J. King Saud Univ., Vol. 10, *Admin. Sci.* (1), pp. 13-28 (A.H. 1418/1998)

A Method for Estimating Simultaneous Equations Models with Time-series and Cross-section Data

Khalid I. AlDakhil
Assistant Professor, Department of Economics,
College of Administrative Sciences, King Saud University,
Riyadh, Saudi Arabia

(Received on 27/11/1416; accepted for publication on 11/6/1417)

Abstract. The main objective of this paper is to develop a method of estimation for a general simultaneous equations model with time-series and cross-section data. The most common estimation method used in the literature is the 3SLS, and it accounts for the endogeneity issue and disturbance correlation across equations for a given time period. The suggested 2SLS/SUR method accounts for the endogeneity issue and the disturbance correlation across time periods for a given equation. This method is a novel econometric procedure for estimation and we are unaware of any prior published study that has used it in the context of a simultaneous-equations approach.

Introduction

Most economic theories are based on sets or systems of relationships, therefore, they are expressed in terms of multiple-equations. One of the main advantages of expressing economic theory in a system of equations is to show the interdependency that characterizes some of these theories. This interdependency or simultaneity issue between some equations or in the whole system, which is called simultaneous equations system, is a practical problem in the subject of Econometrics. The presence of this simultaneity has an implication for the estimation of each equation and the complete system because of the feedback effects and the dual causality between variables from one side and the disturbance correlation across equations and across time periods from the other side. The main objective of this paper is to present a new econometric procedure for estimation of simultaneous equations models with time-series and cross-sectional data.

This paper shows the most common econometric procedures used in the literature to estimate such simultaneous equations models. The first section of the paper gives a general form of simultaneous equations model for M equations, N observations, and T time periods. In the second section of the paper, an analysis of some possible disturbance correlations between equations, observations, and time periods is shown. The most common estimation method for this kind of models is discussed in section three. Section four shows the suggested estimation technique for simultaneous equations models with time-series and cross-section data and discusses a study that applies these two methods of estimation. The last part discusses the possibility of pooling the data in the whole system as the best option.

1. A General Form of Time-series and Cross-section Simultaneous Equations Model

Consider a set of M equations, K exogeneous variables, N observations, and T time periods. This means we have a number of N observations for each equation and M equations for each time period. Therefore, each time period has the same endogenous and exogenous variables and the same number of observations. For each time period, let the general structural form of the model be as follow:[1]

$$\begin{aligned} &\mathbf{y}_1\gamma_{11}+\mathbf{y}_2\gamma_{21}+\cdots+\mathbf{y}_M\gamma_{M1}+\mathbf{x}_1\beta_{11}+\mathbf{x}_2\beta_{21}\ldots+\mathbf{x}_k\beta_{K1}+\varepsilon_1=0\\ &\mathbf{y}_1\gamma_{12}+\mathbf{y}_2\gamma_{22}+\ldots+\mathbf{y}_M\gamma_{M2}+\mathbf{x}_1\beta_{21}+\mathbf{x}_2\beta_{22}\cdots+\mathbf{x}_k\beta_{K2}+\varepsilon_2=0 \qquad (1.1)\\ &\vdots\\ &\mathbf{y}_1\gamma_{1M}+\mathbf{y}_2\gamma_{2M}+\ldots+\mathbf{y}_M\gamma_{M}+\mathbf{x}_1\beta_{1M}+\mathbf{x}_2\beta_{2M}\ldots+\mathbf{x}_k\beta_{KM}+\varepsilon_M=0,\end{aligned}$$

where:

Each of the vectors $\mathbf{y}_1\ldots\mathbf{y}_M$ represents the N observations on the M endogenous variables.

Each of the vectors $\mathbf{x}_1\ldots\mathbf{x}_K$ represents the N observation on the K exogenous variables.

Each of the vectors $\varepsilon_1\ldots\varepsilon_M$ represents the N observations on the M random errors.

γ's and β's are the structural parameters of the system that will be estimated.

In matrix notation, this system could be written as

$$[\mathbf{y}_1\ \mathbf{y}_2\cdots\mathbf{y}_M]\begin{bmatrix}\gamma_{11} & \gamma_{22}\cdots\gamma_{1M}\\ \gamma_{21} & \gamma_{22}\cdots\gamma_{2M}\\ \vdots & \\ \gamma_{M1} & \gamma_{M2}\cdots\gamma_{MM}\end{bmatrix}+[\mathbf{x}_1\ \mathbf{x}_2\ldots\mathbf{x}_K]\begin{bmatrix}\beta_{11} & \beta_{12}\cdots\beta_{1M}\\ \beta_{21} & \beta_{22}\cdots\beta_{2M}\\ & \\ \beta_{K1} & \beta_{K2}\cdots\beta_{KM}\end{bmatrix}+[\varepsilon_1\ \ \varepsilon_2\ldots\varepsilon_M]=0 \qquad (1.2)$$

For the full set of N observations, the structure is given by

[1] Bold faced lower case letters refer to vectors and bold faced upper case letters refer to matrices.

$$\underset{N\times M}{\mathbf{Y}}\ \underset{M\times M}{\Gamma}\ \underset{N\times K}{\mathbf{X}}\ \underset{K\times M}{\mathrm{B}} = \underset{N\times M}{\mathrm{E}}$$

$$\text{Where } \mathbf{X} = \begin{bmatrix} x_{11} \cdots \cdots x_{1K} \\ x_{21} \cdots \cdots x_{2K} \\ \vdots \\ x_{N1} \cdots \cdots x_{NK} \end{bmatrix}_{N\times K}$$

$$\text{and } \mathbf{Y} = \begin{bmatrix} Y_{11} \cdots Y_{1M} \\ \vdots \\ Y_{N1} \cdots Y_{NM} \end{bmatrix}_{N\times M},$$

$$\mathbf{E} = \begin{bmatrix} \varepsilon_{11}\ \varepsilon_{12} \cdots \varepsilon_{1M} \\ \vdots \\ \varepsilon_{N1} \cdots \cdots \varepsilon_{NM} \end{bmatrix}_{N\times M}.$$

Given that Γ is a square matrix and assuming it is nonsingular, then we could get the reduced form for all the N observations and all the M jointly dependent variables in the conventional way,

$$\mathbf{Y} = -\mathbf{X}\mathbf{B}\Gamma^{-1} + \mathbf{E}\Gamma^{-1} \tag{1.3}$$

Let, $\Pi = -\mathbf{B}\Gamma^{-1}$ and $\mathbf{V} = \mathbf{E}\Gamma^{-1}$, then the above expression could be written as

$$\mathbf{Y} = \mathbf{X}\Pi + \mathbf{V} \tag{1.4}$$

Assume that the disturbances are correlated across equations and uncorrelated across time period. This means we need to estimate the whole system for each time period separately as long as that equation's disturbances are not correlated among time periods. Therefore, we want to estimate the reduced form matrix (Π) and from it we can retrieve estimates for the elements of B and Γ. The most common estimation method used in the literature for this kind of models is the three stage least squares (3SLS).

2. Analysis of Disturbance Correlations

Before starting the discussion of the estimation methods, it is worthwhile to explain the equations disturbance correlations in some detail. Some possible disturbance correlations between observations, equations, and time periods are shown in Fig. 1. To understand these correlations, let the t and s subscripts indicate two different time periods in the system, let the m and l subscripts indicate two different equations in the same system, and let the n and g subscripts indicate two different observations in that system. Within each time period t and s, there are two different equations. These equations are y_{tl} and y_{tm} for period t and y_{sl} and y_{sm} for period s. Each equation contains N observations, and for any given two observations n and g, there will be eight different equations: y_{tln}, y_{tlg}, y_{tmn}, Y_{tmg}, Y_{sln}, , y_{slg}, y_{smn} and y_{smg}, and their corresponding disturbances are: ε_{tln}, ε_{tlg}, ε_{tmn}, ε_{tmg}, ε_{sln}, ε_{smn}, and ε_{smg}, respectively.

This paper as shown in Fig. 1 is mainly concerned with two kinds of disturbance correlations. These two kinds of disturbance correlations are indicated in Fig. 1 by their corresponding roman numerals. The first kind is shown by I and the second by II. The first correlation, which is accounted for by the 3SLS method discussed in the next section is between disturbances of different equations, but of the same observation and time period (e.g., ε_{tlg} and ε_{tmg}). The second correlation which is accounted for by the suggested 2SLS/SUR method discussed in the last section is between two disturbances of different time periods, but for the same equation and observation (e.g., ε_{tln} and ε_{sln}). This means any other kind of disturbance correlation is assumed to equal zero.

3. Three Stage least Squares (3SLS)

The 3SLS applies the generalized least-squares estimation to all equations that have been estimated individually by 2SLS. Therefore, it considers information on the complete structure of the model. The 2SLS provides a very useful estimation procedure to get a unique estimated value for each structural equation parameter.

The real advantage of using the 3SLS method over the 2SLS is that it is asymptotically more efficient when the structural equations disturbances are correlated. Moreover, 2SLS, which estimates the system equation by equation, accounts for all exogenous variables in the system and only the endogenous variables in each specific equation, whereas 3SLS accounts for all exogenous and endogenous variables. Therefore, the first is called a full-information method and the second is a limited information method.[2]

In this study we assume that each structural equation is over-identified. Therefore, the whole system is assumed to be over-identified and consistent estimation of all parameters is possible. To see the 3SLS procedure, let us write each of the structural equations in the above general form as follows:

---

[2] Some of the studies that have used either method in Simultaneous-equations Models are Greenwood [1], and Metwally [2].

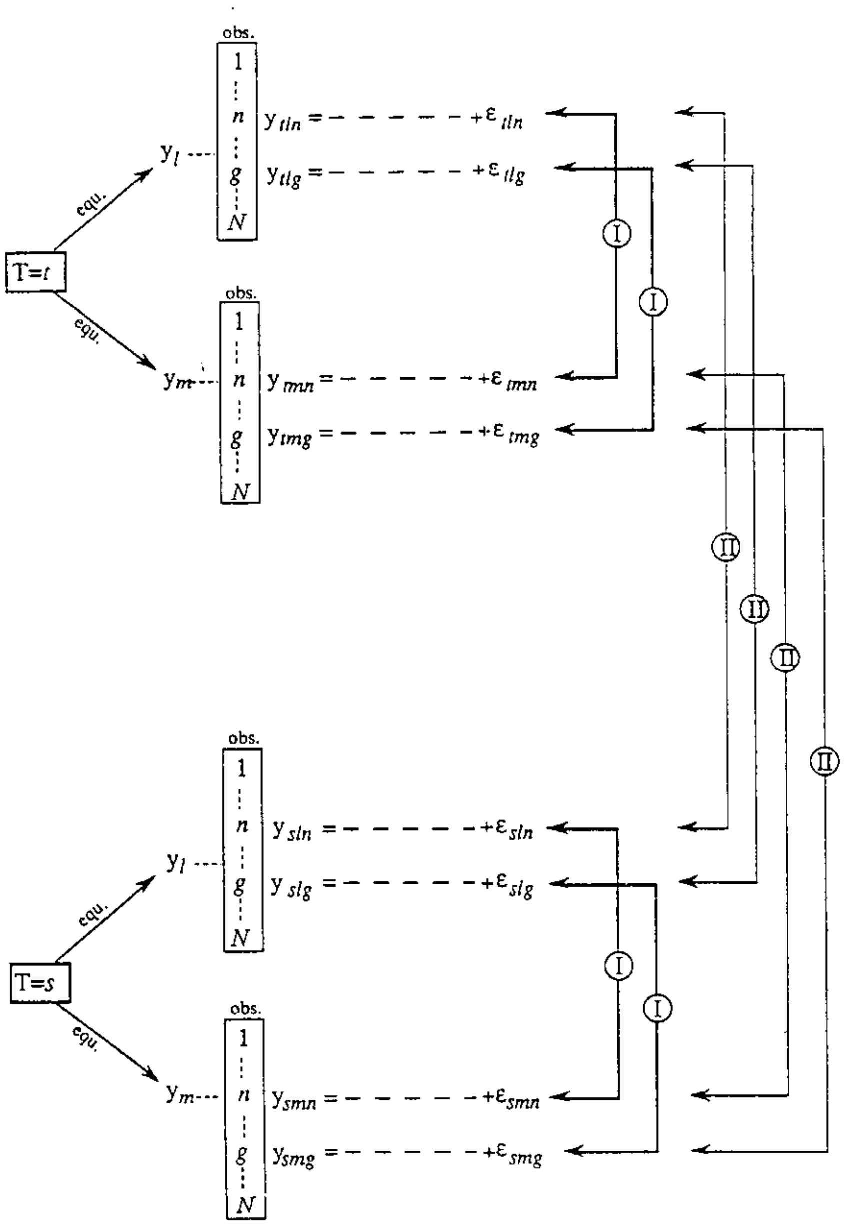

Fig. 1. Analysis of disturbance correlations.

$$\begin{aligned} \mathbf{y}_1 &= Z_1 \ \delta_1 + \varepsilon_1 \\ \mathbf{y}_2 &= Z_2 \ \delta_2 + \varepsilon_2 \\ &\vdots \\ \mathbf{y}_M &= Z_m \ \delta_m + \varepsilon_m \end{aligned} \tag{3.1}$$

where

$$\underset{N\times(M+K)}{\mathbf{Z}_j} = [\mathbf{Y}_j \ \ \mathbf{X}_j]$$

and

$$\underset{(M+K)\times 1}{\delta_j} = \begin{bmatrix} \gamma_j \\ \beta_j \end{bmatrix}$$

This $\mathbf{Z}_j$ matrix contains all the explanatory variables (both jointly dependent and predetermined) in the right-hand side of the jth equation. By stacking the above individual equations, we get the M structural equations for all N observations as

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_M \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 & & \\ & \mathbf{Z}_2 & \\ & & \ddots \\ & & & \mathbf{Z}_M \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_M \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix};$$

or, compactly, as

$$\underset{NMx1}{y} = \underset{NMx(M+K)M}{Z'} \ \underset{(M+K)Mx1}{\delta} + \underset{NMx1}{\varepsilon} \tag{3.2}$$

where

$$E(\varepsilon) = 0;$$

This means each element of the disturbance vector ε has a zero mean, and

$$\underset{NMxNM}{E}(\varepsilon\varepsilon`) = E\begin{bmatrix} \varepsilon_1\varepsilon_2` & \varepsilon_1\varepsilon_2` & \ldots\varepsilon_1\varepsilon_M` \\ \varepsilon_2\varepsilon_1` & \varepsilon_2\varepsilon_2` & \ldots\varepsilon_2\varepsilon_M` \\ \vdots & \vdots & \vdots \\ \varepsilon_M\varepsilon_1` & \varepsilon_M\varepsilon_2` & \ldots\varepsilon_M\varepsilon_M` \end{bmatrix} \quad (3.3)$$

$$= \begin{bmatrix} \sigma_{11}I & \sigma_{12}I & \ldots & \sigma_{1M}I \\ \sigma_{21}I & \sigma_{22}I & \ldots & \sigma_{2M}I \\ \vdots & \vdots & & \vdots \\ \sigma_{M1}I & \sigma_{M2}I & \ldots & \sigma_{MM}I \end{bmatrix} = \underset{MxM}{\sum}\otimes I$$

This is the matrix of variances and contemporaneous covariances, and it contains the disturbances of different equations for the same observation N and the same time period. Using this stacked model and the instrumental variable (IV) estimator, it can be shown that the 3SLS estimator is[3]

$$\hat{\delta}_{3SLS} = \left[[\hat{Z}'\left(\hat{\Sigma}^{-1}\otimes I\right)\hat{Z}]^{-1}\hat{Z}'\left(\hat{\Sigma}^{-1}\otimes I\right)y\right] \quad (3.4)$$

The first stage in this method is to estimate the KxM matrix Π in equation (1.4) by least squares and compute the predicted values for each equation $(\hat{Y}j)$. Then it is possible to form a $(\hat{Z}_j)$ matrix that contains the jointly dependent variables' predicted values and the exogenous variables in the right-hand side of the j th equation: $(\hat{Z}_j)$

$$\hat{\mathbf{Z}}\mathbf{j} = \left[\hat{\mathbf{Y}}_j \;\; \mathbf{X}_j\right]$$

The second stage is to estimate δj, which appears in equation (3.2) by the regression of Y_j on $\hat{Z}_j$. This gives the 2SLS estimator of $\delta_{j,2SLS}$. Then $\hat{\sigma}_{ij}$ can be calculated by:

$$\hat{\sigma}_{ij} = \frac{e_i e_j}{N} = \frac{(y_i - Z_i\hat{\delta}_i)`(y_j - Z_j\hat{\delta}_j)}{N},$$

where N is the total number of observations.

From this stage we get the estimated $\hat{\Sigma}$ as:

[3] For the derivation of the 3SLS estimator, see Greene (3, p. 611).

$$\bar{\Sigma} = \begin{bmatrix} \hat{\sigma}_{11}I & \hat{\sigma}_{12}I & \dots & \hat{\sigma}_{1M}I \\ \hat{\sigma}_{21}I & \hat{\sigma}_{22}I & \dots & \hat{\sigma}_{2M}I \\ \vdots & \vdots & & \vdots \\ \hat{\sigma}_{M1}I & \hat{\sigma}_{M2}I & \dots & \hat{\sigma}_{MM}I \end{bmatrix} = \hat{\Sigma} \otimes I$$

The last stage is to compute the 3SLS estimator of equation 3.4 using the above estimated $\hat{Z}$ and $\hat{\Sigma}$.

Finally, it is worthwhile to mention that the 3SLS method is preferable to other methods of estimation such as ordinary least squares (OLS), Indirect least squares (ILS), and 2SLS because 3SLS has greater asymptotic efficiency for this kind of model and that because the equations disturbances are expected to be correlated. The only shortcoming of this method is that it accounts for the equations' disturbance correlation within each time period, but not among different time periods.

4. Two State Least Squares/Seemingly Unrelated Regression (2SLS/SUR)

This suggested method is based on doing the Zellner's SUR in the context of a simultaneous-equations system. One way to do this is to apply both the 2SLS and SUR methods. Applying the 2SLS method will account for the endogeneity or simultaneity issue in the system, and then applying the SUR method will account for the disturbance correlation between different time periods. In order to be able to apply the 2SLS/SUR method, a balanced data set is required; that is, each time period is required to have the same number of observations.

The first step in this method is similar to the first two stages in the previous method (3SLS). This is to apply 2SLS for each of the assumed balanced data sets (time periods) to obtain the predicted values of the right-side endogenous variable of the model. This gives the Z_j matrix, defined above, but now it has to be obtained based on balanced data sets[4] :

$$\hat{Z}_m = \begin{bmatrix} \hat{Y}_m & X_m \end{bmatrix}$$

Now, since we assume that each time period has the same endogenous and exogenous variables and the same number of observations, then it is possible to establish a single new data set tht contains all different time periods variables and observations.

[4] Schmidt (4) shows an estimation of SUR with unequal number of observation.

The next step is to apply the SUR using the above computed predicted values on the right-hand side of each equation, to account for the error correlation across the corresponding equations from each time period.

Since we assume different time periods in this study, we will refer to each one by t, and to any given equation by m. Then the corresponding m equations within the t different time periods could be written as[5]

$$\begin{aligned} \mathbf{y}_{m1} &= \hat{\mathbf{y}}_{m1}\,\gamma_{m1} + \mathbf{x}_{m1}\,\beta_{m1} + \varepsilon_{m1} \\ \mathbf{y}_{m2} &= \hat{\mathbf{y}}_{m2}\,\gamma_{m2} + \mathbf{x}_{m2}\,\beta_{m2} + \varepsilon_{m2} \\ &\vdots \\ \mathbf{y}_{mT} &= \hat{\mathbf{y}}_{mT}\,\gamma_{mT} + \mathbf{x}_{mT}\,\beta_{mT} + \varepsilon_{mT} \end{aligned} \tag{4.1}$$

where

$$E(\varepsilon_{mt}) = 0, \qquad m = 1, ---, M, \qquad t = 1, --- T$$

and

$$E(\varepsilon_m \varepsilon'_m) = E\begin{bmatrix} \varepsilon_1\varepsilon_1 & \varepsilon_1\varepsilon_2 \ldots \varepsilon_1\varepsilon_T \\ \varepsilon_2\varepsilon_1 & \varepsilon_2\varepsilon_2 \ldots \varepsilon_2\varepsilon_T \\ \vdots & \\ \varepsilon_T\varepsilon_1 & \varepsilon_T\varepsilon_2 \ldots \varepsilon_T\varepsilon_T \end{bmatrix} = \begin{bmatrix} \sigma_{11} I & \sigma_{12} I \ldots \sigma_{1T} I \\ \sigma_{21} I & \sigma_{22} I \ldots \sigma_{2T} I \\ \vdots & \\ \sigma_{T1} I & \sigma_{T2} I \ldots \sigma_{TT} I \end{bmatrix}$$

Assume that the disturbances are correlated across the different t time periods and uncorrelated across the equations. To see that, let m and l subscripts indicate two different equations in the system, and t and s subscripts indicate two different time periods in the system. Then

$$E(\varepsilon_{tm}\,\varepsilon_{sl}) = \sigma_{ts} \qquad \text{if } m = l, \text{ and 0 otherwise,}$$

Therefore, the variance-covariance matrix for the m[th] equation could be written as:

$$E(\underset{T\times 1}{\varepsilon_t}\ \underset{1\times T}{\varepsilon_s'}) = \sigma_{ts} I_T = \begin{bmatrix} \sigma_{11}I & \sigma_{12}I \ldots \sigma_{1T}I \\ \sigma_{21}I & \sigma_{22}I \ldots \sigma_{2T}I \\ \sigma_{T1}I & \sigma_{T2}I \ldots \sigma_{TT}I \end{bmatrix} = \sum \otimes I$$

The above m equations shown in (4.1) for the t different time periods could be written, compactly, as

[5] Each of the **y**, **x**, and ε (Nx1) vectors is defined in the beginning of section I.

$$\underset{TNx1}{Y} = \underset{TNx(T+k)T}{\hat{Z}} \underset{(T+k)Tx1}{\gamma} + \underset{TNx1}{\varepsilon}$$

where

$$y = \begin{bmatrix} y_{M1} \\ y_{M2} \\ \vdots \\ y_{MT} \end{bmatrix}, \qquad \hat{Z} = \begin{bmatrix} \underset{Nx(T+K)}{\hat{Z}_{M1}} & & & \\ & \hat{Z}_{M2} & & \\ & & \ddots & \\ & & & \hat{Z}_{MT} \end{bmatrix}$$

$$\gamma = \begin{bmatrix} \gamma_{M1} \\ \gamma_{M2} \\ \vdots \\ \gamma_{MT} \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_{M1} \\ \varepsilon_{M2} \\ \vdots \\ \varepsilon_{MT} \end{bmatrix}$$

A good candidate to estimate this system of SUR is the Feasible Generalized Least Squares (FGLS) estimator [6]:

$$\hat{\gamma} = \left[[\hat{Z}' \left(\hat{\Sigma}^{-1} \otimes I \right) \hat{Z}]^{-1} \hat{Z}' \left(\hat{\Sigma}^{-1} \otimes I \right) y \right]$$

where the estimator $\hat{\Sigma}$ is based on least squares residuals; that is,

$$\hat{e}_{mt} = \mathbf{Y}_{mt} - \mathbf{X}_{mt}\, b_{mt},$$

and its elements are given by

$$\hat{\sigma}_{ts} = \frac{\hat{e}'_{mt} \hat{e}_{ms}}{T}$$

and, as in Judge *et al.* (5), the above estimator $\hat{\gamma}$ is called the Zellner's seemingly unrelated regression estimator.

Clearly, in comparing this suggested method of estimation (2SLS/SUR) with the previous method (3SLS), we realize that the 2SLS/SUR accounts for the disturbance correlation between time periods, which is shown by II in Fig. 1, and assumes a zero disturbance correlation between equations in a given time period, while the 3SLS does exactly the reverse: It accounts for the correlations between equations in a given time

---

[6] We suggested the FGLS estimator and not the GLS estimator because the variance-covariance matrix is unknown. For more details, see judge, *et al.* (5, p. 321) and Amemiya [6, p. 186].

period, which is shown by I in the same figure, but assumes a zero disturbance correlation among time periods. The method that is more appropriate to estimate the assumed system of equations depends on which of the above two correlations is more important, and that, in turn, depends on the empirical results that will be derived from applying each method of estimation.

5. Panel Data Model (Pooling the Data)

The procedure is based on pooling, or combining, cross-sectional and time-series data. The idea is to create one data set that has all observations (N) of all time periods (T). Since each time period has a balanced number of observations N, we end up with TxN observations for each variable in the new data set. The model associated with the TxN observations could be written as:

$$\underset{Nx1}{y_{11}} = \hat{y}_{11}\,\gamma_{11} + X_{11}\,\beta_{11} + \varepsilon_{11}$$

$$\underset{Nx1}{y_{12}} = \hat{y}_{12}\,\gamma_{12} + X_{12}\,\beta_{12} + \varepsilon_{12}$$

$$\vdots$$

$$\underset{Nx1}{y_{1T}} = \hat{Y}_{1T}\,\gamma_{1T} + X_{1T}\,\beta_{1T} + \varepsilon_{12}$$

$$\vdots$$

$$\underset{Nx1}{y_{M,1}} = \hat{Y}_{M,1}\,\gamma_{M,1} + X_{M,1} + \beta_{M,1} + \varepsilon_{M,1}$$

$$\underset{Nx1}{Y_{M,2}} = \hat{Y}_{M,2}\,\gamma_{M,2} + X_{M,2} + \varepsilon_{M,2}$$

$$\vdots$$

$$\underset{Nx1}{Y_{M,T}} = \hat{Y}_{M,T}\,\gamma_{M,T} + X_{M,T}\,\beta_{M,T} + \varepsilon_{M,T}$$

or compactly, as

$$\begin{bmatrix} \underset{NTx1}{y} \\ \underset{NTx1}{y_2} \\ \vdots \\ \underset{NTx1}{y} \end{bmatrix}_{NTMx1} = \begin{bmatrix} \underset{NTxM}{\hat{y}} & & & \\ & \underset{NTxM}{\hat{y}} & & \\ & & \ddots & \\ & & & \underset{NTxM}{\hat{y}} \end{bmatrix}_{NTMxMM} \begin{bmatrix} \underset{Mx1}{\gamma_1} \\ \underset{Mx1}{\gamma_2} \\ \vdots \\ \underset{Mx1}{\gamma_M} \end{bmatrix}_{MMx1} +$$

$$\begin{bmatrix} \underset{NTxK}{x_1} & & & \\ & \underset{NTxK}{x_2} & & \\ & & \ddots & \\ & & & \underset{NTxK}{x_M} \end{bmatrix}_{NTMxKM} \begin{bmatrix} \underset{Kx1}{\beta_1} \\ \underset{Kx1}{\beta_2} \\ \vdots \\ \underset{kx1}{\beta_M} \end{bmatrix}_{KMx1} + \begin{bmatrix} \underset{NTx1}{\varepsilon_1} \\ \underset{NTx1}{\varepsilon_2} \\ \vdots \\ \underset{NTx1}{\varepsilon_M} \end{bmatrix}_{NTMx1}$$

Alternatively, that could be written as

$$\underset{NTMx1}{Y} = \underset{NTMxMM}{\hat{Y}} \underset{MMx1}{\gamma} + \underset{NTMxKM}{X} \underset{KMx1}{\beta} + \underset{NTMx1}{\varepsilon}$$

Assuming the availability of data for each observation over the T time periods, then the practical question is whether or not the cross-section parameters of the statistical model remain constant over time. If they remain constant, then it is possible to pool the data of the T time periods to get more efficient parameter estimates. But if the cross-section parameter shift over time, then pooling is not an appropriate procedure.

In order to answer this important question of whether or not pooling is an appropriate procedure, the following null and alternative hypotheses need to be tested.

$$Ho: \beta_1 = \beta_2 = \ldots = \beta_K$$

$$H_1: \text{ at least one } \beta_i \neq \beta_j$$

where

K = the number of parameters in the m^{th} equation

i and j = 1, …, K.

According to the null hypothesis (Ho), the coefficients are going to be the same and constant over time periods and over cross-section observations, while the alternative hypothesis (H_1) means the opposite. Therefore, if Ho is not rejected, then pooling is an appropriate procedure, but if it is rejected, then we cannot pool because that will misspecify the statistical model.

For testing the above two hypotheses, the following test statistics could be applied:

$$\frac{\{e'e-(e'_1 e_1 + e'_2 e_2 + e'_3 e_3)\}/K}{e'_1 e_1 + e'_2 e_2 + e'_3 e_3/(N-K)} = F(K, N-K)$$

where

$e'e$ = residual sums of squares resulted from the pooling regression;

$e'_t e_t$ = residual sums of squares resulted from each individual regression, t = 1, …, T.

alternatively, this test could be written as

$$\frac{(ESS1 - ESS2)/2K}{ESS2/(N - 2K)} = F(K, N-K) \tag{5.1}$$

where

ESS1 = $e'e$ = residual sum of squares resulting from the pooling regression.

ESS2 = $\sum_{t=1}^{\tau} e_{tt}$ = total residual sum of squares of the individual regressions.

The test shows how much bigger is the pooled (ESS) than the sum of each of the three periods (ESS). The left-hand side of equation (5.1) gives the calculated test statistic, while the right-hand side gives the tabulated F statistic for a given level of significance (e.g., 5%). If the tabulated F statistic is greater than the tabulated test statistic, then the null hypothesis will not be rejected. The ESS1 is obtained for the pooled data by running a least squares regressions for each time period using the predicted values of the endogenous variables. The ESS2 is also obtained by running separate least squares regressions for each time period using the predicted values of the endogenous variables.

Finally, in a study that develops a large simultaneous-equations model of metropolitan employment growth and migration in the U.S.A., Aldakhil [7, p. 140] uses the above three econometric procedures for estimation purposes. The study model consists of 17 equations (10 structural equations and 7 identities), 17 jointly dependent (endogenous) variables and 19 independent (exogenous variables). The model has been estimated in double-logarithmic form for three different time periods (decades). The study's empirical results suggest that the 2SLS/SUR method is more appropriate than the 3SLS to estimate the system and that implies the importance of the disturbance correlation within time periods for the same equation in that kind of study. For both methods, the number of estimated coefficients in the structural equations was 88 in each time period. The 2SLS/SUR method dominates the 3SLS in terms of the number of significant coefficients. Moreover, it has a greater number of signs that are not unexpected signs and corresponding coefficients that are significant in the firs two periods and almost the same number as 3SLS in the last period.[7]

In order to pool the above study data it has been found that each structural equation in the system has a calculated test statistic greater than the tabulated F statistic. Clearly, this means the null hypothesis, that the coefficients are going to be the same and constant over the periods and over cross-section observations, must be rejected. Therefore, the data cannot be pooled over the three time periods, because fitting them together will misspecify the model, since the coefficients, as concluded, are not the same.

Summary and Conclusion

This paper develops a method of estimation for a general simultaneous equations model with time-series and cross-section data. The model consists of M equations, K exogenous variables, N observations, and T time periods. There are two main issues in this kind of models, the first is the endogeneity of variables in each equation and in the whole system and the second is the disturbance correlation across equations and across time periods. The 3SLS is the most common estimation method in the literature and it accounts for the endogeneity issue and the disturbance correlation across equations for a given time period. The suggested 2SLS/SUR method accounts for the endogeneity issue and the disturbance correlation across time periods for a given equation. This method is a novel econometric procedure for estimation and we are unaware of any published study that has used this suggested econometric methodology in the context of a simultaneous-equations approach. The decision regarding which method to use depends on the relative importance of each kind of disturbance correlation and this is based on each method's empirical results. It is worthwhile to mention that the most preferred and approprite estimation method for simultaneous equations models with time-series and cross sectional data is the 3SLS/SUR because it will take into account the endogeneity and both kinds of disturbance corelations. Therefore, one way to push this study forward is to develop and employ the 3SLS/SUR estimation technique by considering computer and program limitations.

References

[1] Greenwood, Michael J. "Simultaneous-Equation Model of Urban Growth and Migration." *Journal of the American Statistical Association, 70 (1975),* 797-810.

[2] Metwally, M.M. "Trade Relationship Between Saudi Arabia and Japan." *Journal of Administrative Sciences*, King Saud University, 13 (1980), 3-15.

[3] Greene, William H. *Econometric Analysis*. 2nd ed. New York: Macmillan Publishing Company, (1993).

[7] Not unexpected sign for a coefficient means that it either has a prior expected sign, or not specified.

[4] Schmidt, P. "Estimation of Seemingly Unrelated Regressions with Unequal Numbers of Observations." *Journal of Econometrics*, 5, No. 3 (1977), 765-377.

[5] Judge, G., Hill, C., Griffiths, W., Lee, T., and Lutkepol, H. *An Introduction to the Theory and Practice of Econometrics*. 2nd ed. New York: John Wiley and Sons, 1989.

[6] Amemiya, Takeshi. *Advanced Econometrics*. Cambridge, Massachusetts: Harvard University Press, 1985.

[7] AlDakhil, Khalid. "The Role of Employment, Migration and Amenities in the Growth of Large U.S. Metropolitan Areas: 1950-1980": *Unpublished Ph.D. dissertation*, University of Colorado, Boulder, 1994.

طريقة لتقدير نماذج المعادلات الآنية ذات السلاسل الزمنية والمقاطع المستعرضة

خالد بن ابراهيم الدخيل
قسم الاقتصاد، كلية العلوم الإدارية، جامعة الملك سعود
الرياض ـ المملكة العربية السعودية

(قدم للنشر في ٢٧/ ١١/ ١٤١٦هـ؛ وقبل للنشر في ١١/ ٦/ ١٤١٧هـ)

ملخص البحث. في نماذج المعادلات الآنية القائمة على بيانات المقاطع المستعرضة Cross-Sectional Data وبيانات السلاسل الزمنية Time-Series Data توجد مشكلتان رئيسيتان الأولى تتمثل في وجود العلاقة السببية أو التبادلية بين المتغيرات المستقلة والمتغيرات التابعة، والثانية تتمثل في وجود ارتباط بين حدود الخطأ المختلفة، وتجدر الإشارة إلى أن الطريقة الدارجة الاستخدام في تقدير مثل هذه النماذج هي طريقة المربعات الصغرى ذات المراحل الثلاث (3SLS) والتي تأخذ بعين الاعتبار المشكلة الأولى ولكنها تقتصر في حل المشكلة الثانية على معالجة الارتباط بين حدود الخطأ في المعادلات المختلفة، وذلك لنفس المشاهدات والفترة الزمنية.

يهدف هذا البحث إلى اقتراح وتقديم طريقة تقدير قياسية جديدة لهذا النوع من النماذج تقوم على التنسيق والجمع بين طريقة المربعات الصغرى ذات المرحلتين (2SLS) وطريقة نظم المعادلات غير المرتبطة ظاهريا (SURE) لزيلنر، وتتميز هذه الطريقة بأنها تأخذ في الحسبان تلك العلاقة التبادلية بين المتغيرات المستقلة والمتغيرات التابعة من جهة، وكذلك الارتباط بين حدود الخطأ في الفترات الزمنية المختلفة وذلك لنفس المشاهدة والمعادلة من جهة أخرى.