1//2013

VI. Simultaneous Equation Models

1. An introduction to some problems associated with simultaneous equation models

2. Basic Notation and alternative representations
- **a. Structural representation**
- **b. Reduced form representation**
- **c. Transfer function representation**

3. Identification
- **a. Identification as a mapping (3 examples)**
- **b. Necessary conditions for a structural equations to be identified**

4. Estimation of reduced form parameters
- **a. Least squares no restrictions (LSNR)**
- **b. Estimators derived from structural estimators**
- **c. SURE–generally the same as LSNR**

5. Estimation of Structural Models
- **a. Notation**
- **b. Single equation methods: OLS, IV, 2SLS, LIML, k-class, GMM, other**
- **c. Simultaneous equation methods: FIML and 3SLS**

6. Statistical inference
- **a. Identification**
- **b. Reduced form coefficients**
- **c. Structural parameters**

7. Exercises

Structural Models

	Structural representation	Reduced form representation
Identification		
Estimation		
Interpreting coefficients		
Statistical inference		
Forecasting		

3/2011

VI. Simultaneous Equation Models

1. An introduction to some problems associated with simultaneous equation models

a. A graphical presentation of the identification problem. See "What Do Statistical Demand Curves Show" by E. J. Working in the QJE, 1927, pp. 212-235.

b. The inconsistency of OLS estimators of the marginal propensity to consume in a simple national income model. See the article by Haavelmo in the March 1947 issue of JASA.

2. Formulation of linear economic models - basic notation and concepts

Linear economic models can be represented as a system of linear equations using matrices as reviewed in this section. The general specification will allow lagged values of the variables.

a. Structural Representation (G equations, G dependent variables)

$$\boxed{BY'_t + B_1Y'_{t-1} + \ldots + B_sY'_{t-s} + \Gamma X'_t + \varepsilon_t' = 0} \tag{2.1}$$

where

$Y_t = (Y_{t1}, \ldots, Y_{tG})$ is a row vector of observations on G endogenous variables at time t;

$X_t = (X_{t1} \ldots, X_{tK})$ is a row vector of observations on K exogenous variables at time t;

B is a G x G matrix of unknown coefficients and is assumed to be nonsingular;

Γ is a G x K matrix of unknown coefficients; and

$\varepsilon_t = (\varepsilon_{t1}\ \varepsilon_{t2} \ldots\ \varepsilon_{tG})$ is a row vector of random disturbances which is distributed N(0,Ω) where Ω is a G x G positive definite matrix.

Let $\boldsymbol{L}$ denote the "lag" operator, $LY_t = Y_{t-1}$ $L^i Y_t = Y_{t-i}$.

Using the lag operator, we can rewrite equation (2.1) as

(2.1)`

$$\boxed{\left(B + B_1 L + \ldots + B_s L^s\right) Y_t' + \Gamma X_t' + \varepsilon_t' = 0}$$

or as

(2.1)``

$$\boxed{B(L) Y_t' + \Gamma X_t' + \varepsilon_t' = 0}$$

where $B(L) = \left(B + B_1 L + \ldots + B_s L^s\right)$ is a matrix polynomial in

"matrix coefficients."

b. **Reduced Form Representation.** If $|B| \neq 0$, then the previously discussed structural representation of an economic model is observationally equivalent to the *reduced form representation*

$$\boxed{Y'_t = \Pi_1 Y'_{t-1} + \ldots \Pi_s Y'_{t-s} + \Pi X'_t + \eta'_t} \qquad (2.2)$$

where

$\Pi_i = -(B)^{-1} B_i$ $(i = 1, 2, \ldots, s)$ are G x G

$\Pi = -(B)^{-1}\Gamma$ is G x K, and

$\eta'_t = (B)^{-1} \varepsilon_t'$ $\sim N(0, \Sigma = (B)^{-1}\Omega(B^{-1})')$.

<u>Note</u>:

- The reduced form representation, expresses the current value of each endogenous variable in terms of predetermined (exogenous and lagged endogenous) variables.
- (2.2) is obtained by multiplying (2.1) by the matrix B^{-1} and solving Y_t' r
- (2.1) is said to represent a dynamic model if at least one of the B_i matrices is not a null matrix.
- Static models contain no lagged endogenous variables ($B_i = 0$ for all i): whereas, dynamic models contain lagged endogenous variables. .
- The matrix Π, in the reduced form representation, contains the impact multipliers $\left(\frac{dY_t}{dX_t} = \Pi = B^{-1}\Gamma\right)$,the instantaneous response of Y to changes i

c. The Final Form or Transfer Function Representation

If the modulus of the roots of $\left|B(z)\right| = 0$ are greater than one, then $B^{-1}(L)$ exists and the transfer function corresponding to (2.1), or (2.1)'' is given by

(2.3) $$Y_t' = -B^{-1}(L)\Gamma X_t' - B^{-1}(L)\varepsilon_t'$$

which expresses Y_t in terms of current and lagged values X_t' . (2.3) is obtained by multiplying (2.1)'' by $B^{-1}(L)$ and so Y_t' ıg for . An inspection of the coefficients of X_t X_{t-1} $X_{t-2},\ldots$ in (2.3) yields the impact and interim multipliers. Two results are quite easily obtained. The impact multiplier can be obtained from the equation

$$\frac{dY_t}{dX_t} = -B^{-1}(L=0)\Gamma = -B^{-1}\Gamma = \Pi \quad ,$$

which is the same result obtained from the reduced form representation. The sum of the impact and interim multipliers gives the long run cumulative multiplier which is equal to $-B^{-1}(L=1)\Gamma = -(B + B_1 + \ldots B_s)^{-1}\Gamma$. T obtained by expressing $B^{-1}(L)$ as a matrix polynomial of infinite order and selecting the appropriate term in the expansion of $-B^{-1}(L)\Gamma X_t'$. The exercises illustrate applications of these results.

VI. Simultaneous Equation Models

3. Identification–logically precedes estimation

a. Identification as a Mapping Problem

Identification is a property of the mapping between the structural parameter space $\boldsymbol{B} = \{(B, B_1, \ldots, B_s, \Gamma, \Omega),$ where B, Ω, and B_i are G x G matrices, Γ is G x K, Ω is positive definite and symmetric, $|B| \neq 0\}$, and the corresponding reduced form parameters space $\boldsymbol{A} = \{(\Pi_1, \ldots, \Pi_s, \Pi, \Sigma) | \Pi_i = -B^{-1}B_i \;\; i = 1, 2, \ldots, s$

$\Pi = -B^{-1}\Gamma, \Sigma = (B)^{-1}\, \Omega(B')^{-1}$ and $(B, B_1, \ldots, B_s, \Omega, \Gamma)\ \varepsilon\ \boldsymbol{B}\}$

A structural economic hypothesis can be represented as a definite proper subset $\boldsymbol{B}_1$ of the structural parameter space ($\boldsymbol{B}$). The set $\boldsymbol{B}_1$ can be described by restrictions on the elements of the matrices B, $B_1, \ldots, B_s$, Γ, Ω such as exclusions of variables from equations, functional dependencies among structural parameters, or restrictions on the variance covariance matrices.

Define $\boldsymbol{A}_1 = \{(\Pi_1, \ldots, \Pi_s, \Pi, \Sigma) | \Pi_i = -B^{-1}B_i,$

$\Pi = -B^{-1}\Gamma, \Sigma = (B)^{-1}\Omega\, (B')^{-1}$ <u>and</u>

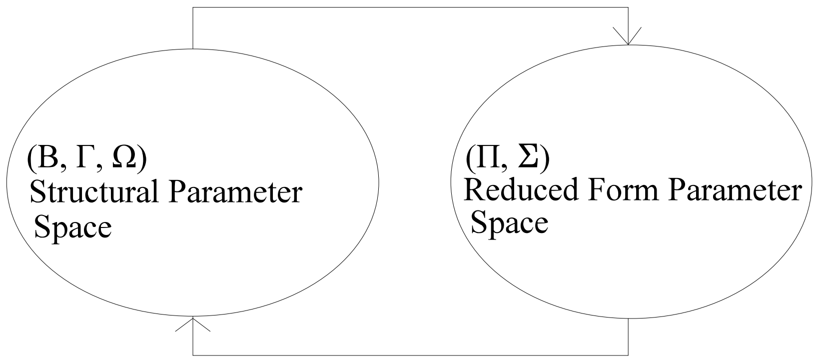
$(B, B_1, \ldots, B_s, \Gamma, \Omega)\ \varepsilon \boldsymbol{B}_1\}$

$\boldsymbol{A}_1$ is then said to be the <u>reduced form hypothesis</u> corresponding to the <u>structural hypothesis</u> $\boldsymbol{B}_1$.

The structural hypothesis (model) $\boldsymbol{B}_1$ is identified if and only if the mapping between $\boldsymbol{B}_1$. and $\boldsymbol{A}_1$ is one-to-one. .

For the static case, the <u>identification problem</u> can be graphically depicted as follows:

$$(1) \quad \Pi = -B^{-1}\Gamma$$
$$\Sigma = B^{-1}\Omega(B^{1})^{-1}$$



(2)Can the equations
$\Pi = -B^{-1}\Gamma$ and
$\Sigma = B^{-1}\Omega(B')^{-1}$ or $\Omega = B\Sigma B^{-1}$
be solved for B, Γ and Ω in terms of Π and Σ?

A **necessary condition** for the mapping between $\boldsymbol{A_1}$ and $\boldsymbol{B_1}$ to be one-to-one is that $\boldsymbol{A_1}$ and $\boldsymbol{B_1}$ have the same dimension (the same number of free parameters).

	Dimension of *B*		**Dimension of *A***
(1) Coefficients			
B:	G^2	Π:	GK
Γ:	$G \bullet K$		
(2) Variance-covariance parameters			
Ω	$G+\frac{G(G-1)}{2}$	Σ	$G+\frac{G(G-1)}{2}$

Total parameters:

$$G^2 + GK + G + \frac{(G-1)(G)}{2} \qquad \text{Total: } GK + G + \frac{(G-1)\ (G)}{2}$$

The structural representation involves G^2 more parameters than the reduced form representation. A necessary condition for a one to one correspondence between $\boldsymbol{B}_1$ and $\boldsymbol{A}_1$ is for $\boldsymbol{B}_1$ to be associated with at least G independent restriction on each structural equation. A common, but not a necessary, approach to this dimensionality problem is to impose restrictions of the following types on the coefficients of each of the structural equations. Ideally, economic theory provides the basis for these restrictions, where exclusions corresponds to specifying some coefficients are zero..

Types of structural restrictions	Number of restrictions
Normalize on one dependent variable	1
# of endogenous variables hypothetically excluded	$G_{\Delta\Delta}$
# of exogenous variables hypothetically excluded	K_2
Total restrictions	$1 + G_{\Delta\Delta} + K_2$

Necessary condition for exclusion restrictions to yield identification[1]:
total number of restrictions $\geq$ G or

- $1 + G_{\Delta\Delta} + K_2 \geq G$, or
- $K_2 \geq G - G_{\Delta\Delta} - 1$, or
- $K_2 \geq G_\Delta - 1$ where G_Δ = the number of included endogenous variables
  $G_\Delta - 1$ = the number of **endogenous regressors**

If $K_2 > G_\Delta - 1$, then $v = K_2 - G_\Delta + 1$ is said to be the number of overidentifying restrictions. In the terminology of the instrumental variables literature v is the number of extra instruments, v = number of instruments - number of endogenous regressors.

In summary, a **necessary condition (sometimes referred to as the order condition)** for each structural equation to be identified, by solving $B\Pi + \Gamma = 0$, based on coefficient restrictions is

- $\boxed{\mathbf{K_2 \geq G_\Delta - 1}}$
- $\boxed{\mathbf{K_2 - G_\Delta + 1 \geq 0}}$
- $\boxed{\textbf{\#instruments} \geq \textbf{\#endogenous regressors}}$

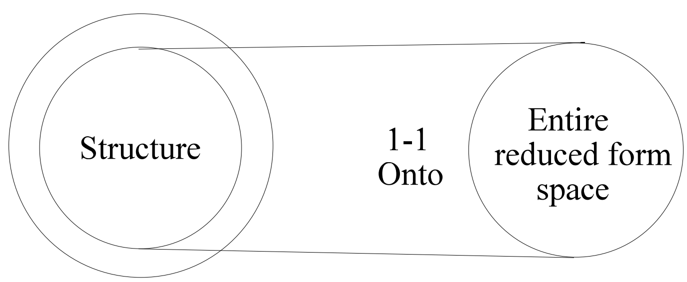
[1]**There are other types of restrictions which can be imposed on the structural coefficients to make the mapping one-to-one. For example, structural vector auto-regressive models (SVAR) use restrictions on the variance covariance matrix to achieve identification.**

Another way of thinking about the condition $K_2 \geq G_\Delta - 1$ is that, for each structural equation, there be at least as many excluded explanatory (predetermined) variables as there are right hand side endogenous variables.

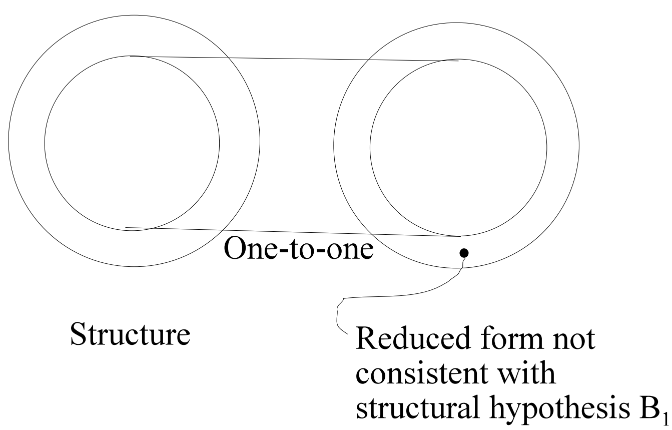
Alternatively, a necessary condition for a structural equation to be identified is that one predetermined variable must be excluded for each endogenous regressor, which is saying that you need at least one instrument for each endogenous regressor, $\nu = K_2 - (G_\Delta - 1) \geq 0$.

If the mapping between $\boldsymbol{B}_1$ and $\boldsymbol{A}_1$ is 1-1 and

(1) $\nu = 0$, the structural equation is said to be <u>exactly identified</u>



(2) If v> 0, those equations are said to be over identified



Note: overidentification is not bad, but it imposes constraints on the reduced form coefficients the validity of which can be tested. using the **estat overid** command in Stata.

(3) If $v < 0$ ($K_2 < G\Delta$ -1), the structural equation is said to be under identified and more than one structure can correspond to the same reduced form. The necessary condition for identification is not satisfied.

Note: $v = K_2 - G_\Delta + 1 \geq 0$ is a necessary condition, but is not sufficient. Sufficient conditions (order conditions) are developed on a case-by-case basis. In the case where identifying restrictions are exclusion restrictions the sufficient condition is $rank[\Pi_{\Delta 2}] = G_\Delta - 1 \text{ or rank}[B_2\ \Gamma_2] = G - 1$ $\Pi_{\Delta 2}$ reduced form coefficients corresponding to the endogenous variables in the structural equation being estimated and the excluded predetermined variables and $[B_2\ \Gamma_2]$ denotes the matrix of structural coefficients in all other structural equations corresponding to the excluded dependent and independent or predetermined variables. Derivations of this condition can by found in older econometrics books, such as those by Johnston or Kmenta or in earlier versions of my class notes.

The following three examples will illustrate each of these situations: :

Example 1: Supply-Demand Model: exactly identified

Structural Model:

Demand: $-Q_t - \beta_{12}P_t + \gamma_{11} + \gamma_{12}Y_t + \varepsilon_{t1} = 0$

Supply: $-Q_t + \beta_{22}P_t + \gamma_{21} - \gamma_{23}FC_t + \varepsilon_{t2} = 0$

$$\begin{bmatrix} -1 & -\beta_{12} \\ -1 & \beta_{22} \end{bmatrix} \begin{bmatrix} Q_t \\ P_t \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & 0 \\ \gamma_{21} & 0 & -\gamma_{23} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ FC_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Where P, Q, Y and FC, respectively, denote price, quantity, income, and factor costs.

Reduced Form Representation:

$$\begin{bmatrix} Q_t \\ P_t \end{bmatrix} = -\begin{bmatrix} -1 & -\beta_{12} \\ -1 & \beta_{22} \end{bmatrix}^{-1} \left\{ \begin{bmatrix} \gamma_{11} & \gamma_{12} & 0 \\ \gamma_{21} & 0 & -\gamma_{23} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ FC_T \end{bmatrix} + \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{bmatrix} \right\}$$

$$= \frac{1}{\beta_{12} + \beta_{22}} \begin{bmatrix} \beta_{22}\gamma_{11} + \beta_{12}\gamma_{21} & \beta_{22}\gamma_{12} & -\beta_{12}\gamma_{23} \\ \gamma_{11} - \gamma_{21} & \gamma_{12} & \gamma_{23} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ FC_t \end{bmatrix} + \begin{bmatrix} \beta_{22}\varepsilon_{t1} + \beta_{12}\varepsilon_{t2} \\ \varepsilon_{t1} - \varepsilon_{t2} \end{bmatrix}$$

$$= \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ FC_t \end{bmatrix} + \begin{bmatrix} \eta_{t1} \\ \eta_{t2} \end{bmatrix}$$

Note: There are six β's and γ's and six π_{ij}'s

Identification: Identification involves solving for the β_{ij}'s and γ_{ij}'s in terms of the π_{ij}'s. There is a one-to-one mapping between the structural and reduced form parameters. **Check each equation for identifiability (rank and order).**

Example 2. Supply-Demand Model: an over identified structural equation

Structural Model:

$$\begin{bmatrix} -1 & -\beta_{12} \\ -1 & \beta_{22} \end{bmatrix} \begin{bmatrix} Q_t \\ P_t \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & 0 \\ \gamma_{21} & 0 & 0 & -\gamma_{24} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ P_{st} \\ FC_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where P_{st} = price of substitutes.

Reduced Form Representation:

$$\begin{bmatrix} Q_t \\ P_t \end{bmatrix} = -\begin{bmatrix} -1 & -\beta_{12} \\ -1 & \beta_{22} \end{bmatrix}^{-1} \left\{ \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & 0 \\ \gamma_{21} & 0 & 0 & -\gamma_{24} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ P_{st} \\ FC_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{bmatrix} \right\}$$

$$= \frac{1}{\beta_{12}+\beta_{22}} \left\{ \begin{bmatrix} \beta_{22}\gamma_{11}+\beta_{12}\gamma_{21} & \beta_{22}\gamma_{12} & \beta_{22}\gamma_{13} & -\beta_{12}\gamma_{24} \\ \gamma_{11}-\gamma_{21} & \gamma_{12} & \gamma_{13} & \gamma_{24} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ P_{st} \\ FC_t \end{bmatrix} \right\} + \begin{bmatrix} \eta_{t1} \\ \eta_{t2} \end{bmatrix}$$

$$= \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \\ P_{st} \\ FC_t \end{bmatrix} + \begin{bmatrix} \eta_{t1} \\ \eta_{t2} \end{bmatrix}$$

Note: **There are seven β_{ij}'s and γ_{ij}'s and eight π_{ij}'s**

Identification. Solve for the β_{ij}'s and γ_{ij}'s in terms of the π_{ij}'s. This example illustrates an overidentified model where restrictions are imposed on the reduced form parameters, e.g.

$$\beta_{22} = \frac{\pi_{12}}{\pi_{22}} = \frac{\pi_{13}}{\pi_{23}}$$

Check the necessary (order) condition for each structural equation. Check the sufficient condition (rank) condition for each structural equation.

Example 3. Supply-Demand Model: an under identified structural equation

Structural Model:

$$\begin{bmatrix} -1 & -\beta_{12} \\ -1 & \beta_{22} \end{bmatrix} \begin{bmatrix} Q_t \\ P_t \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Reduced Form Representation:

$$\begin{bmatrix} Q_t \\ P_t \end{bmatrix} = -\begin{bmatrix} -1 & -\beta_{12} \\ -1 & \beta_{22} \end{bmatrix}^{-1} \left\{ \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{bmatrix} \right\}$$

$$= \left[\frac{1}{\beta_{12} + \beta_{22}}\right] \begin{bmatrix} \beta_{22}\gamma_{11} + \beta_{12}\gamma_{21} & \beta_{22}\gamma_{12} \\ \gamma_{11} - \gamma_{21} & \gamma_{12} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \end{bmatrix} + \begin{bmatrix} \eta_{t1} \\ \eta_{t2} \end{bmatrix}$$

$$= \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} 1 \\ Y_t \end{bmatrix} + \begin{bmatrix} \eta_{t1} \\ \eta_{t2} \end{bmatrix}$$

Note: There are five β_{ij}'s and γ_{ij}'s and and only four π_{ij}'s

Identification: Attempt to solve for the β_{ij}'s and γ_{ij}'s in terms of the π_{ij}'s. In this model β_{22} can be uniquely expressed in terms of the reduced form parameters, but β_{12} can not.

Check the necessary (order) condition for each structural equation. Check the sufficient condition (rank) condition for each structural equation.

VI. Simultaneous Equation Models

4. Estimation of Reduced Form Parameters

Given the Structural model

$YB' + X\Gamma' + \varepsilon = 0$ where $\varepsilon_{t\cdot} \sim N(0, \Omega)$ for all t=1, 2, . . ., N and

Y and X, respectively, have dimension NxG and NxK, then

the associated reduced form can be written in the form

$Y = X\Pi' + \eta$ whe $\eta_{t.}$ $\sim N(0, \Sigma)$.

Thus, each reduced form equation expresses the equilibrium value of an endogenous variable in terms of explanatory variables. There are two main methods of estimating the reduced form coefficients.

a. Reduced form estimators using Least Squares No Restrictions (LSNR)*

$$\hat{\Pi}' = (X'X)^{-1} X'Y$$

where $\Pi = \begin{bmatrix} \pi_1 \\ . \\ . \\ . \\ \pi_G \end{bmatrix}_{GxK}$

and π_i denotes the coefficients in the reduced form equation for Y_{ti}, $\pi_i = (\pi_{i1}, \pi_{i2}, \ldots, \pi_{iK})$.

Let Π_{vec} denote the (GK) x 1 column vector of reduced form coefficents obtained by stacking the Π's in a column, i.e.,

$\Pi'_{vec} = [\pi_1, \pi_2, \ldots, \pi_G]$.

The distribution of $\hat{\Pi}_{vec}$, using LSNR, is given by

$$\hat{\Pi}_{vec} \sim N[\Pi_{vec}, \Sigma \otimes (X'X)^{-1}]$$

and Σ is estimated by

$$\hat{\Sigma} \quad \left(\frac{1}{N}\right)\left(Y - X\hat{\Pi}\right)'\left(Y - X\hat{\Pi}\right)$$

where $\hat{\Sigma}$ has a Wishart distribution. (N-K) is often used as a divisor.

Note:

<u>(1) LSNR Estimators of Π</u> are consistent, unbiased, and normally distributed; however,LSNR doesn't take account of overidentifying restrictions. (Schmidt (1976)) and need not be asymptotically efficient.

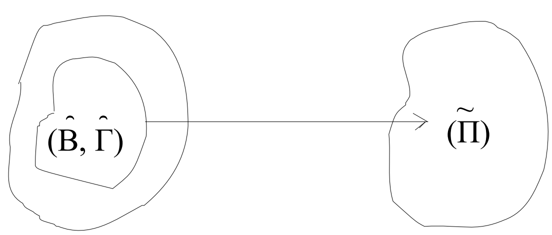
(2) This approach merely amounts to applying least squares to each reduced form equation, regressing each of the dependent variables on the predetermined variables.

Stata commands:
reg y x's

.

b. Reduced form estimators obtained from estimators of (B, Γ).

This approach might be visualized by the following figure:



and involves two steps:

(1) obtain consistent estimates of B, Γ:

$\hat{B}, \hat{\Gamma}$.

(2) calculate $\tilde{\Pi} = -\hat{B}^{-1}\hat{\Gamma}$

If each structural equation is exactly identified, then these estimators will be identical to least squares no restrictions estimators (LSNR). For models containing over identified structural equations, LSNR of II doesn't take account of the overidentifying restrictions imposed on the population reduced form parameters; whereas

$$\tilde{\Pi} = -\hat{B}^{-1}\hat{\Gamma}$$

does. We will comment on the relative merits of LSNR of II and alternative estimators of the form

$$\tilde{\Pi} = -\hat{B}^{-1}\hat{\Gamma}$$

later. The reduced estimators obtained in this way are often named after the method used to estimate (B, Γ). For example, if two stage least squares is used to estimate (B, Γ), then corresponding estimator of Π would be referred to as two stage least squares estimation of Π. The asymptotic distribution of

$$\sqrt{N}\,(\tilde{\Pi}_{vec} - \Pi_{vec})$$

is given by

$$N\,[0, D'VD]$$

where

$$D = (B^{-1})' \otimes \begin{pmatrix} \Pi \\ I_K \end{pmatrix}$$

$$\sqrt{N}\,\mathrm{vec}\left\{ \begin{bmatrix} \hat{\beta} \\ \hat{\Gamma} \end{bmatrix} - \begin{bmatrix} \beta \\ \Gamma \end{bmatrix} \right\} \rightarrow N\,[0, V]$$

See Schmidt (pp. 237-8) for a proof using a different notation.

Notes:

(1) Restricted or derived reduced form estimators of the form

$$\tilde{\Pi} = -\hat{B}^{-1}\hat{\Gamma}$$

will be consistent if $\hat{B}$ and $\hat{\Gamma}$ are consistent estimators

Proof: $\operatorname{plim}\tilde{\Pi} = -\operatorname{plim}\hat{B}^{-1}\hat{\Gamma}$

$$= -(\operatorname{plim}\hat{B}^{-1})^{-1}\operatorname{plim}\hat{\Gamma}$$

$$= -B^{-1}\Gamma = \Pi$$

(2) 3SLS and FIML derived reduced form estimators of π are asymptotically efficient relative to LSNR; however, this doesn't imply that the corresponding estimators will have smaller (exact) variances than LSNR for any given sample size. (McCarthy, IER, 1971, pp. 757-751)

(3) Moments (e.g. means and variances) of restricted reduced form estimators need not exist corresponding to 2SLS. This has implications for forecasts. (Dhrymes, (1973) Econometrica, pp. 119-134)

(4) 2SLS and LIML restricted estimators of the reduced form coefficients are not necessarily asymptotically efficient relative to LSNR.

- LSNR uses all sample information, not all restrictions.
- Restricted estimation uses restrictions--not all sample information.

(5) If all equations are exactly identified, then LSNR, 2SLS, LIML, 3SLS, FIML estimates of the reduced form will be identical.

References: Schmidt, Peter. Econometrics, Marcel Dekker; 1976.

c. SURE–(not so fast, SURE are generally the same as LSNR)

One might be tempted to use SURE to estimate the reduced form coefficients. If the random distrubances in the reduced form equations are correlated across equations, then SURE would yield estimators with smaller variance than LSNR if overidentifying restictions imply that some of the exogenous variables should be deleted for some equations. Otherwise, SURE with identical regressors yields the same estimators as LSNR. This follows from the homework assignment which explored conditions under which SURE and OLS give identical results. One of these conditions is when the separate equations being estimated include identical regressors as is the case in estimating reduced form parameters without restrictions.

VI. Simultaneous Equation Models

5. Estimation of Structural Parameters (Β, Γ, Ω)

This section will introduce alternative methods of estimating structural parameters. These methods can be viewed as being of two types: (1) those in which the structural parameters are estimated one equation at a time and (2) those in which all structural parameters are estimated simultaneously.

a. Review of Notation

$Y = (y_{ti})$ t = 1,2, . . .,N; i = 1,2, . . .,G;

$X = (X_{tj})$ t = 1,2, . . .,N; j = 1,2, . . .,K;

$y_{\cdot i} = (y_{1i}, \ldots, y_{Ni})'$ column vector, i^{th} column of Y matrix;

$y_{t\cdot} = (y_{t1}, \ldots, y_{tG})$ row vector, t^{th} row of Y matrix;

$$Y_1 = = \begin{bmatrix} y_{12} & \cdots & y_{1G_\Delta} \\ y_{22} & \cdots & y_{2G_\Delta} \\ & \vdots & \\ y_{N2} & \cdots & y_{NG_\Delta} \end{bmatrix}$$

is an N x G_Δ-1 matrix of observations of the dependent variables (endogenous regressors) $Y_{t2}, \ldots, Y_{tG_\Delta}$ appearing on the right hand side of the structural equation under consideration;

X_1 = (X_{tj}) t = 1,2, . . .,N, j = 1,2, . . .,K_1; X_1 is the N x K_1 matrix of observations on the exogenous included in the structural equation under consideration.

$$X_2 = \left(X_{ti}\right) \quad t = 1,2,\ldots,N,\ j = K_{1+1},\ldots,K$$

X_2 is the NxK_2 matrix of observations of the exogenous excluded from the particular structural equation under consideration.

$\beta_{i.}$ = coefficients of <u>nonnormalized</u> endogenous variables appearing in the i^{th} structural equation (endogenous regressors); and

$\gamma_{i.}$ = coefficients of exogenous variables which are hypothetically included in the i^{th} structural equation.

Let $y_{t1} = \beta_{12} y_{t2} + \ldots + \beta_{1G_\Delta} y_{tG_\Delta} + \gamma_{11} X_{t1} + \ldots + \gamma_{1K_1} X_{tK_1} + \varepsilon_{t1}$

denote the first structural equation from the system of structural equations,

$$\boxed{\begin{aligned} &\mathrm{B}y'_{t\bullet} + \Gamma X'_{t\bullet} + \varepsilon'_{t\bullet} = 0 \\ &Y\mathrm{B}' + X\Gamma' + \varepsilon = 0 \end{aligned}} \qquad (5.2)$$

Equation (5.1) can also be written in the following form in terms of the matrices

$$y_{11} = [y_{12} \;\cdots\; y_{1G_\Delta}] \begin{bmatrix} \beta_{12} \\ \vdots \\ \beta_{1G_\Delta} \end{bmatrix} + [X_{11} \;\cdots\; X_{1K_1}] \begin{bmatrix} \gamma_{11} \\ \vdots \\ \gamma_{1K_1} \end{bmatrix} + \varepsilon_{11}$$

$$\vdots$$

$$y_{N1} = [y_{N2} \;\cdots\; y_{NG_\Delta}] \begin{bmatrix} \beta_{12} \\ \vdots \\ \beta_{1G_\Delta} \end{bmatrix} + [X_{N1} \;\cdots\; X_{NK_1}] \begin{bmatrix} \gamma_{11} \\ \vdots \\ \gamma_{1K_1} \end{bmatrix} + \varepsilon_{N1}$$

or, equivalently,

$$\boxed{y_{\cdot 1} = Y_1 \beta'_{1\cdot} + X_1 \gamma'_{1\cdot} + \varepsilon_{\cdot 1}} \qquad (5.1)'$$

or

$$\boxed{y_{\cdot 1} = (Y_1 X_1) \begin{pmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \end{pmatrix} + \varepsilon_{\cdot 1}} \qquad (5.1)''$$

This notation will facilitate a formal presentation of estimation procedures. The *concentration parameter* associated with the structural equation (5.1) is defined by

$$\mu^2 = (X_2\Pi_{22})'\left(I - X_1\left(X_1'X_1\right)^{-1} X_1'\right)(X_2\Pi_{22}) / \left(Var\left(\varepsilon_{t1}\right)\right)$$

and provides a measure of the spread of the estimator around the true value and is related to the "strength" of the instruments (excluded exogneous variables). An unbiased estimator of this parameter is provided by

$$\hat{\mu}^2 = K_2\left(\tilde{F} - 1\right)$$

where $\tilde{F}$ denotes the Chow statistic associated testing the hypothesis that the instruments $\left(Z = X_2\right)$ or exogenous variables excluded from structural equation have significant explanatory power for the endogenous regressor(s) in equation (5.1). Generally speaking, the larger the value of the concentration parameter, the smaller the variance (where defined) of the consistent structural estimators and the better the limiting t and F distributions will match the corresponding exact finite sample distributions.

b. Single Equation Techniques: OLS, 2SLS, LIML, k-class, Sawa's combined estimator.

(1) Ordinary Least Squares Estimators (OLS)

$$y_{.1} = Y_1\beta'_{1.} + X_1\gamma'_{1.} + \varepsilon_{.1} \qquad (5.3)$$

$$= [Y_1X_1]\begin{bmatrix}\beta'_{1.}\\ \gamma'_{1.}\end{bmatrix} + \varepsilon_{.1}$$

The corresponding sum of squared errors can be written as

$$SSE = (y_{.1} - Y_1\beta_{1.}' - X_1\gamma_{1.}')'(y_{.1} - Y_1\beta_{1.}' - X_1\gamma_{1.}')$$
$$= (-1, \beta_{1.})Y_\Delta'\left(I - X_1\left(X_1'X_1\right)^{-1} X_1'\right)Y_\Delta\begin{pmatrix}-1\\ \beta_{1.}'\end{pmatrix}$$

where Y_Δ denotes the lG_Δ matrix of observations on the endogenous variables included in the structural equation being estimated.

The ordinary least squares estimators (OLS) of $\beta_{1.}$ and $\gamma_{1.}$ in (5.1) minimize (5.4) and are given by

$$\begin{bmatrix} \beta_{1.}' \\ \gamma_{1.}' \end{bmatrix}_{OLS} = \left[(Y_1 X_1)'(Y_1 X_1)\right]^{-1}(Y_1 X_1)' y_{.1}$$

$$= \begin{bmatrix} Y_1' Y_1 & Y_1' X_1 \\ X_1' Y_1 & X_1' X_1 \end{bmatrix} \begin{bmatrix} Y_1' y_{.1} \\ X_1' y_{.1} \end{bmatrix}$$

The OLS structural coefficient estimators are biased and inconsistent. Note that only observations on the variables appearing in the structural equation being estimated are required to obtain OLS estimators.

STATA commands–you usually don't want to estimate a structural equation with OLS

reg y1 Y1 X1

(2) Indirect Least Squares (ILS)–included for pedagogical purposes only .

For structural equations which are exactly identified we can obtain consistent estimators (biased) by using the technique of ILS which can be thought of as consisting of two steps

(a) Obtain the least squares No restrictions estimates (LSNR) of Π in $Y = X\Pi' + V$, say $\hat{\Pi}$.

(b) Solve $\hat{\Pi}B + \Gamma = 0$ for the corresponding estimates of the structural parameters in the exactly identified structural equations. These estimators are referred to as the ILS estimates of B and Γ.

Example: Consider the simple macro structural model defined by

$$C_t = \alpha + \beta Y_t + \varepsilon_t$$
$$Y_t = C_t + Z_t$$

The corresponding reduced form representation is given by

$$C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} Z_t + \frac{\gamma_t}{1-\beta} = \pi_{11} + \pi_{12} Z_t + \upsilon_{t1}$$
$$Y_t = \frac{\alpha}{1-\beta} + \frac{\beta Z_t}{1-\beta} + \frac{\gamma_t}{1-\beta} = \pi_{11} + \pi_{12} Z_t + \upsilon_{t1}$$

The consumption function is exactly identified. Using Haavelmo's data (1947, JASA, pp. 105-122) we obtain the LSNR estimates of Π:

$$\hat{\Pi} = \begin{bmatrix} \hat{\Pi}_{11} & \hat{\Pi}_{12} \\ \hat{\Pi}_{21} & \hat{\Pi}_{22} \end{bmatrix} = \begin{bmatrix} 344.70 & 2.048 \\ 344.70 & 3.048 \end{bmatrix}$$

and since $\beta = \dfrac{\Pi_{12}}{\Pi_{22}}$,

we obtain $\hat{\beta} = \dfrac{\hat{\pi}_{12}}{\hat{\pi}_{22}} = \dfrac{2.048}{3.048} = .672$ as the ILS of β.

Recall that the OLS estimator of β is .732.

(3) Two Stage Least Squares (2SLS). For overidentified structural equations the

technique of ILS is not applicable and the technique of OLS yields inconsistent estimators. Perhaps the most commonly used technique of consistent estimation in such cases is that of 2SLS. For the case of an exactly identified structural equation the 2SLS estimator is equal to the ILS estimator; hence, ILS need never be performed.

2SLS estimates can readily be obtained using STATA with the command

STATA commands

ivreg y1 (Y1=X1 X2) X1, options or
ivreg y1 (Y1=X2) X1 or
ivregress 2sls (Y1=X1 X2) X1 or
ivregress 2sls y1 (Y1=X2) X1

2SLS estimation was independently developed by H. Theil and R.L. Basmann. 2SLS estimators are asymptotically normally distributed.

(a) Theil's Development of 2SLS.

Theil's approach to two stage least squares involves running two separate regressions; hence, the name two stage least squares. Denote the reduced form representation corresponding to Y_1 (The dependent variables on the right side of the structural equation of interest) by

$$Y_1 = X\Pi_1' + V_1 \quad (NxG_\Delta - 1)$$

where Π_1' is $Kx(G_\Delta - 1)$.

The LSNR estimator of Π_1 is given by

$$\hat{\Pi}'_1 = (X'X)^{-1}X'Y_1.$$

The 2SLS least squares estimators of $\beta_{1.}$ and $\gamma_{1.}$ in

$$y_{\cdot 1} = Y_1\beta_{1.}' + X_1\gamma_{1.}' + \varepsilon_{1.}$$

can be thought of as being obtained by performing the following two-step process:

(1) replace Y_1 in (1.1)' by its least squares estimate

$$\hat{Y}_1 = X\hat{\Pi}_1 \qquad Y_1 = X\hat{\Pi}_1 + \hat{V}_1 \qquad)$$

$$Y_{.1} = \hat{Y}_1\beta'_{1.} + X_1\gamma'_{1.} + \varepsilon^*_{.1} \qquad \varepsilon^* = \varepsilon_{.1} + \hat{V}_1\beta_{1.} \qquad \text{where}$$

$$= (\hat{Y}_1, X_1)\begin{pmatrix}\beta'_{1.}\\ \gamma'_{1.}\end{pmatrix} + \varepsilon^*_{.1}$$

and then

(2) apply least squares to this result to yield, i.e.,

$$\begin{bmatrix}\beta'_{1.}\\ \gamma'_{1.}\end{bmatrix}_{2SLS} = ([\hat{Y}_1, X_1]'[\hat{Y}_1, X_1])^{-1}[\hat{Y}_1, X_1]'y_{.1}$$

$$= \begin{bmatrix}\hat{Y}'_1\hat{Y}_1 & \hat{Y}'_1X_1\\ X'_1\hat{Y}_1 & X'_1X_1\end{bmatrix}^{-1}\begin{bmatrix}\hat{Y}'_1 & y_{.1}\\ X'_1 & y_{.1}\end{bmatrix} \qquad (5.6)$$

Compare this formula to that obtained for OLS, (5.5).

<u>Exercises</u>: Demonstrate that

$$\begin{bmatrix}\beta'_{1.}\\ \gamma];_{1.}\end{bmatrix}_{2SLS} = \begin{bmatrix}Y'_1Y_1 - \hat{V}'_1\hat{V}_1 & Y'_1X_1\\ X'_1Y_1 & X'_1X_1\end{bmatrix}^{-1}\begin{bmatrix}(Y'_1 - \hat{V}'_1)y_{.1}\\ X'_1y_{.1}\end{bmatrix}$$

$$= \begin{bmatrix}Y'_1X(X'X)^{-1}\ X'Y_1 & Y'_1X_1\\ X'_1Y_1 & X'_1X_1\end{bmatrix}^{-1}\begin{bmatrix}Y'_1X(X'X)^{-1}X'y_{.1}\\ X'_1y^{.1}\end{bmatrix}$$

Hint: $(\hat{Y}_1 = Y_1 - \hat{V}_1 = X\hat{\Pi}_1 = X(X'X)^{-1}X'Y_1)$

$X'_1\hat{Y}_1 = (X_1X(X'X)^{-1}X'Y_1) = X'Y_1$

Comment:

The second expression for $\begin{bmatrix} \beta_1' \\ \gamma_{.1}' \end{bmatrix}_{2SLS}$ can be applied directly and doesn't require a two-stage process. It is this approach that is generally used in computer programs.

(b) Basmann's Development

Recall that the sum of squared errors associated with (5.1), can be written as

$$G_1(\beta_{1.}) = (y_{.1} - Y_1\beta_{1.}' - X_1\gamma_{1.}')'(y_{.1} - Y_1\beta_{1.}' - X_1\gamma_{1.}')$$

$$= (-1, \beta_{1.})Y_\Delta'\left(I - X_1\left(X_1'X_1\right)^{-1}X_1'\right)Y_\Delta\begin{pmatrix} -1 \\ \beta_{1.}' \end{pmatrix}$$

OLS estimates of $\beta_{1.}$ can be shown to be equivalent to minimizing $G_1(\beta_{1.})$ over β_1 and using

$$\hat{\gamma}'_{1.} = \hat{\Pi}_{1\Delta}\begin{pmatrix} -1 \\ \beta'_{1.} \end{pmatrix}_{OLS}$$

to estimate $\gamma_{1.}$. Now consider

$$\boxed{G_2(\beta_{1.}) = (-1, \beta_{1.})Y'_\Delta(I - X(X'X)^{-1}X')Y_\Delta\begin{pmatrix} -1 \\ \beta'_{1.} \end{pmatrix}} \qquad (5.7)$$

$G_2(\beta_{1.})$ can be thought of as the sum of squared errors associated with

$$y_{.1} = Y_1\beta'_{1.} + X_1\gamma'_{1.} + X_2\gamma'_{2.} + \varepsilon_{.1} \qquad (5.8)$$

where all of the explanatory variables in the structural model are included. As discussed earlier, we hypothesize that, based on economic theory, $\gamma_{2.} = 0$ and this provides the basis for "identifying" the structural equation under consideration.

If the hypothesis $\gamma_{2.} = 0$ is "true," we might expect that the sum of squared errors associated with (5.1) and (5.8) would be very similar. Equivalently, the reduction in sum of squared error associated with adding "X_2" to the structural equation (5.1) would be small.

It can be shown that 2SLS estimators of $\beta_{1.}$ and $\gamma_{1.}$ are defined by

$$\boxed{\min_{\beta_{1.}}[G_1(\beta_{1.}) - G_2(\beta_{1.})]} \tag{5.9}$$

and $$(\hat{\gamma}_{1.})_{2SLS} = \hat{\pi}_{1\Delta}\begin{pmatrix} -1 \\ \beta'_{1.} \end{pmatrix}_{2SLS} \tag{5.10}$$

Two stage least squares are constructed so as to minimize the reduction in the sum of squared errors which would result from including the variables which are hypothesized to not be included in the structural equation of interest, kind of like minimizing the Chow test associated with the hypothesis $H_o : \gamma_2 = 0$.

$G_1(\beta_1.)$ and $G_2(\beta_1.)$ are sometimes used to define estimators of the variance of ε_{t1}. For this reason 2SLS estimators are also referred to as <u>least variance</u> difference estimators. Note that 2SLS requires observations on <u>all</u> of the <u>independent variables</u> in the <u>structural</u> model and the <u>dependent variables included in the structural equation of interest</u>. OLS only requires observations on the independent variables of the structural equation being estimated.

(c) $E\left(\beta_{2SLS}^{h}\right)$ is only $h \le \nu$ ed for . Hence, the 2SLS estimator doesn have a finite mean in an exactly identified $(\nu = 0)$ structural equation. The bias, where defined, is asymptotically proportional to $\left(\frac{\nu-1}{\mu^2}\right)$ μ^2 /here denotes the concentration parameter defined earlier. Thus, the bias of 2SLS can be substantial if $\left(\frac{\nu-1}{\mu^2}\right)$ is large. An unbiased estimator of the concentration parameter is given by $\hat{\mu}^2 = K_2\left(\tilde{F}-1\right)$ $\tilde{F}$ where = the F-statistic associa with testing the hypothesis that the coefficients of the instruments (X_2) in the estimated reduced form for Y_1 are equal to zero. Thus weak instruments (small F-stat and concentration parameter) can be associated with significant bias.

(d) Recall that in a structural model with one endogenous regressor that the bias of the 2SLS estimator will be less than 10 percent of the bias of the OLS estimator if 10<$\tilde{F}$

(e) Distribution of Two-Stage Least Squares Estimator.

The exact (finite) sample distribution of

$$\begin{pmatrix} \hat{\beta}_{1\cdot} \\ \hat{\gamma}_{1\cdot} \end{pmatrix}_{2SLS}$$

is not normal. In fact, the exact distributions of the 2SLS estimators were not derived until the 1970's. The distribution function involves multiple infinite series. The main findings from these studies will be summarized in another section. However, it is known that the distribution function of the 2SLS structural coefficient estimators approaches a normal distribution function as N grows indefinitely large with <u>mean</u> $(\beta_{1\cdot}, \gamma_{1\cdot})'$ and covariance matrix (Σ_{2SLS}).

$$\boxed{\begin{pmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \end{pmatrix}_{2SLS} \overset{a}{\sim} N\left[\begin{pmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \end{pmatrix} ; \ \Sigma_{2SLS} = \left[\frac{\omega_{11}{}^2}{N} \right] \left(\text{plim}_{N\to\infty} \left\{ \frac{1}{N} \begin{bmatrix} \hat{Y}'_1\hat{Y}_1 & Y'_1X_1 \\ X'_1Y_1 & X'_1X_1 \end{bmatrix} \right\} \right)^{-1} \right]}$$

The covariance matrix of

$$\begin{pmatrix} \hat{\beta}'_{1\cdot} \\ \hat{\gamma}'_{1\cdot} \end{pmatrix}_{2SLS} \qquad \text{in the structural model}$$

$$y_{\cdot 1} = [Y_1, X_1] \begin{bmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \end{bmatrix} + \varepsilon_{\cdot 1}$$

$$= [\hat{Y}_1, X_1] \begin{bmatrix} \beta'_{1\cdot} \\ \hat{\gamma}'_{1\cdot} \end{bmatrix} + \varepsilon^*_{\cdot 1} = Z \begin{bmatrix} \hat{\beta}'_{1\cdot} \\ \hat{\gamma}'_{1\cdot} \end{bmatrix} + \varepsilon'_{\cdot 1}$$

is generally estimated by

$$\hat{\omega}_{11}(Z'Z)^{-1} = \hat{\omega}_{11} \begin{bmatrix} \hat{Y}'_1\hat{Y}_1 & Y'_1X_1 \\ X'_1\hat{Y}_1 & X'_1X_1 \end{bmatrix}^{-1}$$

where

$$\hat{\omega}_{11} = \left(\frac{1}{N - K + \upsilon} \right) \sum_{t-1}^{n} (e_t^2)$$

$$= \left(\frac{1}{N - K + \upsilon} \right) G_1(\hat{\beta}_{1\cdot})$$

is used as an estimator of var $(\varepsilon_{t1}) = \omega_{11}$. This is the formula used in most computer programs. Two other estimators of ω_{11} have been considered:

$$\frac{G_2(\hat{\beta}_{1\cdot})}{N-K+G_\Delta+1} \qquad \text{and} \qquad \frac{G_1(\hat{\beta}_{1\cdot})-G_2(\hat{\beta}_{1\cdot})}{\upsilon}.$$

Dhrymes suggested the second estimator. The choice of the estimator of ω_{11} impacts the degrees of freedom used in performing t-tests. For example, in using Dhrymes' suggested estimator, one should use a t-table with ν degrees of freedom.

(4) Limited Information Maximum Likelihood (LIML)

(a) Development from a likelihood function
First we note that

$$y_{\cdot 1} = Y_1\beta_{1\cdot} + X_1\gamma_{1\cdot} + \varepsilon_{\cdot 1}$$

$$Y_\Delta = (y_{\cdot 1}Y_1) = X\Pi_\Delta + (v_{\cdot 1}, V_1)\ .$$

Consider the solution to the constrained optimization problem:

Maximize the likelihood function of $(y_{\cdot 1}Y_1)$ over the parameters $(\beta_{1\cdot}, \gamma_{1\cdot})$
Subject to the restrictions: rank ($\Pi_{\Delta 2}$ G_Δ = - 1

The solution to this optimization problem yields the LIML estimates of $\beta_{1\cdot}$ and $\gamma_{1\cdot}$
.

(b) Development of the LIML estimators as a least variance ratio (LVR) estimator.

Define $W_{\Delta\Delta} = Y'_\Delta(I - X(X'X)^{-1}X')Y_\Delta$

$$W^*_{\Delta\Delta} = Y'_\Delta(I - X_1(X'_1X_1)^{-1}X'_1)Y_\Delta.$$

As noted previously, estimators of the variance $\omega_{11} = \text{var}(\varepsilon_{t1})$ can be defined in terms of

the quadratic forms $G_1(\beta_{1\cdot}) = (-1, \beta_{1\cdot})\ (W^*_{\Delta\Delta})\ (-1, \beta_{1\cdot})'$ ar

$G_2(\beta_{1\cdot})$

The LIML (LVR) estimator of $\beta_{1\cdot}$, $\tilde{\beta}_{1\cdot}$, can be obtained from

$$\underset{\beta_{1\cdot}}{\text{minimize}} \left[\frac{G_1(\beta_{1\cdot}) - G_2(\beta_{1\cdot})}{G_2(\beta_{1\cdot})} \right]$$

i.e., minimize

$$\frac{b'W^*_{\Delta\Delta}b}{b'W_{\Delta\Delta}b} \text{ s.t. } b'W_{\Delta\Delta}b = 1.$$

Solution:

The corresponding Lagrangian function is defined by

$$L(b,\lambda) = b'W^*_{\Delta\Delta}b + \lambda(1 - b'W_{\Delta\Delta}b)$$

$$\frac{\partial L}{\partial b} = 2(W^*_{\Delta\Delta} - \lambda W_{\Delta\Delta})b = 0.$$

In order to obtain a nontrivial solution for b we require the determinant

$$|W^*_{\Delta\Delta} - \lambda W_{\Delta\Delta}| = 0.$$

Premultiplying the first order condition by b' we note that

$$\hat{\lambda} = \frac{G_1(\tilde{\beta}_{1\cdot})}{G_2(\tilde{\beta}_{1\cdot})} - 1.$$

*The LIML (LVR) estimator of $\beta_{1\cdot}$ corresponds to the characteristic vector, normalized on $y_{\cdot 1}$, which is associated with the smallest characteristic root of $W^*_{\Delta\Delta}$ in the metric $W_{\Delta\Delta}$ (where $|W^*_{\Delta\Delta} - \lambda W_{\Delta\Delta}| = 0$).

STATA Version 12 includes LIML capability using the command:

ivregress liml y1 (Y1=X1 X2) X1 or
ivregress liml y1 (Y1=X2) X1

Note:

(a) $$\tilde{F} = \left(\frac{N-K}{\upsilon}\right)\left(\frac{G_1(\tilde{\beta}_{1.}) - G_2(\tilde{\beta}_{1.})}{G_2(\tilde{\beta}_{1.})}\right) \qquad (5.11)$$

can be used to perform a statistical test of the necessary conditions for identification. F is approximately distributed F(ν, N-K). Note that this statistic is similar in form to the statistic associated with the Chow test.

If the command **estat overid** follows an **ivregress** command, then tests of overidentification are performed. If 2sls is used, Basmann's chi-square test is reported as is Woodridge's robust score test. If LIML is used, Anderson and Rubin's chi square test is reported as is Basmann's F test (above). If GMM is used, Hansen's J statistic and chi square test are given. Statistically significant tests indicate that the exclusion restrictions are not valid, instruments are not valid.

(b) $\tilde{\beta}_{1.}$ and $\tilde{\gamma}_{1.}$ are consistent es $\beta_{1.}$ and $\gamma_{1.}$ but are bi

(c) $\mathrm{Plim}_{N\to\infty}\,(\min\hat{\lambda}) = 1$

(d) $E\left(\hat{\beta}_{LIML}\right)$ is not defined for a fixed normalization (a particular endogenous y appears on the left hand side of the structural equation); hence, the associated pdf has thick tails, Mariano, R. and T. Sawa (1972, The exact finite sample distribution of the LIML estimator in the case of two included endogenous variables, JASA, 67, 159-163). Anderson (2010, The LIML estimator has finite moments!, Journal of Econometrics, 157, 359-361) demonstrates that LIML moments exist if the normalization is $b'\Phi b = 1$.

(e) The interquartile range of LIML can be considerably larger than that for 2SLS because of having thicker tails; however, the pdf of the LIML estimtor is often better "centered" than for 2SLS, particularly for large values of $\left(\frac{\nu-1}{\mu^2}\right)$, weak instruments and large number of overidentifying restrictions. The there are tradeoffs between bias and variance (MSE). Hansen, Heaton, and Yaron (1996, Finite sample properties of some alternative GMM estimators, *Journal of Business and Economic Statistics*, 14(3), 262-280) use a continuous updating (CUE) of a GMM-like generalization of LIML to address this problem. Hausmann, Menzel, Lewis, and Newey (2007, A reduced bias GMM-like estimator with reduced estimator dispersion, MIT manuscript) modify the CUE to solve the no moments/large dispersion problem.

(f) Anderson, Kunitomo, and Matsushita (2010, "On the asymptotic optimality of the LIML estimator with possibly many instruments," *Journal of Econometrics,* 157, 191-204) remind us that while 2SLS and LIML are asymptotically equivalent with large sample sizes, they are quite different with large K_2 (Many instruments) and argue that LIML may be an attractive option over the semiparametric methods of GMM and EL in situations with many instruments or many weak instruments.

(5) General k-Class Estimators.

The k-class estimators of $\tilde{\beta}_{1\cdot}$ and $\tilde{\gamma}_{1\cdot}$ are defined by

$$\begin{bmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \end{bmatrix}_k = \begin{bmatrix} Y'_1 Y_1 - k\hat{V}'_1 \hat{V}_1 & Y'_1 X_1 \\ X'_1 Y_1 & X'_1 X_1 \end{bmatrix}^{-1} \begin{bmatrix} (Y'_1 - k\hat{V}'_1) y_{\cdot 1} \\ X'_1 Y_{\cdot 1} \end{bmatrix}$$

where

$$\hat{V}_1 = Y_1 - X\hat{\Pi}_1 = Y_1 - X(X'X)^{-1}X'Y_{1\cdot}$$

The k-class estimators can be obtained by

$$\min_{\beta_{1\cdot}} [G_1(\beta_{1\cdot}) - k\, G_2(\beta_{1\cdot})]$$

with

$$(\hat{\gamma}_{1\cdot})_k = \hat{\pi}'_{1\Delta} \begin{bmatrix} -1 \\ \beta'_{1\cdot} \end{bmatrix}_k$$

, McDonald [1977, Econometrica]

It should be noted that k = 0, 1, respectively, yield OLS and 2SLS estimators. If k is selected to be λ from the previous section, then LIML estimates are obtained.

The k-class estimator of $\beta_{1\cdot}$ and $\gamma_{1\cdot}$ will be consistent if and only if plim k=1. Note that it then follows that 2SLS and LIML estimates will be consistent; whereas, OLS estimates are inconsistent.

Stata will perform k-class estimators with the command
ivreg2 y1 (Y1=X1 X2 or X2) X1, kclass(#) fuller(#) liml

Exercises and Notes:

1. That if we select $Z = [Y_1 - k\hat{V}_1, X_1]$, then the associated instrumental variables estimator corresponding to Z is the k-class estimator.

Hint: Solve $$Z'y_{.1} = Z'[Y_1, X_1] \begin{bmatrix} \beta'_{1.} \\ \gamma'_{1.} \end{bmatrix}_k$$

for $(\beta_{1.}, \gamma_{1.})'$ and manipulate the derived expression to obtain the desired result.

2. Nagar demonstrates that if we select k to be $1 + \frac{v-1}{N}$, the k-class estimator is "almost" unbiased. Demonstrate that this estimator is consistent. It can be shown that the expected value of k-class estimators (integer moments) are not defined if $k > 1$.

3. Zellner (Journal of Econometrics), Estimation of Functions of Population Means and Regression Coefficients Including Structural Coefficients (MELO), 8(1978), 127-158 (esp. p. 141).

   Selecting k to be

$$k = 1 - \frac{K}{N - K - G_\Delta + 1}$$

   yields the minimum expected loss Bayesian estimator when based on a diffuse prior on (Π, Σ). The random disturbances in the reduced form are assumed to be normally distributed. Note that this estimator is consistent (plim $k = 1$).

4. The k-class estimators ($0 \le k \le 1$) of $\beta_{1.}$ and $\gamma_{1.}$ in (1.1) can also be obtained using least squares algorithms by regressing

$$y_{.1} - a\hat{V}_1 \text{ on } (\hat{Y}_1 - a\hat{V}_1) \text{ and } X_1$$

$$\text{where } a = 1 - \sqrt{1-k},\ \hat{v}_1 = y_{.1} - \hat{y}_{.1} \text{ and}$$

$$\hat{V}_1 = Y_1 - \hat{Y}_1$$

Thus least squares computer programs can be used to obtain k-class estimators, McDonald and Maynes (1980, Journal of Statistical Computation and Simulation) .

(6) Sawa's combined estimator, Journal of Econometrics (1973):

$$\begin{pmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \end{pmatrix}_{comb} = \left(1 + \frac{\nu - 1}{N - K}\right) \begin{pmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \end{pmatrix}_{2SLS} - \left(\frac{\nu - 1}{N - K}\right) \begin{pmatrix} \beta_{1\cdot} \\ \gamma'_{1\cdot} \end{pmatrix}_{OLS}$$

This estimator is "almost" unbiased and consistent. For equations having a single overidentifying restriction ($\nu = 1$), Sawa's combined estimator is merely the two stage least squares estimator.

(7) GMM and Instrumental variables have been discussed earlier and can be formulated to include the previous estimators as special cases. Depending on our schedule, we may spend more time on them. As noted earlier , there are many variations of GMM. The basic GMM form is to minimize a quadratic form $(\varepsilon' Z)\hat{Q}^{-1}(Z'\varepsilon)$ over the unknow $\hat{Q}$ arameters where denotes an estimate weighting matrix. The optimal weighting matrix is $\hat{Q} = Var(Z'\varepsilon)$ $\hat{Q}$. If ac with each iteration, the estimators are referred to as continuous updating estimators (CUE) and reduce the bias of GMM. Hansen, McDonald, and Newey (2010) consider a variation of the IV (GMM) estimators previously described where the objective function is $(\rho(\varepsilon)'Z)\hat{Q}^{-1}(Z'\rho(\varepsilon))$ $\hat{Q}=Var(Z'\rho(\varepsilon))$ ere $\rho(\varepsilon)$ a vector of the derivatives of the log pdf evaluated at the vector of the estimated disturbances. If the errors are normal then this nolinear general instrumental variables (NLIV) estimator includes regular GMM as a special case and has the potential to provide improved estimators for non normal errors distributions.

c. Simultaneous Equation Estimation Methods: FIML and 3SLS

(1) Full Information Maximum Likelihood (FIML) Estimation.

$$BY_{t\cdot} + \Gamma X_{t\cdot} + \varepsilon_{t\cdot} = 0$$

$$\varepsilon_{t\cdot} \sim N(O,\Omega).$$

The associated likelihood function is given by

$$L(Y;B,\Gamma,\Omega) = \frac{e^{-\frac{1}{2}\Sigma_t (BY_{t\cdot} + \Gamma X_{t\cdot})'\Omega^{-1}(BY_{t\cdot} + \Gamma X_t)}}{(2\pi)^{-\frac{NG}{2}} |\Omega|^{\frac{N}{2}} |B|^{N}}$$

<u>Solution</u>: Maximize L( ), or ℓ( ) = ln L( ) with respect to B, Γ, Ω subject to any restrictions on the parameters, i.e., solve

$$\frac{\partial L}{\partial B} = 0$$

$$\frac{\partial L}{\partial \Gamma} = 0$$

$$\frac{\partial L}{\partial \Omega} = 0$$

The first order conditions are a system of nonlinear equations in the unknown structural parameters. These estimators are asymptotically normal and efficient under quite general conditions.

The 2011 release of Stata allows FIML estimation of some structural models.

(2) Three-Stage Least Squares. References (3SLS)

Let the hth structural equation in the system of structural equations

$$YB' + X\Gamma' + \varepsilon = 0$$

be denoted

$$y_{\cdot h} = Y_1^{(h)}\beta'_{h\cdot} + X_1^{(h)}\gamma'_{h\cdot} = \varepsilon_{\cdot h} \qquad (5.13)$$

where $Y_1^{(h)}$ denotes the endogenous variables included on r.h.s. of hth structural equation and $X_1^{(h)}$ denotes the exogenous variables included in hth structural equation

Most econometric packages can perform 3SLS estimation. The format for STATA 3SLS estimation is:

. . .

STATA Commands

reg3 (depvar1 rhs_varlist1) (depvar2 rhs_varlist2) ...(depvarG rhs_varlistG), endog(list of endogenous variables)

reg3 (depvar1 rhs_varlist1) (depvar2 rhs_varlist2) ...(depvarG rhs_varlistG), endog(list of endogenous variables) ireg3 "iterates until estimates converge"

2SLS adjusts for endogenous regressors, but does not take account of possible correlation between the error terms in the different equations. 3SLS takes account of both the endogeniety problem and possible correlation between the structural random disturbances. Hence, the 3SLS estimators are asymptotically efficient.

Development of 3SLS estimators: estimation equations and motivation.

Equation (5.13) can be rewritten as

(5.13)'

$$y_{\cdot h} = [Y_1^{(h)} X_1^{(h)}] \begin{bmatrix} \beta'_{h\cdot} \\ \gamma'_{h\cdot} \end{bmatrix} + \varepsilon_{\cdot h}$$

or

(5.13)''

$$y_{\cdot h} = Z_h\delta_h + \varepsilon_{\cdot h} \quad h = 1, 2, \ldots G \text{ where } \varepsilon_{\cdot h} \sim N[0, \omega_{hh} I].$$

The G-equations given in (5.13)" can be rewritten in matrix form as

(5.14)

$$y^* = \begin{bmatrix} y_{\cdot 1} \\ y_{\cdot 2} \\ \vdots \\ y_{\cdot G} \end{bmatrix} = \begin{bmatrix} Z_1 & 0 & 0 & \dots & 0 \\ 0 & Z_2 & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & Z_G \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix} + \begin{bmatrix} \varepsilon_{\cdot 1} \\ \varepsilon_{\cdot 2} \\ \vdots \\ \varepsilon_{\cdot G} \end{bmatrix}$$

$$= Z^*\delta^* + \varepsilon^* \qquad \text{or} \qquad (5.14)'$$

$$\boxed{y^* = Z^*\delta^* + \varepsilon^*}$$

where

$$\boxed{\varepsilon^* \sim N[0; \Omega \otimes I],}$$

$$\mathrm{var}(\varepsilon^*) = \begin{bmatrix} \mathrm{var}(\varepsilon_{\cdot 1}) & \mathrm{cov}(\varepsilon_{\cdot 1}, \varepsilon_{\cdot 2}) & \dots & \mathrm{cov}(\varepsilon_{\cdot 1}, \varepsilon_{\cdot G}) \\ & \mathrm{var}(\varepsilon_{\cdot 2}) & \dots & \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{cov}(\varepsilon_{G\cdot}, \varepsilon_{\cdot 1}) & \dots & & \mathrm{var}(\varepsilon_{\cdot G}) \end{bmatrix}$$

$$= \begin{bmatrix} \omega_{11}I & \omega_{12}I & \dots & \omega_{1G}I \\ \omega_{21}I & \omega_{22}I & \dots & \omega_{2G}I \\ & \vdots & & \\ \omega_{G1}I & \dots & & \omega_{GG}I \end{bmatrix}$$

$$= \Omega \otimes I.$$

<u>Note</u>:

(a)

$$Z^* = \begin{bmatrix} Z_1 & 0 & 0 & \cdots & 0 \\ 0 & Z_2 & 0 & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & Z_G \end{bmatrix}$$

(b)

$$Z^* = \begin{bmatrix} Y_1^{(1)} & X_1^{(1)} & & 0 & 0 & \cdots & 0 \\ 0 & & Y_1^{(2)} & X_1^{(2)} & 0 & \cdots & 0 \\ \vdots & & & & & & \\ 0 & & \cdots & 0 & \cdots & Y_1^{(G)} & X_1^{(G)} \end{bmatrix}$$ and

(c)

$$\delta^* = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix} = \begin{bmatrix} \beta'_{1\cdot} \\ \gamma'_{1\cdot} \\ \beta'_{2\cdot} \\ \gamma'_{2\cdot} \\ \vdots \\ \beta'_{G\cdot} \\ \gamma'_{G\cdot} \end{bmatrix}$$

δ^* is a column vector containing all structural coefficients in the entire model. The representation in (5.14) can be used to represent OLS, 2SLS, as well as 3SLS estimators.

The <u>OLS structural coefficient estimators</u> of δ in (5.14) can collectively be represented as

(5.15)

$$\delta_{OLS} = \begin{bmatrix} \delta_{1(OLS)} \\ \vdots \\ \delta_{G(OLS)} \end{bmatrix} = (Z^{*\prime} Z^{*})^{-1} Z^{*\prime} y^{*}$$

which can be shown to reduce to (5.2)

$$(\delta_h)_{OLS} = \begin{bmatrix} \beta'_{h\cdot} \\ \gamma'_{h\cdot} \end{bmatrix}_{OLS} = (Z'_h Z_h)^{-1} X'_h y_{\cdot h} \qquad (5.15)'$$

$$= \begin{bmatrix} Y_1^{(h)\prime} Y_1^{(h)} & Y_1^{(h)\prime} X_1^{(h)} \\ X_1^{(h)\prime} Y_1^{(h)} & X_1^{(h)\prime} X_1^{(h)} \end{bmatrix}^{-1} \begin{bmatrix} Y_1^{(h)\prime} y_{\cdot h} \\ X_1^{(h)\prime} y_{\cdot h} \end{bmatrix}$$

Recall that the OLS estimators are biased and inconsistent because the explanatory variables (right-hand side variables) are correlated with the error terms.

The <u>two-stage least squares</u> estimators are consistent. They can be thought of as being obtained by replacing the right-hand side endogenous variables by their LSNR predicted values and then applying least squares, i.e.,

(5.16)

$$\hat{\delta}_{2SLS} = (\hat{Z}^{*\prime} \hat{Z}^{*})^{-1} \hat{Z}^{*} y^{*}$$

where

$$\hat{Z}^{*} = \begin{bmatrix} \hat{Z}_1 & 0 & 0 & \dots & 0 \\ 0 & \hat{Z}_2 & 0 & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & \hat{Z}_G \end{bmatrix}$$

$$= \begin{bmatrix} \hat{Y}_1^{(1)} X_1^{(1)} & 0 & 0 & \ldots & 0 \\ 0 & \hat{Y}_1^{(2)} & X_1^{(2)} \, 0 & \ldots & 0 \\ \ldots & & & & \\ \ldots & & & & \\ \ldots & & & & \\ 0 & \ldots & & \hat{Y}_1^{(G)} X_1^{(G)} & \end{bmatrix};$$

hence,

$$(\hat{\delta}_h)_{2SLS} = \begin{bmatrix} \beta'_{h\cdot} \\ \hat{\gamma}'_{h\cdot} \end{bmatrix}_{OLS} = (\hat{Z}'_h \hat{Z}_h)^{-1} \hat{Z}'_h y_{\cdot h}$$

$$= \begin{bmatrix} \hat{Y}_1^{(h)\prime} \hat{Y}_1^{(h)} & \hat{Y}_1^{(h)\prime} X_1^{(h)} \\ X_1^{(h)\prime} \hat{Y}_1^{(h)} & X_1^{(h)\prime} X_1^{(h)} \end{bmatrix}^{-1} \begin{bmatrix} \hat{Y}_1^{(h)\prime} y_{\cdot h} \\ X_1^{(h)\prime} y_{\cdot h} \end{bmatrix} \qquad (5.16)'$$

which is in the same form as in (5.2).

The 2SLS estimator of δ^* can also be expressed as

$$\delta^*_{2SLS} = [Z^{*\prime} X^* (X^{*\prime} X^*)^{-1} X^{*\prime} Z^*]^{-1} Z^{*\prime} X^* (X^{*\prime} X^*)^{-1} X^{*\prime} y^*$$

since

$$\hat{Z}^* = X^* (X^{*\prime} X^*)^{-1} X^{*\prime} Z^*$$

where $X^* = \begin{bmatrix} X & 0 & 0 & \ldots & 0 \\ 0 & X & 0 & \ldots & 0 \\ \ldots & & & & \\ \ldots & & & & \\ \ldots & & & & \\ 0 & 0 & 0 & \ldots & X \end{bmatrix} = I \otimes X.$

The two-stage least squares estimators make corrections for the correlation between the right- hand side dependent variable and the structural error term, but do not take account of the covariances between the error terms in different equations.

Three-stage least squares makes adjustments for:

(1) the correlation between the right-hand side endogenous variables and the error terms, and

(2) correlations between the error terms in different structural equations. This is done by multiplying (5.14) by $X^{*\prime} = I \otimes X'$:

$$\boxed{X^{*\prime}y^{*} = X^{*\prime}Z^{*}\delta^{*} + X^{*\prime}\varepsilon^{*}} \tag{5.17}$$

where $X^{*\prime}\varepsilon^{*} \sim N[0; \Omega \otimes X'X]$

since $X^{*\prime}(\Omega \otimes I)X^{*} = \Omega \otimes X'X$.

Note that the 2SLS estimator of δ^{*} can be obtained by applying least squares to (5.17), three-stage least squares estimators can be obtained by applying a generalized least squares formula to (5.17) to yield

$$(\delta^{*})_{3SLS} = [(X^{*\prime}Z^{*})'[X^{*\prime}(\Omega \otimes I)X^{*}]^{-1}X^{*\prime}Z^{*}]^{-1}(X^{*\prime}Z^{*})'[X^{*\prime}(\Omega \otimes I)X^{*}]^{-1}X^{*\prime}y^{*}$$

$$= [Z^{*\prime}(\Omega^{-1} \otimes X(X'X)^{-1}X')Z^{*}]^{-1}Z^{*\prime}(\Omega^{-1} \otimes X'(X'X)^{-1}X)y^{*}$$

Hint: Recall properties of $\otimes$

Comments:

(1) In practice Ω is not known and is estimated from 2SLS residuals.

(2) If Ω is diagonal or each structural equation is exactly identified, then 3SLS = 2SLS.

(3) $\hat{\delta}_{3SLS}$ is consistent. The asymptotic distributi $\hat{\delta}_{3SLS}$ is

$$N[\delta; [Z^{*\prime}(\Omega^{-1} \otimes X(X'X)^{-1}X')Z^{*}]^{-1}$$

Exercise: Demonstrate that 3SLS estimators of δ can be obtained as instrumental variables estimators.

Hint: Select the instrumental variables, $Z = (\hat{\Omega}^{-1} \otimes X(X'X)^{-1}X')Z^*$ equation (3.3) solve $Z'y^* = Z'Z^*\delta_{IV}$ for δ_{IV}.

The asymptotic distribution of

$$\sqrt{N}\left[\begin{bmatrix}\tilde{B}_{vec}\\ \tilde{\Gamma}_{vec}\\ \tilde{\Omega}_{vec}\end{bmatrix}_{3SLS} - \begin{bmatrix}B_{vec}\\ \Gamma_{vec}\\ \Omega_{vec}\end{bmatrix}\right] \qquad \text{is}$$

$$N\left[0,\ \Sigma = \left[-\operatorname{plim}\frac{1}{N}\begin{bmatrix}\ell_{BB} & \ell_{B\Gamma} & \ell_{B\Omega}\\ \ell_{\Gamma B} & \ell_{\Gamma\Gamma} & \ell_{\Gamma\Omega}\\ \ell_{\Omega B} & \ell_{\Omega\Gamma} & \ell_{\Omega\Omega}\end{bmatrix}^{-1}\right]\right]$$

where ℓ denotes the log likelihood function and subscripts denote derivatives. Thus, both 3SLS and FIML are asymptotically efficient.

Note:

1. David Hendry demonstrated that almost all simultaneous equation estimators (3SLS, 2SLS, instrumental variables, . . .) can be obtained as approximate solutions to the necessary conditions defining the FIML estimators.
   Hendry, "The Structure of Simultaneous Equation Estimators," Journal of Econometrics 4 (1976), pp. 51-88.

6. Statistical Inference

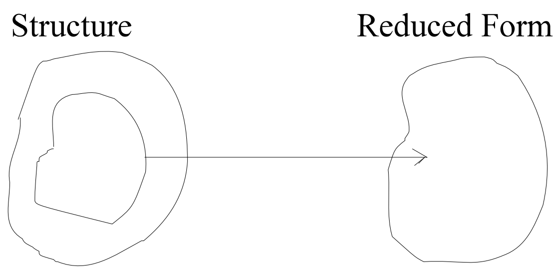
This section briefly summarizes procedures for testing hypotheses about (1) identification, (2) structural and reduced form coefficients, and (3) restrictions on structural coefficients.

a. Identification -- Logically Prior to Estimation

(1) System tests of the validity of all overidentifying restrictions.
Hendry (1972, IER) Let

$$LR = \tilde{N}\ln\left(|\tilde{\Sigma}|/|\hat{\Sigma}|\right)$$

where



$$\tilde{\Sigma} = \tilde{B}^{-1}\tilde{\Omega}(\tilde{B}^{-1})'$$

is the restricted estimator of the reduced form variance covariance matrix and $\hat{\Sigma}$ is the unrestricted estimator of the reduced form variance covariance matrix.

$$\hat{\Sigma} = [Y'(I - X(X'X)^{-1}X')Y]/(N - K)$$

The asymptotic distribution of LR is

χ^2 (total number of overidentifying restrictions)

Also see Hendry, The Econometric Analysis of Time Series, p. 338.

Questions:

1. Which estimation technique? Hendry suggests FIML or 3SLS.

2. How large sample size for the asymptotic distribution to be an accurate approximation?

(2) Single equation tests of over identifying restrictions

The null hypothesis is that the instruments (deleted exogenous variables) should not be included in the structural equation being estimated.

A common test statistic of this hypothesis is defined as follows:

$$\frac{(N-K)}{(\upsilon)} \frac{G_1(\tilde{\beta}_{1.}) - G_2(\tilde{\beta}_{1.})}{G_2(\tilde{\beta}_{1.})} \overset{a}{\sim} F(\upsilon, N-K)$$

where $\tilde{\beta}_{1.}$ denotes an estimator $\beta_{1.}$.

If the Stata command **estat overid** follows an **ivregress** command, then tests of overidentification are performed. If 2SLS is used, Basmann's chi-square test is reported as is Woodridge's robust score test. If LIML is used, Anderson and Rubin's chi square test is reported as is Basmann's F test (above). If GMM is used, Hansen's J statistica and chi square test are given. Statistically insignficicance is consistent with the instruments not being included in the structural equation of interest; whereas, statistically significant tests indicate that the exclusion restrictions are not valid and the instruments are not valid.

ivregress 2sls or liml y1 (Y1=X1 X2) X1
estat overid

Some theoretical references on the exact distribution of the distribution of the identifiability test statistic. McDonald (1972, Econometrica, G_Δ = 2, LIML), Basmann, Richardson (G_Δ =2, 2SLS) (1973, Econometrica), and Rhodes (1981, Econometrica, G_Δ arbitrary, LIML)

Simulation results suggest the following about the quality of approximation

$F(\upsilon, N-K)$ $\chi^2(\nu)$ an

OLS "F statistic" provides a poor approximation
2SLS and LIML identifiability test statistics are more nearly approximated by the F distribution than the OLS test statistic.

LIML identifiability test statistic seems to be closer to F than for 2SLS.

b. Reduced form coefficients

(1) LSNR: $\hat{\Pi}_{Vec} \sim N(\Pi_{Vec};\ \Sigma \otimes (X'X)^{-1})$

(a) t, F appropriate for LSNR
(b) Chow, LR, and Wald tests are appropriate

(2) Derived RF: $\tilde{\Pi}_{Vec} \overset{A}{\sim} N(\Pi_{Vec}; \Sigma_{\tilde{\Pi}})$

t and F statistics provide the *asymptotic* distributions for derived reduced from estimators

$$\frac{\tilde{\Pi}_{ij} - \Pi_{ij}}{S_{\tilde{\Pi}_{ij}}} \overset{a}{\sim} N(0,1)$$

c. Structural coefficients

(1) Ho: $\beta_{ij} = \beta^0_{ij}$

$$\frac{\hat{\beta}_{ij} - \beta_{ij}}{s_{\hat{\beta}_{ij}}} \sim_a N[0,1] \text{ or } t(\ )$$

(2) $H_o: \gamma_{ij} = \gamma^0_{ij}$

$$\frac{\hat{\gamma}_{ij} - \gamma_{ij}}{s_{\hat{\gamma}_{ij}}} \sim_a N[0,1] \text{ or } t(\ )$$

Some observations based upon Monte Carlo simulations and analytic results suggest:

(1) OLS "t-statistics" do not seem to have a distribution which is closely approximated by a t density.

(2) The density functions of the 2SLS and LIML "t-statistics" seem to be much more closely approximated by a t density. (It all depends.)

- "2SLS and LIML "t-statistics" seem reliable to use in testing significance of exogenous variables"

- "For an endogenous variable, the distribution of the 2SLS t-statistic deviates far from Student's t distribution if the non-centrality parameter is small (<13) and degree of over-identification (>7)."

- "Student's t approximation to the LIML t-statistic is the most accurate, and the two-sided test may be most appropriate in empirical studies (modest skewness)."

- Morimune, K, t Test in a Structural Equation, Econometrica, 57 (1989), 1341-1360.

(3) Caution: while the "t-ratios" for consistent structural coefficient estimators are asymptotically N [0, 1], this does not imply that either the t or the N(0,1) will give accurate results for sample sizes encountered in practice.

(4) An alternative which may be practical in simple models is to transform the hypothesis to the reduced form and test using LSNR.

(5) Still another approach to determining the distribution of the test statistic is to use the bootstrap.

c. "Testing a Subset of Coefficients in a Structural Equation"

Consider the structural model

$$y_1 = y_1\beta'_{10} + X_1\gamma'_{10} + \varepsilon_1$$

H_0: Some of the coefficients in β_{10} and γ_{10} are zero.

(1) Chow statistics based upon a comparison of SSE's are **<u>INVALID</u>**.

(2) Wald tests can be fairly accurate–based on asymptotic results (sample size)

These tests can be implemented in Stata by obtaining consistent estimators of the structural coefficients and then using the test command, e.g., ivregress (2sls or liml) y1 (Y1=Z) X1, followed by test x3=0

(3) Morimune and Tsukuda found that "likelihood ratio" tests based on 2SLS or LIML estimators were quite closely approximated with a chi square density.

$$\tilde{\lambda}_1 = \frac{G_1(\tilde{\beta}_{1\cdot}) - G_2(\tilde{\beta}_{1\cdot})}{G_2(\tilde{\beta}_{1\cdot})}$$

$$\tilde{\lambda}_2 = \frac{G_1(\tilde{\beta}_{2\cdot}) - G_2(\tilde{\beta}_{2\cdot})}{G_2(\tilde{\beta}_{2\cdot})}$$

where the subscript "1" corresponds to the case in which the variables with hypothesized zero coefficients have been deleted and "2" to the inclusion of the variables.

The test statistic is then given by

$$\frac{N - \#\ \text{coeff. in unrestricted equation}}{\#\ \text{restrictions}} \qquad \frac{\tilde{\lambda}_1 - \tilde{\lambda}_2}{1 + \tilde{\lambda}_2}$$

This statistic is asymptotically distributed as

F (# restrictions, N - # Coefficients in unrestricted model)

If there is a single restriction, as in the previous section, the corresponding LR test statistic is asymptotically as an F(1, N - # Coefficients in unrestricted model) or as a $\chi 2$ (N - # Coefficients in unrestricted model) statistic. Morimune, K. and Y. Tsukuda, "Testing a Subset of Coefficients in a Structural Equation," Econometrica, 52 (1984), 427-448.

(4) Testing Structural Parameters when using Instrumental Variables
Kleibergen, F, "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," Econometrica 70(2002), 1781-1803

Kleibergen proposes a test statistic, based on a quadratic form of the score of the concentrated log-likelihood, which can be used for performing joint tests on all of the structural parameters in instrumental variables regression. The test statistic is independent of "nuisance" parameters (pivotal statistic) and has an asymptotic chi distribution.

7. Simultaneous Equations Exercises

A. Identification

1. Using the supply and demand example (number 1) from section VI.3.a, express each of the six structural parameters in terms of reduced form parameters.

2. Attempt to replicate (1) for example 2 in section VI.3.a.

3. Attempt to replicate (1) for example 3 in section VI.3.a.

4. Verify that

   (1) $\nu = K_2 - G_\Delta + 1 \geq 0,$

   (2) $K_2 + G_{\Delta\Delta} + 1 \geq G,$ and

   (3) $K_2 \geq G_\Delta - 1$

   are equivalent expressions for the order (necessary) conditions for identifying structural equations.

5. Obtain the reduced form of the following set of structural equations.

$$
\begin{aligned}
y_{1t} &= -2y_{2t} + 7x_{1t} + 4x_{2t} + x_{3t} - 8x_{4t} + u_{1t} \\
y_{2t} &= 2y_{1t} + y_{3t} - x_{1t} + 7x_{3t} - 9x_{5t} + u_{2t} \\
y_{3t} &= 2y_{1t} - 7x_{2t} + 7x_{3t} + 14x_{4t} + u_{3t}
\end{aligned}
$$

   Investigate the identifiability of each equation, both (a) by using only the structural equations and (b) by using the reduced form equations. (L.S.E. 1966)

6. Consider the model defined by

$$
\begin{aligned}
C_t &= \alpha_o + \alpha_1 Y_t + u_t & (1) \\
I_t &= \beta_0 + \beta_1 Y_t + \beta_2 I_{t-1} + v_t & (2) \\
Y_t &= C_t + I_t + G_t & (3)
\end{aligned}
$$

   Discuss the identification of the above equations if G is the only exogenous variable (apart from the dummy variable for the constant term)

   a. as written

   b. as in (a) but over the sample period G is constant.

7. Discuss the restrictions, if any, which are implied for the reduced form by each of the following structural equations, assuming nothing is known about the covariance matrix of the disturbance terms.

$$y_{1t} = \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + u_{1t}$$

$$y_{2t} = \beta_{21}y_{1t} + \gamma_{21}x_{1t} + \gamma_{22}x_{2t} + \gamma_{23}x_{3t} + u_{2t}$$

$$y_{3t} = \beta_{31}y_{1t} + \beta_{32}y_{2t} + u_{3t}$$

What additional restrictions, if any are implied by

a. $\gamma_{12} = 0$ (L.S.E. 1969)

B. A Little Theory

1. Verify that

$$\hat{\pi}_{vec} \sim N\left[\pi_{vec}, \sum \otimes (X'X)^{-1}\right]$$

simplifies to the regular regression result in the case of a model with a single reduced form equation.

2. What do you think the relationship between 2SLS and OLS estimators will be in the case of the reduced form equations having high R^2's. Hint: Compare (5.5) and (5.6) in VI.5.

3. Verify that

$$Y'_1 Y_1 - \hat{V}'_1 \hat{V}_1 = \hat{Y}'_1 \hat{Y}_1$$

in the discussion of 2SLS.

4. One estimator of $\omega_{11} = \text{var}(\varepsilon_{t\bullet})$ is given by

$$\hat{\omega}_{11} = \left(\frac{1}{N - K + \nu}\right) G_1(\hat{\beta}_{1\bullet}).$$

explain why the denominator $N - K + \nu$ is used. See the discussion of 2SLS.

5. In the discussion of LIML verify that

$$\hat{\lambda} = \frac{G_1\ (\hat{\beta}_{1\bullet})}{G_2\ (\hat{\beta}_{1\bullet})} - 1.$$

6. In the discussion of k-class estimators

   verify Min_{β} $[G_1\ (\beta_{1\bullet}) - k\ G_2\ (\beta_{1\bullet})]$ yields

   (1) OLS for k = 0
   (2) 2SLS for k = 1
   (3) LIML for k = λ (problem 5 above)

7. Verify that $Z = [Y_1 - k\hat{V}_1\ , X_1]$ yields an instrumental variables equivalent to k- class estimators.

8. What value of k yields Zellner's MELO estimator for $N = 25$, $K = 5$, $K_2 = 1$, $G_\Delta = 2$, $\nu = 1$.

9. Verify that $\Omega = I$ implies that 3SLS = 2SLS.

C. Application: An exactly identified case

Consider the following Supply and Demand Model:

Demand: $Q_t = a_{11} + B_{12} P_t + a_{12} Y_t + e_{t1}$

Supply: $Q_t = a_{21} + B_{22} P_t + a_{23} FC_t + e_{t2}$

Where Q_t, P_t, Y_t and FC_t denote quantity, price, income and factor costs.

Observations on these variables are given by:

P_t	185	215	275	279	310	330	400	360	450	515
Q_t	320	360	460	460	480	540	600	570	680	780
Y_t	100	120	160	164	180	200	240	220	280	320
FC_t	10	12	14	15	20	16	24	20	28	30

1. Express the reduced form representation in terms of the structural coefficients.

2. Determine which of the structural coefficients can be expressed in terms of the reduced form coefficients and make this relationship explicit where possible.

3. Determine whether the supply and demand equations are identified. Use the necessary (order) conditions for identification in your analysis.

4. Estimate the reduced form equations for P and Q using the technique of Least Squares No Restrictions (LSNR), i.e., just use regular least squares {reg P D FC and reg Q D FC) .

a) Test for the presence of autocorrelation.

b) Test for heteroskedasticity.

5. Estimate the supply and demand equations using OLS.

6. Estimate the supply and demand equations using 2SLS.

7. Comment on the properties of the estimators associated with (5) and (6).

8. Indicate how you could test the following hypotheses and discuss any related problems.

a) $\beta_{12} = -2$

b) $a_{12} = 0$

c) $\Pi_{12} = 2.5$

d) $\Pi_{22} = 0$

9. What implication does $\Pi_{22} = 0$ have with respect identification of any of the structural equations?

10. Indicate and perform a predictive test of your model when the last two observations are used to test the predicative ability of the model under consideration. Do these two observations lie in 95% confidence intervals?

11. Indicate how you would go about making price and quantity predictions for the next two periods after the sample data. Construct appropriate confidence intervals.

D. Application: an overidentified case

Consider the following model to explain variation in consumption and prices of food:

Demand: $Q_t = \beta_{12}P_t + \gamma_{11} + \gamma_{12}D_t + \varepsilon_{t1}$

Supply: $Q_t = \beta_{22}P_t + \gamma_{21} + \gamma_{23}F_t + \gamma_{24}A_t + \varepsilon_{t2}$

where is P = relative prices of food to consumer prices
Q=per captia food consumption
D=per capita disposable income
F=prices received by farmers last year/general consumer prices
A=time index in years

1. Check the necssary (order)conditions for each of the structural equations to be identified.

2. Estimate the structural equations using the methods of OLS, 2SLS, LIML, and 3SLS.

3. Estimate the corresponding reduced form equations using the methods of (1) LSNR and (2) 2SLS . Recall that the 2SLS estimates of the reduced form are obtained by estimating the structure using 2SLS and then deriving the corresponding reduced form estimates.

4. Do the variables D, F, and A have statistically significant explanatory power in the reduced form equation for P? Use F and t-tests. What implications do these results have for the strength of the instruments?

5. Perform and interpret a statistical test of the overidentifying restrictions for the demand equation

6. Obtain predictions for Q corresponding to 2SLS and LSNR for (D.F,A) = (140, 110, 25).

6. Run separate tests of the hypotheses $H_o: \beta_{12} = 0$ $H_o: \gamma_{12} = 0$

Q	P	D	F	A
98.485	100.323	87.4	98	1
99.187	104.264	97.6	99.1	2
102.163	103.435	96.7	99.1	3
101.504	104.506	98.2	98.1	4
104.24	98.001	99.8	110.8	5
103.243	99.456	100.5	108.2	6
103.993	101.066	103.2	105.6	7
99.9	104.763	107.87	109.8	8
100.35	96.446	96.6	108.7	9
102.82	91.228	88.9	100.6	10
95.435	93.085	75.1	81	11
92.424	98.801	76.9	68.6	12
94.5358	102.908	84.5	70.9	13
98.757	98.756	90.6	81.4	14
105.797	95.119	103.1	102.3	15
100.225	98.451	105.1	105	16
103.522	86.498	96.4	110.5	17
99.929	104.016	104.4	92.5	18
105.223	105.769	110.7	89.3	19
106.232	113.49	127.1	93	20

E. A simple dynamic model

A very simple example of a dynamic model is the famous cobweb model defined by the **structural model**:

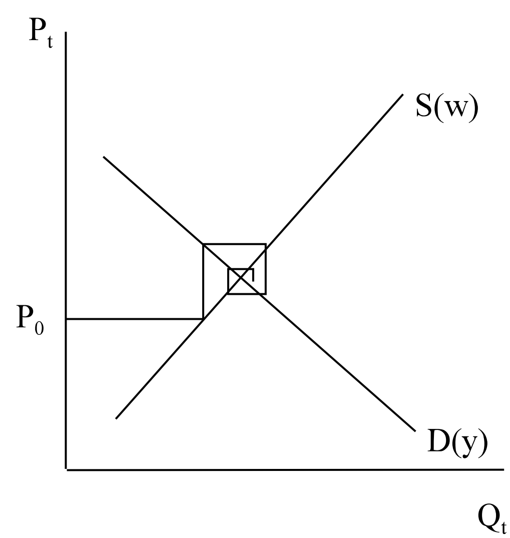
Demand: $Q_t = \alpha + \beta P_t + \xi Y_t$ (E.1 a-b)

Supply: $Q_t = \gamma + \delta P_{t-1} + \rho W_t$

The variables appearing in the model can be classified as endogenous, and predetermined (exogenous and lagged endogenous) as follows:

Endogenous: Q_t, P_t

$$\left.\begin{array}{l}\text{Exogenous: } Y_t, W_t \\ \text{Lagged Endog: } P_{t-1}\end{array}\right\} \text{Predetermined}$$



If the initial price (P_0) differs from the equilibrium price, then the time path of (Q_t, P_t) will resemble a "cobweb" -- converging to the equilibrium if

$|slope\ demand\ curve| <$ *slope of the supply curve*, or

$|1/\beta| < 1/\delta$ $\quad |\delta/\beta| < 1$ or $\quad |\delta/\beta| > 1$ diverging for

Note:

- Economic theory suggests that δ and ξ are positive and β and ρ are negative.
- For convenience, the random disturbances have been deleted.

1. Demonstrate that the supply and demand equations satisfy the necessary conditions for identification. Recall that the necessary condition is that the number of excluded predetermined (exogenous and predetermined) variables appearing in the model must be at least as large as the number of current endogenous regressors in each equation.

2. Demonstrate that the structural model can also be written as:

(E.2)

$$\begin{bmatrix} -1 & \beta \\ -1 & 0 \end{bmatrix} \begin{bmatrix} Q_t \\ P_T \end{bmatrix} + \begin{bmatrix} \alpha & 0 & 0 & \xi \\ \gamma & \delta & \rho & 0 \end{bmatrix} \begin{bmatrix} 1 \\ P_{t-1} \\ W_t \\ Y_t \end{bmatrix} = 0 \qquad ,$$

(E.3 a-c)

Structural equations	**Generic form** (Section VI.2)
$\begin{bmatrix} -1 & \beta \\ -1 & 0 \end{bmatrix} \begin{bmatrix} Q_t \\ P_T \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \delta \end{bmatrix} \begin{bmatrix} Q_{t-1} \\ P_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha & 0 & \xi \\ \gamma & \rho & 0 \end{bmatrix} \begin{bmatrix} 1 \\ W_t \\ Y_t \end{bmatrix} = 0$	$BY_t' + B_1 Y_{t-1}' + \Gamma X_t' = 0$
$\left\{ \begin{bmatrix} -1 & \beta \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \delta \end{bmatrix} L \right\} \begin{bmatrix} Q_t \\ P_t \end{bmatrix} + \begin{bmatrix} \alpha & 0 & \xi \\ \gamma & \rho & 0 \end{bmatrix} \begin{bmatrix} 1 \\ W_t \\ Y_t \end{bmatrix} = 0$	$(B + B_1 L) Y_t' + \Gamma X_t' = 0$
$\begin{bmatrix} -1 & \beta \\ -1 & \delta L \end{bmatrix} \begin{bmatrix} Q_t \\ P_t \end{bmatrix} + \begin{bmatrix} \alpha & 0 & \xi \\ \gamma & p & 0 \end{bmatrix} \begin{bmatrix} 1 \\ W_t \\ Y_t \end{bmatrix} = 0 \ .$	$B(L) Y_t' + \Gamma X_t' = 0$

3. Demonstrate that the reduced form representation can be obtained from (E.2)

(E.4)

$$\begin{bmatrix} Q_t \\ P_t \end{bmatrix} = -\begin{bmatrix} -1 & \beta \\ -1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \alpha & 0 & 0 & \xi \\ \gamma & \delta & \rho & 0 \end{bmatrix} \begin{bmatrix} 1 \\ P_{t-1} \\ W_t \\ Y_t \end{bmatrix}$$

$$= \begin{bmatrix} \gamma + \delta P_{t-1} + \rho W_t \\ \frac{\gamma-\alpha}{\beta} + \frac{\delta}{\beta} P_{t-1} + \frac{\rho}{\beta} W_t - \frac{\xi}{\beta} Y_t \end{bmatrix}$$

Note: The reduced form expresses each current dependent variable in terms of predetermined (exogenous and lagged endogenous) variables.

4. Demonstrate that the final form or transfer functions for P_t and Q_t can be written as are

Generic form

$$\begin{bmatrix} Q_t \\ P_t \end{bmatrix} = -\begin{bmatrix} -1 & \beta \\ -1 & \delta L \end{bmatrix}^{-1} \begin{bmatrix} \alpha & 0 & \xi \\ \gamma & \rho & 0 \end{bmatrix} \begin{bmatrix} 1 \\ W_t \\ Y_t \end{bmatrix}$$

$$Y_t' = -B^{-1}(L)\Gamma X_t'$$

$$= \begin{bmatrix} \frac{\gamma\beta+\alpha\delta}{\delta-\beta} + \rho \sum_{i=0} (\delta/\beta)^i W_{t-i} - \frac{\xi\delta}{\beta} \sum_{i=0} (\delta/\beta)^i Y_{t-i-1} \\ \frac{\alpha-\gamma}{\delta-\beta} - \frac{\rho}{\beta} \sum_{i=0} (\delta/\beta)^i W_{t-i} - \frac{\xi}{\beta} \sum_{i=0} (\delta/\beta)^i Y_{t-i} \end{bmatrix}$$

Hint: This result follows from the following relationship.

$$B^{-1}(L) = \begin{bmatrix} -1 & \beta \\ -1 & \delta L \end{bmatrix}^{-1} = \frac{1}{-\delta L + \beta} \begin{bmatrix} \delta L & -\beta \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{\beta(1 - (\delta/\beta)L)} \begin{bmatrix} \delta L & -\beta \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{\beta} \Sigma(\delta/\beta)^i L^i \begin{bmatrix} \delta L & -\beta \\ 1 & -1 \end{bmatrix}.$$

Note: In the final form, each current dependent variable is expressed in terms of current and lagged values of exogenous variables.

Further note that:

$$\frac{\partial Q_t}{\partial W_t} = \rho<0 \qquad \frac{\partial Q_t}{\partial W_{t-1}} = \frac{\rho\delta}{\beta}>0,$$

$$\frac{\partial Q_t}{\partial Y_t} = 0 \qquad \frac{\partial Q_t}{\partial Y_{t-1}} = \frac{-\xi\delta}{\beta}>0, \; \ldots.$$

$$\frac{\partial P_t}{\partial W_t} = \frac{\rho}{\beta}>0 \qquad \frac{\partial P_t}{\partial W_{t-1}} = \frac{\rho\delta}{\beta^2}<0, \; \ldots.$$

$$\frac{\partial P_t}{\partial Y_t} = \frac{-\xi}{\beta}>0 \qquad \frac{\partial P_t}{\partial Y_{t-1}} = \frac{-\xi\delta}{\beta^2}<0, \; \ldots. \; .$$

5. Show that the short-run (impact) multipliers can be obtained from the reduced form coefficients as well as from the transfer function form. Hint: $-B^{-1}(L=0)\Gamma = -B^{-1}\Gamma$

6. Demonstrate that the <u>long-run cumulative multipliers</u> are given by

$$-\begin{bmatrix} -1 & \beta \\ -1 & \delta(L=1) \end{bmatrix}^{-1} \begin{bmatrix} \alpha & 0 & \xi \\ \gamma & \rho & 0 \end{bmatrix} = \begin{bmatrix} \dfrac{\alpha\delta-\beta\gamma}{\delta-\beta} & \dfrac{-\beta\rho}{\delta-\beta} & \dfrac{\delta\xi}{\delta-\beta} \\ \dfrac{\alpha-\gamma}{\delta-\beta} & \dfrac{-\rho}{\delta-\beta} & \dfrac{\xi}{\delta-\beta} \end{bmatrix}$$

Hint: $-B^{-1}(L=1)\Gamma = -(B+B_1)^{-1}\Gamma$

7. What STATA commands would you use to obtain consistent estimators of
   a. the reduced form coefficients
   b. the structural coefficients in the cobweb model?

8. The cobweb formulations corresponds to expectations formed as follows $E\left(P_t \middle| P_{t-1,\ldots}\right) = P_t^* = P_{t-1}$. Alternative models for forming expectations have been c
including adaptive expectations, rational expectations, and the use of ARIMA models.

F. Consider the following model to describe returns to education for married working women:

$$\boxed{\ln(wage) = \beta_1 + \beta_2 education + \varepsilon}$$

This model can be estimated using data from Mroz (Econometrica 54 (1986, 765-799))

1. Estimate the model using OLS
2. Estimate the model using IV with Z=father's education
3. Estimate the model using IV with Z=(father's education and mother's education)
   Use and interpret the ***estat overid*** command
4. Compare and interpret the different estimates of the impact of education on ln(wage).
5. Which estimate would you feel most comfortable with and why?
6. Comment on additional modifications you might include in the model.

The data is found in mroz.raw and is described as follows:

inlf hours kidslt6 kidsge6 age educ wage repwage
hushrs husage huseduc huswage faminc mtr motheduc
fatheduc unem city exper nwifeinc lwage expersq

Obs: 753

1. inlf	=1 if in labor force, 1975
2. hours	hours worked, 1975
3. kidslt6	# kids < 6 years
4. kidsge6	# kids 6-18
5. age	woman's age in yrs
6. educ	years of schooling
7. wage	wife's estimated wage from earns., hours
8. repwage	reported wage at interview in 1976
9. hushrs	hours worked by husband, 1975
10. husage	husband's age
11. huseduc	husband's years of schooling
12. huswage	husband's hourly wage, 1975
13. faminc	family income, 1975
14. mtr	fed. marginal tax rate facing woman
15. motheduc	mother's years of schooling
16. fatheduc	father's years of schooling
17. unem	unem. rate in county of resid.
18. city	=1 if live in SMSA
19. exper	actual labor mkt exper
20. nwifeinc	(faminc - wage*hours)/1000
21. lwage	log(wage)
22. expersq	exper^2

G. An Application of Transfer Functions

Maloney and Ireland ("Fiscal Versus Monetary Policy: An Application of Transfer Functions," Journal of Econometrics, 1980, pp. 253-266) use transfer functions to study the relative importance of monetary and fiscal policy upon aggregate production. They examine the coefficients (dynamic multipliers) of the lagged exogenous variables in the transfer function representation. Maloney and Ireland do not begin their analysis by specifying a structural model, as done in Kmenta and Smith (RESTAT,1973), but rather estimate a transfer function directly. Their model is given by

$$y_t = \frac{w_1(L)}{\delta_1(L)} m_t + \frac{w_2(L)}{\delta_2(L)} g_t + \zeta_t$$

where w_i (L) and δ_i (L) are assumed to be polynomials in the lag operator (L) and y_t, g_t, and m_t respectively, denote the percentage change in real GDP, change in real government purchases of goods and services, and in the real monetary base. Using quarterly data from 1953 to 1975, they estimate the model to be

$$y_t = \frac{.243}{1 - 1.4847L + .70886L^2} m_t + \frac{.06648}{1 - 1.0311L + .83244L^2} g_t + .005 + \zeta_t \, .$$

The impact multipliers corresponding to *m* and *g* are .243 and .06648.

a. What is the interpretation of the impact multipliers corresponding to *m* and g?

b. What is the interpretation of the long-run multipliers corresponding to *m* and *g?*

c. Evaluate the impact (short run) multipliers corresponding to m and g.

Helpful notes on the calculation of lagged weights for m and g in the Maloney-Ireland transfer function representation. Consider the calculation of the distributed lag coefficients for *m*.

$$\frac{w_1(L)}{\delta_1(L)} =$$

$$.24 + .36\,\mathrm{L} + .37\mathrm{L}^2 + .25L^3 + \ldots..$$

$$1 - .1.5L + .7L^2 \,|\, .24$$

The short run or impact multiplier for m is given by

$$\frac{w_1(L=0)}{\delta_1(L=0)} = .24$$

which is the leading coefficient of the expansion of $\frac{w_1(L)}{\delta_1(L)}$.

The long run multiplier is given by

$$\frac{w_1(L=1)}{\delta_1(L=1)} = \frac{.243}{1 - 1.4847 + .70886} = 1.08$$

which is the sum of the coefficients in the expansion of $\frac{w_1(L)}{\delta_1(L)}$.

Summary Table

	Impact multiplier	Long run multiplier
m	.243	1.08
G	.066	.083

Polynomial distributed lags could have been used rather than a ratio of polynomials.