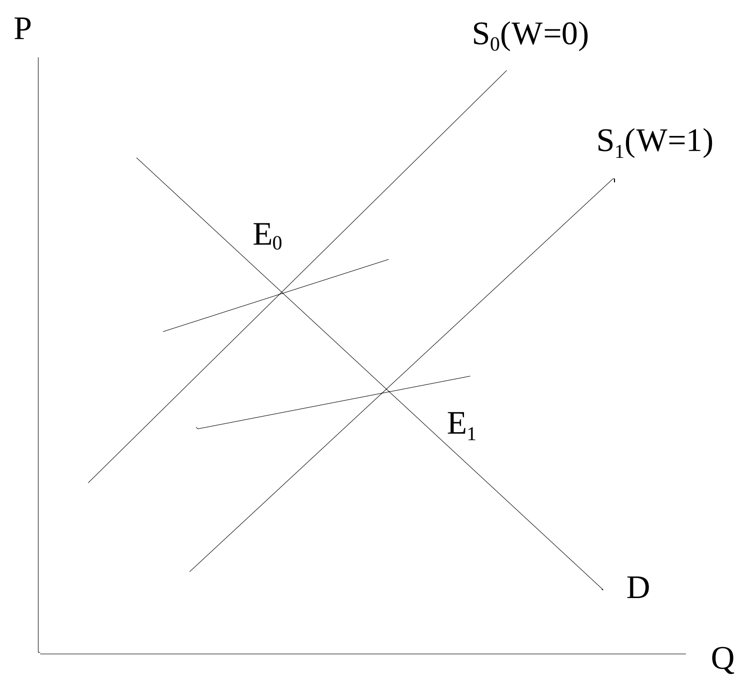
$$\rightarrow \quad \begin{aligned} Q_t &= \pi_{11} + \pi_{21} W_t + u_{t1} = \frac{\alpha b - a\beta}{b-\beta} + \frac{b\gamma}{b-\beta} W_t + \frac{b\varepsilon_{t2} - \beta\varepsilon_{t1}}{b-\beta}; \\ P_t &= \pi_{12} + \pi_{22} W_t + u_{t2} = \frac{\alpha - a}{b-\beta} + \frac{\gamma}{b-\beta} W_t + \frac{\varepsilon_{t2} - \varepsilon_{t1}}{b-\beta}. \end{aligned}$$

→ We can estimate π_{11}, π_{21}, π_{12} and π_{22}. From them, can recover the structural parameters α, b, α, β and γ?

→ No. four equations and 5 knowns.

→ What equation can be identified?

- Intuition for Example 2:



- Question: Can we recover structural parameters from reduced-form parameters?
  - Note that we have MK reduced form parameters:

    $y_{t1} = \pi_{11}x_{t1} + \pi_{12}x_{t2} + ... + \pi_{1K}x_{tK} + v_{t1}$;

    :

    $y_{tM} = \pi_{M1}x_{t1} + \pi_{M2}x_{t2} + ... + \pi_{MK}x_{tK} + v_{tM}$.

  - But, Γ contains M(M-1) [the coefficient of the dependent variable in each equation = -1], and B contains KM.

    → In total, M(M-1) + KM.

  - So, there are M(M-1) more structural parameters.
  - Can't recover structural parameters without any restrictions on structural parameters.

- Question: What kinds of restrictions do we need?

1) Exclusive restrictions on coefficients:

   - Continue Example 2: Demand and Supply

     1) $Q_t = a + bP_t + \varepsilon_{t1}$ (Demand)

     2) $Q_t = \alpha + \beta P_t + \gamma W_t + \varepsilon_{t2}$ (Supply), where W_t = weather.

     $\rightarrow$ The coefficient of W_t in 1) = 0.

   - Consider the following model:

     $- y_{1t} + \beta_{11}x_{1t} + \beta_{21}x_{2t} + \varepsilon_{1t} = 0$

     $- y_{2t} + \gamma_{12}y_{1t} + \beta_{32}x_{3t} + \varepsilon_{2t} = 0$

     $\rightarrow$ $y_{1t}(-1) + y_{2t}(0) + x_{1t}\beta_{11} + x_{2t}\beta_{21} + x_{3t}(0) + \varepsilon_{1t}$.

     $y_{1t}(\gamma_{21}) + y_{2t}(-1) + x_{1t}(0) + x_{2t}(0) + x_{3t}\beta_{32} + \varepsilon_{2t}$.

     $\rightarrow$ y_{2t} and x_{3t} are excluded from equation 1): $\gamma_{21} = \beta_{31} = 0$.

     x_{1t} and x_{2t} are excluded from equation 2): $\beta_{12} = \beta_{22} = 0$.

2) linear restrictions on coefficients.

   - In the demand-supply example, $b + \beta = 0$ (a silly example).

3) Covariance restrictions:

   - Σ is diagonal.

- Question: How could we check whether an equation is identified by a certain set of restrictions on the system?

A. Method 1:

- Based on the forms of the reduced form parameters AND given restrictions.
- See Greene or Schmidt.

B. Method 2:

(1) Cases with restrictions on the coefficients only

- Cases where models can be consistently estimated by 2SLS or 3SLS.
- $Y\Gamma + XB + E = 0$, where

$$\Gamma = \begin{pmatrix} -1 & \gamma_{12} & \gamma_{13} & \cdots & \gamma_{1M} \\ \gamma_{21} & -1 & \gamma_{23} & \cdots & \gamma_{2M} \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma_{M1} & \gamma_{M2} & \gamma_{M3} & \cdots & -1 \end{pmatrix}; \mathrm{B} = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1M} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2M} \\ \vdots & \vdots & & \vdots \\ \beta_{K1} & \beta_{K2} & \cdots & \beta_{KM} \end{pmatrix}.$$

Here, we assume that the diagonals of Γ = -1. In fact, it is sufficient to assume that one of entries in each column = -1.

$$\begin{pmatrix} Y & X \end{pmatrix} \begin{pmatrix} \Gamma \\ \mathrm{B} \end{pmatrix} + E = Z\Delta + E = 0,$$

where $\Delta = [\Delta_1, \Delta_2, .., \Delta_M] = [\Delta_1, \Delta_{(1)}]$.

- First equation: $Z\Delta_1 + \varepsilon_1 = 0$.

- Let Φ_1 be a $R_1 \times (M+K)$ known matrix and suppose that a set of restrictions on the first equation is given by $\Phi_1\Delta_1 = 0$.

- Example: Assume 2 endo vars and 3 exo. vars.

  (1) $\gamma_{21} = 0$

  $\rightarrow$ The 2nd endo. var. does not appear in the 1st equation.

  $\rightarrow \Phi_1 = (0, 1, 0, 0, 0)$.

  $\rightarrow \Phi_1\Delta_1 = 0$

  $\rightarrow \gamma_{21} = 0$.

  (2) $\beta_{21} + \beta_{11} = 2$:

  $\rightarrow$ Coeff. of the 2nd pre. var. in the 1st eq. + coeff. of the 1st pre. var. in the 1st eq. = 2.

  $\rightarrow$ $\Phi_1 = (2, 0, 1, 1, 0)$.

  $\rightarrow$ $\Phi_1\Delta_1 = 0$

  $\rightarrow$ $-2 + \beta_{21} + \beta_{11} = 0$.

- Note that $\Phi_1\Delta = (\Phi_1\Delta_1 \; \Phi_1\Delta_{(1)}) = (0, \Phi_1\Delta_{(1)})$

  $\rightarrow \text{rank}(\Phi_1\Delta_{(1)}) \leq \min(R_1, M\text{-}1)$.

Theorem: Rank Condition

A necessary and sufficient condition for the 1st equation to be identified by $\Phi_1\Delta_1 = 0$ is:

$$\text{rank}(\Phi_1\Delta) = M - 1.$$

<Proof> Schmidt, p. 134.

Theorem: Order Condition

A necessary (but not sufficient) condition for identification is $R_1 \geq M - 1$.

- Example 1:

  $C_t = \alpha + \beta INC_t + v_t$

  $INC_t = C_t + INV_t$

  $\rightarrow \quad \begin{pmatrix} C_t & INC & 1 & INV \end{pmatrix} \begin{pmatrix} -1 & 1 \\ \beta & -1 \\ \alpha & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \varepsilon & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.$

  $\rightarrow \quad \beta_{21} = 0.$

  $\rightarrow \quad \Phi_1 = (0, 0, 0, 1).$

  $\rightarrow \quad \Phi_1\Delta = (0, 1) \rightarrow \text{rank}(\Phi_1\Delta) = 1 = M - 1 \rightarrow$ The first eq. is identified.

  $\rightarrow \quad R_1 = 1 = M - 1 \rightarrow$ exactly identified.

- Example 2: Demand and Supply

  D: $Q = a + bP + \varepsilon_1$

  S: $Q = \alpha + \beta P + \gamma W + \varepsilon_2$

$$\rightarrow \quad \begin{pmatrix} Q & P & 1 & W \end{pmatrix} \begin{pmatrix} -1 & -1 \\ b & \beta \\ a & \alpha \\ 0 & \gamma \end{pmatrix} + \begin{pmatrix} \varepsilon_1 & \varepsilon_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

  - Restrictions on the 1st equation: $\beta_{21} = 0$

    → $\Phi_1 = (0, 0, 0, 1)$

    → $\text{rank}(\Phi_1\Delta) = \text{rank}[(0, \gamma)] = 1 = M - 1$, if $\gamma \neq 0$.

    → $R_1 = 1 = M - 1$

    → The 1st equation is exactly identified.

  - Restrictions on the 2nd equation: no restriction

    → $R_2 = 0 < 1 = M - 1$

    → Not identified.

- Example 3:

  $y_1 = \gamma_{21}y_2 + \beta_{11}x_1 + \varepsilon_1$;

  $y_2 = \gamma_{12}y_1 + \gamma_{32}y_3 + \beta_{12}x_1 + \beta_{32}x_3 + \varepsilon_2$;

  $y_3 = \gamma_{23}y_2 + \beta_{23}x_2 + \beta_{33}x_3 + \varepsilon_3$.

$$\rightarrow \quad \begin{pmatrix} y_1 & y_2 & y_3 & x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} -1 & \gamma_{12} & 0 \\ \gamma_{21} & -1 & \gamma_{23} \\ 0 & \gamma_{32} & -1 \\ \beta_{11} & \beta_{12} & 0 \\ 0 & 0 & \beta_{23} \\ 0 & \beta_{32} & \beta_{33} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \varepsilon_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$$

- Restrictions on the 1st equation: $\gamma_{31} = 0$, $\beta_{21} = 0$, $\beta_{31} = 0$.

→ $R_1 = 3 > 2 = M - 1$: Overidentified if identified.

$$\rightarrow \quad \Phi_1 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$\rightarrow \quad rank(\Phi_1\Delta) = rank \begin{pmatrix} 0 & \gamma_{32} & -1 \\ 0 & 0 & \beta_{23} \\ 0 & \beta_{32} & \beta_{33} \end{pmatrix} = 2 = M - 1.$$

→ Overidentified.

- Restrictions on the 2nd eq.: $\beta_{22} = 0$.

→ $R_2 = 1 < 2 = M - 1$: Not identified.

- Restrictions on the 3rd eq.: $\gamma_{13} = 0$, $\beta_{13} = 0$

→ $R_3 = 2 = M - 1$: Just identified if identified.

$$\rightarrow \quad \Phi_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

→ $rank(\Phi_3\Delta) = rank \begin{pmatrix} -1 & \gamma_{12} & 0 \\ \beta_{11} & \beta_{12} & 0 \end{pmatrix} = 2 = M - 1$: Just identified.

- Example 4:

$y_1 = \gamma_{31}y_3 + \beta_{11}x_1 + \varepsilon_1$

$y_2 = \gamma_{12}y_1 + \gamma_{32}y_3 + \beta_{12}x_1 + \beta_{32}x_3 + \varepsilon_2$

$y_3 = \gamma_{13}y_1 + \beta_{13}x_1 + \beta_{23}x_2 + \varepsilon_3$

$$\rightarrow (y_1 \quad y_2 \quad y_3 \quad x_1 \quad x_2 \quad x_3)\begin{pmatrix} -1 & \gamma_{12} & \gamma_{13} \\ 0 & -1 & 0 \\ \gamma_{31} & \gamma_{32} & -1 \\ \beta_{11} & \beta_{12} & \beta_{13} \\ 0 & 0 & \beta_{23} \\ 0 & \beta_{32} & 0 \end{pmatrix} + (\varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3) = (0 \quad 0 \quad 0).$$

- The 1st eq.: Overidentified.
- The 2nd eq.: Not identified
- Restrictions on the 3rd equation: $\gamma_{23} = 0$, $\beta_{33} = 0$.

$\rightarrow$ $R_3 = 2 = M - 1$: Just identified if identified.

$\rightarrow$ $\Phi_3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

$\rightarrow$ $rank(\Phi_3\Delta) = rank\begin{pmatrix} 0 & -1 & 0 \\ 0 & \beta_{32} & 0 \end{pmatrix} = 1 \neq M - 1$: Not identified.

$\rightarrow$ The restrictions on the 3rd equations are a subset of those on the first equations.

(2) Cases with covariance restrictions.

- Cases where models can be consistently estimated by MLE (FIML).
- Example 1:

  1) $y_1 = \beta_{11}x_1 + \varepsilon_1$

  2) $y_2 = \gamma_{12}y_1 + \beta_{12}x_1 + \varepsilon_2$

  $\rightarrow$ 1) is identified and 2) is not identified.

  $\rightarrow$ If $cov(\varepsilon_1,\varepsilon_2) = 0$, then 2) is also identified.

- Suppose: i) R_1 linear restrictions on the first equations given by $\Phi_1\Delta_1 = 0$.

  ii) J_1 zero restrictions on σ_{1j}:

  e.g., $\sigma_{1,M-J_1+1} = \sigma_{1,M-J_2+2} = ... = \sigma_{1,M} = 0$:

  That is, the error term of the 1st equation is uncorrelated with the errors of the last J_1 equations.

- Then, partition:

$$\Sigma = \begin{pmatrix} \Sigma_{M-J_1} \\ \Sigma_{J_1} \end{pmatrix}.$$

where Σ_{M-J_1} has $(M-J_1)$ rows and Σ_{J_1} has J_1 rows.

Theorem: Generalized Rank Condition (GRC)

A necessary (but not sufficient) condition for the identification of the first equation is

$$rank\begin{pmatrix}\Phi_1\Delta \\ \Sigma_{J_1}\end{pmatrix}_{(R_1+J_1)\times M} = M-1$$

Theorem:

The sufficient conditions for the identification of the first equation are:

i) GRC holds.

ii) The last J_1 equations are identified.

<Proof>See Schmidt.

- Example 2: Continue Example 1.

  1) $y_1 = \beta_{11}x_1 + \varepsilon_1$

  2) $y_2 = \gamma_{12}y_1 + \beta_{12}x_1 + \varepsilon_2$

  $\rightarrow$ The 1st equation is identified by the exclusive restriction, $\gamma_{21} = 0$.

  $\rightarrow$ The 2nd equation: $R_2 = 0$ and $J_2 = 1$.

  $\rightarrow \Sigma = \begin{pmatrix}\sigma_{11} & 0 \\ 0 & \sigma_{22}\end{pmatrix} \rightarrow \Sigma_{M-J_2} = (0\ \sigma_{22});\ \Sigma_{J_2} = (\sigma_{11}\ 0)$.

  $\rightarrow rank\begin{pmatrix}\Phi_2\Delta \\ \Sigma_{J_2}\end{pmatrix} = rank\left(\Sigma_{J_2}\right) = 1 = M-1$:

  $\rightarrow$ So, GRC holds for the 2nd equation and the first equation is identified.

  $\rightarrow$ So, the 2nd equation is identified!!!

- Example 3:

  1) $y_1 = \gamma_{21}y_2 + \varepsilon_1$

  2) $y_2 = \gamma_{12}y_1 + \varepsilon_2$

  with $\sigma_{12} = 0$.

  → No restrictions on coefficients.

  → The 1st equation: $J_1 = 1$.

  → $\Sigma = \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix}$ → $\Sigma_{J_1} = (0\ \sigma_{22})$.

  → $rank\begin{pmatrix} \Phi_1\Delta \\ \Sigma_{J_1} \end{pmatrix} = rank\left(\Sigma_{J_1}\right) = 1 = M-1$: GRC holds.

  → But we can't determine whether the 1st equation is identified, because we can't determine whether the 2nd equation is identified.

  → This problem is symmetric. So, we can't determine whether the two equations are identified.

  → As a matter of fact, we can't identify the two equations, because one equation is a linear function of the other. They are the same equations.

- Example 4:

  $y_1 = \gamma_{21}y_2 + \gamma_{31}y_3 + \beta_{11}x_1 + \beta_{31}x_3 + \varepsilon_1$ ($R_1 = 1$)

  $y_2 = \gamma_{12}y_1 + \beta_{12}x_1 + \varepsilon_2$ ($R_2 = 3$)

  $y_3 = \gamma_{13}y_1 + \beta_{23}x_2 + \beta_{22}x_3 + \varepsilon_3$ ($R_3 = 2$),

  with $\sigma_{ij} = 0$ for $i \neq j$.

→ You can show that the last two equations are identified by exclusive restrictions.

→ Can't identify the 1st equation by exclusive restrictions alone. ($R_1 = 1 < M - 1$)

→ $$\Sigma = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \rightarrow \Sigma_{J_1} = \begin{pmatrix} 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}.$$

→ $\Phi_1 = (0, 0, 0, 0, 1, 0) \rightarrow \Phi_1\Delta = (0, 0, \beta_{23})$.

→ $$rank\begin{pmatrix} \Phi_1\Delta \\ \Sigma_{J_1} \end{pmatrix} = rank\begin{pmatrix} 0 & 0 & \beta_{23} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} = 2 = M - 1: \text{ GRC holds.}$$

→ The first equation is identified!!!

[3] Single Equation Estimators

- How could we estimate each individual equation?

(1) IV and 2SLS Estimation

- Notation:
  - y_{t1} = the dependent variable of the first equation in a simultaneous equations system:
    - $y_1 = (y_{11}, y_{21}, \ldots , y_{T1})'$.
  - Y_{t1} = the $(m_i-1)\times 1$ vector of endogenous regressors in the first equation.
    - $Y_1 = (Y_{11}, Y_{21}, \ldots , Y_{T1})'$.
  - X_{t1} = the $k_i\times 1$ vector of exogenous (or predetermined) regressors in the first equation.
    - $X_1 = (X_{11}, X_{21}, \ldots , X_{T1})'$.

- Idea:
  - Consider the estimation of the first equation:

$$y_1 = Y_1\gamma_1 + X_1\beta_1 + \varepsilon_1 \equiv Z_1\delta_1 + \varepsilon_1.$$

  - Be aware of the difference between Δ_1 and δ_1:
    - Let $Y = (y_1, Y_1, Y_1^*)$, where Y_1^* is the matrix of the endogenous variables that are excluded from the first equation.
    - Let $X = (X_1, X_1^*)$, where X_1^* is the matrix of the exogenous variables that are excluded from the first equation.

- Δ_1 = vector of coefficients on all of the endogenous and exogenous variables in the first equation.

  $\rightarrow$ vector of coefficients on (Y,X) = $(y_1,Y_1,Y_1^*,X_1,X_1^*)$.

- δ_1 = vector of coefficients on the endogenous and exogenous included in the first equation.

  $\rightarrow$ vector of coefficients on $Z_1 = (Y_1,X_1)$.

- The OLS estimator of $\delta_1 = (\gamma_1',\beta_1')'$ will be inconsistent because Y_1 and ε_1 are correlated.
- We can estimate the equation using the IV method.
  - Note that by our assumptions, $E(X'\varepsilon_1) = 0_{K\times 1}$. Further, the reduced form equations indicate that the veriables in X are correlated with Y_1. So, variables in X can be used as instrumental variables.

- Two possible cases:

1) If $k_e = m_i - 1$ (# of variables in X $(k_e+(K-k_e))$ = # of variables in Y_1 and X_1 $(m_i-1+(K-k_e))$):
   - $\hat{\delta}_{1,IV} = (X'Z_1)^{-1}X'y_1$;
   - estimated $Cov(\hat{\delta}_{1,IV}) = s_{11}(X'Z_1)^{-1}X'X(Z_1'X)^{-1}$,

     where $s_{11} = (y_1 - Z_1\hat{\delta}_{1,IV})'(y_1 - Z_1\hat{\delta}_{1,IV})/T$;
   - $\hat{\delta}_{1,IV}$ is consistent and asymptotically normal.

2) If $k_e > m_i - 1$, then # of instruments > # of parameters (p_1):

- $$\hat{\delta}_{1,2SLS} = (Z_1'X(X'X)^{-1}X'Z_1)^{-1}Z_1'X(X'X)^{-1}X'y_1 = [Z_1'P(X)Z_1]^{-1}Z_1'P(X)y_1.$$

- Estimated $Cov(\hat{\delta}_{1,2SLS}) = s_{11}(Z_1'P(X)Z_1)^{-1}$,

  where $s_{11} = (y_1 - Z_1\hat{\delta}_{1,2SLS})'(y_1 - Z_1\hat{\delta}_{1,2SLS})/T$.

- The 2SLS estimator is consistent and asymptotically normal.

- The 2SLS estimator minimizes:

$$(y - Z_1\delta_1)'X(\hat{\sigma}_{11}X'X)^{-1}X'(y - Z_1\delta_1) = \frac{(y - Z_1\delta_1)'X(X'X)^{-1}X'(y - Z_1\delta_1)}{\hat{\sigma}_{11}},$$

  where $\hat{\sigma}_{11}$ is any consistent estimator of σ_{11}.

  - Notice that the OLS estimator of δ_1 minimizes

$$(y_1 - Z_1\delta_1)'(y_1 - Z_1\delta_1).$$

Theorem:

2SLS is consistent if the rank order condition is satisfied. If the order condition is satisfied but the rank condition is not, 2SLS is computable, but it is inconsistent.

<proof> See the Technical Notes below.

Comment: A case where the order condition is satisfied and the rank condition is not.

$y_1 = \gamma_{31}y_3 + \beta_{11}x_1 + \varepsilon_1$

$y_2 = \gamma_{12}y_1 + \gamma_{32}y_3 + \beta_{12}x_1 + \beta_{32}x_3 + \varepsilon_2$

$y_3 = \gamma_{13}y_1 + \beta_{13}x_1 + \beta_{23}x_2 + \varepsilon_3$

$\rightarrow$ The 3rd equation is not identified.

$\rightarrow$ Recall that 2SLS on the 3rd equation = OLS of y_3 on $\hat{y}_1$ (fitted value from regression of y_1 on x_1, x_2 and x_3), x_1 and x_2.

$\rightarrow$ If we derive the reduced form equation for y_1, we can see that y_1 depends only on x_1 and x_2 (check this by yourself). This means that $\hat{y}_1$ converges the fitted values from regression of y_1 on x_1 and x_2 as $T \rightarrow \infty$. That is, for $T \rightarrow \infty$, $\hat{y}_1$, x_1 and x_2 are perfectly correlated. So, 2SLS cannot be asymptotically defined. The 2SLS estimator does not have the usual asymptotic distribution.

(2) Limited Information ML (LIML)

- Consider the 1st equation: $y_1 = Y_1\gamma_1 + X_1\beta_1 + \varepsilon_1$.
- Consider the following system:

  $y_1 = Y_1\gamma_1 + X_1\beta_1 + \varepsilon_1$ (Structural equation)

  $y_2 = X\pi_2 + u_2$ (Reduced form)

  :

  $y_M = X\pi_M + u_M$ (Reduced form)

- LIML = MLE on this system of equations (Hausman, 1975).

- LIML estimator
  - Recall that the 2SLS estimator of δ_1 minimizes:

$$(y_1 - Z_1\delta_1)'X(\hat{\sigma}_{11}X'X)^{-1}X'(y_1 - Z_1\delta_1).$$

  - LIML estimator of δ_1 (say, $\hat{\delta}_{1,LIML}$) minimizes:

$$(y_1 - Z_1\delta_1)'X\left(\frac{(y_1 - Z_1\delta_1)'(y_1 - Z_1\delta_1)}{T}X'X\right)^{-1}X'(y_1 - Z_1\delta_1)$$

$$= T\frac{(y_1 - Z_1\delta_1)'X(X'X)^{-1}X'(y_1 - Z_1\delta_1)}{(y_1 - Z_1\delta_1)'(y_1 - Z_1\delta_1)}.$$

  - Computation:
    - Let $P(X) = X(X'X)^{-1}X'$.
    - Let $\delta_1^o = \begin{pmatrix} -1 \\ \delta_1 \end{pmatrix}$ and $W_1 = (y_1, Z_1) = (y_1, Y_1, X_1)$ such that

$$W_1\delta_1^o = -(y_1 - Z_1\delta_1).$$

- Then, the objective function of LIML can be written:

$$T\frac{\delta_1^{o\prime}W_1'P(X)W_1\delta_1^o}{\delta_1^{o\prime}W_1'W_1\delta_1^o}.$$

- Let λ_1 be the smallest eigenvalue of the matrix:

$$\left(W_1'W_1\right)^{-1}W_1'P(X)W_1.$$

- Then, the LIML estimator of $\hat{\delta}_{1,LIML}^o = \begin{pmatrix} -1 \\ \hat{\delta}_{1,LIML} \end{pmatrix}$ = the eigenvector

  corresponding to λ_1 normalized so that the first entry equals (-1).

- $Cov(\hat{\delta}_{1,LIML}) = s_{11,LIML}(Z_1'P(X)Z_1)^{-1}$,

  where $s_{11,LIML} = \dfrac{(y_1 - Z_1\hat{\delta}_{1,LIML})'(y_1 - Z_1\hat{\delta}_{1,LIML})}{T}$.

- The minimized value of the objective function = $T\lambda_1$.

  [Why?]

- $\left[W_1'W_1\right]^{-1}\left[W_1'P(X)W_1\right]\lambda_1 = \lambda_1\hat{\delta}_{1,LIML}^o$

  $\rightarrow [W_1P(X)W_1]\hat{\delta}_{1,LIML}^o = \lambda_1\left[W_1'W_1\right]\hat{\delta}_{1,LIML}^o$.

  $\rightarrow \hat{\delta}_{1,LIML}^{o}{}'[W_1P(X)W_1]\hat{\delta}_{1,LIML}^o = \lambda_1\hat{\delta}_{1,LIML}{}'\left[W_1'W_1\right]\hat{\delta}_{1,LIML}^o$.

  $\rightarrow r \equiv T\dfrac{\hat{\delta}_1^{o\prime}W'P(X)W\hat{\delta}_1^o}{\hat{\delta}_1^{o\prime}W'W_1\hat{\delta}_1^o} = T\lambda_1$.

- If the first equation is exactly identified, r = T.

- If the first equation is correctly specified,

  $$T(\lambda_1 - 1) \rightarrow_d \chi^2(k_e - m_i + 1),$$

  where k_e-m_i+1 = # of variables in Z_1 - # of variables in X.

Digression to Eigenvalue and Eigenvector

- Let A be a p×p square matrix.
- Then, there exists a scalar λ and a p×1 vector ξ such that

  $$A\xi = \lambda\xi .$$

  The scalar λ is called "eigenvalue" and ξ is called "eigenvector corresponding to λ.
- The matrix A has p eigenvalues.
- The eigenvalues are the solutions of the equations $\det(A-\lambda I_p) = 0$.

End of Digression

- Properties of LIML
  - LIML = 2SLS, asymptotically.
  - But, in general, LIML has better finite sample properties than 2SLS.
    In addition, LIML is less sensitive to weak instrumental variables.
  - LIML = 2SLS, numerically, if the equation estimated is exactly identified.
  - Inefficient compared to 3SLS or FIML.
  - The specification of the first equation can be tested by T(r-1).

[Technical Note]

Consistency and Asymptotic normality of $\hat{\delta}_{1,2SLS}$:

- $\hat{\delta}_{1,2SLS}$ $= (Z_1'X(X'X)^{-1}X'Z_1)^{-1}Z_1'X(X'X)^{-1}X'y_1$

  $= (Z_1'X(X'X)^{-1}X'Z_1)^{-1}Z_1'X(X'X)^{-1}X'(Z_1\delta_1 + \varepsilon_1)$

  $= \delta_1 + (Z_1'X(X'X)^{-1}X'Z_1)^{-1}Z_1'X(X'X)^{-1}X'\varepsilon_1$

  $= \delta_1 + [(Z_1'X/T)(X'X/T)^{-1}(X'Z_1/T)]^{-1}(Z_1'X/T)(X'X/T)^{-1}X'\varepsilon_1/T.$

$\rightarrow$ plim $Z_1'X/T$, plim $X'X/T$ are finite. And

plim $X'\varepsilon_1/T$ = plim $\Sigma_t X_{t\bullet}\varepsilon_{t1}/T$ = plim $\Sigma_t E(X_{t\bullet}\varepsilon_{t1})/T = 0$.

$\rightarrow$ plim $\hat{\delta}_{1,2SLS} = \delta_1$ (Consistent)

- $$\sqrt{T}\left(\hat{\delta}_{1,2SLS} - \delta_1\right) = \left(\left(\frac{Z_1'X}{T}\right)\left(\frac{X'X}{T}\right)^{-1}\left(\frac{X'Z_1}{T}\right)\right)^{-1}\left(\frac{Z_1'X}{T}\right)\left(\frac{X'X}{T}\right)^{-1}\left(\frac{\Sigma_t X_{t\bullet}\varepsilon_{t1}}{\sqrt{T}}\right)$$

$$\rightarrow_d N\left(0_{p_1\times 1}, \sigma_{11}\left(\left(p\lim\frac{Z_1'X}{T}\right)\left(p\lim\frac{X'X}{T}\right)^{-1}\left(p\lim\frac{X'Z_1}{T}\right)\right)^{-1}\right)$$

Why?

$$\frac{\Sigma_t X_{t\bullet}\varepsilon_{t1}}{\sqrt{T}} \rightarrow_d N\left(0, \lim\frac{1}{T}\Sigma_t Cov(X_{t\bullet}\varepsilon_{t1})\right) = N\left(0, \lim\frac{1}{T}\Sigma_t \sigma_{11} X_{t\bullet}X_{t\bullet}'\right)$$

$$= N\left(0, \sigma_{11}\lim\frac{1}{T}X'X\right).$$

- $Cov(\hat{\delta}_{1,2SLS}) \approx \sigma_{11}(Z_1'P(X)Z_1)^{-1}.$

[4] Systems Estimators

- How could we estimate all equations jointly?
- Is there any gain of estimating the whole system?

(1) Three Stage Least Squares (3SLS)

- Equation i: $y_i = Y_i\gamma_i + X_i\beta_i + \varepsilon_i \equiv Z_i\delta_i + \varepsilon_i$, i =1, ... , M.

$$y_* = \begin{pmatrix} y_1 \\ y_2 \\ : \\ y_M \end{pmatrix}; Z_* = \begin{pmatrix} Z_1 & 0 & \ldots & 0 \\ 0 & Z_2 & \ldots & 0 \\ : & : & & : \\ 0 & 0 & \ldots & Z_M \end{pmatrix}.$$

$\rightarrow y_* = Z_*\delta_* + \varepsilon_*$, with $\mathrm{Cov}(\varepsilon_*) = \Sigma\otimes I_T$.

- Consider the following auxiliary regression models:

$$y_i = \hat{Z}_i\delta + err_i \text{ where } \hat{Z}_i = (\hat{Y}_i, X_i) = [P(X)Y_i, X_i] \text{ and i = 1, ... , M.}$$

$\rightarrow y_* = \hat{Z}_*\delta_* + err_*$.

- 3SLS estimator = FGLS on this auxiliary system of equations:

  - $\tilde{\delta}_* = \left(\hat{Z}_*'(S^{-1}\otimes I_T)\hat{Z}_*\right)^{-1}\hat{Z}_*'(S^{-1}\otimes I_T)y_*$,

  where S is a consistent estimate of Σ based on 2SLS residuals: That is,

$$s_{ij} = \frac{(y_i - Z_i\hat{\delta}_{i,2SLS})'(y_i - Z_i\hat{\delta}_{i,2SLS})}{T}.$$

- Alternative Representations of 3SLS

1) $\tilde{\delta}_* = \left(Z_*'(S^{-1}\otimes P(X))Z_*\right)^{-1} Z_*'(S^{-1}\otimes P(X))y_*$.

Why?

$$\hat{Z}_* = \begin{pmatrix} \hat{Z}_1 & 0 & ... & 0 \\ 0 & \hat{Z}_2 & ... & 0 \\ : & : & & : \\ 0 & 0 & ... & \hat{Z}_M \end{pmatrix} = \begin{pmatrix} P(X)Z_1 & 0 & ... & 0 \\ 0 & P(X)Z_2 & ... & 0 \\ : & : & & : \\ 0 & 0 & ... & P(X)Z_M \end{pmatrix}$$

$$= \begin{pmatrix} P(X) & 0 & ... & 0 \\ 0 & P(X) & ... & 0 \\ : & : & & : \\ 0 & 0 & ... & P(X) \end{pmatrix} \begin{pmatrix} Z_1 & 0 & ... & 0 \\ 0 & Z_2 & ... & 0 \\ : & : & & : \\ 0 & 0 & ... & Z_M \end{pmatrix} = (I_M \otimes P(X))Z_*$$

2) $$\tilde{\delta}_* = \begin{pmatrix} s^{11}\hat{Z}_1'\hat{Z}_1 & s^{12}\hat{Z}_1'\hat{Z}_2 & ... & s^{1M}\hat{Z}_1'\hat{Z}_M \\ s^{21}\hat{Z}_2'\hat{Z}_1 & s^{22}\hat{Z}_2'\hat{Z}_2 & ... & s^{2M}\hat{Z}_2'\hat{Z}_M \\ : & : & & : \\ s^{M1}\hat{Z}_M'\hat{Z}_1 & s^{M2}\hat{Z}_M'\hat{Z}_2 & ... & s^{MM}\hat{Z}_M'\hat{Z}_M \end{pmatrix} \begin{pmatrix} \Sigma_j s^{1j}\hat{Z}_1' y_j \\ \Sigma_j s^{2j}\hat{Z}_2' y_j \\ : \\ \Sigma_j s^{Mj}\hat{Z}_M' y_j \end{pmatrix}.$$

3) Consider:

$$(I_M \otimes X')y_* = (I_M \otimes X')Z_*\delta_* + (I_M \otimes X')\varepsilon_* .$$

- 3SLS = FGLS on this system.

  $Cov[(I_M \otimes X')\varepsilon_*] = E[(I_M \otimes X')\varepsilon_*\varepsilon_*'(I_M \otimes X)] = (I_M \otimes X')E(\varepsilon_*\varepsilon_*')E(I_M \otimes X)$

  $= (I_M \otimes X')(\Sigma \otimes I_T)(I_M \otimes X) = \Sigma \otimes X'X.$

  $\rightarrow$ (Estimated $Cov[(I_M \otimes X')\varepsilon_*])^{-1} = S^{-1} \otimes (X'X)^{-1}$

  $\rightarrow$ FGLS $= [Z_*'(I_M \otimes X)(S^{-1} \otimes (X'X)^{-1})(I_M \otimes X')Z_*]^{-1}$

  $\times Z_*'(I_M \otimes X)(S^{-1} \otimes (X'X)^{-1})(I_M \otimes X')y_*$

  $= [Z_*'(S^{-1} \otimes X(X'X)^{-1}X')Z_*]^{-1} Z_*'(S^{-1} \otimes X(X'X)^{-1}X')y_* = \text{3SLS}$

- 2SLS on each equation

  = FGLS using $I_M \otimes (X'X)^{-1}$ instead of $S^{-1} \otimes (X'X)^{-1}$.

Theorem:

If all the equations are identified by exclusive restrictions and if Σ is nonsingular, then 3SLS is consistent and

$$\sqrt{T}(\tilde{\delta}_* - \delta_*) \Rightarrow N\left(0, \left[p\lim \frac{\hat{Z}_*'(S^{-1} \otimes I_T)\hat{Z}_*}{T}\right]^{-1}\right).$$

<Proof>See Schmidt, P. 207.

Comments:

- Use $\left[\hat{Z}_*'(S^{-1}\otimes I_T)\hat{Z}_*\right]^{-1}$ as $Cov(\tilde{\delta}_*)$.
- 3SLS is asymptotically more efficient than 2SLS (Schmidt, p. 209) except the following three exceptional cases:
    a) If Σ is diagonal, 3SLS = 2SLS asymptotically.
    b) If Σ is diagonal, and if $s_{ij} = 0$ for $i \neq j$ are used to estimate Σ, 3SLS = 2SLS, numerically.
    c) If all equations are exactly identified by the restrictions on Γ and B, ILS = 2SLS = LIML = 3SLS, numerically.

- If any equation fails the order condition for identification, 3SLS is not computable.
- If one equation is exactly identified, it does not help the efficiency of 3SLS estimatators for other equations: The 3SLS estimators for the other equations computed jointly with the identified equation are numerically equal to the 3SLS estimators for the other equations computed excluding the identified equation.
- If there are some linear restrictions $R\delta_* = r$, the restricted 3SLS is consistent and efficient relative to the unrestricted 3SLS.
- If Σ is unrestricted, 3SLS is asymptotically efficient.
- The 3SLS estimator explained above is a two-step estimator. Can do iterative 3SLS. But this iterative 3SLS = two-step 3SLS asymptotically.

(2) Minimum Distance Estimator (MD)

- Chamberlain, Handbook, Chapter 22.
- Note that the reduced form parameters Π (or vec(Π) = π_*) are functions of structural parameters: say, $\pi_* = f(\Gamma, B)$.
- MD principle: Find Γ and B which minimize the distance between $\hat{\pi}_*$ and $f(\Gamma, B)$.
- Specifically, minimize

$$\left(\hat{\pi}_* - f(\Gamma, \mathrm{B})\right)' A^{-1} \left(\hat{\pi}_* - f(\Gamma, \mathrm{B})\right),$$

where A is any positive definite matrix.

- Optimal choice of A = $Cov(\hat{\pi}_*)$.
- Optimal MD (OMD) = 3SLS, asymptotically.

(3) Full Information (FIML)

- Assume that error terms are normally distributed.
- FIML of the parameters in YΓ + XB + E are the values of Γ, B and Σ which minimize

$$l_T = -\frac{MT}{2}\ln(2\pi) - \frac{T}{2}\ln[\det(\Sigma)] + T\ln[\det(\Sigma)] - \frac{1}{2} trace\left\{\Sigma^{-1}(Y\Gamma + X\mathrm{B})'(Y\Gamma + X\mathrm{B})\right\},$$

subject to all a priori restrictions (see Schmidt p. 129).

- Let θ be a parameter vector of all of the parameters in Γ, B and Σ.

- Let $H_T(\theta) = \frac{\partial^2 l_T}{\partial\theta\partial\theta'}$ and $I_T(\theta_o) = E[-H_T(\theta_o)]$. Then,

$$\sqrt{T}(\hat{\theta}_{ML} - \theta_o) \Rightarrow N\left(0, \left[p\lim \frac{1}{T} I_T(\theta_o)\right]^{-1}\right).$$

- Can use $\left[I_T(\hat{\theta}_{ML})\right]^{-1}$ as an estimated $Cov(\hat{\theta}_{ML})$.

  Or use $\left[-H_T(\hat{\theta}_{ML})\right]^{-1}$.

- BHHH (Berndt, Hall, Hall and Hausman, 1974, Annals of Economic and Social Managements):
  - Let $l_t = -(M/2)\ln(2\pi) - (.5)\ln[\det(\Sigma)] + \ln[\det(\Gamma)]$

    $$- (.5)(Y_{t\bullet}'\Gamma + X_{t\bullet}'B)'\Sigma^{-1}(Y_{t\bullet}'\Gamma + X_{t\bullet}'B).$$
  - Let $B_T(\theta) = \Sigma_t \frac{\partial l_t}{\partial\theta}\frac{\partial l_t}{\partial\theta'}$.
  - Then, plim (1/T) $B_T(\theta)$ = - plim (1/T) $H_T(\theta)$, if data are iid.
  - So, can use $\left[B_T(\hat{\theta}_{ML})\right]^{-1}$ as an estimated $Cov(\hat{\theta}_{ML})$.

- If error are not normally distributed?
  - $\hat{\theta}_{ML}$ is still consistent, but neither $\left[-H_T(\hat{\theta}_{ML})\right]^{-1}$ nor $\left[B_T(\hat{\theta}_{ML})\right]^{-1}$ is a consistent estimator of $Cov(\hat{\theta}_{ML})$. The correct estimate of $Cov(\hat{\theta}_{ML})$ is:

$$\left[-H_T(\hat{\theta}_{ML})\right]^{-1} B_T(\hat{\theta}_{ML})\left[-H_T(\hat{\theta}_{ML})\right]^{-1}.$$

Theorem

If there is no restriction on Σ and the errors are normally distributed, 3SLS $\Rightarrow$ FIML.

<Proof> Schmidt, p. 224.

- Comments:
  - Without restrictions on Σ, 3SLS is asymptotically efficient.
  - If there are some restrictions on Σ, the FIML with the restrictions is more efficient.

    $\rightarrow$ In SUR, restrictions on Σ do not help efficiency, but they do for SEM.
  - Iterative 3SLS is not numerically identical to FIML.
  - If all equations are identified and no restrictions on Σ, then IV = 2SLS = LIML = 3SLS (both two-step and iterative) = FIML.

[5] Comparisons of Estimators

(1) OLS on individual equations:

- Biased, inconsistent, inappropriate.
- However, have smaller variances. → Less sensitive to data and specification.

(2) Single Equations Estimators (ILS, 2SLS, LIML)

- Advantage: Consistency.
- 2SLS = LIML, numerically, if exactly identified.
- For the over-identified equations, 2SLS and LIML are more efficient.
- Advantage of 2SLS over LIML: Simple
- Advantage of LIML over 2SLS:
  - Invariance.
    - Example: (1) $y_1 = \gamma y_2 + \varepsilon_1 \rightarrow$ (2) $y_2 = (1/\gamma)y_1 - (1/\gamma)\varepsilon_1$
    - 2SLS of γ from (1) $\neq$ 1/(2SLS of $1/\gamma$ from (2)), if overidentified
    - But LIML of γ from (1) = 1/(LIML of $1/\gamma$ from (2)), even if overidentified.
    - If exactly identified, 2SLS of γ from (1) = 1/(2SLS of $1/\gamma$ from (2)).
      - Intuition:
        - Supposed that X contains only a single IV variable.
        - 2SLS of γ from (1) = $(X'y_2)^{-1}X'y_1$.
        - 2SLS of $1/\gamma$ from (2) = $(X'y_1)^{-1}X'y_2$.
  - Better finite sample properties.

(3) System Methods (3SLS, FIML)

- Advantage over single equations estimators:
  - greater efficiency.
    - 3SLS: No gain if all equations are exactly identified or errors are uncorrelated across different equations.
    - FIML: No gain if all equations are exactly identified and no covariance restrictions on Σ. However, great efficiency gain over both 3SLS and 2SLS if the errors are uncorrelated and if this information is used in FIML.

- Disadvantage:
  - Computational complexity.
  - Sensitive to misspecification.

[6] Specification Test

(1) Test Based on LIML (Anderson and Rubin, 1950):

- Model: $y_1 = Y_1\gamma_1 + X_1\beta_1 + \varepsilon_1$.
- $T(\lambda_1 - 1) \to_d \chi^2(k_e - m_i + 1)$.

(2) Test Based on 2SLS

(Hansen, 1982, Econometrica; Hausman, 1984, Handbook)

Digression to Uncentered R^2:

- Consider a single regression model $y = X\beta + \varepsilon$.
- Let $\hat{\beta}$ be any consistent estimator of β; and $e = y - X\hat{\beta}$.
- $R^2 = 1$ - SSE/SST, where $SSE = e'e$ and $SST = \Sigma_t(y_t - \bar{y})^2 = y'y - T\bar{y}^2$.
- $R_u^2 = 1 - \frac{e'e}{y'y}$.
- If $\bar{y} = 0$, $R_u^2 = R^2$.

End of Digression

- Model: $y_1 = Y_1\gamma_1 + X_1\beta_1 + \varepsilon_1 = Z_1\delta_1 + \varepsilon_1$
- Instruments: X (all of predetermined or exogenous regressors).

- Test procedure:
  - Let $\hat{\varepsilon}_1$ be the vector of the 2SLS residuals (= $y_1 - Z_1\hat{\delta}_{1,2SLS}$). (If X_1 contains the vector of ones as a column, then it can be shown that the sample mean of the 2SLS residuals equals zero.)
  - Regress $\hat{\varepsilon}_1$ on X and get R_u^2 (= R^2 if X_1 contains the vector of ones).
  - Then,

    $$J_T \equiv TR_u^2 \to_d \chi^2(k_e-(m_i-1)),$$

    if $E(X_{t\bullet}\varepsilon_{t1}) = 0$.

(3) Test based on FIML

See Greene (pp. 700-701).

(4) Testing Exogeneity of a Single regressor Based on 2SLS

(4.1) Hausman test (1978, Econometrica)

- H_o: My model is correctly specified.
- Consider two types of estimators
  - $\tilde{\theta}$: efficient and consistent under H_o, but inconsistent under H_a.
  - $\hat{\theta}$: consistent both under H_o and H_a, but inefficient under H_o.

- Let $\tilde{\Phi} = Cov(\tilde{\theta})$ and $\hat{\Phi} = Cov(\hat{\theta})$.
- Under H_o, $(\hat{\Phi} - \hat{\Phi})$ is positive semidefinite.
- Under H_o,

$$H_T = \left(\hat{\theta} - \tilde{\theta}\right)' \left(\hat{\Phi} - \tilde{\Phi}\right)^{+} \left(\hat{\theta} - \tilde{\theta}\right) \rightarrow_p \chi^2(df = rank\left(\hat{\Phi} - \tilde{\Phi}\right)).$$

- Intuition:
  - Under H_o, both estimators are consistent, so the distance between the two would be small.
  - However, under H_a, only $\hat{\theta}$ is consistent. Thus, the distance would be large.

(4.2) Testing Exogeneity of a Single Regressor by the Hausman Test

- Model: $y_1 = Y_1\gamma_1 + X_1\beta_1 + \varepsilon_1 = Z_1\delta_1 + \varepsilon_1$.
- $X = [X_1, X_1^*]$ and let $Y_1 = [h, g, ...]$.
- Wish to test H_o: h is an exogenous or predetermined regressor.

- Test Procedure:
  - Let $\hat{\delta}_{1,2SLS}$ be the 2SLS estimator using X as instruments; $\Phi_{1,2SLS} = Cov(\hat{\delta}_{1,2SLS})$. Let $\hat{\delta}^*_{1,2SLS}$ be the 2SLS estimator using [X,h] as instruments; and $\Phi^*_{1,2SLS} = Cov(\hat{\delta}^*_{1,2SLS})$.
  - $H_T = \left(\hat{\delta}_{1,2SLS} - \hat{\delta}^*_{1,2SLS}\right)' \left(\Phi_{1,2SLS} - \Phi^*_{1,2SLS}\right)^+ \left(\hat{\delta}_{1,2SLS} - \hat{\delta}^*_{1,2SLS}\right) \rightarrow_d \chi^2(1)$.

    [Greene (p. 702) assumes that $\Phi_{1,2SLS} - \Phi^*_{1,2SLS}$ is invertible. However, theoretically, $rank\left(\Phi_{1,2SLS} - \Phi^*_{1,SLS}\right) = 1$.]

- Alternative by Newey (1985, JEC) and Eichenbaum, Hasnsen and Singleton (1988, JPE)
  - Let $\hat{\varepsilon}_1$ be the vector of 2SLS residuals using X as instruments.
  - Let $\hat{\varepsilon}_1^*$ be the vector of 2SLS residuals using [X,h] as instruments.
  - $J_T = T\bullet(R_u^2$ from the regression of $\hat{\varepsilon}_1$ on X).
  - $J_T^* = T\bullet(R_u^2$ from the regression of $\hat{\varepsilon}_1^*$ on [X,h]).
  - $D_T \equiv J_T^* - J_T \Rightarrow \chi^2(df = 1)$.
  - This test is asymptotically identical to H_T.

(4.3) Testing Exogeneity of Two Regressors by the Hausman Test

- Model: $y_1 = Y_1\gamma_1 + X_1\beta_1 + \varepsilon_1 = Z_1\delta_1 + \varepsilon_1$.
- $X = [X_1, X_1^*]$ and let $Y_1 = [h, g, ...]$.
- Wish to test H_o: h and g are exogenous regressors.

- Test Procedure:
  - Let $\hat{\delta}_{1,2SLS}$ be the 2SLS estimator using X as instruments; $\Phi_{1,2SLS} = Cov(\hat{\delta}_{1,2SLS})$. Let $\hat{\delta}^*_{1,2SLS}$ be the 2SLS estimator using [X,h,g] as instruments; and $\Phi^*_{1,2SLS} = Cov(\hat{\delta}^*_{1,2SLS})$.
  - $H_T = \left(\hat{\delta}_{1,2SLS} - \hat{\delta}^*_{1,2SLS}\right)'\left(\Phi_{1,2SLS} - \Phi^*_{1,2SLS}\right)^+\left(\hat{\delta}_{1,2SLS} - \hat{\delta}^*_{1,2SLS}\right) \to_p \chi^2(2)$.

- Alternative by Newey (1985, JEC) and Eichenbaum, Hasnsen and Singleton (1988, JPE)
  - Let $\hat{\varepsilon}_1$ be the vector of 2SLS residuals using X as instruments.
  - Let $\hat{\varepsilon}_1^*$ be the vector of 2SLS residuals using [X,h,g] as instruments.
  - $J_T = T \bullet (R_u^2$ from the regression of $\hat{\varepsilon}_1$ on X).
  - $J_T^* = T \bullet (R_u^2$ from the regression of $\hat{\varepsilon}_1^*$ on [X,h,g]).
  - $D_T \equiv J_T^* - J_T \to_p \chi^2(df = 2)$.
  - This test is asymptotically identical to H_T.

[7] Empirical Example

- Use **mwemp.wf1**. You can download this file from my web page.
- This is the data set of working married women in 1981 sampled from PSID. Total number of observations are 923, and 17 variables are observed.

VARIABLES	DEFINITION
LRATE	LOG OF HOURLY WAGE RATE ($)
ED	YEARS OF EDUCATION
URB	URB=1 IF RESIDENT IN SMSA
MINOR	MINOR=1 IF BLACK AND HISPANIC
AGE	YEARS OF AGE
TENURE	MONTHS UNDER THE CURRENT EMPLOYER
EXPP	NUMBER OF YEARS WORKED SINCE AGE 18
REGS	REGS=1 IF LIVES IN THE SOUTH OF U.S.
OCCW	OCCW=1 IF WHITE COLOR
OCCB	OCCB=1 IF BLUE COLOR
INDUMG	INDUMG=1 IF IN THE MANUFACTURING INDUSTRY
INDUMN	INDUMN=1 IF NOT IN MANUFACTURING SECTOR
UNION	UNION=1 IF UNION MEMBER
UNEMPR	% UNEMPLOYMENT RATE IN THE RESIDENT'S COUNTY, 1980
LOFINC	LOG OF OTHER FAMILY MEMBER'S INCOME IN 1980 ($)
HWORK	HOURS OF HOMEWORK PER WEEK
KIDS5	NUMBER OF CHILDREN. 5 YEARS OF AGE
LHWORK	ln(HWORK+1)

- Suppose we wish to estimate the following equations:

$$\begin{aligned} LRATE = {} & \gamma_{21}LHWORK + \beta_{11} + \beta_{21}ED + \beta_{31}ED^2 + \beta_{41}EXPP + \beta_{51}EXPP^2 \\ & + \beta_{61}AGE + \beta_{71}AGE^2 + \beta_{81}OCCW + \beta_{9,1}OCCB \\ & + \beta_{10,1}UNEMPR + \beta_{11,1}REGS + \beta_{12,1}MINOR \\ & + \beta_{13,1}INDUMG + \beta_{14,1}UNION + \beta_{15,1}URB + \varepsilon_1; \end{aligned}$$

$$\begin{aligned} LHWORK = {} & \gamma_{12}LRATE + \beta_{12} + \beta_{22}ED + \beta_{32}ED^2 + \beta_{42}AGE + \beta_{52}AGE^2 \\ & + \beta_{62}REGS + \beta_{72}MINOR + \beta_{8,2}URB + \beta_{9,2}KIDS5 \\ & + \beta_{10,2}LOFINC + \varepsilon_2. \end{aligned}$$

Endogenous: LRATE, LHWORK.

Exogeous: C, ED, ED^2, *EXPP*, *$EXPP^2$*, AGE, AGE^2, *OCCW*, *OCCB*, *UNEMPR*, REGS, MINOR, *INDUMG*, *UNION*, URB, **KIDS5**, **LOFINC**.

<Reduced Form Equation of LHWORK>

Dependent Variable: LHWORK
Method: Least Squares
Date: 05/28/03 Time: 12:47
Sample: 1 923
Included observations: 923

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.901471	0.503746	1.789535	0.0739
ED	0.127457	0.052988	2.405414	0.0164
ED^2	-0.005358	0.002091	-2.562242	0.0106
EXPP	-0.018128	0.010586	-1.712444	0.0872
EXPP^2	0.000278	0.000322	0.862359	0.3887
AGE	0.062152	0.016795	3.700552	0.0002
AGE^2	-0.000629	0.000207	-3.039743	0.0024
OCCW	-0.093918	0.053605	-1.752023	0.0801
OCCB	-0.000177	0.075990	-0.002323	0.9981
UNEMPR	-0.604820	0.813102	-0.743843	0.4572
REGS	-0.043049	0.045960	-0.936664	0.3492
MINOR	-0.039302	0.048575	-0.809084	0.4187
INDUMG	-0.048607	0.053036	-0.916487	0.3597
UNION	0.025167	0.050859	0.494834	0.6208
URB	-0.018390	0.043783	-0.420033	0.6746
KIDS5	**0.104081**	**0.030488**	**3.413849**	**0.0007**
LOFINC	**0.016716**	**0.029498**	**0.566686**	**0.5711**

R-squared	0.065892	Mean dependent var	2.909292
Adjusted R-squared	0.049395	S.D. dependent var	0.574604
S.E. of regression	0.560233	Akaike info criterion	1.697318
Sum squared resid	284.3579	Schwarz criterion	1.786234
Log likelihood	-766.3122	Durbin-Watson stat	1.614776

- Overall significance test.

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	3.994315	(16, 906)	0.0000
Chi-square	63.90904	16	0.0000

- Testing significance of excluded variables.

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	6.065738	(2, 906)	0.0024
Chi-square	12.13148	2	0.0023

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(16)	0.104081	0.030488
C(17)	0.016716	0.029498

<Reduced Form Equation of LRATE>

Dependent Variable: LRATE
Method: Least Squares
Sample: 1 923
Included observations: 923

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.310456	0.273963	1.133203	0.2574
ED	-0.071412	0.028817	-2.478071	0.0134
ED^2	0.005213	0.001137	4.583704	0.0000
EXPP	0.024203	0.005757	4.203963	0.0000
EXPP^2	-0.000383	0.000175	-2.187337	0.0290
AGE	-0.001903	0.009134	-0.208304	0.8350
AGE^2	-2.56E-05	0.000113	-0.227435	0.8201
OCCW	0.133112	0.029153	4.565945	0.0000
OCCB	0.028524	0.041328	0.690194	0.4902
UNEMPR	-0.550904	0.442207	-1.245806	0.2132
REGS	-0.029157	0.024995	-1.166490	0.2437
MINOR	-0.073556	0.026418	-2.784329	0.0055
INDUMG	0.134493	0.028844	4.662849	0.0000
UNION	0.145738	0.027660	5.268952	0.0000
URB	0.143252	0.023811	6.016148	0.0000
KIDS5	0.011349	0.016581	0.684462	0.4939
LOFINC	0.116074	0.016043	7.235354	0.0000

R-squared	0.429705	Mean dependent var	1.662759
Adjusted R-squared	0.419633	S.D. dependent var	0.399943
S.E. of regression	0.304684	Akaike info criterion	0.479162
Sum squared resid	84.10598	Schwarz criterion	0.568078
Log likelihood	-204.1332	Durbin-Watson stat	1.846204

- Overall significance test:

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	42.66570	(16, 906)	0.0000
Chi-square	682.6512	16	0.0000

- Testing significance of excluded variables:

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	23.54156	(7, 906)	0.0000
Chi-square	164.7909	7	0.0000

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(4)	0.024203	0.005757
C(5)	-0.000383	0.000175
C(8)	0.133112	0.029153
C(9)	0.028524	0.041328
C(10)	-0.550904	0.442207
C(13)	0.134493	0.028844
C(14)	0.145738	0.027660

<2SLS for the LRATE Equation>

Dependent Variable: LRATE
Method: Two-Stage Least Squares
Sample: 1 923
Included observations: 923
Instrument list: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LHWORK	0.329712	0.197458	1.669781	0.0953
C	0.748991	0.394655	1.897837	0.0580
ED	-0.113720	0.043547	-2.611428	0.0092
ED^2	0.007207	0.001754	4.108755	0.0000
EXPP	0.027329	0.008340	3.276763	0.0011
EXPP^2	-0.000454	0.000232	-1.961097	0.0502
AGE	-0.008049	0.016488	-0.488163	0.6256
AGE^2	4.15E-05	0.000188	0.221008	0.8251
OCCW	0.175850	0.041902	4.196713	0.0000
OCCB	0.021316	0.052249	0.407967	0.6834
UNEMPR	-0.313988	0.575756	-0.545348	0.5856
REGS	-0.038220	0.031821	-1.201084	0.2300
MINOR	-0.084969	0.033655	-2.524723	0.0117
INDUMG	0.159190	0.037615	4.232079	0.0000
UNION	0.138186	0.035648	3.876416	0.0001
URB	0.173691	0.030052	5.779629	0.0000

R-squared	0.086977	Mean dependent var	1.662759
Adjusted R-squared	0.071877	S.D. dependent var	0.399943
S.E. of regression	0.385302	Sum squared resid	134.6508
F-statistic	26.42415	Durbin-Watson stat	1.747450
Prob(F-statistic)	0.000000		

GENR TSLSE = RESID.

<Specification Test for the LRATE Equation>

Dependent Variable: TSLSE
Method: Least Squares
Sample: 1 923
Included observations: 923

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.735761	0.340764	-2.159149	0.0311
ED	0.000283	0.035844	0.007908	0.9937
ED^2	-0.000227	0.001415	-0.160603	0.8724
EXPP	0.002851	0.007161	0.398103	0.6906
EXPP^2	-2.03E-05	0.000218	-0.093199	0.9258
AGE	-0.014346	0.011361	-1.262715	0.2070
AGE^2	0.000140	0.000140	1.002455	0.3164
OCCW	-0.011772	0.036262	-0.324643	0.7455
OCCB	0.007267	0.051405	0.141361	0.8876
UNEMPR	-0.037500	0.550031	-0.068178	0.9457
REGS	0.023257	0.031090	0.748055	0.4546
MINOR	0.024371	0.032859	0.741675	0.4585
INDUMG	-0.008670	0.035877	-0.241664	0.8091
UNION	-0.000745	0.034404	-0.021656	0.9827
URB	-0.024375	0.029617	-0.823009	0.4107
KIDS5	-0.022968	0.020624	-1.113653	0.2657
LOFINC	0.110562	0.019954	5.540784	0.0000

R-squared	0.033636	Mean dependent var	7.50E-15
Adjusted R-squared	0.016570	S.D. dependent var	0.382155
S.E. of regression	0.378975	Akaike info criterion	0.915554
Sum squared resid	130.1217	Schwarz criterion	1.004470
Log likelihood	-405.5283	F-statistic	1.970941
Durbin-Watson stat	1.756744	Prob(F-statistic)	0.012548

J_T = 0.033636*923= 31.04 > 3.84 = c at 95% of confidence level.

Model is rejected.

Alternative Model?

$$\begin{aligned} LRATE = {} & \gamma_{2,1}LHWORK + \beta_{1,1} + \beta_{2,1}ED + \beta_{3,1}ED^2 + \beta_{4,1}EXPP \\ & + \beta_{5,1}EXPP^2 + \beta_{6,1}AGE + \beta_{7,1}AGE^2 + \beta_{8,1}OCCW \\ & + \beta_{9,1}OCCB + \beta_{10,1}UNEMPR + \beta_{11,1}REGS + \beta_{12,1}MINOR \\ & + \beta_{13,1}INDUMG + \beta_{14,1}UNION + \beta_{15,1}URB + \beta_{16,1}LOFINC \\ & + \varepsilon_1; \\ LHWORK = {} & \gamma_{1,2}LRATE + \beta_{1,2} + \beta_{2,2}ED + \beta_{3,2}ED^2 + \beta_{4,2}AGE + \beta_{5,2}AGE^2 \\ & + \beta_{6,2}REGS + \beta_{7,2}MINOR + \beta_{8,2}URB + \beta_{9,2}KIDS5 \\ & + \beta_{10,2}LOFINC + \varepsilon_2. \end{aligned}$$

Endogenous: LRATE, LHWORK.

Exogeous: C, ED, ED^2, EXPP, $EXPP^2$, AGE, AGE^2, OCCW, OCCB, UNEMPR, REGS, MINOR, INDUMG, UNION, URB, **KIDS5**, LOFINC.

The LRATE equation is exactly identified. So, can't test for the specification of it.

<2SLS for the LRATE Equation with LOFINC>

Dependent Variable: LRATE
Method: Two-Stage Least Squares
Sample: 1 923
Included observations: 923
Instrument list: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LHWORK	0.109040	0.167052	0.652732	0.5141
C	0.212160	0.337028	0.629501	0.5292
ED	-0.085310	0.036360	-2.346228	0.0192
ED^2	0.005798	0.001470	3.944363	0.0001
EXPP	0.026180	0.006918	3.784292	0.0002
EXPP^2	-0.000413	0.000192	-2.150612	0.0318
AGE	-0.008680	0.013672	-0.634845	0.5257
AGE^2	4.30E-05	0.000156	0.276282	0.7824
OCCW	0.143353	0.035086	4.085742	0.0000
OCCB	0.028543	0.043338	0.658613	0.5103
UNEMPR	-0.484955	0.478111	-1.014315	0.3107
REGS	-0.024463	0.026467	-0.924268	0.3556
MINOR	-0.069270	0.028006	-2.473416	0.0136
INDUMG	0.139793	0.031326	4.462502	0.0000
UNION	0.142994	0.029568	4.836080	0.0000
URB	0.145258	0.025283	5.745377	0.0000
LOFINC	0.114251	0.017152	6.661065	0.0000

R-squared	0.372912	Mean dependent var	1.662759
Adjusted R-squared	0.361837	S.D. dependent var	0.399943
S.E. of regression	0.319495	Sum squared resid	92.48170
F-statistic	38.80162	Durbin-Watson stat	1.821699
Prob(F-statistic)	0.000000		

<2SLS for the LHWORK Equation>

Dependent Variable: LHWORK
Method: Two-Stage Least Squares
Sample: 1 923
Included observations: 923
Instrument list: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LRATE	-0.417752	0.142145	-2.938921	0.0034
C	1.023219	0.472991	2.163296	0.0308
ED	0.095646	0.053044	1.803140	0.0717
ED^2	-0.003258	0.002218	-1.469314	0.1421
AGE	0.056944	0.014409	3.952136	0.0001
AGE^2	-0.000611	0.000181	-3.378723	0.0008
REGS	-0.067556	0.044888	-1.504993	0.1327
MINOR	-0.059135	0.047460	-1.246000	0.2131
URB	0.038790	0.047934	0.809224	0.4186
KIDS5	0.113335	0.030070	3.769026	0.0002
LOFINC	0.069324	0.032290	2.146922	0.0321

R-squared	0.073980	Mean dependent var	2.909292
Adjusted R-squared	0.063826	S.D. dependent var	0.574604
S.E. of regression	0.555964	Sum squared resid	281.8956
F-statistic	5.843729	Durbin-Watson stat	1.630243
Prob(F-statistic)	0.000000		

GENR TSLSE = RESID.

<Specification Test for the LHWORK Equation>

Dependent Variable: TSLSE
Method: Least Squares
Sample: 1 923
Included observations: 923

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007946	0.499782	0.015898	0.9873
ED	0.001979	0.052571	0.037648	0.9700
ED^2	7.79E-05	0.002075	0.037542	0.9701
EXPP	-0.008017	0.010503	-0.763331	0.4455
EXPP^2	0.000118	0.000319	0.368304	0.7127
AGE	0.004413	0.016663	0.264810	0.7912
AGE^2	-2.88E-05	0.000205	-0.140328	0.8884
OCCW	-0.038310	0.053183	-0.720330	0.4715
OCCB	0.011739	0.075392	0.155711	0.8763
UNEMPR	-0.834961	0.806703	-1.035029	0.3009
REGS	0.012327	0.045598	0.270328	0.7870
MINOR	-0.010894	0.048193	-0.226055	0.8212
INDUMG	0.007578	0.052618	0.144024	0.8855
UNION	0.086049	0.050459	1.705336	0.0885
URB	0.002664	0.043438	0.061332	0.9511
KIDS5	-0.004513	0.030248	-0.149200	0.8814
LOFINC	-0.004118	0.029266	-0.140712	0.8881

R-squared	0.007080	Mean dependent var	-6.22E-15
Adjusted R-squared	-0.010455	S.D. dependent var	0.552941
S.E. of regression	0.555824	Akaike info criterion	1.681516
Sum squared resid	279.8998	Schwarz criterion	1.770432
Log likelihood	-759.0196	F-statistic	0.403760
Durbin-Watson stat	1.631399	Prob(F-statistic)	0.981932

$J_T = 0.007080*923 = 6.534 < c = 12.59$ (df = 6).

The model is not rejected.

<Testing Exogeneity of LRATE in the LHWORK Equation>

- 2SLS for the LHWORK Equation Using [X, LRATE] as Instruments

Dependent Variable: LHWORK
Method: Two-Stage Least Squares
Sample: 1 923
Included observations: 923
Instrument list: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB
UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC
LRATE

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LRATE	-0.294699	0.055614	-5.299027	0.0000
C	0.994632	0.470753	2.112854	0.0349
ED	0.105456	0.051871	2.033064	0.0423
ED^2	-0.003979	0.002076	-1.917036	0.0555
AGE	0.054478	0.014130	3.855400	0.0001
AGE^2	-0.000583	0.000178	-3.276454	0.0011
REGS	-0.059622	0.043971	-1.355931	0.1755
MINOR	-0.051056	0.046553	-1.096743	0.2730
URB	0.020359	0.043634	0.466573	0.6409
KIDS5	0.112525	0.029977	3.753647	0.0002
LOFINC	0.056628	0.029258	1.935435	0.0532

R-squared	0.078925	Mean dependent var	2.909292
Adjusted R-squared	0.068825	S.D. dependent var	0.574604
S.E. of regression	0.554478	Sum squared resid	280.3904
F-statistic	7.814705	Durbin-Watson stat	1.622138
Prob(F-statistic)	0.000000		

GENR TSLSES = RESID

<Hausman Test for the Exogeneity of LRATE in the LHWORK Equation>

Dependent Variable: TSLSES
Method: Least Squares
Sample: 1 923
Included observations: 923

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.008618	0.498819	-0.017277	0.9862
ED	0.002554	0.052610	0.048555	0.9613
ED^2	4.04E-05	0.002093	0.019304	0.9846
EXPP	-0.011537	0.010577	-1.090793	0.2757
EXPP^2	0.000173	0.000319	0.542686	0.5875
AGE	0.007156	0.016620	0.430563	0.6669
AGE^2	-5.34E-05	0.000205	-0.260755	0.7943
OCCW	-0.057669	0.053650	-1.074904	0.2827
OCCB	0.007591	0.075214	0.100926	0.9196
UNEMPR	-0.754840	0.805268	-0.937378	0.3488
REGS	0.008633	0.045512	0.189686	0.8496
MINOR	-0.008276	0.048271	-0.171444	0.8639
INDUMG	-0.011982	0.053106	-0.225620	0.8215
UNION	0.064854	0.051091	1.269372	0.2046
URB	0.000261	0.044181	0.005913	0.9953
KIDS5	-0.005353	0.030176	-0.177399	0.8592
LOFINC	-0.008303	0.030020	-0.276567	0.7822
LRATE	0.022382	0.060448	0.370270	0.7113

R-squared	0.008094	Mean dependent var	-7.47E-15
Adjusted R-squared	-0.010538	S.D. dependent var	0.551463
S.E. of regression	0.554361	Akaike info criterion	1.677307
Sum squared resid	278.1209	Schwarz criterion	1.771453
Log likelihood	-756.0771	F-statistic	0.434422
Durbin-Watson stat	1.622910	Prob(F-statistic)	0.977495

$J_T^* = 0.008094*923 = 7.467$

$D_T = J_T^* - J_T = 7.467 - 6.534 = 0.933 < 3.84$ (c at 95%).

Do not reject the exogeneity of LRATE.

The system of LRATE and LHWORK may be a triangular system with uncorrelated errors.

Possibly,

$$
\begin{aligned}
LRATE = {} & \beta_{11} + \beta_{21}ED + \beta_{31}ED^2 + \beta_{41}EXPP + \beta_{51}EXPP^2 \\
& + \beta_{61}AGE + \beta_{71}AGE^2 + \beta_{81}OCCW + \beta_{9,1}OCCB \\
& + \beta_{10,1}UNEMPR + \beta_{11,1}REGS + \beta_{12,1}MINOR + \beta_{13,1}INDUMG \\
& + \beta_{14,1}UNION + \beta_{15,1}URB + \beta_{16,1}LOFINC + \varepsilon_1; \\
LHWORK = {} & \gamma_{1,2}LRATE + \beta_{1,2} + \beta_{2,2}ED + \beta_{3,2}ED^2 + \beta_{4,2}AGE + \beta_{5,2}AGE^2 \\
& + \beta_{6,2}REGS + \beta_{7,2}MINOR + \beta_{8,2}URB + \beta_{9,2}KIDS5 \\
& + \beta_{10,2}LOFINC + \varepsilon_2.
\end{aligned}
$$

And $cov(\varepsilon_{t1},\varepsilon_{t2}) = 0$.

<3SLS Estimation>

- Go to **\objects\New Objects..**.
- Choose **System** and click on the **ok** button.
- Then, an empty window will pop up.
- Type the followings on the window:

```
LRATE = C(1)*LHWORK+C(2)+C(3)*ED+C(4)*ED^2+C(5)*EXPP+C(6)*EXPP^2
+C(7)*AGE+C(8)*AGE^2+C(9)*OCCW+C(10)*OCCB+C(11)*UNEMPR+C(12)*REGS
+C(13)*MINOR+C(14)*INDUMG+C(15)*UNION+C(16)*URB

LHWORK = C(17)*LRATE+C(18)+C(19)*ED+C(20)*ED^2+C(21)*AGE+C(22)*AGE^2
+C(23)*REGS+C(24)*MINOR+C(25)*URB+C(26)*KIDS5+C(27)*LOFINC

INST  C ED  ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB UNEMPR REGS MINOR
INDUMG UNION URB KIDS5 LOFINC
```

- Click on **proc\Estimate**.
- Then, you will see the menu for etimation of systems of equations. Choose **Three-Stage Least Squares**.

  - For Two-Step 3SLS, choose **Iterate Coefs**.
  - For Iterative 3SLS, choose **Sequential**.

<Two-Step 3SLS>

System: UNTITLED
Estimation Method: Three-Stage Least Squares
Included observations: 923
Total system (balanced) observations 1846
Linear estimation after one-step weighting matrix

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.316955	0.195506	1.621202	0.1051
C(2)	0.754988	0.388683	1.942424	0.0522
C(3)	-0.110795	0.043099	-2.570726	0.0102
C(4)	0.007136	0.001737	4.107821	0.0000
C(5)	0.025214	0.007765	3.247241	0.0012
C(6)	-0.000419	0.000212	-1.979395	0.0479
C(7)	-0.006258	0.016132	-0.387917	0.6981
C(8)	2.61E-05	0.000184	0.142412	0.8868
C(9)	0.161359	0.039119	4.124792	0.0000
C(10)	0.025640	0.047418	0.540715	0.5888
C(11)	-0.557631	0.524880	-1.062397	0.2882
C(12)	-0.035491	0.031407	-1.130047	0.2586
C(13)	-0.088882	0.033154	-2.680910	0.0074
C(14)	0.158562	0.034772	4.560014	0.0000
C(15)	0.161606	0.032894	4.913005	0.0000
C(16)	0.174344	0.029721	5.865941	0.0000
C(17)	-0.399495	0.141257	-2.828148	0.0047
C(18)	0.608250	0.464156	1.310443	0.1902
C(19)	0.094434	0.052727	1.791010	0.0735
C(20)	-0.003400	0.002204	-1.542229	0.1232
C(21)	0.049287	0.014256	3.457428	0.0006
C(22)	-0.000530	0.000179	-2.956594	0.0032
C(23)	-0.052656	0.044539	-1.182262	0.2373
C(24)	-0.041961	0.047074	-0.891365	0.3729
C(25)	0.021175	0.047542	0.445395	0.6561
C(26)	0.099741	0.029789	3.348190	0.0008
C(27)	0.129521	0.030201	4.288649	0.0000

Determinant residual covariance	0.036334

Equation: LRATE=C(1)*LHWORK+C(2)+C(3)*ED+C(4)*ED^2+C(5)
*EXPP+C(6)*EXPP^2+C(7)*AGE+C(8)*AGE^2+C(9)*OCCW
+C(10)*OCCB+C(11)*UNEMPR+C(12)*REGS+C(13)*MINOR
+C(14)*INDUMG+C(15)*UNION+C(16)*URB
Instruments: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB
UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC
Observations: 923

R-squared	0.105855	Mean dependent var	1.662759
Adjusted R-squared	0.091068	S.D. dependent var	0.399943
S.E. of regression	0.381297	Sum squared resid	131.8667
Durbin-Watson stat	1.749121		

Equation: LHWORK = C(17)*LRATE+C(18)+C(19)*ED+C(20)*ED^2
+C(21)*AGE+C(22)*AGE^2+C(23)*REGS+C(24)*MINOR+C(25)
*URB+C(26)*KIDS5+C(27)*LOFINC
Instruments: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB
UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC
Observations: 923

R-squared	0.070479	Mean dependent var	2.909292
Adjusted R-squared	0.060287	S.D. dependent var	0.574604
S.E. of regression	0.557014	Sum squared resid	282.9615
Durbin-Watson stat	1.631175		

<Iterative 3SLS>

System: UNTITLED
Estimation Method: Iterative Three-Stage Least Squares
Included observations: 923
Total system (balanced) observations 1846
Sequential weighting matrix & coefficient iteration
Convergence achieved after: 4 weight matrices, 5 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.316830	0.193444	1.637837	0.1016
C(2)	0.754961	0.384474	1.963620	0.0497
C(3)	-0.110732	0.042643	-2.596706	0.0095
C(4)	0.007134	0.001719	4.150618	0.0000
C(5)	0.025199	0.007656	3.291489	0.0010
C(6)	-0.000419	0.000209	-2.007779	0.0448
C(7)	-0.006241	0.015952	-0.391260	0.6957
C(8)	2.59E-05	0.000181	0.142954	0.8863
C(9)	0.161069	0.038578	4.175172	0.0000
C(10)	0.025737	0.046679	0.551363	0.5815
C(11)	-0.560987	0.516852	-1.085391	0.2779
C(12)	-0.035466	0.031070	-1.141494	0.2538
C(13)	-0.088955	0.032795	-2.712470	0.0067
C(14)	0.158474	0.034270	4.624252	0.0000
C(15)	0.161884	0.032415	4.994106	0.0000
C(16)	0.174367	0.029406	5.929625	0.0000
C(17)	-0.398776	0.141547	-2.817258	0.0049
C(18)	0.591898	0.464794	1.273463	0.2030
C(19)	0.094386	0.052836	1.786402	0.0742
C(20)	-0.003405	0.002209	-1.541561	0.1234
C(21)	0.048986	0.014281	3.430018	0.0006
C(22)	-0.000527	0.000180	-2.933159	0.0034
C(23)	-0.052069	0.044626	-1.166775	0.2435
C(24)	-0.041284	0.047167	-0.875276	0.3815
C(25)	0.020481	0.047635	0.429955	0.6673
C(26)	0.099205	0.029846	3.323923	0.0009
C(27)	0.131894	0.030159	4.373317	0.0000
Determinant residual covariance		0.036316		

Equation: LRATE=C(1)*LHWORK+C(2)+C(3)*ED+C(4)*ED^2+C(5)
*EXPP+C(6)*EXPP^2+C(7)*AGE+C(8)*AGE^2+C(9)*OCCW
+C(10)*OCCB+C(11)*UNEMPR+C(12)*REGS+C(13)*MINOR
+C(14)*INDUMG+C(15)*UNION+C(16)*URB
Instruments: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB
UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC
Observations: 923

R-squared	0.106021	Mean dependent var	1.662759
Adjusted R-squared	0.091237	S.D. dependent var	0.399943
S.E. of regression	0.381262	Sum squared resid	131.8422
Durbin-Watson stat	1.749099		

Equation: LHWORK = C(17)*LRATE+C(18)+C(19)*ED+C(20)*ED^2
+C(21)*AGE+C(22)*AGE^2+C(23)*REGS+C(24)*MINOR+C(25)
*URB+C(26)*KIDS5+C(27)*LOFINC
Instruments: C ED ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB
UNEMPR REGS MINOR INDUMG UNION URB KIDS5 LOFINC
Observations: 923

R-squared	0.070137	Mean dependent var	2.909292
Adjusted R-squared	0.059941	S.D. dependent var	0.574604
S.E. of regression	0.557117	Sum squared resid	283.0655
Durbin-Watson stat	1.631210		

<FIML without covariance restrictions>

- Go to **\objects\New Objects..**.
- Choose **System** and click on the **ok** button.
- Then, an empty window will pop up.
- Type the followings on the window:

```
LRATE = C(1)*LHWORK+C(2)+C(3)*ED+C(4)*ED^2+C(5)*EXPP+C(6)*EXPP^2
+C(7)*AGE+C(8)*AGE^2+C(9)*OCCW+C(10)*OCCB+C(11)*UNEMPR+C(12)*REGS
+C(13)*MINOR+C(14)*INDUMG+C(15)*UNION+C(16)*URB

LHWORK = C(17)*LRATE+C(18)+C(19)*ED+C(20)*ED^2+C(21)*AGE+C(22)*AGE^2
+C(23)*REGS+C(24)*MINOR+C(25)*URB+C(26)*KIDS5+C(27)*LOFINC

INST  C ED  ED^2 EXPP EXPP^2 AGE AGE^2 OCCW OCCB UNEMPR REGS MINOR
INDUMG UNION URB KIDS5 LOFINC
```

- Click on **proc\Estimate**.
- Then, you will see the menu for etimation of systems of equations. Choose **Full Information Maximum Likelihood**. Choose **Sequential**.
- Click on the **Options button**.
    - For algorithm, choose **Marquardt**.
    - For Max iteration, set 1000.
    - For convergence, set 0.0001.
- Click on **ok**.
- In the **System Estimation** window, click on **ok**.

<FIML Results>

System: UNTITLED
Estimation Method: Full Information Maximum Likelihood (Marquardt)
Included observations: 923
Total system (balanced) observations 1846
Convergence achieved after 1 iteration

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	2.284228	1.592785	1.434110	0.1515
C(2)	-1.696298	2.310763	-0.734086	0.4629
C(3)	-0.342766	0.238420	-1.437657	0.1505
C(4)	0.017219	0.010007	1.720693	0.0853
C(5)	0.049180	0.021833	2.252528	0.0243
C(6)	-0.000794	0.000436	-1.819601	0.0688
C(7)	-0.116266	0.093043	-1.249597	0.2114
C(8)	0.001191	0.001012	1.177279	0.2391
C(9)	0.271869	0.114339	2.377752	0.0174
C(10)	0.061459	0.089243	0.688666	0.4910
C(11)	-0.970614	0.941853	-1.030537	0.3028
C(12)	0.052540	0.137036	0.383401	0.7014
C(13)	-0.052722	0.119191	-0.442333	0.6582
C(14)	0.266327	0.117146	2.273455	0.0230
C(15)	0.276173	0.117061	2.359226	0.0183
C(16)	0.216676	0.115664	1.873317	0.0610
C(17)	-0.428337	0.147982	-2.894527	0.0038
C(18)	1.015221	0.516965	1.963811	0.0496
C(19)	0.089032	0.061366	1.450845	0.1468
C(20)	-0.003048	0.002510	-1.214174	0.2247
C(21)	0.050166	0.015105	3.321149	0.0009
C(22)	-0.000551	0.000194	-2.845307	0.0044
C(23)	-0.052129	0.047714	-1.092527	0.2746
C(24)	-0.042890	0.046995	-0.912660	0.3614
C(25)	0.031039	0.045756	0.678358	0.4975
C(26)	0.026766	0.017521	1.527597	0.1266
C(27)	0.095756	0.032759	2.923039	0.0035
Log Likelihood		-969.0466		
Determinant residual covariance		0.109549		

Equation: LRATE=C(1)*LHWORK+C(2)+C(3)*ED+C(4)*ED^2+C(5)
*EXPP+C(6)*EXPP^2+C(7)*AGE+C(8)*AGE^2+C(9)*OCCW
+C(10)*OCCB+C(11)*UNEMPR+C(12)*REGS+C(13)*MINOR
+C(14)*INDUMG+C(15)*UNION+C(16)*URB
Observations: 923

R-squared	-10.541951	Mean dependent var	1.662759
Adjusted R-squared	-10.732832	S.D. dependent var	0.399943
S.E. of regression	1.369934	Sum squared resid	1702.184
Durbin-Watson stat	1.633455		

Equation: LHWORK = C(17)*LRATE+C(18)+C(19)*ED+C(20)*ED^2
+C(21)*AGE+C(22)*AGE^2+C(23)*REGS+C(24)*MINOR+C(25)
*URB+C(26)*KIDS5+C(27)*LOFINC
Observations: 923

R-squared	0.064066	Mean dependent var	2.909292
Adjusted R-squared	0.053803	S.D. dependent var	0.574604
S.E. of regression	0.558932	Sum squared resid	284.9137
Durbin-Watson stat	1.647929		