THE CLASSICAL SIMULTANEOUS-EQUATION MODEL AND THE 2SLS ESTIMATION OF A STRUCTURAL EQUATION

The classical simultaneous-equation model of econometrics is a system of M structural equations which may be compiled to give the following equation:

$$\begin{aligned}(1)\qquad &[y_{t1}, y_{t2}, \ldots, y_{tM}][\gamma_{.1}, \gamma_{.2}, \ldots, \gamma_{.M}] + x_{t.}[\beta_{.1}, \beta_{.2}, \ldots, \beta_{.M}]\\ &\qquad + [\varepsilon_{t1}, \varepsilon_{t2}, \ldots, \varepsilon_{tM}] = [0, 0, \ldots, 0],\end{aligned}$$

This can be written in summary notation as

$$(2)\qquad y_{t.}\Gamma + x_{t.}B + \varepsilon_{t.} = 0,$$

where $\Gamma = [\gamma_{.1}, \gamma_{.2}, \ldots, \gamma_{.j}]$. The elements of the vector $\varepsilon_{t.} = [\varepsilon_{t1}, \varepsilon_{t2}, \ldots, \varepsilon_{tM}]$, of the M structural disturbances are assumed to be distributed independently of time such that, for every t, there are

$$(3)\qquad E(\varepsilon_{t.}) = 0 \quad \text{and} \quad D(\varepsilon_{t.}) = E(\varepsilon_{t.}'\varepsilon_{t.}) = \Sigma_{\varepsilon\varepsilon}.$$

It is also assumed that the structural disturbances are distributed independently of the exogenous variables so that $C(\varepsilon_{t.}, x_{s.}) = 0$ for all t and s.

The reduced form of the system is obtained from equation (2) by postmultiplying it by the inverse of the matrix Γ. This gives

$$(6)\qquad y_{t.} = x_{t.}\Pi + \eta_{t.} \quad \text{with} \quad \Pi = -B\Gamma^{-1} \quad \text{and} \quad \eta_{t.} = -\varepsilon_{t.}\Gamma^{-1}.$$

It follows that the vector $\eta_{t.} = -\varepsilon_{t.}\Gamma^{-1}$ of reduced-form disturbances has

$$(7)\qquad E(\eta_{t.}) = 0 \quad \text{and} \quad D(\eta_{t.}) = \Gamma'^{-1}D(\varepsilon_{t.})\Gamma^{-1} = \Gamma'^{-1}\Sigma_{\varepsilon\varepsilon}\Gamma^{-1} = \Omega.$$

It is assumed that the statistical properties of the data can be described completely in terms of its first and second moments. The dispersion matrices of $x_{t.}$ and $y_{t.}$can be denoted by $D(x_{t.}) = \Sigma_{xx}$ and $D(y_{t.}) = \Sigma_{yy}$ and their covariance matrix by $C(x_{t.}, y_{t.}) = \Sigma_{xy}$. By combining the reduced-form regression relationship of (6) with a trivial identity in $x_{t.}$, we get the following system:

$$(8)\qquad [\,y_{t.} \quad x_{t.}\,]\begin{bmatrix} I & 0\\ -\Pi & I\end{bmatrix} = [\,\eta_{t.} \quad x_{t.}\,].$$

Given the assumptions that $D(\eta_{t.}) = \Omega$ and that $C(\eta_{t.}, x_{t.}) = 0$, it follows that

$$(9)\qquad \begin{bmatrix} I & -\Pi'\\ 0 & I\end{bmatrix}\begin{bmatrix} \Sigma_{yy} & \Sigma_{yx}\\ \Sigma_{xy} & \Sigma_{xx}\end{bmatrix}\begin{bmatrix} I & 0\\ -\Pi & I\end{bmatrix} = \begin{bmatrix} \Omega & 0\\ 0 & \Sigma_{xx}\end{bmatrix}.$$

Premultiplying this system by the inverse of the leading matrix gives an equivalent equation in the form of

$$(10)\qquad \begin{aligned}\begin{bmatrix}\Sigma_{yy} & \Sigma_{yx}\\ \Sigma_{xy} & \Sigma_{xx}\end{bmatrix}\begin{bmatrix}I & 0\\ -\Pi & I\end{bmatrix} &= \begin{bmatrix}I & \Pi'\\ 0 & I\end{bmatrix}\begin{bmatrix}\Omega & 0\\ 0 & \Sigma_{xx}\end{bmatrix}\\ &= \begin{bmatrix}\Omega & \Pi'\Sigma_{xx}\\ 0 & \Sigma_{xx}\end{bmatrix}.\end{aligned}$$

From this system, the equations $\Sigma_{yy} - \Sigma_{yx}\Pi = \Omega$ and $\Sigma_{xy} - \Sigma_{xx}\Pi = 0$ may be extracted, from which are obtained the parameters that characterise the reduced-from relationship:

$$(11)\qquad \Pi = \Sigma_{xx}^{-1}\Sigma_{xy} \qquad \text{and} \qquad \Omega = \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}.$$

These parameters can be estimated provided that the empirical counterparts of the moment matrices Σ_{xx}, Σ_{yy} and Σ_{xy} are available in the form of $M_{xx} = T^{-1}\sum_t x'_{t.}x_{t.}$, $M_{yy} = T^{-1}\sum_t y'_{t.}y_{t.}$ and $M_{xy} = T^{-1}\sum_t x'_{t.}y_{t.}$.

Now consider combining the structural equation of (2) with a trivial identity to form the counterpart of equation (8). This is the equation

$$(12)\qquad [\,y_{t.} \quad x_{t.}\,]\begin{bmatrix}\Gamma & 0\\ B & I\end{bmatrix} = [\,\varepsilon_{t.} \quad x_{t.}\,].$$

Given that $D(\varepsilon) = \Sigma_{\varepsilon\varepsilon}$ and that $C(\varepsilon, x) = 0$, it follow that

$$(13)\qquad \begin{bmatrix}\Gamma' & B'\\ 0 & I\end{bmatrix}\begin{bmatrix}\Sigma_{yy} & \Sigma_{yx}\\ \Sigma_{xy} & \Sigma_{xx}\end{bmatrix}\begin{bmatrix}\Gamma & 0\\ B & I\end{bmatrix} = \begin{bmatrix}\Sigma_{\varepsilon\varepsilon} & 0\\ 0 & \Sigma_{xx}\end{bmatrix},$$

and, from this, an equivalent expression can be obtained in the form of

$$(14)\qquad \begin{aligned}\begin{bmatrix}\Sigma_{yy} & \Sigma_{yx}\\ \Sigma_{xy} & \Sigma_{xx}\end{bmatrix}\begin{bmatrix}\Gamma & 0\\ B & I\end{bmatrix} &= \begin{bmatrix}\Gamma'^{-1} & \Pi'\\ 0 & I\end{bmatrix}\begin{bmatrix}\Sigma_{\varepsilon\varepsilon} & 0\\ 0 & \Sigma_{xx}\end{bmatrix}\\ &= \begin{bmatrix}\Omega\Gamma & \Pi'\Sigma_{xx}\\ 0 & \Sigma_{xx}\end{bmatrix}.\end{aligned}$$

This identity, together with those of (11), provides the fundamental equations that relate the structural parameters Γ, B to the moment matrices of the data variables:

$$(15)\qquad \begin{bmatrix}0\\ 0\end{bmatrix} = \begin{bmatrix}\Pi'\Sigma_{xy} & \Pi'\Sigma_{xx}\\ \Sigma_{xy} & \Sigma_{xx}\end{bmatrix}\begin{bmatrix}\Gamma\\ B\end{bmatrix}.$$

By setting $\Pi' = \Sigma_{yx}\Sigma_{xx}^{-1}$, we can express the matrix of the equation in terms of the data moments alone.

Equation (15) is the basis from which the values the structural parameters Γ and B must be inferred. As it stands, the system contains insufficient information for the purpose. In particular, the constituent equation $\Pi'\Sigma_{xy}\Gamma + \Pi'\Sigma_{xx}B = 0$ is a transformation of its companion $\Sigma_{xy}\Gamma + \Sigma_{xx}B = 0$; and, therefore, it contains no additional information. In order for the parameters of the structrural equations to be identifiable, sufficient prior information regarding the structure must be available.

In practice, the prior information commonly takes the form of normalisation rules that set the diagonal elements of Γ to -1 and the exclusion restrictions that set certain of the elements of Γ and B to zeros. If none of restrictions affect more than one equation, then it is possible to treat each equation in isolation.

If the restrictions on the parameters of the jth equation are in the form of exclusion restrictions and a normalisation rule, then they can be represented by the equation

$$\text{(17)} \qquad \begin{bmatrix} R'_\diamond & 0 \\ 0 & R'_* \end{bmatrix} \begin{bmatrix} \gamma_{.j} \\ \beta_{.j} \end{bmatrix} = \begin{bmatrix} r_j \\ 0 \end{bmatrix} \qquad \text{or} \qquad \begin{bmatrix} R'_\diamond & 0 \\ 0 & R'_* \end{bmatrix} \begin{bmatrix} \gamma_{.j} + e_j \\ \beta_{.j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where R_* comprises a selection of columns from the identity matrix I_K of order K, $R_\diamond$ comprises, likewise, a set of columns from the identity matrix I_M of order M, and r_j is a vector containing zeros and an element of minus one corresponding to the normalisation rule. The vector e_j is the jth column of I_M whose unit cancels with the normalised element of $\gamma_{.j}$.

The general solution to these restrictions is

$$\text{(18)} \qquad \begin{bmatrix} \gamma_{.j} \\ \beta_{.j} \end{bmatrix} = \begin{bmatrix} S_\diamond & 0 \\ 0 & S_* \end{bmatrix} \begin{bmatrix} \gamma_{\diamond j} \\ \beta_{*j} \end{bmatrix} - \begin{bmatrix} e_j \\ 0 \end{bmatrix},$$

where $\gamma_{\diamond j}$ and β_{*j} are composed of the M_j and K_j unrestricted elements of $\gamma_{.j}$ and $\beta_{.j}$ respectively, and where $S_\diamond$ and S_* are the complements of $R_\diamond$ and R_* within I_M and I_K respectively.

On substituting the solution of (18) into the equation $\Sigma_{xy}\gamma_{.j} + \Sigma_{xx}\beta_{.j} = 0$, which is from the jth equation of (16), we get

$$\text{(19)} \qquad \Sigma_{xy}S_\diamond\gamma_{\diamond j} + \Sigma_{xx}S_*\beta_{*j} = \Sigma_{xy}e_j.$$

This is a set of K equations in $M_j + K_j$ unknowns; and, given that the matrix $[\Sigma_{xy}, \Sigma_{xx}]$ is of full rank, it follows that the necessary and sufficient condition for the identifiability of the parameters of the jth equation is that $K \geq M_j + K_j$.

If this condition is fulfilled, then any subset of $M_j + K_j$ of the equations of (19) will serve to determine $\gamma_{\diamond j}$ and β_{*j}. However, we shall be particularly interested in a set of $M_j + K_j$ independent equations in the form of

$$\text{(20)} \qquad \begin{bmatrix} S'_\diamond\Pi'\Sigma_{xy}S_\diamond & S'_\diamond\Pi'\Sigma_{xx}S_* \\ S'_*\Sigma_{xy}S_\diamond & S'_*\Sigma_{xx}S_* \end{bmatrix} \begin{bmatrix} \gamma_{\diamond j} \\ \beta_{*j} \end{bmatrix} = \begin{bmatrix} S'_\diamond\Pi'\Sigma_{xy}e_j \\ S'_*\Sigma_{xy}e_j \end{bmatrix},$$

which are derived by premultiplying equation (19) by the matrix $[\Pi S_\diamond, S_*]'$.

The so-called two stage least-square estimates are derived from these equations by substituting the empirical moments M_{xx}, M_{xy} and the estimate $\hat{\Pi} = M_{xx}^{-1}M_{xy}$ in place of Σ_{xx}, Σ_{xy} and $\Pi = \Sigma_{xx}^{-1}\Sigma_{xy}$ repectively and solving the resulting equations for $\gamma_{\diamond j}$ and β_{*j}.

Two-Stage Least Squares and Instrumental Variables Estimation

The 2SLS estimating equations were derived independently by Theil and by Basmann, who followed a different line of reasoning from the one which we have pursued above. Their approach was to highlight the reason for the failure of ordinary least-squares regression to deliver consistent estimates of the parameters of a structural equation.

The failure is due to the violation of an essential condition of regression analysis which is that the disturbances must be uncorrelated with the explanatory variables on the RHS of the equation. Within the equation $y_j = Y_\diamond \gamma_{\diamond j} + X_* \beta_{*j} + \varepsilon_j$, there is a direct dependence of $Y_\diamond$ on the structural disturbances of ε. However, the disturbances are independent of the exogenous variables in X_*.

The original derivations of the 2SLS estimator were inspired by the idea that, if it were possible to purge the variables of $Y_\diamond$ of their dependence on ε, then ordinary least-squares regression would become the appropriate method of estimation. Thus, if $X\Pi_{X\diamond}$ were available, then this could be put in place of $Y_\diamond$; and the problem of dependence would be overcome.

Although $X\Pi_{X\diamond}$ is an unknown quantity, a consistent estimate of it is available in the form of $\hat{Y}_\diamond = X\hat{\Pi}_{X\diamond}$. Finding the estimate $\hat{\Pi}_{X\diamond}$ represents the first stage of the 2SLS procedure. Applying ordinary least-squares regression to the equation $y_j = \hat{Y}_\diamond \gamma_{\diamond j} + X_* \beta_{*j} + e$ is the second stage.

An alternative approach which leads to the same 2SLS estimator is via the method of instrumental-variables estimation. The method depends upon finding a set of variables which are correlated with the regressors yet uncorrelated with the disturbances.

In the case of the structural equation, the appropriate instrumental variables are the exogenous variables of the system as a whole which are contained in the matrix X. Premultiplying the structural equation by X' gives

$$X'y_j = X'Y_\diamond \gamma_{\diamond j} + X'X_* \beta_{*j} + X'\varepsilon. \tag{21}$$

Within this system, the cross products correspond to a set of moment matrices which have the following limiting values:

$$\begin{aligned}
&\text{plim}(T^{-1}X'y_j) = \Sigma_{xy} e_j,\\
&\text{plim}(T^{-1}X'Y_\diamond) = \Sigma_{xy} S_\diamond,\\
&\text{plim}(T^{-1}X'X_*) = \Sigma_{xx} S_*,\\
&\text{plim}(T^{-1}X'\varepsilon) = 0.
\end{aligned} \tag{22}$$

When the moment matrices are replaced by their limiting values, we obtain the equation

$$\Sigma_{xy} e_j = \Sigma_{xy} S_\diamond \gamma_{\diamond j} + \Sigma_{xx} S_* \beta_{*j}, \tag{23}$$

which has been presented already as equation (19). In this system, there are K equations in $M_j + K_j$ parameters. We may assume that $[\Sigma_{xy}, \Sigma_{xy}]$ is of full rank. In that case, the necessary condition for the indentifiablity of the parameters $\gamma_{\diamond j}$ and β_{*j} is that $K \geq$

$M_j + K_j$, which is to say that the number of exogenous variables in the system as a whole must be no less that the number of structural parameters that need to be estimated.

The empircal counterpart of (23) is the equation

$$X'y_j = X'Y_\diamond \gamma_{\diamond j} + X'X_* \beta_{*j}. \tag{24}$$

If $K = M_j + K_j$, then this equation can be solved directly to provide the estimates. However, if $K > M_j + K_j$, then the equation is bound to be algebraically inconsistent and the parameters are said to be overidentified. To resolve the inconsistency, we may apply to (21) the method of generalised least-squares regression. The disturbance term in (21), which is $X'\varepsilon$, had a dispersion matrix $D(X'\varepsilon) = \sigma^2 X'X$. When this is used in the context of the generalised least-squares estimator, we obtain, once again, the 2SLS estimates in the form of

$$\begin{bmatrix} Y_\diamond' X(X'X)^{-1}X'Y_\diamond & Y_\diamond' X(X'X)^{-1}X'X_* \\ X_*' X(X'X)^{-1}X'Y_\diamond & X_*' X(X'X)^{-1}X'X_* \end{bmatrix} \begin{bmatrix} \gamma_{\diamond j} \\ \beta_{*j} \end{bmatrix} = \begin{bmatrix} Y_\diamond' X(X'X)^{-1}X'y_{.j} \\ X_*' X(X'X)^{-1}X'y_{.j} \end{bmatrix}, \tag{25}$$

With the help of the identities $Y_\diamond' X(X'X)^{-1} = \hat{\Pi}'_{X\diamond}$ and $X_*' X(X'X)^{-1}X' = X_*'$, the equation can be rewritten as

$$\begin{bmatrix} \hat{\Pi}'_{X\diamond} X'Y_\diamond & \hat{\Pi}'_{X\diamond} X'X_* \\ X_*'Y_\diamond & X_*'X_* \end{bmatrix} \begin{bmatrix} \gamma_{\diamond j} \\ \beta_{*j} \end{bmatrix} = \begin{bmatrix} \hat{\Pi}'_{X\diamond} X'y_{.j} \\ X_*'y_{.j} \end{bmatrix}, \tag{26}$$

which is evidently a sample analogue of equation (20).