DEGREE OF OVER-IDENTIFICATION AND ITS EFFECTS ON THE PERFORMANCE OF ESTIMATORS

A DISSERTATION

SUBMITTED

IN

PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF PHILOSOPHY IN ECONOMICS

SUPERVISED BY

DR. S. K. MISHRA

PROF., DEPT. OF ECONOMICS

N. E. H. U. SHILLONG

SUBMITTED BY

MADHUCHHANDA DAS GUPTA

TO

THE DEPARTMENT OF ECONOMICS

NORTH EASTERN HILL UNIVERSITY

SHILLONG

1992.

PHONE :
GRAMS · NEHU

North-Eastern Hill University

Mayurbhanj Complex, Nongthymmai, Shillong-793014

Dr. S.K. Mishra.
Professor
Department of Economics

CERTIFICATE

This is to certify that Ms. **MADHUCHHANDA DAS GUPTA** has worked under my supervision for her M. Phil Dissertation entitled "Degree of Over Identification, And Its Effects on The Performance of Estimators" and no part of it has been submitted elsewhere for the award of any degree. This dissertation, in my opinion, is worthy of an award of the degree of Master of Philosophy in Economics.

Shillong,
The 30th October 1992

(Dr. S.K. MISHRA)
Supervisor

Dr. S. K. Mishra
PROFESSOR
DEPT OF ECONOMICS
NEHU, SHILLONG

ACKNOWLEDGEMENTS

It is my pleasure to record my heartfelt gratitude and indebtedness to my Supervisor, Dr. S.K. Mishra, Professor, Department of Economics, N.E.H.U. for his able guidance, valuable advice, inspiration and constant ungrudging help in every stage of my research work. I shall be failing in my duty if I do not mention here that it is Professor Mishra who has instilled in me the spirit of dedication and sincerity of purpose required for conducting the research.

I also express my gratitude to my teacher, Professor K. Bez, Head of the Department of Economics, N.E.H.U. and other teachers of the department for their advice and encouragement.

My long-standing thanks are due to Mrs. B. Sharma and Mr. G. Pathaw of the Department of Economics and my friend Ms. Haimanti Ghose for their sustained encouragement.

I owe a special debt of gratitude to my parents who have not only encouraged me but also sacrificed a lot to give me comfort and solace during the course of my work.

Finally, I would like to thank the Institute of Computer Education, Shillong for the careful typing of the manuscript.

Dated: 30th October, 1992
Shillong.

Madhuchhanda Das Gupta
(Madhuchhanda Das Gupta)

CONTENTS

CHAPTER - I

INTRODUCTION

INTRODUCTION

1.1 THE CONCEPT OF SIMULTANEOUS EQUATION MODELS

The greater part of empirical work in econometrics deals with structural estimation, ie. the estimation of the parameters of the behavioural and technological relationships whose nature is defined by the mathematical economic theory. Such work may vary from the estimation of the cost curve of a particular firm to large and complex models of an entire economy.

To start with an econometric enquiry it is necessary to first design an econometric model which abstractly represents reality by bringing out what is relevant to a particular question neglecting all other aspects. These models may be of different sizes. Some models may involve single relationship among economic variables and therefore can be summed up in a single equation. While others may consist of an impressive array of relations. In the latter type, the economic variables are determined by a complete system of equations. By a complete system is meant one in which there are as many equations as endogenous variables, ie, variables whose formation is to be 'explained' by the equations. The equations are usually of four types : equations of economic behaviour, institutional rules, technological laws of transformation and identities. The term 'structural equation' will be used to comprise all four types of equations.

Systems of structural equation may be composed entirely on the basis of economic theory or on the dual basis of theory combined with systematically collected statistical data for the relevant variable for a given period. These dynamic models serve as an indespensible aid to forecasting and an invaluable guide to policy-making whether for a firm or a governmental agency.

Since Economic models frequently involve a set of relationships designed to explain the behaviour of certain variables, the use of simultaneous equation model will be more meaningful and logical. It is clear from Marshall's classical analysis of determination of price and quantity by supply and demand. The demand function can be represented as

$$Q_D = \alpha_1 + \beta_1 + U_1 \qquad \ldots\ldots \text{(a)}$$

This describes the behaviour of buyers in determining the quantity they are willing to buy, Q_D, as a function of price P. Likewise,

$$Q_S = \alpha_2 + \beta_2 P + \delta_2 W + u_2 \qquad \ldots\ldots \text{(b)}$$

as the supply function.

In the equations (a) and (b) Q_D = quantity demanded; Q_S = quantity supplied: W = weather conditions and U_1 and u_2 are random disturbances with mean zero.

A complete model of the market should presumably include an analysis of the process whereby buyers and sellers arrive at the price that equates both quantities. The Cob-web theory gives the best description by assuming that either functions operate alternatingly as if buyers and sellers announce in turn what price they set or what quantities they are willing to trade. The result is an iterative process which under certain conditions converge to equilibrium values of P and Q at which the market is cleared. These values are determined by the conditions that they satisfy both (a) and (b) or by ..

$$Q_D = Q_S \qquad \ldots\ldots \text{(c)}$$

We thus see that price and quantity are simultaneously determined by demand and supply or as Marshall puts it - that "we might as reasonably dispute whether it is the upper or the lower blade of a pair of scissors that cuts a piece of paper, as whether value is governed by utility or cost of production."

Marshall's model of demand and supply generating market data is the simplest example of simultaneous equations in economics.

Let us take another example, namely a simple income determination model which consists solely of a consumption function and the national income identity.

$$C_t = \alpha + \beta Y_t + U_t \qquad \text{.....(d)}$$

$$Y_t = C_t + I_t \qquad \text{.....(e)}$$

where C = consumption, Y=income, I=investment U=error term and t= time.

In this model investment I is autonomous, ie. independent of income Y and consumption C. Since C and Y are determined within the model, they are endogenous variables and I is an exogenous variable.

We will make the following assumptions ...

$$E(U_t) = 0 \quad \text{for all } t$$

$$E(U_t^2) = \sigma_u^2 \quad \text{for all } t$$

$$E(U_s U_t) = 0 \quad \text{for all } s, t;\ s \neq t$$

and I_t and U_t are independent.

The explicit dependence of C and Y on I and U is now shown by solving the system into ...

$$C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I_t + \frac{1}{1-\beta} U_t \quad \text{.....(f)}$$

$$Y_t = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I_t + \frac{1}{1-\beta} U_t \quad \text{.....(g)}$$

Given a sample of joint observations on C, Y and I, our interest is in estimating the parameters of the consumption function (d). Now, in equation (d), the regressor and the disturbance are not uncorrelated, not even contemporaneously. The convariance of Y and U may be found by multiplying equation (g) throughout by U_t and taking expectation.

$$E(Y_t U_t) = \frac{\alpha}{1-\beta} E(U_t) + \frac{1}{1-\beta} I_t E(U_t) + \frac{1}{1-\beta} E(U_t^2) \text{.....(h)}$$

$$= \frac{\sigma_u^2}{1-\beta} \neq 0 \quad \text{(using the assumptions).}$$

Thus, if least square method is applied directly over the consumption function, the estimates of α and β will not be consistent.

Let us consider this explicitly. The classical least-square estimator of β in (d) is...

$$b = \frac{\Sigma C (Y - \bar{Y})}{\Sigma (Y - \bar{Y})^2} = \frac{\Sigma (\alpha + \beta Y + U_t)(Y - \bar{Y})}{\Sigma (Y - \bar{Y})^2}$$

$$= \beta + \frac{\Sigma U_t (Y - \bar{Y})}{\Sigma (Y - \bar{Y})^2}$$

$$= \beta + \frac{\Sigma U_t (Y - \bar{Y}) / T}{\Sigma (Y - \bar{Y})^2 / T}$$

Now $\Sigma U_t (Y - \bar{Y}) / T$ is a sample covariance so that under general conditions ....

$$\text{plim} \frac{\Sigma U_t (Y - \bar{Y})}{T} = E\, U_t\, y_t = E\, U_t\, y_t = (1 - \beta)^{-1} \sigma^2 ;$$

similarly,

$\Sigma (Y - \bar{Y})^2 / T$ is a sample variance so that under general conditions ...

$$\text{Plim} \frac{\Sigma (Y - \bar{Y})^2}{T} = E(Y - EY)^2 = E(y - Ey)^2 = \sigma_{YY} \text{ (say)}$$

Then,

$$\text{Plim } b = \beta + \frac{(1 - \beta)^{-1} \sigma^2}{\sigma_{YY}}$$

This shows that the least-squares estimator will not be consistent. The direction of the asymptotic bias is clear if we employ the economic information that the marginal propensity to consume lies between zero and one : with 0 β 1, plim b β.

we should therefore go for alternative methods of estimation for such cases.

The above model is a simple one and can be extended further. Suppose investment is influenced by interest rate, then investment becomes endogenous. Now, the model will consist of three endogenous variables. For generalisation, we should specify as many relations as there are endogenous variables in the model. The classification of endogenous and pre-determined variables which includes exogenous and lagged endogenous variables, therefore, depends upon the assumptions underlying the construction of the model. The values of the exogenous variables are completely determined outside the system under consideration, whereas the values of the lagged endogenous variables are represented by the past values of the endogenous variables of the model.

1.2 <u>THE GENERAL SIMULTANEOUS EQUATION MODEL</u> :

Before proceeding further let us develop the general simultaneous equation model. For this let us take a linear simultaneous equation model containing G structural relations. The model may be written as :

$$a_{11}Y_{1t} + a_{12}Y_{2t} + \dots + a_{1G}Y_{Gt} + b_{11}X_{1t} + b_{12}X_{2t} + \dots + b_{1k}X_{kt} = U_{1t}$$

$$a_{21}Y_{1t} + a_{22}Y_{2t} + \dots + a_{2G}Y_{Gt} + b_{21}X_{1t} + b_{22}X_{2t} + \dots + b_{2k}X_{kt} = U_{2t}$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$a_{G1}Y_{1t} + a_{G2}Y_{2t} + \dots + a_{GG}Y_{Gt} + b_{G1}X_{1t} + b_{G2}X_{2t} + \dots + b_{Gk}X_{kt} = U_{Gt}$$

where Y's are endogenous variables, the X's are pre-determined variables, U's are the stochastic disturbance terms and t is time. The a's and b's are the structural coefficients. There are G endogenous and k pre-determined variables in the system. The underlying theory in general will satisfy that only a small number of all the variables of the model will occur in any one equation, and the matrices A and B will have a considerable number of zeros. Further in each structural equation, one endogenous variable can by convention be regarded as the dependent variable, the coefficient of which is equal to unity. This is known as the normalization rule. Again, some equations may be identities which means all their coefficients are known and that they contain no stochastic disturbances.

The model may be written in the matrix form as ...

$$AY_t + BX_t = U_t \qquad (1)$$

where A is a G x G matrix of coefficients of current endogenous variables. B is a G x K matrix of coefficients of pre-determined variables. Y_t is G x 1 column vector, X_t is a K x 1 column vector and U_t is also a column vector of order G x 1.

That is...

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1G} \\ a_{21} & a_{22} & \dots & a_{2G} \\ a_{G1} & a_{G2} & \dots & a_{GG} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1K} \\ b_{21} & b_{22} & \dots & b_{2K} \\ b_{G1} & b_{G2} & \dots & b_{GK} \end{bmatrix}$$

$$Y_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{Gt} \end{bmatrix} \qquad X_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Kt} \end{bmatrix} \qquad U_t = \begin{bmatrix} U_{1t} \\ U_{2t} \\ \vdots \\ U_{Gt} \end{bmatrix}$$

If there are constant terms in any of the equations one of the x's will be equal to unity for all t=1, 2,...,T.

In this model, A is a square matrix and is assumed to be non-singular, since if it were not, one or more of the structural relations would merely be a linear combination of other structural relations, thus being redundant, or if the rows of B did not obey the same linear restrictions as the rows of A the G structural equations would be inconsistent. The B matrix is generally not a square matrix. The G equations in the structural form jointly determine for each observation, the value of G endogenous variables, given K pre-determined variables, G stochastic disturbance terms and $G^2 + GK$ coefficients of the system. The stochastic disturbance term U_t are identically and independently distributed over the samples with zero mean and constant convariance matrix.

Since A is assumed to be a non-singular matrix, we can solve for the vector of endogenous variables Y_t by pre-multiplying (1) by A^{-1}. Thus we get...

$$A^{-1} AY_t = -A^{-1} BX_t + A^{-1}U_t$$

$$\underset{G\times 1}{Y_t} = \underset{G\times k}{P}\ \underset{K\times 1}{X_t} + \underset{G\times 1}{V_t} \qquad \ldots\ldots (2)$$

where,

$$\underset{G\times k}{P} = \underset{G\times G}{-A^{-1}}\ \underset{G\times K}{B} \quad (\text{or } AP = -B) \qquad \ldots\ldots (3)$$

$$\underset{G\times 1}{V_t} = \underset{G\times G}{A^{-1}}\ \underset{G\times 1}{U_t} \quad (\text{or } AV_t = U_t) \qquad \ldots\ldots (4)$$

In matrix form the P and V_t matrices can be written as ....

$$P = \begin{bmatrix} P_{11} & P_{12} & \ldots & P_{1K} \\ P_{21} & P_{22} & \ldots & P_{2K} \\ P_{G1} & P_{G2} & \ldots & P_{GK} \end{bmatrix} \qquad V_t = \begin{bmatrix} V_{1t} \\ V_{2t} \\ \vdots \\ V_{Gt} \end{bmatrix}$$

Equation (2) is the reduced form which expresses each of the endogenous variable Y_t as a linear function of all pre-determined variables X_t and stochastic disturbance term U_t. The coefficient matrix P is the matrix of the reduced form coefficients and V_t is the vector of reduced form stochastic disturbance terms.

1.3 THE IDENTIFICATION PROBLEM :

In simultaneous equation model the ordinary least square method fails to give consistent estimators because in the equations the endogenous variables are present among the explanatory variables, at least in general. However, in the reduced form equation, the explanatory variables are represented by the pre-determined variables of the system so that OLS estimators of the reduced form coefficients are consistent. "This suggests that we may try to estimate the structural coefficients via the reduced form. But here again a question may be raised - why do we need to estimate the structural coefficients, when the reduced form coefficients which can be known tell us all we need about the economic process under review by determining the outcome for any given values of the pre-determined variables"

The reasons why we must try to establish the structral coefficients have been given by Marschak (1953). The basic argument is that by its very definition each structural equation describes a specific, distinct and thereby autonomous link in the economic process. This autonomous character of structural relations also means that the structural coefficients are more stable, more like physical constants, than are the reduced form composites. Structural coefficients unlike the reduced form coefficients are more easily judged by intuition and their changes are better capable of reasonable discussion and interpretation. Structural coefficients are more easily comparable with the accumulated empirical evidence which refers to the economic structure and not to the reduced form.

Now, given the structural form ...

$$\underset{G\times G}{A}\ \underset{G\times 1}{Y_t} + \underset{G\times K}{B}\ \underset{K\times 1}{X_t} = \underset{G\times 1}{U_t}$$

and the reduced form ....

$$\underset{G\times 1}{Y_t} = \underset{G\times K}{P}\ \underset{K\times 1}{X_t} + \underset{G\times 1}{V_t}$$

the question that arises is : Can we obtain the estimates of the parameters of the structural form, viz. coefficient matrices A and B and the covariance matrix Σ of ...

$$\mathrm{Cov}\ (U_t) = E(U_t\ U_t') = \underset{G\times G}{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1G} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2G} \\ \vdots & \vdots & & \vdots \\ \sigma_{G1} & \sigma_{G2} & \cdots & \sigma_{GG} \end{bmatrix}$$

given the estimates of the parameters of the reduced form, viz. coefficient matrix P and covariance matrix Ω of...

$$Cov(V_t) = E(V_t V_t') = E[B^{-1} U_t U_t' (B^{-1})'] = B^{-1}\Sigma(B^{-1})' = \Omega ?$$

This is where the problem of identification arises. In order to estimate the parameters, it is necessary that the equations are identified. So the problem of identification is a very important problem in the estimation of parameters for it deals with the problem of finding a unique solution for the structural parameters from the reduced form coefficients.

In case the parameters can be estimated from the reduced form parameters we say that the system is identifiable otherwise we say that the system is underidentified or unidentified. In other words, a system of structural equation is identified if and only if every equation of the system is identified; if any equation is not identified, the system is not identified.

A structural equation that is identified is just or exactly identified if and only if there is a unique way of calculating its parameters. It is over-identified if there is more than one way to calculate its parameters from the reduced form parameters, leading to restrictions on the reduced-form parameters.

1.4 THE IDENTIFICATION OF A STRUCTURAL EQUATION :

It would be desirable to have some general rule for determining the identification status of any given structural equation. Such a rule can be derived in the following way...

The structural equations are :

$$A Y_t + B X_t = U_t$$

and the reduced form equations are :

$$Y_t = P X_t + V_t$$

where,

$$V_t = A^{-1}U_t$$

and $P = -A^{-1}B$

or, $AP = -B$

This is the relation used for determining the identification status of a structural equation. Writing the relation in matrix form, we have ...

$$\begin{bmatrix} a_{11} & a_{12} \cdots & a_{1G} \\ a_{21} & a_{22} \cdots & a_{2G} \\ \vdots & & \\ a_{G1} & a_{G2} \cdots & a_{GG} \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \cdots & P_{1k} \\ P_{21} & P_{22} \cdots & P_{2k} \\ & & \\ P_{G1} & P_{G2} \cdots & P_{Gk} \end{bmatrix} = - \begin{bmatrix} b_{11} & b_{12} \cdots & b_{1k} \\ b_{21} & b_{22} \cdots & b_{2k} \\ & & \vdots \\ b_{G1} & b_{G2} \cdots & b_{Gk} \end{bmatrix}$$

For a single equation of the system, say the first equation, this becomes...

$$\begin{bmatrix} a_{11} & a_{12} \cdots & a_{1G} \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \cdots & P_{1k} \\ P_{21} & P_{22} \cdots & P_{2k} \\ \vdots & \vdots & \\ P_{G1} & P_{G2} \cdots & P_{Gk} \end{bmatrix} = - \begin{bmatrix} b_{11} & b_{12} \cdots & b_{1k} \end{bmatrix} \qquad \ldots\ldots (5)$$

or, $a_1 P = -b_1$ ..... (5a)

If all the endogenous and pre-determined variables of the system do not appear in the first equation, some of the a's and some of the b's will be equal to zero.

Let G^{Δ} = number of endogenous variables which appear in the 1st equation. (ie, the number of nonzero elements in a_1) ;

$$G^{\Delta\Delta} = G - G^{\Delta} \quad ;$$

K^{*} = number of pre-determined variables which appear in the 1st equation;

$$K^{**} = K - K^{**} \quad ;$$

Without any loss of generality, we assume that the elements of a_1 and b_1 are arranged in such a way that the non-zero elements appear first and is followed by the zero elements. Then a_1 and b_1 can be partitioned as...

$$\left.\begin{aligned} a_1 &= \begin{bmatrix} a_{\Delta} & 0_{\Delta\Delta} \end{bmatrix} \\ b_1 &= \begin{bmatrix} b_{*} & 0_{**} \end{bmatrix} \end{aligned}\right\} \qquad \ldots\ldots (6)$$

where, $\underset{(1 \times G^{\Delta})}{a_{\Delta}} = [a_{11} \quad a_{12} \quad \ldots \quad a_{1G}]$

$\underset{(1 \times G^{\Delta\Delta})}{0_{\Delta\Delta}} = [\,0 \quad 0 \quad \ldots \quad 0\,]$

$$B_{*\,(1\times k^{*})} = [\, b_{11} \quad b_{12} \quad \ldots \quad b_{1k^{*}} \,]$$

$$O_{**\,(1\times k^{**})} = [\, 0 \quad 0 \quad \ldots \quad 0 \,]\,.$$

The (G x K) matrix P can be partitioned correspondingly. It need not contain any zeros but its rows correspond to the endogenous variables and its columns to pre-determined variables, so we write...

$$P = \begin{bmatrix} P_{\Delta *} & P_{\Delta **} \\ P_{\Delta\Delta *} & P_{\Delta\Delta **} \end{bmatrix} \quad \ldots \quad (7)$$

where,

$$P_{\Delta *} \rightarrowtail (G^{\Delta} \times K^{*})$$
$$P_{\Delta **} \rightarrowtail (G^{\Delta} \times K^{**})$$
$$P_{\Delta\Delta *} \rightarrowtail (G^{\Delta\Delta} \times K^{*})$$
$$P_{\Delta\Delta **} \rightarrowtail (G^{\Delta\Delta} \times K^{**})$$

By using (6) and (7), we can rewrite (5a) as...

$$\begin{bmatrix} a_{\Delta} & 0_{\Delta\Delta} \end{bmatrix} \begin{bmatrix} P_{\Delta *} & P_{\Delta **} \\ P_{\Delta\Delta *} & P_{\Delta\Delta **} \end{bmatrix} = -\,[\, b_{*} \quad 0_{**} \,] \quad \ldots\ldots \quad (8)$$

This leads to the following equalities :

$$a_{\Delta} \; P_{\Delta *} = -b_{*} \quad \ldots\ldots \quad (9)$$

$$a_{\Delta} \; P_{\Delta **} = 0_{**} \quad \ldots\ldots (10)$$

Since one of the a's in each structural equation is equal to unity, the equalities (9) and (10) involve $(G^{\Delta} -1)$ unknown a's and k^{*} unknown b's. The equality (10) is particularly important for us as it does not involve any b's. If we can find the solution of a_{Δ} from (10), we can solve for b_{*} easily from (9) by substituting the values of a_{Δ} in (9). The equality (10) contains altogether K^{**} equations, one for each element of the $(1\times K^{**})$ vector. So, if we want to obtain a solution for the $(G^{\Delta} -1)$ unknown elements of a_{Δ}, we need at least $(G^{\Delta} -1)$ equations. That means we require...

$$K^{**} >= G^{\Delta} - 1 \quad \ldots\ldots \quad (11)$$

This is known as the order condition of identification. This condition which is necessary for the identification of a structural equation states that the number of pre-determined variables excluded from the given equation is at least as large as the number of endogenous variables included in the equation less one. It should be noted that this condition is only a necessary and not a sufficient condition for identification since the k^{**} equations in (10) may not be independent. That is the equations in (10) may contain less than $G^{\Delta}-1$ different pieces of information about the relation between the a's and the P's. Therefore, the necessary and sufficient condition for identification is that the number of independent equations in (10) is $G^{\Delta}-1$. This will be the case if and only if the order of the largest non-zero determinant that can be formed from all square submatrices of $P_{\Delta**}$ is $G^{\Delta}-1$, that is if and only if

$$rank\ (P_{\Delta**}) = G^{\Delta}-1 \qquad \ldots\ldots (12)$$

This is known as the rank condition for identifiability.

The rank of $P_{\Delta**}$, can be determined by partitioning the structural coefficient matrices as follows:

$$A = \begin{bmatrix} a_{\Delta} & 0_{\Delta\Delta} \\ A_{\Delta} & A_{\Delta\Delta} \end{bmatrix} \qquad B = \begin{bmatrix} b_{*} & 0_{**} \\ B_{*} & B_{**} \end{bmatrix}$$

where a_{Δ}, b_{*}, $0_{\Delta\Delta}$ and 0_{**} are row vectors defined as in (6) and

where,

$$A_{\Delta} \quad \rightarrowtail\!\!\!\!\rightarrow \quad (G-1) \times G^{\Delta} \quad ,$$
$$A_{\Delta\Delta} \quad \rightarrowtail\!\!\!\!\rightarrow \quad (G-1) \times G^{\Delta\Delta} \quad ,$$
$$B_{*} \quad \rightarrowtail\!\!\!\!\rightarrow \quad (G-1) \times K^{*} \quad ,$$
$$B_{**} \quad \rightarrowtail\!\!\!\!\rightarrow \quad (G-1) \times K^{**} \quad ,$$

It should be noted here that $A_{\Delta\Delta}$ and B_{**} are matrices of the structural coefficients for the variables omitted from the 1st equation but included in other structural equations. Let us now form a new matrix Θ defined as...

$$\Theta = [A_{\Delta\Delta} \quad B_{**}],$$

then,

$$\text{rank}\,(P_{\Delta**}) = \text{rank}\,(\theta) - G^{\Delta\Delta} \qquad .....(12a)$$

This can be proved as follows. Let θ_* be defined as...

$$\theta_* = \begin{bmatrix} 0_{**} & 0_{\Delta\Delta} \\ B_{**} & A_{\Delta\Delta} \end{bmatrix}$$

The rank of θ_* is the same as that of θ since the rank of a matrix is not affected by switching any columns or by adding a row of zeros to the matrix.

Now, θ_* can be written as ...

$$\theta_* = \begin{bmatrix} a_{\Delta} & 0_{\Delta\Delta} \\ A_{\Delta} & A_{\Delta\Delta} \end{bmatrix} \begin{bmatrix} -P_{\Delta**} & 0_{\Delta,\Delta\Delta} \\ -P_{\Delta\Delta**} & I_{\Delta\Delta} \end{bmatrix} .$$

where $0_{\Delta,\Delta\Delta}$ is a G^{Δ} x $G^{\Delta\Delta}$ matrix of zeros, and $I_{\Delta\Delta}$ is an identity matrix of order $G^{\Delta\Delta}$ x $G^{\Delta\Delta}$. To see that, we carry out multiplication in the above equality to obtain...

$$\theta_* = \begin{bmatrix} -a_{\Delta}P_{\Delta**} & 0_{\Delta\Delta} \\ (-A_{\Delta}P_{\Delta**} - A_{\Delta\Delta}P_{\Delta\Delta**}) & A_{\Delta\Delta} \end{bmatrix}$$

But by (11),

$$-a_{\Delta}P_{\Delta**} = 0_{**} .$$

and from the equality $AP = -B$ it follows that...

$$-A_{\Delta}P_{\Delta**} - A_{\Delta\Delta}P_{\Delta\Delta**} = B_{**}$$

Using the theorem that if a matrix C is multiplied by a non-singular matrix, the product has the same rank as C, we can write...

$$\begin{aligned} \text{rank}\,(\theta_*) &= \text{rank}\,(A^{-1}\theta_*) \\ &= \text{rank} \begin{bmatrix} -P_{\Delta**} & 0_{\Delta,\Delta\Delta} \\ -P_{\Delta\Delta**} & I_{\Delta\Delta} \end{bmatrix} \\ &= \text{rank} \begin{bmatrix} -P_{\Delta**} & 0_{\Delta,\Delta\Delta} \\ -P_{\Delta\Delta**} & I_{\Delta\Delta} \end{bmatrix} \begin{bmatrix} I_{\Delta**} & 0_{**,\Delta\Delta} \\ P_{\Delta\Delta**} & I_{\Delta\Delta} \end{bmatrix} \end{aligned}$$

$$= \text{rank} \begin{bmatrix} -P_{\Delta**} & 0_{\Delta.\Delta\Delta} \\ 0_{\Delta\Delta.**} & I_{\Delta\Delta} \end{bmatrix}$$

$$= \text{rank} \; (P_{\Delta\Delta**}) + G^{\Delta\Delta} \; .$$

where $0_{**,\Delta\Delta}$ and $0_{\Delta\Delta,**}$ are zero matrices of order $K^{**} \times G^{\Delta\Delta}$ and $G^{\Delta\Delta} \times K^{**}$, respectively, and I_{**} is an identity matrix of order $K^{**} \times K^{**}$. This completes the proof of (i)a.

A convenient way to check the rank condition is to write down the matrix of all coefficients in the structural form, cross out the columns in which there are non-zero entries in the 1st row, ie. if the equation to be identified is the first equation, strike out the 1st row itself and then check the rank of the resulting matrix.

The order and rank condition enables us to set up the following general rule for determining the identification status of a structural equation.

1. If $K^{**} = G^{\Delta} - 1$ and rank $(P_{\Delta**}) = G^{\Delta} - 1$, then the structural equation is exactly identified.

2. If $K^{**} > G^{\Delta} - 1$ and rank $P_{\Delta**} = G^{\Delta} - 1$, then the structural equation is over-identified.

3. If $K^{**} = G^{\Delta} - 1$ and rank $P_{\Delta**} < G^{\Delta} - 1$, then the structural equation is under-identified.

4. If $K^{**} < G^{\Delta} - 1$, then the structural equation is unidentified.

The above rank and order conditions for identifiability have been stated in terms of population parameters on the assumption that none of the structural parameters is equal to zero. But we do not know what the values of these parameters really are and would like to test the hypothesis that they are equal to zero. If this hypothesis were not rejected in every case, some variables would be considered irrelevant and therefore should not be counted while checking the order condition for identifiability of a structural equation. This means that a structural equation that appears to satisfy the order condition a priori, may not, in fact satisfy this condition when the irrelevant variables have been discarded. Therefore, we may want to consider the identifiability of an equation as a hypothesis to be tested instead of relying on prior specification.

1.5 <u>IDENTIFICATION THROUGH RESTRICTION ON THE DISTURBANCE VARIANCE-COVARIANCE MATRIX</u>:

In the above discussion, the examination of the identifiability conditions has been confined to the specification of the structural equation and no reference has been made to the variance-covariance matrix of the disturbance. We have shown that identification of a structural equation can be achieved by zero restrictions on some of the coefficients. Although we have not discussed, but identification can also be achieved by non-zero restrictions on the structural coefficients e.g. by specifying that some coefficients are equal to given numbers that are not necessarily zero, or by specifying the ratio or ratios between coefficients in a linear equation. If this is the case, then it should also be possible to achieve identification by prior restrictions on the variance-covariance matrix of the disturbances.

Let $\Sigma = E(u_t u_t')$

Σ is a $G \times G$ matrix, the terms on the principal diagonal indicating the variances (assumed constant) of the disturbances and the off diagonal elements, the covariances between the pairs of disturbances. If specific restrictions can be placed on some of these elements, they constitute an additional source of identifying power.

We will first examine the restrictions on the covariances. Let us consider the model

$$Y_1 + \gamma_{11} X_1 = U_1$$

$$\beta_{21} Y_1 + Y_2 + \gamma_{21} X_1 = U_2$$

$$\beta_{31} Y_1 + \beta_{32} Y_2 + Y_3 + \gamma_{31} X_1 = U_3$$

Without further restrictions only the first equation is identifiable and the second and third is not. We shall, however, examine the identifiability of the model again by considering admissible transformation matrices, as this approach facilitates the study of restrictions on variances and covariances.

It should be noted here that a transformation matrix which is non-singular is admissible if the transformed structure satisfies all the apriori restrictions on the original structure.

Using transformation matrix

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

the transformed first equation becomes...

$$(f_{11} + f_{12}\beta_{21} + f_{13}\beta_{31}) Y_1 + (f_{12} + f_{13}\beta_{32}) Y_2 + f_{13} Y_3 + (f_{11}\gamma_{11} + f_{12}\gamma_{21} + f_{13}\gamma_{31}) X_1 = f_{11} U_1 + f_{12} U_2 + f_{13} U_3$$

If the coefficients of the transformed equation are to obey the same restrictions as those of the original equation, we must have,

$$f_{11} + f_{12}\beta_{21} + f_{13}\beta_{31} = 1$$

$$f_{12} + f_{13}\beta_{32} = 0$$

$$f_{13} = 0$$

giving $f_{11} = 1$, $f_{12} = 0$ and $f_{13} = 0$. The admissible transformation matrix is then given by ...

$$F = \begin{bmatrix} 1 & 0 & 0 \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

showing that the first equation is identified and the second and third equation is not.

Suppose, we can now postulate...

$$\Sigma = \begin{bmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix}$$

The vector of disturbances in the transformed structure is $F u_t$ and so the variance-covariance matrix for the disturbances of the transformed structure is...

$$\rho = E\, F u_t u_t' F'$$
$$= F \Sigma F' \qquad \ldots\ldots \quad 10$$

Let us now consider...

$$F_4 = \begin{bmatrix} 1 & 0 & 0 \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \gamma_{11} \\ \beta_{21} & 1 & 0 & \gamma_{21} \\ \beta_{31} & \beta_{32} & 1 & \gamma_{31} \end{bmatrix}$$

The normalization condition on Y_2 in the second equation and on Y_3 in the third give...

$$f_{22} + f_{23}\ \beta_{32} = 1$$

$$f_{33} = 1$$

and the exclusion of Y_3 from the second equation gives...

$$f_{23} = 0$$

which also implies $\quad f_{22} = 1$

Thus F is now...

$$F = \begin{bmatrix} 1 & 0 & 0 \\ f_{21} & 1 & 0 \\ f_{31} & f_{32} & 1 \end{bmatrix}$$

We have not yet considered the effect of zero convariance restrictions $\sigma_{12} = \sigma_{13} = \sigma_{23} = 0$. These must be satisfied by the transformed structure (13). Hence...

$$f_1 \Sigma f_2' = 0$$

$$f_1 \Sigma f_3' = 0$$

$$f_2 \Sigma f_3' = 0$$

The first of these gives...

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix} \begin{bmatrix} f_{21} \\ 1 \\ 0 \end{bmatrix} = f_{21} \sigma_{11} = 0$$

so that, $f_{21} = 0$

and in the same way the second and the third conditions give $f_{31} = 0$ and $f_{32} = 0$. Therefore, the only admissible transformation matrix is...

$$F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and all the three equations are identified.

The above model has two features, viz. a triangular B matrix and a diagonal Σ matrix. The presence of these two features defines a recursive system. All the equations of the recursive system can be identified and simple estimation procedures are available for such models.

Let us now consider the restrictions on the terms on the principal diagonal of Σ. For this purpose, let us assume that σ_{11} is equal to zero. This implies that the variance of the first equation is zero and this equation is exact rather than stochastic. To consider its identifiability we use two conditions. Firstly, any transformed first equation must satisfy the apriori restriction of the first equation that is...

$$f_1 A \phi = 0$$

Secondly, the disturbance of any transformed first equation must be zero, that is...

$$f_1 \Sigma f_1' = 0$$

Putting these two conditions together gives...

$$f_1 [A\phi \;\; \Sigma] = 0 \qquad \ldots\ldots \quad (14)$$

If the first equation is to be identified, the f_1 vectors satisfying (14) must be scalar multiples of one another. Thus the necessary and sufficient condition for the identifiability of the first equation under the additional restriction $\sigma_{11} = 0$ is...

$$\rho [A\phi \;\; \Sigma] = G - 1 \qquad \ldots\ldots \quad (15)$$

It should be noted that if all the other G disturbances are non zero $\rho(\Sigma) = G-1$ and the rank condition of (15) will be satisfied even if there are no apriori restrictions on the β and γ coefficients.

Let the model be ...

$$Y_1 - \gamma_{11} X_1 = u_1$$

$$\beta_{21} Y_1 + Y_1 + \gamma_{21} X_1 = u_2$$

Without restrictions on Σ the second equation is not identified. If however, the second equation becomes...

$$\beta_{21} Y_1 + Y_2 + \gamma_{21} X_1 = 0$$

then,

$$\Sigma = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

and for the second equation...

$$[A\ \Phi] = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

with rank 1 = G-1, so that the second equation is identified.

It should be remembered while solving the identification problem that a non-stochastic equation which occurs frequently in the econometric model is the identity, where the disturbance term is zero and the coefficients of the variables are known. For instance,

$$\equiv \quad C + \quad I + \quad G$$

National Income	Cons exp	Inv. exp.	Govt expenditure on goods and services.

and $$Q_D \equiv Q_S$$

Quantity demanded	Quantity supplied

In such cases problem of identification does not arise.

In our discussion we have dealt only with models which are linear in variables and parameters. Many realistic models, however, may be non-linear in variables and/or a priori restrictions. Identification theory for such model is difficult and has only been partially developed. It is for this reason that we have not summarized such cases here.

CHAPTER - II

OVER - IDENTIFICATION

OVER-IDENTIFICATION

2.1 MEANING OF OVER-IDENTIFICATION:

A model is said to be identified if it is in a unique statistical form, ie, if it enables us to find unique estimates of its parameters from the sample data. Given the model...

$$AY_t + BX_t = U_t$$

consisting of G structural relations and the corresponding reduced form...

$$Y_t = PX_t + V_t$$

where,

$$P = - A^{-1} B$$

and $$V_t = A^{-1}U_t$$

an equation of the model is identified if it satisfies the order and rank conditions. viz..

(1) The number of pre-determined variables dropped from the given equation is at least as great as the number of endogenous variables included in the equation less one.

(2) The equation is identified if and only if it is possible to form at least one non-zero determinant of order (G-1) from the coefficients of the variables excluded from that particular equation but contained in other equations of the model.

The entire model or system is identified if every equation of the system is identified. If any equation is not identified, the system is said to be unidentified. Though the above model of G structural equations may be identified, yet exact identification may not be possible because of a number of inadequacies present in the system. For instance, there may be more equations than the number of unknowns to be estimated in the system $AP = -B$. If this is the case, we can hardly expect the system to be consistent. The result in such cases is that there will not be a unique solution to the parameters estimated. On the contrary, there will be more than one way to calculate the structural parameters from the reduced form parameters giving multiple estimates which need not be identical. This is the case of over-identification. A particular equation of the system is said to be over-identified if the number of pre-determined variables excluded from the equation is greater than the number of endogenous variables included in the equation less one.

A case of overidentification is illustrated below :

$$C_t = a_0 + a_1 Y_t + a_2 C_{t-1} + U_{1t} \quad \text{(consumption)}$$

$$I_t = b_0 + b_1 r_t + b_2 I_{t-1} + U_{2t} \quad \text{(investment)}$$

$$r_t = c_0 + c_1 Y_t + c_2 M_t + U_{3t} \quad \text{(money market)}$$

$$Y_t = C_t + I_t + G_t \quad \text{(income identity)}$$

Let us first test the consumption function for identification.

1. Order condition :

The consumption function includes 3 variables and excludes 3 variables.

Hence, $K^{**} > G-1$. The order condition therefore is fulfilled.

2. Rank condition :

Deleting the first row and the first, second and fifth columns of the structural coefficient matrix we obtain a matrix of coefficients of excluded variables

$$\begin{bmatrix} 1 & -b_1 & -b_2 & 0 & 0 \\ 0 & 1 & 0 & -c_1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The rank of the above matrix is 3. The rank condition is satisfied since we can construct at least one non-zero determinant of order 3 = (G-1).

In the order condition we have seen that the inequality sign holds as 3>1. Hence the consumption function is over-identified. The identification of the other equations would follow along the same lines.

2.2 DEGREE OF OVER-IDENTIFICATION :

A model if identified, is either exactly identified or over-identified. Now, a question arises about the degree of identification, the meaning of which surpasses the mere counting of restrictions or determining the rank of a matrix.

If identification is achieved by excluding certain variables from some equations which are otherwise present in the system, these variables must make a difference to the statistical explanation of the system in order to give meaning to their exclusion from particular equations. The reason being everything depends on everything else in an interrelated system. Although some equations may be identified as they satisfy the formal conditions or even strongly identified due to the significant specification of restrictions, the system as a whole may not be identified because other individual equations of the system are not identified.

We should also point out here that the rank condition involves the properties of determinants of structural parameters which may be satisfied for some values of parameters and not for others. Since we do not know the true values of the parameters, which are to be estimated through statistical inference, it is not possible to determine the true values of a determinant involved in the rank condition, and therefore it cannot also be stated whether the system is definitely identified. It is therefore not the rank condition but the order condition which gives us certain fair information.

Hence it is the number of endogenous variables present or pre-determined variables absent which will determine the degree of identification, rather over-identification which is of our concern. The larger the number of pre-determined variables dropped than the number of endogenous variables included, the greater will be the degree of over-identification. But here again, we cannot say with certainity that no pre-determined variables are missing and no irrelevant endogenous variables are present. Some of the endogenous variables in the equation may be irrelevant or some pre-determined variables may be missing and if this be so it will be a case of wrong specification. Thus over-identification may be considered as a case of irrelevant endogenous variables or a case of missing pre-determined variables or a mix of both the cases.

It may be pointed out here that the presence of irrelevant endogenous variables under certain conditions do not affect the estimation of the parameters provided that they are uncorrelated, but missing exogenous variables have an adverse effect on the estimation. They always make the result inefficient. That it so happens, can be shown by taking a regression equation and then omitting a relevant variable or by including an irrelevant variable.

Let us suppose that the true regression equation is ...

$$Y = b_1 x_1 + b_2 x_2 + u$$

Instead we estimate the equation $Y = b_1 x_1 + v$

The least square estimator of b_1 from the equation is

$$\hat{b}_1 = \frac{\Sigma y x_1}{\Sigma x_1^2}$$

$$= \frac{\Sigma x_1 (b_1 x_1 + b_2 x_2 + u)}{\Sigma x_1^2}$$

$$= b_1 + b_2 \frac{\Sigma x_1 x_2}{\Sigma x_1^2} + \frac{\Sigma x_1 u}{\Sigma x_1^2}$$

Taking expectation

$$E(\hat{b}_1) = b_1 + b_2 \frac{\Sigma x_1 x_2}{\Sigma x_1^2} \quad [\text{since } E(U) = 0]$$

Hence, the bias equals the true coefficient of the excluded variable times the regression coefficient of the omitted variable on the variable that is included. Now, if the x's are random variables we take plim instead of expectation, and we get a similar result about asymptotic bias. The bias however will be zero if x_1 and x_2 are independent. But in such a case the residuals will include x_2 and the estimated variance of b_1 will be upward biased.

Let us now consider the case of inclusion of an irrelevant variable. Suppose the true equation is $y = b_1 x_1 + u$ but we estimate the equation $y = b_1 x_1 + b_2 x_2 + v$

The least square estimates of b_1 and b_2 from the misspecified equation are...

$$b_1 = \frac{\Sigma x_2^2 \, \Sigma x_1 y - \Sigma x_1 x_2 \, \Sigma x_2 y}{\Sigma x_1^2 \, \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

$$= \frac{\Sigma x_2^2 \, \Sigma x_1 (b_1 x_1 + u) - \Sigma x_1 x_2 \, \Sigma x_2 (b_1 x_1 + u)}{\Sigma x_1^2 \, \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

$E(b_1) = b_1$ [since $E(u) = 0$]

$$b_2 = \frac{\Sigma x_1^2 \, \Sigma x_2 y - \Sigma x_1 x_2 \, \Sigma x_1 y}{\Sigma x_1^2 \, \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

$$= \frac{\Sigma x_1^2 \, \Sigma x_2 (b_1 x_1 + u) - \Sigma x_1 x_2 \, \Sigma x_1 (b_1 x_1 + u)}{\Sigma x_1^2 \, \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

$E(b_2) = 0$ [since $E(u) = 0$]

Thus we get unbiased estimates of the parameters b_1 and b_2 Since omission of relevant variables lead to biased estimates, one may believe that it is better to include variables in the equation when in doubt rather than exclude them. But this is not a right thing to do, because though the inclusion of irrelevant variables has no effect on the bias of the estimators, it does affect the variances.

If b_1^* is the estimator of b_1 from the correct equation, then ...

$$Var\ (b_1^*) = \sigma^2 / \Sigma x_1^2.$$ But...

$$Var\ (\hat{b}_1) = \sigma^2 / \Sigma x_1^2 (1-r_{12}^2)$$

where r_{12} is the correlation between x_1 and x_2.

Thus Var $(\hat{b}_1)$ =Var (b_1^*), the equality holding for $r_{12} = 0$. Therefore, if we include irrelevant variables we will get unbiased but inefficient estimates. It can be shown, however, that the estimator for the residual variance we use is an unbiased estimator for σ^2. Thus, no further bias arises from the use of estimated variances from misspecified equation.

2.3 THE ESITMATION OF OVERIDENTIFIED EQUATIONS

From the discussion of identification, it follows that estimation of an over-identified equation is possible. Since the model is a simultaneous equation model, OLS cannot be used in estimating the parameters of an equation as there will be correlation between regressors and the residuals giving inconsistent and biased estimates. we therefore, have to use other alternative methods of estimation such as the indirect least squares method (ILS), two stage least squares method (2SLS), limited information maximum-likelihood method (LIML),three stage least squares method (3SLS) and full information maximum likelihood method (FIML), ILS, 2SLS and LIML are essentially single equation methods in which each equation is estimated separately using only the information about the restrictions on the coefficients of that particular equation. On the other hand, 3SLS and FIML are system methods where all the equations of the fully specified model are simultaneously estimated.

The Indirect Least Squares method proceeds by least square estimation of the reduced form parameters and then uses the prior restrictions to transform the reduced form coefficients to structural parameters. Now, if the equation to be estimated is an exactly identified equation, there is one way to do this and we get unique solutions with this method and if the equation is underidentified, there will be no solutions. On the other hand, if the equation is over-identified, there will be more than one way to go from the reduced form to the structural equation and as a result there will be multiple solutions. Again, in over-identified equations there are more than required restrictions than is necessary to estimate the parameters of the equation in question. The difficulty that arises is that in a finite sample there is no guarantee that estimates obtained by a set of restrictions will obey those that are not used in the derivation.

The difficulty with ILS is that it does not use the restrictions to estimate the reduced form parameters and as such the reduced form estimates will in general not be compatible with the restrictions that are used in estimating the structural parameters. On the other hand, over-identified equations imply the existence of restrictions not only on the structural parameters, but also on the parameters of the reduced form equations. Therefore, ILS cannot be used to estimate the parameters of over-identified equations as it will fail to give consistent estimates. The technique is only feasible when the equation is just identified.

The most widely used method for estimating the parameters of an over-identified equation is the two stage least squares developed by H. Theil and independently by Basmann who calls it Generalised Classical linear estimation method. The basic idea of the method is to substitute for the endogenous variables which are correlated with the residuals, linear functions of all exogenous variables. Since the variables are uncorrelatedwith the residuals, the estimates of the parameters will be consistent.

The Limited - information Maximum likelihood method is another alternative approach which gives consistent estimates and was developed by Anderson and Rubin. The estimates of the parameters of an equation are arrived at by maximising the likelihood function for the observations on the endogenous variables included in that equation disregarding the over-identifying restrictions on other strutural equations. This method assumes that the structural disturbances are normally distributed.

Though the single equation estimation methods 2SLS and LIML gives consistent estimates, they are not asymptotically efficient as they do not consider the correlation of disturbances across the equation. This deficiency can be overcome by estimating all the equations of the system simultaneously. For this purpose the full information methods such as 3SLS or FIML can be used.

The three-stage least squares method devised by Zellner and Theil begins by estimating each structural equation of the model separately by two-stage least squares, subject only to the identifying restrictions on that equation. The estimates obtained are then used to calculate the residuals in each stochastic structural equation, and those residuals are used in turn to calculate the estimates of the variance - covariance matrix of the structural disturbance. In the 'third stage' the coefficients of all the equations are simultaneously estimated, by means of Aitken's [1934] generalized least-squares method using the estimated variance-covariance matrix and the identifying restrictions on all the coefficients in the model.

Another approach which gives consistent estimates of the parameters of the structural model is the full information maximum likelihood method. It involves the application of maximum likelihood principle on the equations of the system simultaneously under the specification that the structural disturbances are normally distributed, using all the restrictions on the structural equations.

A study of experiments carried out by different econometricians reveals that in the case of an over-identified equation, the results obtained by 2SLS and 3SLS are identical on the one hand, and LIML and FIML on the other. However, it may be noted that if the model to be estimated contains one exactly identified equation and more than one over-identified equation, the results obtained by the single equation methods would differ from the system methods.

CHAPTER - I I I

SIMULATION OF OVER-IDENTIFIED MODELS & ESTIMATION

SIMULATION OF OVER-IDENTIFIED MODELS AND ESTIMATION

We have used the two stage least squares method (2SLS) developed by Theil and the generalised indirect least squares (GILS) method developed by Mishra, in estimating the parameters of over-identif--ied equations of different models. It will therefore be intelligible to describe the methods at the very outset.

3.1 TWO STAGE LEAST SQUARES

Suppose the equation to be estimated is the first equation of the general interdependent system of equations ...

$$y = Y_1 a_1 + X_1 b_1 + u_1 \qquad \text{..... (1)}$$

where $y_1 =$ (T x 1) vector of endogenous variables whose coefficient in the first equation is 1

$Y_1 =$ (T x (G^{Δ} -1)) matrix of the remaining endogenous variables in the first equation.

$X_1 =$ (T x K^{*}) matrix of the pre-determined variables in the equation

$u_1 =$ (T x 1) vector of the disturbances in the equation

In the first stage the matrix $\hat{Y}$ is computed by regressing each variable in Y_1 on all the pre-determined variables in the system and then replacing the actual observations on the y variables by the corresponding regression values. Therefore,

$$\hat{Y}_1 = X(X'X)^{-1}X'Y_1 \qquad \text{..... (2)}$$

In the second stage the regression of y on $\hat{Y}_1$ and X_1 yields the estimating equations

$$\begin{bmatrix} \hat{Y}_1'\hat{Y}_1 & \hat{Y}_1'X_1 \\ X_1'\hat{Y}_1 & X_1'X_1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \hat{Y}_1'y \\ X_1'y \end{bmatrix} \qquad \text{...(3)}$$

where the vector $\begin{bmatrix} c \\ d \end{bmatrix}$ will now denote the 2SLS estimator of $\begin{bmatrix} a \\ b \end{bmatrix}$ The above system contains $G^{\Delta} -1 + K^{*}$ equations in

$G^{\Delta}-1+K^{*}$ unknowns which will in general have a unique solution.

The form in which the 2SLS equations are usually presented can be derived in the following way :
The matrix Y_1 can be written as...

$$Y_1 = \hat{Y}_1 + V_1$$

where V_1 is the TX $(G^{\Delta}-1)$ matrix of OLS residuals. Now, since the OLS residuals are orthogonal to the estimated value of the dependent variable and to each of the explanatory variables, therefore...

$$\hat{Y}_1' V_1 = 0 \text{ and } X_1' V_1 = 0$$

$$\hat{Y}_1'\hat{Y}_1 = (Y_1 - V_1)'(Y_1 - V_1)$$

$$= Y_1'Y_1 - V_1'Y_1 - Y_1'V_1 + V_1'V_1$$

$$= Y_1'Y_1 - (\hat{Y}_1 + V_1)'V_1 - V_1'(\hat{Y}_1 + V_1) + V_1'V_1$$

$$= Y_1'Y_1 - \hat{Y}_1'V_1 - V_1'V_1 - V_1'\hat{Y}_1 - V_1'V_1 + V_1'V_1$$

$$= Y_1'Y_1 - V_1'V_1$$

$$\hat{Y}_1'X_1 = Y_1'X_1 - V_1'X_1$$

$$= Y_1'X_1$$

$$X_1'\hat{Y}_1 = X_1'Y_1 - X_1'V_1 = X_1'Y_1$$

and $$\hat{Y}_1'y_1 = (Y_1 - V_1)'y_1 = Y_1'y_1 - V_1'y_1$$

Thus the 2SLS estimator can now be written as

$$\begin{bmatrix} Y_1'Y_1 - V_1'V_1 & Y_1'X_1 \\ X_1'Y_1 & X_1'X_1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} Y_1'y_1 - V_1'y_1 \\ X_1'y_1 \end{bmatrix}$$

The form shows clearly how the 2SLS estimator differs from the inconsistent OLS estimator which is given by

$$\begin{bmatrix} Y_1' Y_1 & Y_1' X_1 \\ X_1' Y_1 & X_1' X_1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} Y_1' y_1 \\ X_1' y_1 \end{bmatrix}$$

3.2 GENERALISED INDIRECT LEAST SQUARES :

The system of structural equations is written in matrix form as

$$AY + BX = U$$

The reduced form of the model is

$$Y = PX + V$$

where $V = A^{-1}U$

and $P = -A^{-1}B$

or $A\,P = -B$

Using OLS method, we estimate the reduced form coefficient matrix P. Now the above relation becomes $A\hat{P} = -B$. Writing this relation for the c^{th} equation we get

$$a_c \hat{P} = -\,b_c \qquad \ldots\ldots (1)$$

Where a_c is the c^{th} row vector of matri A and b_c is the c^{th} row vector of matrix B.

Now, if we arrange a_c and b_c such that each of them can be partitioned into two sub-vectors

$$a_c = (a_{1c} \mid a_{2c})$$

$$b_c = (b_{1c} \mid b_{2c})$$

where all elements of a_{2c} and b_{2c} are known and all elements of a_{1c} and b_{1c} are unknown. Accordingly we reshuffle the rows and columns of $\hat{P}$. Then,

$$\hat{P} = \begin{bmatrix} \hat{P}_{11} & \hat{P}_{21} \\ \hat{P}_{12} & \hat{P}_{22} \end{bmatrix}$$

Writing (1) in matrix form

$$[\,a_{1c} \mid a_{2c}\,] \left[\begin{array}{c|c} \hat{P}_{11} & \hat{P}_{21} \\ \hline \hat{P}_{12} & \hat{P}_{22} \end{array}\right] = -\,[\,b_{1c} \mid b_{2c}\,]$$

whence, $a_{1c}\ \hat{P}_{11} + a_{2c}\ \hat{P}_{12} = -\ b_{1c}$ ..... (2)

$$a_{1c}\ \hat{P}_{21} + a_{2c}\ \hat{P}_{22} = -\ b_{2c} \quad \text{...... (3)}$$

Since a_{2c} and b_{2c} are known.

$$a_{1c} = (-a_{2c}\ \hat{P}_{22} - b_{2c})\quad (\hat{P}_{21})^{-1}$$

$\hat{P}_{21}$ can only be inverted iif it is a square matrix and has full rank.

Thus,

$$a_{1c} = -\ (a_{2c}\ \hat{P}_{22} + b_{2c})\ (\hat{P}_{21})^{-1}$$

gives ILS

Once a_{1c} is known it can be substituted in (2) and b_{1c} can be known.

However, in case of over-identified equations, $\hat{P}_{21}$ is not a square matrix. Hence it is not possible to obtain its ordinary inverse.

We may however, find the generalised inverse of $\hat{P}_{21}$ which exists in a case of over-identified equations because the number of equations formed by (3) is larger than the number of unknowns to be estimated. Thus for over-identified equations

$$a_{1c} = -\ (a_{2c}\ \hat{P}_{22} + b_{2c})\ (\hat{P}_{21})^{-g} \quad \text{..... (4)}$$

system (4) has a solution and the technique (3) of obtaining a_{1c} may be called GILS.

$(\hat{P}_{21})^{-g}$ has a least square g inverse given by

$$[(\hat{P}_{21})' (\hat{P}_{21})]^{-1} (\hat{P}_{21})' = (\hat{P}_{21})^{-g}$$

Thus. $a_{1c} = -(a_{2c} \hat{P}_{22} + b_{2c}) (\hat{P}_{21})^{-g}$

This estimator of a_{1c} will be calledGeneralised Indirect Least Square Estimator or GILS.

3.3 METHODOLOGY AND ANALYSIS :

In our endeavour to find out the effect of various degrees of over-identification on the performance of estimators, we have made use of Monte Carlo experiments. The essence of a Monte Carlo study is that different sets of parameter values are specified for postulated distribution that is underlying the model. Large number of samples of finite sizes are generated by repeated numerical drawings from the distribution. The estimating technique is then applied to the samples, and the resulting sampling distribution of the estimates are compared with the true value of the parameters. The estimating techniques used in our study are the two stage least squares (2SLS) method and the generalised indirect least squares (GILS) method.

With the help of a computer, we have generated 5000 random numbers in the universe which gives the raw data for the exogenous variables ie. the X s and the endogenous variables, ie. the Y's. Normally, distributed errors have also been generated and included in the models under study. Since an increase in the size of the standard deviation of error did not bring about any noticeable change in the estimates, we have abandoned the idea and restricted the study to the normal distribution of errors ie., with mean zero and standard deviation one. Three sample sizes have been taken, viz, 20, 60 and 100, and for each size 30 experiments have been done with numbers randomly picked from the universe to estimate the parameters of the model under consideration.

We have improvised 16 models of sizes (4.8), (5.10), (6.9) and (7.5). The first figure inside the parenthesis denotes the number of equations in the model and the second shows the number of exogenous variables present, the last of which is the constant. Each size has four models - the first with a given endogenous and exogenous variable matrices; the second constructed retaining the endogenous variable matrix of the first model but changing the exogenous variable matrix; the third model formed by changing the endogenous variable matrix but retaining the original exogenous variable matrix; the fourth consisting of the changed endogenous and exogenous variable matrices. The first equaion of each model is an exactly identified equation and the subsequent ones are over-identified equations of different degrees.

The deviations of the 2SLS and GILS estimates of the different models from their true values are shown in tabular form. We shall discuss each model with the help of these tables separately, starting with the smallest one consisting of four equations.

TABLE I

Number of Samples: 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±20	±20 to ±25	±25 to ±30	±45 to ±55	±200 to ±250	Total frequency
0			1	[illegible]		3			1	1			8
1				[illegible]	[illegible]	[illegible]							7
2			1	1	2	[illegible]							5
3	1		1										[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±20	±20 to ±25	±25 to ±30	±45 to ±55	±200 to ±250	Total frequency
0			[illegible]	[illegible]		1	1				1	[illegible]	[illegible]
1				[illegible]	[illegible]	[illegible]	[illegible]						7
2			[illegible]		[illegible]	1		1					[illegible]
3	[illegible]		[illegible]		1								[illegible]

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±20	±20 to ±25	±25 to ±30	±45 to ±55	±200 to ±250	Total frequency
0			[illegible]	[illegible]	1	[illegible]	[illegible]	[illegible]					[illegible]
1				3	[illegible]	[illegible]							[illegible]
2			[illegible]	[illegible]	[illegible]	[illegible]							[illegible]
3	1	[illegible]			1								[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±20	±20 to ±25	±25 to ±30	±45 to ±55	±200 to ±250	Total frequency
0			1	[illegible]	[illegible]	[illegible]							3
1				[illegible]	[illegible]	[illegible]							7
2			1	1	2	1							5
3	1	1			1								[illegible]

(Table I cont'd)

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±20	±20 to ±25	±25 to ±30	±45 to ±55	±200 to ±250	Total frequency
0			[illegible]	4		[illegible]							9
1				3	2	[illegible]							[illegible]
3			1	1	[illegible]	1							5
5	1		1		1								[illegible]

FIML

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±20	±20 to ±25	±25 to ±30	±45 to ±55	±200 to ±250	Total frequency
0			1	[illegible]	1	[illegible]							[illegible]
1				[illegible]	[illegible]	[illegible]							
[illegible]			1	1	[illegible]	1							[illegible]
5	1		1		1								[illegible]

An analysis of table I shows that when 2SLS method is used in estimating the parameters of the equations of the model, we find that two estimates of equation 1 and one estimate of equation 2 have improved with increase in sample size. While one GILS estimate of equation 1 has improved with increase in sample size, the magnitude of deviation of one GILS estimate from its true value has decreased in sample size 60 but has registered a rise in size 100. Though one GILS estimate of equation 2 has improved in sample size 60, it has not recorded any further improvement in sample size 100. While there has been no improvement in the 2SLS estimates of equation 3, one GILS estimate of the same equation has worsened and one has improved in sample size 60. The frequency distribution of sample size 100 however has remained the same as in sample size 60. We also find that though one 2SLS and one GILS estimate of equation 4 has improved in sample size 60, the magnitude of its deviation from the true value has increased in sample size 100. When the parameters of the last equation of the model is estimated by the 2SLS method, we find that one estimate has deviated further away from its true value. In general howe v er, it appears that as the degree of over-identification increases, the tendency of the estimates to converge to their respective true values also increases. While the bias of 6(26%) 2SLS and 13(57%) GILS estimates have decreased with increase in sample size, that of 3(13%) 2SLS and 1(4%) GILS estimates have increased with increase in sample size. 9(39%) 2SLS and 6(26%) GILS estimates in sample size 100 have bias greater than in sample size 60 but smaller than in sample size 20, and 5(21%) 2SLS and 3(13%) GILS estimates in sample size 100 are more biased than in sample size 20 but less biased than in sample size 60. The variances of 15(65%) 2SLS and 17(74%) GILS estimates has decreased with increase in sample size. The remaining 2SLS and GILS estimates have slightly higher variances in the largest sample as compared to sample size 60 but are less than in sample size 20. The t values of 2SLS and GILS estimates are significant.

TABLE II

Number of Samples: 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to +1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±25	±25 to ±30	±60 to ±70	±70 to ±80	±80 to ±90	Total frequency
0		1	2	2	2	1							8
1			2		4	1							7
3				2	1	2							5
5	1	1			1								3

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±25	±25 to ±30	±60 to ±70	±70 to ±80	±80 to ±90	Total frequency
0			4	1	1	2							8
1			1	1	4	1							7
3				2	2	1							5
5	1	1			1								3

Number of Samples: 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±25	±25 to ±30	±60 to ±70	±70 to ±80	±80 to ±90	Total frequency
0			3	2	1	2							[illegible]
1			2		4	1							[illegible]
3				2	3	1							[illegible]
5		1			1								[illegible]

(Table II cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±25	±25 to ±30	±60 to ±70	±70 to ±80	±80 to ±90	Total frequency
0	1			2	1	1	1			1		1	8
1			1	1	4	1							7
3			1	1	1	1					1		5
5		1			1								2

Number of Samples : 100

2SLS

Degree of over-identificat.	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	[illegible] to [illegible]	Total frequency
0			[illegible]	[illegible]	[illegible]	1	1						[illegible]
[illegible]			[illegible]		4	1							
[illegible]				[illegible]	[illegible]	[illegible]		1					[illegible]
5	[illegible]	1			1								[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±25	±25 to ±30	±60 to ±70	±70 to ±80	±80 to ±90	Total frequency
0			2	[illegible]	[illegible]	1		1					8
1		[illegible]	1		4	1							[illegible]
[illegible]				1	1	1			1				[illegible]
5	1	[illegible]			1								

Table II shows that some of the 2SLS estimates of the exactly identified equation have deviated further from its true value with increase in sample size. These deviations have been further accentuated when estimated by the GILS method. There has not been any improvement in the 2SLS estimates of equation 2 with increase in sample size. The magnitude of deviation of one GILS estimate of the equation in the largest sample size however has become smaller. Increase in the size of the sample has not brought about any improvement in the 2SLS and GILS estimates of equation 1. The frequency distribution of the estimates of the last equation shows that both the estimating methods give more or less same results. Except for one estimate of equation 3, the tendency of the estimates is to come closer to their true values as the degree of over-identification increases. The bias of 3(13%) 2SLS and 5(22%) GILS estimates decrease with increase in sample size and that of 10(43%) 2SLS and 13(57%) GILS estimates increase with increase in sample size. Of the remaining estimates, 5(22%) 2SLS and 1(4%) GILS estimates have bias greater in sample size 100 than in sample size 50 but smaller than in sample size 20 and the bias of 2(9%) 2SLS and 4(17%) GILS estimates in sample size 100 are greater than in sample size 20 but smaller than in sample size 50. The variance of 17(74%) 2SLS and 13(57%) GILS estimates have decreased with increase in sample size. The remaining estimates have slightly higher variance in sample size 100 than in sample size 50. Both 2SLS and GILS estimates have significant t values.

TABLE : III

Number of Samples : 20

2SLS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	[illegible]	[illegible]	[illegible]		[illegible]
1	1				4	[illegible]	1	[illegible]
3				[illegible]		2		5
5	1	1			1			[illegible]

GILS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		[illegible]	[illegible]	[illegible]	[illegible]		1	[illegible]
1		1			4	1	1	7
[illegible]				1		1		[illegible]
5		1			1			2

Number of Samples : 60

2SLS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
[illegible]		1	1	[illegible]	[illegible]	[illegible]		[illegible]
[illegible]	[illegible]				1	1	1	[illegible]
[illegible]				[illegible]	[illegible]	1		[illegible]
[illegible]		1			1			[illegible]

(Table III cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	4	[illegible]	1		[illegible]
1				1	4	1	1	7
2				[illegible]		1	1	5
3		[illegible]			[illegible]			[illegible]

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1	[illegible]	[illegible]		1		[illegible]
1					4	[illegible]	1	[illegible]
2				[illegible]		[illegible]		[illegible]
3	1	1			1			3

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			[illegible]	[illegible]	[illegible]	2		[illegible]
1			[illegible]		4	1	1	[illegible]
2				[illegible]	[illegible]	1		[illegible]
3	1		1		1			[illegible]

An analysis of table III shows that while two 2SLS estimates of the exactly identified equation have deviated further from their true values in sample size 60, one of these estimates has improved in sample size 100. The equation when estimated by the GILS method shows that some of the estimates have worsened in the larger sized samples when compared to those in sample size 20. There has not been any improvement in the 2SLS estimates of equation 2. One GILS estimate of the same equation has worsened in sample size 60 but has recorded an improvement in sample size 100. The frequency distribution of the deviation of 2SLS estimates of equation 3 and 4 are same in sample sizes 20 and 100. While one 2SLS estimate of equation 3 has improved in sample size 60, one 2SLS estimate of equation 4 has become equal to its true value. Again, we find that while one GILS estimate has worsened in sample size 60, two GILS estimates have worsened in sample size 100 when compared to those in sample size 20. As the degree of over-identification increases, both 2SLS and GILS estimates come closer to their true values. The bias of 6(26%) 2SLS and 5(22%) GILS estimates have decreased with increase in sample size and that of 10(43%) 2SLS and 13(57%) GILS estimates have increased with increase in sample size. 5(22%) 2SLS and 1(4%) GILS estimates have bias greater in sample size 100 than in sample size 60, but smaller than in sample size 20. 2(9%) 2SLS and 4(17%) GILS estimates in sample size 100 have bias greater than in sample size 20, but smaller than in sample size 60. The variances of 20(87%) 2SLS and 18(78%) GILS estimates have decreased with increase in sample size. The remaining estimates have slightly higher variances in sample size 100 when compared to sample size 60. The t values of both 2SLS and GILS estimates are significant.

TABLE : IV

Number of Samples : 20

2SLS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±15	±30 to ±35	Total freq-uency
0		1	1	2	2	1		1		8
1			1	2	3	1				7
3				1	1				1	3
5	1	1			1					3

GILS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±15	±30 to ±35	Total freq-uency
0			4	1	1	1				[illegible]
1				3	[illegible]	1				
3			1	1	1					3
5	1	1			1					3

Number of Samples : 60

2SLS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±15	±30 to ±35	Total freq uency
0			3	3	[illegible]					8
1			2	1	[illegible]	1				7
3				1	1		[illegible]			3
5	[illegible]	1			1					[illegible]

GILS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±12 to ±15	±30 to ±35	Total freq-uency
0			3	2	2	1				8
1			1	1	4	1				[illegible]
3				3	1	1				5
5		1			1					2

(Table IV cont'd)

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	+3 to +6	±6 to ±9	+9 to +12	±12 to ±15	±30 to ±35	Total frequency
[illegible]			4	[illegible]	1	1				8
[illegible]			2	1	[illegible]	1				[illegible]
[illegible]				4	1					5
5	1	1			1					3

GILS

Degree of over-identificat.	±.001 to +.01	+.01 to +.1	+.1 to +1	+1 to +3	+3 to +6	+6 to +9	+9 to +12	+12 to ±15	±30 to +35	Total frequency
[illegible]			4	1	[illegible]	[illegible]				8
[illegible]			[illegible]	1	[illegible]	[illegible]				
[illegible]				4	1					5
5	1	1			1					3

Table IV reveals that while one 2SLS estimate of exactly identified equation has improved with increase in sample size, one GILS estimate has worsened. One 2SLS estimate and one GILS estimate of equation 2 has improved with increase in sample size. While one GILS estimate of equation 3 has deviated further from its true value in sample size 60, it has improved in sample size 100. One 2SLS estimate of the same equation on the other hand, has come closer to its true value with increase in sample size. The deviation of the 2SLS estimates of equation 4 lie in the same frequency classes in samples sizes 20 and 100. One GILS estimate of equation 5 in sample size 60, however, has become equal to its true value. A scrutiny of the three sample sizes show that both 2SLS and GILS estimates have a tendency to converge to their respective true values with increase in degree of over-identification. The bias of 11(47%) 2SLS and 10(43%) GILS estimates decrease with increase in sample size and that of 8(35%) 2SLS and 8(35%) GILS estimates increase with increase in sample size. 2(9%) 2SLS and 4(18%) GILS estimates in sample size 100 have smaller bias than in sample size 20 and greater bias than in sample size 60. 2(9%) 2SLS and 1(4%) GILS estimates in sample size 100 have bias greater than in sample size 20 but smaller than in size 60. While the t values of both 2SLS and GILS estimates are significant, the variance of 15(65%) 2SLS and 18(78%) GILS estimates decrease with increase in sample size. 8(35%) 2SLS and 5(22%) GILS estimates in sample size 100 have slightly higher variance than in sample size 60 but smaller variance than in sample size 20.

we shall now analyse models consisting of five equations and ten exogenous variables, the last of which is the constant.

TABLE : V

Number of Samples : 10

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	[illegible]	4	[illegible]		10
1			1	[illegible]	5	1		[illegible]
[illegible]				[illegible]	[illegible]	[illegible]		[illegible]
5				[illegible]	1	[illegible]		[illegible]
	[illegible]		[illegible]					[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	[illegible]	4	[illegible]		[illegible]
1			1	2	[illegible]	1		[illegible]
3					4	[illegible]		[illegible]
5				[illegible]		[illegible]		[illegible]
7	1		1	1				[illegible]

Number of Samples : 50

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	[illegible]	4	[illegible]		10
1				[illegible]	5	[illegible]		[illegible]
[illegible]					4	[illegible]		[illegible]
5				2	1	[illegible]		5
7	1	1		1				[illegible]

(Table V cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	3	4	2		10
1		1		2	5	1		9
3					4	3		7
5				2	1	2		5
7	1	1		1				3

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	3	4	2		10
1				2	6	1		9
3				1	3	3		7
5				3		2		5
7	1	2						3

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	5	3	1		10
1			1	1	6	1		9
3					3	4		7
5				3		1	1	5
7	1	1	1					3

The first model of this size represented by Table V reveals that there has not been any improvement in the 2SLS estimates of the exactly identified equation with increase in sample size. Though the GILS estimates of the equation are the same as the 2SLS estimates in sample size 20 and 60, one estimate has improved in the largest sample size. Although the magnitude of deviation of one estimate of equation I derived separately by the 2SLS and GILS methods has lessended in sample size 60, it has registered a rise in sample size 100. 2SLS method has given better estimates for equations 3 and 4 than the GILS method. We also find that one 2SLS estimate of the largest degree over-identified equation has worsened in sample size 60 when compared to that in size 20, but its deviation has been considerably reduced in the largest sample. A GILS estimate of the equation, on the other hand, has shown improvement with increase in sample size. The model is not very conclusive as regards the effect of degree of over-identification on the parameter estimates of the equations of the model. While the bias of 11(32%) 2SLS and 12(35%) GILS estimates have decreased with increase in sample size, that of 15(44%) 2SLS and 11(32%) GILS estimates have increased with increase in sample size. Of the remaining estimates, the bias of 3(9%) 2SLS and 4(12%) GILS estimates in sample size 100 are smaller than in sample size 20 but greater than in sample size 60, and that of 5(15%) 2SLS and 7(21%) GILS estimates are greater than in sample size 20 but smaller than in sample size 60. The variances of 27(79%) 2SLS estimates and 22(68%) GILS estimates have decreased with increase in sample size and that of 9(26%) GILS estimates have increased with increase in sample size. The remaining 2SLS and GILS estimates have slightly higher variances in sample size 100 than in size 60. The t values of both 2SLS and GILS estimates are significant.

TABLE : VI

Number of Samples: 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				4	[illegible]	[illegible]		10
1			1	[illegible]	[illegible]	1		[illegible]
3				[illegible]	[illegible]		[illegible]	[illegible]
5		1		1	[illegible]	[illegible]		[illegible]
7		1	1			1		[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				1	[illegible]			[illegible]
1			1	[illegible]	1	1		[illegible]
3				[illegible]	[illegible]	1		
5				[illegible]	1	2		5
7		1				[illegible]		[illegible]

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				4	[illegible]	[illegible]		10
1			[illegible]	[illegible]	[illegible]	[illegible]		[illegible]
3				[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5				[illegible]	[illegible]	[illegible]		[illegible]
7		1				1		[illegible]

(Table VI cont'd)

GILS

Degree of over-identificat.	+.001 to +.01	+.01 to +.1	+.1 to +1	+1 to +3	+3 to +6	+6 to +9	+9 to +12	Total frequency
0				4	[illegible]	4		10
1			[illegible]	4	1	[illegible]		9
3				[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5					[illegible]	[illegible]		[illegible]
[illegible]	1	[illegible]				[illegible]		[illegible]

Number of Samples : 10

[illegible]

Degree of over-identificat.	+.001 to +.01	+.01 to +.1	+.1 to +1	+1 to +3	+3 to +6	+6 to +9	+9 to +12	Total frequency
0				[illegible]	[illegible]	[illegible]		10
1			[illegible]	[illegible]	[illegible]	1		[illegible]
[illegible]				[illegible]	[illegible]		[illegible]	[illegible]
5			[illegible]	1	[illegible]	[illegible]		[illegible]
[illegible]	[illegible]		[illegible]			[illegible]		[illegible]

GILS

Degree of over-identificat.	+.001 to +.01	+.01 to +.1	+.1 to +1	+1 to +3	+3 to +6	+6 to +9	+9 to +12	Total frequency
0				4	[illegible]	4		10
1			[illegible]	4	[illegible]	1		9
3				[illegible]	[illegible]	1	1	[illegible]
5				[illegible]	1	[illegible]		5
[illegible]	[illegible]	[illegible]				1		[illegible]

Table VI reveals that though the deviations of both 2SLS andGILS estimates of equation 1 lie in the same frequency classes, an increase in sample size has, however, worsened one estimate. The GILS estimates of equation 2 in sample size 20 are better than the 2SLS estimates. Though two parameter estimates computed separately by 2SLS and GILS method, one each of equations 2 and 3 have worsened in sample size 60, their deviation from the true value have lessened in sample size 100. The magnitude of deviations of the 2SLS estimates of equation 4 are smaller than the GILS estimates for all sample sizes. While one 2SLS estimate of equation 5 in sample size 60 has become equal to its true value, it has recorded a slight deviation in sample size 100. Two GILS estimates of the same equation on the other hand, have improved with increase in sample size. Higher degree of over - identification seem to have favourable effect on both 2SLS and GILS estimates. The bias of 10(29%) 2SLS estimates and 9(26%) GILS estimates have decreased with increase in sample size, and that of 2(5%) 2SLS and 7(21%) GILS estimates have increased with increase in sample size. 11(32%) 2SLS and 11(32%) GILS estimates in sample size 100 have greater bias than in sample size 60 but smaller bias than in size 20, and that of another 11(32%) 2SLS estimates and 7(21%) GILS estimates in the sample size 100 are greater than in sample size 20 but smaller than in size 60, while the t values of both 2SLS and GILS estimates are significant, the variances of 33(97%) 2SLS and 25(74%) GILS estimates have decreased with increase in sample size. The remaining estimates of both the methods have greater variances in sample size 100 than in sample size 60 but smaller variances than in sample size 20.

TABLE : VII

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1		2	6	1		10
1			1	3	4	1		9
3				4	3			7
5			1		1	3		5
7	1	1			1			3

G1LS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1		2	6	1		10
1			1	3	4	1		9
3				4	3			7
5			1		[illegible]	[illegible]		5
7	1	1		1				3

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1	1	1	6	1		10
1			1	3	4	1		9
3			1	4	[illegible]			[illegible]
5			1		2	2		5
7			1		1			2

G1LS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1		1	7	1		10
1			1	3	4	1		9
3				3	4			7
5			1		2	2		5
7		1			1			2

(Table VII cont'd)

Number of Samples : 100

ISLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1		2	6	1		10
1			1	3	4	1		9
[illegible]				4	[illegible]			7
5			1		[illegible]	[illegible]		5
[illegible]	1	1			1			3

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
[illegible]		1		[illegible]	[illegible]	[illegible]		[illegible]
1			1	[illegible]	4	1		9
3				5	[illegible]			[illegible]
5			1		[illegible]	[illegible]		[illegible]
[illegible]	1	1			1			3

An analysis of table VII shows that though one 2SLS estimate of equation 1 has improved in sample size 60, it has worsened in sample size 100 to make the frequency distribution same as in sample size 20. The frequency distribution of the deviations of 3SLS estimates of the equation is the same as 2SLS estimates except for one estimate of sample size 60 which has become worse to lie in a higher frequency class. Equations 2 and 4 have the deviations of both 2SLS and 3SLS estimates lying in the same frequency classes for all sample sizes except for one estimate of equation 4 in sample size 20, the deviation of which has increased. Though one 2SLS estimate of equation 3 has improved in sample size 60, it has registered a rise in sample size 100. One the other hand, one 3SLS estimate has worsened in sample size 60, but has improved in size 100. 3SLS estimates of the 5th equation are better than the 2SLS estimates in sample sizes 20 and 60. Except for 2 estimates of equation 4, the estimates have a tendency to come closer to their true value with increase in degree of over-identification. The bias of 8(24%) 2SLS and 8 (24%) 3SLS estimates decrease with increase in sample size, and that of 5 (15%) 2SLS and 3 (9%) 3SLS estimates increase with increase in sample size. 7 (20%) 2SLS estimates and 9 (26%) 3SLS estimates in sample size 100 have greater bias than in sample size 60 but smaller bias than in size 20 and that of 14(41%) 2SLS and 9 (26%) 3SLS estimates in the largest sample are greater than in sample size 20 but smaller than in sample size 60. Of the 34 estimates, 5 (15%) 2SLS and 2 (6%) 3SLS estimates have variances not decreasing with increase in sample size. On the contrary, their variances in the largest sample is slightly higher than in sample size 60. The t values of 2SLS and 3SLS estimates are significant.

TABLE VIII

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to +.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1	2	2	5			10
1			1	3	2	2		8
3				4	3			7
5			1	2	2			5
7	1		1					2

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to +3	±3 to +6	+6 to ±9	±9 to ±12	Total frequency
0		1	2	2	5			10
1			1	4	2	2		9
3				4	3			7
5			1	2	2			5
7	1	1			1			3

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to +.1	±.1 to ±1	±1 to +3	±3 to +6	+6 to +9	+9 to ±12	Total frequency
0			2	2	5			9
1			1	4	2	2		9
3				4	3			7
5			1	1	3			5
7	2	1						3

(Table VIII cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1		2	2	5			10
1	1		1	3	2	2		9
3				4	3			7
5			1	2	2			5
7			1		1			2

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1	2	2	5			10
1	1			4	2	2		9
3				4	3			7
5	1			2	2			5
7	1	1						2

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1	2	2	5			10
1			1	4	2	2		9
3				4	3			7
5			1	2	2			5
7		1	1					2

Table VIII shows that there has been no improvement in either 2SLS or GILS estimates of the exactly identified equation with increase in sample size except for one parameter estimate in sample size 60, the 2SLS estimate of which has become equal to its true value and its GILS estimate has recorded a negligible deviation. Though the magnitude of deviations of two 2SLS estimates, one each of equations 2 and 4 has increased in sample size 60, their magnitude have decreased in sample size 100. The GILS estimates of equation 2 has however, improved in sample size 60 but worsened in sample size 100. There has been no improvement in 2SLS estimates of equation 3 and GILS estimates of equations 3 and 4 with increase in sample size. The 2SLS estimates of equation 5 are better than the GILS estimates as the magnitude of deviations of the former is smaller. As the degree of over-identification increases, the magnitude of deviation of both 2SLS and GILS estimates from their true values decreases. While the bias of 6(18%) 2SLS and 9(27%) GILS estimates decrease with increase in sample size, that of 7(21%) 2SLS estimates and 9(27%) GILS estimates increase with increase in sample size. Sample size 100 has bias of 11(32%) 2SLS and 10(29%) GILS estimates smaller than in sample size 20 but greater than in sample size 60, and that of 10(29%) 2SLS estimates and 6(17%) GILS estimates greater than in sample size 20 but smaller than in sample size 60. Both 2SLS and GILS estimates have significant t values. The variances of 29(85%) 2SLS estimates and 28(82%) GILS estimates decrease with increase in sample size. Sample size 100 has 5(15%) 2SLS and 6(18%) GILS estimates with variances greater than in sample size 60 but smaller than in sample size 20.

We shall now discuss models consisting of six equations and nine exogenous variables, the last variable being the constant

TABLE : IX

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	[illegible]	3		[illegible]
1			1	2	[illegible]	[illegible]		8
2			1	[illegible]	[illegible]	1		[illegible]
3				[illegible]	[illegible]	1		6
4				3	2			5
5				1	2			[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				2	[illegible]	3		8
1			1	[illegible]	4	1		6
2			1	[illegible]	2	1		7
3				3	2	1		6
4				[illegible]	[illegible]			5
5				1	[illegible]			3

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			2	[illegible]	[illegible]		9
1			1	2	3	2		8
2			2	2	2	1		7
3				3	2	1		6
4				3	2			5
5	1		1		2			4

(Table 1x cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	[illegible]	[illegible]		9
1			1	[illegible]	4	1		8
2			[illegible]	[illegible]	[illegible]	[illegible]		[illegible]
3				[illegible]	3	1		5
4				[illegible]	[illegible]			5
5	[illegible]	1			1			[illegible]

Number of Samples : [illegible]

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	[illegible]	[illegible]		[illegible]
1				[illegible]	[illegible]	[illegible]		[illegible]
2			[illegible]	[illegible]	[illegible]	1		[illegible]
3				[illegible]	[illegible]	[illegible]		[illegible]
4				[illegible]	[illegible]			[illegible]
5		[illegible]			[illegible]			[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	[illegible]	[illegible]		[illegible]
1			1	[illegible]	4	1		[illegible]
2			[illegible]	[illegible]	[illegible]	1		[illegible]
3				[illegible]	[illegible]	1		6
4				[illegible]	[illegible]			[illegible]
5		1			[illegible]			[illegible]

Table IX shows that there has been neither any improvement in the 2SLS estimates nor in the GILS estimates of equations 3,4 and 5 with increase in sample size. The frequency distribution of the deviation of estimates from their true values continues to be in the same classes for all sample sizes. while there has been no improvement in the 2SLS estimates of the exactly identified equation with increase in sample size, the deviation of the GILS estimates are in the same classes as the 2SLS estimates except for one estimate of the smallest sample which has become equal to its true value. The deviation of GILS estimates of equation 2 have remained in the same classes for all sample sizes, one 2SLS estimate of the equation has worsened in the largest sample. While one GILS estimate of equation 5 has worsened in the largest sample, one 2SLS estimate has recorded a slight deviation in sample size 60 but has become equal to its true value in sample size 100. An increase in the degree of over-identification seems to cause lesser deviation in the estimates derived by the two methods. While the bias of 17(44%) 2SLS and 10(26%) GILS estimates have decreased with increase in sample size, that of 13(33%) 2SLS and 15(38%) GILS estimates have increased with increase in sample size. The bias of 3(8%) 2SLS and 6(15%) GILS estimates in sample size 100 are greater than in sample size 60 but smaller than in size 20, and that of 6(15%) 2SLS and 8(21%) GILS estimates are greater than in sample size 20 but smaller than in size 60. Both 2SLS and GILS estimates have significant t values. The variances of 37(95%) 2SLS and 37(95%) GILS estimates have decreased with increase in sample size. 2(5%) 2SLS and 1(2.5%) GILS estimate in sample size 100 have greater variance than in sample size 60 but smaller variance than in sample size 20. The variance of 1(2.5%) GILS estimate have increased with increase in sample size.

TABLE : [illegible]

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	[illegible]	[illegible]		[illegible]
1			[illegible]	[illegible]	[illegible]	[illegible]		[illegible]
2	1			4	[illegible]	1		7
3				3				4
4		1		[illegible]		[illegible]		5
5	1							[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	4	2		9
1			[illegible]	2	[illegible]	2		8
2			[illegible]	4	1	1		[illegible]
3				[illegible]	[illegible]	[illegible]		6
4		1		[illegible]		[illegible]	[illegible]	5
5	2		[illegible]		1			4

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	4	[illegible]		[illegible]
1			1	3	[illegible]	2		8
2				4		[illegible]		6
3				[illegible]		4		6
4	1			2	1	1		5
5		1	1		1			3

(Table X cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			2	4	2		9
1			1	2	3	2		8
2			1	4	1	1		7
3				2	1	3		6
4	1			2		2		5
5		1	1		1			3

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				2	4	2		8
1			1	3	2	2		8
2			1	4	1	1		7
3				2		4		6
4	1			2		2		5
5	1		1	1	1			4

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				2	4	2		8
1			1	3	2	2		8
2			1	4	1	1		7
3				2	1	3		6
4	1			2		2		5
5	1		1	1	1			4

An analysis of table [illegible] shows that one 2SLS estimate and one GILS estimate of the exactly identified equation has become equal to its true value in the largest sample. Only one estimate (when computed by both 2SLS and GILS method separately) of equation 2 has improved in sample size 100. While one 2SLS estimate of equation 3 has become equal to its true value in sample size 60, it has recorded a deviation in sample size 100. GILS estimates of equation 3 and 4 have, however, worsened in sample sizes 60 and 100. In equation 5 we find, while one GILS estimate has improved with increase in sample size, one 2SLS estimate has improved in sample size 60 but worsened in sample size 100. The 2SLS estimates of equation 5 in the larger samples have worsened when compared to the smallest sample. Though one GILS estimate of the equation has become equal to its true value in sample size 60, it has deviated further from the true value in the largest sample. The model is not very conclusive as regards the effect of degree of over-identification on the estimates. 10.26% 2SLS and 8.21% GILS estimates have bias decreasing with increase in sample size; [illegible]% 2SLS and [illegible]% GILS estimates have bias increasing with increase in sample size; [illegible]% 2SLS and [illegible]% GILS estimates in sample size 100 have greater bias than in sample size 60 but smaller bias than in sample size 20; and 14.[illegible]% 2SLS and 17.[illegible]% GILS estimates have bias smaller in sample size 100 than in sample size 60 but greater than in sample size 20. While the t values of both 2SLS and GILS estimates are significant, the variances of 37.[illegible]% 2SLS and [illegible]% GILS estimates have decreased with increase in sample size. The remaining estimates have slightly larger variances in sample size 100 when compared to sample size 60.

TABLE : XI

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	1	[illegible]		[illegible]
1			[illegible]	[illegible]	[illegible]	[illegible]		8
2			[illegible]	[illegible]	[illegible]	1		7
3				[illegible]	[illegible]			[illegible]
4				[illegible]	[illegible]			[illegible]
5	1	[illegible]			[illegible]			[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	1	4		9
1			1	[illegible]	[illegible]	[illegible]		9
2			[illegible]	[illegible]	[illegible]	1		[illegible]
3				[illegible]	[illegible]			[illegible]
4				[illegible]	[illegible]			[illegible]
5	[illegible]			[illegible]	[illegible]			4

Number of Samples : 60

3SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				[illegible]	[illegible]	4		[illegible]
1			[illegible]	[illegible]	[illegible]	[illegible]		[illegible]
2			[illegible]	[illegible]	[illegible]	1		7
3				[illegible]	[illegible]			6
4				[illegible]	[illegible]			[illegible]
5	[illegible]	1			[illegible]			[illegible]

(Table XI cont'd)

GILS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				[illegible]	1	4		[illegible]
1			1	[illegible]	[illegible]	[illegible]		[illegible]
2			[illegible]	2	[illegible]	1		[illegible]
[illegible]				[illegible]	[illegible]			6
4				[illegible]	[illegible]			[illegible]
[illegible]			[illegible]		[illegible]			[illegible]

Number of Samples : [illegible]

[illegible]

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
[illegible]	1			[illegible]	1	4		[illegible]
1			[illegible]	[illegible]	[illegible]	[illegible]		[illegible]
[illegible]			[illegible]	[illegible]	[illegible]	1		[illegible]
[illegible]				[illegible]	[illegible]			[illegible]
4				[illegible]	[illegible]			[illegible]
[illegible]			1		[illegible]			[illegible]

[illegible]

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			[illegible]	1	4		[illegible]
1			1	[illegible]	[illegible]	[illegible]		[illegible]
[illegible]			[illegible]	[illegible]	[illegible]	1		[illegible]
[illegible]								[illegible]
4				[illegible]	[illegible]			[illegible]
5				1	[illegible]			[illegible]

Table XI reveals that there has been neither any improvement in the 2SLS estimates nor in the GILS estimates of equation 3,4 and 5 with increase in sample size. The frequency distribution of the deviations of both 2SLS and GILS estimates of the exactly identified equation is the same in sample sizes 20 and 100. One 2SLS estimate and one GILS estimate of the equation has however, become equal to its true value in sample size 60. While one 2SLS estimate has deviated further from its true value in the larger samples, one GILS estimate of the same equation has improved in sample size 60 but the size of its deviation has increased in sample size 100. One 2SLS estimate of equation 6 has become equal to its true value and there has been a reduction in the size of deviation of another estimate of the same equation in the largest sample size. The parameters of the same equation when estimated by the GILS method, has made two estimates, one each in sample sizes 60 and 100 equal to their respective true values, and has worsened one estimate in sample size 100 when compared to that in sample size 60. Higher the degree of over-identification, better the 2SLS and GILS estimates. The bias of 14(36%) 2SLS and 9(23%) GILS estimates have decreased with increase in sample size and that of 7(18%) 2SLS and 10(26%) GILS estimates have increased with increase in sample size. 9(23%) 2SLS and 11(28%) GILS estimates in sample size 100 have greater bias than in sample size 60 but smaller bias than in sample size 20, and 9(23%) 2SLS and 9(23%) GILS estimates in sample size 100 have bias smaller than in sample size 60 but greater than in sample size 20. Except for 5(13%) 2SLS and 4(10%) GILS estimates, which have slightly higher variances in sample size 100 than in sample size 60, the variances of the remaining estimates have decreased with increase in sample size. The t values of both 2SLS and GILS estimates are significant.

TABLE : XII

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0				3	2	3		8
1				4	1	3		8
2				4	2	1		7
3				2	2	2		6
4		1		2		2		5
5	1		1	1	1			4

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±5 to ±9	±9 to ±12	Total frequency
0				3	2	3		8
1				4	1	3		8
2				4	2	1		7
3				2	2	2		6
4		1		2		1	1	5
5			1	1	1			3

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	+.01 to +.1	±.1 to ±1	±1 to ±3	±3 to +6	±6 to ±9	±9 to ±12	Total frequency
0				3	2	3		8
1				4	1	3		8
2				5	1	1		7
3				2	2	2		6
4	1			2		2		5
5	1	1	1		1			4

(Table XII cont'd)

[illegible]

Degree of over-identificat.	+.001 to ±.01	+.01 to ±.1	-.1 to ±[illegible]	+1 to +[illegible]	+[illegible] to +5	+6 to +9	+9 to +12	Total frequency
0				[illegible]	[illegible]	[illegible]		8
1				4	[illegible]	[illegible]		8
[illegible]				[illegible]	[illegible]	1		
[illegible]				[illegible]	[illegible]	[illegible]		6
4	1			[illegible]		[illegible]	[illegible]	[illegible]
[illegible]	[illegible]		[illegible]		[illegible]			4

Number of Samples : [illegible]

[illegible]

Degree of over-identificat.	+.001 to +.01	+.01 to -.1	+.1 to +1	+1 to +[illegible]	+[illegible] to +6	-6 to +[illegible]	+ to -12	Total frequency
0				[illegible]	[illegible]	[illegible]		8
[illegible]				4	[illegible]	[illegible]		8
[illegible]				4	[illegible]	[illegible]		[illegible]
				[illegible]	[illegible]	[illegible]		[illegible]
[illegible]		[illegible]		[illegible]		[illegible]		[illegible]
[illegible]	[illegible]		[illegible]	[illegible]	1			4

[illegible]

Degree of over-identificat.	±.001 to +.01	+.01 to -.1	+.1 to +1	+1 to +[illegible]	+[illegible] to +6	+6 to ±[illegible]	+[illegible] to -[illegible]	Total frequency
[illegible]				[illegible]	[illegible]	[illegible]		[illegible]
1				4	1	[illegible]		[illegible]
[illegible]				4	[illegible]	[illegible]		
				[illegible]	[illegible]	[illegible]		6
4		1		[illegible]				[illegible]
[illegible]	[illegible]		[illegible]		1			4

Table XII shows that the frequency distribution of the deviation of estimates from their true values, of equations 1,2 and 4 is the same for all sample sizes. We also find that there has been no improvement in the GILS estimates of equations 1,2,3 and 4 with increase in sample size. Though three 2SLS estimates, 1 each of equations 2,3 and 5, have improved in sample size 50, the size of their deviation from the true values have increased in sample size 100. Again, we find that the magnitude of deviation of the GILS estimate of equation 5 has decreased in the largest sample. The GILS estimate of the last equation of the model has worsened with increase in sample size. The 2SLS estimate however, has improved in sample size 50, but worsened in sample size 100. That the estimates get better as the degree of over-identification increases can be seen clearly when the model is estimated by the 2SLS method. The bias of 10 (2[illegible]%) 2SLS and 5 (2[illegible]%) GILS estimates have decreased with increase in sample size and that of 6 (15%) 2SLS and 5 (22%) GILS estimates have increased with increase in sample size. 10 (26%) 2SLS and 10 (2[illegible]%) GILS estimates in sample size 100 have greater bias than in sample size 50 but smaller bias than in sample size 20, and the bias of 13 (33%) 2SLS and 11 (28%) GILS estimates in sample size 100 are greater than in sample size 20 but smaller than in sample size 50. Both 2SLS and GILS estimates have significant t values. [illegible] ([illegible]%) 2SLS and [illegible] ([illegible]%) GILS estimates have variances decreasing with increase in sample size. 2 (5%) 2SLS and 3 ([illegible]%) GILS estimates in sample size 100 have variances slightly greater than in sample size 50 but smaller than in sample size 20.

Models consisting of seven equations and eight exogenous variables, the last variable being the constant will now be discussed.

TABLE XIII

Number of Samples: 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±25	±30 to ±50	±55 to ±90	±110 to ±180	±250 to ±800	±1200 to ±6000	Total frequency
0			1	2	1	1							5
1		1		1	1	1		1					[illegible]
2				1	1						1	1	[illegible]
3			1	1				1					3
4				1			1						[illegible]
5					1	1							[illegible]
6		1			1								2

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±25	±30 to ±50	±55 to ±90	±110 to ±180	±250 to ±800	±1200 to ±6000	Total frequency
0	1		1	1	1	1							[illegible]
1				1	3	1					1		[illegible]
2				2	1		1	1			1		5
3				4			1						5
4				2	1		1						4
5				1	1	1							[illegible]
6			1		1								[illegible]

Number of Samples : 60

3SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±25	±30 to ±50	±55 to ±90	±110 to ±180	±250 to ±800	±1200 to ±6000	Total frequency
0			1	1	2	3	1						8
1			2	1	2	1				1			7
2				3	1						1	1	6
3				2	1		1		1				5
4				2	1		1						4
5					2	1							3
6		1			1								2

'Table XIII cont'd'

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±25	±30 to ±50	±55 to ±90	±110 to ±180	±250 to ±800	±1200 to ±6000	Total frequency
0	2	1			2	3							8
1			1	1	1		3	1					7
2				2	1			1		1		1	6
3				4			1						5
4				1	1		1						4
5					2	1							3
6			1		1								2

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±25	±30 to ±50	±55 to ±90	±110 to ±180	±250 to ±800	±1200 to ±6000	Total frequency
0			1		2	3	1	1					8
1			1	3		2				1			7
2				2	1	1				2			6
3				1	1	1	1					1	5
4				1	1		1						4
5					2	1							3
6		1			1								2

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±25	±30 to ±50	±55 to ±90	±110 to ±180	±250 to ±800	±1200 to ±6000	Total frequency
0	1	1		1	2	1	1						8
1			1	1	1			1	2			1	7
2				1	1							3	6
3				3			2						5
4				2	1		1						4
5					2	1							3
6			1		1								2

The first model of this size represented by Table XIII shows that the magnitude of deviations of some 2SLS as well as some GILS estimates from their respective true values is very large. They lie in two, three and four digit classes. The frequency distribution of the deviations of 2SLS estimates from their true values of the last three equations has remained the same for all sample sizes. While the frequency distribution of the deviation of GILS estimates of equation 5 and 7 has remained the same, one estimate of equation 6 has deviated further away from its true value in sample sizes 60 and 100. We also find that two 2SLS and two GILS estimates, one each of equations 1 and 4, have worsened with increase in sample size. while two 2SLS estimates of equation 3 have improved in sample size 100, three GILS estimates of the equation have worsened in the largest sample. Both 2SLS and GILS estimates have a tendency to converge to their respective true values with increase in the degree of over-identification. 12(34%) 2SLS and 3(9%) GILS estimates have bias decreasing with increase in sample size: 20(57%) 2SLS and 22(63%) GILS estimates have bias increasing with increase in sample size; 3(9%) 2SLS and 4(11%) GILS estimates have bias greater in sample size 100 than in sample size 60 but smaller than in sample size 20: and 6(17%) GILS estimates in sample size 100 have greater bias than in sample size 20 but smaller bias than in sample size 60. Both 2SLS and GILS estimates have significant t values. The variances of the estimates are in general very high (of three or four digits) and that of only 15(43%) 2SLS and 9(26%) GILS estimates have decreased with increase in sample size.

TABLE XIV

Number of Samples: 20

2SLS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±17	±20 to ±40	±70 to ±80	±120 to ±160	Total freq-uency
0				2	4	1		1				8
1				2	3	1						7
2		2			1	2	1					6
3				3		2						5
4			1	2	1							4
5			1		1	1						3
6		1		1								2

[illegible]

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±17	±20 to ±40	±70 to ±80	±120 to ±150	Total freq-uency
0				3	4	1						8
1				3	1	3						7
2			2		1	3						6
3				2	1	1						5
4			1	2	1							4
5		1			1	1						3
6			1	1								2

Number of Samples : 50

2SLS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±17	±20 to ±40	±70 to ±80	±120 to ±150	Total freq-uency
0				4	2	2						8
1				3	4	1						7
2		1	1	1	1	2						6
3				2	1	2						5
4			1	2	1							4
5				1	1	1						3
6		1		1								2

(Table XIV cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±17	±20 to ±40	±70 to ±80	±120 to ±160	Total frequency
0				5	2	1						8
1				2	4	1						7
2			2	1	1	2						6
3				2	1	1	1					5
4			1	1	1							4
5					1	1						3
6		1		1								2

Number of Samples : 100

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±17	±20 to ±40	±70 to ±80	±120 to ±160	Total frequency
0				5	2	1						8
1				3	2	1	1					7
2			1		1	3						6
3				2	2	1						5
4			1	2	1							4
5				1	1	1						3
6		1		1								2

3SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	±15 to ±17	±20 to ±40	±70 to ±80	±120 to ±160	Total frequency
0				3	2	1			1		1	8
1				1	1	1				1		7
2			1	1	2	1					1	6
3			1	2	1				1			5
4			1	2	1							4
5				1	1	1						3
6			1	1								2

Table XIV reveals that while some of the 2SLS estimates of the exactly identified equations have improved with increase in sample size, some of the GILS estimates of the same equation have worsened with increase in sample size. The deviations of the 2SLS estimates of equations 5 and 7 and GILS estimates of equation 5 continue to be in the same class intervals irrespective of an increase in sample size. While the frequency distribution of the GILS estimates of equation 7 is the same for sample sizes 20 and 100, one GILS estimate of the equation has improved in sample size 60. Again we find that one 2SLS estimate of equation 6 and one GILS estimate of the same equation has worsened when compared to the smallest sample size. One 2SLS estimate of equation 2 has worsened in sample 100. Its deviation from the true value has been further accentuated when estimatd by the GILS method. One parameter of equation 3 estimated by the 2SLS and the GILS method has recorded a smaller deviation in sample size 60 but has registered a rise in its deviation in sample size 100. Though one 2SLS estimate of equation 4 has improved with increase in sample size, one GILS estimate of the same equation has deviated further away from its true value with increase in sample size. The higher the degree of over-identification the better the estimates of both the methods. The bias of 1·3%) 2SLS and 2·6%) GILS estimates have decreased with increase in sample size and that of 10·29%) 2SLS and 14·40%) GILS estimates have increased with increase in sample size. 12·34%) 2SLS and 11·31%) GILS estimates in sample size 100 have greater bias than in sample size 60 but smaller bias than in sample size 20, and 12·34%) 2SLS and 8·33%) GILS estimates in sample size 100 have smaller bias than in sample size 60 but greater bias than in sample size 20. The t values of both 2SLS and GILS estimates are significant. The variances of 34·97% 2SLS and 24(69%) GILS estimates have decreased with increase in sample size. 8·33%) GILS estimates have variance increasing with increase in sample size. The variances of 1(3%) 2SLS and 3(8% GILS estimates in samcle size 100 are slightly greater than in sample size 60 but smaller than in size 20.

TABLE XV

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			1	[illegible]	[illegible]	2		8
1				[illegible]	1	1		7
2				[illegible]	2		1	6
3				[illegible]	[illegible]		1	5
4				1	[illegible]			4
5			[illegible]		[illegible]			[illegible]
6								[illegible]

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		[illegible]	1	1	[illegible]	2		8
1				4	[illegible]	1		[illegible]
2			[illegible]	[illegible]	[illegible]		1	6
3				[illegible]	1	[illegible]	1	[illegible]
4				1	[illegible]			4
5			[illegible]	1	[illegible]			[illegible]
6			[illegible]					[illegible]

Number of Samples : 50

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		1	[illegible]	1	[illegible]	[illegible]		8
1				4	2	1		7
2			[illegible]	[illegible]	1		1	6
3				[illegible]	[illegible]	[illegible]	[illegible]	5
4				1	[illegible]			4
5			1		[illegible]			[illegible]
6		1	1					2

(Table XV cont'd)

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0		2		1	3	2		8
1				3	3	1		7
2			1	3	1		1	6
3				3		1	1	5
4				1	3			4
5			1		2			3
6		1	1					2

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0			2	1	3	2		8
1				4	2	1		7
2				4	1		1	6
3			2	1	1		1	5
4				1	3			4
5			1		2			3
6			2					2

GILS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	2			1	4	1		8
1				3	3	1		7
2			1	3	1		1	6
3				2	1	1	1	5
4				1	3			4
5			1		2			3
6			2					2

An analysis of Table XV shows that there has been no improvement in the 2SLS estimates of equation 5 and 6 and GILS estimates of equation 3 and 5. While some of the GILS estimates of the exactly identified equation have improved with increase in sample size, one 2SLS estimate of sample size 100 of the same equation has become worse when compared to that in sample size 60. Two 2SLS estimates of equation 4 have become better in sample size 100 than in the smaller sized samples, but the frequency distribution of the GILS estimates of the same equation in the sample size 100 is the same as in sample size 20 indicating no improvement in the estimates of the larger sample. Though one parameter estimated separately by the GILS and 2SLS methods, has improved in sample size 60, it has registered an increase in the magnitude of deviation in sample size 100. The frequency distribution of the parameter estimates of the highest degree over-identified equation is the same when computed by the two different methods. In this case also we find that the estimates get better as the degree of over-identification increases. The bias of 13(37%) 2SLS and 11(31%) GILS estimates have decreased with increase in sample size and that of 11(31%) 2SLS and 9(26%) GILS estimates have increased with increase in sample size. In sample size 100, there are 4(12%) 2SLS and 8(23%) GILS estimates with bias greater than in sample size 60 but smaller than in sample size 20, and 7(21%) 2SLS and 7(20%) GILS estimates with bias greater than in sample size 20 but smaller than in sample size 60. The variances of 27(77%) 2SLS and 34(97%) GILS estimates have decreased with increase in sample size. 8(23%) 2SLS estimates and 1(3%) GILS estimate in sample size 100 have variances greater than in sample size 60 but smaller than in sample size 20. Both 2SLS and GILS estimates have significant t values.

TABLE : XVI

Number of Samples : 20

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			4		2		7
1				1	6			7
2				3	2	1		6
3				[illegible]	3	1		[illegible]
4				3	1			[illegible]
5		1			1			[illegible]
6			[illegible]		[illegible]			[illegible]

3SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	[illegible]			4		[illegible]		[illegible]
1				1	6			[illegible]
2				[illegible]	[illegible]	1		6
3				[illegible]	[illegible]	[illegible]		5
4				[illegible]	[illegible]			4
5		[illegible]	[illegible]		[illegible]			[illegible]
6		[illegible]			[illegible]			[illegible]

Number of Samples : 60

2SLS

Degree of over-identificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total frequency
0	1			4		[illegible]		7
1				1	6			7
2				[illegible]	2	1		6
3				1	3	1		[illegible]
4				3	1			4
5			[illegible]		1			3
6			1		1			2

(Table XVI cont'd)

GILS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total freq-uency
0		2		4		2		8
1				1	6			7
2				3	2	1		6
3				1	3	1		5
4				3	1			4
5		1	1		1			3
6		1			1			2

Number of Samples : 100

3SLS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total freq-uency
0	2			4		2		8
1				1	6			7
2				3	2	1		6
3				1	3	1		5
4				3	1			4
5		1	1		1			3
6		1			1			2

GILS

Degree of over-iden tificat.	±.001 to ±.01	±.01 to ±.1	±.1 to ±1	±1 to ±3	±3 to ±6	±6 to ±9	±9 to ±12	Total freq-uency
0		2		4		2		8
1				1	6			7
2				3	2	1		6
3			1	1	3			5
4				3	1			4
5		1	1		1			3
6			1		1			2

Table XVI shows that there has not been any improvement in the 2SLS estimates of equation 2, 3 and 5, and GILS estimates of equations 2, 3 and 6, One 2SLS estimate of the exactly identified equation in sample sizes 20 and 60 has become equal to its true value but the estimate has deviated slightly in sample size 100. Two GILS estimates of equation 1 have worsened in sample sizes 60 and 100. While the frequency distribution of the deviation of 2SLS estimates from their respective true values is the same for the equations 4, 6 and 7, one estimate of each of these three equations has worsened in sample size 60. Again we find that two GILS estimates of equation 4 and one GILS estimate of equation 7 in the largest sample size has worsened when compared to the sample sizes 20 and 60. One GILS estimate of equation 4 has improved when compared to that of the smallest sample size. The estimates computed by both the methods have a tendency to come closer to their respective true values as the degree of over-identification increases. 13(37%) 2SLS and 10(29%) GILS estimates have bias decreasing with increase in sample size; 7(20%) 2SLS and 14(40%) GILS estimates have bias increasing with increase in sample size; 8(23%) 2SLS and 5(14%) GILS estimates in sample size 100 have greater bias than in sample size 60 but smaller bias than in sample size 20; and 7(20%) 2SLS and 6(17%) GILS estimates in sample size 100 have greater bias than in sample size 20 but smaller bias than in sample size 60. While the t values of both 2SLS and GILS estimates are significant, the variances of 32(94%) 2SLS and 29(83%) GILS estimates have decreased with increase in sample size. The variances of remaining 3(6%) 2SLS and 6(17%) GILS estimates in sample size 100 are greater than in sample size 60 but smaller than in sample size 20.

CHAPTER - I V

SUMMARIZATION OF RESULTS

SUMMARIZATION OF RESULTS

In the previous chapter we have discussed at length the behaviour of estimates of exactly identified and over-identified equations of different degrees computed by the 2SLS and GILS methods in three different sample sizes. The study has however been restricted to correct specification of the model. The behaviour of estimates in the mis-specification case has been left out from the purview of our study

In the course of our experiments, we have found that some of the parameter estimates in the different samples have become negative when their true values are positive. This may be due to the fact that the mean of the errors in the randomly picked samples is not zero as has been assumed in the universe. If the sample error deviates further away from its universe mean which is zero, there is a possibility of a change in sign of the estimates.

For the purpose of summarization of findings we shall consider the four different-sized models separately.

A survey of the models consisting of four equations and eight exogenous variables, the last of which is the constant, reveals that the parameter estimates of the equations comprising the first model derived by the 2SLS and GILS methods separately not only come closer to their respective true values with increase in the size of the samples, but also with increase in degree of over-identification. That the estimates get better with increase in the degree of over-identification can be seen in all three sample sizes. The second model which has been constructed by changing the exogenous variable matrix and retaining the endogenous variable matrix of the first model shows that although some of the estimates computed separately by the 2SLS and GILS methods have deviated further away from their respective true values with increase in the degree of over-identification, the magnitude of deviations of some GILS estimates are larger than the 2SLS estimates. The third model constructed using the exogenous variable matrix of the first model but changing the endogenous variable matrix shows, when computed by the 2SLS method, that the frequency distribution of the deviation of estimates of the over-identified equations from their respective true values is the same when the number of samples taken are 20 and 100. This indicates that there has been no improvement in the estimates with increase in the size of the sample. The estimates however, have a tendency to converge to their true values with increase in the degree of over-identification. When the GILS method is used we find that some of the estimates of higher degree over-identified equations in the largest sample size have worsened compared to those in the smallest sample size. The fourth model, constructed using the changed exogenous and endogenous variable matrices of the second and third models shows that the 2SLS as well as GILS estimates come closer to their respective true values with increase in the degree of over-identification and in the increase of the number of samples taken. On an average we find that while 34% 2SLS estimates have bias decreasing with increase in sample size, 35% GILS estimates have bias following the same pattern. 30% of both 2SLS and GILS estimates have bias increasing with increase in the size of the samples. Of the remaining estimates we find that while an average of 21% 2SLS estimates have greater bias in sample size 100 when compared to those in sample size 60 but smaller bias than those in sample size 20, only 15% GILS estimates have behaved in the similar manner. Again we find that the bias of 15% 2SLS and 20% GILS estimtes in sample size 100 are greater than those in sample size 20 but smaller than those in sample size 60. On an average, 74% 2SLS and 63% GILS estimates have variances decreasing with increase in sample size. The t values of both 2SLS and GILS estimates are significant.

The first model, of the group of models comprising of five equations and ten exogenous variables, shows that the frequency distribution of deviation of 2SLS and GILS estimates from their true values are same for equations 1 and 2 in sample size 20, and for all the equations of the model when the number of samples taken is 60. The 2SLS estimates of the last three equations are however, better than the GILS estimates when the number of samples taken are 20 and 100. Some of the GILS estimates of equations 1 and 2 in the largest sample are better than the 2SLS estimates. We also find that two 2SLS estimates - one each of equations 4 and 5 and four GILS estimates - two estimates of equation 1 and one each of equations 4 and 5 have improved with increase in sample size. The highest degree over-identified equation has estimates (computed by both 2SLS and GILS methods) with very small deviations. The second model reveals that both 2SLS and GILS estimates have the same frequency distribution for the exactly identified equation. While some of the estimates of the over-identified equations, derived by the two different methods, have worsened in sample size 60, they have recorded lesser deviations in sample size 100. Again we find that, while one 2SLS estimate has become equal to its true value in sample size 60, one GILS estimate has improved with increase in sample size. In general, the estimates get better as the degree of over-identification increases. The third model shows that the frequency distribution of 2SLS and GILS estimates is same except for equation 4 in sample size 20 and equation 3 in sample size 100, where one GILS estimate has recorded lesser deviation than the 2SLS estimate. In sample size 60, we find that while 2SLS method has given better estimates for equations 1 and 3, GILS method has given better estimates for the highest degree over-identified equation. Except for two estimates of equation 4, both 2SLS and GILS estimates have a tendency to come closer to their true values with increase in degree of over-identification. The last model of this size reveals that the frequency distribution of deviation of 2SLS and GILS estimates is the same for all sample sizes. As for the over-identified equations, we find that GILS gives better estimates when the number of samples taken is 60 and 2SLS gives better estimates when the number of samples taken is 100. That the estimates come closer to their respective true values with increase in degree of over-identification can be seen best in this model. On an average we find tht 26% 2SLS and 28% GILS estimates have bias decreasing with increase in sample size; 21% 2SLS and 22% GILS estimates have bias increasing with increase in sample size; 24% 2SLS and 25% GILS estimates have bias greater in sample size 100 than in sample size 60 but smaller than in sample size 20; and 29% 2SLS and 25% GILS estimates in sample size 100 have greater bias than in sample size 20 but smaller bias than in sample size 60. While both 2SLS and GILS estimates have significant t values, on an average 88% 2SLS and 79% GILS estimates have variances decreasing with increase in the size of samples.

The first of the four six equation models shows that there has been no improvement in the 2SLS and the GILS estimates with increase in sample size except for the last equation of the model. While two 2SLS estimates of the highest degree over-identified equation have improved with increase in sample size, two GILS estimates have come closer to their true values in sample size 60 but recorded increase in deviation in sample size 100. The frequency distribution of both 2SLS and GILS estimates of equations 3,4 and 5 is the same. Both 2SLS and GILS estimates have a tendency to come closer to their true values with increase in the degree of over-identification. The second model shows that the magnitude of deviation of both 2SLS and GILS estimates of the various equations is the same when the sample size is 100 except of equation 4. We also find that some estimates have deviated further away from their true values even when the degree of over-identification increases. While some estimates have improved with increase in sample size, some have worsened. The third model reveals that while the 2SLS estimates of the last equation are better than the GILS estimates when the number of samples taken are 20 and 100, one GILS estimate of the equation has become equal to its true value in sample size 60. Except for equation 3 in sample sizes 20 and 60 and equation 6 in all sample sizes, the frequency distribution of both 2SLS and GILS estimates are the same. We find that both 2SLS and GILS estimates have a tendency to converge to their true values with increase in the degree of over-identification. The fourth model shows that the magnitude of deviation of some GILS estimates of over-identified equations is smaller than that of the 2SLS estimates. But in the case of exactly identified equation, 2SLS method has given better estimates. While some of the estimates computed by the two different methods have worsened with increase in sample size, the estimates seem to deviate less with increase in the degree of over - identification in all sample sizes. On an average, 33% 2SLS and 23% GILS estimates have bias decreasing with increase in sample size; 18% 2SLS and 23% GILS estimtes have bias increasing with increase in sample size; 20% 2SLS and 25% GILS estimates have greater bias in sample size 100 than in sample size 60 but smaller bias than in sample size 20 and 27% 2SLS and 26% GILS estimates have greater bias in sample size 100 than in sample size 20 but smaller bias than in sample size 60. While the estimates have significant t values, 88% 2SLS and 79% GILS estimates have variances decreasing with increase in sample size.

Of the four seven equation models, the first one reveals that the magnitude of deviation of some 2SLS and GILS estimates of equations 1 and 2 are very high. The 2SLS estimates of these equations are however better than the GILS estimates. While some of the GILS estimates of equations 4,5 and 6 have smaller deviations than the 2SLS estimates, the estimates of the last equation are better when computed by the 2SLS method. We find that both 2SLS and GILS estimates have come closer to their true values with increase in degree of over identification. The second model reveals that while a GILS estimate of equation 3 is better than the 2SLS estimate in sample size 20, it has deviated further away from the true value in sample sizes 60 and 100 making the 2SLS estimate better. When the number of samples taken is 100 we find that 2SLS gives better estimates for all equations except for two over-identifies equations of degree four and five. The frequency distribution of deviation of estimates of these two equations is the same when estimated by the 2SLS and the GILS methods. Although some of the 2SLS estimates have improved with increase in sample size, some GILS estimates have deviated further away from their true values in the largest sample. Both 2SLS and GILS estimates have a tendency to come closer to their true values as the degree of over-identification increases. The third model shows that the frequency distribution of deviation of both 2SLS and GILS estimates of equations 5 and 7 is the same in the different sample sizes. While some GILS estimates of the exactly identified equation are better than the 2SLS estimates, some 2SLS estimates of equation 2 have lesser deviations than that of the GILS estimates. One estimate of the last equation computed by the 2SLS and GILS method has improved in sample size 60 but recorded an increase in deviation in sample size 100. Though some 2SLS and GILS estimates of the model have worsened with increase in sample size, in general the estimates computed by both the methods have a tendency to converge to their true values as the degree of over-identification increases. The last model of this group shows that in the case of the exactly identified equation, 2SLS method gives better estimates than the GILS method. The frequency distribution of 2SLS and GILS estimates of equations 1 and 2 is the same for all sample sizes. While one 2SLS estimate of the last equation has improved with increase in sample size, one GILS estimate has improved in sample size 60 but worsened in sample size 100. We find that some of the estimates, derived by the two different methods, have improved with increase in the degree of over-identification and sample size. On an average, 28% 2SLS and 19% GILS estimates have bias decreasing with increase in sample size; 34% 2SLS and 42% GILS estimates have bias increasing with increase in sample size; 19% 2SLS and 20% GILS estimates have greater bias in sample size 100 than in sample size 60 but smaller bias than in sample size 20; and 19% of both 2SLS and GILS estimates have greater bias in sample size 100 than in sample size 20 but smaller bias than in sample size 60. The t values of both 2SLS and GILS estimates of the models are significant. When an average of variances of estimates of all four models is taken it is found that 89% 2SLS and 69% GILS estimates have variances decreasing with increase in sample size.

CHAPTER - V

CONCLUSION

CONCLUSION

The research has been conducted to find out the effect of various degrees of over-identification on the performance of estimators. For this purpose we have made use of simultaneous equation models.

In sections 1.1 and 1.2 of Chapter I, we have discussed in detail the concept of simultaneous equation models and its necessity. We have also shown that the ordinary least squares method when applied to such models fails to give consistent estimates because of the presence of endogenous variables among the explanatory variables. The OLS method, however, will give consistent estimates if the model is transformed into a reduced form model because in the reduced form equations the explanatory variables are represented by the pre-determined variables of the model. This suggests that we can derive the estimates of the structural coefficients via the reduced form coefficients. Thus, the problem is one of identification.

In sections 1.3, 1.4 & 1.5 of the same chapter, we have delineated the identification problem which is the problem of finding a unique solution for the structural coefficients from the reduced form coefficients. If the structural coefficients cannot be estimated from the reduced form coefficients, the equation is unidentified or under-identified. A structural equation, if identified, is either exactly identified or over-identified.

An equation is over-identified if the number of exogenous variables which enter into the system but not into the equation is greater than the number of endogenous variables which enter into the equation. It is the number of endogenous variable present and the number of exogenous variables absent that will determine the degree of over-identification. The larger the number of exogenous variables included than the number of endogenous variables included, the greater will be the degree of over-identification. This concept of over-identification and its degree have been discussed in detail in Chapter II.

Models of four different sizes have been taken. Size indicates the number of equations and variables constituting the model. Each of the four sizes has four models. Monte Carlo experiments have been conducted with these sixteen models using the estimating techniques two stage least squares (2SLS) and generalised indirect least squares (GILS). In Chapter III we have analysed the results of each model separately. The criteria on the basis of which inferences about the behaviour of the 2SLS and the GILS estimates have been drawn are : the magnitude of deviation of estimates from their true values, bias and variance.

In Chapter IV we have summarised the performance of both 2SLS and GILS estimates of exactly identified and over-identified equations of different degrees, when the four different-sized models are taken separately.

In the course of our study, we have found that some of the parameter estimates computed by the two different methods, in the three different sample sizes have become negative when their true values are positive. This may be due to the fact that the mean of the errors in the randomly picked samples have deviated further away from the universe mean which has been assumed to be zero.

It has also been found that the GILS method when used in estimating the parameters of over-identified equations gives estimates which, if not better, are as competent as those derived by the 2SLS method.

On the basis of the criteria used in analysing the results, we have found that only some 2SLS as well as GILS estimates have improved with increase in sample sizes. Some of the estimates, however, have recorded smaller deviations from their true values as the number of samples taken is increased from 20 to 60, but their magnitude of deviation has increased as the number of samples taken is increased further to 100. We also find that some other estimates have deviated further away from their true values as the number of samples taken is increased from 20 to 60, but there has been a reduction in the magnitude of deviation of these estimates as the number of samples taken is further increased to 100. The magnitude of deviation of some estimates from their respective true values have also been found to increase with increase in sample sizes. This behaviour of both 2SLS and GILS estimates may be due to the fact that when we generate random numbers the values of the majority are around the mean and few values are quite far away from the mean. So when random samples are taken from these randomly generated numbers, there is a possibility that in some samples, a few values quite away from the mean value have been selected. Both 2SLS and GILS estimates have in general a tendency to converge to their respective true values as the degree of over-identification increases.

Due to the limitations imposed by the computer, the study had to be restricted to the use of only fifteen variables. If we could use more variables, the proposition that both 2SLS and GILS estimates tend to converge to their true values with increase in the degree of over-identification would have possibly been more clear.

It is to be noted that fool-proof generalisation cannot be done on the basis of Monte Carlo method. This is because the method does not give the true probability distribution of the estimates but rather an empirical estimate of that distribution.

The research has been conducted using only the correct model, that is, using apriori restrictions that correctly describe the artificial structure that has been used. The behaviour of the estimates in the mis-specified cases has been left out from the purview of the study.

When a model is mis-specified, some endogenous variables may be irrelevant or some relevant exogenous variables may be missing or it may be a case of both missing exogenous and irrelevant endogenous variables. Mis-specification may also arise due to the missing of relevant equations which in turn may lead to the missing of relevant endogenous variables. We know that the presence of irrelevant endogenous variables under certain conditions do not affect the estimation of the parameters provided they are uncorrelated. But missing endogenous variables have an adverse effect on the estimates. There is, therefore, scope for further research to find out the extent of effect of degree of over-identification on the performance of estimators in mis-specified models.

APPENDIX

P R O G R A M

```
D(10,15),UCOR(10,10),YS(100,1)
          DIMENSION ZS(100,15),Y(100,10),Z(100,15),ZDZ(15,15)
          DIMENSION ZDY(15,10),PIE(15,10),YHAT(100,10),YER(100,10)
          DIMENSION PA(10),PH(10),P(10 [illegible]LSTY(10),LSTI [illegible]
          DIMENSION E(100,10),[illegible](15),E[illegible](100,1[illegible]),ZZ(15)[illegible]10)
          DIMENSION A(15,15),B(15,10),PX(15)[illegible]PPX(15,10)
          CHARACTER DESC(10)*[illegible]
          CHARACTER *6 FI,FY,FO,FC,CF
          DOUBLE PRECISION IX,IY,YPL
C --- PARAMETERS ------------------------------------------------------------
          NV=100
          NHV=10
          N[illegible]V=15
          NH=NHV
          IP[illegible]=[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible]
[illegible] = REDUCED FORM COEFF
[illegible] = STRUCTURAL  COEFF
[illegible]
[illegible]
C ---------------------------------------------------------------------------
          [illegible] TWO STAGE LEAST SQUARES [illegible] PROGRAM [illegible]*****
          [illegible]
          [illegible]
          [illegible] EST[illegible]
          CF='[illegible]
          WRITE(*,*) 'OUTPUT DEVICE ? IF PRINTER SEE THAT IT IS READ[illegible]
          READ(*,*) NC
          WRITE(*,*) 'READ NO. OF EQ [illegible] IN SYSTEM [illegible] OF EXOG VARIABLES
          READ(*,*) [illegible]NH,N
          [illegible]
          IPY=IPL[illegible]
          [illegible]I=IPL*[illegible]
          WRITE(*,*) 'ARE YORD AND ORD ALREADY STORED ? [illegible]
          READ(*,*) IRR
          IF(IRR.[illegible]) THEN
          WRITE(*,*) 'FILE TO STORE YORD AND CORD MATRICES ?'
          READ(*,*) FO
          WRITE(*,*) NOW [illegible] FOR ROW [illegible] ALL ELEMENTS
          OPEN(1[illegible],FILE=FO,STATUS='NEW')
          DO 55 I=1,NH
          READ(*,*) (YORD(I,J),J=1,NH),(CORD(I,J),J=1,N)
          WRITE(11,*) (YORD(I,J),J=1,NH),(CORD(I,J),J=1,N)
       55 CONTINUE
          CLOSE(11)
          ELSE
```

```
      WRITE(*,*)' FILE IN WHICH [illegible] ARE STORED ?'
      READ(*,*) FC
      OPEN(11,FILE=FC,[illegible])
      DO 40 I=1,[illegible]
      READ(11,*) [illegible]
   40 CONTINUE
      CLOSE(11)
      ENDIF
      WRITE(*,*) [illegible]
    [illegible] ELSE [illegible]
      READ(*,*) [illegible]
      [illegible]
-----------------------------------------------------------------------
      [illegible] STRUCTURAL COEFF [illegible]
      [illegible] X
      WRITE [illegible] D E'
      READ(*,*) [illegible] CHOICE
      DO 44 [illegible] NH
      WRITE(NZ,[illegible]) [illegible]
   44 CONTINUE
   45 FORMAT [illegible]
      OPEN(11,FILE=[illegible],STATUS='NEW',[illegible],RECL=[illegible])
      [illegible]
      DO 33 [illegible] NH
      DO 33 [illegible]=1,NH
      [illegible]
      [illegible]
      [illegible]
      [illegible]
   [illegible] CONTINUE
      WRITE [illegible] COEFF [illegible]
      DO [illegible]
      [illegible]
   [illegible] CONTINUE
      [illegible]
      [illegible]
      CALL [illegible]
      [illegible] (B)
   [illegible] CONTINUE
      [illegible]
      [illegible]
      DO [illegible] NK
      [illegible]
      [illegible]
      [illegible]
   [illegible] CONTINUE
   [illegible] FORMAT [illegible]
      WRITE [illegible] MATRIX
      [illegible]
      WRITE(NZ,[illegible]) [illegible] J=1,NK
      DO 42 I=1,NH
      IRCL=IRCL+1
      WRITE(11,REC=IRCL) [illegible]
   42 CONTINUE
      DO 22 I=1,NH
```

```
      DO 22 J=1,NA
      IRCL=IRCL+1
      WRITE(11,REC=IRCL) [illegible]
 22   CONTINUE
      WRITE(*,*) ' COEFFICIENT MATRIX'
      DO 42 I=1,NA
      WRITE(*,[illegible]) [illegible]
      DO 42 J=1,NA
      IRCL=IRCL+1
      WRITE(12,REC=IRCL) [illegible]
 42   CONTINUE
      WRITE(12,REC=1) [illegible]
      CLOSE(12)
      WRITE(*,*) ' HOW MANY OBSERVATIONS TO GENERATE IN THE UNIVERSE'
      READ(*,*) NN
      CALL DATA(NN,NK,NA,Y,[illegible],[illegible],[illegible],IFI,NC)
      STOP [illegible] GENERATION OVER [illegible] PROGRAM [illegible] EXPERIMENTS
      [illegible]
      WRITE(*,*) [illegible] EFFICIENCY [illegible]
      WRITE(*,*) ' NO. OF OBSERVATIONS IN THE UNIVERSE [illegible]'
      READ(*,*) NN
      WRITE(*,*) [illegible] SEED [illegible] SAMPLE FOR [illegible]
      WRITE(*,*) [illegible] SAMPLES FROM THE UNIVERSE
      READ(*,*) IX
      WRITE(*,*) [illegible] OF ENDOGENOUS EXOGENOUS VARIABLES [illegible]
      READ(*,*) NA,NK
      [illegible]=[illegible]
      [illegible]=[illegible]
      WRITE(*,*) [illegible] THE [illegible] TO INCLUDE [illegible]
      READ(*,*) [illegible]
      WRITE(*,*) [illegible] INCLUDE(1),EXCLUDE( [illegible]
      READ(*,*) [illegible],NC)
      WRITE(*,*) [illegible] EXPERIMENT [illegible]
      READ(*,*) NE)
      WRITE(*,*) [illegible] PRINT [illegible]
      READ(*,*) [illegible]
      WRITE(*,*) [illegible]
      READ(*,*) [illegible]
      [illegible]
      DO 170 [illegible]
      [illegible]=NC
      WRITE(*,*) [illegible] EXPERIMENTS WITH [illegible] OBSERVATIONS = [illegible]
      OPEN [illegible],STATUS='OLD',ACCESS='DIRECT',RECL=IFX
      OPEN(8,FILE=[illegible],STATUS='OLD',ACCESS='DIRECT',RECL=IFZ
      OPEN(11,FILE=[illegible],STATUS='NEW',ACCESS='DIRECT',RECL=IFL
      OPEN(14,FILE=[illegible],STATUS='NEW',ACCESS='DIRECT',RECL=IFL
      IRCL=1
      DO 110 I=1,NE
      WRITE(*,*)'EXPERIMENT NO. =',I
      DO 88 J=1,N
 177  CALL RANDU(IX,IY,YFL)
      JY=ABS(YFL*NN)
      IF((JY.EQ.0).OR.(JY.GT.NN)) GOTO 177
      READ(7,REC=JY)(Y(J),J=1,NA)
```

```
   89 CONTINUE
      READ(8,REC=JY)(ZZ(J),J=1,NHU)
   90 CONTINUE
      IND=0
      DO 1 J=1,NHU
      IF(LSTY(J).NE.0) THEN
      IND=IND+1
      Y(I,IND)=YY(J)
      ENDIF
    1 CONTINUE
      IND=0
      DO 2 J=1,NHU
      IF(LSTZ(J).NE.0) THEN
      IND=IND+1
      Z(I,IND)=ZZ(J)
      ENDIF
    2 CONTINUE
C     WRITE(*,*)(YY(J),J=1,NHU),(ZZ(J),J=1,[illegible]U)
C     WRITE(*,*)(Y(I,J),J=1,NH),(Z(I,J),J=1,N[illegible])
   69 CONTINUE
      IF(IFF.EQ.1) CALL [illegible](N,NH,[illegible],I,I
   37 FORMAT(I5,7F10.3)
      CALL [illegible]SELS(N,NH,NH,[illegible],[illegible],ZO[illegible],FIE,[illegible],[illegible],FA,F[illegible],F,[illegible],
     [illegible]ZORI,LOCF,YEF[illegible],[illegible],TT,YS,ZS,YH,IFCL,RK,FRK
  100 CONTINUE
      CLOSE(7)
      CLOSE(8)
      WRITE(11,REC=1) FLOAT(IRCL)
      WRITE(14,REC=1) FLOAT(IRCL)
      CLOSE(14)
      CLOSE(11)
      WRITE(NO,*) * * * * * * * * * * * * * * * * * * * * * * * *
      WRITE(NO,*)           RESULTS FOR TWO-STAGE LEAST SQUARES
      WRITE(NO,*) NO. OF OBSERVATIONS ',N,  EXPERIMENTS  ',NEX
      CALL TEST(FIE,TT,YH,YHAT,YORD,ZORD,LSTY,LSTZ,[illegible],YER,NH,NH,NEX,
     [illegible]IRCL,DESC,NHU,NHU,1)
      DO 2000 I=1,5
      WRITE(NO,*)
 2000 CONTINUE
      WRITE(NO,*) RESULTS FOR GENERALIZED   (DIRECT LEAST SQUARES)
      WRITE(NO,*)  NO. OF OBSERVATIONS ',N,  EXPERIMENTS ',NEX
      CALL TEST(FIE,TT,[illegible],YHAT,YORD,ZORD,LSTY,LSTZ,[illegible],YER,N[illegible],NH,NEX,
     [illegible],IRCL,DESC,NHU,NHU,[illegible]
   [illegible] CONTINUE
      END
C     --- --- ------ TEST BLOCK -------------------------------------
      SUBROUTINE TEST(FIE,FI,A,B,YORD,ZORD,LSTY,LSTZ,PP,SP,NH,NK,
     1NEX,FC,IRL,DESC,NHU,NHU,JCH)
      COMMON NHV,NHV,NV,NHX,NO
      DIMENSION FIE(NHV,NHV),FI(NHV,NHV),A(NV,NHV),B(NV,NHV)
      DIMENSION LSTY(NHV),LSTZ(NHV),PP(NV,NHV),SP(NV,NHV)
      DIMENSION YORD(NHV,NHV),ZORD(NHV,NHV),LY(10),LZ(15)
      CHARACTER *6 FC,DESC(10) *20,DESCR *60
      NHK=NH+NK
      VRN=SQRT(FLOAT(NEX))
```

```
      VR=FLOAT(NEX)/FLOAT(NEX-1)
      DO 10 J=1,NH
      DO 10 JJ=1,NH
      FIE(J,JJ)=0.0
      EF(J,JJ)=0.0
   10 CONTINUE
      DO 11 J=1,NH
      DO 11 JJ=1,NH
      [illegible]
      [illegible]
   11 CONTINUE
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      LI(I)=[illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      CALL [illegible]
      DO [illegible]
      DO [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
      [illegible]
    [illegible] CONTINUE
      DO [illegible] I=1,NH[illegible]
      LSTZ(I)=[illegible]
      IF LSTZ(J).NE.0 THEN
      IND=IND+1
```

```
      Z(I,IND)=ZZ(J)
      ENDIF
    1 CONTIN[illegible]      A(I,JJ)=(A(I,J)[illegible] I=1,NHL
      IF(LSTY(J).NE.0) THEN
      IND=IND+1
      Y(I,IND)=Y(I,J)
      ENDIF
    1 CONTINUE
         [illegible]
      READ [illegible]     Y[illegible]
   5F CONTINUE
      READ [illegible]
    [illegible]
      [illegible]
      [illegible]
      WRITE [illegible]
      DO 16 [illegible]
   16 WRITE [illegible]
      DO [illegible]
      DO [illegible]
      [illegible]
   1[illegible] CONTINUE
      WRITE [illegible]
      DO [illegible]
  1[illegible] WRITE [illegible]
      DO 11[illegible] JJ=1,NH
      DO [illegible] J=1,NH[illegible]
      PF(J,JJ)=PF(J,JJ)-E(J,JJ)
  11[illegible] CONTINUE
      WRITE [illegible] DELUF DESC[illegible]
      DO 14 JJ=1,NH
   14 WRITE [illegible]
      WRITE [illegible] DESC[illegible]
      DO [illegible] JJ=1,NH
   [illegible] WRITE [illegible] J=1,NH[illegible]
      WRITE [illegible] DESC[illegible]
      DO [illegible] J=1,NH
   1[illegible] WRITE [illegible] J=1,NH
      WRITE [illegible] DESC[illegible]
      DO 1[illegible] J=1,[illegible]
   1[illegible] WRITE [illegible] (PF(J,JJ),J=1,NHR
      DO [illegible] J=1,NH
      DO [illegible] JJ=1,NH
      TV=0.0
      IF(FIE(J,JJ).NE.0.0) THEN
      TV=PF(J,JJ)/SQRT(FIE(J,JJ))
      ENDIF
      FIE(J,JJ)=TV*[illegible]
   [illegible] CONTINUE
      DO 1[illegible] J=1,NHR
      DO 1[illegible] JJ=1,NH
      TV=0.0
      IF(A(J,JJ).NE.0.0) THEN
      TV=PF(J,JJ)/SQRT(A(J,JJ))
      ENDIF
```

```
      A(J,JJ)=TV*VRN
   24 CONTINUE
      WRITE(NO,*) DESCR(DESC,7,JCH,NO)
      DO 38 JJ=1,NH
   38 WRITE(NO,100)(FIE(J,JJ),J=1,NH)
      WRITE(NO,*) DESCR(DESC,5,JCH,NO)
      DO 39 JJ=1,NH
   39 WRITE(NO,100)(A(J,JJ),J=1,NH)
      DO 73 I=1,NHV
      LSTY(I)=LY(I)
   73 CONTINUE
      DO 74 I=1,NH
      [illegible]
   74 CONTINUE
  100 FORMAT(1X,6E12.5)
      WRITE(NO,*)'================================================================'
      RETURN
      END
      FUNCTION DESCR(DESC,I,J,NO)
      CHARACTER DESCR *60,DESC(10) *20
      DESCR=DESC(INT(I/2,+.5))//DESC(7-I,7-I [illegible] DESC 7-I
      WRITE(NO,*)'----------------------------------------------------------------'
      RETURN
      END
      SUBROUTINE FIC(NHU,NHU,NH,NV,FIE,A,IRCL,IV,LY,YORD,ZORD,LZ,LL)
      COMMON NHV,NHV,IV,NHZ,NO
      DIMENSION FIE(NH,NH),A(NV,NHV),YORD(NHV,NHV),ZORD(NH,NV)
      DIMENSION LY(NHV),LZ(NHV)
      NH1=NH+1
      [illegible]
      DO 18 JJ=1,NH
      DO 17 J=1,NH1
      A(J,JJ)=0
   17 CONTINUE
      DO 18 I=1,NH
      FIE(I,JJ)=0
   18 CONTINUE
      II=0
      DO 2 I=1,NHU
          IF(LY(I).NE.0) THEN
          II=II+1
          JJ=0
          DO 22 J=1,NHU
              IF(LZ(J).NE.0) THEN
              JJ=JJ+1
              IRCL=IRCL+1
              READ(IV,REC=IRCL) FIE(JJ,II)
              ELSE
              IRCL=IRCL+1
              ENDIF
   22     CONTINUE
          ELSE
          IRCL=IRCL+NHU
          ENDIF
    2 CONTINUE
```

```
      II=0
      DO 3 I=1,NHU
          IF(LY(I).NE.0) THEN
          II=II+1
          JJ=0
          DO 33 J=1,NHU
              IF(LY(J).NE.0) THEN
              JJ=JJ+1
                  IF(L.EQ.0) THEN
                      IF(YCFD(II,JJ).GT.0.0) THEN
                      IRCL=IRCL+1
                      READ(IV,REC=IRCL) A(JJ,II)
                      ENDIF
                  ELSE
                  IRCL=IRCL+1
                  READ(IV,REC=IRCL) A(JJ,II)
                  ENDIF
              ELSE
                  IF(L.NE.0) THEN
                  IRCL=IRCL+1
                  ENDIF
              ENDIF
 33           CONTINUE
      ELSE
          IF(L.NE.0) THEN
          IRCL=IRCL+NHU
          ENDIF
      ENDIF
  3   CONTINUE
      JW=JJ
      II=0
      DO 4 I=1,NHU
          IF(LY(I).NE.0) THEN
          II=II+1
          JJ=0
          DO 44 J=1,NHU
              IF(LZ(J).NE.0) THEN
              JJ=JJ+1
                  IF(L.EQ.0) THEN
                      IF(ZOFD(II,JJ).GT.0.0) THEN
                      IRCL=IRCL+1
                      READ(IV,REC=IRCL) A(JJ+JW,II)
                      ENDIF
                  ELSE
                  IRCL=IRCL+1
                  READ(IV,REC=IRCL) A(JJ+JW,II)
                  ENDIF
              ELSE
                  IF(L.NE.0) THEN
                  IRCL=IRCL+1
                  ENDIF
              ENDIF
 44       CONTINUE
          ELSE
              IF(L.NE.0) THEN
```

```
            IRCL=IRCL+NKU
            ENDIF
         ENDIF
    4 CONTINUE
C     WRITE(NO,*) IRCL,' IS THE ENDING RECORD NUMBER'
C     DO 5 I=1,NH
C     WRITE(NO,100)(PIE(J,I),J=1,NH)
C   5 CONTINUE
C     DO 6 I=1,NH
C     WRITE(NO,100)(A(J,I),J=1,NH+NH)
C   6 CONTINUE
  100 FORMAT(1X,7E11.4)
      RETURN
      END
C---- -- ----- DATA GENERATION BY SIMULATION BLOCK -----------------
      SUBROUTINE DATA(NM,NH,[illegible]FIE,A,E,EN,FY,FZ,IFY,IFZ,NC)
      COMMON [illegible],NHALF,NC
      DIMENSION [illegible]
      DIMENSION A(NHK,NH),FIE(NH,NH),VE(10,10),VEN(10,10),E(10)
      DIMENSION ZE(15,15),ZP(15),AVZ(10),AVEN(10),AVE(15)
      CHARACTER *6 FZ,FY
      DOUBLE PRECISION IX,IY,YFL
      WRITE(*,*)' FEED SEED TO GENERATE RANDOM NUMBERS,MEAN AND SD'
      READ(*,*) IX,AM,ASD
      I=1
      NH1=NH-1
      DO 7 J=1,NH
      DO 80 JJ=1,NH
      VE(J,JJ)=0.0
   80 VEN(J,JJ)=0.0
      AVE(J)=0.0
      AVEN(J)=0.0
    7 CONTINUE
      DO 9 J=1,NH1
      AVZ(J)=0.0
      DO 9 JJ=1,NH1
      ZE(J,JJ)=0.0
    9 CONTINUE
      OPEN(6,FILE=FZ,STATUS='NEW',RECL=IRL,ACCESS='DIRECT')
      OPEN(7,FILE=FY,STATUS='NEW',RECL=[illegible],ACCESS='DIRECT')
      DO 100 II=1,NN
      DO 31 J=1,NH
      CALL RANDD(IX,[illegible],YFL)
      Z(I,J)=[illegible](YFL*NC)
   31 CONTINUE
      DO 8 J=1,NH1
      AVZ(J)=AVZ(J)+Z(I,J)
      DO 8 JJ=J,NH1
      ZE(J,JJ)=ZE(J,JJ)+Z(I,J)*Z(I,JJ)
    8 CONTINUE
      Z(I,NH1)=1.0
      DO 36 J=1,NH
      Y(I,J)=0
      DO 37 K=1,NH
      Y(I,J)=Y(I,J)+Z(I,K)*PIE(K,J)
```

```
 37 CONTINUE
 36 CONTINUE
    DO 70 J=1,NH
    E(I,J)=RANDN(IX,AM,ASD)
    AVE(J)=AVE(J)+E(I,J)
 70 CONTINUE
C   WRITE(*,*)(E(I,J),J=1,NH)
    DO 51 J=1,NH
    EN(I,J)=0.0
    DO 51 K=1,NH
    EN(I,J)=EN(I,J)-E(I,K)*A(K,J)
    AVEN(J)=AVEN(J)+EN(I,J)
 51 CONTINUE
C   WRITE(*,*)(EN(I,J),J=1,[illegible])
    DO 73 J=1,NH
    DO 73 JJ=1,NH
    [illegible]E(J,JJ)=VE(J,JJ)+E(I,J)[illegible]
    VEN(J,JJ)=VEN(J,JJ)+EN[illegible]
 73 CONTINUE
    DO 52 J=1,NH
    Y(I,J)=[illegible](I,J)+EN[illegible]
 52 CONTINUE
 3[illegible] CONTINUE
    IY=II
    [illegible]
    IZ=II
    WRITE(1,REC=IZ)(Y(I,J),J=1,NH)
 100 CONTINUE
    CLOSE(7)
    CLOSE(8)
    DO 77 J=1,NH
    [illegible]
    [illegible]
    [illegible]
    [illegible]
 77 CONTINUE
    DO 4 J=1,NH
    AVE(J)=AVE(J)/NN
    AVEN(J)=AVEN(J)/NN
    [illegible]
    [illegible]
    [illegible]
    [illegible]
    VEN(JJ,J)=VEN(J,JJ)
  4 CONTINUE
    DO 10 J=1,NH
    DO 10 JJ=1,NH
    [illegible]E(J,JJ)=[illegible]E(J,JJ)/NN-A[illegible]/SQRT[illegible]
 10 CONTINUE
    DO 5 J=1,NH
    DO 5 JJ=1,NH
    VE(J,JJ)=(VE(J,JJ)/NN-AVE(J)*AVE(JJ))/(SQRT(VE(J,J)*VE(JJ,JJ)))
    VEN(J,JJ)=(VEN(J,JJ)/NN-AVEN(J)*AVEN(JJ))/SQRT(VEN(J,J)*VEN(JJ,JJ))
  5 CONTINUE
    WRITE(NO,*)'ERROR CORRELATION MATRIX FOR STRUCTURAL EQNS'
```

```
      DO 74 J=1,NH
      WRITE(NO,5)(VE(J,JJ),JJ=1,NH)
   74 CONTINUE
      WRITE(NO,*)'ERROR CORRELATION MATRIX FOR REDUCED FORM EQNS'
      DO 75 J=1,NH
      WRITE(NO,5)(VEN(J,JJ),JJ=1,NH)
   75 CONTINUE
      WRITE(NO,*)' CORRELATION MATRIX AMONG INDEPENDENT VAR'
      DO 11 J=1,NH
      WRITE(NO,5)(DD(J,JJ),JJ=1,NH)
   11 CONTINUE
      WRITE(NO,*)'AVERAGES OF EFFECT OF EQNS AND INDEP VARS'
      WRITE(NO,*)(AV(J),J=1,NH)
      WRITE(NO,*)(AVEN(J),J=1,NH)
      WRITE(NO,*)(AV(J),J=1,N)
      WRITE(NO,*)'STD DEVIATIONS OF STRUCT & RF EQNS AND INDEP VAR'
      WRITE(NO,*)(SQRT(DD(J,J)),J=1,NH)
      WRITE(NO,*)(SQRT(VE(J,J)),J=1,NH)
      WRITE(NO,*)(SQRT(A[illegible](J)),J=1,NH)
C     TABLE [illegible]
    5 FORMAT(1X,10F7.4)
      RETURN
      END
C ----- GENERALIZED [illegible] LEAST SQUARES [illegible] -----
      SUBROUTINE TAI(IE,CORI,YORD,NH,NV,DD,[illegible],R
      COMMON NH,NHV,NV,[illegible],NO
      DIMENSION ID(10),CD(15),TI(15,10),F(15,10),FF(15,10),R[illegible]
      DIMENSION CD(NH,NH),COY(NH,NHV),YORD(NHV,NHV)
      DIMENSION FLE(NH,NH),CURD(NHV,NHV),RS(15),RH(10),R(NHV)
      CHARACTER STR*7
      STR='***'
        DO 20 I=1,NH
        ID(I)=0
        FL(I)=0.0
   20   CONTINUE
        DO 21 J=1,NH
        [illegible]
        [illegible]
    1   CONTINUE
      NHF=0
      DO 22 J=1,NH
      IF(YORD(I,J).NE.0.0) NHF=NHF+1
   22 CONTINUE
      IN=0
      JN=NHF
        NHF=NHF-1
      DO 1 J=1,NH
      IF(YORD(I,J).EQ.0.0) THEN
        JN=JN+1
        ID(JN)=J
        ELSE
        IF(YORD(I,J).EQ.-1.0) THEN
          ID(NHF)=J
          ELSE
          IN=IN+1
```

```
          ID(IN)=J
        ENDIF
      ENDIF
    2 CONTINUE
C     IF(NHP.LT.NH) WRITE(NO,*)(ID(JJ),JJ=1,NHPP),STR,ID(NHP),STR,
C    1(ID(JJ),JJ=NHP+1,NH)
C     IF(NHP.EQ.NH) WRITE(NO,*)(ID(JJ),JJ=1,NHPP),STR,ID(NHP)
      NHQ=NH-NHP
      DO 3 J=1,NH
      DO 3 K=1,NH
      F(K,J)=FIE(K,ID(J))
    3 CONTINUE
      NKF=0
      DO 92 I=1,NK
      IF(IDRI(I,3).NE.1.0) NKF=NKF+1
   92 CONTINUE
      IN=0
      JN=NKF
      DO 4 I=1,NK
      IF(IDRI(I,3).EQ.1.0) THEN
        JN=JN+1
        ID(JN)=I
      ELSE
        IN=IN+1
        ID(IN)=I
      ENDIF
    4 CONTINUE
C     IF(NKF.LT.NK) WRITE(NO,*)(ID(JJ),JJ=1,NKF),STR,
C    1 (ID(JJ),JJ=NKF+1,NK)
C     IF(NKF.EQ.NK) WRITE(NO,*)(ID(JJ),JJ=1,NKF)
C     WRITE(NO,*)' REARRANGED FIE MATRIX'
      NKQ=NK-NKF
      DO 5 J=1,NH
      DO 55 K=1,NK
   55 FK(K,J)=F(ID(K),J)
C     WRITE(NO,56)(F(K,J),K=1,NK)
    5 CONTINUE
   56 FORMAT(1X,7E10.3)
C     WRITE(NO,*)' FF MATRIX AND FV VECTOR'
      DO 6 J=1,NHQ
      DO 6 K=1,NHPP
      FF(J,K)=F(J+NHP,K)
      FV(J)=F(J+NHP,NHP)
C     WRITE(NO,56)(FF(J,K),K=1,NHPP),FV(J)
    6 CONTINUE
      DO 8 J=1,NHPP
      DO 8 K=J,NHPP
      ZDZ(J,K)=0.0
      DO 9 L=1,NHQ
    9 ZDZ(J,K)=ZDZ(J,K)+FF(L,J)*FF(L,K)
      ZDZ(K,J)=ZDZ(J,K)
    8 CONTINUE
      DO 10 J=1,NHPP
      ZDY(J,1)=0.0
      DO 10 L=1,NHQ
```

```
10 ZDY(J,1)=ZDY(J,1)+PP(L,J)*FV(L)
   CALL MINV(ZDZ,NHPP)
   DO 11 J=1,NHPP
   PA(J)=0.0
   DO 11 [illegible]=1,NHPP
   [illegible]
11 CONTINUE
   PA(NHP[illegible]
[illegible]
   WRITE(MC,[illegible]
12 FORMAT(1[illegible]ERSE METHOD IN EQN[illegible]
13 FORMAT([illegible]
   DO 14 [illegible]
   [illegible]
   DO 14 J=1,[illegible]
   [illegible]
14 CONTINUE
   CALL [illegible]
   CALL [illegible]
   WRITE([illegible]
   WRITE([illegible]
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   ENDIF
   CONTINUE
   DO [illegible] J=1,[illegible]
   IF([illegible].NE.0) THEN
   [illegible]
   [illegible]
   ENDIF
   CONTINUE
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   IF [illegible]))) THEN
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   [illegible]
   ENDIF
 1 CONTINUE
   RETURN
   END
C ------------- DATA PRINTING BLOCK ----------------------------------
   SUBROUTINE [illegible]
   COMMON [illegible]
   DIMENSION Y(NX,NH[illegible]),Z([illegible])
   WRITE(*,*) 'DEVICE TO PRINT Y AND Z'
```

```
      READ(*,*) NR
      DO 1 I=1,N
      WRITE(NR,2) (Y(1,J),J=1,NH),(Z(I,J),J=1,NH)
    1 CONTINUE
    2 FORMAT 1[illegible]F3.2)
      PAUSE 'TO PROCEED PRESS ENTER KEY'
      RETURN
      END
C ------------- RANDOM NUMBER GENERATION BLOCK -------------------
      FUNCTION RANDN(IX,A,B)
      DOUBLE PRECISION T1,PI
      DOUBLE PRECISION IX,IY,YFL
C     THE CONSTANT IS 2*PI = 2*DATAN(1.0) = 2*3.14159265358979
      CALL RANDO(IX,IY,YFL)
      T1=DABS(YFL)
      CALL RANDO(IY,IX,YFL)
      PI=DABS(YFL)
      RANDN=DSQRT(-2.0*DLOG(PI))*DCOS(6.283185307179586*T1)
      RANDN=RANDN*B+A
      RETURN
      END
      SUBROUTINE RANDO(IX,IY,YFL)
      DOUBLE PRECISION IX,IY,YFL
      COMMON [illegible]
      CALL RANDOM(IX,IY,YFL)
      IX=IY
      RETURN
      END
C --------------- TWO-STAGE LEAST SQUARES BLOCK -------------------
      SUBROUTINE TWOSLS(N,NH,NHV,Y,Z,ZDI,ZDY,PIE,YHAT,YER,PA,[illegible]
     1[illegible],UCOF,YER2,INDEX,[illegible]
      DIMENSION [illegible]
      DIMENSION [illegible]
      DIMENSION [illegible]
      DIMENSION ZS(NV,NHV),[illegible]
      DIMENSION ZDY(NHV,NHV),PIE(NHV,NH),YHAT(NV,NHV),YER[illegible]
      DIMENSION RA(NHV),RH(NHV),[illegible]
C     WRITE(NO,*)' FIRST STAGE OF TWO-STAGE STARTS'
      CALL OLS(N,NH,NHV,[illegible])
      DO 1 J=1,NH
      DO 1 I=1,NH
      [illegible]
      WRITE [illegible]
      WRITE [illegible]
    1 CONTINUE
      DO 22 I=1,NH
      CALL FAB(PIE,ZORD,YORD,NV,NH,ZDY,ZDI,I,R)
      DO 22 J=1,NH
      RHX(J,I)=RX(J)
   22 CONTINUE
      DO 31 I=1,NH
      INDEX(I)=NH
   31 CONTINUE
      CALL ERMAT(N,NH,YER,UCOF,INDEX)
C     WRITE(NO,*)' SECOND STAGE OF TWO-STAGE STARTS'
```

```
      NU=1
      IEQN=0
      DO 20 I=1,NH
      L=I
      IEQN=IEQN+1
      DO 18 J=1,NH
      IF(YORD(I,J).LT.0) L=J
   18 CONTINUE
      DO 11 J=1,N
   11 Y1(J,1)=Y(J,L)
      CALL ORDER(I,NH,NP,IN,ZORD,YORD,N,IS,YHAT,Z)
      INDEX(I)=IN
C     WRITE(1,17) IEQN
      CALL OLS(N,IN,[illegible])
      DO 2 JJ=1,IN
      IFCL=IFCL+1
      WRITE(11,REC=IFCL) [illegible]
      WRITE(14,REC=IFCL) [illegible](JJ,I)
    2 CONTINUE
      DO 29 J=1,N
      YER(J,I)=YER(J,N)
   29 CONTINUE
   20 CONTINUE
      CALL ERMAT(N,NH,YER,UCOR,INDEX)
   17 FORMAT('  REGRESSION EQUATION FOR 2SLS FOR ',I2,'TH STRUCTURAL
     1EQUATION')
      RETURN
      END
      SUBROUTINE ORDER(I,NH,NP,IN,ZORD,YORD,N,IS,YHAT,Z)
      COMMON NHV,NHV,NV,NH,NC
      DIMENSION ZORD(NH,NV),YORD(NH,NH),Z(NV,[illegible]),YHAT([illegible],NH)
      DIMENSION IS(NV,NH)
      IN=0
      DO 12 J=1,NH
      IF(YORD(I,J).GT.0) THEN
      IN=IN+1
      DO 13 K=1,N
   13 IS(K,IN)=YHAT(K,J)
      ENDIF
   12 CONTINUE
      DO 14 J=1,NV
      IF(ZORD(I,J).NE.[illegible]) THEN
      IN=IN+1
      DO 15 K=1,N
   15 IS(K,IN)=Z(K,J)
      ENDIF
   14 CONTINUE
      RETURN
      END
      SUBROUTINE ERMAT(N,NH,YER,UCOR,INDEX)
      COMMON NHV,NHV,NV,NH,NC
      DIMENSION YER(NV,NHV),UCOR(NHV,NHV),INDEX(NHV),SD(20)
      DO 1 I=1,NHV
      DO 1 K=1,NHV
      UCOR(I,K)=0.0
```

```
      IF(I.EQ.K) UCOR(I,K)=1.0
    1 CONTINUE
      DO 22 I=1,NH
      DO 22 J=I,NH
      UCOR(I,J)=0.0
      DO 23 K=1,N
   23 UCOR(I,J)=UCOR(I,J)+YER(K,I)*YER(K,J)
      UCOR(I,J)=UCOR(I,J)/FLOAT(N)
      UCOR(J,I)=UCOR(I,J)
      IF(I.EQ.J) SD(I)=UCOR(I,J)
   22 CONTINUE
C     WRITE(*,*)'ERROR CORRELATION MATRIX'
C     DO 24 I=1,NH
C     WRITE(*,25)(UCOR(I,J),J=1,NH)
C  24 CONTINUE
C     WRITE [illegible] AT [illegible]
C     DO 27 [illegible]
C     WRITE(*,25) [illegible]
C  27 CONTINUE
   25 FORMAT [illegible]
      RETURN
      END
C [illegible]
      [illegible]
      [illegible]
      [illegible]
      DIMENSION [illegible] ZDZ(NH,NH) [illegible]
      DIMENSION YHAT [illegible] ,RA(NH) [illegible]
   [illegible] FORMAT [illegible]
      DO 2 [illegible]
      DO 2 [illegible]
      [illegible]
      [illegible]
    2 CONTINUE
      [illegible]
      [illegible]
      [illegible]
      DO 3 [illegible]
    3 [illegible] =ZDZ [illegible]
      [illegible]
    [illegible] CONTINUE
   [illegible] FORMAT [illegible]
C     DO 26 [illegible]
C     WRITE(*,35) [illegible]
C  26 CONTINUE
      DO 3 I=1,NH
      DO 3 J=1,NH
      ZDY(I,J)=0.0
      DO 4 K=1,N
    4 ZDY(I,J)=ZDY(I,J)+Z(K,I)*Y(K,J)
    3 CONTINUE
      CALL MINV(ZDZ,NH)
C     DO 27 JJ=1,NH
C     WRITE(*,35)(ZDZ(JJ,JU),JU=1,NH)
C  27 CONTINUE
```

```
      DO 5 I=1,NK
      DO 5 J=1,NH
      PIE(I,J)=0.0
      DO 6 K=1,NK
    6 PIE(I,J)=PIE(I,J)+ZDZ(I,K)*ZDY(K,J)
    5 CONTINUE
      DO 7 I=1,N
      DO 7 J=1,NH
      YHAT(I,J)=0.0
      DO 8 K=1,NK
    8 YHAT(I,J)=YHAT(I,J)+Z(I,K)*PIE(K,J)
    7 CONTINUE
      DO 9 I=1,N
      DO 9 J=1,NH
      YER(I,J)=Y(I,J)-YHAT(I,J)
    9 CONTINUE
      DO 10 I=1,NH
      RA(I)=0.0
      RH(I)=0.0
      DO 28 J=1,N
      RA(I)=RA(I)+Y(J,I)**2
   28 RH(I)=RH(I)+YER(J,I)**2
      R(I)=1.0-(RH(I)/(RA(I)-ZDY(NK,I)**2/N))
   10 CONTINUE
C     WRITE(NO,29)(R(I),I=1,NH)
   29 FORMAT(' R SQUARE VECTOR',6F9.3)
C     WRITE(NO,*)' BETA VECTORS FOR VARIOUS EQUATIONS'
C     DO 19 J=1,NH
C     WRITE(NO,41) J
C     WRITE(NO,18)(PIE(I,J),I=1,NK)
C  19 CONTINUE
   18 FORMAT(1X,7E10.3)
   41 FORMAT(' BETA COEFFICIENTS FOR EQN ',I3)
C     WRITE(*,*)' EXPECTED VALUES OF DEPENDENT VAR AND RES VECTORS'
C     DO 20 I=1,N
C     WRITE(*,21)(I,YHAT(I,J),YER(I,J),J=1,NH)
C  20 CONTINUE
   21 FORMAT(1X,I4,8F9.3)
      ANK=N-NK
      DO 32 I=1,NK
      DO 33 J=1,NH
      S=SQRT((RH(J)/ANK)*(ZDZ(I,I)))
      TT(I,J)=S
   33 CONTINUE
   32 CONTINUE
C     WRITE(NO,*)' STANDARD ERROR AND t VALUES FOR BETA'
C     DO 34 J=1,NH
C     WRITE(NO,42) J
C     WRITE(NO,36)(TT(I,J),PIE(I,J)/TT(I,J),I=1,NK)
C  34 CONTINUE
   36 FORMAT(1X,7E10.3)
   42 FORMAT(' STD ERRORS AND t VALUES FOR COEEFICIENTS OF EQN',I3)
      RETURN
      END
C ----------- MATRIX INVERSION BLOCK ------------------------------
```

```
      SUBROUTINE MINV(A,N)
      COMMON NHV,NHV,NV,NHH,NO
      DIMENSION A(NHH,NHH)
      DO 1 I=1,N
      A(I,I)=1.0/A(I,I)
      DO 2 J=1,N
      IF(J.NE.I) A(J,I)=A(J,I)*A(I,I)
    2 CONTINUE
      DO 4 J=1,N
      DO 6 K=1,N
      IF(J.NE.I.AND.K.NE.I) A(J,K)=A(J,K)-A(J,I)*A(I,K)
    6 CONTINUE
    4 CONTINUE
      [illegible]
      [illegible]
    5 CONTINUE
    1 CONTINUE
      RETURN
      END
```

BIBLIOGRAPHY

Christ, C.F. (1966) *Econometric Models and Methods*, New York, John Wiley & Sons.

Cramer, J.S. (1971) *Empirical Econometrics*. Amsterdam, North-Holland Publishing Company.

Dhrymes, P.J. (1970) *Econometrics: Statistical Foundations And Applications*. New York: Harper & Row.

Fisher, F.M. (1959) "*Generalisation of the Rank and Order Conditions for Identifiability*". Econometrica, Vol. 27, No. 3 (July), pp 431 - 447.

Fisher, F.M. (1966) *The Identification Problem in Econometrics*. New York, McGraw-Hill.

Goldberger, A.S(1964) *Econometric Theory*. New York, John Wiley & Sons.

Hood, W.C., and *Tjalling C.Koopmans* (Editors) (1953) *Studies in Econometric Method, Cowels Foundation Monograph 14*. New York, Wiley.

Intriligator, M.D. *(1978)* *Econometric Models: Techniques and Applications*. Prentice Hall, Inc., Englewood Cliffs, New Jersey.

Johnston, J. (1963) *Econometric Methods*. New York, McGraw-Hill.

Kmenta, J. (1971) *Elements of Econometrics*. New York, Macmillan.

Klein, L.R. (1953) *A Textbook of Econometrics*. Evanston, I ll, Row, Peterson.

Koutsoyiannis, A. *(1977)* *Theory of Econometrics: An Introductory Exposition of Econometric Methods*. 2nd Ed. London, Macmillan.

Maddala, G.S. (1977) *Econometrics*. New York, McGraw-Hill.

Malinvaud, E. (1966) *Statistical Methods of Econometrics*. Rand McNally.

Mishra, S.K. (1991) *"Generalisation of Indirect Least Square For Over-Identified Equations."* Dept. of Economics, N.E.H.U. Shillong.

Murphy, J.L. (1973) *Introductory Econometrics,* New York. Richard D. Irwin, Inc.

Theil, H. (1971) *Principles of Econometrics.* New York, John Wiley & Sons.

Tintner, G. (1952) *Econometrics.* New York, John Wiley & Sons.

Thomas, R.L. (1985) *Introductory Econometrics. Theory and Applications.* New York, Longman Group Limited.

Zarembka, P. (Ed) (1974) *Frontiers in Econometrics.* New York, Academic Press, Inc.