UNIT 5 ESTIMATION OF SYSTEM OF EQUATIONS*

Structure

5.0 Objectives

5.1 Introduction

5.2 Seemingly Unrelated Regression Equations (SURE)

5.3 Generalized Least Squares (GLS)

5.4 Feasible Generalized Least Squares (FGLS)

5.5 Maximum Likelihood Estimates

5.5.1 Hypothesis Testing

5.5.2 Treating Autocorrelation

5.6 Interrelated Factor Demand

5.7 Let Us Sum Up

5.8 Key Words

5.9 Some Useful Books

5.10 Answer/Hints to Check Your Progress Exercises

5.0 OBJECTIVES

After going through this unit you should be in a position to

- explain the underlying ideas behind the seemingly unrelated regression equations (SURE) model;
- estimate the parameters of the SURE models using generalized least squares (GLS) method;
- compare the efficiency of the GLS estimators with the OLS estimators;
- describe the advantages of the feasible generalized least squares (FGLS) estimators;
- apply the maximum likelihood estimators (MLE) for SURE model;
- test hypotheses in the SURE framework;
- suggest remedial measures for treatment of autocorrelation in the SURE model; and
- apply of the SURE model for estimation of inter-related factor demand.

* Dr. Samraj Sahay, Independent Researcher, Delhi

5.1 INTRODUCTION

So far we have been dealing with single equation regression models. In these models there is as explained (or, dependent) variable; and there are one or more explanatory (or, independent) variables. Such models mainly use a set of predictors or explanatory variables to describe the behaviour of a specific dependent variable.

There are however a different category of econometric models where we need to explain the behaviour of a whole system of equations. In such cases a single equation will not serve the purpose; rather we need to have more than one equation. In such a system, each equation explains certain economic phenomenon.

Multiple equations could be of two types: (i) The individual equations interact with each other as the random error components of some of the equations may be correlated. However, none of the variables in the system behaves both as an explanatory and explained variable. In other words, the equations within the system may be statistically linked but not structurally. Such system of equation is referred to as seemingly unrelated regression equations (SURE) which are related to each other in reality but may not appear to be so. (ii) A variable acts as dependent or endogenous variable in an equation; in another equation the same variable acts as explanatory or exogenous variable. This type of equation system is known as simultaneous equation model.

5.2 SEEMINGLY UNRELATED REGRESSION EQUATIONS (SURE)

The SURE model has several important applications in economics. It has been used extensively in the estimation of system of demand equations. As you know, demands for various commodities by households are inter-related. Demand equations represent a system of equations explaining consumption of different commodities by households. Thus, it is very likely that the equations in such system are correlated due to specific household attributes. Let us take the case of estimation of a cost function along with input share equations using firm level data. The correlation between the equations may arise due to certain firm specific and unobservable effects, which influence the choice of inputs and the cost of production.

Let us try and understand the concept of SURE using a simple example. Suppose we have data of five firms for 20 years. Our objective is to study investment demand. We have 5 equations for 5 firms. Each of these equations can be estimated separately using OLS, which provide consistent estimators. Although each of these equations appears to be independent of each other, there may be some relationship among the equations. It is logical to assume that the error terms associated with each of these equations are correlated. These equations, which are 'seemingly unrelated regression equations', are in fact not independent of one

another. To develop an understanding of SURE let us consider the following model which has a set of M multiple regression equations.

$$y_i = \beta_i X_i + \varepsilon_i \qquad i = 1, 2, ..M \qquad ... (5.1)$$

where y_i is $(T X 1)$, X_i is $(T X K_i)$, β_i is $(Ki X 1)$ and ε_i is a $(T X 1)$ vector.

In equation (5.1), T represents the number of cross sections, i.e., the number of observations in each of the regression equations.

Further, to estimate these equations as a system of equations we need to combine these Seemingly Unrelated Regression Equations (SURE) in one stacked model which is expressed as

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & X_M \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix} = X\beta + \varepsilon \qquad ... (5.2)$$

In equation (5.2), each of the M equations is treated as a classical regression model fulfilling all the assumptions. All the *M* equations are related to each other through the error structure which is given by

$$E(\varepsilon\varepsilon') = \Omega = \sum \otimes I_T = \begin{bmatrix} \sigma_{11} I_T & \sigma_{12} I_T & \cdots & \sigma_{1M} I_T \\ \sigma_{21} I_T & \sigma_{22} I_T & \cdots & \sigma_{2M} I_T \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} I_T & \sigma_{M2} I_T & \cdots & \sigma_{MM} I_T \end{bmatrix}_{MTXMT} \qquad ... (5.3)$$

In the above equation we have

$$\sum = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \sigma_{M2} & \cdots & \sigma_{MM} \end{bmatrix}_{MXM} \qquad ... (5.4)$$

which is the variance covariance matrix for each t =1,....., T. The symbol $\otimes$ represents the 'Kronecker product operator'.

For the error structure given at equation (5.4), we have $E(\varepsilon\varepsilon') = \Omega = \sum \otimes I_T$. The error structure has three implications: (i) for all t the ε_{it} has constant variance, (ii) the covariance between the error terms of the equations ε_{ti} and ε_{tj} is constant, and (iii) intertemporal covariance or the covariance between error term of the observations within the equation is equal to 0 (i.e., covariance between ε_{ti} and ε_{t*j} ($t \neq t$*) is equal to 0).

Thus we see that the SURE model is a specific case of simultaneous equations model where there are *M* structural equations with *M* joint dependent variable and K independent variables which are exogenous. None of the structural

equations have any current or lagged endogenous variable as one of the regressors.

5.3 GENERALIZED LEAST SQUARES (GLS)

We know that each of the equation in the system is a classical regression equation and OLS can be used to estimate consistent, if not efficient parameters for each equation separately. As the generalized regression model can be applied to the stacked model (Eq. 5.2), the efficient estimator is considered as generalized least squares (GLS). Here, the GLS estimator of β is

$$\hat{\beta} = \left(X\Omega^{-1}X\right)^{-1}\left(X\Omega^{-1}Y\right) = \left[X'\left(\Sigma^{-1}\otimes I\right)X\right]^{-1}\left[X'\left(\Sigma^{-1}\otimes I\right)Y\right] \qquad \text{... (5.5)}$$

It is evident from equation (5.5) that the GLS estimator is different from the OLS estimator. If we assume that all classical assumptions hold for the GLS, then the GLS estimators are BLUE. Now, the question that arises here is how efficient is the GLS estimator as compared to OLS in this case. The following conditions have been identified where GLS does not yield efficiency gain over OLS:

(i) Using GLS for estimation of the set of equation is not efficient in case the equation in the system are unrelated i.e., $\sigma_{ij} = 0$ for $i \neq j$. Here, GLS estimation is equivalent to OLS estimation carried out for each equation.

(ii) GLS and OLS are identical in case the regressors are identical in the equations, i.e., $X_i = X_j$

(iii) GLS no longer remain efficient and gets identical with OLS in case of estimation of the regressors of equations which is subset of the regressors in other block of equation within the system of equation.

However, efficiency of GLS increases over OLS in case there is greater correlation between the error terms and less correlation between the regressors of the equations in the system.

5.4 FEASIBLE GENERALIZED LEAST SQUARES (FGLS)

As we have seen in the preceding sections it has been assumed that $\sum$ is known. However, as usual $\sum$ is unknown and is estimated using estimators devised for this known as feasible generalized least squares (FGLS). Elements of $\sum$ are consistently estimated using the least squares residual

$$\hat{\sigma}_{ij} s_{ij} = \frac{e_i' e_j}{T} \qquad \text{… (5.4)}$$

Where, e represents the OLS residuals. Following two possibilities have been recommended for estimation

$$s_{ij}^{*} = \frac{e_i' e_j}{\left(T - K_i\right)\left(T - K_j\right)^{1/2}} \qquad \text{and} \qquad s_{ij}^{**} = \frac{e_i' e_j}{T - \max\left(K_i, K_j\right)}$$

The first possibility is unbiased only in case i is equal to j whereas the second possibility is unbiased if either i is equal to j or K_i is equal to K_j. However, the asymptotic properties of the FGLS estimator do not require estimates of $\sum$ to be unbiased. The estimates just need to be consistent. The consistent estimator which gives us biased estimator is obtained as

$$\hat{s}_{ij} = \frac{e_{it} e_{ij}}{T} \qquad \text{for } i, j = 1, 2 \qquad \ldots (5.5)$$

Thus, the feasible GLS estimates obtained using this consistent estimator of Σ is asymptotically efficient. Iterating this procedure i.e., computing this feasible GLS residuals followed by estimating second round of estimates of Σ using GLS residual (5.5) will give us asymptotically efficient FGLS estimates. The iteration when carried out till convergence results in maximum likelihood estimates of regression coefficients.

5.5 MAXIMUM LIKELIHOOD ESTIMATION

For SURE model maximum likelihood estimates can be obtained by iterating the procedure used for FGLS estimation. It is noteworthy that at each iteration, σ_{ij} is estimated using equation 5.4. As such maximum likelihood does not have any specific advantage over FGLS with reference to its asymptotic properties. Although maximum likelihood is asymptotically equivalent to FGLS, it serves as an alternative to FGLS in case of small samples where it gives different result. Thus it has been a preferred choice for small sample size depending on the type of data set.

Let us consider only one observation for each of the M dependent variables and the associated independent variables. The model for this observation if arranged horizontally can be written as

$$\left[y_1, y_{2,} \ldots\ldots\ldots y_M\right] = \left[X_t^*\right]\left[\pi_1 \pi_2 \ldots\ldots\ldots\ldots \pi_M\right] + \left[\varepsilon_1 \varepsilon_2 \ldots\ldots\ldots\ldots \varepsilon_M\right]_t = \left[X_t^*\right]\pi' + E$$

$$\ldots (5.6)$$

Where X_t^* is the row vector of the full set of all K^* different independent variables of the model. π_i is the column vector of coefficients for the i^{th} equation (in case all the variables are not present in all the equations, π_i will have zero value allowing exclusion restriction). The matrix for the parameter thus has one column for each equation. Let ε_t be the vector of M error. Then $E(\varepsilon\varepsilon_t^{'}) = \sum$. The log of joint normal density of these M errors is

$$LogL = -\frac{M}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2}\varepsilon_i^{'}\Sigma^{-1}\varepsilon_t \qquad \ldots (5.7)$$

As the errors are assumed to be normally distributed, the log likelihood function is

$$LogL = \sum_{t=1}^{T} \log L_t = -\frac{MT}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2}\sum_{t=1}^{T} \varepsilon_i' \Sigma^{-1} \varepsilon_t \qquad \ldots (5.8)$$

Taking the derivatives of equation 5.8 with respect Π and Σ after setting them to 0, the maximum likelihood estimator can be determined.

As we are aware that SURE model is generalized regression model, hence FGLS and the maximum likelihood estimator must be equivalent. Thus to see the difference between the two we need to get a step further and introduce the estimates of Σ in the likelihood function which gives us the concentrated log-likelihood as

$$LogL_c = -\frac{T}{2}\left[M\left(1 + \log(2\pi)\right) + \log|W|\right] \qquad \ldots (5.9)$$

where W is the maximum likelihood estimator of Σ.

5.5.1 Hypothesis Testing

We have seen in the preceding chapters that F statistic is used to test the hypothesis about β, Here likelihood ratio test is used as an alternative for it to test the hypothesis. The likelihood ratio statistics is given as

$$\lambda = -2\left(\log l - \log l_u\right) = T\left(\log\left|\hat{W}_r\right| - \log\left|\hat{W}_u\right|\right) \qquad \ldots (5.10)$$

where $\hat{W}_r$ and $\hat{W}_u$ are the residual sums of the squares and cross product matrices using constrained and unconstrained estimators respectively. The distribution of likelihood ratio statistics is asymptotic with chi-square distribution having degree of freedom equal to number of restrictions.

In the case of estimation of SURE, two types of hypothesis testing need to be carried out:

(i) There is no correlation between the errors of the equations in the system or Σ is a diagonal matrix, thus estimating one equation after the other using OLS would be efficient.

To test the hypothesis that there is no correlation between the errors (Σ is a diagonal), the likelihood ratio statistic for an unrestricted model is given as

$$\lambda_{LR} = T\left[\sum_{i=1}^{M} \log \hat{\sigma}_i^2 - \log\left|\hat{\Sigma}\right|\right] \qquad \ldots (5.11)$$

Where $\hat{\sigma}_i^2$ is $\frac{e_i' e_j}{T}$/T from the OLS of the individual regression and $\hat{\Sigma}$ denotes the maximum likelihood estimator of Σ. The chi-square distribution of the statistic is has M (M-1)/2 degrees of freedom. Breusch and Pagan (1980) suggested an alternative statistic for testing the hypothesis known as Lagrange multiplier (LM) statistic

$$\lambda_{LM} = T\sum_{i=2}^{M}\sum_{J=1}^{i-1} r_{ij}^2 \qquad \ldots (5.12)$$

where r_{ij}^2 is the estimated correlation $\frac{\hat{\sigma}_{ij}}{\left(\hat{\sigma}_{ij}\hat{\sigma}_{jj}\right)^{\frac{1}{2}}}$ and this statistic also has a chi-squared distribution with M (M-1)/2 degrees of freedom. The biggest advantage of this test is that unlike the likelihood ratio test it does not require the maximum likelihood estimator of the Σ, as it is based on the residuals obtained of the OLS estimation of the equations.

(ii) Test of linear restriction on the coefficients.

The general F statistic for hypothesis testing is valid only asymptotically in case of SURE system as it requires using $\hat{\Sigma}$. However, it is possible to test the coefficient restrictions across the equations in the SURE system of equations using the following test statistic

$$W = \left(R\hat{\beta}_{FGLS} - q\right)'\left[R\,\mathrm{var}\left(\hat{\beta}_{FGLS}\right)R'\right]^{-1}\left(R\hat{\beta}_{FGLS} - q\right) \qquad \ldots (5.13)$$

The statistic has an asymptotic chi-square distribution with J degrees of freedom.

5.5.2 Treatment of Autocorrelation

In SURE model, each of the equation in the system may have time series observations which are likely to have autocorrelation. The SURE framework can be extended to treat the problem of autocorrelation. Let us take the case of a SUR equations (5.1). Suppose that in the equation the error term is given as

$$\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + u_{it}$$

where u_{it} is not correlated across observations.

$$E(\varepsilon\varepsilon') = \Omega = \begin{bmatrix} \sigma_{11}\Omega_{11} & \sigma_{12}\Omega_{12} & \cdots & \sigma_{1M}\Omega_{13} \\ \sigma_{21}\Omega_{21} & \sigma_{22}\Omega_{22} & \cdots & \sigma_{2M}\Omega_{13} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1}\Omega_{M1} & \sigma_{M2}\Omega_{M2} & \cdots & \sigma_{MM}\Omega_{M3} \end{bmatrix}_{MTXMT}$$

$$\Omega = \begin{bmatrix} 1 & \rho_j & \cdots & \rho_j^{T-1} \\ \rho_i & 1 & \cdots & \rho_j^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_i^{T-1} & \rho_i^{T-1} & \cdots & 1 \end{bmatrix}_{TXT}$$

The treatment for autocorrelation can be carried out using the following 3 steps:

(i) Each equation in the system should be estimated first using OLS followed by computation of any consistent estimators of ρ. Data for each of the

equation is then transformed using Prais–Winsten or Cochrane-Orcutt to remove autocorrelation.

(ii) Σ is estimated using the transformed data and OLS

(iii) FGLS is used to estimate the parameters based on the estimated Σ and transformed data.

5.6 INTERRELATED FACTOR DEMAND EQUATIONS

It has been observed that in the recent time multivariate regression model have been mainly applied in estimation of demand equations related to either commodity demand or factor demand in production studies. Here we will see application of SUR framework discussed in the preceding sections in estimation of system of inter-related factor demand equations. To develop an understanding on this let us examine in detail a production model.

Let us suppose production being characterized by a particular production function given as $Y = f$ (x). The exercise for minimizing the production cost of producing a specific output at specified rate for a given set of factor prices results into cost minimizing set of factor demands $x_i = x_i$ (Y, p). Then the following cost function will give the total cost of production

$$C = \sum_{i=1}^{M} p_i x_i(Y, p) = C(Y, p) \qquad \ldots (5.14)$$

The cost minimizing factor demands are obtained by Shephard's Lemma, according to which if minimum total cost of production is given by $C(Y, p)$, then in that case cost-minimizing factor demands is given by

$$x_i^* = \frac{\partial C(Y, p)}{\partial p_i} = \frac{Y \partial C(p)}{\partial p_i} \qquad \ldots (5.15)$$

Suitable parameterization of cost function equation (5.14) and cost share equation 5.15 give us M or $M + 1$ equations that can be used to estimate these. Further by expanding the equation we get the translog cost function (for detail see Green (2003), pp. 366-369) which can be written as econometric model

$$- \beta_1 \log p_1 + \ldots\ldots\ldots + \beta_M \log p_M + \delta_{11}\left(\frac{1}{2}\log^2 p_1\right) + \delta_{12} \log p_i \log p_2 + \delta_{12}\left(\frac{1}{2}\log^2 p_2\right) + \ldots\ldots$$

$$\left.\log^2 p_M\right)$$

… (5.16)

The cost share is given as a system of equation which forms a SUR model allowing for estimations as discussed in Section 5.3 to Section 5.5. The cost shares are given by

$$s_1 = \frac{\partial \log c}{\partial \log p_1} = \beta_1 + \delta_{11} \log p_1 + .\delta_{12} \log p_2 . + ... + \delta_{1M} \log p_M$$

$$s_2 = \frac{\partial \log c}{\partial \log p_2} = \beta_2 + \delta_{12} \log p_1 + .\delta_{22} \log p_2 . + ... + \delta_{2M} \log p_M$$

$$\vdots$$

$$s_M = \frac{\partial \log c}{\partial \log p_M} = \beta_M + \delta_{1M} \log p_1 + .\delta_{2M} \log p_2 . + ... + \delta 3_{MM} \log p_M$$

… (5.17)

In addition, a restriction is imposed on the model as the cost share must sum up to 1,

$$\beta_1 + \beta_2 + ... \beta_M = 1$$

$$\sum_{i=1}^{M} \delta_{ij} = 0 \text{ (coloumn sum is equal to 0)}$$

$$\sum_{i=1}^{M} \delta_{ij} = 0 \text{ (row sum is equal to 0)} \quad \text{… (5.18)}$$

Before estimation, the problem arising due to imposition of the restriction (5.18) is handled by dividing the prices of (M – 1) equations by the price of the M^{th} equation. This eliminates the last row and the last column of each of the parameter matrix. Further, by dropping the M^{th} equation, we obtain a non-singular system, thereby avoiding the problem of singularity of the 'error covariance matrix'. The parameters estimated is used to compute the elasticities of substitution

$$\phi_{ij} = \frac{\delta_{ij} + s_i s_j}{s_i s_j}, \phi_{ii} = \frac{\delta_{ii} + s_i (s_i - 1)}{s_i^2} \quad \text{… (5.19)}$$

As the elasticities computed may differ at different data points, one should select the central point of the data or mean of the data to compute the elasticities.

5.7 LET US SUM UP

In this unit we have discussed about system of multivariate regression equation which are essential in explaining some economic phenomenon. The equations in the system may behave independently and also have regressors which are completely different, but still they may be related to each other through the error terms. Such system is known as seemingly unrelated regression (SUR) equations.

The equations are considered to be statistically linked to each other even though structurally they are independent.

Parameters of each of the equations in the system can be estimated individually using OLS. However, as it has been shown in this unit the OLS is not the appropriate estimator for the SUR model. Parameters of the SUR models can be estimated using generalized least squares (GLS), feasible generalized least squares (FGLS) and Maximum likelihood estimation (MLE). Likelihood ratios and Breusch-Pagan test which uses Lagrange multiplier statistics is used for testing the most essential hypothesis that the errors of the equation in the SUR model are unrelated. The problem of autocorrelation in each of the equations can be treated using Prais –Winsten or Cochrane-Orcutt transformation for data of each equation followed by estimation using FGLS. It has been shown that estimators of the SUR model can be applied for estimation of system of inter-related factor demand equation. The cost share equations which typically forms the SUR econometric model is estimated using the FGLS and the estimators are used for computing the substitution elasticities.

5.8 KEY WORDS

Lagrange Multiplier Statistic : A test statistic with large sample justification that can be used to test for omitted variables, heteroskedasticity, and serial correlation, among other model specification problems.

Likelihood Ratio Statistic : A statistic that can be used to test single or multiple hypotheses when the constrained and unconstrained models have been estimated by maximum likelihood.

Maximum Likelihood Method of Estimation (MLE) : A broadly applicable estimation method where the parameter estimates are chosen to maximize the log-likelihood function.

Restricted Model : In hypothesis testing, the regression model obtained after imposing all of the restrictions required under the null.

Seemingly Unrelated Regression : System of equations where individual equations behaves independently but the individual equations may interact with each other as the random error components of some of the equations may be correlated.

Simultaneous Equations Model : Econometric models where the equations are related and cannot be solved independently.

5.9 SOME USEFUL BOOKS

Baltagi, B. H. (2008) "*Econometrics*." Springer. 4th ed., Chapter 10

Greene, W. H., 2003, "*Econometric Analysis.*" Prentice Hall, New Jersy, USA. 5th ed., Chapter 14

Maddala, G. S., 1992, "*Introduction to Econometrics*", Macmillan Publishing Company, New York. 2d ed.,

Wooldridge, Jeffrey M. 2012, "*Introductory Econometrics*", South-Western College Publishing, 2012.

5.10 EXERCISE FOR PRACTICE

1. Explain with example the difference between seemingly unrelated regression equations (SURE) and simultaneous equations.
2. In case of SUR model GLS is considered as efficient estimator compared to OLS. Explain the reason for it. Enumerate the conditions when GLS does not yield efficiency gain over OLS.
3. How would you test the hypothesis that the errors of the equations are not correlated or Σ is a diagonal?
4. Explain the strategy adopted to deal with autocorrelation in data for each of the equation in SUR model.
5. Suppose that we have two equation systems

$$y_1 = \beta_1 x_1 + \varepsilon_1,$$
$$y_2 = \beta_2 x_2 + \varepsilon_2$$

Both this equation satisfies all the assumptions of the classical multivariate regression model and all variables have zero means. From a sample of 20 observations, following sample second-moment matrix is obtained

	y_1	y_2	x_1	x_2
y_1	20	6	4	3
y_2	6	10	3	6
x_1	4	3	5	2
x_2	3	6	2	1

(i) Compute the FGLS estimates of $\beta 1$ and $\beta 2$.

(ii) Compute the maximum likelihood estimates of the model parameters.

(iii) Test the hypothesis that $\beta 1 = \beta 2$ using the likelihood ratio test.

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 6 INTRODUCTION TO SIMULTANEOUS EQUATIONS MODELS*

Structure

6.0 Objectives

6.1 Introduction

6.2 Simultaneous Equations Model (SEM)

6.3 Structural Form and Reduced Form

6.4 Identification Problem

6.4.1 Order Condition

6.4.2 Rank Condition

6.4 General Structure of SEM

6.5 Simultaneity Bias

6.6 Let Us Sum Up

6.7 Key Words

6.8 Some Useful Books

6.9 Answers/Hints to Check Your Progress Exercises

6.0 OBJECTIVES

After reading this unit, you will be able to:

- distinguish between a 'single equation linear regression model' and a 'simultaneous equation model' (SEM);
- specify the 'types of variables' in a SEM with illustrations;
- outline the components of the 'structural form' and the 'reduced form' of SEM, specifying the types of variables in each;
- discuss the 'identification problem' in the SEM;
- describe, with an illustration, the meaning of 'observationally equivalent structural rule';
- present the general structure of equations in a 'simultaneous equation model (SEM)' using notations from linear algebra; and
- explain the phenomenon of 'simultaneous equations bias' in a SEM.

* Prof. Poulomi Roy, Jadavpur University, Kolkata

6.1 INTRODUCTION

In the previous Unit we mentioned about regression models having multiple equations. In these types of regression models, the equations are inter-dependent. For example, in the two-sector Keynesian model, income depends upon consumption and investment ($Y = C + I$). Further, consumption depends upon income($C = \alpha + \beta Y$). Therefore, estimation of a single equation, without considering the other, will lead to inefficient estimators. We need to consider both the equations simultaneously.

Single equation linear regression models express the relationship between one endogenous (dependent) variable and one or more exogenous (independent) variables. But most economic models include certain dependent variables which interact simultaneously on other dependent variables. This means, economics is interdependent in the sense that 'everything depends on everything else'. Hence, economic data is "a product of economic system, described as a 'system of simultaneous relations' among many random variables". These relations involve both the current and past values of many interacting variables.

A single equation model considers one-way causality. The explanatory variables in a single equation model influence the dependent variable but the explanatory variables are not reciprocally influenced by the dependent variable. A restricted framework like this is inconsistent with the data generation process for the more involved interactive economic relations. For instance, in a partial equilibrium framework for a single commodity, we visualize price, quantity-demanded and quantity-supplied as determined simultaneously. We know that it is inappropriate to say that price causes consumption or consumption causes price since the values that these variables take are jointly determined. Such an interactive mechanism can be better explained by a system of relations or equations. Such a system of equations gives due importance to the underlying process of complex economic relations. Such an analytical process is known as a 'simultaneous equations system (SES)' or simultaneous equations model (SEM).

6.2 SIMULTANEOUS EQUATIONS MODEL (SEM)

A simultaneous equations model (SEM) includes the following types of variables:

(i) Endogenous Variables: Endogenous variables are 'jointly determined variables'. They are those variables whose outcome values are determined within the system, through the joint interaction with other variables. A simultaneous equations model (SEM) is formulated with the objective of analysing the behaviour of such endogenous variables. Examples of such variables are price, consumption, production, income, etc.

(ii) Exogenous Variables: These are variables which are determined entirely from outside the system but are known to affect the outcome of the endogenous variables. They are assumed to condition the outcome values

of the endogenous variables; but exogenous variables are not reciprocally affected. The reciprocal effect implies that no feedback relationship is assumed. Examples of this type of variables are weather related variables (e.g. temperature and rainfall), international price of a commodity, etc.

(iii) Lagged Endogenous Variables: These are past values of the endogenous variables. For example, if Y_t is an endogenous variable in time 't', Y_{t-1}, Y_{t-2}, etc. are past values for lagged Y_t's. It is also called lagged dependent variable.

(iv) Pre-determined Variables: These are current and past values of the exogenous variables and lagged values of the endogenous variables. You should remember that the number of pre-determined variables in a SEM is the sum of the number of exogenous variables and lagged endogenous variables.

(v) Error Variables: These are residual variables which appear in the final specification of the model to account for non-observable random errors for shocks or disturbances. We make some specifications about these unobservable random variables to imagine an ideal condition. These conditions are taken as assumptions, and they are in respect of the mean, variances and inter-correlations of the exogenous and pre-determined variables.

A system of simultaneous equations is said to be 'complete' if it has as many equations as there are endogenous variables. Let us consider two examples of simultaneous equations model (SEM).

Example 6.1: Demand-Supply Model

Let q_{dt} be the demand for rice in period t, q_{st} be the supply of rice in period t and q_t be the equilibrium quantity demanded and supplied in period t. The demand-supply model can now be explained with the help of the following system of equations:

Demand: $q_{dt} = \alpha_0 + \alpha_1 p_t + \alpha_2 y_t + \alpha_3 p_{t-1} + u_t$ … (6.1)

Supply: $q_{st} = \beta_0 + \beta_1 p_t + \beta_2 r_t + \beta_3 p_{t-1} + e_t$ … (6.2)

Equilibrium condition is: $q_{dt} = q_{st} = q_t$

Here, p_t is the price in period t, y_t is the income in period t, r_t is the rainfall in period t, p_{t-1} is the price in period $(t-1)$, i.e., the lagged endogenous variable. Note that q_t and p_t are endogenous variables while y_t, r_t, p_{t-1} are predetermined variables. u_t and e_t are the error terms of demand and supply equations respectively.

Example 6.2: **Keynesian Model**

A simple Keynesian model comprises consumption and investment equations and is expressed as:

Consumption function: $c_t = \alpha_0 + \alpha_1 y_t + u_t$

Autonomous investment function: $i_t = \bar{i}$

Equilibrium condition: $y_t = c_t + i_t$

where c_t is consumption in period t, $y_t$is income in period $t$and i_t is autonomous investment in time t taken as equal to $\bar{i}$. In this model, c_t and y_t are endogenous variables, i_t is the exogenous variable and u_t is the random error term.You should note there are two endogenous variables (c_t and y_t) and two regression equations in this model.

6.3 STRUCTURAL FORM AND REDUCED FORM

The set of equations which actually describes the economic phenomenon is called the 'structural form'. In structural equations, the variation of a current endogenous variable is described by:

(i) current endogenous variables,

(ii) past endogenous variables,

(iii) current and past exogenous variables, and

(iv) random error term.

The demand-supply model and the simple Keynesian model discussed above are examples of structural form model. An equation of a structural model is known as the *structural equation* and its coefficients are known as *structural coefficients*. The error terms of the equations are known as *structural disturbances*.

The reduced form of the SEM is obtained from the structural form after eliminating the current endogenous variables from the right-hand side of the structural equations. It implies that in an equation, an endogenous variable is expressed as a function of pre-determined variables and error terms. Thus, in the reduced form, on the right-hand side of an equation we have

(i) past endogenous variables,

(ii) current and past exogenous variables, and

(iii) random error terms.

Let us consider the demand-supply model (**Example 6.1**) once again. To obtain the reduced form of this model, we first equate the demand equation (6.1) and the supply equation (6.2). We obtain

$$\alpha_0 + \alpha_1 p_t + \alpha_2 y_t + \alpha_3 p_{t-1} + u_t = \beta_0 + \beta_1 p_t + \beta_2 r_t + \beta_3 p_{t-1} + e_t$$

On re-arrangement of terms, we obtain the reduced form equation of p_t as follows:

$$(\alpha_1 - \beta_1)p_t = (\beta_0 - \alpha_0) + \beta_2 r_t - \alpha_2 y_t + (\beta_3 - \alpha_3)p_{t-1} + (e_t - u_t)$$

Or, $p_t = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{\beta_2}{\alpha_1 - \beta_1} r_t - \frac{\alpha_2}{\alpha_1 - \beta_1} y_t + \frac{\beta_3 - \alpha_3}{\alpha_1 - \beta_1} p_{t-1} + \frac{e_t - u_t}{\alpha_1 - \beta_1}$ … (6.3).

We can write the above equation as

$$p_t = \pi_{10} + \pi_{11} r_t + \pi_{12} y_t + \pi_{13} p_{t-1} + v_{1t} \quad \text{… (6.4)}$$

where

$\pi_{10} = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}$

$\pi_{11} = \beta_2/(\alpha_1 - \beta_1)$

$\pi_{12} = -\alpha_2/(\alpha_1 - \beta_1)$

$\pi_{13} = (\beta_3 - \alpha_3)/(\alpha_1 - \beta_1)$

$v_{1t} = (e_t - u_t)/(\alpha_1 - \beta_1)$

Look into equation (6.4) carefully. It expresses the endogenous variable p_t as a function of pre-determined variables and error terms only; there is no current endogenous variable on the right-hand side of the equation. Further, note that the coefficients π_{10}, π_{11}, π_{12} and π_{13} are the reduced form coefficients.

Now, let us substitute equation (6.4) in either the demand equation to obtain the reduced form of q_t(we can substitute equation (6.4) in the supply equation also – we will obtain the same result).

From equation (6.1), we get

$$q_t = \alpha_0 + \alpha_1(\pi_{10} + \pi_{11} r_t + \pi_{12} y_t + \pi_{13} p_{t-1} + v_{1t}) + \alpha_2 y_t + \alpha_3 p_{t-1} + u_t.$$

By re-arranging terms in the above, we get

$$q_t = (\alpha_0 + \alpha_1\pi_{10}) + \alpha_1\pi_{11} r_t + (\alpha_1\pi_{12} + \alpha_2) y_t + (\alpha_1\pi_{13} + \alpha_3) p_{t-1} + (\alpha_1 v_{1t} + u_t) q_t \quad \text{… (6.5)}$$

Note that we again have expressed the endogenous variable q_t in terms of pre-determined variables and error terms only. Thus, we can write the reduced form of q_t as

$$q_t = \pi_{20} + \pi_{21} r_t + \pi_{22} y_t + \pi_{23} p_{t-1} + v_{2t} \quad \text{… (6.6)}$$

where

$\pi_{20} = \alpha_0 + \alpha_1\pi_{10}$ $\quad \pi_{21} = \alpha_1\pi_{11}$ $\quad \pi_{22} = \alpha_1\pi_{12} + \alpha_2$

$\pi_{23} = \alpha_1\pi_{13} + \alpha_3$ $\quad v_{2t} = \alpha_1 v_{1t} + u_t$

The reduced form of p_t and q_t are functions of pre-determined variables only (i.e., current and past values of exogenous variables and past values of endogenous variables). The current endogenous variables do not appear on the right-hand side

of the reduced form model. All the predetermined variables of the system appear only on the right-hand side of the reduced form equations.

Since the reduced form equations have only exogenous variables as explanatory variables, we can estimate the reduced form parameters by applying OLS method. Our objective, however, is to estimate the structural equations given at equations. Thus, we should be in a position to obtain estimators for the structural parameters (a_1, b_1, c_1, a_2, b_2 and c_2) from the reduced form parameters (π_1, π_2, π_3, π_4, π_5, and π_6). By using the estimated values of the parameters in the reduced form model, we can obtain the estimated structural form parameters in certain cases; not always. We present some examples below to bring forward this point.

Example: 6.3: Exactly identified model

Let us consider the following model:

$q_d = a_1 + b_1 p + c_1 y + u_1$ … (6.7)

$q_s = a_2 + b_2 p + c_2 r + u_2$ … (6.8)

where q is quantity, p is price, y is income, r is rainfall andu_1 and u_2 are disturbance terms. Here, p and q are jointly determined endogenous variables and y and r are exogenous variables. The exogenous variables should be uncorrelated with the random error terms. Since p is correlated with the residuals u_1 and u_2 (by assumption), we get inconsistent estimators for the parameters. To see how this is, let us now derive the reduced form of p and q as:

$a_2 + b_2 p + c_2 r + u_2 = a_1 + b_1 p + c_1 y + u_1$

Or,$(b_2 - b_1)p = (a_1 - a_2) + c_1 y - c_2 r + (u_1 - u_2)$

Or, $p = \frac{a_1 - a_2}{b_2 - b_1} + \frac{c_1}{b_2 - b_1} y - \frac{c_2}{b_2 - b_1} r + v_2$

We can write the above equation as

$p = \pi_4 + \pi_5 y + \pi_6 r + v_2$ … (6.9)

Substituting (6.9) in equation (6.7) we get

$q = a_1 + b_1(\pi_4 + \pi_5 y + \pi_6 r + v_2) + c_1 y + u_1$

Or, $q = a_1 + b_1\pi_4 + (b_1\pi_5 + c_1)y + b_1\pi_6 r + (u_1 + b_1 v_2)$

Or, $q = \frac{a_1 b_2 - a_2 b_1}{b_2 - b_1} + \frac{c_1 b_2}{b_2 - b_1} y - \frac{c_2 b_1}{b_2 - b_1} r + (u_1 + b_1 v_2)$

We can write the above equation as

$q = \pi_1 + \pi_2 y + \pi_3 r + v_1$ … (6.10)

where$\pi_1 = \frac{a_1 b_2 - a_2 b_1}{b_2 - b_1}$,$\pi_2 = \frac{c_1 b_2}{b_2 - b_1}$,$\pi_3 = \frac{-c_2 b_1}{b_2 - b_1}$, and $v_1 = (u_1 + b_1 v_2)$.

Since the reduced form equations have only exogenous variables as explanatory variables, we can estimate the reduced form parameters by applying OLS method.

Our objective, however, is to estimate the structural equations of the SEM. Thus, we should be in a position to obtain estimators for the structural parameters (a_1, b_1, c_1, a_2, b_2 and c_2) from the reduced form parameters (π_1, π_2, π_3, π_4, π_5, and π_6). By using the estimated values of the parameters in the reduced form model, we can obtain the estimated structural form parameters as follows:

$\hat{b}_1 = \frac{\hat{\pi}_3}{\hat{\pi}_6}$, $\hat{b}_2 = \frac{\hat{\pi}_2}{\hat{\pi}_5}$, $\hat{c}_2 = -\hat{\pi}_6(\hat{b}_1 - \hat{b}_2)$, $\hat{c}_1 = \hat{\pi}_5(\hat{b}_1 - \hat{b}_2)$, $\hat{a}_1 = \hat{\pi}_1 - \hat{b}_1\hat{\pi}_4$ and $\hat{a}_2 = \hat{\pi}_1 - \hat{b}_2\hat{\pi}_4$.

This method of obtaining the structural parameters from reduced form parameters is known as the 'indirect least squares(ILS) method' (we will discuss further on this estimation method in Unit 16).

In the above example we could obtain unique estimators of all the structural parameters. It is however not always possible to get the structural parameters from the reduced form parameters. There could be two different situations: (i) we do not get estimators of some of the structural parameters, and (ii) we get multiple estimators for some of the structural parameters. These two cases are illustrated in Example 6.4 and Example 6.5 respectively.

Example 6.4: Under-identified model

$q_d = a_2 + b_2p + u_2$ … (6.11)

$q_s = a_1 + b_1p + c_1y + u_1$ … (6.12)

In equilibrium, $q_d = q_s$

Or, $a_2 + b_2p + u_2 = a_1 + b_1p + c_1y + u_1$

Or, $(b_2 - b_1)p = (a_1 - a_2) + c_1y + (u_1 - u_2)$

Or, $p = \frac{a_1 - a_2}{b_2 - b_1} + \frac{c_1}{b_2 - b_1}y + \frac{1}{b_2 - b_1}(u_1 - u_2)$ ... (6.13)

The above equation can be written as

$p = \pi_3 + \pi_4y + v_2$ … (6.14)

Now, substituting (6.14) in any of the structural equations we get the reduced form of q (since, $q_s = q_d$). Let us substitute (6.14) in (6.12).

$q_s = a_2 + b_2(\pi_3 + \pi_4y + v_2) + u_2 = (a_2 + b_2\pi_3) + b_2\pi_4y + (u_2 + b_2v_2)$

If we express the reduced form parameters (π_3 and π_4) in terms of structural parameters, as in (6.14), and re-arrange terms, we obtain

$q = \frac{a_1b_2 - a_2b_1}{b_2 - b_1} + \frac{c_1b_2}{b_2 - b_1}y + v_1$ … (6.15)

The above equation can be written as

$q = \pi_1 + \pi_2 y + v_1$ ... (6.16)

Solving for the structural form parameters, in terms of the reduced form parameters, we get:

$\hat{b}_2 = \frac{\hat{\pi}_2}{\hat{\pi}_4}$, $\hat{a}_2 = \hat{\pi}_1 - \hat{b}_2\hat{\pi}_3$.

There is now no way of getting the estimates of a_1, b_1 and c_1. We therefore have a situation where the demand function is estimable, but the supply function is not.

Example 6.5: Under-identified model

Let us consider the following two equations:

$q_d = a_1 + b_1 p + u_1$ … (6.17)

$q_s = a_2 + b_2 p + c_2 r + u_2$ … (6.18)

From equations (6.17) and (6.18) we get the reduced form of p as

$a_1 + b_1 p + u_1 = a_2 + b_2 p + c_2 r + u_2$

Or, $(a_1 - a_2) - c_2 r + (u_1 - u_2) = (b_2 - b_1)p$

Or, $p = \frac{a_1 - a_2}{b_2 - b_1} - \frac{c_2}{b_2 - b_1} r + \frac{1}{b_2 - b_1}(u_1 - u_2)$

We can write the above as

$p = \pi_3 + \pi_4 r + v_2$ … (6.19)

Now substituting (6.19) in (6.18) we get the reduced form of q as

$q = a_2 + b_2(\pi_3 + \pi_4 r + v_2) + c_2 r + u_2$

$= (a_2 + b_2\pi_3) + (b_2\pi_4 + c_2)r + (b_2 v_2 + u_2)$

We can write the above as

$q = \pi_1 + \pi_2 r + v_1$ … (6.20)

Now, we can estimate $\hat{b}_1 = \frac{\hat{\pi}_2}{\hat{\pi}_4}$, $\hat{a}_1 = \hat{\pi}_1 - \hat{b}_1\hat{\pi}_3$ but we cannot estimate a_2, b_2, c_2.

Hence, in this situation, the demand function is estimable, but the supply function is not.

Example 6.6: Over-identified Model

Now, let us consider another demand supply model:

$q_s = a_1 + b_1 p + c_1 y + d_1 r + u_1$ … (6.21)

$q_d = a_2 + b_2 p + u_2$ … (6.22)

By equating (6.21) and (6.22) we get the reduced form of p as follows:

$a_1 + b_1 p + c_1 y + d_1 r + u_1 = a_2 + b_2 p + u_2$

Or, $(a_1 - a_2) + c_1 y + d_1 r + (u_1 - u_2) = (b_2 - b_1)p$

Or, $p = \frac{a_1 - a_2}{b_2 - b_1} + \frac{c_1}{b_2 - b_1} y + \frac{d_1}{b_2 - b_1} r + \frac{1}{b_2 - b_1}(u_1 - u_2)$

We can write the above equation as

$$p = \pi_4 + \pi_5 y + \pi_6 r + v_2 \quad \ldots (6.23)$$

Now, substituting (6.23) in (6.22) we get

$$\begin{aligned} q &= a_2 + b_2(\pi_4 + \pi_5 y + \pi_6 r + v_2) + u_2 \\ &= (a_2 + b_2\pi_4) + b_2\pi_5 y + b_2\pi_6 r + (b_2 v_2 + u_2) \\ &= \pi_1 + \pi_2 y + \pi_3 r + v_1 \quad \ldots (6.24) \end{aligned}$$

If we try to solve the structural form parameters, in terms of the above reduced form parameters, then we get two estimates of b_2 as

$\hat{b}_2 = \frac{\hat{\pi}_2}{\hat{\pi}_5}$ and $\hat{b}_2 = \frac{\hat{\pi}_3}{\hat{\pi}_6}$.

For each of these $\hat{b}_2$, we get one estimate of $\hat{a}_2 = \hat{\pi}_1 - \hat{b}_2\hat{\pi}_4$ and we get no estimate of the parameters of the demand function, i.e.,a_1, b_1, c_1, d_1.

In Example 6.3 we could find estimators for parameters of the demand function and the supply function. In Example 6.4 and Example 6.5 we could find estimators for parameters of the demand function while estimators for the parameters of the supply function could not obtained. In these cases the demand function is identifiable but the supply function is not identifiable. In the case of Example 6.6 we could find more than one estimator for the parameters of the demand function but no estimator for the supply function. In this case the demand function is over-identified while the supply function is under-identified. This leads us to the problem of identification which we discuss in Section 6.4.

Check Your Progress 1 [answer within the space given in about 50-100 words]

1) Distinguish between a 'single equation linear regression model' and a 'simultaneous equations regression model'.

……
……
……

2) Specify the 'types of variables' in a 'simultaneous equations system'.

……
……
……

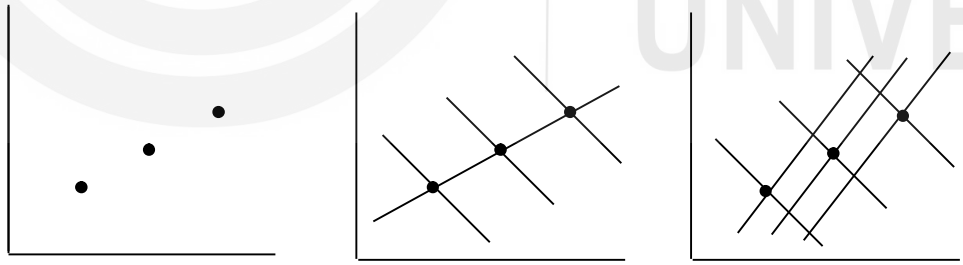
3) Consider the following simultaneous equations system, where c_t and y_t are endogenous variables while i_t is exogenous.
Consumption function: $c_t = \alpha_0 + \alpha_1 y_t + u_t$

Autonomous investment function: $i_t = \bar{i}$
Equilibrium condition: $y_t = c_t + i_t$
Find out the reduced form of the model.

……
……
……
……
……

6.4 IDENTIFICATION PROBLEM

Let us consider the demand function$q = \alpha + \beta p + u$where we assume that p and u are correlated. To measure the coefficients of the demand equation, suppose we use time series data on the quantity bought of the commodity at different prices. Quantity bought, however, is identical with the quantity sold at the corresponding price. Thus, market data register points of interaction of equilibrium supply and demand at prevailing prices at certain points of time. That is, it only shows the points of interactions of demand and supply. If we use these data for estimation, we actually measure the coefficients of a function of the form Q=(p). This equation may be either the demand function or the supply function. How can we be sure that this equation represents demand function or supply function? If one is interested in measurement of the demand function, then (s)he can use the data. Similarly, a person interested in measurement of the supply equation will also be using the same data. Therefore, there is a need for some criteria, which will enable us to verify whether the estimated coefficients belong to a particular relationship or another. Such criteria are known as 'identification conditions' of an equation in a simultaneous equations model.



(a) (b) (c)

Fig. 6.1: Supply and Demand Functions

We begin with the concept of identification using Fig. 6.1. The observed data consist of the market outcomes as shown in Fig. 6.1. Fig. 6.1(a) shows three data points. From the Fig. 6.1(b) and 14.1(c) we can see that we can get the same data for two different situations. In Fig. 6.1(b) the supply curve is stable, and the demand curve is shifting. The data represents points on the supply curve under the assumption that the supply function is fairly stable and the demand curve shifts. In Fig. 14(c) both demand curve and supply curve are shifting randomly. If

we wish to estimate $Q = f(p)$ from the same data set then we estimate neither the demand nor the supply curve; rather we end up estimating a mongrel equation (i.e., a combination of two functions).

A second inference that we draw from Fig. 15(b) is very important: the shift in the demand curve helps us in identification of the supply curve. There is a certain variable in the demand equation (that is excluded from the supply equation), which causes a shift in the demand curve. The result is paradoxical – the exclusion of certain variables from the supply curve helps us in identification of the supply curve. This is often termed as the 'paradox of identification'. To identify an equation, we have to look at the variables that are absent from the equation (while being present in other equations of the model). In identification the relevant variables are those which are absent from the equation!

We can make three different statements on the status of identification of equations in a SEM.

(i) If we do not get estimator of the structural form parameters in terms of the reduced form parameters, then the equation is **under-identified**.

(ii) If we get unique estimator of the structural form parameters in terms of the reduced form parameters of an equation, then the equation is **exactly identified**.

(iii) If we get multiple estimators of the structural form parameters in terms of the reduced form parameters, then the equation is **over-identified.**

It is quite tedious and time consuming to transform the structural form of the SEM to its reduced form and then examine whether structural parameters can be expressed in terms of reduced form parameters. There is a counting rule, which tells us the identification status of an equation on the basis of variables excluded from an equation. There are two conditions which should be checked: (i) order condition, and (ii) rank condition.

6.4.1 Order Condition for Identification

For an equation to be identified the total number of variables (endogenous and exogenous) excluded from the equation must be equal to or greater than the number of endogenous variables in the model less one. Since the number of endogenous variables is equal to the number of equations in a SEM, the order condition for identification can be stated as follows:

For an equation to be identified the total number of variables excluded from it but included in other equations should be equal to more than the number of equations of the system less one.

Let G =total number of equations (= total number of endogenous variables)

K =total number of variables in the model (endogenous and predetermined)

M = the number of variables (endogenous and exogenous) included in a particular equation.

Then the order condition for identification is

$$(K - M) \geq (G - 1) \quad \text{... (6.25)}$$

If there is equality (i.e., $(K - M) = (G - 1)$) the equation is exactly identified. If there is inequality (i.e., $(K - M) > (G - 1)$) the equation is over identified.

Example 6.8

There are10 equations in a simultaneous equations model with 15 variables (10 endogenous and 5 exogenous).

(a) Check the identification status of the first equation in the model, if it contains 11 variables.

For the first equation, G = 12, K = 15 and M = 11.

The **order condition** is given by$(K - M) \geq (G - 1)$. By substituting the values in the equation, we obtain (15–11) < (10–1). Thus the first equation in the model is not identified.

(b) Check the identification status of the second equation if it includes 5 variables.

G **= 10** K **= 15** M **= 5**

Order condition: $(K - M) \geq (G - 1)$

(15–5) > (10–1)

Thus, the second equation is over-identified.

The order condition for identification is a necessary condition for an equation to be identified; but it is not a sufficient condition. It means that the order condition may be fulfilled in an equation; yet the relation may not be identified.

6.4.2 The Rank Condition for Identification

According to the rank condition, in a system of G equations a particular equation is identified if and only if the rank of the matrix of parameters of the excluded variables (endogenous and predetermined) from that equation is equal to $(G - 1)$. In other words, an equation is identified if it is possible to construct at least one non-zero determinant of order $(G - 1)$ from the coefficients of the variables excluded from that particular equation but contained in the other equations of the model. The rank condition is a necessary and sufficient condition for identification.

When you are required to check the identification status of an equation in a SEM, there are certain steps you should follow. For illustration purpose we take the following SEM.

$$y_1 = 3y_2 + 2x_1 + x_2 + u_1$$

$$y_2 = y_3 + x_3 + u_2$$

$$y_3 = y_1 - y_2 - 2x_3 + u_3 \quad \text{... (6.26)}$$

There are 3 endogenous variables in the model (G=3); total number of variables is 6 (implies, K=6).

Step1: Write down the parameters of all the equations of the model in a separate table, noting that the parameter of a variable excluded from an equation is equal to zero.

Thus, the SEM given at (6.26) can be written as

$$-y_1 + 3y_2 + 0y_3 - 2x_1 + x_2 + 0x_3 + u_1 = 0$$

$$0y_1 + y_2 + y_3 - 0x_1 + 0x_2 + x_3 + u_2 = 0$$

$$y_1 + y_2 + y_3 - 0x_1 + 0x_2 + 2x_3 + u_3 = 0 \quad \text{… (6.27)}$$

Table 6.1: Parameters of the Equation System

Equations	Variables					
	y_1	y_2	y_3	x_1	x_2	x_3
1st equation	–1	3	0	–2	1	0
2nd equation	0	–1	1	0	0	1
3rd equation	1	–1	–1	0	0	–2

You should prepare a table similar to Table 6.1. Note that we ignore the error terms while checking the identification status of an equation. Let us assume that we have to check the identification of the second equation in the above SEM.

Step2: Eliminate the row of coefficients of the equation, which is being examined for identification.

For example, for checking the identification status of the second equation, we strike out the second row of the coefficients in Table 6.1.

Step3: Eliminate the columns in which a non-zero coefficient of the equation being examined appears. By deleting the relevant row and columns we are left with the coefficients of variables not included in the particular equation, but contained in the other equations of the model.

For example, for identification the second equation of the system, we will eliminate the second, third and the sixth columns of Table 6.1. After carrying out Steps 2 and 3, we obtain Table 6.2.

Table 6.2: Strike-out the Rows and Columns

Equations	Variables					
	y_1	y_2	y_3	x_1	x_2	x_3
1st equation	–1	3	0	–2	1	0
2nd equation	0	–1	1	0	0	1
3rd equation	1	–1	–1	0	0	–2

Step4: Form the determinant(s) of order $(G-1)$ and examine their value. If at least one of these determinants is non-zero, the equation is identified. If all the determinants of order $(G-1)$ are zero, the equation is under-identified.

Table 6.3: Remaining Parameters

y_1	x_1	x_2
−1	−2	1
1	0	0

We have three determinants of order $(G-1) = 3-1=2$ on the basis of Table 6.3.

They are

$$\Delta_1 = \begin{vmatrix} -1 & -2 \\ 1 & 0 \end{vmatrix} \neq 0$$

$$\Delta_2 = \begin{vmatrix} -2 & 1 \\ 0 & 0 \end{vmatrix} = 0$$

$$\Delta_3 = \begin{vmatrix} -1 & 1 \\ 1 & 0 \end{vmatrix} \neq 0$$

(the symbol 'Δ' stands for determinant). We see that we can form two non-zero determinants of order $G-1 = 3-1 = 2$; hence the second equation of our system is identified.

Step5: To see whether the equation is exactly identified or over-identified we use the order condition $(K-M) > (G-1)$.With this criterion, if the equality sign is satisfied, that is, if $(K-M) = (G-1)$, the equation is exactly identified. If the inequality sign holds, that is, if $(K-M) > (G-1)$, then the equation is over-identified.

In the case of the second equation we have

$G = 3$ $K = 6$ $M = 3$

and the counting rule $(K-M) \geq (G-1)$ gives

$(6-3) > (3-1)$.

Therefore the second equation of the model is over-identified.

Check Your Progress 2[answer within the space given in about 50-100 words]

1) What is meant by the 'identification problem' in a SES?

……
……
……
……
……

2) What is a 'necessary condition' for identification in a SEM?

……
……
……
……

……

3) What is a 'sufficient condition' for an equation of SES to be identified?

.....
.....
.....
.....

4) Consider the three-equation model

$$Y_1 = \beta_{13}Y_3 + \gamma_{12}X_2 + u_1$$
$$Y_2 = \beta_{21}Y_1 + \beta_{23}Y_3 + \gamma_{21}X_1 + \gamma_{22}X_2 + u_2$$
$$Y_3 = \gamma_{33}X_3 + u_3$$

where Y_1, Y_2 and Y_3 are endogenous, and X_1, X_2 and X_3 are exogenous variables. Discuss the identification status of each equation of the model, based on the order and the rank condition.

.....
.....
.....
.....

4) The structure of a model with four endogenous and three exogenous variables is as follows (1 indicates presence and 0 absence of the variable in the equation):

1 0 1 1 1 0 0

1 1 1 0 0 1 1

0 0 1 0 1 0 0

1 0 1 1 0 1 0

Which of the above four equations is/are identified?

.....
.....
.....
.....

6.4 GENERAL STRUCTURE OF SEM

Suppose there are G number of endogenous variables, Y_1, Y_2,….,Y_G and K number of exogenous variables, X_1,X_2,….,X_K in a simultaneous equations system.

Thus, the total number of equations in the system are G. Let $\beta_1, \beta_2, \ldots\ldots, \beta_G$ be the coefficients associated with the endogenous variables and $\gamma_1, \gamma_2, \ldots\ldots, \gamma_k$ be the coefficients associated with the exogenous variables. The system of SEM at any time point t can be written as:

$$\begin{aligned} Y_1\beta_{11} + Y_2\beta_{21} + \ldots + Y_G\beta_{G1} + X_1\gamma_{11} + X_2\gamma_{21} + \ldots + X_K\gamma_{K1} + e_1 &= 0 \\ Y_1\beta_{12} + Y_2\beta_{22} + \ldots + Y_G\beta_{G2} + X_1\gamma_{12} + X_2\gamma_{22} + \ldots + X_K\gamma_{K2} + e_2 &= 0 \\ \vdots & \\ Y_1\beta_{1G} + Y_2\beta_{2G} + \ldots + Y_G\beta_{GG} + X_1\gamma_{1G} + X_2\gamma_{2G} + \ldots + X_K\gamma_{KG} + e_G &= 0 \end{aligned}$$

... (6.28)

We assume that the structural disturbances are generated by a stationary multivariate process such that:

(i) $E(e_i) = 0$, $i = 1,2,3,\ldots,G$

(ii) $E(e_i e_i') = \sigma_{ii} I_T$ for $i = 1,2,3,\ldots,G$

(iii) $E(e_i e_j') = \sigma_{ij} I_T$, for $i \neq j$, I,j=1,2,3,..,G

Thus, if $e' = \left(e_1' e_2' e_3' \ldots .. e_G'\right)$

Then $E(ee') = E\begin{bmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_G \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_G \end{pmatrix}' \end{bmatrix} = \begin{bmatrix} \sigma_{11}I & \sigma_{12}I & \cdots & \cdots & \sigma_{1G}I \\ \sigma_{21}I & \sigma_{22}I & \cdots & \cdots & \sigma_{2G}I \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \sigma_{G1}I & \sigma_{G2}I & \cdots & \cdots & \sigma_{GG}I \end{bmatrix} = \sum \otimes I_T$

... (6.29)

The contemporaneous covariance matrix $\sum$ is a $G \times G$ symmetric positive semi-definite matrix. It may have less than full rank as some of the equations may appear in the form of identities with null error vectors. In a matrix form, the system of simultaneous equations model, for all time points, can be written as:

$YB + X\Gamma + E = 0$

where 0 is a TxG matrix of zeros,

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & \cdots & y_{1G} \\ y_{21} & y_{22} & \cdots & \cdots & y_{2G} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ y_{T1} & y_{T2} & \cdots & \cdots & y_{TG} \end{bmatrix} = (y_1 y_2 \ldots\ldots y_G)$$

and

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1K} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2K} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ x_{T1} & x_{T2} & \cdots & \cdots & x_{TK} \end{bmatrix} = (x_1 x_2 \ldots\ldots x_K)$$

Y and X are sample values taken by the jointly determined variables and the pre-determined variables, respectively, and

$$E = \begin{bmatrix} e_{11} & e_{12} & \cdots & \cdots & e_{1G} \\ e_{21} & e_{22} & \cdots & \cdots & e_{2G} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ e_{T1} & e_{T2} & \cdots & \cdots & e_{TG} \end{bmatrix} = (e_1 e_2 \ldots . e_G)$$

is the matrix of unobservable values taken by the random error vectors. B is a $G \times G$ matrix of coefficients of the current endogenous variables, where each column refers to the coefficients of a particular equation. Specifically:

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1G} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2G} \\ \vdots & \vdots & \cdots & \vdots \\ \beta_{G1} & \beta_{G2} & \cdots & \beta_{GG} \end{bmatrix} = [\beta_1 \beta_2 \cdots \cdots \beta_G] \quad \text{... (6.30)}$$

is the $G \times G$ matrix of structural coefficients associated with the endogenous variables and each column contains coefficients of a particular equation. Further,

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1G} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2G} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{K1} & \gamma_{K2} & \cdots & \gamma_{KG} \end{bmatrix} = [\gamma_1 \gamma_2 \cdots \cdots \gamma_K] \quad \text{... (6.31)}$$

is the $K \times G$ matrix of structural coefficients associated with the exogenous-predetermined variables where each column contains the coefficients of a particular equation. For the system to be complete, B must be of order G. In addition, if it is non-singular, then we may rewrite the system of equations as:

$$YB + X\Gamma + E = 0 \quad \text{... (6.32)}$$

If we post-multiply the above equation by B^{-1}, we obtain

$$YBB^{-1} + X\Gamma B^{-1} + EB^{-1} = 0$$

Or, $Y + X\Gamma B^{-1} + EB^{-1} = 0$

Or, $Y = -X\Gamma B^{-1} - EB^{-1}$

We can write the above equation as

$$Y = X\pi + V \quad \text{… (6.33)}$$

Equation (6.33) is the reduced form of the simultaneous equations model

where

$$\pi = -\Gamma B^{-1} = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1G} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{K1} & \pi_{K2} & \cdots & \pi_{KG} \end{bmatrix} = (\pi_1 \pi_2 \cdots \pi_G)$$

is a $K \times G$ matrix of reduced form coefficients.

$V = -EB^{-1}$ is a $T \times G$ matrix of error terms. That is:

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1G} \\ v_{21} & v_{22} & \cdots & v_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ v_{T1} & v_{T2} & \cdots & v_{TG} \end{bmatrix} = (v_1 v_2 \cdots v_G)$$

The G equations in (6.33) are called the *reduced form equations.* The matrix $\pi = -\Gamma B^{-1}$is the matrix of reduced form parameters. The matrix V represents the reduced form disturbances. With this transformation, we can express each endogenous variable as a linear function of the predetermined variables in the system.

The system of reduced form equations may be rewritten as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_G \end{bmatrix} = \begin{bmatrix} X & & & \\ & X & & \\ & & \ddots & \\ & & & X \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_G \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_G \end{bmatrix}$$

Or, $y = (I \otimes X)\pi + v$ ... (6.34)

6.5 SIMULTANEITY BIAS

In the reduced form we keep one of the endogenous variables on the left-hand side while the predetermined variables (exogenous variables and lagged dependent variables) are kept on the right-hand side in an equation. The coefficient of the endogenous variables on the left-hand side is 1. To keep the dependent variable of the i^{th} structural equation on the left-hand side and all other variables on the right-hand side, we set the coefficient of the i^{th} endogenous variables y_i to be equal to –1. This is called the 'normalization rule'.

The i^{th} equation can therefore be written as:

$YB_i + X\Gamma_i + e_i = 0$ ... (6.35)

$$y_i = Y_i\beta_i + Y_i^*\beta_i^* + X_i\gamma_i + X_i^*\gamma_i^* + e_i$$

$$y_i = Y_i\gamma_i + X_i\beta_i + e_i = (Y_i \quad X_i)\begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} + e_i$$

Or, $y_i = Z_i\delta_i + e_i$ ... (6.36)

Here, $B_i = \begin{bmatrix} -1 \\ \beta_i \\ \beta_i^* \end{bmatrix} = \begin{bmatrix} -1 \\ \beta_i \\ 0 \end{bmatrix}$, $\Gamma_i = \begin{bmatrix} \gamma_i \\ \gamma_i^* \end{bmatrix} = \begin{bmatrix} \gamma_i \\ 0 \end{bmatrix}$

$Y = \begin{bmatrix} y_i & Y_i & Y_i^* \end{bmatrix}$ and $X = \begin{bmatrix} y_i & Y_i & Y_i^* \end{bmatrix}$. $G = G_i^{\Delta} + G_i^{\Delta\Delta}, K = K_i^* + K_i^{**}$, Y_i^* are the $G_i^{\Delta\Delta}$ endogenous variables that appear in the i^{th} equation with zero coefficients, i.e.,they do not appear in the i^{th} equation. X_i^{**} are those exogenous or predetermined variables that appear in the i^{th} equation with zero coefficients. $y_i \quad Y_i \quad X_i$ are the endogenous variables that appear in the i^{th} equation with non zero coefficients. Consequently, y_i is a $T \times 1$ vector, Y_i is a $T \times (G_i^{\Delta} - 1)$ matrix,

β_i is a $(G_i^{\Delta} \times 1)$ vector, e_i is an unobservable $(T \times 1)$ random vector with mean vector zero and covariance $\sigma_i^2 I_T$, $Z_i = (Y_i \quad X_i)$ and $\delta_i = \begin{bmatrix} \beta_i \\ \gamma_i \end{bmatrix}$.

Under this specification, the least square estimator of δ_i is

$$\hat{\delta}_i = (Z_i Z_i')^{-1} Z_i' y_i \text{m} \qquad \text{... (6.37)}$$

The mean of random vector $\hat{\delta}_i$ is

$$E(\hat{\delta}_i) = E\left[(Z_i Z_i')^{-1} Z_i' y_i\right] = E\left[(Z_i Z_i')^{-1} Z_i' (Z_i \delta_i + e_i)\right] = \delta_i + E\left[(Z_i Z_i')^{-1} Z_i' e_i\right]$$

The last term does not vanish because Z_i contains endogenous variables that are jointly determined with y_i. They are thus not independent of e_i. Consequently, the estimator $\hat{\delta}_i$ is biased, i.e., $E(\hat{\delta}_i) \neq \delta_i$. Further, as the sample size increases, $\hat{\delta}_i$ does not converge to δ_i even in probability, since:

$$\text{plim}(\hat{\delta}_i) = \delta_i + p\text{lim} T(Z_i Z_i')^{-1} \text{ plim} T^{-1} Z_i' e_i \neq \delta_i$$

Specifically, plim $T^{-1} Z_i' e_i$ does not converge to zero vector because $Z_i = (Y_i \quad X_i)$ contains Y_i that is not independent of the error vector e_i. Therefore, the least squares estimator applied directly to a structural equation containing two or more endogenous variables will be inconsistent. More specifically, the statistical property of unbiasedness and consistency are not fulfilled.

Example 6.9

To illustrate the above, let us consider a simple macro model of income determination consisting solely of a consumption function and national income identity.

$C_t = \alpha + \beta Y_t + u_t$

$Y_t \equiv C_t + I_t$

Here, C_t= aggregate consumption expenditure at time t,

Y_t=national income at time point t,

I_t=investment expenditure at time point t and

u_t=error term at time point t.

We classify C_t and Y_t as endogenous variables and I_t as exogenous variable. We make two assumptions, namely: $u \sim N(0, \sigma_u^2 I)$ and I and u are independent. The first assumption implies that:

(i) $E(u_t) = 0$

(ii) $V(u_t) = \sigma_u^2$

(iii) $\text{cov}(u_t, u_s) = 0, t \neq s$

The second assumption will be satisfied if either z is a set of fixed numbers, or I is a random variable distributed independently of u. Hence, $E(u_t, I_t) = 0$. We

shall now derive the reduced form of this SEM for which we will substitute the consumption function in the income identity.

$Y_t = C_t + I_t = \alpha + \beta Y_t + u_t + I_t$

Or, $(1-\beta)Y_t = \alpha + u_t + I_t$

Or, $Y_t = \frac{\alpha}{1-\beta} + \frac{I_t}{1-\beta} + \frac{u_t}{1-\beta}$ …(6.38)

Substituting (6.38) in the 'consumption function' we get:

$C_t = \alpha + \beta(\frac{\alpha}{1-\beta} + \frac{I_t}{1-\beta} + \frac{u_t}{1-\beta}) + u_t$

Or, $C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I_t + \frac{1}{1-\beta} u_t$

We can write the above equation as

$C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I_t + v_t$ … (6.39)

where $v_t = \frac{1}{1-\beta} u_t$. Now from reduced form of Y_t is given by:

$Y_t = \frac{\alpha}{1-\beta} + \frac{I_t}{1-\beta} + \frac{u_t}{1-\beta}$ ... (6.40)

If we take expected value of equation (6.40) we obtain

$E(Y_t) = \frac{\alpha}{1-\beta} + \frac{I_t}{1-\beta}$

Therefore, $Y_t - E(Y_t) = \frac{u_t}{1-\beta}$

$cov(Y_t, u_t) = E\left(\left(\frac{1}{1-\beta} u_t\right) u_t\right) = \frac{1}{1-\beta} E(u_t^2) = \frac{\sigma_u^2}{1-\beta} \neq 0$ ... (6.41)

Let us now consider the consumption function in its structural form, i.e.,

$C_t = \alpha + \beta Y_t + u_t.$

On applying OLS to this equation, we have:

$\widehat{\beta} = \frac{\sum(c_t - \overline{c})(Y_t - \overline{Y})}{\sum(Y_t - \overline{Y})^2} = \frac{\sum(c_t - \overline{c}) y_t}{\sum y_t^2}$ ... (6.42)

where

$y_t = Y_t - \overline{Y} = \frac{\sum c_t y_t}{\sum y_t^2} - \overline{c} \frac{\sum y_t}{\sum y_t^2} = \frac{\sum c_t y_t}{\sum y_t^2}$

Since $\sum y_t = 0$, we have

$\frac{\sum c_t y_t}{\sum y_t^2} = \frac{\sum(\alpha + \beta Y_t + u_t) y_t}{\sum y_t^2}$

$= \frac{\alpha \sum y_t + \beta \sum Y_t y_t + \sum u_t y_t}{\sum y_t^2}$

$= \frac{\alpha \sum y_t + \beta \sum (y_t + \overline{Y}) y_t + \sum u_t y_t}{\sum y_t^2} = \frac{\alpha \sum y_t + \beta \sum y_t^2 + \beta \overline{Y} \sum y_t + \sum u_t y_t}{\sum y_t^2}$

$\widehat{\beta} = \beta . 1 + \frac{\sum u_t y_t}{\sum y_t^2}$ ... (6.43)

Let us take *plim* on both sides of equation (6.43) to check the consistency property of the estimators. We get

$$\text{plim}\widehat{\beta} = \text{plim}\beta + \text{plim}\left(\frac{\sum\frac{y_t u_t}{n}}{\sum\frac{y_t^2}{n}}\right)$$

$$=\beta + \frac{plim\sum\frac{u_t y_t}{n}}{plim\sum\frac{y_t^2}{n}} = \beta + \frac{\frac{\sigma_u^2}{1-\beta}}{\sigma_y^2} \quad \text{... (6.44)}$$

Now, $\frac{1}{1-\beta}\frac{\sigma_u^2}{\sigma_y^2} \neq 0$.

Therefore, $\left(\text{plim}\hat{\beta} - \beta\right) \neq 0$ even if $n \to \infty$.Thus, the bias will not vanish even if we increase the sample size to ∞.This bias is known as the 'simultaneous equation bias' or simultaneity bias.

Check Your Progress 2 [answer within the space given in about 50-100 words]

1) If the structural model in matrix form for a system of simultaneous equations is given by $\boldsymbol{YB + X\Gamma + E = 0}$, write down the reduced form equations.

.....
.....
.....
.....
.....

2) What is meant by 'simultaneity bias'?

.....
.....
.....
.....
.....

6.6 LETUSSUMUP

In estimating the parameters of a system of simultaneous equations, several problems arise. First, the direct application of the least squares rule to a structural equation may result in an estimator that is biased. Further, the ordinary least square estimator will not converge to the true parameter even when the sample size is increased indefinitely. In searching for consistent estimator, one possibility is to: (i) find out the reduced form model and then used the estimated values of the reduced form parameters (obtained by the least square method) and then (ii) use the reduced form parameters to solve the structural form parameters. But this also may not always work as we may not get unique solution of the structural form parameters. This raises the problem of identification.

Identification is a problem of model formulation, rather than of model estimation or evaluation. An econometric model is identified if it is in a unique statistical

form, enabling unique estimates of its parameters to be subsequently made from sample data. If a model is not identified then we cannot say what relationship we are estimating.

The counting rule (necessary condition) for an equation with a system of G equations to be just identified is that there should be exactly G–1 variables not appearing in the equation in order to differentiate that equation from the remaining G–1 equations. The rank condition requires that the excluded variables of the i^{th} equation should appear in the remaining G–1 equations in such a way that no other equation (or linear combination of equations) appears in any other equation. Finally, the unit explains how in the presence of simultaneity, the LSEs produce biased estimates of parameter which even with $n \rightarrow \infty$, continues to remain. This is termed as the 'simultaneous equation bias' or simply 'simultaneity bias'.

6.7 KEYWORDS

Simultaneous Equations Model (SEM) : This is a model which involves different type of variables like endogenous variables, lagged endogenous variables, exogeneous variables, pre-determined variables which could be both the exogeneous variables and the endogenous variables with lags, along with the error term.

Structural and Reduced Form Models : Structural form model refers to the system or the set of equations which describes the actual economic phenomena. An equation of the structural model is known as the structural equation; its coefficients are known as the structural coefficients and its error term called the structural disturbance. The reduced form of the model is obtained from the structural form model by eliminating the current endogenous variables from the structural equations.

Identification Problem : We consider the reduced form models for arriving at the estimates of the structural parameters. This method, being indirect, is called the method of 'indirect least squares'. However, getting unique estimates of the structural parameters is possible only when the system of simultaneous equations is 'complete'. This is the case when the number of equations in the system are exactly as many as there are endogenous variables. When this requirement is not fulfilled, we get situations in which either there are no estimates of structural parameters or there are multiple estimates of the

same. Both these situations are referred to as the problem of identification. A 'no estimate' situation is termed as the 'under identification problem' and 'multiple estimates' situation termed as the problem of 'over identification'. We get unique estimates of the structural parameters only when the equations are exactly identified or the when the system of simultaneous equation is 'complete'.

Simultaneity Bias or Simultaneous Equation Bias : Refers to a situation where, in the presence of simultaneity, for the OLS estimate of β, $\hat{\beta}$, the $\left(plim\hat{\beta} - \beta\right) \neq 0$ even if $n \rightarrow \infty$. Thus, if the LSE bias refers to the violation of the property of unbiasedness of the estimate, the simultaneity bias refers to the violation of the property of the consistency of the estimated value of β.

Exogenous Variables : These are variables determined outside the model. The values of the exogenous variables are assumed to be given.

Endogenous Variables : The values of the endogenous variables are determined within the model.

6.8 SOME USEFUL BOOKS

Wooldridge, Jaffery (2010) *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

A. Collin Cameron and P. K. Trivedi (2005). *Micro Econometric Methods and Applications*, Cambridge University Press.

W. H. Greene (2018). *Econometric Analysis*, Eighth Edition, Pearson Education.

6.9 ANSWERS/HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) The former expresses the relationship between one endogenous variable and one or more exogenous variables. It considers one-way causality. This means explanatory variables exert influence on the dependent variable, but they are not reciprocally influenced by it. The latter, i.e., the SES or the SEM considers the reciprocity. SEM allows for an interactive mechanism by taking into account the feedback among interdependent variables. In other words, SEM allows for two-way causality.

2) Endogenous variables (or jointly determined variables), exogenous variables (which are from outside the system without any reciprocal effect on the endogenous variable), lagged endogenous variables, pre-determined variables and an error variable to account for all residual effect.

3) A SES is said to be complete if it has as many equations as there are endogenous variables. Such a situation also helps us in getting unique estimates. Identification problem relates to getting unique estimates of structural parameters from the reduced equations parameters. This is possible only when the equations are exactly identified. In cases of over identification and under identification, we either get multiple estimates or no estimates respectively. In both these cases, we say that there is an identification problem.

Check Your Progress 2

1) Substitute the consumption function $\boldsymbol{c_t = \alpha_0 + \alpha_1 y_t + u_t}$ in $\boldsymbol{y_t = c_t + i_t}$. The reduced form equations are $\boldsymbol{y_t = \frac{a_0}{1-a_1} + \frac{1}{1-a_1} i_t + \frac{1}{1-a_1} u_t}$ and $\boldsymbol{c_t = \frac{a_0}{1-a_1} + \frac{a_1}{1-a_1} i_t + \frac{1}{1-a_1} u_t}$.

2) To form a matrix of rank *G–1*, we must have at least *G–1* number of columns in the Δ matrix. Thus, there should be at least (*G–1*) missing variables from the equation under consideration. This is known as the 'order condition' which is a necessary condition for identification.

3) A sufficient condition for an equation in the system to be identified is that the matrix Δ obtained after applying the rule for checking is of rank (G–1), where G is the number of endogenous variables in the equation system or the number of equations. This is known as the rank condition.

4) Arrange the equations as in equation (6.27). Prepare a table similar to Table 6.1. Strike out the relevant rows and columns. Compare the relationship between (K–M) and (G–1). This will give you the order condition. Prepare the determinants for the rank condition. The first and third equations are over-identified while the second equation is under-identified.

5) There are four equations. Hence, (G–1) = 3. Thus, at least 3 variables should be excluded from an equation for it to be identified. The first and fourth equations are exactly identified; the second equation is under-identified; and the third equation is over-identified.

Check Your Progress 3

1) Assuming that *B* is a non-singular matrix, the reduced form equations for the structural model indicated are: $\boldsymbol{YB + X\Gamma + E = 0}$, $\boldsymbol{YBB^{-1} + X\Gamma B^{-1} + EB^{-1} = 0}$, $\boldsymbol{Y + X\Gamma B^{-1} + EB^{-1} = 0}$, $\boldsymbol{Y = -X\Gamma B^{-1} - EB^{-1}}$ or $\boldsymbol{Y = X\pi + V}$ where *V* is the matrix of disturbance terms.

2) In the presence of a structural equation containing two or more endogenous variables, applying the least squares approach yields biased estimates. More specifically, to take an illustration, if we apply the OLS to the consumption

function [$\boldsymbol{C_t = \alpha + \beta Y_t + u_t}$], if we consider the *plim* for the estimate of $\hat{\boldsymbol{\beta}}$ obtained, even as n→∞ we find that $\left(\boldsymbol{plim\hat{\beta} - \beta}\right) \neq \boldsymbol{0}$. This violates the consistency property. This is known as the 'simultaneous equation bias' or the 'simultaneity bias'.

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 7 ESTIMATION OF SIMULTANEOUS EQUATIONS MODELS*

Structure

7.0 Objectives

7.1 Introduction

7.2 Limited Information Systems

7.2.1 Indirect Least Squares (ILS) Method

7.2.2 Instrumental Variable (IV) Method

7.2.3 Two-stage Least Squares (2SLS) Method

7.3 Full Information Systems

7.3.1 Full Information Maximum Likelihood (ML) Method

7.3.2 Three-stage Least Squares (3SLS) Method

7.4 Let Us Sum Up

7.5 Key Words

7.6 Suggested Books for Further Reading

7.7 Answers/Hints to Check Your Progress Exercises

7.0OBJECTIVES

After reading this unit, you will be able to:

- distinguish between the 'limited information method' and the 'full information method' of estimating simultaneous equations models (SEM);
- explain the indirect least squares (ILS) method of estimation of a simultaneous equations system;
- show that the ILS estimators are consistent;
- obtain the estimators of the structural parameters in a simultaneous equations system by the 'instrumental variable (IV)' method;
- derive the expression for the estimated coefficients of the structural parameters in a 'simultaneous equations system' by the 2SLS method;
- describe the 'full information maximum likelihood (FIML)' method; and

*Prof. Poulomi Roy, Jadavpur University, Kolkata

- write a note on LIML and 3SLS methods of estimation distinguishing them for the type of simultaneous equations systems they represent.

7.1 INTRODUCTION

The method of estimation of a simultaneous equations system can be classified into two categories: (i) single-equation method (also known as 'limited-information method') and (ii) systems method (also called 'full-information method'). In this unit, we will discuss some methods under both the categories. In the single-equation methods, we estimate each equation separately using the information contained in the equation under consideration. The information pertains to restrictions on the parameters of the equation. Let us consider the following example of a simultaneous equations model.

$$\beta_{11}y_1 + \beta_{13}y_3 + \gamma_{11}z_1 + \gamma_{13}z_3 + u_1 = 0 \quad \ldots (7.1)$$

$$\beta_{21}y_1 + \beta_{22}y_2 + \gamma_{21}z_1 + \gamma_{23}z_3 + u_2 = 0 \quad \ldots (7.2)$$

$$\beta_{32}y_2 + \beta_{33}y_3 + \gamma_{31}z_1 + \gamma_{32}z_2 + u_3 = 0 \quad \ldots (7.3)$$

When we estimate equation (7.1), we make use of the fact that the variable y_2 and z_2 are missing from the equation and z_1 and $z_3$are exogenous. We are not concerned about what variables are missing from the other equations. The restrictions on other equations are used only for identification and not for estimation. Remember that we can never estimate an under-identified equation.

A single equation method requires assumption about the variance and covariance of the disturbances only in the equation to be estimated. The systems method, on the other hand, considers information about variance and covariance of disturbances in all the equations as also the covariance between the error terms of different equations. This is the reason why the single equation methods are called 'limited information methods'. In the full information methods, we use information on the restrictions in all the equations.

The usual methods of obtaining consistent estimates of the parameters in simultaneous equations models are: (i) indirect least squares (ILS) method, (ii) instrumental variable (IV) method, (iii) two-stage least squares (2SLS) method, method. All these methods are single equation methods. The full information maximum likelihood (FIML) method and the 'three-stage least squares (3SLS)' method are the two 'systems method' of estimating the parameters of a simultaneous equations system. In FIML, we estimate all the equations together. In the present unit, you will study about both these methods.

A system method has some advantages over the single equation method which is applied to one identified equation at a time.

(i) A system method takes account of identifying restrictions on all the structural equations in the system while the single equation method

accommodates only the identifying restrictions for the equation to be estimated.

(ii) A single equation method requires assumptions about the variance and covariance of the disturbances only in the equation to be estimated whereas a system method considers information about variance and covariance of disturbances in all the equations and also the covariance between errors of different equations.

If all the equations of the system are not identified, then we can use any one of the single-equation methods to estimate the parameters of the identified equations. This is the advantage of the single equation method over the system method.

7.2 LIMITED INFORMATION SYSTEMS

In this section, we discuss five methods viz. (i) Indirect Least Squares Method, (ii) Instrumental Variables (IV) Method, (iii) Two-Stage Least Squares (2 SLS) Method, (iv) Generalised Least Squares (GLS) Method and (v) Limited Information Maximum Likelihood (LIML) Method. Before going for estimation, we should first check whether the equation is identified or not. If the equation is identified, then only we can proceed for estimation. It should be noted that if an equation is over-identified, then we cannot use the ILS method of estimation. However, we can use any one of the other methods of estimation in such cases. For the exactly-identified or the just-identified equation we can use the ILS method to estimate the structural form parameters.

7.2.1 Indirect Least Squares (ILS) Method

The method of Indirect Least Squares is applied to an exactly-identified structural equations system. The method contains the following three steps: (i) obtain the reduced form model, (ii) derive the estimators of the reduced form coefficients by applying the OLS method to the reduced form model and (iii) using the relation between the structural and reduced form parameters, estimate the structural form parameters[by using the reduced form parameters obtained in step (ii)]. Let us now consider the general form of the system of linear simultaneous equations involving G endogenous variables and K exogenous variables. For any time period t, the structural form of such a model is like:

$$Y_t B + X_t \Gamma + e_t = 0 \qquad \ldots (7.4)$$

where

$Y_t = [Y_{t1} Y_{t2} \ldots\ldots Y_{tG}]$,

$X_t = [X_{t1} X_{t2} \ldots\ldots X_{tK}]$, and

$e_t = [e_{t1} e_{t2} \ldots\ldots e_{tG}]$are three vectors, and B and G are matrices [B is a $G \times G$ matrix and Γ is a $K \times G$ matrix]. The i^{th} structural equation is

$$Y_i B + X_i \Gamma + e_i = 0 \quad \ldots (7.5)$$

The i^{th} observation can also be written as:

$$y_i = Y_i\beta_i + Y_i^{\cdot}\beta_i^{\cdot} + X_i\gamma_i + X_i^{\cdot}\gamma_i^{\cdot} + e_i$$

$$= Y_i\beta_i + X_i\gamma_i + e_i$$

$$= (Y_i X_i)\begin{pmatrix}\beta_i \\ \gamma_i\end{pmatrix} + e_i$$

$$= Z_i\delta_i + e_i \quad \ldots (7.6)$$

where $Z_i = (Y_i X_i)$ and $\delta_i = \begin{pmatrix}\beta_i \\ \gamma_i\end{pmatrix}$. Further, y_i is the $T\times 1$ vector of observations on the dependent endogenous variable in the equation, and Y_i is the $T\times(G_i^{\Delta}-1)$ matrix of observations on the other $(G_i^{\Delta}-1)$ current endogenous variables included in the equation. Likewise, X_i is the $T\times K_i^*$ matrix of observations on the K_i^* predetermined variables,and e_i is the $T\times 1$ vector of disturbances (included in the model). $Y_i^{\bullet}$ is the $\{T\times (G - G_i^{\Delta})\}$ matrix of observations on excluded endogenous variables and $X_i^{\bullet}$ is the $\{T\times (K - K_i^*)\}$ matrix of observations on excluded exogenous variables.To reiterate, there are $(G - G_i^{\Delta})$ number of endogenous variables excluded from the equations and$(K - K_i^*)$ number of predetermined variables excluded from the system of equations. We assume that the exclusion restriction ensures the exact identification of the equation. Now, incorporating these into the system of equations in (7.6) we get:

$$y_i = Y_i\beta_i + Y_i^{\cdot}\tilde{0} + X_i\gamma_i + X_i^{\cdot}\tilde{0} + e_i$$

$$\left(y_i \; Y_i \; Y_i^{\cdot} \; X_i \; X_i^{\cdot}\right)\begin{pmatrix}1 \\ -\beta_i \\ 0 \\ -\gamma_i \\ 0\end{pmatrix} = e_i$$

$$\left(y_i \; Y_i \; Y_i^{\cdot}\right)\begin{pmatrix}1 \\ -\beta_i \\ 0\end{pmatrix} + \left(X_i \; X_i^{\cdot}\right)\begin{pmatrix}-\gamma_i \\ 0\end{pmatrix} = e_i$$

$$YB_i^{'} + X\Gamma_i^{'} = e_i$$

where,

$$Y = \left(y_i \; Y_i \; Y_i^*\right), X = \left(X_i \; X_i^*\right) B_i^{'} = \begin{pmatrix}1 \\ -\beta_i \\ 0\end{pmatrix}$$

and

$$\Gamma_i^{'} = \begin{pmatrix}-\gamma_i \\ 0\end{pmatrix}$$

$(Y_i X_i)\begin{pmatrix} Bi' \\ \Gamma_i' \end{pmatrix} = e_i$

where $Y_i = \begin{pmatrix} y_i & Y_i \end{pmatrix}$

Thus, the relationship between the reduced form parameters and structural form parameters is:

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

$$\Rightarrow \pi' = -\Gamma'(B')^{-1}$$

$$\Rightarrow \pi' B' = \Gamma'$$

Hence, for the i^{th} structural equation we have:

$$\Pi' B_i' = -\Gamma_i' \qquad \ldots (7.7)$$

The estimated reduced form parameter matrix is:

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

Replacing the estimated value of the reduced form parameters in (7.7), we get:

$$(X'X)^{-1}X'YB_i' = -\Gamma_i'$$

$$(X'X)^{-1}X'Y\begin{pmatrix} 1 \\ -\beta_i \\ 0 \end{pmatrix} = -\begin{pmatrix} -\gamma_i \\ 0 \end{pmatrix}$$

$$(X'X)^{-1}X'(y_i - Y_i\beta_i) = \begin{pmatrix} \gamma_i \\ 0 \end{pmatrix}$$

Pre-multiplying both sides by $(X'X)$we have

$$X'(y_i - Y_i\beta_i) = (X'X)\begin{pmatrix} \gamma_i \\ 0 \end{pmatrix}$$

The matrix X may be partitioned as $X = \begin{bmatrix} X_i & X_i^* \end{bmatrix}$ so that $X' = \begin{bmatrix} X_i' \\ X_i^{*'} \end{bmatrix}$

Hence, $(X'X) = \begin{bmatrix} X_i' \\ X_i^{*'} \end{bmatrix}\begin{bmatrix} X_i & X_i^* \end{bmatrix} = \begin{bmatrix} X_i'X_i & X_i'X_i^* \\ X_i^{*'}X_i & X_i^{*'}X_i^* \end{bmatrix}$

$$\therefore X'(y_i - Y_i\beta_i) = \begin{bmatrix} X_i'X_i & X_i'X_i^* \\ X_i^{*'}X_i & X_i^{*'}X_i^* \end{bmatrix}\begin{pmatrix} \gamma_i \\ 0 \end{pmatrix}$$

$$\begin{bmatrix} X_i' \\ X_i^{*'} \end{bmatrix}(y_i - Y_i\beta_i) = \begin{bmatrix} X_i'X_i & X_i'X_i^* \\ X_i^{*'}X_i & X_i^{*'}X_i^* \end{bmatrix}\begin{pmatrix} \gamma_i \\ 0 \end{pmatrix}$$

$$\begin{bmatrix} X_i'y_i - X_i'Y_i\beta_i \\ X_i^{*'}y_i - X_i^{*'}Y_i\beta_i \end{bmatrix} = \begin{bmatrix} X_i'X_i\gamma_i \\ X_i^{*'}X_i\gamma_i \end{bmatrix}$$

$$\therefore X_i' y_i = X_i' Y_i \beta_i + X_i' X_i \gamma_i$$

$$X_i^{*'} y_i = X_i^{*'} Y_i \beta_i + X_i^{*'} X_i \gamma_i \quad \ldots (7.8)$$

The above system of equations contains K equations. The number of unknowns to be determined is the elements of β_i and γ_i. The number of elements in β_i is $G_i^{\Delta} - 1$ and that in γ_i is K_i^*. Therefore, the number of unknowns to be determined from the system is $G_i^{\Delta} - 1 + K_i^*$. For a unique solution to exist, the number of unknowns must be exactly the same as the number of equations. Hence, we must have:

$$K = G_i^{\Delta} - 1 + K_i^*$$

$$K - K_i^* = G_i^{\Delta} - 1$$

$$K_i^{**} = G_i^{\Delta} - 1$$

Thus, the number of excluded exogenous variables must be the same as the number of included endogenous variables less one. This is same as the 'order condition' for identification. This implies that for ILS method to be applicable to the structural equation, it must be exactly-identified. The ILS estimator of the parameter vector β_i and γ_i in the i^{th} structural equation (7.8) can be written as:

$$X_i' y_i = X_i' Y_i \beta_{iILS} + X_i' X_i \gamma_{iILS}$$

$$X_i^{*'} y_i = X_i^{*'} Y_i \beta_{iILS} + X_i^{*'} X_i \gamma_{iILS}$$

$$\begin{bmatrix} X_i' \\ X_i^{*'} \end{bmatrix} y_i = \begin{bmatrix} X_i' \\ X_i^{*'} \end{bmatrix} Y_i \beta_{iILS} + \begin{bmatrix} X_i' \\ X_i^{*'} \end{bmatrix} X_i \gamma_{iILS}$$

$$X' y_i = X' \begin{pmatrix} Y_i & X_i \end{pmatrix} \begin{pmatrix} \beta_{iILS} \\ \gamma_{iILS} \end{pmatrix}$$

$$X' y_i = X' Z_i \delta_i$$

We can solve the above equation for δ_i as

$$\delta_{iILS} = \left(X' Z_i\right)^{-1} X' y_i \quad \ldots (7.9)$$

Equation (7.9) is the ILS estimator of equation (7.6). We can now show that the ILS estimator of the parameters of an exactly-identified system of structural equations is consistent. To verify this, let us consider the i^{th} form of the structural equation with G number of endogenous and K number of exogenous variables given by:

$$y_i = Y_i \beta_i + X_i \gamma_i + e_i$$

$$y_i = \begin{bmatrix} Y_i & X_i \end{bmatrix} \begin{bmatrix} \beta_i \\ \gamma_i \end{bmatrix} + e_i$$

$$y_i = Z_i \delta_i + e_i$$

where $\delta_{iILS} = \left(X'Z_i\right)^{-1} X'\left(Z_i\delta_i + e_i\right)$

$= \left(X'Z_i\right)^{-1} X'Z_i\delta_i + \left(X'Z_i\right)^{-1} X'e_i$

$= \delta_i + \left(X'Z_i\right)^{-1} X'e_i$

Taking *plim* on both sides of the above equation we have:

$$p\lim \delta_{iILS} = \delta_i + p\lim\left(\frac{X'Z_i}{T}\right)^{-1} p\lim\left(\frac{X'e_i}{T}\right)$$

But, $p\lim\left(\frac{X'e_i}{T}\right) = 0$.

$$\therefore p\lim \delta_{iILS} = \delta_i + p\lim\left(\frac{X'Z_i}{T}\right)^{-1}.0 = \delta_i$$

Hence, the ILS estimator is a consistent estimator of δ_i.

7.2.2 Instrumental Variables (IV) Method

The IV method consists of: (i) choosing instruments for regressors correlated with disturbances, (ii) replacing these regressors by the chosen instruments, and (iii) proceed with the estimation of the structural coefficients. The instruments are so chosen that they are uncorrelated with the disturbances and they have a strong association with the regressors they replace. To obtain the estimates of the coefficients by this method, let us consider the model:

$$y_i = Y_i\beta_i + X_i\gamma_i + e_i$$

$$y_i = (Y_iX_i)\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix} + e_i$$

$$y_i = Z_i\delta_i + e_i \qquad \ldots (7.10)$$

where $Z_i = (Y_iX_i)$, $\delta_i = \begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}$, y_i is the $T\times1$ vector of observations on the dependent endogenous variable in the equation, Y_i is the $T\times$ (G_i^A-1) matrix of observations on the other (G_i^A-1) current endogenous variables included in the equation, X_i is the $T\times K_i^*$ matrix of observations on the K_i predetermined variables included in the model and e_i is the $T\times1$ vector of disturbances. Note that, we cannot apply OLS to the equation (7.10) since the regressors in Y_i, which is a sub matrix of Z_i, are correlated with e_i. If we apply OLS, the estimators of δ_i generated would be:

$$\delta_{iOLS} = \left(Z_i'Z_i\right)^{-1} Z_i'y_i = \left(Z_i'Z_i\right)^{-1} Z_i'\left(Z_i\delta_i + e_i\right) = \delta_i + \left(Z_i'Z_i\right)^{-1} Z_i'e_i$$

Then, $P\lim \delta_{iOLS} = \delta_i + P\lim\left(\frac{Z_i'Z_i}{T}\right)^{-1} P\lim\left(\frac{Z_i'e_i}{T}\right)$

For consistency of δ_{iOLS} we require that $P\lim\left(\frac{Z_i'e_i}{T}\right)=0$.

$$P\lim\left(\frac{Z_i'e_i}{T}\right)=P\lim\left(\frac{1}{T}\begin{pmatrix}Y_i'\\X_i'\end{pmatrix}e_i\right)=P\lim\begin{pmatrix}\frac{Y_i'e_i}{T}\\\frac{X_i'e_i}{T}\end{pmatrix}$$

Since X_i are exogenous, we have $P\lim\left(\frac{X_i'e_i}{T}\right)=0$. But $P\lim\left(\frac{Y_i'e_i}{T}\right)\neq 0$ because of the endogeneity of the variables in Y_i .Therefore, $P\lim\left(\frac{Z_i'e_i}{T}\right)\neq 0$ and the OLSE of δ_i is inconsistenti.e. $P\lim\delta_{iOLS}\neq\delta_i$. Therefore, an alternative method of estimation in cases like this is to consider the instrumental variables method. For this, suppose we are able to identify an IV matrix W which is of the same order as Z_i, i.e., W is of order T × ($G_i^{\Delta}-1+K_i^{*}$). Now, for W to qualify as an instrument to be used in place of Z_i, the following conditions must be satisfied.

(i) $p\lim\left(\frac{W'e_i}{T}\right)=0$ i.e. W is uncorrelated with e

(ii) $p\lim\left(\frac{W'Z_i}{T}\right)=\sum_{WZ_i}$ i.e. $\sum_{WZ_i}$ exist and is a non-singular matrix.

(iii) $p\lim\left(\frac{W'y_i}{T}\right)=\sum_{Wy_i}$ i.e. $\sum_{Wy_i}$ is a non-null vector.

(iv) $p\lim\left(\frac{W'y_i}{T}\right)=\sum_{WW}$ i.e. $\sum_{WW}$ is a finite positive definite matrix

Let us assume that the above conditions are satisfied by a W so that it qualifies as an instrument for Z_i. Pre-multiplying both sides of (7.10) by W' we get:

$$W'y_i=W'Z_i\delta_i+W'e_i$$

Dividing both sides of the above equation by T we get:

$$\frac{W'y_i}{T}=\frac{W'Z_i}{T}\delta_i+\frac{W'e_i}{T}$$

Taking *plim* on both sides of the above equation we get:

$$plim\,\frac{W'y_i}{T}=p\lim\left(\frac{W'Z_i}{T}\right)\delta_i+p\lim\left(\frac{W'e_i}{T}\right)$$

i.e., $\sum_{Wy_i}=\left(\sum_{WZ_i}\right)\delta_i+0$

i.e., $\delta_i=\sum_{WZ_i}^{-1}\sum_{Wy_i}$

$\sum_{WZ_i}$ *and* $\sum_{Wy_i}$ include the population cross moments of the variables appearing in the two matrices W and Z_i. If we replace the population cross moments by the sample cross moments, we get the IV estimator of δ_i as:

$$\delta_{iIV} = \left(\frac{W'Z_i}{T}\right)^{-1}\left(\frac{W'y_i}{T}\right)$$

$$\text{or } \delta_{iIV} = \left(W'Z_i\right)^{-1}\left(W'y_i\right) \quad \ldots (7.11)$$

Equation (7.11) gives the IV estimator of δ_i with W as the instrument for Z_i.

7.2.3 Two-Stage Least Squares (2SLS) Method

The 2 SLS method is another single equation method which can be applied to an identified simultaneous structural equation model. The approach involves application of OLS in two stages as follows.

Stage I: In the first stage, we derive the reduced form of the model and apply the OLS method to obtain the estimates of reduced form parameters. This stage helps in eliminating the randomness in the explanatory and endogenous variables in the model.

Stage II: In the second stage, we reformulate the structural equations by replacing the matrix of 'explanatory endogenous variables' with their estimated coefficients (obtained in Stage I above). We then apply the OLS method to the reformulated equations to obtain the 2SLS estimator of the coefficients.

To illustrate the method of 2SLS, let us consider the i^{th} structural equation of a simultaneous equations model involving G number of endogenous and K number of exogenous variables as:

$$y_i = Y_i\beta_i + X_i\gamma_i + e_i \quad \ldots (7.12)$$

where y_i is the $T\times1$ vector of observations on the dependent endogenous variable in the equation. We assume that the exclusion restrictions imposed on the coefficients of the equation ensures the over-identification of the relationship. Note that, in its absence, the method of OLS is inapplicable due to the presence of Y_i as a regressor. This is also the reason why we replace Y_i by some non-stochastic matrix. Now, the reduced form of the model for a T sample observation is:

$$Y = X\Pi' + V \quad \ldots (7.13)$$

where, Y is a $T\times G$ matrix of observations on the endogenous variables, X is the $T\times K$ matrix of observations on the K predetermined variables (both Y and X considered over the entire sample period), Π' is the $(K\times G)$ matrix of reduced form coefficients and V is the matrix of reduced form disturbances of order $(T\times G)$.The OLSE of $\hat{\Pi}' = \left(X'X\right)^{-1}X'Y$. The estimated value of Y is $\hat{Y} = X\hat{\Pi}'$. Since $\hat{Y} = X\left(X'X\right)^{-1}X'Y$, using this we can get estimate of Y_i as

$\hat{Y}_i = X(X'X)^{-1}X'Y_i$. This completes the 1st stage of 2SLS procedure. In the second stage, we replace Y_i by $\hat{Y}_i$ in equation (7.12) to obtain:

$$y_i = \hat{Y}_i\beta_i + X_i\gamma_i + e_i \quad \ldots (7.14)$$

Since $\hat{Y}_i$ is uncorrelated with e_i, we can replace $\hat{Y}_i$ in place of Y_i and apply OLS to the equation (7.14). This gives us the estimated value of the parameters β_i and γ_i. These are known as the 2SLS estimators. To obtain these expressions, as before, we write $y_i = Y_i\beta_i + X_i\gamma_i + e_i$ as:

$$y_i = (Y_iX_i)\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix} + e_i$$

$$y_i = Z_i\delta_i + e_i$$

where $Z_i = (Y_iX_i)$, $\delta_i = \begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}$

Now, the equation (7.14) can be re-written as:

$$y_i = \begin{bmatrix}\hat{Y}_i & X_i\end{bmatrix}\begin{bmatrix}\beta_i\\ \gamma_i\end{bmatrix} + e_i$$

or $y_i = \hat{Z}_i\delta_i + e_i \quad \ldots (7.15)$

Applying OLS to equation (7.15), we get the 2SLS estimates of δ_i as:

$$\delta_{i2SLS} = \left(\hat{Z}_i'\hat{Z}_i\right)^{-1}\hat{Z}_i'y_i$$

Since $\hat{Z}_i = \left(\hat{Y}_i \quad X_i\right)$, $\hat{Z}_i' = \begin{pmatrix}\hat{Y}_i'\\ X_i'\end{pmatrix}$

$$\therefore \hat{Z}_i'\hat{Z}_i = \begin{pmatrix}\hat{Y}_i'\hat{Y}_i & \hat{Y}_i'X_i\\ X_i'\hat{Y}_i & X_i'X_i\end{pmatrix} \text{ and } \hat{Z}_i'y_i = \begin{pmatrix}\hat{Y}_i'\\ X_i'\end{pmatrix}y_i = \begin{pmatrix}\hat{Y}_i'y_i\\ X_i'y_i\end{pmatrix}$$

Therefore, $\delta_{i2SLS} = \begin{bmatrix}\beta_{i2SLS}\\ \gamma_{i2SLS}\end{bmatrix} = \begin{pmatrix}\hat{Y}_i'\hat{Y}_i & \hat{Y}_i'X_i\\ X_i'\hat{Y}_i & X_i'X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i'y_i\\ X_i'y_i\end{pmatrix} \quad \ldots (7.16)$

In this expression for δ_{i2SLS}, $\hat{Y}_i$ is the estimated value of Y_i obtained from the reduced form of variables in Y_i. To express δ_{i2SLS} only in terms of observed values, we recall that the reduced form of all endogenous variables in the model is:

$$Y = X\Pi' + V$$

Applying OLS to this equation we get:

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

The reduced form equation for Y_i is

$$Y_i = X\Pi_i' + V_i$$

From this reduced form equation, the estimated value of Y_i can be obtained as:

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

Denoting the residuals by R_i we have:

$$Y_i = \hat{Y}_i + R_i$$

Now,

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

$$\hat{Y}_i'X_i = (Y_i - R_i)'X_i = Y'_iX_i - R_i'X_i$$

Now,

$$R_i'X = (Y_i' - \hat{Y}_i')X = Y_i'X - \hat{Y}_i'X = Y_i'X - \{Y_i'X(X'X)^{-1}X'X\} = Y_i'X - Y_i'X = 0$$

Since X_i is only a sub-matrix of X, we have $R_i'X_i = 0$. We need to arrive at δ_{i2SLS} by substituting for $\hat{Y}_i'y_i$ and $X_i'y_i$ in (7.16).

For this, let us note that $\hat{Y}_i'X_i = Y'_iX_i$ (since $R_i'X_i = 0$) and hence $X_i'\hat{Y}_i = X_i'Y_i$.

Further,

$\hat{Y}_i'y_i = Y_i'X(X'X)^{-1}X'y_i$ (from 7.17).

Therefore,

$$\delta_{i2SLS} = \begin{pmatrix} Y_i'X(X'X)^{-1}X'Y_i & Y_i'X_i \\ X_i'Y_i & X_i'X_i \end{pmatrix}^{-1} \begin{pmatrix} Y_i'X(X'X)^{-1}X'y_i \\ X_i'y_i \end{pmatrix} \quad \ldots (7.18)$$

Equation (7.18) expresses the 2SLS estimates of β_i and γ_i in terms of the observed values only.

Check Your Progress 1[answer within the space given in about 50-100 words]

1) Distinguish between 'single equation method' and 'systems method' of solving a simultaneous equations model (SEM).

.....
.....
.....
.....

2) Specify the two steps involved in the ILS method.

.....
.....
.....
.....

3) State the condition required to be met for the application of the ILS method.

.....
.....
.....
.....
.....

4) Specify the three steps involved in the application of the IV method of estimating the parameters of a SEM.

.....
.....
.....
.....
.....

5) What are the two stages involved in the application of the 2 SLS method?

……
……
……
……

7.3 FULL INFORMATION MODELS

In this section, we discuss two methods viz. (i) Full Information Maximum Likelihood (FIML) Method, and (ii) Three-stage Least Squares (3 SLS) Method.

7.3.1 Full Information Maximum Likelihood (FIML) Method

This is a systems method of estimation in which we obtain the estimators of parameters in the entire system of equations at a time. This is done by maximizing a likelihood function framed by incorporating all identifying restrictions in the structural equations. Let us consider a SEM involving G number of endogenous and K number of exogenous variables. Then the i^{th} structural equation may be written as:

$$y_i = Z_i\delta_i + e_i \qquad \ldots (7.19)$$

where $Z_i = \left(Y_i \;\; X_i\right)$, $\delta_i = \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix}$

For i=1,2,.....,G the above equation may be written as:

$$\begin{aligned} y_1 &= Z_1\delta_1 + e_1 \\ y_2 &= Z_2\delta_2 + e_2 \\ &\vdots \\ y_G &= Z_G\delta_G + e_G \end{aligned}$$

$$\text{i.e.} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_G \end{bmatrix} = \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ 0 & 0 & \cdots & Z_G \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_G \end{bmatrix}$$

$$\text{i.e. } y = Z\delta + e \qquad \ldots (7.20)$$

where $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_G \end{bmatrix}_{TGx1}$, $Z = \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ 0 & 0 & \cdots & Z_G \end{bmatrix}_{TGx\sum_i G_i^{\Delta} - 1 + K_i^*}$, $\delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix}_{\sum_i (G_i^{\Delta} - 1 + K_i^*)x1}$,

$$e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_G \end{bmatrix}_{TGx1}$$

Let us assume that the disturbances *e*'s are normally distributed with common mean=0 and common variance-covariance matrix $\sum$, i.e., $E(e_t) = 0$ and $E(e_t e_t') = \sum$ .When we have *T* number of observations, the vector of all *TG* number of disturbances in the model follow a multivariate normal distribution with mean=0 and variance-covariance matrix $\sum \otimes I_T$, i.e., $e \sim N(0, \sum \otimes I_T)$.On the basis of this assumption, we can find out the joint pdf of *e* as:

$$g(e) = \frac{1}{\left(\sqrt{2\pi}\right)^{GT}} \frac{1}{\left|\sum \otimes I_T\right|} e^{\frac{-1}{2} e'(\sum \otimes I_T)^{-1} e} \qquad \ldots (7.21)$$

To obtain the joint pdf of *y*, we use the Jacobian transformation. The structural form of the model for any time period *t* can be written as:

$$BY_t + \Gamma X_t = e_t$$

The joint probability density function of elements in Y_t given X_t is:

$$f(Y \mid X) = g(e)\left|B\right|^T = \frac{1}{\left(\sqrt{2\pi}\right)^{GT}} \frac{1}{\left(\sum \otimes I_T\right)} Exp\left[-\frac{1}{2} e' \left(\sum \otimes I_T\right)^{-1} e\right] \left|B\right|^T .$$

The likelihood function (*L*) [i.e. the logarithmic function] of the above probability density function is given by:

$$\log\left[f(Y \mid X)\right] = \log\left[\frac{1}{\left(\sqrt{2\pi}\right)^{GT}} \frac{1}{\left(\sum \otimes I_T\right)}\right] - \frac{1}{2} e' \left(\sum \otimes I_T\right)^{-1} e + T \log\left|B\right|$$

$$= \log\left[\frac{1}{\left(\sqrt{2\pi}\right)^{GT}} \frac{1}{\left(\sum \otimes I_T\right)}\right] - \frac{1}{2} (y - Z\delta)' \left(\sum \otimes I_T\right)^{-1} (y - Z\delta) + T \log\left|B\right|$$

$$\sum \otimes I_T = \begin{bmatrix} \sigma_{11}I_T & \sigma_{12}I_T & \cdots & \sigma_{1G}I_T \\ \sigma_{21}I_T & \sigma_{22}I_T & \cdots & \sigma_{2G}I_T \\ \vdots & \vdots & \vdots\vdots & \vdots \\ \sigma_{G1}I_T & \sigma_{G2}I_T & \cdots & \sigma_{GG}I_T \end{bmatrix} \quad \ldots (7.22)$$

where

$$\sigma_{11}I_T = \begin{pmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{11} & \cdots & 0 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{11} \end{pmatrix} = \sigma_{11}{}^{T}.$$

$$\therefore \sum \otimes I_T = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1G} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2G} \\ \vdots & \vdots & \vdots\vdots & \vdots \\ \sigma_{G1} & \sigma_{G2} & \cdots & \sigma_{GG} \end{bmatrix}^{T} \sum{}^{T}.$$

We will maximize the likelihood function subject to restrictions prevailing in the system. Then, by solving the condition for the maximum of log L, we obtain the FIML estimators of δ and Σ.

You should note that the distribution of asymptotic variance-covariance matrix Σ as obtained under FIML is exactly equal to the asymptotic variance-covariance matrix as obtained under 3SLS.

The asymptotic distribution of δ as obtained under FIML is exactly equal to the asymptotic distribution of δ as obtained under 3SLS if and only if no constraints are available on Σ. To obtain the solution values of δ and Σ from FIML we require the application of the numerical optimisation technique.

7.3.2 Three-Stage Least Squares (3SLS) Method

One of the system methods of estimation of the coefficients in a simultaneous equations system is the method of three-stage least squares (3SLS). We apply this method to all the structural equations of the SEM given that they are identified.

Let us consider a system of simultaneous equations involving G number of endogenous variables and K number of exogenous variables. The i[th] structural equation is written in the form

$$y_i = Y_i\beta_i + X_i\gamma_i + e_i$$

$$= (Y_iX_i)\begin{pmatrix}\beta_i \\ \gamma_i\end{pmatrix} + e_i$$

$$= Z_i\delta_i + e_i \quad \ldots (7.23)$$

where $Z_i = (Y_i\, Xi)$ and $\delta_i = \begin{pmatrix}\beta_i \\ \gamma_i\end{pmatrix}$.

You should note that y_i is the $T \times 1$ vector of observations on the dependent endogenous variable in the equation and Y_i is the $T\times(G_i^{\Delta}-1)$ matrix of observations on the other $(G_i^{\Delta}-1)$ current endogenous variables included in the equation. Likewise, X_i is the $T \times K_i^*$ matrix of observations on the K_i^* predetermined variables and e_i is the $T\times1$ vector of disturbances (included in the model).

Now pre-multiplying both sides of (7.23) by X' we get:

$$X'y_i = X'Z_i\delta_i + X'e_i \qquad \ldots (7.24)$$

Thus, for $i=1,2,\ldots,G$ we have

$$X'y_1 = X'Z_1\delta_1 + X'e_1$$

$$X'y_2 = X'Z_2\delta_2 + X'e_2$$

. . .

. . .

. . .

$$X'y_G = X'Z_G\delta_G + X'e_G$$

Combining these equations in matrix notations we have

$$\begin{bmatrix} X'y_1 \\ X'y_2 \\ \vdots \\ X'y_G \end{bmatrix} = \begin{bmatrix} X'Z_1 & 0 & \cdots & 0 \\ \cdots & X'Z_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X'Z_G \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix} + \begin{bmatrix} X'e_1 \\ X'e_2 \\ \vdots \\ X'e_G \end{bmatrix}$$

Or,

$$X'y = X'Z\delta + X'e \qquad \ldots (7.25)$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_G \end{bmatrix}$$

is the vector of $TG\times1$ observations on the dependent endogenous variables, and

y_i is the $T\times1$ vector of observations on the dependent variable of the i[th] structural equation, $i=1,2,\ldots,G$.

$$Z = \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Z_G \end{bmatrix}_{TGx\sum(G_i^{\Delta}-1+K_i^*)}$$

Now, Z_1 is $_{T\,x}\,G_1^{\Delta}-1+K_1^*$,

Z_2 is $_{T\,x}\,G_2^{\Delta}-1+K_2^*$,

.

.

Z_G is $_{T\,x}\,G_G^{\Delta}-1+K_G^*$.

Thus number of rows in Z is TG and number of columns in Z is $\sum(G_i^{\Delta}-1+K_i^*)$.

$$e=\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_G \end{bmatrix} \text{ is } T\times 1$$

and $\delta=\begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix}$ is $\sum(G_i^{\Delta}-1+K_i^*)x1$.

The equation (7.25) may be written in the following form:

$$I\otimes X'y=I\otimes X'Z\delta+I\otimes X'e. \qquad \ldots (7.26)$$

The 3SLS estimator of the parameter vector δ is obtained in the following three stages:

Stage-1:

Regress the matrix of explanatory endogenous variables in each of the structural equations on all predetermined variables of the system. The reduced form of the i[th] structural equation of Y_i may be written as

$$Y_i=X\Pi_i'+v_i$$

The OLSE of $\hat{\Pi}_i'$ is

$$\hat{\Pi}_i'=\left(X'X\right)^{-1}X'Y_i.$$

For each of the structural equation we estimate the Π_i' matrices. This completes the first stage.

Stage-2:

Using the estimator generated in stage –I we obtain the estimated value of Y_i.

Now, $\hat{Y}_i=X\hat{\Pi}_i'=X(X'X)^{-1}X'Y_i$

In the second stage, we replace Y_i by $\hat{Y}_i$ in equation (7.23) to obtain

$$y_i=\hat{Y}_i\beta_i+X_i\gamma_i+e_i$$

Or, $y_i = \begin{bmatrix} \hat{Y}_i & X_i \end{bmatrix} \begin{bmatrix} \beta_i \\ \gamma_i \end{bmatrix} + e_i$

Or,

$$y_i = \hat{Z}_i \delta_i + e_i \quad \ldots (7.27)$$

Applying OLS to equation (7.27), we derive the 2SLS estimator of δ_i as:

$$\delta_{i2SLS} = \left(\hat{Z}_i' \hat{Z}_i\right)^{-1} \hat{Z}_i' y_i . \quad \ldots (7.28)$$

The 2SLS estimators are then applied to find the residuals of the regression model. These residuals are then used for estimation of variance and covariance of the errors in the matrix. Let the estimated sigma matrix be:

$$\hat{\Sigma} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \cdots & \hat{\sigma}_{1G} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} & \cdots & \hat{\sigma}_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{G1} & \hat{\sigma}_{G2} & \cdots & \hat{\sigma}_{GG} \end{bmatrix}$$

where $\hat{\sigma}_{ij} = \dfrac{\left(y_i - Z_i \delta_{i2SLS}\right)'\left(y_i - Z_i \delta_{i2SLS}\right)}{\tau_{ij}}$, and τ_{ij} is the degrees of freedom.

Stage-3:

In the third stage, we use the estimated sigma matrix for application of generalized least squares (GLS) to equation (7.26). The matrix of variance and covariance of the disturbances (error terms) in equation (7.26) is given by

$$E\left[\left(I \otimes X'\right) ee' \left(I \otimes X\right)\right]$$
$$= \sum \otimes E\left(X'X\right) = \sum \otimes \left(X'X\right)$$

The structure of the variance-covariance matrix suggests that an appropriate method of estimation will be the method of generalized least squares (GLS). To apply GLS to equation (7.26), we replace the unknown matrix $\sum$ by the estimated matrix $\hat{\sum}$ from stage 2. This gives us δ_{3SLS}.

Applying GLS to (7.26) we get

$$\delta^* = \left[Z'\left(I \otimes X'\right)'\left[\sum^{-1} \otimes \left(X'X\right)^{-1}\right]\left(I \otimes X'\right)Z\right]^{-1} Z'\left(I \otimes X'\right)'\left[\sum^{-1} \otimes \left(X'X\right)^{-1}\right]\left(I \otimes X'\right)y \quad \ldots (7.29)$$

Now replacing the $\sum$ by $\hat{\sum}$ as obtained under stage-2, we obtain the 3SLS estimators as:

$$\delta_{3SLS} = \left[Z'\left(I \otimes X'\right)'\left[\hat{\sum}^{-1} \otimes \left(X'X\right)^{-1}\right]\left(I \otimes X'\right)Z\right]^{-1} Z'\left(I \otimes X'\right)'\left[\hat{\sum}^{-1} \otimes \left(X'X\right)^{-1}\right]\left(I \otimes X'\right)y \quad \ldots (7.30)$$

Check Your Progress 2 [answer within the space given in about 50-100 words]

1) Specify the underlying principle behind the FIML method.

.....
.....
.....
.....
.....
.....

2) State the i^{th} structural equation (with G number of endogenous variables and K number of exogenous variables).

.....
.....
.....
.....
.....
.....

3) State the i^{th} structural equation (with G number of endogenous variables and K number of exogenous variables).

.....
.....
.....
.....
.....
.....

7.4 LET US SUM UP

An important feature of SEMs is that, by fully specifying the system, it is clear which variables are assumed to be exogenous, and which ones appear in each equation. Given such a full system, we can determine which equations can be identified (and hence their structural parameters can be estimated). The method of instrumental variables (IV) can consistently estimate the parameters in a linear model when one or more explanatory variables are endogenous. An instrumental variable must have two properties: (i) it must be exogenous i.e. uncorrelated with the error term of the structural equation and (ii) it must be partially correlated with the endogenous explanatory variable. Finding a variable with these two properties is usually challenging. The method of two-stage least squares, which allows for using more instrumental variables than the number of explanatory variables that we have, can be used in such cases. When used properly, it enables us to estimate the effect on the dependent variables in the presence of endogenous explanatory variables. This holds valid for all cases i.e. cross-

sectional, time series, and panel data sets. But when the instruments used are poor (i.e., they are correlated with the error term), or weakly correlated with the endogenous explanatory variables, or both, then the 2SLS method can be worse than the OLS.

7.5 KEYWORDS

Indirect Least Squares Method In this method, we obtain the estimates of the reduced form coefficients by applying the OLS method and then indirectly obtain the estimated values of the structural parameters. The latter is done by using the relationship between the parameters of the reduced form equations and the original structural equations.

Instrumental Variable Method The method uses some appropriate exogenous variable as an instrument (or instruments) to reduce the dependence between the error term and the explanatory variables.

2SLS Method This is generally applicable to over-identified models or systems. In the first stage, we derive the reduced form of the model and apply the OLS method to obtain the estimates of reduced form parameters. In the second stage, we reformulate the structural equations by replacing the matrix of 'explanatory endogenous variables' with the estimates of their coefficients (obtained in the first stage).

LIML (Limited Information Maximum Likelihood) Method This is also known as the Least Variance Ratio (LVR) method. This is the first of the single-equation methods suggested for simultaneous equations models.

FIML (Full Information Maximum Likelihood Method) This is a system method of estimation in which we obtain the estimators of parameters in the entire system of equations at a time. This is done by maximizing a likelihood function which is framed by incorporating all identifying restrictions in the structural equations.

7.6 SOME USEFUL BOOKS

Cameron, A. Collin and P. K. Trevedi (2005). *Micro Econometric Methods and Applications*, Cambridge University Press.

Green, W. H. (2017). *Econometric Analysis*, Eighth Edition, Pearson.

Wooldridge, J.W. *Econometric Analysis of Cross section and Panel Data*, MIT Press, 2010.

7.7 ANSWERS/ HINTS TO CHECK YOUR PROGRESS EXERCISES

Check Your Progress 1

1) In the 'single equation method', we deal with one equation at a time. This means we concern ourselves with the variables that are present and absent in that particular equation. In other words, we work with the information limited to the equation under focus. For this reason, single equation methods are also called as the 'limited information methods'. In the systems method, on the other hand, we take into account information about the variance and covariance of disturbances in all the equations (as also the covariance between the error terms of different equations). For this reason, the systems method is also called as the 'full information method'. In the systems method, we obtain the estimators of parameters in the entire system at a time.

2) The method has the following three steps: (i) obtain the reduced form model, (ii) derive the estimates of the reduced form coefficients by applying the OLS method to the reduced form model and (iii) using the relation between the structural and reduced form parameters, estimate the structural form parameters [with the help of the reduced form parameters obtained in step (ii)].

3) The structural equation must be exactly-identified.

4) (i) Choice of instruments for regressors correlated with disturbances, (ii) replacing the regressors by their chosen instruments, and (iii) proceed with the estimation of the structural coefficients.

5) The two stages are: (i) derive the reduced form of the model and apply the OLS method to obtain the estimates of reduced form parameters and (ii) reformulate the structural equations by replacing the estimated coefficients (obtained in Stage I above). Then apply the OLS method to the reformulated equations to obtain the 2SLS estimator of the coefficients.

Check Your Progress 2

1) The principle is to maximise a likelihood function framed by incorporating all the identified restrictions in the structural equations.

2) $y_i = Z_i\delta_i + e_i$ where $Z_i = \begin{pmatrix} Y_i & X_i \end{pmatrix}$, $\delta_i = \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix}$

3) $$\log\left[f\left(Y \mid X\right)\right] = \log\left[\frac{1}{\left(\sqrt{2\pi}\right)^{GT}}\frac{1}{\left(\sum \otimes I_T\right)}\right] - \frac{1}{2}e'\left(\sum \otimes I_T\right)^{-1}e + T\log|B|$$

$$= \log\left[\frac{1}{\left(\sqrt{2\pi}\right)^{GT}}\frac{1}{\left(\sum \otimes I_T\right)}\right] - \frac{1}{2}\left(y - Z\delta\right)'\left(\sum \otimes I_T\right)^{-1}\left(y - Z\delta\right) + T\log|B|$$