Simultaneous (and Recursive) Equation Systems

Robert Dixon
Department of Economics at the University of Melbourne

In their History of Macroeconometric Model-Building, Bodkin, Klein and Marwah give "pride of place" to Jan Tinbergen (1991, p 31). He was the first to estimate and use a system of equations to depict economic behaviour and to evaluate policy options. (Tinbergen was attempting to model the macro-behaviour of the Dutch and US economy in the twenties and early thirties.) Since that time a great deal of research effort has been devoted to the theory of estimating simultaneous equations, to the construction of large-scale macroeconometric (and microeconometric) models and to the development of algorithms which will solve such models (especially when they contain non-linearities).

We are familiar with the usual single equation regression model in which one variable (y) is the dependent or determined variable and x_1, x_2, ..., etc are the independent or determining variables. Simultaneous equation models are fundamentally different to single equation models. In simultaneous equation models there are a number of exogenous variables and many endogenous or jointly determined variables. Variables that are considered to be jointly dependent on each other while being affected by some other variables are called *endogenous* variables. These variables are the object of the explanation sought by a simultaneous equation system. Some endogenous variables affect the jointly dependent variables with a lag; they are called *lagged endogenous variables*. The variables that enter the system and affect the endogenous variables, but are not affected by them, are called *exogenous* variables. Typically, in a simultaneous equation model, variables which are on the L.H.S. of some equations in the system will appear on the R.H.S. of other equations in the system. In a truly simultaneous system all of the left-hand-side variables are jointly determined.

It is customary to call exogenous (whether current or lagged) and lagged endogenous variables together 'predetermined variables'. Thus any model will have endogenous variables and predetermined variables. The latter being made up of (truly) exogenous variables and lagged endogenous variables. Lagged values of endogenous variables are treated as exogenous variables because for determination of the current period's values of the endogenous variables they are 'pre' in the sense of being prior in time or as given from outside the model. For this reason the exogenous and lagged endogenous variables are often called *predetermined* variables.

We call the individual equations which make up the simultaneous equation system, '*structural equations'*. A structural equation can be specified to include endogenous and predetermined variables of the system. The variables that are not included in an equation can be said to be present in the equation but with a coefficient of zero. If there are 'g' endogenous variables, there must be 'g' equations to explain these variables. A model is complete when there are as many structural equations as there are endogenous variables, so that the endogenous variables can be solved for in terms of the predetermined variables.

Structural relations or equations may be classified into four types: (1) behavioural, such as demand and supply equations, (2) 'technological', such as production functions (but note the inverted commas around technological), (3) institutional, such as income tax-income relations, and (4) identities, these are mere definitional or accounting statements.

Consider a linear system of 'g' structural difference equations which relate a vector of 'g' endogenous variables (Y) to a vector of 'k' exogenous variables (X). We will allow the endogenous variables to enter also with a lag. (Thus we have two sets of predetermined variables: the exogenous variables and the lagged endogenous variables.) For the moment, we will assume for simplicity that all exogenous variables enter without a lag and that endogenous variables enter with only a one-period lag.

In matrix notation, such a system (after we set the random disturbances equal to their expected value of zero) may be written as:

$$\mathbf{A}\mathbf{Y}_t = \mathbf{B_1}\mathbf{X}_t + \mathbf{B_2}\mathbf{Y}_{t-1}$$

where **A** is a g x g matrix of the (structural) coefficients on the current endogenous variables, $\mathbf{B_1}$ is a g x k matrix of the (structural) coefficients on the current exogenous variables and $\mathbf{B_2}$ is a g x g matrix of the coefficients of the lagged endogenous variables.

It may help if we go through some examples so that you can see exactly what these matrices (**A**, $\mathbf{B_1}$ and $\mathbf{B_2}$) are. [But N.B. we need to distinguish carefully between the vector of current values of the endogenous variables denoted by the symbol **Y** and the symbol for the single variable 'national income' or GDP which is denoted as Y.]

<u>First Example</u>

Consider first the structural equations of a simple keynesian model:

$$C_t = \alpha_0 + \alpha_1(Y_t - T_t)$$

$$I_t = \beta_0 + \beta_1(Y_t - Y_{t-1})$$

$$T_t = \gamma_0 + \gamma_1 Y_t$$

$$Y_t = C_t + I_t + G_t$$

where C = consumption expenditure
I = investment
T = taxes
Y = income
G = government expenditure (exogenous)

In this model the endogenous variables are C_t, I_t, T_t and Y_t. The predetermined variables are G (an exogenous variable) and Y_{t-1}, a lagged endogenous variable.

We will now proceed to write this simultaneous equation system in matrix form. The first step is to rearrange all of the equations of the system so that all terms involving (current values of) the endogenous variables appear on the left hand side (and in such a way as to preserve the ordering of the variables) and all predetermined variables appear on the right. If we do this we have:

$$C_t + \alpha_1 T_t - \alpha_1 Y_t = \alpha_0$$

$$I_t - \beta_1 Y_t = \beta_0 - \beta_1 Y_{t-1}$$

$$T_t - \gamma_1 Y_t = \gamma_0$$

$$Y_t - C_t - I_t = G_t$$

In this form the system can be given matrix representation as follows (note that I have included the constants and that I regard them as related to an exogenous 'variable' which has the value of unity):

$$\begin{bmatrix} 1 & 0 & \alpha_1 & -\alpha_1 \\ 0 & 1 & 0 & -\beta_1 \\ 0 & 0 & 1 & -\gamma_1 \\ -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} C_t \\ I_t \\ T_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \alpha_0 & 0 \\ \beta_0 & 0 \\ \gamma_0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ G \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\beta_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_{t-1} \\ I_{t-1} \\ T_{t-1} \\ Y_{t-1} \end{bmatrix}$$

Our system of structural equations is now in the form:

$$\mathbf{AY}_t = \mathbf{B_1 X}_t + \mathbf{B_2 Y}_{t-1}$$

Longer Lags

Note, as an aside, that sometimes lagged endogenous variables enter with a lag greater than one and that the exogenous variables may enter with a lag. Suppose, as an example, that we have a model with current exogenous variables and with endogenous variables entering with both a single period lag and with a two-period lag. In this event the system could be represented by the following:

$$\mathbf{AY}_t = \mathbf{B_1 X}_t + \mathbf{B_2 Y}_{t-1} + \mathbf{B_3 Y}_{t-2}$$

A Second Example

Consider the following multiplier-accelerator model:

$$C_t = 0.7Yt$$

$$I_t = 2.5(Y_{t-1} - Y_{t-2})$$

$$Y_t = C_t + I_t + G_t$$

Notice that we have three endogenous variables (C_t, I_t and Y_t), one exogenous variable (G) and two lagged endogenous variables (Y_{t-1} and Y_{t-2}).

To write this model in matrix notation we proceed as before, by first collecting all endogenous variables on the L.H.S. of the structural equations. This results in:

$$C_t - 0.7Y_t = 0$$

$$I_t = 2.5Y_{t-1} - 2.5Y_{t-2}$$

$$Y_t - C_t - I_t = G_t$$

In matrix form we now have:

$$\begin{bmatrix} 1 & 0 & -.7 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} C_t \\ I_t \\ Y_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} [G] + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2.5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_{t-1} \\ I_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -2.5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_{t-2} \\ I_{t-2} \\ Y_{t-2} \end{bmatrix}$$

which is of the form

$$\mathbf{AY}_t = \mathbf{B_1X}_t + \mathbf{B_2Y}_{t-1} + \mathbf{B_3Y}_{t-2}$$

Recursive Models

A system of equations is recursive rather than simultaneous if each of the endogenous variables can be determined sequentially rather than jointly. To put the same point a little differently: a recursive model is one where the matrix of coefficients of the endogenous variables ie. the matrix **A** in our notation) is triangular. By 'triangular' we mean that the matrix has zero's in all of the cells above the diagonal.

Because the **A** matrix is triangular, a recursive model is one in which there is unidirectional dependency among the endogenous variables. The equations can be ordered such that the first endogenous variable is determined only by exogenous variables, the second is determined only by the first endogenous variable and exogenous variables, the third by only the first two endogenous variables and exogenous variables, and so forth. There must be no feedback from an endogenous variable to one lower in the casual chain.

To see the nature of these models consider the following system:

$$Y_{1t} = \gamma_{11}X_{1t}$$

$$Y_{2t} = \gamma_{21}X_{1t} + \beta_{21}Y_{1t}$$

$$Y_{3t} = \gamma_{31}X_{1t} + \beta_{31}Y_{1t} + \beta_{32}Y_{2t}$$

where, as usual, the Y's and the X's are, respectively, the endogenous and exogenous variables.

From the structure of this system, it is clear that there is a sequential (one-way) rather than a simultaneous (ie. two-way) relationship between the endogenous variables. Thus, Y_1 affects Y_2, but Y_2 does not affect Y_1. Similarly, Y_1 and Y_2 influence Y_3 without, in turn, being influenced by Y_3. In other words, each equation exhibits a unilateral casual dependence, hence they are called recursive, triangular or casual models.

Notice that if we form the matrix of coefficients on the endogenous variables (ie. the **A** matrix) by grouping all of the endogenous variables on the L.H.S. of the equations, the **A** matrix would be written as:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -\beta_{21} & 1 & 0 \\ -\beta_{31} & -\beta_{32} & 1 \end{bmatrix}$$

Many models, and especially models used for teaching purposes, are recursive.

On the next page you will find an example of a recursive model taken from an econometrics textbook.

An Example of a Recursive Model

Gujarati (1995, p. 681) gives an example of a recursive system. He postulates the following model of wage and price determination:

$$\dot{P}_t = \beta_{10} + \beta_{11}\dot{W}_{t-1} + \beta_{12}\dot{R}_t + \beta_{13}\dot{M}_t + \beta_{14}\dot{L}_t$$

$$\dot{W}_t = \beta_{20} + \beta_{21}U_1 + \beta_{22}\dot{P}_t$$

where $\dot{P}$ = rate of change of price per unit of output

$\dot{W}$ = rate of change of wages per employee

$\dot{R}$ = rate of change of price of capital

$\dot{M}$ = rate of change of import prices

$\dot{L}$ = rate of change of labour respectively

U = unemployment rate, %

The price equation postulates that the rate of change of price in the current period is a function of the rates of change in the prices of capital and of raw materials, together with the rate of change in labour productivity and rate of change in wages in the previous period. The wage equation shows that the rate of change in wages in the current period, thes rate of change in price and the unemployment rate. It is clear that the causal chain runs from

$$\dot{W}_{t-1} \rightarrow \dot{P}_t \rightarrow \dot{W}_t.$$

In simultaneous equation matrix notation this system maybe written as:

$$\begin{bmatrix} 1 & 0 \\ -\beta_{22} & 1 \end{bmatrix} \begin{bmatrix} \dot{P} \\ \dot{W} \end{bmatrix} = \begin{bmatrix} \beta_{10} & \beta_{12} & \beta_{13} & \beta_{14} & 0 \\ \beta_{20} & 0 & 0 & 0 & \beta_{21} \end{bmatrix} \begin{bmatrix} 1 \\ \dot{R}_t \\ \dot{M}_t \\ \dot{L}_t \\ \dot{U}_t \end{bmatrix} + \begin{bmatrix} 0 & \beta_{11} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{P}_{-1} \\ \dot{W}_{-1} \end{bmatrix}$$

Notice that the **A** matrix is triangular, in other words, that the system is recursive.

References:

R. Bodkin, L. Klein and K. Marwah, A History of Macroeconometric Model-Building, Edward Elgar, 1991.

D. Challen & A. Hagger, Modelling the Australian Economy, Longman Cheshire, 1979, Ch. 6

D. Challen & A. Hagger, Macroeconometric Systems: Construction, Validation and Applications, Macmillan, 1983. Ch. 2.

D. Gujarati, Basic Econometrics, Third edition, McGraw-Hill, 1995.

R. Pindyck & D. Rubinfeld, Econometric Models and Economics Forecasts, Fourth edition, Prentice-Hall, 1998, Chapters 12 - 14.

J. Tinbergen, 'An Economic Policy for 1936' in, H. Klaassen (et al) eds, Jan Tinbergen Selected Papers, North Holland, 1959, pp. 37-84.

J. Tinbergen, Statistical Testing of Business-Cycle Theories, League of Nations, 1939.