3. SEEMINGLY UNRELATED REGRESSIONS (SUR)

[1] Examples

- Demand for some commodities:

  $y_{Nike,t} = x_{Nike,t}'\beta_{Nike} + \varepsilon_{Nike,t}$

  $y_{Reebok,t} = x_{Reebok,t}'\beta_{Reebok} + \varepsilon_{Reebok,t}$;

  where $y_{Nike,t}$ is the quantity demanded for Nike sneakers, $x_{Nike,t}$ is an $1\times k_{Nike}$ vector of regressors such as the unit price of Nike sneakers, prices of other sneakers, income ... , and t indexes time.

- Grunfeld's investment data:
  - I: gross investment ($million)
  - F: market value of firm at the end of previous year.
  - C: value of firm's capital at the end of previous year.
  - $I_{it} = \beta_{1i} + \beta_{2i}F_{it} + \beta_{3i}C_{it} + \varepsilon_{it}$, where i = GM, CH (Chrysler), GE, etc.
  - Notice that although the same regressors are used for each i, the values of the regressors are different across different i.

- CAPM (Capital Asset Pricing Model)
  - $r_{it} - r_{ft}$: excess return on security i over a risk-free security.
  - $r_{mt}-r_{ft}$: excess market return.
  - $r_{it}-r_{ft} = \alpha_i + \beta_i(r_{mt}-r_{ft}) + \varepsilon_{it}$.
  - Notice that the values of regressors are the same for every security.

- VAR (Vector Autoregressions)
  - $g_{CPI,t}$ = growth rate of CPI (inflation rate)
  - $g_{GDP,t}$ = growth rate of norminal GDP
  - $g_{CPI,t} = \alpha_{CPI} + \beta_{CPI} g_{CPI,t-1} + \gamma_{CPI} g_{GDP,t-1} + \varepsilon_{CPI,t}$

    $g_{GDP,t} = \alpha_{GDP} + \beta_{GDP} g_{CPI,t-1} + \gamma_{GDP} g_{GDP,t-1} + \varepsilon_{GDP,t}$
  - Notice that the values of regressors are the same.

- Cobb-Douglas Cost function system
  - Cost function is a function of output and input prices.
  - Assume M inputs (labor, capital, land, material, energy, etc).
  - Cobb-Douglas Cost function:

$$\ln C = \alpha + \beta_y \ln Y + \Sigma_{j=1}^{M} \beta_j \ln(P_j) + \varepsilon_c . \qquad \text{(CD.1)}$$

  - Share functions ($s_j = p_j x_j / C$, where x_j = quantity of input j used)

$$s_j = \frac{\partial \ln C}{\partial \ln P_j} = \beta_j + \varepsilon_j \text{, j = 1, ..., M.} \qquad \text{(CD.2)}$$

  - Observe that $\Sigma_{j=1}^{M} s_j = 0$. This imples that $\Sigma_{j=1}^{M} \beta_j = 0$ and $\Sigma_{j=1}^{M} \varepsilon_j = 0$.

- Translog Cost function System
  - Assume three inputs (j = 1, 2, and 3 index inputs, capital (k), labor (l), and fuel (f), respectively.

- Translog production function:

$$\ln C = \alpha + \Sigma_{j=1}^{3}\beta_j \ln P_j + .5\Sigma_{j=1}^{3}\Sigma_{i=1}^{3}\delta_{ji}\log P_j \log P_i + \Sigma_{j=1}^{3}\gamma_{y,j}\ln P_j \ln Y + \theta_y \ln Y + .5\theta_{yy}(\ln Y)^2 + \varepsilon_c, \quad \text{(T.1)}$$

where $\delta_{ji} = \delta_{ij}$. (See Greene Ch. 14.)

- This a generalization of Cobb-Douglas cost function.
  - If the production function is homothetic, $\gamma_{y,1} = \gamma_{y,2} = \gamma_{y,3} = 0$.
  - If the production function is Cobb-Douglas,

    $\delta_{ji} = 0$, $\gamma_{y,j} = 0$, $\theta_{yy} = 0$, for all j and i.
  - If the production function is Cobb-Douglas and constant returns to scale,

    $\delta_{ji} = 0$, $\gamma_{y,j} = 0$, $\theta_{yy} = 0$, for all j and i; and $\theta_y = 1$.
- Share functions:

$$s_1 = \frac{\partial \ln C}{\partial \ln P_1} = \beta_1 + \Sigma_{i=1}^{3}\delta_{1i}\ln P_i + \gamma_{y,1}\ln Y + \varepsilon_1;$$
$$s_2 = \frac{\partial \ln C}{\partial \ln P_2} = \beta_2 + \Sigma_{i=1}^{3}\delta_{2i}\ln P_i + \gamma_{y,2}\ln Y + \varepsilon_2; \quad \text{(T.2)}$$
$$s_3 = \frac{\partial \ln C}{\partial \ln P_3} = \beta_3 + \Sigma_{i=1}^{3}\delta_{3i}\ln P_i + \gamma_{y,3}\ln Y + \varepsilon_3.$$

- Note that $s_1 + s_2 + s_3 = 1$. Thus,

$$\Sigma_{j=1}^{3}\beta_j = 1; \Sigma_{j=1}^{3}\Sigma_{i}^{3}\delta_{ji} = 0; \Sigma_{j=1}^{3}\gamma_{y,j} = 0; \Sigma_{j=1}^{3}\varepsilon_j = 0.$$

[2] Basic Model

- Model: $y_1 = X_1\beta_1 + \varepsilon_1$;

  $y_2 = X_2\beta_2 + \varepsilon_2$;

  :

  $y_n = X_n\beta_n + \varepsilon_n$,

  where $y_i = T\times 1$, $X_i = T\times k_i$, $\beta_i = k_i\times 1$, $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})' = T\times 1$, and $i = 1, \ldots, n$.

- Assumptions on the model:

  1) SIC hold. (In fact, WIC are sufficient).

  2) $cov(\varepsilon_{it},\varepsilon_{jt}) = E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij}$ for any t, i and j. $\rightarrow$ $cov(\varepsilon_{it},\varepsilon_{it}) = var(\varepsilon_{it}) = \sigma_{ii}$.

  3) $cov(\varepsilon_{it},\varepsilon_{js}) = E(\varepsilon_{it}\varepsilon_{js}) = 0$ for any $t \neq s$.

  4) The ε_{it} are normally distributed. (For simplicity. Not required)

- Implication of nonzero σ_{ij}:

  A unobservable macro shock at time t can influence all of the y_{it}.

- Form of $E(\varepsilon_i\varepsilon_j')$

$$E(\varepsilon_i\varepsilon_j') = E\left(\begin{pmatrix}\varepsilon_{i1}\\ \varepsilon_{i2}\\ \vdots \\ \varepsilon_{iT}\end{pmatrix}\begin{pmatrix}\varepsilon_{j1} & \varepsilon_{j2} & \dots & \varepsilon_{jT}\end{pmatrix}\right) = E\begin{pmatrix}\varepsilon_{i1}\varepsilon_{j1} & \varepsilon_{i1}\varepsilon_{j2} & \dots & \varepsilon_{i1}\varepsilon_{jT}\\ \varepsilon_{i2}\varepsilon_{j1} & \varepsilon_{i2}\varepsilon_{j2} & \dots & \varepsilon_{i2}\varepsilon_{jT}\\ \vdots & \vdots & & \vdots\\ \varepsilon_{iT}\varepsilon_{j1} & \varepsilon_{iT}\varepsilon_{j2} & \dots & \varepsilon_{iT}\varepsilon_{jT}\end{pmatrix}$$

$$= \begin{bmatrix}\sigma_{ij} & 0 & \dots & 0\\ 0 & \sigma_{ij} & \dots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \dots & \sigma_{ij}\end{bmatrix} = \sigma_{ij} I_T$$

- Matrix representation of the model

$$y_1 = X_1\beta_1 + \varepsilon_1$$

$$y_2 = X_2\beta_2 + \varepsilon_2$$

$$\vdots$$

$$y_n = X_n\beta_n + \varepsilon_n$$

$$\rightarrow \quad \begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}_{nT\times 1} = \begin{pmatrix}X_1 & 0_{T\times k_2} & \dots & 0_{T\times k_n}\\ 0_{T\times k_1} & X_2 & \dots & 0_{T\times k_n}\\ \vdots & \vdots & & \vdots\\ 0_{T\times k_1} & 0_{T\times k_2} & \dots & X_n\end{pmatrix}_{nT\times k_*}\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_n\end{pmatrix}_{k_*\times 1} + \begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{pmatrix}_{nT\times 1},$$

where $k_* = \Sigma_i k_i$.

$$\rightarrow \quad y_* = X_*\beta_* + \varepsilon_*$$

Digression to Kronecker Products

- Let $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{p\times q}$. The two matrices do not have to be of the same dimensions. Then,

$$A\otimes B = \begin{pmatrix} a_{11}B & a_{12}B & ... & a_{1n}B \\ a_{21}B & a_{22}B & ... & a_{2n}B \\ : & : & & : \\ a_{m1}B & a_{m2}B & ... & a_{mn}B \end{pmatrix}_{mp\times nq}$$

- Example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}; B = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$$

$$\rightarrow A\otimes B = \begin{pmatrix} 1\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} & 2\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} & 3\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \\ 1\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} & 1\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} & 2\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 2 & 4 & 4 & 6 & 6 \\ 1 & 1 & 2 & 2 & 3 & 3 \\ 2 & 2 & 2 & 2 & 4 & 4 \\ 1 & 1 & 1 & 1 & 2 & 2 \end{pmatrix}$$

- Facts:
  - Let A, B and C be conformable matrices. Then,

    $(A+B)\otimes C = (A\otimes C) + (B\otimes C)$;

    $(A\otimes B)(C\otimes D) = AC\otimes BD$, if AC and BD can be defined;

    $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$, if both A and B are invertible;

    $(A\otimes B)' = (A'\otimes B')$.

End of Digression

- Covariance matrix of ε_* in the system $y_* = X_*\beta_* + \varepsilon_*$
  - Let $\Sigma = [\sigma_{ij}]_{n\times n}$. (Note that $\sigma_{ij} = \sigma_{ji}$)

$$Cov(\varepsilon_*) = E(\varepsilon_*\varepsilon_*') = E\begin{pmatrix} \varepsilon_1\varepsilon_1' & \varepsilon_1\varepsilon_2' & \ldots & \varepsilon_1\varepsilon_n' \\ \varepsilon_2\varepsilon_1' & \varepsilon_2\varepsilon_2' & \ldots & \varepsilon_2\varepsilon_n' \\ \vdots & \vdots & & \vdots \\ \varepsilon_n\varepsilon_1' & \varepsilon_n\varepsilon_2' & \ldots & \varepsilon_n\varepsilon_n' \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_{11}I_T & \sigma_{12}I_T & \ldots & \sigma_{1n}I_T \\ \sigma_{21}I_T & \sigma_{22}I_T & \ldots & \sigma_{2n}I_T \\ \vdots & \vdots & & \vdots \\ \sigma_{n1}I_T & \sigma_{n2}I_T & \ldots & \sigma_{nn}I_T \end{pmatrix} = \Sigma \otimes I_T \equiv \Omega$$

  - If we let $\varepsilon_{\bullet t} = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$, the $\varepsilon_{\bullet t}$ are iid $N(0_{n\times 1}, \Sigma)$.

[3] OLS and GLS

- Two possible OLS

  - OLS on individual i: $\hat{\beta}_i = (X_i'X_i)^{-1}X_i'y_i$, for i = 1, ... , n.

  - OLS on $y_* = X_*\beta_* + \varepsilon_*$:

$$\hat{\beta}_* = \begin{pmatrix} \hat{\beta}_1^* \\ \hat{\beta}_2^* \\ \vdots \\ \hat{\beta}_n^* \end{pmatrix} = \left(X_*'X_*\right)^{-1} X_*'y_*.$$

Proposition 1:

$\hat{\beta}_j^* = \hat{\beta}_j$, for i = 1, 2, ... , n.

<Proof>

$$X_*'X_* = \begin{pmatrix} X_1' & 0 & \cdots & 0 \\ 0 & X_2' & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & X_n' \end{pmatrix} \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & X_n \end{pmatrix}$$

$$= \begin{pmatrix} X_1'X_1 & 0 & \cdots & 0 \\ 0 & X_2'X_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & X_n'X_n \end{pmatrix}.$$

And

$$(X_*'X_*)^{-1} = \begin{pmatrix} (X_1'X_1)^{-1} & 0 & \dots & 0 \\ 0 & (X_2'X_2)^{-1} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & (X_n'X_n)^{-1} \end{pmatrix};$$

$$X_*'y_* = \begin{pmatrix} X_1'y_1 \\ X_2'y_2 \\ \vdots \\ X_n'y_n \end{pmatrix}.$$

Thus,

$$\hat{\beta}_* = \begin{pmatrix} \hat{\beta}_1^* \\ \hat{\beta}_2^* \\ \vdots \\ \hat{\beta}_n^* \end{pmatrix} = (X_*'X_*)^{-1}X_*'y_*$$

$$= \begin{pmatrix} (X_1'X_1)^{-1} & 0 & \dots & 0 \\ 0 & (X_2'X_2)^{-1} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & (X_n'X_n)^{-1} \end{pmatrix} \begin{pmatrix} X_1'y_1 \\ X_2'y_2 \\ \vdots \\ X_n'y_n \end{pmatrix}$$

$$= \begin{pmatrix} (X_1'X_1)^{-1}X_1'y_1 \\ (X_2'X_2)^{-1}X_2'y_2 \\ \vdots \\ (X_n'X_n)^{-1}X_n'y_n \end{pmatrix} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_n \end{pmatrix}$$

- Implication:
  - Note that $Cov(\varepsilon_*) = \Sigma \otimes I_T \neq \sigma_\varepsilon^2 I_{nT}$. So, in general, the OLS on $y_* = X_*\beta_* + \varepsilon_*$ would not be efficient, and so are individual OLS estimators.
  - You can use individual OLS estimators $\hat{\beta}_i$ and $Cov(\hat{\beta}_i) = \sigma_{ii}(X_j'X_j)^{-1}$. But they would be inefficient.

Proposition 2:

Let $\Omega^{-1} = (\Sigma \otimes I_T)^{-1} = \Omega^{-1} = (\Sigma \otimes I_T)^{-1} = \Sigma^{-1} \otimes I_T$. Then, the (infeasible) GLS estimator $\tilde{\beta}_* = (X_*'\Omega^{-1}X_*)^{-1}X_*'\Omega^{-1}y_*$ is unbiased, efficient, consistent and asymptotically efficient

<Proof> Obvious.

- Structure of the GLS estimator
  - Denote $\Sigma^{-1} = [\sigma^{ij}]$.
  - Then,

$$X_*'\Omega^{-1}X_* = \begin{pmatrix} \sigma^{11}X_1'X_1 & \sigma^{12}X_1'X_2 & \cdots & \sigma^{1n}X_1'X_n \\ \sigma^{21}X_2'X_1 & \sigma^{22}X_2'X_2 & \cdots & \sigma^{2n}X_2'X_n \\ \vdots & \vdots & & \vdots \\ \sigma^{n1}X_n'X_1 & \sigma^{n2}X_n'X_2 & \cdots & \sigma^{nn}X_n'X_n \end{pmatrix};$$

$$X_*'\Omega^{-1}y_* = \begin{pmatrix} \Sigma_{j=1}^{n}\sigma^{1j}X_1'y_j \\ \Sigma_{j=1}^{n}\sigma^{2j}X_2'y_j \\ \vdots \\ \Sigma_{j=1}^{n}X_n'y_j \end{pmatrix}.$$

(Justify this by yourself.)

- Efficiency gain of GLS over OLS:

Digression 1

- Suppose

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix};\ A^{-1} = \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix}$$

where A_{11} and A_{22} are invertible square matrices. Then,

$$A^{11} = (A_{11}-A_{12}(A_{22})^{-1}A_{21})^{-1}$$
$$= (A_{11})^{-1} + (A_{11})^{-1}A_{12}(A_{22}-A_{21}(A_{11})^{-1}A_{12})^{-1}A_{21}(A_{11})^{-1}$$
$$A^{12} = - (A_{11})^{-1}A_{12}(A_{22}-A_{21}(A_{11})^{-1}A_{12})^{-1}$$
$$= - (A_{11}-A_{12}(A_{22})^{-1}A_{21})^{-1}A_{12}(A_{22})^{-1}$$
$$A^{21} = - (A_{22}-A_{21}(A_{11})^{-1}A_{12})^{-1}A_{21}(A_{11})^{-1}$$
$$= - (A_{22})^{-1}A_{21}(A_{11}-A_{12}(A_{22})^{-1}A_{21})^{-1}$$
$$A^{22} = (A_{22}-A_{21}(A_{11})^{-1}A_{12})^{-1}$$
$$= (A_{22})^{-1} + (A_{22})^{-1}A_{21}(A_{11}-A_{12}(A_{22})^{-1}A_{21})^{-1}A_{12}(A_{22})^{-1}$$

End of Digression 1

Digression 2 (Review):

- Let $B = [b_{ij}]_{n \times n}$ be a symmetric matrix, and $c = [c_1, \ldots, c_n]'$. Then, the scalar $c'Bc$ is called a quadratic form of B.
- If $c'Bc > (<) 0$ for any nonzero vector c, B is called positive (negative) definite.
- If $c'Bc \geq (\leq) 0$ for any nonzero c, B is called positive (negative) semidefinite.

- Let B be a symmetric and square matrix given by:

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{12} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{1n} & b_{2n} & \dots & b_{nn} \end{pmatrix}.$$

Define the principal minors by:

$$|B_1| = b_{11}; |B_2| = \begin{vmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{vmatrix}; |B_3| = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{vmatrix}; \dots$$

B is positive definite iff *B_1*, *B_2*, ... , *B_n* are all positive. B is negative definite iff *B_1* < 0, *B_2* > 0, *B_3* < 0, ...

End of Digression 2

Digression 3 (Review):

- Let $\tilde{\theta}$ and $\hat{\theta}$ be two p×1 unbiased estimators. Let c = [c_1, ... , c_p]′ be any nonzero vector. Then, $\hat{\theta}$ is said to be efficient relative to $\tilde{\theta}$ iff

$\text{var}(c'\tilde{\theta}) \geq \text{var}(c'\hat{\theta})$.

$$\leftrightarrow c'Cov(\tilde{\theta})c - c'Cov(\hat{\theta})c \geq 0.$$

$$\leftrightarrow c'[Cov(\tilde{\theta}) - Cov(\hat{\theta})]c \geq 0.$$

$\leftrightarrow [Cov(\tilde{\theta}) - Cov(\hat{\theta})]$ is positive semidefinite.

- If $\hat{\theta}$ is more efficient than $\tilde{\theta}$, $\text{var}(\hat{\theta}_j) \leq \text{var}(\tilde{\theta}_j)$, for any j = 1, ... , p. But, the reverse is not true.

End of Digression 3

- Return to Efficiency gain of GLS over OLS:
  - Consider the cases of two equations:

    $y_1 = X_1\beta_1 + \varepsilon_1$

    $y_2 = X_2\beta_2 + \varepsilon_2$,

    with $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$; $\Sigma^{-1} = \begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{12} & \sigma^{22} \end{pmatrix}$.

  - Σ must be positive definite; that is, $\sigma_{11} > 0$ and $\sigma_{11}\sigma_{12}-(\sigma_{12})^2 > 0$.

  - $$Cov(\tilde{\beta}_*) = Cov\begin{pmatrix} \tilde{\beta}_1 \\ \tilde{\beta}_2 \end{pmatrix} = \left(X_*'\Omega^{-1}X_*\right)^{-1} = \begin{pmatrix} \sigma^{11}X_1'X_1 & \sigma^{12}X_1'X_2 \\ \sigma^{12}X_2'X_1 & \sigma^{22}X_2'X_2 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} (say)$$

$$\to Cov(\tilde{\beta}_1) = A^{11};\ Cov(\tilde{\beta}_2) = A^{22}.$$

- Using the fact that $A^{11} = (A_{11} - A_{12}(A_{22})^{-1}A_{21})^{-1}$,

$$Cov(\tilde{\beta}_1) = [\sigma^{11}X_1'X_1 - \{(\sigma^{12})^2/\sigma^{22}\}X_1'X_2(X_2'X_2)^{-1}X_2'X_1]^{-1}$$

$$= [\sigma^{11}X_1'X_1 - \{(\sigma^{12})^2/\sigma^{22}\}X_1'X_1 + \{(\sigma^{12})^2/\sigma^{22}\}X_1'X_1$$

$$- \{(\sigma^{12})^2/\sigma^{22}\}X_1'X_2(X_2'X_2)^{-1}X_2'X_1]^{-1}$$

$$= [\{\sigma^{11}-(\sigma^{12})^2/\sigma^{22}\}X_1'X_1 + \{(\sigma^{12})/\sigma^{22}\}X_1'M(X_2)X_2]^{-1},$$

where $M(X_2) = I_T - X_2(X_2'X_2)^{-1}X_2'$

$$= [(1/\sigma_{11})X_1'X_1 + \{(\sigma^{12})^2/\sigma^{22}\}X_1'M(X_2)X_1]^{-1}$$

$$= \sigma_{11}[X_1'X_1 + \{\sigma_{11}(\sigma^{12})^2/\sigma^{22}\}X_1'M(X_2)X_1]^{-1}$$

$$= \sigma_{11}[X_1'X_1 + \{(\sigma_{12})^2/(\sigma_{11}\sigma_{22}-(\sigma_{12})^2)\}X_1'M(X_2)X_1]^{-1}$$

$$= \sigma_{11}[X_1'X_1 + m^2X_1'M(X_2)X_1]^{-1}, \text{ where } m = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2}}$$

($m^2 \geq 0$).

- Note that $Cov(\hat{\beta}_1) = \sigma_{11}(X_1'X_1)^{-1}$. Thus, $\tilde{\beta}_1$ is more efficient than $\hat{\beta}_1$, because:

$$[Cov(\tilde{\beta}_1)]^{-1} - [Cov(\hat{\beta}_1)]^{-1} = \frac{1}{\sigma_{11}} m^2 X_1' M(X_2) X_1 \text{ is psd.}$$

$\rightarrow$ $Cov(\hat{\beta}_1) - Cov(\tilde{\beta}_1)$ is psd.

$\rightarrow$ $\tilde{\beta}_1$ is more efficient.

Similarly, we can show that $\tilde{\beta}_2$ is more efficient than $\hat{\beta}_2$.

- There are three possible cases in which $\hat{\beta}_1$ is as efficient as $\tilde{\beta}_1$

  ($m^2 X_1' M(X_2) X_1 = 0$):

  1) $\sigma_{12} = 0 \rightarrow m = 0$.

  2) $X_2 = X_1 \rightarrow X_1' M(X_2) X_1 = X_1' M(X_1) X_1 = 0$.

  3) $X_2 = [X_1, W] \rightarrow X_1' M(X_2) X_1 = X_1' 0 = 0$.

  For 1) and 2), $\hat{\beta}_2$ and $\tilde{\beta}_2$ are equally efficient. But for 3), $\tilde{\beta}_2$ is still more efficient than $\hat{\beta}_2$.

- Three Cases in which OLS = GLS:
  - Case I: $\sigma_{ij} = 0$ for any $i \neq j$.
    - $\Sigma = \text{diag}(\sigma_{11}, \ldots, \sigma_{nn}) \rightarrow \Sigma^{-1} = \text{diag}(1/\sigma_{11}, \ldots, 1/\sigma_{nn})$.

$$X_*'\Omega^{-1}X_* = \begin{pmatrix} \sigma^{11}X_1'X_1 & 0 & \cdots & 0 \\ 0 & \sigma^{22}X_2'X_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma^{nn}X_n'X_n \end{pmatrix};$$

$$X_*'\Omega^{-1}y_* = \begin{pmatrix} \sigma^{11}X_1'y_1 \\ \sigma^{22}X_2'y_2 \\ \vdots \\ \sigma^{nn}X_n'y_n \end{pmatrix}.$$

- Thus,

$$\tilde{\beta}_* = (X_*'\Omega^{-1}X_*)^{-1}X_*'\Omega^{-1}y_*$$

$$= \begin{pmatrix} \sigma^{11}X_1'X_1 & 0 & \ldots & 0 \\ 0 & \sigma^{22}X_2'X_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \sigma^{nn}X_n'X_n \end{pmatrix}^{-1} \begin{pmatrix} \sigma^{11}X_1'y_1 \\ \sigma^{22}X_2'y_2 \\ \vdots \\ \sigma^{nn}X_n'y_n \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sigma^{11}}\left(X_1'X_1\right)^{-1} & 0 & \ldots & 0 \\ 0 & \frac{1}{\sigma^{22}}\left(X_2'X_2\right)^{-1} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \frac{1}{\sigma^{nn}}\left(X_n'X_n\right)^{-1} \end{pmatrix} \begin{pmatrix} \sigma^{11}X_1'y_1 \\ \sigma^{22}X_2'y_2 \\ \vdots \\ \sigma^{nn}X_n'y_n \end{pmatrix}$$

$$= \begin{pmatrix} (X_1'X_1)^{-1}X_1'y_1 \\ (X_2'X_2)^{-1}X_2'y_2 \\ \vdots \\ (X_n'X_n)^{-1}X_n'y_n \end{pmatrix} = \hat{\beta}_*$$

- Case II: $X_1 = X_2 = \ldots = X$.
  - This is the case where the values of regressors are the same for all equations.

$$X_* = \begin{pmatrix} X & 0 & \ldots & 0 \\ 0 & X & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & X \end{pmatrix} = \begin{pmatrix} 1X & 0X & \ldots & 0X \\ 0X & 1X & \ldots & 0X \\ \vdots & \vdots & & \vdots \\ 0X & 0X & \ldots & 1X \end{pmatrix} = I_n \otimes X$$

- Then,

$\tilde{\beta}_*$ $= [X_*'\Omega^{-1}X_*]^{-1}X_*'\Omega^{-1}y_*$

$= [(I_n\otimes X)'(\Sigma^{-1}\otimes I_T)(I_n\otimes X)]^{-1}(I_n\otimes X)'(\Sigma^{-1}\otimes I_T)y_*$

$= [\Sigma^{-1}\otimes X'X]^{-1}(\Sigma^{-1}\otimes X')y_*$

$= (\Sigma\otimes(X'X)^{-1})(\Sigma^{-1}\otimes X')y_*$

$= (I_n\otimes(X'X)^{-1}X')y_*$

$$= \begin{pmatrix} (X'X)^{-1}X' & 0 & \dots & 0 \\ 0 & (X'X)^{-1}X' & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & (X'X)^{-1}X' \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} (X'X)^{-1}X'y_1 \\ (X'X)^{-1}X'y_2 \\ \vdots \\ (X'X)^{-1}X'y_n \end{pmatrix}$$

$= \hat{\beta}_*$

- Case III: $X_1, X_2, \dots, X_m \subset X_{m+1}, X_{m+2}, \dots, X_n$

$\hat{\beta}_i = \tilde{\beta}_i$ for i = 1, 2, ... , m. But $\tilde{\beta}_j$ are still more efficient than $\hat{\beta}_j$ for j = m+1, ... , n.

[4] Feasible GLS Estimator

- The GLS estimator defined above is not feasible in the sense that it depends on the unknown covariance matrix Σ.
- Feasible GLS estimator is a GLS estimator obtained by replacing Σ by a consistent estimate of it.
- Feasible GLS estimators cannot be said to be unbiased. But they are consistent and asymptotically equivalent to the infeasible counterpart.

(1) Two-Step Feasible GLS

- Let $s_{ij} = \frac{1}{T}(y_i - X_i\hat{\beta}_i)'(y_j - X_j\hat{\beta}_j) = \frac{1}{T}\Sigma_{t=1}^{T} e_{ti}e_{tj}$; and $S = [s_{ij}]_{n\times n}$.
- It can be shown that plim $s_{ij} = \sigma_{ij}$.
- The two-step feasible GLS estimator is then given by:

$$\tilde{\beta}_* = (X_*'\hat{\Omega}^{-1}X_*)^{-1}X_*'\hat{\Omega}^{-1}y_*,$$

where $\hat{\Omega} = S \otimes I_T$.

(2) Iterative Feasible GLS.

- Using the two-step GLS estimator, recompute S.
- Using this recomputed S, recompute the feasible GLS estimator.
- Repeat this procedure, until the value of FGLS does not change.

(3) Facts:

- The two-step and iterative feasible GLS are asymptotically equivalent to the MLE under the normality assumption about the ε's.
- In fact, iterative feasible GLS = MLE, numerically.

[4] MLE of β and Σ

- Consider a simple regression model:

  $y = X\beta + \varepsilon$; $\varepsilon \sim N(0_{T\times 1},\Omega)$.

  $\rightarrow l_T = -(T/2)\ln(2\pi) - (1/2)\ln[\det(\Omega)] - (1/2)(y-X\beta)'\Omega^{-1}(y-X\beta)$.

- The log-likelihood function of the SUR model:

  $y_* = X_*\beta + \varepsilon_*$; $\varepsilon_* \sim N(0_{nT\times 1},\Sigma\otimes I_T)$.

  $l_* = -(nT/2)\ln(2\pi) - (1/2)\ln[\det(\Sigma\otimes I_T)] - (1/2)(y_*-X_*\beta_*)'(\Sigma^{-1}\otimes I_T)(y_*-X_*\beta_*)$.

  - Let A and B are p×p and q×q matrices. Then,

    $\det(A\otimes B) = [\det(A)]^q[\det(B)]^p$ (Theil, p. 305).

  - $(y_*-X_*\beta_*)'(\Sigma^{-1}\otimes I_T)(y_*-X_*\beta_*) = \Sigma_i\Sigma_j\sigma^{ij}(y_i-X_i\beta_i)'(y_j-X_j\beta_j)$.

    $$\begin{aligned}\rightarrow l_* &= -(nT/2)\ln(2\pi) + (T/2)\ln[\det(\Sigma^{-1})] \\ &\qquad - (1/2)(y_*-X_*\beta_*)'(\Sigma^{-1}\otimes I_T)(y_*-X_*\beta_*) \\ &= -(nT/2)\ln(2\pi) + (T/2)\ln[\det(\Sigma^{-1})] - (1/2)\Sigma_i\Sigma_j\sigma^{ij}(y_i-X_i\beta_i)'(y_j-X_j\beta_j) \\ &= -(nT/2)\ln(2\pi) + (T/2)\ln[\det(\Sigma^{-1})] - (1/2)\Sigma_i\sigma^{ii}(y_i-X_i\beta_i)'(y_i-X_i\beta_i) \\ &\qquad -\Sigma_i\Sigma_{j<i}\sigma^{ij}(y_i-X_i\beta_i)'(y_j-X_j\beta_j)\end{aligned}$$

  - $\dfrac{\partial \ln[\det(A^{-1})]}{\partial a^{ij}} = a_{ij}$ for $A = [a_{ij}]_{n\times n}$ and $A^{-1} = [a^{ij}]$.

  - For a symmetric A, $\dfrac{\partial \ln[\det(A^{-1})]}{\partial a^{ii}} = a_{ii}$; and $\dfrac{\partial \ln[\det(A^{-1})]}{\partial a^{ij}} = 2a_{ij}$ for $j \neq i$.

- Maximize l_* w.r.t. β_* and σ^{ii} and σ^{ij} to get MLE of β_* and Σ.
  - Let A be a symmetric p×p matrix and; x be a p×1 vector. Then,

    $$\frac{\partial x'Ax}{\partial x} = 2Ax.$$

  - FOC:
    - $\frac{\partial l_*}{\partial \beta_*} = X_*'(\Sigma^{-1} \otimes I_T)(y_* - X_*\beta_*) = 0$;
    - $\frac{\partial l_*}{\partial \sigma^{ii}} = \frac{T}{2}\sigma_{ii} - \frac{1}{2}(y_i - X_i\beta_i)'(y_j - X_j\beta_j) = 0$;
    - $\frac{\partial l_*}{\partial \sigma^{ij}} = T\sigma_{ij} - (y_i - X_i\beta_i)'(y_j - X_j\beta_j) = 0$.

- MLE estimators, $\hat{\beta}_*$ and $\hat{\Sigma}$ solves:
  - $\hat{\beta}_* = [X_*'(\hat{\Sigma}^{-1} \otimes I_T)X_*]^{-1}X_*'(\hat{\Sigma}^{-1} \otimes I_T)y_*$;
  - $\hat{\sigma}_{ij} = \frac{1}{T}(y_i - X_i\hat{\beta}_i)'(y_j - X_j\hat{\beta}_j)$.

[5] Testing Hypotheses

- Let $\tilde{\beta}_*$ be a feasible GLS estimator; S be a consistent estimator of Σ; and $C_* = Cov(\tilde{\beta}_*) = [X_*'(S^{-1}\otimes I_T)X_*]^{-1}$.

(1) Testing linear hypotheses:

- H_o: $R\beta_* = r$, where R and r are known matrices with m rows.
- Under H_0, $W_T = (R\tilde{\beta}_* - r)'[RC_*R']^{-1}(R\tilde{\beta}_* - r) \Rightarrow \chi^2(m)$.

- Example 1:

  $\ln(q_{coke,t}) = \beta_{coke,1} + \beta_{coke,2}\ln(p_{coke,t}) + \beta_{coke,3}\ln(p_{pep,t}) + \varepsilon_{coke,t}$;

  $\ln(q_{pep,t}) = \beta_{pep,1} + \beta_{pep,2}\ln(p_{coke,t}) + \beta_{pep,3}\ln(p_{pepe,t}) + \varepsilon_{pep,t}$.

$$\rightarrow \beta_* = \begin{pmatrix} \beta_{coke,1} \\ \beta_{coke,2} \\ \beta_{coke,3} \\ \beta_{pep,1} \\ \beta_{pep,2} \\ \beta_{pep,3} \end{pmatrix}.$$

$\rightarrow H_0$: Own-price elasticities are the same.

$\rightarrow H_0$: $\beta_{coke,2} - \beta_{pep,3} = 0$

$\rightarrow R = (0,1,0,0,0,-1)$, $r = 0$.

- Example 2:
  - Assume that $X_1, \ldots, X_n$ contain the same number (k) of variables.
  - H_0: $\beta_1 = \beta_2 = \ldots = \beta_n$.

$$\to R = \begin{pmatrix} I_k & -I_k & 0_{k\times k} & \ldots & 0_{k\times k} & 0_{k\times k} \\ 0_{k\times k} & I_k & -I_k & \ldots & 0_{k\times k} & 0_{k\times k} \\ : & : & : & & : & : \\ 0_{k\times k} & 0_{k\times k} & 0_{k\times k} & \ldots & I_k & -I_k \end{pmatrix};\ r = 0_{k(n-1)\times 1}.$$

(2) Testing nonlinear hypotheses:

- H_o: $w(\beta_*) = 0$, where w is a m×1 vecor of functions of β_*.
- Under H_0, $W_T = w(\tilde{\beta}_*)'[W(\tilde{\beta}_*)C_*W(\tilde{\beta}_*)']^{-1}w(\tilde{\beta}_*) \Rightarrow \chi^2(m)$, where

$$W(\beta_*) = \frac{\partial w}{\partial \beta_*'}.$$

- Example:

  $\ln(q_{coke,t}) = \beta_{coke,1} + \beta_{coke,2}\ln(p_{coke,t}) + \beta_{coke,3}\ln(p_{pep,t}) + \varepsilon_{coke,t}$;

  $\ln(q_{pep,t}) = \beta_{pep,1} + \beta_{pep,2}\ln(p_{coke,t}) + \beta_{pep,3}\ln(p_{pep,t}) + \varepsilon_{pep,t}$.

$\to H_0$: $\beta_{coke,2}\beta_{pep,3} = 1$

$$\to w(\tilde{\beta}_*) = \tilde{\beta}_{coke,2}\tilde{\beta}_{pep,3} - 1.$$

$$\to W(\tilde{\beta}_*) = (0, \tilde{\beta}_{pep,3}, 0, 0, 0, \tilde{\beta}_{coke,2}).$$

(3) Testing diagonality of Σ

Digression to LM test:

- Let θ be a unknown parameter vector ($p\times1$).
- H_o: $w(\theta) = 0$.
- Let $\hat{\theta}_R$ be the restricted MLE wihch max. l_T subject to $w(\theta) = 0$.
- Let $s_T(\theta) = \frac{\partial l_T}{\partial \theta}$ and $I_T(\theta) = -E\left(\frac{\partial^2 l_T(\theta_o)}{\partial\theta\partial\theta'}\right)$.
- Then, $LM = s_T(\hat{\theta}_R)'[I_T(\hat{\theta}_R)]^{-1}s_T(\hat{\theta}_R)$.

End of Digression

- Note that under H_o: Σ is diagonal, the restricted MLE of β_i's = OLS of β_i's; restricted MLE of $\sigma_{ii} = s_{ii} = (y_i - X_i\hat{\beta}_i)'(y_i - X_i\hat{\beta}_j)/T$; and restricted MLE of $\sigma_{ij} = 0$.

- Breusch and Pagan (1979, Restud):
  - Let $r_{ij} = s_{ij}/(s_{ii}s_{jj})^{1/2}$ (estimated correlation coefficient between e_i and e_j).
  - LM_T for $H_o = T\Sigma_i\Sigma_{j<i}r_{ij}^2 \Rightarrow \chi^2[n(n-1)/2]$.
  - Do not need to compute unrestricted MLE.
  - This statistic is obtained under the assumption of normal errors.
  - Question: Is this statistic still chi-squared even if the errors are not normal?

[6] Autocorrelation:

(1) AR(1)

- $y_{it} = x_{it}'\beta_i + \varepsilon_{it}$; $\varepsilon_{it} = \rho_i\varepsilon_{i,t-1} + v_{it}$; $(v_{1t}, \ldots, v_{nt})' \sim N(0_{n\times 1}, \Sigma)$.
- Use OLS estimates of β_i to estimate ρ_i's as we did in ECN 525.
- Transform each equation using Prais-Winsten or Cochrane-Orcutt methods:

$$\sqrt{1-\hat{\rho}_i^2}\,y_{i1} = \sqrt{1-\hat{\rho}_i^2}\,x_{i1}'\beta_i + v_{i1};$$

$$y_{i2} - \hat{\rho}_i y_{i1} = (x_{i2} - \hat{\rho}_i x_{i1})\beta_i + v_{i2};$$

$$\vdots$$

$$y_{iT} - \hat{\rho}_i y_{i,T-1} = (x_{iT} - \hat{\rho}_i x_{i,T-1})'\beta_i + v_{iT}.$$

- Then, do SUR.

(2) AR(p)

- Similar to the above procedures.

(3) MA(q):

- Procedures are complicated.

[6] Application:

- Use taba15_1a.wf1 (EVIEWS data set) from the CD attached to the textbook.
- Grunfeld's investment data:

    - I: gross investment ($million)
    - F: market value of firm at the end of previous year.
    - CS: value of firm's capital at the end of previous year.
    - $I_{it} = \beta_{1i} + \beta_{2i}F_{it} + \beta_{3i}CS_{it} + \varepsilon_{it}$,
      where i = GM (1), CH (Chrysler, 2), GE (3), WE (Westinghouse, 4), and US (U.S. Steel, 5).

- GLS
    - Read the work file using EVIEWS.
    - 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:

```
i1 = c(1)+c(2)*f1+c(3)*cs1
i2 = c(4)+c(5)*f2+c(6)*cs2
i3 = c(7)+c(8)*f3+c(9)*cs3
i4 = c(10)+c(11)*f4+c(12)*cs4
i5 = c(13)+c(14)*f5+c(15)*cs5
```

    - Click on **proc\Estimate**.
    - Then, you will see the menu for etimation of systems of equations. Choose **Seeingly Unrelated Regression**.

        - For Two-Step GLS, choose **Iterate Coefs**.
        - For Iterative GLS, choose **Sequential**.
        - Do not use "One-Step Coefs" nor "Simultaneous".

<Two-Step GLS>

- Estimation Results:

System: SUR
Estimation Method: Seemingly Unrelated Regression (Marquardt)
Sample: 1935 1954
Included observations: 20
Total system (balanced) observations 100
Linear estimation after one-step weighting matrix

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-162.3641	89.45923	-1.814951	0.0731
C(2)	0.120493	0.021629	5.570868	0.0000
C(3)	0.382746	0.032768	11.68047	0.0000
C(4)	0.504304	11.51283	0.043804	0.9652
C(5)	0.069546	0.016898	4.115732	0.0001
C(6)	0.308545	0.025864	11.92971	0.0000
C(7)	-22.43891	25.51859	-0.879316	0.3817
C(8)	0.037291	0.012263	3.040936	0.0031
C(9)	0.130783	0.022050	5.931272	0.0000
C(10)	1.088877	6.258804	0.173975	0.8623
C(11)	0.057009	0.011362	5.017416	0.0000
C(12)	0.041506	0.041202	1.007400	0.3166
C(13)	85.42325	111.8774	0.763543	0.4473
C(14)	0.101478	0.054784	1.852344	0.0674
C(15)	0.399991	0.127795	3.129956	0.0024

Determinant residual covariance 6.18E+13

Equation: I1 = C(1)+C(2)*F1+C(3)*CS1
Observations: 20

R-squared	0.920742	Mean dependent var	608.0200
Adjusted R-squared	0.911417	S.D. dependent var	309.5746
S.E. of regression	92.13828	Sum squared resid	144320.9
Durbin-Watson stat	0.936490		

Equation: I2 = C(4)+C(5)*F2+C(6)*CS2
Observations: 20

R-squared	0.911862	Mean dependent var	86.12350
Adjusted R-squared	0.901493	S.D. dependent var	42.72556
S.E. of regression	13.40980	Sum squared resid	3056.985
Durbin-Watson stat	1.917509		

Equation: I3 = C(7)+C(8)*F3+C(9)*CS3
Observations: 20

R-squared	0.687636	Mean dependent var	102.2900
Adjusted R-squared	0.650887	S.D. dependent var	48.58450
S.E. of regression	28.70654	Sum squared resid	14009.12
Durbin-Watson stat	0.962757		

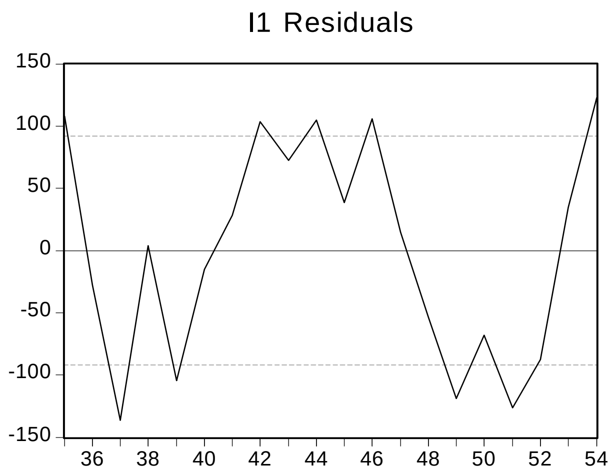
Equation: I4 = C(10)+C(11)*F4+C(12)*CS4
Observations: 20

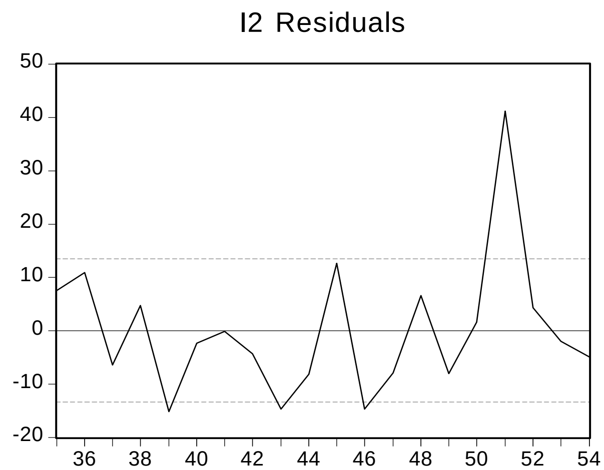
R-squared	0.726429	Mean dependent var	42.89150
Adjusted R-squared	0.694244	S.D. dependent var	19.11019
S.E. of regression	10.56701	Sum squared resid	1898.249
Durbin-Watson stat	1.259005		

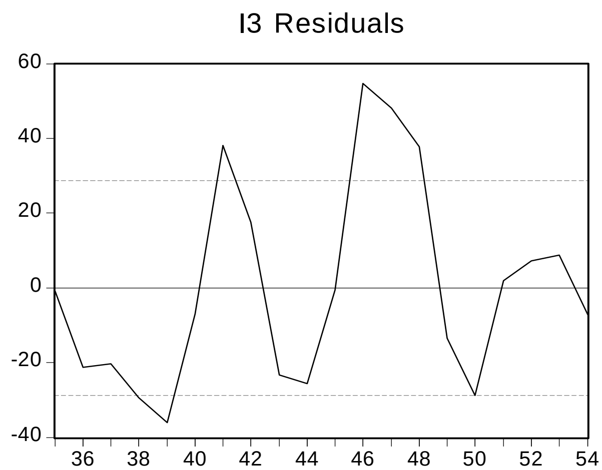
Equation: I5 = C(13)+C(14)*F5+C(15)*CS5
Observations: 20

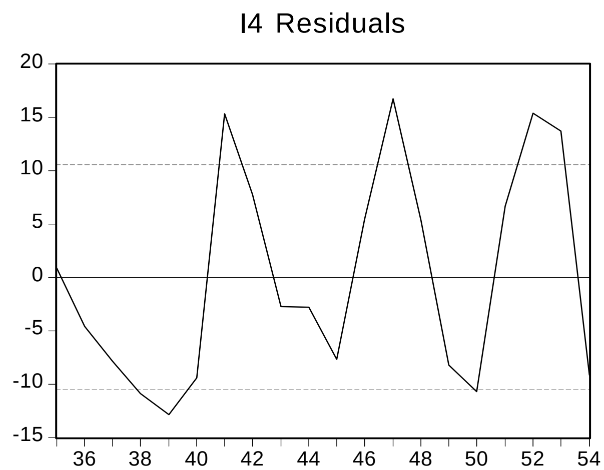
R-squared	0.421959	Mean dependent var	405.4600
Adjusted R-squared	0.353954	S.D. dependent var	129.3519
S.E. of regression	103.9692	Sum squared resid	183763.0
Durbin-Watson stat	1.017982		

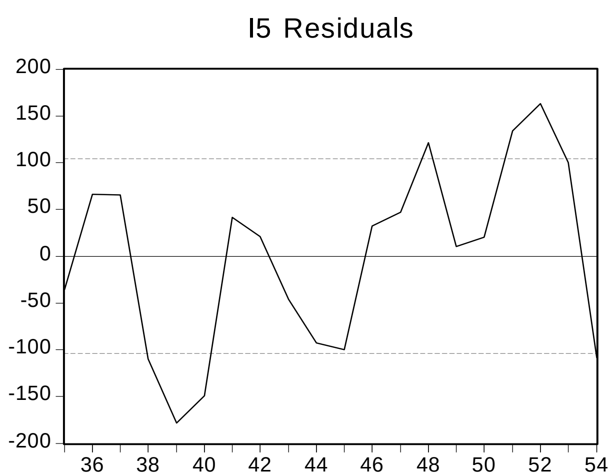
- views/residuals/graphs











- views/residulas/correlation matrix

	I1	I2	I3	I4	I5
I1	1.000000	-0.298702	0.269251	0.156947	-0.329933
I2	-0.298702	1.000000	0.006257	0.138324	0.384018
I3	0.269251	0.006257	1.000000	0.776898	0.482637
I4	0.156947	0.138324	0.776898	1.000000	0.698954
I5	-0.329933	0.384018	0.482637	0.698954	1.000000

- Testing H_o: c(1) = c(4), c(2) = c(5), c(3) = c(6)

  - **views/Wald Coefficient Tests**.
  - Type:

    c(1) = c(4), c(2) = c(5), and c(3) = c(6).

Wald Test:
System: SUR

Null Hypothesis:	C(1)=C(4) C(2)=C(5) C(3)=C(6)		
Chi-square	8.631969	Probability	0.034606

<Iterative GLS>

System: SUR
Estimation Method: Iterative Seemingly Unrelated Regression
(Marquardt)
Date: 08/29/02 Time: 18:57
Sample: 1935 1954
Included observations: 20
Total system (balanced) observations 100
Sequential weighting matrix & coefficient iteration
Convergence achieved after: 13 weight matrices, 14 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-173.0379	84.27963	-2.053141	0.0431
C(2)	0.121953	0.020243	6.024448	0.0000
C(3)	0.389451	0.031852	12.22679	0.0000
C(4)	2.378341	11.63135	0.204477	0.8385
C(5)	0.067451	0.017102	3.943997	0.0002
C(6)	0.305066	0.026067	11.70320	0.0000
C(7)	-16.37654	24.96084	-0.656089	0.5135
C(8)	0.037019	0.011770	3.145121	0.0023
C(9)	0.116954	0.021731	5.381926	0.0000
C(10)	4.488934	6.022064	0.745415	0.4581
C(11)	0.053861	0.010294	5.232286	0.0000
C(12)	0.026469	0.037038	0.714661	0.4768
C(13)	138.0101	94.60801	1.458757	0.1483
C(14)	0.088600	0.045278	1.956796	0.0537
C(15)	0.309302	0.117830	2.624987	0.0103

Determinant residual covariance 5.97E+13

Equation: I1 = C(1)+C(2)*F1+C(3)*CS1
Observations: 20

R-squared	0.919702	Mean dependent var	608.0200
Adjusted R-squared	0.910255	S.D. dependent var	309.5746
S.E. of regression	92.74073	Sum squared resid	146214.3
Durbin-Watson stat	0.936717		

Equation: I2 = C(4)+C(5)*F2+C(6)*CS2
Observations: 20

R-squared	0.910565	Mean dependent var	86.12350
Adjusted R-squared	0.900043	S.D. dependent var	42.72556
S.E. of regression	13.50807	Sum squared resid	3101.956
Durbin-Watson stat	1.885111		

Equation: I3 = C(7)+C(8)*F3+C(9)*CS3			
Observations: 20			
R-squared	0.669021	Mean dependent var	102.2900
Adjusted R-squared	0.630083	S.D. dependent var	48.58450
S.E. of regression	29.54949	Sum squared resid	14843.93
Durbin-Watson stat	0.898029		

Equation: I4 = C(10)+C(11)*F4+C(12)*CS4			
Observations: 20			
R-squared	0.701750	Mean dependent var	42.89150
Adjusted R-squared	0.666661	S.D. dependent var	19.11019
S.E. of regression	11.03336	Sum squared resid	2069.496
Durbin-Watson stat	1.124739		

Equation: I5 = C(13)+C(14)*F5+C(15)*CS5			
Observations: 20			
R-squared	0.390335	Mean dependent var	405.4600
Adjusted R-squared	0.318610	S.D. dependent var	129.3519
S.E. of regression	106.7753	Sum squared resid	193816.3
Durbin-Watson stat	0.967353		

<GLS with AR(1)>

- EVIEWS estimates β's and ρ's jointly using nonlinear GLS method.
- Type:

  i1 = c(1)+c(2)*f1+c(3)*cs1+[ar(1)=c(4)]
  i2 = c(5)+c(6)*f2+c(7)*cs2+[ar(1)=c(8)]
  i3 = c(9)+c(10)*f3+c(11)*cs3+[ar(1)=c(12)]
  i4 = c(13)+c(14)*f4+c(15)*cs4+[ar(1)=c(16)]
  i5 = c(17)+c(18)*f5+c(19)*cs5+[ar(1)=c(20)]

- The following results are from two-step nonlinear GLS (Iterate Coefs).

System: SUR
Estimation Method: Seemingly Unrelated Regression (Marquardt)
Date: 08/29/02 Time: 19:00
Sample: 1936 1954
Included observations: 20
Total system (balanced) observations 95
Iterate coefficients after one-step weighting matrix
Convergence achieved after: 1 weight matrix, 50 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-93.56461	90.03532	-1.039199	0.3021
C(2)	0.097228	0.017643	5.510973	0.0000
C(3)	0.422374	0.046941	8.997976	0.0000
C(4)	0.522621	0.172481	3.030015	0.0034
C(5)	-9.993516	12.79400	-0.781109	0.4372
C(6)	0.080515	0.018744	4.295565	0.0001
C(7)	0.328708	0.022774	14.43375	0.0000
C(8)	-0.217537	0.212416	-1.024109	0.3091
C(9)	-26.10222	32.46218	-0.804081	0.4239
C(10)	0.041268	0.013352	3.090671	0.0028
C(11)	0.120261	0.038715	3.106289	0.0027
C(12)	0.483310	0.179013	2.699858	0.0086
C(13)	4.827474	8.871484	0.544156	0.5879
C(14)	0.048916	0.011829	4.135373	0.0001
C(15)	0.059925	0.052753	1.135943	0.2596
C(16)	0.386245	0.161864	2.386240	0.0195
C(17)	134.4019	193.5550	0.694386	0.4896
C(18)	0.157174	0.040462	3.884510	0.0002
C(19)	-0.024437	0.251078	-0.097330	0.9227
C(20)	0.771540	0.161074	4.789968	0.0000

Determinant residual covariance	1.61E+13

Equation: I1 = C(1)+C(2)*F1+C(3)*CS1+[AR(1)=C(4)]
Observations: 19

R-squared	0.948140	Mean dependent var	623.3053
Adjusted R-squared	0.937768	S.D. dependent var	310.2069
S.E. of regression	77.38514	Sum squared resid	89826.90
Durbin-Watson stat	1.397833		

Equation: I2 = C(5)+C(6)*F2+C(7)*CS2+[AR(1)=C(8)]
Observations: 19

R-squared	0.908871	Mean dependent var	88.53579
Adjusted R-squared	0.890646	S.D. dependent var	42.47399
S.E. of regression	14.04562	Sum squared resid	2959.193
Durbin-Watson stat	1.782442		

Equation: I3 = C(9)+C(10)*F3+C(11)*CS3+[AR(1)=C(12)]
Observations: 19

R-squared	0.737358	Mean dependent var	105.9316
Adjusted R-squared	0.684830	S.D. dependent var	47.02801
S.E. of regression	26.40152	Sum squared resid	10455.61
Durbin-Watson stat	1.219829		

Equation: I4 = C(13)+C(14)*F4+C(15)*CS4+[AR(1)=C(16)]
Observations: 19

R-squared	0.723566	Mean dependent var	44.46842
Adjusted R-squared	0.668279	S.D. dependent var	18.24806
S.E. of regression	10.51001	Sum squared resid	1656.906
Durbin-Watson stat	1.464709		

Equation: I5 = C(17)+C(18)*F5+C(19)*CS5+[AR(1)=C(20)]
Observations: 19

R-squared	0.568551	Mean dependent var	415.7526
Adjusted R-squared	0.482261	S.D. dependent var	124.1974
S.E. of regression	89.36507	Sum squared resid	119791.7
Durbin-Watson stat	1.565597		