EC402: MEI. Spring Term: The Analysis of Time Series*

Lecture Notes #5: Simultaneous Equations

1. Multivariate Regressions

This is characterized by a set of N reduced form regressions. The first building block is to extend single to multivariate regression. So y_t becomes an N vector,

$$y_t = X_t\beta + \varepsilon_t, \qquad t = 1, ..., T$$

where

$$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Nt} \end{bmatrix}, \; X_t = \begin{bmatrix} x'_{1t} & 0 & \cdots & 0 \\ 0 & x'_{2t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x'_{Nt} \end{bmatrix}, \; \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix}, \; \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}.$$

where each β_i is a vector of coefficients. Note that each of the i equations need not have the same regressors.

Assume that the ith equation has K_i regressors. Hence X_t is $N \times K$ where $K = \sum_i K_i$, and β is of size $K \times 1$.

The errors are assumed correlated across the i,

$$E\left(\varepsilon_t\varepsilon'_t\right) = \Omega$$

an $N \times N$ matrix, but uncorrelated across t. Just because we are using the t suffix does not mean that this is a good model for time series. It is a better model where we have multiple observations on each unit in a random sample: multiple demands by a household or firm for example.

This model is known as SUR: seemingly unrelated regressions. The equations appear to be unrelated but are in fact related through the errors.

A possible estimator for β is least squares on each of the i equations individually,

$$\hat{\beta} = \left(\sum_{t=1}^{T} X_t' X_t \right)^{-1} \sum_{t=1}^{T} X_t' y_t.$$

If the regressors are *process independent*, the estimates are unbiased with standard small sample properties. However they are not in general efficient.

An alternative if Ω is known is GLS. It is reasonably straightforward to see that the GLS estimator is

$$\hat{\beta}_{gls} = \left(\sum_{t=1}^{T} X_t' \Omega^{-1} X_t \right)^{-1} \sum_{t=1}^{T} X_t' \Omega^{-1} y_t.$$

Again if regressors are process independent the GLS estimator is optimal.

Of course in general Ω is not known and we need to fall back on feasible GLS. The approach is absolutely standard. Estimate the N equations separately using least squares. Calculate the least squares residuals. Estimate Ω as

$$\hat{\Omega} = \frac{1}{T} \sum_{t} \hat{\varepsilon}_t \hat{\varepsilon}_t'.$$

Then use the GLS formula, with $\hat{\Omega}$ for Ω to obtain the F-GLS estimator.

Note. under two circumstances the GLS estimator simplifies to least squares:

1. (Obviously) if Ω is known to be diagonal.

2. (Less obviously) if the regressors in each of the N equations are the same.[1]

[1] This result will also extend to the VAR models we'll discuss later on.

2. The General Simultaneous Equations Model

Suppose y_t is an $N \times 1$ vector of endogenous variables. Each of these is a linear function of the other endogenous variables and of a set of K exogenous variables, x_t. Usually x_t are contemporaneously independent of the errors or may be process independent. x_t may include lagged endogenous variables. The set of N equations is written as

$$\Gamma y_t = B x_t + \varepsilon_t \qquad t = 1, ..., T, \tag{2.1}$$

where Γ is $N \times N$, B is $N \times K$, ε_t is $N \times 1$, x_t is $K \times 1$, y_t is $N \times 1$. Expression (2.1) is called the *Structural Form*.

For example, suppose

$$\begin{aligned} y_{1t} &= \gamma_{12} y_{2t} + \beta_{11} x_{1t} + \beta_{14} x_{4t} + \varepsilon_{1t} \\ y_{2t} &= \gamma_{23} y_{3t} + \beta_{22} x_{2t} + \beta_{23} x_{3t} + \varepsilon_{2t} \\ y_{3t} &= \gamma_{32} y_{2t} + \beta_{31} x_{1t} + \beta_{33} x_{3t} + \beta_{34} x_{4t} + \varepsilon_{3t}. \end{aligned}$$

Then, in structural form, we have

$$\underbrace{\begin{bmatrix} 1 & -\gamma_{12} & 0 \\ 0 & 1 & -\gamma_{23} \\ 0 & -\gamma_{32} & 1 \end{bmatrix}}_{\Gamma} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \underbrace{\begin{bmatrix} \beta_{11} & 0 & 0 & \beta_{14} \\ 0 & \beta_{22} & \beta_{23} & 0 \\ \beta_{31} & 0 & \beta_{33} & \beta_{34} \end{bmatrix}}_{B} \begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \\ x_{4t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}.$$

The solution for y_t in terms of x_t is called the *Reduced Form*, that is,

$$y_t = \Gamma^{-1} B x_t + \Gamma^{-1} \varepsilon_t$$

or

$$y_t = \Pi x_t + v_t \tag{2.2}$$

where $\Pi = \Gamma^{-1} B$ is $N \times K$ and $v_t = \Gamma^{-1} \varepsilon_t$. Note that each element of v_t is generally a function of all the elements of ε_t, so each endogenous variable depends on all the errors.

The data may be thought of as being generated by the reduced form (2.2). So, given the data, we can make estimates of the elements of Π and the variance matrix of v.

2.1. Identification

Focussing on Π, if we know the elements of Π, can we find out about the parameters Γ and B? This is a problem of identification. In general, the answer is no.

The basic equation is

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

or

$$B = \Gamma\Pi.$$

Looking at each element in turn, we have $N \times K$ equations. Assuming that the diagonal elements of Γ are all restricted to be equal to one,[2] there are $N \times K$ unknowns in B and $N \times (N-1)$ unknowns in Γ. That is, many more unknowns than equations!

So what are the conditions for identification?

i) Conditions apply to one equation at a time.

ii) A necessary condition (the *order condition*) is that the number of omitted x, y variables in the equation has to be $\geq N - 1$. Note that *a linear restriction counts as a omitted variable.*

iii) A necessary and sufficient condition (the *rank condition*) is that the matrix obtained from *all* the Γ, B coefficients in the other equations corresponding to the zero's in the equation concerned, is of rank $N - 1$. (Note this matrix has $(N-1)$ rows.).

iv) If the nec./suff. conditions is passed and the number of zero's + linear restrictions $> N - 1$, the equation is over-identified; if equal to $N - 1$, it is just identified.

Note: an identity (e.g. $y = c + i + g$) is <u>always</u> identified.

[2] This is an innocuous normalization since we can always re-scale each equation by a constant.

Example 1.

$$\begin{aligned}
y_{1t} &= \gamma_{12}y_{2t} + \beta_{11}z_{1t} + \beta_{13}z_{3t} + \beta_{14}z_{4t} + u_{1t} \\
y_{2t} &= \gamma_{24}y_{4t} + \beta_{22}z_{2t} + \beta_{24}z_{4t} + u_{2t} \\
y_{3t} &= \gamma_{32}y_{2t} + \beta_{32}z_{2t} + \beta_{33}z_{3t} + u_{3t} \\
y_{4t} &= \gamma_{42}y_{2t} + \gamma_{43}y_{3t} + \beta_{41}z_{1t} + \beta_{42}z_{2t} + \beta_{44}z_{4t} + u_{4t}.
\end{aligned}$$

First, write in matrix form

$$\begin{array}{cccccccc} y_{1t} & y_{2t} & y_{3t} & y_{4t} & z_{1t} & z_{2t} & z_{3t} & z_{4t} \end{array}$$

$$\begin{bmatrix} 1 & -\gamma_{12} & 0 & 0 & -\beta_{11} & 0 & -\beta_{13} & -\beta_{14} \\ 0 & 1 & 0 & -\gamma_{24} & 0 & -\beta_{22} & 0 & -\beta_{24} \\ 0 & -\gamma_{32} & 1 & 0 & 0 & -\beta_{32} & -\beta_{33} & 0 \\ 0 & -\gamma_{42} & -\gamma_{43} & 1 & -\beta_{41} & -\beta_{42} & 0 & -\beta_{44} \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \\ z_{1t} \\ z_{2t} \\ z_{3t} \\ z_{4t} \end{bmatrix} = \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \end{bmatrix}$$

$N = 4$, $N - 1 = 3$.

<u>*Equation 1*</u>

<u>*Order*</u>: *No. of zeros* $= 3 = N - 1$

<u>*Rank*</u>: $\begin{bmatrix} 0 & -\gamma_{24} & -\beta_{22} \\ 1 & 0 & -\beta_{32} \\ -\gamma_{43} & 1 & -\beta_{42} \end{bmatrix} = 3$ *(note* $\det \neq 0$*)*

So just identified.

<u>*Equation 2*</u>

<u>*Order*</u>: *No. of zeros* $= 4 > N - 1$

<u>*Rank*</u>: $\begin{bmatrix} 1 & 0 & -\beta_{11} & -\beta_{13} \\ 0 & 1 & 0 & -\beta_{33} \\ 0 & -\gamma_{43} & -\beta_{41} & 0 \end{bmatrix} = 3$ *(det. of 1st sub-matrix* $\neq 0$*)*

So over-identified.

<u>*Equation 3*</u>

<u>*Order*</u>: *No. of zeros* $= 4 > N - 1$

$$\underline{\text{Rank:}} \quad \begin{bmatrix} 1 & 0 & -\beta_{11} & -\beta_{14} \\ 0 & -\gamma_{24} & 0 & -\beta_{24} \\ 0 & 1 & -\beta_{41} & -\beta_{44} \end{bmatrix} = 3 \quad \textit{(det. of 1st sub-matrix} \neq 0\textit{)}$$

Equation 4

No. of zeros $= 2 < N - 1$.

So not identified.

Note, if $\beta_{11} = \beta_{13}$*, equivalent to one extra zero in Equation 1 so it becomes over identified.*

Section 2.4 presents a more extensive discussion of identification issues. That section is meant for your *information only*, and will not be examined.

2.2. Estimation

$$\text{Structural Form} \quad : \qquad \Gamma y_t = B x_t + \varepsilon_t \tag{2.3}$$

$$\text{Reduced Form:} \qquad y_t = \Gamma^{-1} B x_t + \Gamma^{-1} \varepsilon_t = \Pi x_t + v_t. \tag{2.4}$$

Consider first equation in the structural from (2.3)

$$y_{1t} = Y_{1t}' \gamma_1 + X_{1t}' \beta_1 + \varepsilon_{1t} \tag{2.5}$$

where Y_{1t}' = row vector of y variables included in first equation excluding y_{1t}, X_{1t}' = row vector of x variables included in first equation. We can rewrite this in matrix form as

$$y_1 = Y_1 \gamma_1 + X_1 \beta_1 + \varepsilon_1 \tag{2.6}$$

where

$$y_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1T} \end{bmatrix}, \qquad Y_1 = \begin{bmatrix} Y_{11}' \\ Y_{12}' \\ \vdots \\ Y_{1T}' \end{bmatrix} \quad \text{etc.}$$

$y_1 = T \times 1$, $Y_1 = T \times (N_1 - 1)$, $X_1 = T \times K_1$, where N_1 is the number of endogenous variables in equation (2.5), K_1 is number of exogenous variables in equation (2.5). Note

$(N - N_1)$ = number of excluded y variables, $K - K_1$ = number of excluded x variables. Let $X_1^T = T \times (K - K_1)$ matrix of excluded exogenous variables. Recall, from (2.4), each y is correlated with all errors. So, to estimate (2.6), we need to use IV (instrumental variables).

We have $(N_1 - 1)$ right hand side endogenous variables correlated with the error. And we have $(K - K_1)$ spare (excluded) X variables to use as instruments. But we need at least as many instruments as endogenous variables, i.e. we need

$$(K - K_1) \geq (N_1 - 1)$$

or equivalently, adding $N - N_1$ on each side

$$(N - N_1) + (K - K_1) \geq N - 1.$$

But the LHS is the number of zeros in first equation – so this is the *order condition* for identification. So, if the equation is identified – i.e. this order condition is satisfied – we can do IV.

Proceeding to estimation, form the matrices

$$Z_1 = [Y_1, X_1], \qquad X = [X_1, X_1^T], \qquad \delta_1 = \begin{pmatrix} \gamma_1 \\ \beta_1 \end{pmatrix}.$$

So the first equation (2.6) can be written

$$y_1 = Z_1 \delta_1 + \varepsilon_1 \tag{2.7}$$

Using X as the set of instruments, the IV estimator is

$$\hat{\delta}_1 = \left(Z_1' X (X'X)^{-1} X' Z_1 \right)^{-1} Z_1' X (X'X)^{-1} X' y_1 \tag{2.8}$$

where

$$\begin{aligned} var(\hat{\delta}_1) &= s_{iv}^2 \left((Z_1' X)(X'X)^{-1}(X' Z_1) \right)^{-1} \\ s_{iv}^2 &= \frac{1}{(T - N_1 - K_1)} \sum_t \left(y_t - Z_{1t}' \hat{\delta}_1 \right)^2. \end{aligned}$$

Note that if $Z_1'X$ is square (i.e. $N_1 + K_1 - 1 = K$, so just identified), then $\hat{\delta}_1$ can be simplified, i.e.

$$\begin{aligned}\hat{\delta}_1 &= \left(Z_1'X(X'X)^{-1}X'Z_1\right)^{-1} Z_1'X(X'X)^{-1}X'y_1 \\ &= (X'Z_1)^{-1}(X'X)(Z_1'X)^{-1}(Z_1'X)(X'X)^{-1}X'y_1 \\ &= (X'Z_1)^{-1}X'y_1. \end{aligned} \tag{2.9}$$

In the case of over-identification, the IV estimator (2.8) is known as Two Stage Least Squares.[3] In the just-identified case, (2.9) is known as Indirect Least Squares.

Notes:

1. The above seems to imply that if an equation passes the order condition then we can estimate it. We know that you cannot estimate an underidentified equation. But the above *seems* to imply that you can estimate an equation which passes the order condition but fails the rank condition. An estimate of the variance of the limiting distribution of $\sqrt{T}(\hat{\delta}_1 - \delta_1)$ is

$$\sigma^2 \left[\frac{1}{T}Z_1'X\left(\frac{1}{T}X'X\right)^{-1}\frac{1}{T}X'Z_1\right]^{-1}.$$

   If an equation fails the rank condition, then in the limit $\frac{1}{T}Z_1'X$ tends to matrix with rank less than the number of columns of Z_1. This in turn makes the matrix in square brackets singular and non-invertible. In finite samples you get very large estimated standard errors; the estimates are very poorly defined.

[3] To see why this is a two stage least square, note that if we regress Z_1 on X we obtain the least square coefficients $(X'X)^{-1}X'Z_1$. We can therefore construct $\hat{Z}_1 := X(X'X)^{-1}X'Z_1$. Regressing then y_1 on $\hat{Z}_1$via least squares we get the coefficients

$$\begin{aligned}\left(\hat{Z}_1'\hat{Z}_1\right)^{-1}\hat{Z}_1 y_1 &= \left[Z_1'X(X'X)^{-1}X'X(X'X)^{-1}X'Z_1\right]^{-1} X(X'X)^{-1}X'Z_1 y_1 \\ &= \left[Z_1'X(X'X)^{-1}X'Z_1\right]^{-1} X(X'X)^{-1}X'Z_1 y \\ &= : \hat{\delta}_1\end{aligned}$$

2. The IV estimator is consistent but not efficient. There are more complex techniques which are asymptotically efficient, in particular Full Information Maximum Likelihood (FIML) and Three Stage Least Squares (3SLS).

   The underlying rationale of ML is exactly as you would expect. 3SLS is basically a feasible GLS technique. It uses 2SLS for the initial estimates, estimates Ω from the residuals and then does a final iteration of GLS based on the estimated Ω.

2.3. Dynamic Models

The standard notation for a simultaneous system can be extended to simultaneous stochastic difference models,

$$A_0 y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + B_0 x_t + ... + B_s x_{t-s} + \varepsilon_t. \tag{2.10}$$

Note that lagged y variables are contemporaneously independent so long as ε_t is iid. Many results from the static framework carry through. If the ε_t are serially independent, then the standard criteria for identification are valid. The reduced form is

$$y_t = A_0^{-1} A_1 y_{t-1} + ... + A_0^{-1} A_p y_{t-p} + A_0^{-1} B_0 x_t + ... + A_0^{-1} B_s x_{t-s} + A_0^{-1} \varepsilon_t. \tag{2.11}$$

Note that the lagged y variables are treated as exogenous variables in the previous model. Clearly (2.10) can be re-expressed using lag polynomials,

$$A(L) y_t = B(L) x_t + \varepsilon_t.$$

The key question is whether $A(L)$ can be inverted. In general

$$A^{-1}(L) = \frac{A^*(L)}{|A(L)|} \tag{2.12}$$

where $A^*(L)$ is the adjoint of A, the transposed matrix of cofactors.[4] For $A(L)$ to be invertible, $|A(L)| \neq 0$. If $A(L)$ is invertible, (2.10) can be written as

$$y_t = A^{-1}(L) B(L) x_t + A^{-1}(L) \varepsilon_t. \tag{2.13}$$

[4] A review of some linear algebra definition is presented in the appendix.

This is the Final Form and it expresses y_t as a function of current and lagged exogenous variables and errors only (these lags are usually infinite). (2.13) can also be written as

$$|A(L)|\, y_t = A^*(L)B(L)x_t + A^*(L)\varepsilon_t. \tag{2.14}$$

This is the Autoregressive Final Form. Now $|A(L)|$ is just a scalar polynomial of order $p \times N$. Each equation in (2.14) is a stochastic difference equation in lagged values of the equation dependent variable, current and lagged values of the exogenous variables, and the vector of errors. Further in all N equations the coefficients on the lagged dependent variables are the same. This implies a common pattern of dynamic behavior for the endogenous variables in the system.

In particular it implies that if one equation is stable, then the whole system is stable. As usual the condition for stability is that the characteristic equation associated with the lag polynomial $|A(L)|$ should have roots with absolute values less than one.

In view of the importance of stability, one of the first things that you want to know about any dynamic system is its autoregressive final form and hence the roots of $|A(L)|$. How do we do this? Two ways: one is to do the formal matrix algebra. However in small systems without too many lags it is sometimes easier to solve by hand.

Example 2. *Consider the simple income-expenditure model in obvious notation,*

$$c_t = \beta_{11} + \gamma_{13} y_t + \gamma_{14} y_{t-1} + \varepsilon_{1t} \tag{2.15}$$

$$i_t = \beta_{21} + \beta_{22} i_{t-1} + \varepsilon_{2t} \tag{2.16}$$

$$y_t = c_t + i_t. \tag{2.17}$$

Identification

c_t	i_t	y_t	y_{t-1}	i_{t-1}	
1	0	$-\gamma_{13}$	$-\gamma_{14}$	0	$N = 3$
0	1	0	0	$-\beta_{22}$	
-1	-1	1	0	0	

(2.15) *Order: Zeros* $= 2 = N - 1$. *Order condition satisfied.*

$$\textit{Rank} = \begin{bmatrix} 1 & -\beta_{22} \\ -1 & 0 \end{bmatrix} = 2. \qquad \textit{Rank condition satisfied.}$$

Just identified.

(2.16) *Order: Zeros* $= 3 > N - 1$. *Order condition satisfied.*

$$\textit{Rank} = \begin{bmatrix} 1 & -\gamma_{13} & -\gamma_{14} \\ -1 & 1 & 0 \end{bmatrix} = 2. \qquad \textit{Rank condition satisfied.}$$

Over identified.

(2.17) *Identify (it's an identity).*

To derive the reduced form, substitute (2.15), (2.16) into (2.17).

$$\begin{aligned} y_t &= \beta_{11} + \gamma_{13} y_t + \gamma_{14} y_{t-1} + \varepsilon_{1t} \\ &\quad + \beta_{21} + \beta_{22} i_{t-1} + \varepsilon_{2t} \end{aligned}$$

or

$$y_t = \frac{\beta_{11} + \beta_{21}}{1 - \gamma_{13}} + \frac{\gamma_{14}}{1 - \gamma_{13}} y_{t-1} + \frac{\beta_{22}}{1 - \gamma_{13}} i_{t-1} + \frac{\varepsilon_{1t} + \varepsilon_{2t}}{1 - \gamma_{13}} \tag{2.18}$$

$$i_t = \beta_{21} + \beta_{22} i_{t-1} + \varepsilon_{2t}. \tag{2.19}$$

From the identity,

$$c_t = y_t - i_t,$$

using (2.18), (2.19), we have

$$c_t = \frac{\beta_{11} + \beta_{21}\gamma_{13}}{1 - \gamma_{13}} + \frac{\gamma_{14}}{1 - \gamma_{13}} y_{t-1} + \frac{\beta_{22}\gamma_{13}}{1 - \gamma_{13}} i_{t-1} + \frac{\varepsilon_{1t} + \varepsilon_{2t}\gamma_{13}}{1 - \gamma_{13}}. \tag{2.20}$$

These are the reduced form equations. To estimate (2.15), use i_{t-1} *as instrument for* y_t *which is correlated with* ε_{1t}. *(2.16) is already a reduced form equation, so it may be estimated by OLS.*

To estimate the stability of the model, generate a condensed model consisting of only y *and* i *equations, eliminating* c_t *(and, in general, all its lags, although there are no lags of* c *in this model).*

So substituting (2.15) into (2.17), we have a condensed model,

$$y_t = \beta_{11} + \gamma_{13} y_t + i_t + \gamma_{14} y_{t-1} + \varepsilon_{1t}$$

or

$$\begin{aligned} y_t &= \frac{\beta_{11}}{(1-\gamma_{13})} + \frac{1}{(1-\gamma_{13})} i_t + \frac{\gamma_{14}}{(1-\gamma_{13})} y_{t-1} + \frac{\varepsilon_{1t}}{(1-\gamma_{13})} \\ i_t &= \beta_{21} + \beta_{22} i_{t-1} + \varepsilon_{2t}. \end{aligned} \tag{2.21}$$

Writing this condensed model in matrix lag operator form gives

$$\underbrace{\begin{bmatrix} 1 - \frac{\gamma_{14}L}{(1-\gamma_{13})} & \frac{-1}{1-\gamma_{13}} \\ 0 & (1-\beta_{22}L) \end{bmatrix}}_{A(L)} \begin{bmatrix} y_t \\ i_t \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\beta_{11}}{1-\gamma_{13}} \\ \beta_{21} \end{bmatrix}}_{B(L)} + \begin{bmatrix} \frac{\varepsilon_{1t}}{1-\gamma_{13}} \\ \varepsilon_{2t} \end{bmatrix}$$

Note that the x variables include only the constant terms in this case. The determinant,

$$\begin{aligned} |A(L)| &= (1-\beta_{22}L)\left(1 - \frac{\gamma_{14}L}{(1-\gamma_{13})}\right) \\ &= 1 - \left[\beta_{22} + \frac{\gamma_{14}}{1-\gamma_{13}}\right] L + \frac{\beta_{22}\gamma_{14}}{(1-\gamma_{13})} L^2. \end{aligned}$$

The adjoint matrix

$$A^*(L) = \begin{bmatrix} (1-\beta_{22}L) & \frac{1}{1-\gamma_{13}} \\ 0 & 1 - \frac{\gamma_{14}L}{(1-\gamma_{13})} \end{bmatrix}.$$

So the Autoregressive Final Form, based on (2.14), is

$$\left(1 - \left(\beta_{22} + \frac{\gamma_{14}}{1-\gamma_{13}}\right) L + \frac{\beta_{22}\gamma_{14}}{(1-\gamma_{13})} L^2\right) \begin{bmatrix} y_t \\ i_t \end{bmatrix} =$$

$$\begin{bmatrix} (1-\beta_{22}L) & \frac{1}{1-\gamma_{13}} \\ 0 & 1 - \frac{\gamma_{14}L}{(1-\gamma_{13})} \end{bmatrix} \begin{bmatrix} \frac{\beta_{11}}{(1-\gamma_{13})} \\ \beta_{21} \end{bmatrix} + \begin{bmatrix} (1-\beta_{22}L) & \frac{1}{1-\gamma_{13}} \\ 0 & 1 - \frac{\gamma_{14}L}{(1-\gamma_{13})} \end{bmatrix} \begin{bmatrix} \frac{\varepsilon_{1t}}{1-\gamma_{13}} \\ \varepsilon_{2t} \end{bmatrix}.$$

Multiplying out gives

$$\begin{aligned} y_t &= \left(\beta_{22} + \frac{\gamma_{14}}{1-\gamma_{13}}\right) y_{t-1} - \frac{\beta_{22}\gamma_{14}}{(1-\gamma_{13})} y_{t-2} + \frac{(1-\beta_{22})}{(1-\gamma_{13})}\beta_{11} + \frac{\beta_{21}}{(1-\gamma_{13})} + \frac{\varepsilon_{1t}}{(1-\gamma_{13})} \\ &\quad - \frac{\beta_{22}\varepsilon_{1t-1}}{(1-\gamma_{13})} + \frac{\varepsilon_{2t}}{(1-\gamma_{13})} \\ i_t &= \left(\beta_{22} + \frac{\gamma_{14}}{1-\gamma_{13}}\right) i_{t-1} - \frac{\beta_{22}\gamma_{14}}{(1-\gamma_{13})} i_{y-2} + \left(1 - \frac{\gamma_{14}}{1-\gamma_{13}}\right)\beta_{21} + \varepsilon_{2t} - \frac{\gamma_{14}}{(1-\gamma_{13})}\varepsilon_{2t-1}. \end{aligned}$$

The model is stable if the roots of $A(L)$ are less than 1 in absolute value. From $|A(L)|$, we see that the roots satisfy the quadratic equation

$$\lambda^2 - \left(\beta_{22} + \frac{\gamma_{14}}{1-\gamma_{13}}\right)\lambda + \frac{\beta_{22}\gamma_{14}}{(1-\gamma_{13})} = 0.$$

In this case, this factorises and the roots are $\frac{\gamma_{14}}{1-\gamma_{13}}$, β_{22}. *So the model is stable iff* $\left|\frac{\gamma_{14}}{1-\gamma_{13}}\right| < 1$, $|\beta_{22}| < 1$.

2.4. More on Identification

[This section is for information only. Details will not be examined.]

A system of simultaneous equations can be written,

$$\Gamma y_t = Bx_t + \varepsilon_t \qquad t = 1, ..., T$$

where there are N equations and K exogenous variables. Thus Γ is $N \times N$, with diagonal elements normalized to 1, and B, $N \times K$. Γ is assumed non-singular, hence the reduced form is

$$\begin{aligned} y_t &= \Gamma^{-1}Bx_t + \Gamma^{-1}\varepsilon_t \\ &= \Pi x_t + \nu_t \end{aligned}$$

Π is an $N \times K$ matrix. The identification problem arises because the most that can be determined from data on y_t and x_t are the elements of Π and the variance covariance matrix of the ν's. Assuming that nothing is known about the covariance matrix of the ε's and hence of the ν's, the problem is whether it is possible to solve back from Π to obtain estimates of the structural coefficients, Γ and B. From the definition of the reduced form

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

and

$$B = \Gamma\Pi. \tag{2.22}$$

Thus there are $N \times K$ equations in $N \times K + N \times (N-1)$ unknowns (allowing for the restrictions on the diagonal elements in Γ). As things stand there are far more unknowns than equations and hence if the elements of Γ and B are unrestricted he structural coefficients are underidentified and cannot be recovered.

However, it is usually the case that the elements of Γ and B are subject to restrictions, in particular to *exclusion restrictions*: we know on *a priori* grounds that some variables do not occur in a particular equation (economic theory can provide guidance in choosing the exclusion restrictions). So the problem becomes, given a set of restrictions can the structural coefficients be identified. This problem can be tackled equation by equation. So consider the first equation. Transposing the first row of restrictions from (2.22),

$$\begin{array}{llll} \beta_{11} & = \pi_{11}\gamma_{11} + & \cdots & +\pi_{N1}\gamma_{1N} \\ \beta_{12} & = \pi_{12}\gamma_{11} + & \cdots & +\pi_{N2}\gamma_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{1K} & = \pi_{1K}\gamma_{11} + & \cdots & +\pi_{NK}\gamma_{1N} \end{array} .$$

Rearranging and using the normalizing restriction $\gamma_{11} = 1$

$$\begin{array}{llllll} -\pi_{11} = & -\beta_{11} & & +\pi_{21}\gamma_{12} + & \cdots & \pi_{N1}\gamma_{1N} \\ -\pi_{12} = & & -\beta_{12} & +\pi_{22}\gamma_{12} + & \cdots & \pi_{N2}\gamma_{1N} \\ \vdots & & \ddots & \vdots & & \vdots \\ -\pi_{1K} = & & -\beta_{1K} & +\pi_{2K}\gamma_{12} + & \cdots & \pi_{NK}\gamma_{1N} \end{array} . \tag{2.23}$$

This gives us K linear non-homogeneous equations in $K + N - 1$ unknowns.

Now recall, for a system of non-homogeneous linear equations in n unknowns, $Bx = r$ has a unique solution $x = B^{-1}r$ if and only if B contains n linearly independent equations, otherwise B is not invertible. If there are fewer than n equations, there will be multiple solutions. If there are more than n linearly independent equations, the system will in general be inconsistent with no solutions.

Thus, referring back to (2.23) if the coefficients of the first rows of Γ and B are unrestricted, the structural coefficients cannot be recovered; there are far too few equations. However if we can find another $N-1$ linearly independent restrictions, then these, together with the K restrictions in (2.23), implied by the reduced form, form a system which can be solved for the structural coefficients.

Expressing (2.23) in matrix terms, let β_1' be the first row of B and γ_1' be the first row of Γ. The first element of γ_1 is unity. Define γ_1^* to omit the first element,

$$\gamma_1 = \begin{bmatrix} 1 \\ \cdots \\ \gamma_1^* \end{bmatrix} .$$

Define

$$\alpha_1 = \begin{bmatrix} \gamma_1 \\ \cdots \\ \beta_1 \end{bmatrix}, \qquad \alpha_1^* = \begin{bmatrix} \gamma_1^* \\ \cdots \\ \beta_1 \end{bmatrix}.$$

Partition Π so that

$$\Pi = \begin{bmatrix} \pi_1' \\ \cdots \\ \Pi_2' \end{bmatrix}, \qquad \Pi_2 \text{ is } K \times (N-1).$$

Then (2.23) can be rewritten,

$$-\pi_1 = [\Pi_2 : -I_K]\, \alpha_1^*.$$

Exclusion Restrictions The simplest way to express an exclusion restriction in matrix notation is to use a selection vector. Thus the restriction

$$\gamma_{1j} = 0$$

can be expressed in vector notation as

$$\alpha_1' \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0.$$

If there are q such restrictions then the set can be written as

$$\alpha_1' \Phi = 0$$

where Φ is $(N+K) \times q$ matrix of selection vectors.

Some points:

1. Linear homogeneous restrictions involving several coefficients are easily accommodated within the framework.

2. γ_{11} is already restricted to be one, hence it cannot be subject to an exclusion restriction. On the other hand restricting a second coefficient to be a linear function of γ_{11} generates an implicit non-homogeneous restriction.

3. The q columns of Φ are assumed linearly independent. If the columns are linearly dependent it implies that the same restriction has been entered twice and hence that at least one restriction is redundant.

Partition

$$\Phi = \begin{bmatrix} \phi_1' \\ \cdots \\ \Phi_2' \end{bmatrix}.$$

If restrictions are homogeneous $\phi_1 = 0$.

Adding the exclusion restrictions to (2.23)

$$\begin{bmatrix} -\pi_1 \\ -\phi_1 \end{bmatrix} = \begin{bmatrix} \Pi_2 : -I_K \\ \Phi_2 \end{bmatrix} \alpha_1^*. \tag{2.24}$$

It is straightforward from (2.24) to obtain the basic condition for identification: that the central matrix

$$\begin{bmatrix} \Pi_2 : -I_K \\ \Phi_2 \end{bmatrix} \tag{2.25}$$

should be non-singular. As this matrix has $K + N - 1$ columns a necessary condition for invertibility is that (2.25) is square and has $K + N - 1$ rows. This in turn implies that Φ_2 must have $N - 1$ rows and hence that there are $N - 1$ exclusion restrictions. This is the order condition: the number of variables excluded from an equation must be at least as great as the number of equations minus one. The necessary and sufficient condition for identification of the first equation is the rank condition; the matrix (2.25) must be square and of rank $K + N - 1$.

What happens if there are more than $N - 1$ exclusion restrictions? In this case the number of restrictions on Γ and B is so large that they imply restrictions on Π. Such a situation is called over-identification. In this case the restrictions appear not only in the Φ matrix but also in the Π matrix. Thus the same restriction appears twice. Once the redundant restrictions are removed you are left with $K + N - 1$ rows.

If you examine (2.25) you will see that checking the rank condition is something of a nightmare. In practice we use an alternative and equivalent (and slightly less nightmarish) condition. The first equation passes the rank condition if rank $(A\Phi) = N - 1$, where

$A = [\Gamma : -B]$. Assuming that Φ contains $(N-1)$ restrictions then $A\Phi$ is $N \times (N-1)$ and is the product of the matrix of coefficients with the selection matrix. Its columns are the columns of the coefficient matrix that are subject to exclusion restrictions in the first equation. As a result the row associated with the current equation, in our case the first will be identically zero. The equation passes the rank condition if the remaining $(N-1)\times(N-1)$ matrix is non-singular.

If Φ contains more than $(N-1)$ restrictions then provided that $A\Phi$ contains one sub-matrix of rank $(N-1)$ the equation is identified. If the equation is identified and Φ contains more than $(N-1)$ restrictions then the system is over-identified.

3. Appendix

Here we recall some basic definitions from linear algebra.

Minors of a matrix. Given a $n \times n$ matrix A, consider the A_{ij} matrix obtained by deleting row i and column j. The $(i,j) - th$ *minor* of a A is defined as

$$M_{ij} = \det A_{ij}.$$

Cofactors of a matrix. The $(i,j) - th$ *cofactor* of a matrix A is defined as

$$C_{ij} = (-1)^{i+j} M_{ij}$$

where M_{ij} is the $(i,j) - th$ minor.

Adjoint matrix Given a $n \times n$ matrix A, the *adjoint matrix* of A ($adj\,(A)$) is the $n \times n$ matrix whose $(i,j) - th$ entry is the $(j,i) - th$ cofactor of A.

We can now state the following Theorem from linear algebra:

Theorem. *For any $n \times n$ matrix A, we have that*

$$A \times adj\,(A) = \det A \times I$$

where I is the $n \times n$ identity matrix.

This implies that, if A is invertible

$$A^{-1} = \frac{adj\,(A)}{\det A}.$$