Monte Carlo Study of Econometric Estimators for Small Samples

George Donatos

A Thesis

in

The Department

of

Mathematics

Presented in Partial Fulfillment of the Requirements for the degree of Master of Science in Mathematics at
Concordia University
Montréal, Québec, Canada

October 1984

ABSTRACT

Monte Carlo Study of Econometric Estimators for Small Samples

George Donatos

The main purpose of this thesis is to provide a comprehensive examination of the study of Monte Carlo experiments utilised in the study of small sample properties of econometric estimators in simultaneous equation model.

Chapter I describes the simultaneous equation model and some well known methods of estimation.

Chapter II reviews the empirical works on small-sample properties of estimators which have been published in the last ten eyars.

Finally, Chapter III describes the Monte Carlo experiments performed in this study for the data from the Greek economy and presents the results and the conclusions.

ACKNOWLEDGEMENTS

I wish to thank Professor T.D. Dwivedi for his assistance during the preparation of this work.

DEDICATION

I wish to dedicate this thesis to the memory of my grandfather, Athanasios.

TABLE OF CONTENTS

CHAPTER I

SYSTEM OF SIMULTANEOUS EQUATIONS

1. INTRODUCTION

Economists have continually searched for appropriate statistical tools to provide meaningful inferences about the structure of the economic system. In many cases, this could not be achieved by routine application of statistical techniques developed in other disciplines such as astronomy and biology. Ronald A. Fisher's development of statistical techniques in the 1920's, which allowed inferences to be made from small samples, greatly enhanced capabilities in the field. However, the non-reproducible nature of economic observations, the simultaneity of the structural relations making up the economic system, and the dynamic characteristics of the economic process required a new approach.

The breakthrough in the investigation of simultaneous equations was made by Haavelmo (1943, 1944, 1947). He postulated that the economic activity could be analyzed as a simultaneous system of stochastic equations. Instead of determining a single dependent variable in an equation (as in OLS), Haavelmo obtained a joint distribution of dependent variables from the simultaneous structure. It is to be noted that if more than one dependent variable appeared in a particular equation, least square estimation yielded biased and inconsistent estimators of the population parameter.

In order to obtain better estimates, several alternatives to ordinary least squares have been developed in the econometrics literature. For example,

a) Full information maximum likelihood (FIML)

b) Limited information maximum likelihood (LIML)

c) Two stage least squares (2SLS)

d) Three stage least squares (3SLS)

e) K-class double K-class and h-class estimator.

In this chapter, we introduce the simultaneous equation model and show the inability of ordinary least squares method in estimating the parameters of simultaneous equation model. Further, we derive a few well-known methods of estimation of the parameters of simultaneous equation model for the further readability of this thesis.

2. SIMULTANEOUS EQUATION MODEL

Let us assume a linear model containing T structural relations. The ith equation at time t can be written as

$$\sum_{j=1}^{n} \beta_{ij} y_{tj} + \sum_{k=1}^{m} \gamma_{ki} z_{tk} = u_{ti} \qquad (1.1)$$

$$t = 1,2,\ldots,T \qquad i = 1,2,\ldots,n$$

where y_{tj} denote the endogenous variables at time t to be explained in the system, z_{tk} is either an exogenous variable or a lagged value of an endogenous variable. The exogenous and lagged variables are called predetermined variables. The model may then be regarded as a theory explaining the determination of the jointly dependent variables $y_{tj}(j = 1,2,\ldots,n\ ,\ t = 1,2,\ldots,T)$ in terms of the predetermined variables $z_{tk}(k = 1,2,\ldots,m\ ,\ t = 1,2,\ldots,T)$ and the disturbances $u_{ti}(t = 1,2,\ldots,T\ ,\ i = 1,2,\ldots,n)$. β and γ are known as structural coefficients and some of them may be equal to zero.

For T observations (1.1) can be written as:

$$\begin{matrix} y_{11} & y_{12} & y_{1n} \\ y_{21} & y_{22} & y_{2n} \\ \\ y_{T1} & y_{T2} & y_{Tn} \end{matrix} \quad \begin{matrix} \beta_{11} & \beta_{12} & \beta_{1n} \\ \beta_{21} & \beta_{22} & \beta_{2n} \\ \\ \beta_{n1} & \beta_{n2} & \beta_{nn} \end{matrix} \quad +$$

$$\begin{matrix} z_{11} & z_{12} & z_{1n} \\ z_{21} & z_{22} & z_{2n} \\ \\ z_{T1} & z_{T2} & z_{Tm} \end{matrix} \quad \begin{matrix} \gamma_{11} & \gamma_{12} & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \gamma_{2n} \\ \\ \gamma_{m1} & \gamma_{m2} & \gamma_{mn} \end{matrix} \quad =$$

$$\begin{matrix} u_{11} & u_{12} & u_{1n} \\ u_{21} & u_{22} & u_{2n} \\ \\ u_{T1} & u_{T2} & u_{Tn} \end{matrix} , \tag{1.2}$$

$$YB + Z\Gamma = U . \tag{1.3}$$

Assuming that B is a non-singular matrix, we mulitply (1.3) by B^{-1} to obtain

$$Y = -Z\Gamma B^{-1} + UB^{-1} , \text{ or}$$

$$Y = Z\Pi + V \tag{1.4}$$

where $\Pi = -\Gamma B^{-1}$ and $V = UB^{-1}$.

The equation (1.4) is called the reduced form. We can get to (1.4) if the system is such that the predetermined variables together with the errors uniquely determine the current endogeneous variables.

Let us consider (1.1) from the point of view of estimating its parameter from a sample of size T . For the moment we focus our attention on only one equation in (1.1), say the first. The equation will contain as explanatory variables, some current endogeneous variables say n_1 , and some predetermined ones, say m_1 where $n_1 < n$ and $m_1 \leq m$. We consider the current endogeneous variables appearing as explanatory variables $y_2 , y_3 \ldots y_{n_1+1}$; the dependent variable is, of course, numbered y_1 . The explanatory varialbes appearing in the first equation are numbered as $z_1 , z_2 \ldots z_{m_1}$. Now we can write the T observations in the first equation as

$$y_1 = Y_1 \beta_1 + Z_1 \gamma_1 + U_1 , \tag{1.5}$$

where y_1 is a $T \times 1$ vector, Y_1 and Z_1 are $T \times n_1$ and $T \times m_1$, matrices respectively. β_1 is a subvector of $(\beta_{11} \; \beta_{12} \; \beta_{n1})'$; γ_1 is a subvector of $(\gamma_{11} \; \gamma_{21} \; , \gamma_{n1})'$ and U_1 is a $T \times 1$ vector of disturbances.

3. FUNDAMENTAL ASSUMPTIONS

$E(u_{ti}) = 0$

$E(u_{ti}\ u_{t'j}) = \delta_{tt'}\sigma_{ij}$ where $\delta_{tt'}$ is the Kronker-delta

$E(z_{ti}\ u_{tj}) = 0$

$E(y_{ti}\ u_{tj}) \neq 0 \quad i,j = 1,2,\ldots,n$ (1.6)

4. METHODS OF ESTIMATION

Without any loss of generality we may consider equation (1.5) $y_1 = Y_1\beta_1 + Z_1\gamma_1 + U_1$ which can be written as:

$$y_1 = [Y_1 \vdots Z_1]\binom{\beta_1}{\gamma_1} + U_1 \tag{1.7}$$

or

$$y_1 = X_1\delta_1 + u_1 , \tag{1.8}$$

where

$$X_1 = [Y_1 \vdots Z_1] \text{ and } \delta_1 = \binom{\beta_1}{\gamma_1} . \tag{1.9}$$

The simplest method for estimating (1.8) is to employ least squares. So the least square estimator $\hat{\delta}_1$ can be written as follows: $\hat{\delta}_1 = (X_1'X_1)^{-1}X_1'y_1$, (provided $X_1'X_1$ is a non-singular matrix), which is the same as:

$$\begin{matrix} \hat{\beta}_1 \\ \hat{\gamma}_1 \end{matrix} = \begin{matrix} Y_1'Y_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{matrix}^{-1} \begin{matrix} Y_1'y_1 \\ Z_1'y_1 \end{matrix} \tag{1.10}$$

In the following we will show that the ordinary least squares estimates do not possess the desired properties, since the regressors Y_1 are not, in general, independent of the disturbances U_1. Also the ordinary least square estimates of this model are biased and inconsistant.

Let us consider a simple macroeconomic Keynesian model which is given as follows:

$$C_t = \alpha Z_t + \beta Y_t + U_t \tag{1.11}$$

$$Y_t = C_t + A_t \tag{1.12}$$

where

C_t = consumption at time t

Y_t = income at time t

$Z_t = 1$ for every t

A_t = investment at time t

U_t = disturbance

and $t = 1, 2, \ldots, T$.

In this model C_t and Y_t, consumption and income respectively, represent endogenous variables and A_t, the investment, is an exogeneous variable. It is easily seen that the Keynesian model can also be written as follows:

$$Y_t B + Z_t \Gamma = U_t , \tag{1.13}$$

where

$$Y_t = (C_t \quad y_t) , \qquad B = \begin{matrix} 1 & 1 \\ -\beta & 1 \end{matrix}$$

$$Z_t = (1 \quad A_t) , \qquad \Gamma \quad \begin{matrix} -\alpha & 0 \\ 0 & -1 \end{matrix} .$$

The disturbance or the error term U_t has the same standard specification

$$E(U_t) = 0 \quad \text{and} \quad E(U_t U_1') = \delta_{tt'} \sigma^2 \quad \text{for all} \quad t, t' .$$

Now as it appears that (1.12) does not have any parameter and therefore one would like to estimate the parameters of (1.11) by using ordinary least squares method.

Suppose we have a sample of size T, then the LS estimators of α and β in (1.11) are given by

$$\hat{\beta} = \frac{\Sigma(C_t - \bar{C})(Y_t - \bar{Y})}{\Sigma(Y_t - \bar{Y})^2} \quad \text{and} \quad \hat{\alpha} = \bar{C} - \hat{\beta}\bar{Y} \tag{1.14}$$

where $\bar{C}$, $\bar{Y}$ are respectively the sample means of consumption and income. Also $\bar{C}$ can be written as follows:

$$\bar{C} = \alpha + \beta\bar{Y} + \bar{U} . \qquad (1.15)$$

Now replacing the value of C_t and $\bar{C}$ we have:

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

Thus

$$P\lim \hat{\beta} = \beta + \frac{P\lim[\Sigma(Y_t - \bar{Y})(U_t - \bar{U})/T]}{P\lim[\Sigma(Y_t - \bar{Y})^2/T]} .$$

One notices that the probability limit of the numerator is the sample covariance between Y_t and U_t and the denominator probability limit is the same variance of Y_t. It is well known that the sample moments converge in probability to the population moments, we need to know the variance of Y_t and the covariance between Y_t and U_t.

In order to determine the required quantity, we solve (1.11) and (1.12) to obtain

$$Y_t = \frac{\alpha}{1 - \beta} + \frac{1}{1 - \beta} A_t + \frac{1}{1 - \beta} U_t ,$$

therefore $V(Y_t) = \frac{1}{(1 - \beta)^2} V(A_t) + \frac{1}{(1 - \beta)^2} V(U_t)$.

Now if we denote $V(A_t) = \sigma_A^2$ and $V(U_t) = \sigma^2$ then

$$V(Y_t) = \frac{\sigma_A^2 + \sigma^2}{(1 - \beta)^2} \quad \text{and} \quad Cov(Y_t\ U_t) = \frac{\sigma^2}{(1 - \beta)^2} .$$

Therefore it is easily seen that the probability limit of $\hat{\beta}$ exceeds β which implies the inconsistency of the ordinary least square estimator when applied to such models.

Actually one can directly show that $\hat{\delta}_1$ as given in (1.10) is inconsistant and the proof of this statement can be found in any text-book on econometric estimation. There are several methods of estimation described in literature of such a model, which produce a consistant estimator of the parameters. In the following, we are going to describe a few of them.

5. TWO STAGE LEAST SQUARES

The two-stage least squares (2SLS) estimation technique was developed by Theil (1953, 1961) and independently by Basmann (1957). The method of estimation clarifies its name. Without any loss of generality let us consider the model given in (1.8):

$$y_1 = X_1 \delta_1 + U_1$$

where $X_1 = [Y_1 \vdots Z_1]$ and $\delta_1 = \begin{matrix} \beta_1 \\ \gamma_1 \end{matrix}$.

Therefore $y_1 = [Y_1 \vdots Z_1] \begin{matrix} \beta_1 \\ \gamma_1 \end{matrix} + u_1$ (1.16)

In this notation the jointly dependent and predetermined variable are distinguished on the right hand side. This is important because the former variables are the cause of all the complication, as they are random, and corelated with the disturbance vector U_1. To solve this problem, we go back to the reduced form which shows that $E(Y) = - X B \Gamma^{-1}$, since $E(U) = 0$ and X is non-random where the reduced form is written as,

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

Thus the matrix $E(Y_1)$ is a submatrix of this $E(Y)$ which corresponds to those which occur on the right hand side of our model. Next we write (1.15) in the following equivalent form,

$$y_1 = [E(Y_1) \; Z_1] \begin{matrix} \beta_1 \\ \gamma_1 \end{matrix} + U_1 + [Y_1 - E(Y_1)] \beta_1 . \quad (1.17)$$

Now the first stage of 2SLS is to estimate $E(Y_1)$ using least squares.

Thus we estimate Π_1 in $Y_1 = X\Pi_1 + V_1$ to obtain,

$$\hat{\Pi}_1 = (X'X)^{-1}X'Y_1 \tag{1.18}$$

and

$$\hat{Y}_1 = X(X'X)^{-1}X'Y_1 . \tag{1.19}$$

Now in the second stage of 2SLS y_1 regressed on $\hat{Y}_1$ and Z_1. This yields the estimating equation as follows:

$$\begin{bmatrix} \hat{Y}_1'\hat{Y}_1 & \hat{Y}_1'Z_1 \\ Z_1'\hat{Y}_1 & Z_1'Z_1 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\gamma}_1 \end{bmatrix} = \begin{bmatrix} \hat{Y}_1'y_1 \\ X_1'y_1 \end{bmatrix} \tag{1.20}$$

Now using (1.19) we have

$$\hat{Y}_1'\hat{Y}_1 = Y_1'X(X'X)^{-1}X'Y_1 , \tag{1.21}$$

and

$$Z_1'\hat{Y}_1 = Z_1'X(X'X)^{-1}X'Y_1 = Z_1'Y_1 \tag{1.22}$$

because

$$Z_1 = [Z_1 \vdots Z_2] \begin{bmatrix} I \\ \cdots \\ 0 \end{bmatrix} = X \begin{bmatrix} I \\ \cdots \\ 0 \end{bmatrix} \tag{1.23}$$

where $X = [Z_1 \vdots Z_2]$ and partitions are of the right order.

Therefore, we can express the 2SLS estimator as

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\gamma}_1 \end{bmatrix}_{2SLS} = \begin{bmatrix} Y_1'X(X'X)^{-1}X'Y_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} Y_1'X(X'X)^{-1}X'y_1 \\ Z_1'y_1 \end{bmatrix} \tag{1.24}$$

Now the OLS estimator of V_1 in $Y_1 = X\Pi_1 + V_1$ is

$$V_1 = Y_1 - X\hat{\Pi}_1 = Y_1 - X(X'X)^{-1}X'Y_1 \; ;$$

and therefore,

$$Y_1'X(X'X)^{-1}X'Y_1 = Y_1'Y_1 - V_1'V_1 \; .$$

Hence we may express the 2SLS estimator as

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\gamma}_1 \end{bmatrix} = \begin{bmatrix} Y_1'Y_1 - V_1'V_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} Y_1' - V_1' \\ Z_1' \end{bmatrix} y_1 \; . \qquad (1.25)$$

If we want to solve for $\hat{\beta}_1$ and $\hat{\gamma}_1$ let us write (1.24) as

$$\begin{bmatrix} Y_1'X(X'X)^{-1}X'Y_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\gamma}_1 \end{bmatrix}_{2SLS} = \begin{bmatrix} Y_1'X(X'X)^{-1}X'y_1 \\ Z_1'y_1 \end{bmatrix} \qquad (1.26)$$

$$Y_1'X(X'X)^{-1}X'Y_1\hat{\beta}_1 + Y_1'Z_1\hat{\gamma}_1 = Y_1'X(X'X)^{-1}X'y_1 \qquad (1.27)$$

$$Z_1'Y_1\hat{\beta}_1 + Z_1'Z_1\hat{\gamma}_1 = Z_1'y_1 \; . \qquad (1.28)$$

Multiplying (1.28) by $(Z_1'Z_1)^{-1}$ we get

$$\hat{\gamma}_1 = (Z_1'Z_1)^{-1} Z_1'y_1 - (Z_1'Z_1)^{-1} Z_1'Y_1 \hat{\beta}_1$$

$$= (Z_1'Z_1)^{-1} Z_1'[y_1 - Y_1\hat{\beta}_1] \qquad (1.29)$$

and substituting this in (1.27) we get

$$Y_1'X(X'X)^{-1}X'Y_1\hat{\beta}_1 + Y_1'Z_1(Z_1'Z_1)^{-1}Z_1'[y_1 - Y_1\hat{\beta}_1] = Y_1'Z(X'X)^{-1}X'y_1$$

or

$$Y_1'X(X'X)^{-1}X'Y_1\hat{\beta}_1 - Y'Z_1(Z_1'Z_1)^{-1}Z_1'Y_1\hat{\beta}_1 = Y_1'X(X'X)^{-1}X'y_1 - Y_1'Z_1(Z_1'Z_1)^{-1}Z_1'y_1$$

or

$$\{Y_1'[X(X'X)^{-1}X' - Z_1(Z_1'Z_1)^{-1}Z_1']Y_1\}\hat{\beta}_1 = Y_1'[X(X'X)^{-1}X' - Z_1(Z_1'Z_1)^{-1}Z_1']y_1 .$$

Therefore we have

$$\hat{\beta}_1 = [Y_1'NY_1]^{-1}Y_1'Ny_1 \tag{1.30}$$

where $N = X(X'X)^{-1}X' - Z_1(Z_1'Z_1)^{-1}Z_1'$. (1.31)

The 2SLS estimator can be derived more elegantly in the following way.

Let us pre-multiply the structural equation

$$y_1 = Y_1\beta_1 + Z_1\gamma_1 + U_1 \text{ by } X' .$$

We have

$$X'y_1 = X'Y_1\beta_1 + X'Z_1\gamma_1 + X'U_1 \tag{1.32}$$

or

$$X'y_1 = (X'Y_1 \quad X'Z_1)\begin{pmatrix}\beta_1\\ \gamma_1\end{pmatrix} + X'U_1 . \tag{1.33}$$

It is to be noted that the transformed disturbances have zero mean.

$$E(X'U_1) = X'E(U_1) = 0 . \tag{1.34}$$

Furthermore $E(U_1U_1') = \sigma^2 I$, where σ^2 is the residual variance. Therefore, the covariance matrix of transformed disturbances is given by

$$E(X'UU'X) = \sigma^2 X'X . \tag{1.35}$$

Now using (1.35), if we apply generalized least squares to (1.33) we obtain

$$\begin{matrix}\hat{\beta}_1\\ \\ \hat{\gamma}_1\end{matrix} = \left(\begin{matrix}Y_1'X\\ \\ Z_1'X\end{matrix}\;(\sigma^2 X'X)^{-1}(X'Y_1 \;\vdots\; X'Z_1)\right)^{-1}\begin{matrix}Y_1'X\\ \\ Z_1'X_1\end{matrix}\;(\sigma^2 X'X)^{-1} X'y_1$$

$$= \begin{matrix}Y_1'X(X'X)^{-1}X'Y_1 & \vdots & Y_1'X(X'X)^{-1}X'Z_1 & & Y_1'X(X'X)^{-1}X'y_1\\ & \vdots & & & \\ Z_1'X(X'X)^{-1}X'Y_1 & \vdots & Z_1'X(X'X)^{-1}X'Z_1 & & Z_1'X(X'X)^{-1}X'y_1\end{matrix}$$

(1.36)

because the unknown scalar parameter σ^2 cancels out in the straight forward manner. Further utilizing (1.23) the estimator in (1.36) can be written as

$$\begin{matrix}\hat{\beta}_1\\ \\ \hat{\gamma}_1\end{matrix} = \begin{matrix}Y_1'X(X'X)^{-1}X'Y_1 & Y_1'Z_1\\ & \\ Z_1'Y_1 & Z_1'Z_1\end{matrix}\quad\begin{matrix}Y'X(X'X)^{-1}X'y_1\\ \\ Z_1'y_1\end{matrix}, \qquad (1.37)$$

which is identical with 2SLS estimator obtained above.

6. FAMILIES OF GENERAL k-CLASS, h-CLASS AND DOUBLE k-CLASS ESTIMATOR

The OLS (ordinary least squares) of $\begin{pmatrix}\beta\\ \gamma\end{pmatrix}$ in the structural equation (1.7) may be expressed as

$$\begin{pmatrix}\hat\beta_1\\ \hat\gamma_1\end{pmatrix} = \left[\begin{pmatrix}Y_1'\\ Z_1'\end{pmatrix}(Y_1\ Z_1)\right]^{-1}\begin{pmatrix}Y_1'\\ Z_1'\end{pmatrix} y_1$$

$$= \begin{pmatrix}Y_1'Y_1 & Y_1'Z_1\\ Z_1'Y_1 & Z_1'Z_1\end{pmatrix}^{-1}\begin{pmatrix}Y_1'\\ Z_1'\end{pmatrix} y_1 \ . \qquad (1.38)$$

Comparing this with 2SLS estimator in (1.25) we find that the two estimators differ only in the leading matrices. In the latter case we have subtracted $V_1'V_1$ from $Y_1'Y_1$ and V_1' from Y_1'. It can be proved that this correction of OLS estimator enables the two stage least squares estimator to be "consistant".

Alternatively, we could subtract only a part of $V_1'V_1$ and a part of V_1' from $Y_1'Y_1$ and Y_1' respectively. Thus for an arbitrary scalar k (stochastic or non-stochastic) we have

$$\begin{pmatrix}\hat\beta_1\\ \hat\gamma_1\end{pmatrix}_k = \begin{pmatrix}Y_1'Y_1 - kV_1'V_1 & Y_1'Z_1\\ Z_1'Y_1 & Z_1'Z_1\end{pmatrix}^{-1}\begin{pmatrix}Y_1' - kV_1'\\ Z_1'\end{pmatrix} y \qquad (1.39)$$

as an estimator of $\begin{pmatrix}\beta_1\\ \gamma_1\end{pmatrix}$. Note that by specifying different values to k we arrive at different estimators. For example, for $k = 0$ (1.39) provides the OLS estimator defined in (1.38), and for $k = 1$

we have the 2SLS estimator defined in (1.25). In general (1.39) provides the family of general k-class estimators of parameters. This method was originally proposed by Theil.

We observe that the 2SLS estimator was obtained by applying OLS to the structural equation (1.7) after replacing Y_1 by

$$\hat{Y}_1 = Y_1 - V_1 = X(X'X)^{-1}X'Y_1 . \tag{1.40}$$

Now instead of replacing Y_1 by $Y_1 - V_1$ let us replace it by $Y_1 - hV_1$, where h is an arbitrary scalar which may be stochasitc or non-stochastic. Then

$$y_1 = (Y_1 - hV_1)\beta_1 + Z_1\gamma_1 + U_1 \tag{1.41}$$

is the new form of the structural equation. Applying OLS to (1.41) we get

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\gamma}_1 \end{pmatrix}_h = \left[\begin{pmatrix} Y_1' - hV_1' \\ Z_1' \end{pmatrix} (Y_1' - hV_1' \vdots Z_1)\right]^{-1} \begin{pmatrix} Y_1' - hV_1' \\ Z_1' \end{pmatrix} y_1$$

$$= \begin{pmatrix} Y_1'Y_1 - (2h - h^2)V_1'V & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{pmatrix}^{-1} \begin{pmatrix} Y_1' - hV_1' \\ Z_1' \end{pmatrix} y_1 . \tag{1.42}$$

This is known as the family of h-class estimators and was also proposed by Theil. We note that for $h = 0$ we have the OLS and for $h = 1$, the 2SLS estimator as a member of h-class family.

We can it seems combine the two families of general k-class and h-class into one family by defining the Double k-class estimator as follows:

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\gamma}_1 \end{pmatrix}_{k_1,k_2} = \begin{bmatrix} Y_1'Y_1 - k_1V_1'V_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{bmatrix}^{-1} \begin{bmatrix} Y_1' - k_2V_1' \\ Z_1' \end{bmatrix} y_1 \quad (1.43)$$

where k_1 and k_2 are arbitrary scalers (which may be stochastic or non-stochastic). This was originally proposed by Nagar.

We note that $k_1 = k_2 = 0$ provides the OLS and $k_1 = k_2 = 1$ provides the 2SLS estimators as members of the family of Double k-class estimators. If $k_1 = k_2 = k$ we have the general family of k-class estimators and $k_1 = 2h - h^2$ and $k_2 = h$ provides the family of h-class estimators.

7. AN INDENTITY CONNECTING THE DOUBLE k-CLASS AND TWO-STAGE LEAST SQUARE ESIMATORS

If we write (1.5) $y_1 = Y_1\beta_1 + Z_1\gamma_1 + U_1$ in the following form

$$y_1 = Z\delta + U_1 \tag{1.44}$$

where $Z = (Y_1 \quad Z_1)$ and $\delta = \binom{\beta}{\gamma}$ (1.45)

the double k-class estimator can be expressed compactly in the following form

$$\hat{\delta}_{k_1k_2} = [Z'(I-k_1M)Z]^{-1}Z'[I-k_2M]y_1 \tag{1.46}$$

where $M = I - X(X'X)^{-1}X'$. (1.47)

The 2SLS estimator

$$\hat{\delta}_{2SLS} = [Z'M^*Z]^{-1}Z'M^*y \tag{1.48}$$

is obtained by writing $k_1 = k_2 = 1$ in (1.46)

where $M^* = X(X'X)^{-1}X'$. (1.49)

Since the following identity is true for any two matrices Q_1 and Q_2

$$(Q_1 + Q_2)^{-1} = [I-(Q_1+Q_2)^{-1}Q_2]Q_1^{-1} \tag{1.50}$$

provided Q_1 and $Q_1 + Q_2$ are nonsingular.

Now in (1.46) consider the matrix $[Z'(I-k_1M)Z]^{-1}$, which can be written as follows:

$$\begin{aligned}[Z'(I-k_1M)Z]^{-1} &= [Z'Z-k_1Z'MZ]^{-1} = [Z'Z-k_1Z'(I-M^*)]Z]^{-1}\\ &= [Z'Z-k_1Z'Z+k_1Z'M^*Z]^{-1}\\ &= [k_1Z'M^*Z+(1-k_1)Z'Z]^{-1}.\end{aligned}$$

Now applying the identity (1.49) to the above matrix we obtain

$$[Z'(I-k_1M)Z]^{-1} = \frac{1}{k_1}[I-(1-k_1)[Z'(I-k_1M)Z]^{-1}Z'Z](Z'M^*Z)^{-1}$$

$$= \frac{1}{k_1}(Z'M^*Z)^{-1} - \frac{1-k_1}{k_1}[Z'(I-k_1M)Z]^{-1}Z'Z(Z'M^*Z)^{-1} \qquad (1.51)$$

Therefore we have

$$[Z'(I-k_1M)Z]^{-1}Z'M^*y_1 =$$

$$\frac{1}{k_1}\hat{\delta}_{2SLS} - \frac{1-k_1}{k_1}[Z'(I-k_1M)Z]^{-1}Z'Z\hat{\delta}_{2SLS} .$$

Now if we express

$$Z'(I-k_2M)y_1 = k_1Z'M^*y_1 - (1-k_1)Z'M^*y_1 + (1-k_2)Z'My_1$$

$$= k_1Z'M^*y_1 - (1-k_1)Z'M^*Z\hat{\delta}_{2SLS} + (1-k_2)Z'My .$$

Thus we arrive at the result

$$\hat{\delta}_{k_1k_2} = \hat{\delta}_{2SL} + [Z'(I-k_1M)Z]^{-1}[(1-k_2)Z'My_1 - (1-k_1)Z'MZ\hat{\delta}_{2SL}] .$$

This identity is due to Srivastava and Tiwari (1977). Dhyremes (1969) derived a similar relationship between the Double k-class and 2SLS estimator. However, what Dhrymes calls the Double k-class is only a subset of the entire family of Double k-class. Therefore his result is not comparable with the more general result of Srivastava and Tiwari.

Furthermore (1.46) can be written as

$$\hat{\delta}_k = [Z'(I-km)Z]^{-1}Z'(I-km)y_1 \qquad \text{for } k_1 = k_2 = k .$$

Notice that it is a k-class estimator. It is easily seen that

$$\hat{\delta}_{k_1k_2} = \hat{\delta}_k + (k_1 - k_2)[Z'(I - k_1M)Z]^{-1}Z'My_1 \quad . \tag{1.52}$$

Also note that

$$[Z'(I-k_1m)Z]^{-1}Z'my_1 = [Z'(I-k_1M)Z]^{-1}Z'(I-M^*)y_1$$

$$= [Z'(I-k_1m)Z]^{-1}Z'y_1 - [Z'(I-k_1m)Z]^{-1}Z'M^*y_1 .$$

Now substituting from (1.51) to the second term on the RHS of the above we obtain:

$$[Z'(I-k_1m)Z]^{-1}Z'my_1 = [Z'(I-k_1m)Z]^{-1}Z'y_1 - \frac{1}{k_1}[I-(1-k_1)$$

$$[[Z'(1-k_1)m)Z]^{-1}Z'Z][Z'M^*Z)^{-1}Z'M^*y_1$$

$$= [Z'(I-k_1m)Z]\; Z'y_1 - \frac{1}{k_1}[I-(1-k_1)$$

$$[[Z'(1-k_1m)Z]^{-1}Z'Z]\hat{\delta}_{2SLS} \quad . \tag{1.53}$$

Again using the same identity we can write the above equation in the following form:

$$= \frac{1}{1-k_1}[I-[Z'(1-k_1M)Z]^{-1}k_1Z'M^*Z]\hat{\delta}_{OLS} - \frac{1}{k_1}[I-(1-k_1)$$

$$[Z'(I-k_1M)Z]^{-1}Z'Z]\;\hat{\delta}_{2SLS} \quad . \tag{1.54}$$

Now using (1.52) and (1.54) we obtain

$$\hat{\delta}_{k_1k_2} = \hat{\delta}_{k_1} + \frac{k_1-k_2}{1-k_1}\, I-[Z'(1-k_1m)Z]^{-1} k_1 Z' M^* Z \quad \hat{\delta}_{OLS} -$$

$$\frac{k_1-k_2}{k_1} \{(-(1-k_1)[Z'(I-k_1m)Z'\}^{-1} Z'Z\}\, \hat{\delta}_{2SLS} \quad . \quad (1.55)$$

The double k-class, k-class, two stage least squares and ordinary least squares estimator of the parameter denoted as $\hat{\delta}_{k_1k_2}$, $\hat{\delta}_{k_1}$, $\hat{\delta}_{2SLS}$ and $\hat{\delta}_{OLS}$ respectively are connectdd by the above relation. This result is due to Dwivedi (1981).

Now other estimators are also of the same form as discussed above and can be seen in good text of econometrics. So we will not present them here.

CHAPTER II

SMALL SAMPLE PROPERTIES

1. INTRODUCTION

The asymptotic properties of econometric estimators are well known and are contained in almost every modern textbook. However, it is not always possible to have large number of data points in the present changing economic systems. Therefore it is important that we know the behaviour of the estimators when they are obtained with moderate sample size.

It is well known that the estimators obtained by using 2-stage least squares method and limited Information Maximum likelihood method, are asymptotically equivalent, also the same is true for 3-stage least squares and Full Information Maximum likelihood estomators. However, no certain conclusions have been drawn yet about the properties of the estimators when they are obtained with moderate sample size irrespective of the methods employed to obtain them. For this the Monte Carlo methodology has been applied to get some feeling of the behaviour of the estimators for moderate sample size.

In the following we will summarize the conclusions that have been reached using the Monte Carlo techniques.

2. A SURVEY OF THE MONTE CARLO METHODS

Wagner (1958) examined certain small sample properties of "limited-information single equation" maximum likelihood estimates for two models by a Monte Carlo approach. That is, 100 sets of observations over 20 time periods are generated from the models and then with these observations using various statistical techniques, estimates of the parameters of an over identified equation are obtained and compared.

The models, differing only in the variance-covariance matrix of the disturbances, consists of three equations, one of which is an identity. The two models investigated differ only in the variance-covariance matrix. The structural equations are:

$$y_1 - \beta_1 y_2 + \qquad - \gamma_1 = U_1 \tag{2.1}$$

$$- \beta_2 y_2 + y_3 - \gamma_2 Z_1 - \gamma_3 = U_2 \tag{2.2}$$

$$y_1 - y_2 + y_3 + Z_2 = 0 \tag{2.3}$$

The y's are endogenous variables and the Z's are predetermined and exogenous variables. The $U(t)$'s are independently and identically normally distributed with mean zero and finite covariance matrix. The model has the additional restriction

$$Z_1(t) = y_2(t-1) .$$

Thus the first predetermined variable is a one period lagged endogenous variable, and Z_2 is "trend variable".

From 2.3

$$y_3 = y_2 - y_1 - Z_2$$

substituting y_3 in (2.2) we obtain

$$-\beta_2 y_2 + y_2 - y_1 - Z_1 - \gamma_2 Z_1 - \gamma_3 = U_2$$

or
$$(1-\beta_2)y_2 - y_1 - Z_2 - \gamma_2 Z_1 - \gamma_3 = U_2$$

or
$$-y_1 + (1-\beta_2)y_2 - Z_2 - \gamma_2 Z_1 - \gamma_3 = U_2 \qquad (2.4)$$

and equation (2.1)

$$y_1 - \beta_1 y_2 - \gamma_1 = U_1$$

we can write the two equations as follows:

$$\begin{pmatrix} 1 & -\beta_1 \\ -1 & (1-\beta_2) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -\gamma_1 \\ -\gamma_2 & -1 & -\gamma_3 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ 1 \end{pmatrix} = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \qquad (2.5)$$

let

$$\beta = \begin{pmatrix} 1 & -\beta_1 \\ -1 & (1-\beta_2) \end{pmatrix} \qquad \Gamma = \begin{pmatrix} 0 & 0 & -\gamma_1 \\ -\gamma_2 & -1 & -\gamma_3 \end{pmatrix} \qquad Z = \begin{pmatrix} Z_1 \\ Z_2 \\ 1 \end{pmatrix}$$

$$U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} .$$

So the model can be written in the following way:

$$-\beta y + \Gamma Z = U . \qquad (2.6)$$

Now the reduced form of this model is as follows:

$$y = -\beta^{-1}\Gamma Z + \beta^{-1} U . \qquad (2.7)$$

As it can be seen that $|\beta| \neq 0$,

letting $\Pi = -\beta^{-1}\Gamma \qquad r = \beta^{-1} U$

we have

$$y = \Pi Z + r . \qquad (2.8)$$

The two models that Wagner considered had the following values for the structural parameters:

$$\beta_1 = .50 \qquad \beta_2 = .10$$

$$\gamma_1 = .25 \qquad \gamma_2 = .30 \qquad \text{and} \qquad \gamma_3 = .15 \,.$$

Thus one can easily see that the matrix Π of the reduced form is as follows:

$$\Pi = \begin{matrix} .375 & 1.25 & .75 \\ .750 & 2.50 & 1 \end{matrix}$$

$$\text{and} \qquad r = \begin{matrix} \dfrac{.9U_1 + .5U_2}{.4} \\ \\ \dfrac{U_1 + U_2}{.4} \end{matrix}$$

Let the variance-covariance of U be

$$\Sigma = \begin{matrix} 1 & .5 \\ .5 & 1 \end{matrix}$$

Then the variance-covariance matrix of r can be found as follows:

$$E(r\,r') = \begin{matrix} \dfrac{.81\sigma_1^2+.25\sigma_2^2+.90\sigma_{12}}{.16} & \dfrac{.9\sigma_1^2+1.4\sigma_{12}+.5\sigma_2^2}{.16} \\ \\ \dfrac{.9\sigma_1^2+1.4\sigma_{21}+.5\sigma_2^2}{.16} & \dfrac{\sigma_1^2+2\sigma_{12}+\sigma_2^2}{.16} \end{matrix}$$

Now substituting the values of σ_1^2, σ_2^2 and σ_{12} we obtain the variance-covariance matrix of r as given below:

$$r = \begin{matrix} 9.4375 & 13.1250 \\ 13.1250 & 18.750 \end{matrix} \qquad (2.9)$$

Here we would like to report a mistake in Wagner's equation (21) where he reports his r to be

$$\Omega = \begin{matrix} 9.4375 & 13.1250 \\ 13.1250 & 19.750 \end{matrix}.$$

The correlation between the random variables U_1 and U_2 is

$$\text{corr}(U_1, U_2) = \frac{\sigma_{12}}{\sigma_1 \sigma_2} = .5$$

For the same model Wagner specifies another variance-covariance matrix for the disturbance vector U as follows:

$$\Sigma = \begin{matrix} 8.40254 & -8.05232 \\ -8.05232 & 7.80599 \end{matrix}. \qquad (2.10)$$

He specifies this as Model II.

This variance-covariance yields the variance-covariance of the reduced form disturbance vector r as follows:

$$\Omega = \begin{matrix} 9.44041 & 1.20018 \\ 1.20018 & .64927 \end{matrix}. \qquad (2.11)$$

In this case the correlation between U_1 and U_2 is

$$\text{corr}(U_1, U_2) = -.9943.$$

The vector Z is specified as $Z_1(1) = 1$ (i.e. $y_2(0) = 1$) and $Z_2 = .25, .50, .75 \ldots\ldots 5$. The values of y_1 and y_2 are obtained using the reduced from (2.8). The disturbance vector r is generated as follows.

To manufacture r' having the variance-covariance matrices for model I and II we first consider a vector r of random normal disturbance having mean zero and variance-covariance matrix

$$\phi = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} = I \tag{2.12}$$

A set of such normal random variables was supplied by the RAND corporation and was used to obtain the components of s . Having secured a particular s we make a transformation T

$$Ts = r$$

where

$$T = \begin{matrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{matrix}$$

and such that T satisfies

$$T\phi T' = \Omega \, . \tag{2.13}$$

Now having generated the vector r , and given the reduced form coefficient of z , y can be generated. Wagner considered $t = 1, 2, \ldots, 20$ i.e. in other words, each set consists of 20 observations for the vector y .

He then estimates the parameters by the methods of limited information single equation maximum likelihood, least squares and instrumental variables, for each 100 sets of observations on each model.

The conclusion of this study shows that estimates obtained by Limited information method differ greatly from the true value. Since Model II has lagre variances it turns out as expected; the variances of the estimates in Model II is greater than the variances of the estimates in Model I. Also the bias of LS estimates shows up not only on the average but also in nearly every individual estimate. In Model I the instrumental variable estimates encompass LISE estimates and are better than LISE estimates (using Z_2 as instrumental variable). The same sort of generalization does not hold for Model II.

In this study Wagner has constructed the distribution of estimates by three different methods to examine the effect of small samples on the distribution of the estimates. However it has not been possible to come up with any definite conclusions.

Neiswanger and Yancey (1959) utilized Monte Carlo techniques to study the properties of estimates under some conditions which violate specifications of the estimation method employed. These experimental estimates are made under controls which reveal the influence of such specification errors. Two methods of estimation namely least squares and limited information simple equation methods are tested on time series data of small model.

Based on the results of this study, the trial inclusion of time as an additional term in the equations of a model seems justified when time series data are used. If autonomous growth, or similar variations exist in the endogenous variables the estimates will be improved and if autonomous growth is not present the estimates are not made worse on the average and the coefficient for time will be very small relative to its standard error.

Nagar (1960) used the Monte Carlo experiment to investigate small sample properties of the estimator obtained by four methods namely, least squares, two stage least squares, unbiased, and minimum second moment. He used WAgner's model for the experiment with the same specific values of the parameters and same variance-covariance structures.

Certain methods of estimation have been disregarded as they have already been considered by Wagner. WAgner considered only one equation in each model and one which is over identified, whereas Nagar also considered the second equation which is just identified. Just identification implies that the two-stage least squares estimator is identical with the limited information maximum likelihood estimator.

As one can see that the two models considered by Wagner differ only in the variance-covariance matrix of the random disturbances. Nagar has omited the constant terms in both equations for simplicity in the analysis.

One hundred samples, each being of size twenty, were generated for Model I and II. Thus he obtains 100 estimates of the parameter for each of the alternative methods of estimation. The following statistics are calculated for the parameter estimates of the first equation of the model:

1) Arithmetic mean of estimates = $\frac{1}{100} \Sigma \hat{\gamma}_1$, where summation is taken over all 100 estimates for each of the estimation methods;

2) Estimated bias = $\frac{1}{100} \Sigma \hat{\gamma}_1 - .5$; 0.5 being the true value of γ_1 ;

3) Estimated second moment about the mean = $\frac{1}{100} \Sigma \hat{\gamma}_1^2 - (\frac{1}{100} \Sigma \hat{\gamma}_1)^2$;

4) Estimated second moment about the true value = estimated second moment about the mean + (estimated Bias)^2 .

Finally for each of the 100 samples the alternative estimation methods have been ranked according to increasing distance from the true parameter value and the sum of the ranks is given. Hence, a low total of the ranks is an indicator of relative success in estimation. The results are reported in the various tables.

The parameter estimates of the second equation were obtained in the same way and similar statistics were calculated and reported in the tabular form.

The main results are summarized as follows:

i) The small sample bias of least squares exceeds the bias of the consistent estimation methods. There are sixteen comparisons regarding this and the bias reduction ranges from about 25 to almost 100 per cent, the median reduction being of the order of 60 per cent.

ii) Least squares has smaller sampling variance around its (biased) expectation than the consistent methods. The relative excess of the variance of the latter methods over that of least squares varies from a negligible to a tenfold difference, the median excess being of the order of 50 per cent of the least squares variance.

iii) The large bias of least squares implies that its second moment about the true parameter value exceeds that of the consistent methods in twelve cases out of sixteen. If we express the second-order sampling moments about the true value for the consistent methods as a percentage of the corresponding mement for least squares, we obtain figures ranging from about 40 to 110, the median being of the order of 90.

iv) The simplest of the consistent methods considered here is the two-stage least squares method, which shows the smallest bias in all cases. Theoretically, this method is only asymptotically unbiased. We also consider a method which is unbiased to the order 1/T (T being the number of observations), but strangely enough, this method shows a consistently larger bias than two-stage least squares. This anomaly can possibly be explained by the fact that the unbiased estimator is only unbiased to order 1/T when there are no lagged endogenous variables, while in Wagner's models there is one variable of this kind.

v) The third consistent method is the so-called minimum-second-moment method, which minimizes the determinant value of the moment matrix of the sampling errors. This method has the smallest second-order sampling moment about the true value in four cases out of six (two models containing three parameters each); in the two other cases the unbiased method - to roder 1/T - ranks first in this respect.

vi) The minimum-second-moment method belongs, just as the other three methods, to the k-class; but it differs from these in that its k varies from sample to sample. The distribution of this k over the sets of samples analyzed shows a large range; in a rather substantial minority of cases k is negative, which appears rather extreme when it is realized that the inconsistent least squares method corresponds with $k = 0$ and that consistency requires that k approaches 1 for large samples. A modified minimum-second-moment approach is considered, which implies that the original minimum-second-moment estimates are rejected in favour of two-stage least squares as soon as its k is below a preassigned critical level. This does not lead to substantial improvements.

vii) The asymptotic standard errors of two-stage least squares give a rather satisfactory picture of the variability of the estimates about the true value. This is not true for least squares in all cases considered. Instead, it seems that the classical least squares standard errors measure the variability of the estimates about the biased expectation, not about the true value. In some cases this makes a very large difference.

Quandt (1965) reports the results of a set of sampling experiments on a four equation model. These experiments involve the computation of k-class estimates for alternative values of k, with special emphasis being given to direct least squares $k = 0$ and two stage least squares for $k = 1$.

The main objective of this Monte Carlo study was to gather evidence concerning the relative performance of direct least squares and two stage least squares. The second objective was to test the hypothesis that the bias, for both methods of estimation, diminishes as the sparseness (the prevalence of a prior zeros in the coefficient matrix of Y in the structural equation) increases, assuming that the number of zeros are such that the reduced form is assured. And the final objective was to test the hypothesis that the bias for both the methods of estimation decreases as the sparseness of the covariance matrix Σ of error terms increases.

Two models were used for the experiments. Model I is

$$\begin{aligned}
y_1 - 0.2y_2 + 2.0y_3 + y_4 - Z_1 - .5Z_2 + Z_3 &= U_1 \\
-y_1 + y_2 + .5y_3 + .1y_4 - 2.0Z_2 - Z_5 &= U_2 \\
1.5y_1 - .5y_2 + y_3 + .2y_4 + .5Z_3 + Z_4 - 2Z_5 &= U_3 \\
.4y_1 + y_2 - .5y_3 + y_4 - Z_3 - .2Z_5 - 3Z_6 &= U_4
\end{aligned} \tag{2.14}$$

Model II is the same as above with the exception that it includes an additional variable Z_7 with the coefficient 0 , .5 , 0 , 0 respectively.

The first equation which is just identified in Model I, is overidentified in Model II.

Predetermined variables $Z_1, \ldots, Z_7$ are truly exogenous and do not include lagged values of endogenous variables. Values of the exogenous variables were fixed for repeated samples, although two different sets of data were used for the exogenous variables in order to examine the effect of multicollinearity. These two data sets have the following correlation matrices:

For data set 1:

1.000	0.872	0.325	0.707	-0.726	0	0.152
	1.000	0.440	0.847	-0.879	0	0.420
		1.000	0.429	-0.469	0	-0.049
			1.000	-0.865	0	0.464
				1.000	0	-0.274
					1.000	0
						1.000

and for data set 2:

1.000	-0.423	0.335	0.589	0.068	0	0.193
	1.000	-0.165	-0.539	0.439	0	0.065
		1.000	0.132	-0.291	0	0.556
			1.000	-0.095	0	-0.030
				1.000	0	0.217
					1.000	0
						1.000

where the sixth variable is a constant for both data sets. The experiments are performed as follows:

1) Twenty vectors $(U_1, U_2, U_3, U_4)'$ are generated, where the elements of the vector are jointly normally distributed with zero mean and covariance matrix Σ. The latter was either

$$\Sigma_1 = \begin{matrix} 1.0 & 0.6 & 0.8 & -1.0 \\ 0.6 & 1.0 & 0 & -0.2 \\ 0.8 & 0 & 2.0 & -0.6 \\ -1.0 & -0.2 & -0.6 & 2.5 \end{matrix}$$

$$\Sigma_2 = \begin{matrix} 1.0 & 0 & 0 & 0 \\ 0 & 1.0 & 0 & 0 \\ 0 & 0 & 2.0 & 0 \\ 0 & 0 & 0 & 2.5 \end{matrix}$$

2) Given the values of 20 z-vectors, the values of the endogenous variables are calculated from

$$y = -B^{-1}\Gamma Z + B^{-1}U,$$

where the B-matrix was altered from run to run. Five B-matrices were used,

$$B_1 = \begin{matrix} 1.0 & -0.2 & 2.0 & -1.0 \\ -1.0 & 1.0 & 0.5 & 0.1 \\ 1.5 & -0.5 & 1.0 & 0.2 \\ 0.4 & 1.0 & -0.5 & 1.0 \end{matrix}$$

$$B_2 = \begin{matrix} 1.0 & -0.2 & 2.0 & -1.0 \\ -1.0 & 1.0 & 0 & 0.1 \\ 0 & -0.5 & 1.0 & 0.2 \\ 0.4 & 0 & -0.5 & 1.0 \end{matrix}$$

$$B_3 = \begin{matrix} 1.0 & -0.2 & 2.0 & -1.0 \\ 0 & 1.0 & 0 & 0.1 \\ 0 & -0.5 & 1.0 & 0 \\ 0.4 & 0 & 0 & 1.0 \end{matrix}$$

$$B_4 = \begin{matrix} 1.0 & -0.2 & 2.0 & -1.0 \\ 0 & 1.0 & 0 & 0 \\ 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 1.0 \end{matrix}$$

$$B_5 = \begin{matrix} 1.0 & -0.2 & 2.0 & -1.0 \\ 0 & 1.0 & 0.5 & 0.1 \\ 0 & 0 & 1.0 & 0.2 \\ 0 & 0 & 0 & 1.0 \end{matrix}$$

3) A run consisted of 100 samples of 20 observations each. Coefficients for the first equation were estimated from each of the hundred samples in each run. Each sample utilizes a new set of normal deviates.

4) One run was obtained for each model-data set B-matrix combination in conjunction with Σ_1, giving a total of twenty runs. In addition, four more runs were obtained by using each of the four possible model-data set combinations in conjunction with B_1 and Σ_2. The four runs which are obtained for each (B, Σ) combination are designated as logical sets. Thus the four runs from (B_1, Σ_2) are Logical Set 0, the runs from (B_1, Σ_1) Logical Set 1, up to the runs from (B_5, Σ_1) designated Logical Set 5. The set of normal deviates is re-used in each run.

5) k-class estimates were obtained in the usual manner. Writing the first equation as

$$y_1 = \beta_{12} y_2 + \beta_{13} y_3 + \beta_{14} y_4 + \gamma_{11} z_1 + \gamma_{12} z_2 + \gamma_{13} z_3 + u_1 ,$$

or more briefly,

$$y_1 = \beta Y + \gamma Z_* + u_1 , \tag{2.15}$$

we obtain the estimates for $(\beta_{12}, \beta_{13}, \beta_{14}, \gamma_{11}, \gamma_{12}, \gamma_{13})$, denoted briefly as $(\hat{\beta}, \hat{\gamma})$,

$$\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} Y'Y - kV'V & Y'Z_* \\ Z_*'Y & Z_*'Z_* \end{bmatrix}^{-1} \begin{bmatrix} Y' - kV' \\ Z_*' \end{bmatrix} y_1 \tag{2.16}$$

where V represents the matrix of residuals from the regressions of y_1, y_3 and y_4, on all the exogenous variables.

The results of this study can be summarized briefly as follows:

a) Two-stage least squares estimates are not unambiguously better than direct least squares estimates in small sample situations.

b) Estimates are relatively poor when there is high multicollinearity among exogenous variables, with two-stage least squares being relatively more affected.

c) The distribution of two-stage least squares has higher density than direct least squares in some neighborhood of the true value, but it also has thicker tails.

d) Estimates generally improve as the sparseness of the B-matrix increases.

e) Triangularity of the B-matrix improves estimates, and recursiveness is thus more favorable to good estimation than an equivalent amount of sparseness; estimates generally improve as the sparseness of the covariance matrix Σ increases; the improvement being somewhat ambiguous for $k = 1$.

f) Estimates are on the whole better in the overidentified model than in the just identified one.

g) k^* - estimates may be considered rational alternatives to both two-stage least squares and direct least squares.

The question of what estimator to use in a concrete small sample situation is far from settled. Although ultimately the user will always have to state the kinds of risk he wishes to avoid, a great deal more needs to be known about the properties of various estimators. It appears likely on the basis of this and other investigations that a considerable ambiguity and uncertainty will continue to adhere to rankings of estimating techniques.

Craig (1966) describes the results of a series of sampling experiments designed to investigate how sensitive the performances of various estimators are to violations of the assumptions of the usual model. The estimators investigated are direct (or "classical") least squares (DLS), two-stage least squares (2SLS), Nagar's unbiased k-class estimator (UBK), limited-information maximum likelihood (LIML), three-stage least squares (3SLS), and full-information maximum likelihood (FIML). The first four methods are all members of the k-class of estimators and estimate the coefficients simultaneously. Standard errors of the estimates of the last five techniques are available, based on the asymptotic distributions of the estimates in the usual model; standard errors can be computed for DLS by the usual formulae.

The usual simultaneous equation model can be summarized as follows. Let z_t be the $K \times 1$ vector of predetermined variables at observation $t (t = 1, \ldots, T)$. z_t is assumed to be nonstochastic (and measured without error). Let B and Γ be $G \times G$ and $G \times K$ matrices of (non-stochastic) structural coefficients, some of whose elements are specified a priori to be zero. Let u_t be a $G \times 1$ vector of stochastic, structural disturbances of observation t. The $G \times 1$ vector of endogenous variables, y_t, is assumed to be generated:

$$By_t = \Gamma z_t + u_t \,. \qquad (2.17)$$

The stochastic assumptions usually employed are:

$$\begin{aligned} E(u_t) &= 0 \quad \text{all } t \,; \\ E(u_t \; u_m') &= \Sigma \quad t = m \; (t,m = 1, \ldots, T) \,; \\ &= 0 \quad t \neq m \\ |\Sigma| &\neq 0 \end{aligned} \qquad (2.18)$$

Four types of violation of the assumptions of the usual model were investigated. These were (1) errors of measurement in the exogenous variables, (2) stochastic coefficients, (3) heteroskedastic disturbances and (4) autocorrelated distrubances. A sampling experiment by Ladd (1956) investigated the effects of errors of measurement on DLS and LIML. Otherwise these subjects have not been investigated before by sampling experiments.

Each sampling experiment involved the following steps. A set of twenty observations of the exogenous and endogenous variables were generated. From these observations the estimates of the structural coefficients were computed and the standard errors of these estimates calculated. This process was replicated fifty times. The fifty estimates of each coefficient by each estimator were then used to draw inferences about the small-sample distributions of the estimators in that experiment and were compared with the estimates of other experiments to investigate the effects of different ways of generating the data.

The structure employed in most experiments, structure 1, was:

$$B = \begin{matrix} 1 & -.89 & -.16 \\ -.54 & 1 & 0 \\ 0 & -.29 & 1 \end{matrix} \qquad (2.19)$$

$$\Gamma = \begin{matrix} 44 & .74 & 0 & 0 & .13 & 0 & 0 \\ 62 & 0 & .70 & 0 & .96 & 0 & .06 \\ 40 & 0 & .53 & .11 & 0 & .56 & 0 \end{matrix}$$

$$\Sigma = \begin{matrix} 35.24 & 34.48 & 31.12 \\ 34.48 & 36.68 & 29.84 \\ 31.12 & 29.84 & 40.65 \end{matrix} \qquad (2.20)$$

Coefficients whose values were zero or unity were not estimated, these values were assigned to them a priori in estimating the structure. Each equation has two more a priori restrictions than are needed to identify it. The first column of Γ is the constants of the equations and the first element of each z_t vector was always unity. Another structure, structure 2, used in two experiments investigating violations of the assumptions had the same values for B and Γ. Its Σ matrix was:

$$\begin{matrix} 29.24 & 3.32 & -1.24 \\ 3.32 & 36.20 & 4.96 \\ -1.24 & 4.96 & 46.60 \end{matrix} \qquad (2.21)$$

The exogenous variables, in the absence of special features, were independently and rectangularly distributed pseudo-random numbers lying between zero and one hundred. The same values for the exogenous variables were used in each replication of an experiment. The actual simple correlations between these variables were:

	z_3	z_4	z_5	z_6	z_7	
z_2	-.15	-.11	-.18	-.37	-.18	
z_3		-.18	-.01	.27	-.03	
z_4			-.15	-.37	.02	
z_5				-.09	.05	
z_6					.56	(2.22)

(z_1 was a dummy variable taking on the value unity at all observations).

The structural disturbances in the absence of special features, were normally and independently distributed random deviates of means zero and variance-covariance matrix Σ , as specified by the structure investigated. The disturbances were formed by taking linear combinations of independent normal deviates of unit variance. Different disturbances were used in different replications of the experiments, but the same basic variates were used in all the different experiments. Having formed the exogenous data and the disturbances, the endogenous variables were then generated. The reduced-form disturbances $(B^{-1}u_t)$ accounted for from eight to twelve per cent of the variances of the endogenous variables in structure 1 and from five to six per cent in structure 2 when there were no special features. The ways in which these basic data were altered to introduce the violations of the usual assumptions are described in the paper.

Gregg summarised his findings as follows. Only one of the violations of the assumptions of the usual model which were studied produced any pronounced changes in the performances of the estimators. Making the coefficients stochastic led to finding several median deviations which were significant. In addition to greatly increasing the dispersions, it changed the ranking of the estimators. The surprising thing about the other experiments was not the changes that occurred but the failure of the violations of the assumptions to alter the performances of the estimators greatly. Errors in the exogenous variables had little effect on the central tendencies of the estimates; heterskedastic and autocorrelated disturbances did not destroy the usefulness of the standard errors. In all the experiments the differences between the estimators were not very great. No method performed either a great deal better or worse than another.

The data and structures used in the experiments can at best be regarded only as very over-simplified representations of economic processes. They did not deal at all with the type of dynamic structure found in most economic models. In consequence, it would be dangerous to try to generalize the findings of the experiments to the performances of the estimators in econometric models. However, despite this shortcoming, the fact that the performances of the estimators were not greatly sensitive to the violations of the assymptions of the usual model is encouraging for the feasibility of using these estimators in empirical investigations.

Cragg (1966) describes the results of a series of sampling experiments designed to investigate how sensitive the performances of various estimators are to violations of the assumpsions of the usual model. The estimators investigated are direct (or "classical") least squares (DLS), two-stage least squares (2SLS), Nagar's unbiased k-class estimator (UBK), limited-information maximum likelihood (LIML), three-stage least squares (3SLS), and full-information maximum likelihood (FIML). The first four methods are all members of the k-class of estimators and estimate the coefficients simultaneously. Standard errors of the estimates of the last five techniques are available, based on the asymptotic distributions of the estimates in the usual model; standard errors can be computed for DLS by the usual formulae.

The usual simultaneous-equation model can be summarized as follows. Let z_t be the $K \times 1$ vector of predetermined variables at observation $t (t = 1, \ldots, T)$. z_t is assumed to be nonstochastic (and measured without error). Let B and Γ be $G \times G$ and $G \times K$ matrices of (nonstochastic) structural coefficients some of whose elements are specified a priori to be zero. Let u_t be a $G \times 1$ vector of stochastic, structural disturbances of observation t. The $G \times 1$ vector of endogenous variables, y_t, is assumed to be generated.

$$By_t = \Gamma z_t + u_t \quad . \tag{2.23}$$

The stochastic assumptions usually employed are:

$$\begin{aligned} E(u_t) &= 0 && \text{all } t; \\ E(u_t u_m') &= \Sigma && t = m \ (t,m = 1, \ldots, T); \\ &= 0 && t \neq m \\ |\Sigma| &\neq 0 \end{aligned} \tag{2.24}$$

Four types of violation of the assumptions of the usual model were investigated. These were:

1) errors of measurement in the exogenous variables,
2) stochastic coefficients,
3) heteroskedastic disturbances and
4) autocorrelated disturbances.

A sampling experiment by Ladd (1956) investigated the effects of errors of measurement on DLS and LIML. Otherwise these subjects have not been investigated before by sampling experiments.

Each sampling experiment involved the following steps. A set of twenty observations of the exogenous and endogenous variables were generated. Form these observations the estimates of the structural coeffiecients were computed and the standard errors of these estimates calculated. This process was replicated fifty times. The fifty estimates of each coefficient by each estimator were then used to draw inferences about the small-sample distributions of the estimators in that experiment and were compared with the estimates of other experiments to investigate the effects of different ways of generating the data.

The structure employed in most experiments, structure 1, was:

$$B = \begin{matrix} 1 & -.89 & -.16 \\ -.54 & 1 & 0 \\ 0 & -.29 & 1 \end{matrix} \qquad (2.25)$$

$$\Gamma = \begin{matrix} 44 & .74 & 0 & 0 & .13 & 0 & 0 \\ 62 & 0 & .70 & 0 & .96 & 0 & .06 \\ 40 & 0 & .53 & .11 & 0 & .56 & 0 \end{matrix}$$

$$\Sigma = \begin{bmatrix} 35.24 & 34.48 & 31.12 \\ 34.48 & 36.68 & 29.84 \\ 31.12 & 29.84 & 40.64 \end{bmatrix} \qquad (2.26)$$

Coefficients whose values were zero or unity were not estimated, these values being assigned to them a priori in estimating the structure. Each equation has two more a priori restrictions than are needed to identify it. The first column of Γ is the constants of the equations and the first element of each z_t vector was always unity. Another structure, structure 2, used in two experiments investigating violations of the assumptions had the same values for B and Γ. Its Σ matrix was:

$$\begin{matrix} 29.24 & 3.32 & -1.24 \\ 3.32 & 36.20 & 4.96 \\ -1.24 & 4.96 & 46.60 \end{matrix} \qquad (2.27)$$

The exogenous variables, in the absence of special features, were independently and rectangularly distributed pseudo-random numbers lying between zero and one hundred. The same values for the exogenous variables were used in each replication of an experiment. The actual simple correlations between these variables were:

	z_3	z_4	z_5	z_6	z_7
z_2	-.15	-.11	-.18	-.37	-.18
z_3		-.18	-.01	.27	-.03
z_4			-.15	-.37	.02
z_5				-.09	.05
z_6					.56

(2.28)

(z_1 was a dummy variable taking on the value unity at all observations).

The structural disturbances, in the absence of special features, ware normally and independently distributed random deviates of means zero and variance-covariance matrix Σ , specified by the structure investigated. The distrubances were formed by taking linear combinations of independent normal deviates of unit variance. Different disturbances were used in different replications of the experiments, but the same basic variates were used in all the different experiments. Having formed the exogenous data and the disturbances, the endogenous variables were then generated. The reduced-form disturbances $(B^{-1} u_t)$ accounted for from eight to tewlve per cent of the variances of the endogenous variables in structure 1 and from five to six per cent in structure 2 when there were no special features. The ways in which these basic data were altered to introduce the violations of the usual assumptions are described in the paper.

Gragg summerized his findings as follows. Only one of the violations of the assumptions of the usual model which were studied produced any pronounced changes in the performances of the estimators. Making the coefficients stochastic led to finding several median deviations which were significant. In addition to greatly increasing the dispersions, it changed the ranking of the estimators. The surprising thing about the other experiments was not the changes that occurred but the failure of the violations of the assumptions to alter the performances of the estimators greatly. Errors in the exogenous variables had little effect on the central tendencies of the estimates; heteroskedastic and autocorrelated disturbances did not destroy the usefulness of the standard errors. In all the experiments the differences between the estimators were not very great. No method performed either a great deal better or worse than another.

The data and structures used in the experiments can at best be regarded only as very over-simplified representations of economic processes. They did not deal at all with the type of dynamic structure found in most economic models. In consequence, it would be dangerous to try to generalize the findings of the experiments to the performances of the estimators in econometric models. However, despite this shortcoming, the fact that the performances of the estimators were not greatly sensitive to the violations of the assymptions of the usual model is encouranging for the feasibility of using these estimators in empirical investigations.

Cragg (1967) reported the results of another Monte Carlo study. He used the same model of structural equations for this study as is described above.

He also employed the same methods of estimation as is reported in earlier papers. However, his concern in this study is concentrated to observe the effect of multicollinearity on the estimates obtained from various methods.

Most of the experiments used structures from a three-equation model of the form:

$$\begin{bmatrix} 1 & \beta_{12} & \beta_{13} \\ \beta_{21} & 1 & 0 \\ 0 & \beta_{32} & 1 \end{bmatrix} y = \begin{bmatrix} \gamma_{11} & \gamma_{12} & 0 & 0 & \gamma_{15} & 0 & 0 \\ \gamma_{21} & 0 & \gamma_{23} & 0 & \gamma_{25} & 0 & \gamma_{27} \\ \gamma_{31} & 0 & \gamma_{33} & \gamma_{34} & 0 & \gamma_{36} & 0 \end{bmatrix} z + u \qquad (2.29)$$

where the β's and the γ's represent coefficients to be estimated. The first column of the Γ matrix contains the constants of the equations, and the first element of the z vector at every observation was unity. Each equation has two more a priori restrictions imposed on it than are necessary to identify it. The coefficients of the structures of the model which were used in the experiments are shown in Table I. Structures 1 through 5 had different values for the structural coefficients given in the paper (page 92, Table I) but the same values for Σ. Structures 6, 7, and 8 used the structural coefficients of structure 1 and different values for Σ.

The experiments gave a number of interesting findings:

a) The differences in the performances of the methods are not at all pronounced.

b) The frequency with which one estimator came closer to the true values of the structural coefficients than another varied significantly over the different coefficients of an experiment. The variation was sufficient to make the rankings of the estimators for the different coefficients different.

c) The frequencies with which one method came closer to the true values of the structural coefficients, than another depended on the exact set of exogenous data used, the true values of the structural coefficients, the correlations between the structural disturbances and the sizes of the structural disturbances. This variability could change the rankings of the estimators for some coefficients.

d) Given the small differences among the estimators and the variability in their relative performances, DLS was usually the poorest method and 3SLS and FIML were better than 2SLS, UBK and LIML.

e) In most cases differences in the central tendencies of the distributions of the consistent estimators from the true values of the coefficients were not very serious, but large disturbances and multicollinearity could change this conclusion. On this criterion, FIML and LIML seemed slightly superior to other methods. The differences of its medians from the true values was a serious problem for DLS. This feature, rather than wide dispersions, was the reason for the poor rankings of DLS. It weighed more heavily against DLS when larger samples were used.

f) Usually use of the standard errors of the consistent methods would lead to reliable inferences, but this was not always the case. The standard errors of DLS were not useful for making inferences about the true values of the coefficients.

The experiments conducted give no clear guide lines for the choice of an estimator for econometric models. They indicate that the ambiguities to be found in earlier sampling experiments genuinely reflect properties of the simultaneous-equation estimators. The results suggest that, because the consistent estimators do not differ greatly and their relative performances are sensitive to the data and structure studied, 2SLS may well be the best estimator to choose since it is the cheapest and easiest method to compute. The choice of DLS also may be sensible, even for very simple models conforming to the assumptions under which the simultaneous-equation estimators were derived and the experiments conducted.

Gragg (1968) provides some sampling experiment results exploring the effects of certain misspecifications on several simultaneous equation estimators. Again he has used the same model with the same specification coefficients and variance-covariance matrix and has employed the same methods of estimation.

He has specified the following types of misspecification of the model:

1. Specification that important coefficients be zero.
2. An unidentifiable structure.
3. Omission of an exogenous variable.
4. Omission of an equation.
5. Failure to specify all zero coefficients.

He has reported the results of the experiment investigating misspecification by comparing with the results of the experiments in which the structure was correctly specified. He summarizes the results of this finding as follows.

As one might expect, the effects of misspecification very much depended on the seriousness of the mistake made, in the context of the structure studied. The specification that coefficients whose true values were almost zero be zero had little effect on the performances of the estimators. Ignoring an unimportant exogenous variable and omitting a structural equation affected the estimators only slightly. These results suggest that basing the specification of models on only approximate or partial knowledge may not seriously compromise the usefulness of the models.

On the other hand, serious failures in specification had pronounced effects on the estimators. The specification that important non-zero coefficients be zero altered the central tendencies of the estimates of the misspecified equations drastically and markedly increased the dispersions of the estimators. For most purpsoes, the effect of the misspecification was probably enough to render the estimates of the structural coefficients useless. In the face of this misspecification, the consistent k-class estimators seemed to be better than FIML and 3SLS in estimating correctly specified equations within the system. Failing to specify as zero all coefficients for which this was correct had serious effects on the central tendencies and dispersions of the estimators when knowledge of these coefficients was important for the identification of an equation. FIML seemed the estimator most sensitive to this danger. These findings suggest that, in the absence of fairly complete and confidently held knowledge about which coefficients in a structure are very small or zero, successful econometric model-building faces a very difficult task in steering between the Sylla of misspecification and the Charybdis of under-specification.

Summers (1965) has carried out extensive investigations in studying small sample properties of econometric estimators. He employed the following hypothetical economic model:

$$y_{1t} + \beta_{12} y_{2t} + \gamma_{11} Z_{1t} + \gamma_{12} Z_{2t} + \gamma_{10} = u_{1t} \qquad (a)$$
$$y_{1t} + \beta_{22} y_{2t} + \gamma_{23} Z_{3t} + \gamma_{24} Z_{4t} + \gamma_{20} = u_{2t} \qquad (b) \qquad (2.30)$$

The y's are jointly determined variables while the z's are predetermined; u's are bivariate normal variables with zero mean and a variance covariance matrix denoted by Σ . The reduced form of the model is given as follows:

$$y_{1t} = \pi_{11} z_{1t} + \pi_{12} z_{2t} + \pi_{13} z_{3t} + \pi_{14} z_{4t} + \pi_{10} v_{1t} \qquad (c)$$
$$y_{2t} = \pi_{21} z_{1t} + \pi_{22} z_{2t} + \pi_{2t} z_{3t} + \pi_{24} z_{4t} + \pi_{20} v_{2t} \qquad (d) \qquad (2.31)$$

The π's are functions of the β's and γ's and the v's are linear combinations of the u's . The absence of z_3 and z_4 from equation (a) and the absence of z_1 and z_2 from (b) imply certain interdependencies among the π's in (c) and (d).

The equations (a) and (b) were used as the basic model of a hypothetical model. In any particular experiment the values of all of the parameters of the model, both the structural coefficients and the variance-covariance matrix of structural disturbances, were specified. The values taken by all of the predetermined variables were specified of each of the T observations periods considered.

(b) Using the random sampling method*, T observations were generated on both y_1 and y_2 , conditioned by the set of values of the

* A random normal deviate generator was used to produce structural disturbances which were then transformed into reduced form disturbances. In actual computing these two processes were combined into one step.

predetermined variables. The T sets of observations on y_1, y_2, z_1, z_2, z_3, and z_4 constitute a sample.

(c) Limited information, single equation (LISE), Two stage least squares (2SLS), ordinary least squares (OLS) and full information maximum likelihood (FIML) were each applied to the sample to estimate the parameters.

(d) Least squares no restriction (LSNR) was applied to the sample to obtain: i) estimates of the reduced form coefficients π's

ii) conditional predictions of y_1 and y_2, for the same prescribed set of z's as in (c)

iii) these predictions are based on the estimates of the π's.

(e) Steps (b), (c) and (d) were then repeated N times to get N different estimates of each of the parameters by each estimating method.

(f) The N different estimates of each parameter obtained for each estimating method were organised into a relative frequency distribution and summary measures were computed for each distribution.

(g) Then a new specification of the values of the parameters was made and steps (b), (c), (d), (e) and (f) were repeated. A variety of specifications was made in order to observe the sensitivity of the relative frequency distributions to the specification of the characteristics of the model.

Five different sets of parameters were used; four were used with a single sample size, T = 20 but the fifth was used with two sample sizes T = 20 and T = 40. If T is thought of as a parameter there were six distinct parameter constellations. Each was used twice, once in an A experiment and once in a B experiment. The A experiments differ from the B ones in the extent to which the predetermined variables were specified to be intercorrelated.

For each set of pseudo-economic data the structural coefficients, the β's and γ's, were estimated by each of the estimating methods. The estimates of the π's were computed as functions of β's and γ's. The conditional predictions of y_1 and y_2 were made on the basis of reduced form coefficient for each of the methods. These were then compared with the population values of y_1 and y_2 as determined by the structural parameters specified in the experiment. A comparison of the absolute difference was then made for each of the different estimating methods. Within each sampling experiment this was repeated fifty times, the number of sets of pseudo-economic data generated per experiment.

Suppose the parameter being estimated is α and an estimate of α produced by the kth estimating method is α^k. Associated with each method is a frequency function $f_k(\alpha^k | \alpha)$. Some measure of dispersion of f_k around α is called for; clearly the estimating method with the frequency function with the smallest dispersion is best. The measure most commonly used for this purpose is the root mean square (RMSE). Alternatives are the mean absolute error (MAE) and $P = \text{Prob}\{|\hat{\alpha} - \alpha| < \delta\}$.

In this case the bias, standard deviation and RMSE are calculated for f_k of the twelve sampling experiments. The bias and variance measures are not very interesting in themselves but they give an idea of the relative sizes of the two components of the RMSE.

Finally a test of hypothesis is designed to test that there is no difference between the methods. The test $H_o: \text{MAE}_{\alpha^k} = \text{MAE}_{\alpha^\ell}$ is equivalent to the null hypothesis $H_o' = P_{k\ell} = \frac{1}{2}$ where

$P_{k\ell} = P\{|\alpha^k - \alpha| > |\alpha^\ell - \alpha|\}$. Since α^k and α^ℓ were highly correlated in virtually every case and therefore were nearly always on the same side of α, $P_{k\ell}$, the proportion of cases in which $|\alpha^k - \alpha|$ exceeded $|\alpha^\ell - \alpha|$ was extremely easy to obtain by visual inspection. The power efficiency of this test is about 2/3. The use $P_{k\ell}$ as a test statistic circumvents the difficulty of measuring the dispersion of f_k with moments which may not be well behaved.

The following conclusions were drawn.

1. The minimum variance property of large sample OLS structural coefficient estimates certainly is preserved for small samples.

2. The structural coefficient bias of OLS was by far the greatest of the four methods examined.

3. For large samples the standard errors of estimates produced by each method are inversely proportional to the square root of the sample size.

4. The difference between the best estimating method and the worst was substantial in virtually every experiment.

5. In A experiments FIML was clearly best with 2SLS (TSLS) second and LISE and OLS last. In B experiments 2SLS passed up FIML and OLS passed up LISE. FIML performed badly in the four misspecification experiments.

Mikhail (1972) points out that over the last few years Monte Carlo studies conducted to simulate the small sample properties of simultaneous equation estimators have used straight forward simulation and produced results which were often indeterminate and sometimes contradictory. This is to a great extent, due to sampling error which can be quite high depending on the number of replications and the way the simulation process is carried out. In his study he has used "Antithetic variates" in simulating the finite sample properties of econometric estimators.

In previous studies, the probabilistic Monte-Carlo method used was a straight forward simulation of the problem which he calls "Direct Simúlation" meaning that no refinement is exercised in the choice and use of the random numbers. Suppose now δ' is a Monte-Carlo estimator of an unknown parameter δ. The basic idea of the two anti-thetic method is to seek another statistic δ'' which is negatively correlated with δ' and whose expectation is the same as the expectation of δ'. If in direct simulation we use a set of random numbers e_i uniformly distributed between zero and one, we can now use the set $e_i^* = 1 - e_i$ to get an antithetic estimator δ'' which is likely to be negatively correlated with δ'.

The information obtained from the two mutually antithetic estimate is then combined to give a better knowledge of the parameters of the distributions, e.g.; if the correlation coefficients between the biases of the two estiamtes is less than zero the average of the two biases gives a better estimate of the bias than does either of them.

The following model is used in this study.

$$YA = XB + U$$

where Y is the $T \times 2$ matrix of endogenous variable, X is a $T \times 7$ matrix of exogenous variable x_7 being identically equal to unity in all observations. A and B are 2×2 and 7×2 matrices of coeffecients and U is the $T \times 2$ matrix of disturbances.

The comparisons are based on estimates of the bias, variance mean square error and mean absolute error obtained by the method of 2SLS. From his Table I it can be easily seen that the two antithetic method does much better than direct simulation in estimating bias. The standard error of the biases are reduced by a factor which varies between 4 and 8. In estimating the variance and the mean absolute error the two anthetic method does not show much superiority over the direct method of simulation.

B. Raj (1980) considered four alternative forms of two parameter error distributions:

(a) normal,

(b) uniform,

(c) lognormal,

(d) Laplace or double exponential

and reported on a Monte-Carlo study of the small sample properties of least squares, two stage least squares, three stage least squares, full Information Maximum Likelihood estimators. The hypothetical Model used for the Monte Carlo simulations has two behavioral equations and an identity as given below.

$$-y_1 + \beta_{21} y_2 + \gamma_{11} x_1 + \gamma_{21} x_2 + u_1 = 0 \qquad (a)$$

$$\beta_{12} y_1 - y_2 + \beta_{32} y_3 + \gamma_{22} x_2 + \gamma_{32} x_3 + \gamma_{42} x_4 + u_2 = 0 \qquad (b) \qquad (2.32)$$

$$y_1 - y_2 - y_3 + x_3 + x_5 + x_6 = 0 \qquad (c)$$

where y's are endogenous variables, x's are exogenous variables and u's are random disturbances. The model is obtained by adding an identity an endogenous variable and two exogenous variables to the over identified model considered by Summers (1965). It is clear that the presence of identity means that the covariance matrix of model disturbances is singular. Now if we substitute the identiyt in (b), we obtain

$$\frac{\beta_{12} + \beta_{22}}{1 + \beta_{32}} y_1 - y_2 + \frac{\gamma_{22}}{1 + \beta_{32}} x_2 + \frac{\gamma_{32} + \beta_{32}}{1 + \beta_{32}} x_3 + \frac{\gamma_{42}}{1 + \beta_{32}} x_4$$

$$+ \frac{\beta_{32}}{1 + \beta_{32}} x_5 + \frac{\beta_{32}}{1 + \beta_{32}} x_6 + \frac{1}{1 + \beta_{32}} u_2 = 0 \ .(d) \qquad (2.33)$$

Now the model composed of (a) + (d) have positive definite covariance matrices of structural and reduced form disturbances so the analytical results hold in the context of (a).

The Monte-Carlo experiments employed two sets of generated samples of 20 observations each. The observations on six exogenous variables used in the simulations were independent random drawings from the uniform distribution in the range -17.321 to 17.321. The set of 20 observations on the six exogenous variables was kept fixed in repeated samples. The reduced form disturbances were generated in the following way:

(a) A set of 1000 samples of size 20 of independently distributed uniform random numbers between 0 and 1 was generated.

(b) The set of independently distributed uniform random numbers was transformed into a set of independently distributed normal or non-normal variates with zero mean and unit variance by the following transformations.

1. Normal variate: The set of uniform random numbers between zero and one was transformed into the set of standard normal variate by

$$e_1 = (-2 \log w_1)^{\frac{1}{2}} \cos 2\pi\, w_2$$

$$e_2 = (-2 \log w_2)^{\frac{1}{2}} \cos 2\pi\, w_1 \qquad (2.34)$$

where w's are uniform random drawings between zero and one and e's are independent standard normal variates.

2. Uniform variates:

$$e_i = a + (b - a)w_i \quad \text{for all } i$$

where w's uniform random numbers between zero and one and e's are independent standard uniform variates.

3. Lognormal variates:

$$e_i = \exp\{u + \sigma(\sum_{j=1}^{12} w_{ij} - 6)\} \tag{2.35}$$

where w's are uniform random numbers between zero and one and e's are lognormal variates with mean $\exp(u + \frac{\sigma^2}{2}) = 1$ and variance $(\exp \sigma^2 - 1) = 1$.

4. Laplace or double exponential variates:

$$e_i = -\ln w_i \quad \text{for all } i \tag{2.36}$$

where w's are uniform random numbers and e's after a random sign has been attached are standard double exponential variates.

(c) The set of independently and identically distributed variates with zero mean and unit variance was transformed into a set of reduced form disturbances, V_i with zero mean and moment matrix Ω_1 or Ω_2 via $V_i = E P_i$ where P_i is a $G \times G$ lower triangular matrix such that $P_i P_i = \Omega_i$ and E is a $T \times G$ matrix containing independently distributed normal or non-normal variates.

The output of 1000 replications of size 20 consists of two sets of sampling distributions:

(a) The estimates of eight structural parameters

(b) The predications of the mean "future" value of each endogenous variable.

The median was computed for sampling distributions of each structural coefficient and of the predicted value of the mean of each endogenous variable. This measure of location was used for claculating the percentage median bias, the difference between the calculated median and the true value of structural coefficient forecast expressed as a percentage of the true absolute value of structural coefficient/forecast. The dispersion of each sampling distribution around its median was calculated and expressed as a percentage of the true value of the structural coefficient/forecast. This measure was termed the percentage quartile deviation (PQD). Finally, an adhoc measure which is the sum of PQD and the absolute value of PMBIAS was calculated. This measure was termed the percentage quartile deviation median bias PQDMB.

Similarly the mean of the sampling distribution of each structural coefficient and of the forecast value of each endogenous variable. This measure of location was used to obtain the percentage mean bias (PBIAS). The dispersion of each sampling distribution around the mean was also obtained and was expressed as a percentage of the square of the true value of the structural coefficient/forecast. This measure was termed as the percentage variance (PVAR). Finally the dispersion of each sampling distribution around the true value of the parameter/forecast was obtained and expressed as a percentage of the square of the true value of the structural coefficient/forecast. This was termed as the percentage mean squared error (PMSE).

Conclusions:

Estimators of structural coefficient:

(1) The LS estimator is more biased than the 2SLS.

(2) The LS and 2SLS estimators of all positive structural coefficients are negatively biaed for all four error distributions.

(3) The FIML estimator is the least biased while LS is the most biased of the four error distributions, etc.

Predictors of the means of endogenous variables:

(1) FIML is the least biased while LS is the most biased predictor of $E(y_1)$ with respect to the criteria used for all four error distributions.

(2) The choice of alternative predictors of $E(y_2)$ for alternative measures of bias and efficiency is not clear in the first experiment. The FIML predictor of $E(y_2)$ is least biased and most efficient for all four error distributions.

(3) The poor performance of 3SLS is surprising.

CHAPTER III

THE EXPERIMENT

MODEL AND DATA

The economic model used for the Monte Carlo simulations has three behavioral equations and three identities known as the Klein model. For the description of this model, we shall follow Theil (1971) text:

The first equation of the model is the consumption function, which is written as:

$$C_\alpha = \beta_0 + \beta_1 P_\alpha + \beta_2 P_{\alpha-1} + \beta_3 (W_\alpha + W'_\alpha) + \varepsilon_\alpha \tag{3.1}$$

where C_α is the aggregate consumption in year (α), P_α the total profits of that year, $W_\alpha + W'_\alpha$ the total wage bill. So, this equation describes the aggregate consumption linearly in terms of the total wage bill (wage bill paid by private industry plus government wage bill), of the same year $(W_\alpha + W'_\alpha)$, of the current profits P_α and of profits lagged one year $P_{\alpha-1}$, apart from a random disturbance ε_α. All variables of the model are measured in million dollars of 1970.

The second equation of the model is:

$$I_\alpha = \beta'_0 + \beta'_1 P_\alpha + \beta'_2 P_{\alpha-1} + \beta'_3 K_{\alpha-1} + \varepsilon'_\alpha \tag{3.2}$$

where I_α is the net investment in year (α), K_α is the stock of capital goods at the end of the year and $K_{\alpha-1}$ is the stock of capital at the beginning of the year.

Finally, the third equation expresses the demand for labor:

$$W_\alpha = \beta''_0 + \beta''_1 X_\alpha + \beta''_2 X_{\alpha-1} + \beta''_3 (\alpha - 1970) + \varepsilon''_\alpha \tag{3.3}$$

where X_α is the total production of private industry in year (α),

$X_{\alpha-1}$ with one-year lag and (α) should be regarded as the time measured in the calendar years.

There are also three definitional equations which are not subjected to the random error and have no unknown parameters:

$$X_{\alpha} = C_{\alpha} + I_{\alpha} + G_{\alpha} + R_{\alpha} \quad (3.4)$$

$$P_{\alpha} = X_{\alpha} - W_{\alpha} - T_{\alpha} \quad (3.5)$$

$$K_{\alpha} = K_{\alpha-1} + I_{\alpha} \quad (3.6)$$

where G_{α} is the government non-wage expenditure and T_{α} the business taxes in year (α). We modified the (3.4) identity of Klein's Model I, adding the variable R_{α}, which is caused to some statistical differences and to the change of reserves. This modification was considered absolutely necessary because the left and the right sides of (3.4) must balance. In order to succeed it, according to the Greek Tables which concern "the gross national expenditure and the gross product" we had to add the variable R_{α}.

There are now six-equations systems in the six endogenous variables: C_{α}, P_{α}, W_{α}, I_{α}, K_{α}, X_{α} and the eight predetermined variables (lagged: $X_{\alpha-1}$, $P_{\alpha-1}$, $K_{\alpha-1}$ and exogenous: α, W'_{α}, G_{α}, T_{α}, R_{α}).

The exogenous variables are by assumption independent of the operation of the system; also the lagged endogenous variables are at least independent of the current operation of the system if the vectors e_{α}, e'_{α}, e''_{α}, are stochastically independent.

Aiming to examine the small sample properties of the k-class estimators we shall use the consumption function of Klein's Model I and data from the Greek economy.

The consumption function can thus be written as:

$$C_1 = Z_1 \delta_1 + E_1 \qquad (3.7)$$

where:

$$C_1 = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_{21} \end{bmatrix}, \quad Z_1 = \begin{bmatrix} p_1 & w_1 + w_1' & 1 & p_0 \\ p_2 & w_2 + w_2' & 1 & p_1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p_{21} & w_{21} + w_{21}' & 1 & p_{20} \end{bmatrix}$$

$$\delta_1 = \begin{bmatrix} \beta_1 \\ \beta_3 \\ \cdots \\ \beta_0 \\ \beta_2 \end{bmatrix} \quad E_1 = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{21} \end{bmatrix}$$

and the subscript 21 refers to the last of the 21 annual observations of the period 1959 to 1979.

The generation of the disturbances was accomplished as follows:

A set of 100 samples of size 21 of independently distributed random numbers between zero and one was generated. This set of independently distributed uniform random numbers was transformed into a set of independently distributed normal or non-normal variates (see Raj, (1980)): "Specifically, the set of the uniform random numbers between zero and one was transformed into:

(a) the set of standard normal variates,

(b) the set of uniform variates,

(c) the set of lognormal variates with mean unit and variance unit.

In the following we present the conclusions of this study.

CONCLUSIONS

The experiment was conducted by the use of IBM 43 - 41 DOS-VSE and IBM 3031 VM/SP computer. Various aspects of the FROTRAN IV and experimental programs (IMSL, TSP version 71, TROLL version 10) were used, which are relevant to this study and they are presented in the Appendix B. While the IBM 43 - 41 DOS-VSE and IBM 3031 VM/SP are computers of large capacity and high speed, both the size of the computer and time consideration restricted the experiments which were conducted. These restrictions limited the study to the performances of the estimators with a small number of observations. Both time and space considerations limited the number of replications in each experiment to a hundred. The summary routines were written and tested in FORTRAN IV. Also, the necessary modifications were made in the form of the data and the results in order to be applicable for the TSP and TROLL packages or computer programs. Further OLS and 2SLS structural estimates were computed by TSP and TROLL, while UBK only by TROLL.

The accuracy of the data translation into the tape was tested many times. But input-output operations are more subject to mechanical failure than the computations.

In the generation of the disturbances we used $\sigma = 1500$ instead of $\sigma = 2520.35$, which we take from the considered formula. We have done it because, if we use $\sigma = 2520.35$, upper limit is lower than $k = 1$ and so the 2SLS and UBK estimators give negative estimates for the structural coefficients. Moreover, for $\sigma = 1500$ the UBK is the estimator which corresponds to the upper limit $(k = 1.047)$.

Investigation of the estimates of the experiment led to a number of findings.

1. Given the Tables C.1, C.2, C.3, the results are summarized as follows:

(a) The biases of each of the methods, for the different disturbance distributions, behave quite similarly. So, the choice of the least or the most biased estimator is not clear. The OLS estimator is the better in the sense that its bias is smaller than the bias which corresponds to the others, when the disturbances follow a particular form of the three distributions, while UBK estimator is the worst. The poor performance of UBK on the criterion of bias is surprising.

(b) The differences of the RMSE of SD's of each of the methods, for the alternative forms of disturbance distribution, are not great. The estimators have the lower RMSE or SD when the disturbances follow the lognormal distribution and generally, they have the highest RMSE or SD, when the disturbances follow the normal distribution. Also, the dispersions of the OLS estimator about the "means" are respectively, similar to their dispersions about the "true value", as might have been expected from the small baises that were found. The OLS estimator is more efficient in the sense that their RMSE's and SD's are lower than those corresponding to 2SLS and UBK estimators, when the disturbances follow a particular form of the three distributions, while UBK is the worst.

2. The values of $|\mathrm{Cov}\hat{\delta}_{21}|$ and $|\overline{\mathrm{Cov}\hat{\delta}}_{21}|$ are given in the Table C.22 (Appendix C). Using the Table C.22 we may rank the three estimators with respect to the measure in (II. 19). The results are summarized in the Tables C.23 to C.30. In the Tables C.29 and C.30

are including the ratios "e" of the "generalized variances", respectively, about the mean and "true" parameter vector. Also, these results agree with the other results, which we have obtained for the dispersions, in Section 1 of the conclusions. We note that in Tables C.29 and C.30 a (+) or (1) means that the denominator estimator is "efficient" relative to the nominator estimator.

3. In the Tables C.4 to C.21 we have tabulated some summary statistics for the endogenous variable, using the three estimators and the three forms of the disturbance distributions. Several points stand out in the Tables:

(a) The biases of the predictions of the endogenous variable are generally large. The OLS has the smallest bias, from the other estimators and the UBK is the most biased, for all the three disturbance distributions. The estimators are least biased, when the disturbances follow the uniform distribution and are most biased for the normal disturbances, except the OLS which is most biased for lognormal disturbances. The biased differences of each of the relative methods, for the different disturbance distributions, are not very great.

(b) The dispersions of the predicted "values" about the "true one" are large. Also, the differences in the relative performance of each estimator, when the disturbances follow the corresponding forms of the distributions are rather small. The OLS estimator has the smallest RMSE and UBK has the highest RMSE from the other estimators, for all the three disturbance distributions. On the other hand, the OLS estimator has the smallest RMSE, from the other estimators, using the uniform disturbance distribution and the highest RMSE using the lognormal disturbance distribution. Moreover the 2SLS estimator has the smallest RMSE, from the other estimators, when the disturbances follow the uniform distribution and the highest RMSE when the disturbances follow the normal distribution. Also, the UBK estimator has the smallest RMSE, from the other estimators, for lognormal disturbances and the highest RMSE for normal disturbances.

(c) The dispersions of the predicted "values" about their means, generally, are not so great as the dispersions about the "true value". Especially, the OLS estimator has the lowest SD and UBK estimator has the highest SD, for all the three disturbance distributions. Finally, the three estimators have the lowest SD using the lognormal disturbance distribution and the highest SD, when the disturbances follow the normal distribution.

BIBLIOGRAPHY

1. Aubry, J.P. (1973). A Monte Carlo Study of the Estimation of an Overidentified Model with Temporary Dependent Residuals. Technical Report I. Bank of Canada.

2. Carter, R.A.L. (1971). Small Sample Properties of Three Econometric Methods of Estimation. Doctoral Thesis, Queen's University. Canada.

3. Carter, R.A.L. (1978). Monte Carlo Methodology and Monte Carlo Results. Unpublished Manuscript, Chapters II and XII, University of Western Ontario.

4. Cragg, J.G. (1966). On the Sensivity of Simultaneous Equation Estimators to the Stochastic Assumptions of the Models. Journal of the American Statistical Association, 61, 136-51.

5. Cragg, J.G. (1967a). Small Sample Performances of Various Simultaneous Equation Estimators in Estimating the Reduced Form. Metroeconomica 19, 77-93.

6. Cragg, J.G. (1967b). On the Relative Small-Sample Properties of Several Structural-Equation Estimators. Econometrica, 35, 89-110.

7. Cragg, J.G. (1968). Some effects of Incorrect Specification on Small-Sample Properties of Several Simultaneous-Equation Estimators. International Economic Review, 9, 63-86.

8. Dwivedi, T.D. (1980). A Mathematical Identity Between Double k-class, Two Stage Least Squares and Ordinary Least Squares. Memeograph. Department of Mathematics, Concordia University.

9. Hendry, D.F. and Harrison, R.W. (1974). Monte Carlo Methodology and the Small Sample Behaviour of Ordinary and Two-Stage Least Squares. Journal of Econometrics, 2, 151-174.

10. King, D.R. (1974). Feedback and Reciprocal Causation: a Monte Carlo Study of the Relative Small Sample Properties of Simultaneous Equation Estimators. Doctoral Thesis. University of North Carolina.

11. McNown, R.F. (1971). A Simulation Study of the Small Sample Properties of Several Distributed Lag Estimators. Doctoral Thesis. University of Californai at San Diego.

12. Mikhail, W.M. (1972). Simulating the Small-Sample Properties of Econometric Estimators. Journal of the American Statistical Association, 67, 620-624.

13. Mosbaek, E.J. and Wold, H.O. (1969). Independent Systems-Structure and Estimation. North-Holland, Amsterdam.

14. Nagar, A.L. (1960). A Monte Carlo Study of Alternative Simultaneous Equation Estimators. Econometrica, 28, 573-590.

15. Neiswagner, W.A. and Yancey, T.W. (1959). Parameter estimates and autonomous growth. Journal of the American Statistical Association, 54, 389-402.

16. Pepin, R.A. (1971). A Monte Carlo Study of the Small Sample Properties of Several Simultaneous Equation Estimating Techniques when Statistical Problems are Present. Doctoral Thesis. Clark University.

17. Quandt, R.F. (1965). On Certain Sample Properties of k-class Estimators. International Economic Review, 6, 92-104.

18. Raj, B. (1980). A Monte Carlo Study of Small-Sample Properties of Simultaneous Equation Estimators with Normal and Non-normal Disturbances. Journal of the American Statistical Association, 75, 221-229.

19. Sharma, J.K. (1965). The Application of Monte Carlo Methods to the Evaluation of Small Sample Properties of Three-Stage Least Squares Procedure. Doctoral Thesis. University of California. Berkeley.

20. Smith, V.K. (1973). Monte Carlo Methods: Their Role in Econometrics. Lexington, Mass.: D.C. Heath.

21. Srivastava, V.K., Dwivedi, T.D., Belinsky, M. and Tiwari, R. (1980). A Numerical Comparison of Exact Large-Sample and Small Disturbance Approximations of Properties of k-class estimators. International Economic Review, 21, 249-252.

22. Summers, R. (1965). A Capital-Intensive Approach to the Small Sample Properties of Various Simultaneous Equation Estimators. Econometrica, 33, 1-41.

23. Tawil, J.J. (1972). The Small-Sample Behaviour of Various Estimators in a Linear Structural and a Linear Functional Relationship: A Monte Carlo Study. Doctoral Thesis. University of California. Los Angeles.

APPENDIX A

This Appendix includes all the data for the experiment.

TABLE A.1

$K_{\alpha-1}$

ANNUAL DATA FROM 1959 TO 1979

1959	354960.	373151.	394745.	418274.
1963	443889.	470700.	504131.	542194.
1967	580856.	617706.	664105.	720329.
1971	774132.	836250.	909027.	986742.
1975	1.037353E+06	1.086818E+06	1.139868E+06	1197468.
1979	1.258768E+06			

TABLE A.2

I_{α}

ANNUAL DATA FROM 1959 TO 1979

1959	18191.	21594.	23529.	25613.
1963	26811.	33431.	38063.	38662.
1967	36850.	46311.	56224.	53803.
1971	62118.	72777.	77715.	50611.
1975	49465.	53050.	57600.	61300.
1979	68070.			

TABLE A.3

C_{α}

ANNUAL DATA FROM 1959 TO 1979

1959	104452.	107808.	115147.	120050.
1963	126115.	137192.	147707.	157687.
1967	167528.	179025.	190089.	205888.
1971	217242.	232324.	250057.	251650.
1975	265242.	279343.	292500.	310250.
1979	319700.			

TABLE A.4

G_{α}

ANNUAL DATA FROM 1959 TO 1979

1959	7984.	8357.	8452.	9147.
1963	8808.	9420.	10293.	10444.
1967	10291.	8875.	10013.	10085.
1971	9850.	10854.	13751.	19334.
1975	20000.	20122.	19636.	17107.
1979	17422.			

TABLE A.5

W'_α

ANNUAL DATA FROM 1959 TO 1979

1959	11335.	12145.	12952.	13685.
1963	14977.	16585.	18043.	19667.
1967	22370.	24224.	25626.	27657.
1971	29757.	30997.	30947.	30762.
1975	36075.	38831.	43164.	47893.
1979	51378.			

TABLE A.6

R_α

ANNUAL DATA FROM 1959 TO 1979

1959	-5502.	-5811.	429.	-5173.
1963	5075.	2970.	7478.	1116.
1967	3372.	990.	4624.	9632.
1971	8445.	8582.	18817.	10507.
1975	11975.	10577.	7510.	7123.
1979	6020.			

TABLE A.7

K_α

ANNUAL DATA FROM 1959 TO 1979

1959	373151.	394745.	418274.	443887.
1963	470700.	504131.	542194.	580856.
1967	617706.	664017.	720329.	774132.
1971	836250.	909027.	986742.	1037353.
1975	1.086818E+06	1.139868E+06	1.197468E+06	1258768.
1979	1.326838E+06			

TABLE A.8

W_α

ANNUAL DATA FROM 1959 TO 1979

1959	30684.	32485.	35651.	36951.
1963	39909.	43468.	48465.	52942.
1967	56781.	61984.	67723.	72610.
1971	79445.	89400.	91670.	69858.
1975	94350.	104660.	116322.	127869.
1979	131267.			

TABLE A.9

P_{α}

ANNUAL DATA FROM 1959 TO 1979

1959	80595.	82575.	93350.	93848.
1963	104073.	111322.	120961.	126275.
1967	131102.	136299.	149642.	163236.
1971	177302.	192845.	218129.	206967.
1975	206684.	217309.	215978.	222832.
1979	230763.			

TABLE A.10

T_{α}

ANNUAL DATA FROM 1959 TO 1979

1959	13846.	16888.	18556.	18838.
1963	22827.	28223.	34115.	28692.
1967	30158.	37006.	43585.	43562.
1971	40908.	42292.	50541.	55277.
1975	45648.	41123.	44946.	45079.
1979	49182.			

TABLE A.11

$P_{\alpha-1}$

ANNUAL DATA FROM 1959 TO 1979

1959	78041.	80595.	82575.	93350.
1963	93848.	104073.	111322.	120961.
1967	126275.	131102.	136299.	149642.
1971	163236.	177302.	192845.	218129.
1975	206967.	206684.	217309.	215978.
1979	222832.			

TABLE A.12

α

ANNUAL DATA FROM 1959 TO 1979

1959	1959.	1960.	1961.	1962.
1963	1963.	1964.	1965.	1966.
1967	1967.	1968.	1969.	1970.
1971	1971.	1972.	1973.	1974.
1975	1975.	1976.	1977.	1978.
1979	1979.			

TABLE A.13

X_{α}

ANNUAL DATA FROM 1959 TO 1979

1959	125125.	131948.	147557.	149637.
1963	166809.	183013.	203541.	207909.
1967	218041.	235201.	260950.	279408.
1971	297655.	324537.	360340.	332102.
1975	346682.	363092.	377246.	395780.
1979	411212.			

TABLE A.14

$X_{\alpha-1}$

ANNUAL DATA FROM 1959 TO 1979

1959	123230.	125125.	131948.	147557.
1963	149637.	166809.	183013.	203541.
1967	207909.	218041.	235289.	260950.
1971	279408.	297655.	324537.	360340.
1975	332102.	346682.	363092.	377246.
1979	395780.			

TABLE A.15

$W_{\alpha}+W'_{\alpha}$

ANNUAL DATA FROM 1959 TO 1979

1959	42019.	44630.	48603.	50636.
1963	54886.	60053.	66508.	72609.
1967	79151.	86208.	93349.	100267.
1971	109202.	120397.	122617.	120620.
1975	130425.	143491.	159486.	175762.
1979	182645.			

TABLE A.16

B

ANNUAL DATA FROM 1959 TO 1979

B = A-1970

1959	-11.	-10.	-9.	-8.
1963	-7.	-6.	-5.	-4.
1967	-3.	-2.	-1.	0.
1971	1.	2.	3.	4.
1975	5.	6.	7.	8.
1979	9.			

TABLE A.17

C

ANNUAL DATA FROM 1959 TO 1979

C = 20120.2+0.226*PA+0.327*PA1+0.962*WA

1959	104276.	108071.	114975.	120567.
1963	127129.	137082.	147840.	158062.
1967	167184.	176726.	188311.	202401.
1971	218621.	237503.	250435.	254259.
1975	259978.	274856.	293417.	310188.
1979	320843.			

APPENDIX B

This Appendix contains the programs and subroutines utilized in the Monte Carlo experiment and in the comparisons of the specific estimates.

In the following programs and subroutines we have put:

KA1 for	$K_{\alpha-1}$	IA for	I_α
C11 or C1 "	C_α	GA "	G_α
WA1 "	W'_α	RA "	R_α
KA "	K_α	WA2 "	W_α
PA "	P_α	TA "	T_α
PA1 "	$P_{\alpha-1}$	A "	α
XA "	X_α	XA1 "	$X_{\alpha-1}$
WA "	$W_\alpha+W'_\alpha$	B "	$\alpha-1970$
C or CC "	C	D1 "	$\hat{\delta}_T$
D2 "	$\hat{C}_T$	B "	$b(\hat{\delta}_T)$
COV "	$Cov(\hat{\delta}_T)$	COVO "	$\overline{Cov}(\hat{\delta}_T)$
VAR "	$Var(\hat{\delta}_j)$	SDD1 "	$SD(\hat{\delta}_j)$
MSE1 "	$MSE(\hat{\delta}_j)$	RMSE1 "	$RMSE(\hat{\delta}_j)$
COV1 "	$Cov(\hat{C}_j)$	COVO1 "	$\overline{Cov}(\hat{C}_j)$
VARC "	$Var(\hat{C}_j)$	SDC2 "	$SD(\hat{C}_j)$
MSE2 "	$MSE(\hat{C}_j)$	RMSE2 "	$RMSE(\hat{C}_j)$
BY "	$b(\hat{C}_T)$		

Where : $T = 21$.

IMSL : THE GENERATION OF THE DISTURBANCES

PROGRAM GEN 74/835 OPT = 1

```
      PROGRAM GEN (INPUT, OUTPUT, TAPE 6 = OUTPUT)
      REAL RNML (21), UNIF (21), GLOG (21)
      DOUBLE PRECISION DSEED
      NR = 21
      DSEED = 123457.DO
      XM = -0.3475
      S  =  0.693
      DO 10 I = 1,100
      CALL GGNML (DSEED, NR, RNML)
      CALL GGUBS (DSEED, NR, UNIF)
      CALL GGNLG (DSEED, NR, XM, S, GLOG)
      WRITE (6,50)
      DO 15 K = 1,21
      WRITE (6,100) RNML (K), UNIF (K), GLOG (K)
15    CONTINUE
      DSEED = DSEED + 555.DO
10    CONTINUE
50    FORMAT ("1", 9X, "NORMAL (0,1)", 33X,
      "UNIFORM (0,1)", 33X, "LOGNORMAL(
     *-0.3475,0.693)", / / 1X)
100   FORMAT (/ 4X,F16. 12,33X,F16. 12,33X,F16.12)
      STOP
      END
```

TROLL: TRANSFORMATION OF THE DISTURBANCES

```
MACEDIT TRANSF1;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO F1&9=1500.00 * E1&9;
QUIT
&IF &IFARG(15) LT 100
    &SET &IFARG(15)=&IFARG(15)+1 &END
    &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT TRANSF2;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO F2&9=(E2&9-ONEHALF)* 1500.00*SQRT(12);
QUIT
&IF &IFARG(15) LT 100
    &SET &IFARG(15)=&IFARG(15)+1 &END
    &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT TRANSF3;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
QUIT
DO F3&9=(E3&9-ONE)*1500.00;
&IF &IFARG(15) &LT 100
    &SET &IFARG(15)=&IFARG(15)+1 &END
    &GOTO START
&IFEND
&EXIT;
|EOF:
```

TSP : THE ESTIMATION OF δ_1^0

ORDINARY LEAST SQUARES

VARIABLES...

CII
C
PA
PA1
WA

INDEPENDENT VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC	MEAN OF VARIABLE
C	20120.22656250	2204.16967773	9.12825584	1.00000000
PA	0.22672480	0.07814854	2.90120316	156289.81250000
PA1	0.32711643	0.07678342	4.26024818	149017.37500000
WA	0.96273762	0.05030215	19.13908386	98264.93750000

R-SQUARED = 0.9989

DURBIN-WATSON STATISTIC (ADJ. FOR 0 GAPS) = 1.5927

NUMBER OF OBSERVATIONS = 21

SUM OF SQUARED RESIDUALS = 0.10700SE 09

STANDARD ERROR OF THE REGRESSION = 2508.87

ESTIMATE OF VARIANCE-COVARIANCE MATRIX OF ESTIMATED COEFFICIENTS

0.486E 07	-0.750E 02	0.576E 01	0.642E 02
-0.750E 02	0.611E-02	-0.512E-02	-0.119E-02
0.576E 01	-0.512E-02	0.590E-02	-0.856E-03
0.642E 02	-0.119E-02	-0.856E-03	0.253E-02

TSP : GENERATION OF C_1 - ESTIMATION OF THE STRUCTURAL COEFFICIENTS WITH OLS AND 2SLS ESTIMATORS

```
LOAD"
SMPL 1 21"
GENR CC=20120.266+0.227*PA+0.327*PA1+0.963*WA"
GENR C11   =CC+E11    "
GENR C12   =CC+E12    "
GENR C13   =CC+E13    "
GENR C14   =CC+E14    "
GENR C15   =CC+E15    "
GENR C16   =CC+E16    "
GENR C17   =CC+E17    "
GENR C18   =CC+E18    "
GENR C19   =CC+E19    "
GENR C110  =CC+E110   "
PLOT PA * PA1 Q WA S C11  + E11  = "
PLOT PA * PA1 Q WA S C12  + E12  = "
PLOT PA * PA1 Q WA S C13  + E13  = "
PLOT PA * PA1 Q WA S C14  + E14  = "
PLOT PA * PA1 Q WA S C15  + E15  = "
PLOT PA * PA1 Q WA S C16  + E16  = "
PLOT PA * PA1 Q WA S C17  + E17  = "
PLOT PA * PA1 Q WA S C18  + E18  = "
PLOT PA * PA1 Q WA S C19  + E19  = "
PLOT PA * PA1 Q WA S C110 + E110 = "
PRINT ID PA PA1 WA C11   E11   "
PRINT ID PA PA1 WA C12   E12   "
PRINT ID PA PA1 WA C13   E13   "
PRINT ID PA PA1 WA C14   E14   "
PRINT ID PA PA1 WA C15   E15   "
PRINT ID PA PA1 WA C16   E16   "
```

```
30   PRINT ID PA PA1 WA C17   E17   "
31   PRINT ID PA PA1 WA C18   E18   "
32   PRINT ID PA PA1 WA C19   E19   "
33   PRINT ID PA PA1 WA C110  E110  "
34   CORREL PA PA1 WA C11   "
35   CORREL PA PA1 WA C12   "
36   CORREL PA PA1 WA C13   "
37   CORREL PA PA1 WA C14   "
38   CORREL PA PA1 WA C15   "
39   CORREL PA PA1 WA C16   "
40   CORREL PA PA1 WA C17   "
41   CORREL PA PA1 WA C18   "
42   CORREL PA PA1 WA C19   "
43   CORREL PA PA1 WA C110  "
44   OLSQ C11    C PA PA1 WA "
45   OLSQ C12    C PA PA1 WA "
46   OLSQ C13    C PA PA1 WA "
47   OLSQ C14    C PA PA1 WA "
48   OLSQ C15    C PA PA1 WA "
49   OLSQ C16    C PA PA1 WA "
50   OLSQ C17    C PA PA1 WA "
51   OLSQ C18    C PA PA1 WA "
52   OLSQ C19    C PA PA1 WA "
53   OLSQ C110   C PA PA1 WA "
54   INST C11    C PA PA1 WA "
     INVR C PA1  WAI KA1 XA1 A GA RA TA "
55   INST C12    C PA PA1 WA
     INVR G PA1  WA1 KA1 XA1 A GA RA TA "
56   INST C13    C PA PA1 WA
     INVR C PA1  WA1 KA1 XA1 A GA RA TA "
57   INST C14    C PA PA1 WA
     INVR C PA1  WA1 KA1 XA1 A GA RA TA "
```

```
58    INST C15   C PA PA1 WA
      INVR C PA1 WA1 KA1 XA1 A GA RA TA "
59    INST C16   C PA PA1 WA
      INVR C PA1 WA1 KA1 XA1 A GA RA TA "
60    INST C17   C PA PA1 WA
      INVR C PA1 WA1 KA1 XA1 A GA RA TA "
61    INST C18   C PA PA1 WA
      INVR C PA1 WA1 KA1 XA1 A GA RA TA "
62    INST C19   C PA PA1 WA
      INVR C PA1 WA1 KA1 XA1 A GA RA TA"
63    INST C110  C PA PA1 WA
      INVR C PA1 WA1 KA1 XA1 A GA RA TA"
64    STOP "
65    END "
```

TROLL : GENERATION OF C_1 - ESTIMATION OF THE STRUCTURAL COEFFICIENTS WITH k-CLASS ESTIMATORS (k=0, k=1, UBK)

```
MACEDIT GDOLSE1;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E1&9;
&KCLASS DON;
USEQ ALL;
KTYPE OLS;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
MACEDIT GDOLSE2;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E2&9;
&KCLASS DON;
USEQ ALL;
KTYPE OLS;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT GDOLSE3;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E3&9;
&KCLASS DON;
USEQ ALL;
KTYPE OLS;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT GDTWOE1;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E1&9;
&KCLASS DON;
USEQ ALL;
KTYPE 2SLS;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT GDTWOE2;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E2&9;
&KCLASS DON;
USEQ ALL;
KTYPE 2SLS;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT GDTWOE3;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E3&9;
&KCLASS DON;
USEQ ALL;
KTYPE 2SLS;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT GDGENE1;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E1&9;
&KCLASS DON;
USEQ ALL;
KTYPE GENERAL 1   ;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:


MACEDIT GDGENE2;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E2&9;
&KCLASS DON;
USEQ ALL;
KTYPE GENERAL 1   ;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT GDGENE3;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E3&9;
&KCLASS DON;
USEQ ALL;
KTYPE GENERAL 1  ;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:


MACEDIT GDNAGE1;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E1&9;
&KCLASS DON;
USEQ ALL;
KTYPE NAGARUB;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
MACEDIT GDNAGE2;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E2&9;
&KCLASS DON;
USEQ ALL;
KTYPE NAGARUB;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:


MACEDIT GDNAGE3;
|EDIT:
T*
|TOF:
&SET &IFARG(15)=1 &END
&START:
&SETC &9=&IFARG(15) &END
DO C1 =C+E3&9;
&KCLASS DON;
USEQ ALL;
KTYPE NAGARUB;
PERIOD 1;
DOEQ &1;
TROLL DELETE DATA C1;
QUIT;
QUIT
&IF &IFARG(15) LT 100
   &SET &IFARG(15)=&IFARG(15)+1 &END
   &GOTO START
&IFEND
&EXIT;
|EOF:
```

```
DOS FORTRAN IV 360N-FO-479 3-8         MAINPGM        SUBROUTINE: SUMMARY STATISTICS I
--------------------------------------------------------------------------------------------
0001              DIMENSION DO(4),PA(4),PA1(4), WA(4), CA(21),B(100,4),Y(100,21),
                 *D1(4),B(4),D2(4),VAR(4),MSE1(4),SDD1(4),Y1(21),Y2(21),AVARC(21),
                 *SDC2 (21),AMSE2(21),RMSE2(21),RMSE1(4),TITL(3,20),TITLE(20),
                 *DD(4),CC(21)
0002              DOUBLE PRECISION Y,Y1,Y2,AVARC,AMSE2,SDC2,RMSE2,VAR,MSE1,SDD1,
                 *RMSE1,DD,CC
0003              DO 15 I=1,3
0004           15 READ(5,3) (TITL(I,J),J=1,20)
0005              READ(5,1) (DO(M),M=1,4)
0006            1 FORMAT(    4F9.3,44X)
0007              READ(5,2) (PA(M),M=1,21)
0008            2 FORMAT(11F7.0)
0009              READ(5,2) (PA1(M),M=1,21)
0010              READ(5,2) (WA(M),M=1,21)
0011              READ(5,2) (CA(M),M=1,21)
0012              READ(5,3) (TITLE(J),J=1,20 )
0013            3 FORMAT(20A4)
0014              WRITE(6,107)
0015          107 FORMAT(1H1,51H)
0016              WRITE(6,4) (TITLE(J),J=1,20)
0017            4 FORMAT(1H1,20A4)
0018              DO 10 I=1,3
0019              WRITE(6,5) (TITL(I,J),J=1,20)
0020            5 FORMAT(1H-,5X,20A4)
0021              DO 7 M=1,4
```

```
      D1(M)=0
      DD(M)=0
    7 D2(M)=0
      DO 8 M=1,21
      Y1(M)=0
      CC(M)=0
    8 Y2(M)=0
      DO 11 J=1,100
      READ(5,6) (D( J,M),M=1,4)
    6 FORMAT(6X,4F9.0,30X)
   11 CONTINUE
      DO 20 J=1,100
      DO 20 M=1,21
   20 Y(J,M)=D(J,1)+D(J,2)*PA(M)+D(J,3)*PA1(M)+D(J,4)*WA(M)
        DO 30 M=1,4
        DO 30 J=1,100
   30 D1(M)=D1(M)+D(J,M)
      DO 60 M=1,4
   60 D1(M)=D1(M)/100.
      DO 40 M=1,4
      DO 40 J=1,100
   40 D2(M)=D2( 1)+(D(J,M)-D1(M))**2
      DO 50 M=1,4
   50 VAR(M)=D2(M)/100.
      DO 55 M=1,4
   55 SDD1( 1)=DSQRT(VAR(M))
      DO 70 M=1,4
   70 B(M)=D1(M)-DD(M)
      DO 67 M=1,4
```

```
      DO 67 J=1,100
   67 DO(M)=DO(M)+(D(J,M)-BO(M))**2
      DO 71 M=1,4
   71 MSE1(M)=DO(M)/100.
      DO 72 M=1,4
   72 RMSE1(M)=DSQRT(MSE1(M))
      DO 80 M=1,21
      DO 80 J=1,100
   80 Y1(M)=Y1(M)+Y(J,M)
      DO 81 M=1,21
   81 Y1(M)=Y1(M)/100.
      DO 90 M=1,21
      DO 90 J=1,100
   90 Y2(M)=Y2(M)+(Y(J,M)-Y1(M))**2
      DO 91 M=1,21
   91 AVARC(M)=Y2(M)/100.
      DO 92 M=1,21
   92 SDC2(M)=DSQRT(AVARC(M))
      DO 83 M=1,21
      DO 83 J=1,100
   83 CC(M)=CC(M)+(Y(J,M)-CA(M))**2
      DO 93 M=1,21
   93 AMSE2(M)=CC(M)/100.
      DO 94 M=1,21
   94 RMSE2(M)=DSQRT(AMSE2(M))
      WRITE(6,101)
      WRITE(6,102)
      WRITE(6,103) ((M,D1(M),B(M),VAR(M),SDB1(M),MSE1(M),RMSE1(M)),M=1,4
     *)
      WRITE(6,104)
      WRITE(6,102)
      WRITE(6,106) ((M,AVARC(M),SDC2(M),AMSE2(M),RMSE2(M)),M=1,21)
  101 FORMAT(1H ,10X,1HM,13X,2HD1,18X,1HB,15X,3HVAR,15X,4HSDB1,10X,
     *4HMSE1,15X,5HRMSE1)
  102 FORMAT(1H ,119H-------------------------------------------------------
     *-------------------------------------------------------------
     *---)
  103 FORMAT((1H ,9X,I2,6(3X,E15.7))/)
  104 FORMAT(1H-,10X,1HM,14X,4HVARC,13X,5HSDC2 ,9X,4HMSE2,14X,5HRMSE2)
  106 FORMAT((1H ,9X,I2,4(3X,E15.7))/)
      WRITE(6,107)
   10 CONTINUE
      STOP
      END
```

SUBROUTINE: SUMMARY STATISTICS II

```
0001          DIMENSION DO(4),PA(4),PA1(4), WA(4), CA(21),B(100,4),Y(100,21),
             *D1(4),B1(4),      VAR(4), MSE(4),SDDS(4),YL(21),          AVARR(21)
             *SDCZ.(21),AMSE2(21),RMSE2(21),RMSE1(4),TITL(3,20),TITLE(20),
             *                 BY(21),D3(4,4),Y3(21,21),D4(4,4),Y4(21,21),
             *A(21),D5(16),D6(16),Y5(441),Y6(441),LL(21),MM(21)
             *,BB(21),ZZ(21)
             *,DA(100,4,4),D2(100,4),DD(100,4),Y2(100,21),CC(100,21),
             *DDD(100,4,4),CCA(100,21,21),YA(100,21,21)
0002          DOUBLE PRECISION Y,Y1,Y2,AVARC,AMSE2,SDCZ,RMSE2,VAR,MSE1,DDD1,
             *AMSE1,DD,CC,BY,D3,D4,Y3,Y4,D5,D6,Y5,Y6,DET1,DET2,DET3,DET4,D2
             *,BB,DA,DDD,CCA,YA
0003          INTEGER A
0004          DO 12 J=1,21
0005       12 A(J)=J
0006          DO 15 I=1,3
0007       15 READ(5,3) (TITL(I,J),J=1,20)
0008          READ(5,1) (DO(M),M=1,4)
0009        1 FORMAT(    4F9.3,44X)
0010          READ(5,2) (PA(M),M=1,21)
0011        2 FORMAT(1 1F7.0)
0012          READ(5,2) (PA1(M),M=1,21)
0013          READ(5,2) (WA(M),M=1,21)
0014          READ(5,2) (CA(M),M=1,21)
0015          READ(5,3) (TITLE(J),J=1,20 )
0016        3 FORMAT(20A4)
0017          WRITE(6,107)
0018      107 FORMAT(1H1,51H
```

```
      WRITE(6,4) (TITLE(J),J=1,20)
    4 FORMAT(1H1,20A4)
      DO 10 I=1,3
      WRITE(6,5) (TITL(I,J),J=1,20)
    5 FORMAT(1H-,5X,20A4)
      DO 7 M=1,4
    7 D1(M)=0
      DO 8 M=1,21
    8 Y1(M)=0
      DO 11 J=1,100
      READ(5,6) (D(J,M),M=1,4)
    6 FORMAT(6X,4F9.0,38X)
   11 CONTINUE
      DO 20 J=1,100
      DO 20 M=1,21
   20 Y(J,M)=D(J,1)+D(J,2)*PA(M)+D(J,3)*PA1(M)+D(J,4)*WA(M)
        DO 30 M=1,4
        DO 30 J=1,100
   30 D1(M)=D1(M)+D(J,M)
      DO 60 M=1,4
   60 D1(M)=D1(M)/100.
      DO 40 M=1,4
      DO 40 J=1,100
   40 D2(J,M)=D(J,M)-D1(M)
      DO 70 M=1,4
   70 B(M)=D1(M)-D0(M)
      DO 43 J=1,100
```

```
      DO 43 M=1,4
      DO 43 N=1,4
   43 DA(J,M,N)=D2(J,M)*D2(J,N)
      DO 44 M=1,4
      DO 44 N=1,4
   44 D3(M,N)=0
      DO 45 J=1,100
      DO 45 M=1,4
      DO 45 N=1,4
   45 D3(M,N)=D3(M,N)+DA(J,M,N)
      DO 46 M=1,4
      DO 46 N=1,4
   46 D3(M,N)=D3(M,N)/100.
      WRITE(6,101)
      WRITE(6,102)
      WRITE(6,103) ((M,D1(M),B(M)),M=1,4)
      WRITE(6,122)
  122 FORMAT(1H-,10X,10HMATRIX COV)
      WRITE(6,102)
      WRITE(6,123) (A(J),J=1,4)
  123 FORMAT(1H ,4(14X,2HJ=,I2))
      DO 98 M=1,4
      WRITE(6,124) M,(D3(M,N),N=1,4)
  124 FORMAT(1H ,2X,2HI=,I2,4(2X,E15.7))
   98 CONTINUE
      CALL ORIZ(D3,4,DET1,BB,ZZ)
      WRITE(6,108) DET1
```

```
  103 FORMAT(1H-,17HDETERMINANT COV =,E15.7)
      DO 67 M=1,4
      DO 67 J=1,100
   67 DD(J,M)=D(J,M)-DO(M)
      DO 71 J=1,100
      DO 71 M=1,4
      DO 71 N=1,4
   71 DDD(J,M,N)=DD(J,M)*DD(J,N)
      DO 72 M=1,4
      DO 72 N=1,4
   72 D4(M,N)=0
      DO 73 J=1,100
      DO 73 M=1,4
      DO 73 N=1,4
   73 D4(M,N)=D4(M,N)+DDD(J,M,N)
      DO 74 M=1,4
      DO 74 N=1,4
   74 D4(M,N)=D4(M,N)/100.
      WRITE(6,120)
  120 FORMAT(1H-,10X,11HMATRIX COVO)
      WRITE(6,102)
      WRITE(6,123) (A(J),J=1,4)
      DO 97 M=1,4
      WRITE(6,124) M,(D4(M,N),N=1,4)
   97 CONTINUE
      CALL ORIZ(D4,4,DET2,88,ZZ)
      WRITE(6,110) DET2
```

```
 11) FORMAT(1H-,17HDETERMINANT COVO=,E15.7)
     DO 80 M=1,21
     DO 80 J=1,100
  80 Y1(M)=Y1(M)+Y(J,M)
     DO 81 M=1,21
  81 Y1(M)=Y1(M)/100.
     DO 90 M=1,21
     DO 90 J=1,100
  90 Y2(J,M)=Y(J,M)-Y1(M)
     DO 84 M=1,21
  84 BY(M)=Y1(M)-CA(M)
     DO 91 J=1,100
     DO 91 M=1,21
     DO 91 N=1,21
  91 YA(J,M,N)=Y2(J,M)*Y2(J,N)
     DO 92 N=1,21
     DO 92 M=1,21
  92 Y3(M,N)=0
     DO 93 J=1,100
     DO 93 M=1,21
     DO 93 N=1,21
  93 Y3(M,N)=Y3(M,N)+YA(J,M,N)
     DO 94 M=1,21
     DO 94 N=1,21
  94 Y3(M,N)=Y3(M,N)/100.
     WRITE(6,104)
     WRITE(6,102)
     WRITE(6,103) ((M,Y1(M),BY(M)),M=1,21)
 101 FORMAT(1H ,10X,1HM,13X,2HD1,18X,1HB)
 102 FORMAT(1H ,108H-----------------------------------------
```

```
     +---------------------------------------------------------------)
 103 FORMAT((1H ,9X,I2,2(3X,E15.7))/)
 104 FORMAT(1H1,10X,1HM,13X,2HY1,17X,2HBY)
     WRITE(6,121)
 121 FORMAT(1H1,10X,11HMATRIX COV)
     WRITE(6,102)
     L=-5
     DO 111 K=1,4
     WRITE(6,114)
     WRITE(6,114)
     L=L+6
     L1=L+5
     IF (L1.GT.21) L1=21
     WRITE(6,106) (A(J),J=L,L1)
 106 FORMAT(1H ,6(14X,2HJ=,I2))
     WRITE(6,114)
 114 FORMAT(1H ,5IH
     DO 112 M=1,21
     WRITE(6,113) M,(Y3(M,N),N=L,L1)
 113 FORMAT(1H ,2X,2HI=,I2,6(2X,E15.7))
 112 CONTINUE
 111 CONTINUE
     DO 83 M=1,21
     DO 33 J=1,100
```

```
   83 CC(J,M)=Y(J,M)-CA(M)
      DO 95 J=1,100
      DO 95 M=1,21
      DO 95 N=1,21
   95 CCA(J,M,N)=CC(J,M)*CC(J,N)
      DO 96 M=1,21
      DO 96 N=1,21
   96 Y4(M,N)=0
      DO 13 J=1,100
      DO 13 M=1,21
      DO 13 N=1,21
   13 Y4(M,N)=Y4(M,N)+CCA(J,M,N)
      DO 14 M=1,21
      DO 14 N=1,21
   14 Y4(M,N)=Y4(M,N)/100.
      WRITE(6,119)
  119 FORMAT(1H1,10X,11HMATRIX COVD)
      WRITE(6,102)
      L=-5
      DO 116 K=1,4
      WRITE(6,114)
      WRITE(6,114)
      L=L+6
      L1=L+5
      IF (L1.GT.21) L1=21
      WRITE(6,106) (A(J),J=L,L1)
      WRITE(6,114)
      DO 117 M=1,21
      WRITE(6,113) M,(Y4(M,N),N=L,L1)
  117 CONTINUE
  116 CONTINUE
      WRITE(6,107)
   10 CONTINUE
      STOP
      END
```

APPENDIX C

This Appendix contains the tables with the summary statistics, which have been found in the Monte Carlo experiment.

TABLE C.1

NORMAL DISTRIBUTION

		$\bar{\hat{\delta}}_{21}$	$b(\hat{\delta}_{21})$	$SD(\hat{\delta}_j)$	$RMSE\ (\hat{\delta}_j)$
OLS	β_o	0.1992561E 05	-0.1946133E 03	0.9736715D 03	0.9928933D 03
	β_1	0.2284805E 00	0.1480520E-02	0.3337196D-01	0.3340510D-01
	β_2	0.3247264E 00	-0.2273560E-02	0.3792298D-01	0.3799101D-01
	β_3	0.9627980E 00	-0.2020001E-03	0.2421741D-01	0.2421823D-01
2SLS	β_o	0.2059857E 05	0.4783438E 03	0.1326359D 04	0.1410068D 04
	β_1	0.1753293E 00	-0.5167073E-01	0.6538572D-01	0.8333728D-01
	β_2	0.3717523E 00	0.4475230E-01	0.6468157D-01	0.7865681D-01
	β_3	0.9734653E 00	0.1046526E-01	0.2816225D-01	0.3004608D-01
UBK	β_o	0.2072444E 05	0.6042148E 03	0.1437423D 04	0.1559335D 04
	β_1	0.1725708E 00	-0.5442923E-01	0.8054132D-01	0.9720779D-01
	β_2	0.3804843E 00	0.5348432E-01	0.7332970D-01	0.9076521D-01
	β_3	0.9751280E 00	0.1212800E-01	0.2961007D-01	0.3200008D-01

TABLE C.2

UNIFORM DISTRIBUTION

		$\bar{\hat{\delta}}_{21}$	$b(\hat{\delta}_{21})$	$SD(\hat{\delta}_j)$	$RMSE\ (\hat{\delta}_j)$
OLS	β_0	0.1993495E 05	-0.1852734E 03	0.1019235D 04	0.1035897D 04
	β_1	0.2275958E 00	0.5958080E-03	0.3202335D-01	0.3202916D-01
	β_2	0.3298907E 00	0.2890766E-02	0.3208399D-01	0.3221444D-01
	β_3	0.9568989E 00	-0.6101072E-02	0.2204284D-01	0.2286983D-01
2SLS	β_0	0.2057511E 05	0.4548867E 03	0.1154428D 04	0.1240906D 04
	β_1	0.1732941E 00	-0.5370593E-01	0.6688799D-01	0.8578004D-01
	β_2	0.3759892E 00	0.4898924E-01	0.6242194D-01	0.7935294D-01
	β_3	0.9668959E 00	0.3895879E-02	0.2473846D-01	0.2504439D-01
UBK	β_0	0.2068621E 05	0.5659844E 03	0.1216790D 04	0.1342074D 04
	β_1	0.1638818E 00	-0.6311822E-01	0.7269725D-01	0.9627430D-01
	β_2	0.3854370E 00	0.5843705E-01	0.6972398D-01	0.9097711D-01
	β_3	0.9689004E 00	0.5900443E-02	0.2574529D-01	0.2641421D-01

TABLE C.3

LOGNORMAL DISTRIBUTION

		$\bar{\bar{\delta}}_{21}$	$b(\hat{\delta}_{21})$	$SD(\hat{\delta}_j)$	$RMSE\ (\hat{\delta}_j)$
OLS	β_0	0.2014865E 05	0.2842578E 02	0.6473220D 03	0.6479595D 03
	β_1	0.2177186E 00	-0.9281397E-02	0.2318499D-01	0.2497275D-01
	β_2	0.3335133E 00	0.6513357E-02	0.2345679D-01	0.2434542D-01
	β_3	0.9641047E 00	0.1104712E-02	0.1396474D-01	0.1400893D-01
2SLS	β_0	0.2068250E 05	0.5622773E 03	0.8075674D 03	0.9841842D 03
	β_1	0.1734817E 00	-0.5351830E-01	0.3913765D-01	0.6630125D-01
	β_2	0.3705263E 00	0.4352629E-01	0.3562092D-01	0.5624717D-01
	β_3	0.9726098E 00	0.9609818E-02	0.1581676D-01	0.1851078D-01
UBK	β_0	0.2077015E 05	0.6499258E 03	0.8535949D 03	0.1073032D 04
	β_1	0.1662198E 00	-0.6078023E-01	0.4370294D-01	0.7486065D-01
	β_2	0.3766056E 00	0.4960567E-01	0.3921826D-01	0.6323895D-01
	β_3	0.9740008E 00	0.1100081E-01	0.1633693D-01	0.1969891D-01

TABLE C.4

OLS: NORMAL DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1041388D 06	-0.3132181D 03
2	0.1079344D 06	0.1264187D 03
3	0.1148645D 06	-0.2825063D 03
4	0.1204346D 06	0.3846306D 03
5	0.1519686D 06	0.2585361D 05
6	0.1566312D 06	0.1943919D 05
7	0.1654398D 06	0.1773277D 05
8	0.1732825D 06	0.1559548D 05
9	0.1848693D 06	0.1734131D 05
10	0.1972207D 06	0.1819567D 05
11	0.2106355D 06	0.2054653D 05
12	0.2223010D 06	0.1641298D 05
13	0.2460415D 06	0.2879950D 05
14	0.2614747D 06	0.2915070D 05
15	0.2772186D 06	0.2716155D 05
16	0.2967910D 06	0.4514097D 05
17	0.3123041D 06	0.4706209D 05
18	0.3330894D 06	0.5374645D 05
19	0.3561582D 06	0.6365820D 05
20	0.3651026D 06	0.5485261D 05
21	0.3841245D 06	0.6442454D 05

TABLE C.5

OLS: NORMAL DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.6089328D 03	0.6847662D 03
2	0.6108330D 03	0.6237777D 03
3	0.6257450D 03	0.6865614D 03
4	0.6446160D 03	0.7506467D 03
5	0.2377671D 04	0.2596271D 05
6	0.2399216D 04	0.1958669D 05
7	0.2461119D 04	0.1790274D 05
8	0.2621267D 04	0.1581424D 05
9	0.2914979D 04	0.1758460D 05
10	0.3062550D 04	0.1845160D 05
11	0.3230950D 04	0.2079901D 05
12	0.3407059D 04	0.1676288D 05
13	0.2419419D 04	0.2890095D 05
14	0.2619319D 04	0.2926814D 05
15	0.2744128D 04	0.2729982D 05
16	0.3011965D 04	0.4524135D 05
17	0.3168997D 04	0.4716867D 05
18	0.3309636D 04	0.5384825D 05
19	0.3514520D 04	0.6375514D 05
20	0.3440520D 04	0.5496040D 05
21	0.3578414D 04	0.6452385D 05

TABLE C.6

OLS: UNIFORM DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1042320D 0[illegible]	-0.2200438D 03
2	0.1080236D 06	0.2156150D 03
3	0.1149310D 06	-0.2160444D 03
4	0.1205443D 06	0.4943012D 03
5	0.1515097D 06	0.2539467D 05
6	0.1561637D 06	0.1897169D 05
7	0.1649477D 06	0.1724073D 05
8	0.1727625D 06	0.1507549D 05
9	0.1846369D 06	0.1710888D 05
10	0.1969379D 06	0.1791291D 05
11	0.2103251D 06	0.2023611D 05
12	0.2219552D 06	0.1606722D 05
13	0.2456214D 06	0.2837936D 05
14	0.2610410D 06	0.2871697D 05
15	0.2767674D 06	0.2671036D 05
16	0.2962938D 06	0.4464377D 05
17	0.3117854D 06	0.4654336D 05
18	0.3325313D 06	0.5318830D 05
19	0.3555433D 06	0.6304325D 05
20	0.3645510D 06	0.5430103D 05
21	0.3835427D 06	0.6384270D 05

TABLE C.7

OLS: UNIFORM DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.5156112D 03	0.5606016D 03
2	0.5158807D 03	0.5591267D 03
3	0.4496423D 03	0.4988521D 03
4	0.4873244D 03	0.6941317D 03
5	0.2052202D 04	0.2547746D 05
6	0.2074085D 04	0.1908473D 05
7	0.2144306D 04	0.1737357D 05
8	0.2251216D 04	0.1524265D 05
9	0.2991508D 04	0.1736845D 05
10	0.3435006D 04	0.1818517D 05
11	0.3294289D 04	0.2050250D 05
12	0.3464224D 04	0.1643643D 05
13	0.2393656D 04	0.2848013D 05
14	0.2599281D 04	0.2883436D 05
15	0.2715990D 04	0.2684809D 05
16	0.2975519D 04	0.4474282D 05
17	0.3124458D 04	0.4664811D 05
18	0.3246804D 04	0.5328731D 05
19	0.3432454D 04	0.6313663D 05
20	0.3349830D 04	0.5440426D 05
21	0.3471580D 04	0.6393702D 05

TABLE C.8

OLS: LOGNORMAL DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1042351D 06	-0.2168763D 03
2	0.1080353D 06	0.2272994D 03
3	0.1148720D 06	-0.2749963D 03
4	0.1205341D 06	0.4841156D 03
5	0.1518574D 06	0.2574242D 05
6	0.1565198D 06	0.1932785D 05
7	0.1653516D 06	0.1764461D 05
8	0.1731026D 06	0.1541564D 05
9	0.1857227D 06	0.1819468D 05
10	0.1980899D 06	0.1906490D 05
11	0.2115402D 06	0.2145122D 05
12	0.2232399D 06	0.1735191D 05
13	0.2464674D 06	0.2922541D 05
14	0.2619768D 06	0.2965283D 05
15	0.2777485D 06	0.2769154D 05
16	0.2973765D 06	0.4572653D 05
17	0.3129257D 06	0.4768367D 05
18	0.3337125D 06	0.5436954D 05
19	0.3567885D 06	0.6428849D 05
20	0.3657664D 06	0.5551642D 05
21	0.3848000D 06	0.6509995D 05

TABLE C.9

OLS: LOGNORMAL DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.3162526D 03	0.3834723D 03
2	0.3139145D 03	0.3875659D 03
3	0.3242647D 03	0.4251712D 03
4	0.2898643D 03	0.5642599D 03
5	0.1384956D 04	0.2577965D 05
6	0.1396467D 04	0.1937823D 05
7	0.1429269D 04	0.1770240D 05
8	0.1523879D 04	0.1549077D 05
9	0.2004144D 04	0.1830473D 05
10	0.2095304D 04	0.1917969D 05
11	0.2205472D 04	0.2156429D 05
12	0.2319993D 04	0.1750632D 05
13	0.1525502D 04	0.2926519D 05
14	0.1670325D 04	0.2969984D 05
15	0.1751066D 04	0.2774685D 05
16	0.1926018D 04	0.4576707D 05
17	0.2029184D 04	0.4772683D 05
18	0.2106827D 04	0.5441034D 05
19	0.2223900D 04	0.6432695D 05
20	0.2191335D 04	0.5555965D 05
21	0.2271870D 04	0.6513958D 05

TABLE C.10

2SLS: NORMAL DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1046461D 06	0.1940656D 03
2	0.1084844D 06	0.6764137D 03
3	0.1149773D 06	-0.1697444D 03
4	0.1210493D 06	0.9993087D 03
5	0.1515836D 06	0.2546856D 05
6	0.1562690D 06	0.1907698D 05
7	0.1652374D 06	0.1753043D 05
8	0.1726553D 06	0.1496833D 05
9	0.1895660D 06	0.2203801D 05
10	0.2020546D 06	0.2302956D 05
11	0.2157155D 06	0.2562652D 05
12	0.2276099D 06	0.2172193D 05
13	0.2488801D 06	0.3163813D 05
14	0.2647785D 06	0.3245448D 05
15	0.2807447D 06	0.3068773D 05
16	0.3006944D 06	0.4904438D 05
17	0.3164690D 06	0.5122701D 05
18	0.3373671D 06	0.5802411D 05
19	0.3605864D 06	0.6808638D 05
20	0.3697603D 06	0.5951030D 05
21	0.3889381D 06	0.6923806D 05

TABLE C.11

2SLS: NORMAL DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.1730047D 04	0.1740898D 04
2	0.1788444D 04	0.1912085D 04
3	0.1712226D 04	0.1720620D 04
4	0.1982857D 04	0.2220437D 04
5	0.2946404D 04	0.2563843D 05
6	0.2991699D 04	0.1931014D 05
7	0.3086242D 04	0.1780002D 05
8	0.3311944D 04	0.1533035D 05
9	0.6324362D 04	0.2292753D 05
10	0.6554121D 04	0.2394404D 05
11	0.6903259D 04	0.2654003D 05
12	0.7228475D 04	0.2289308D 05
13	0.4947977D 04	0.3202270D 05
14	0.5485476D 04	0.3291480D 05
15	0.5818559D 04	0.3123448D 05
16	0.6345727D 04	0.4945321D 05
17	0.6718898D 04	0.5166575D 05
18	0.7008078D 04	0.5844579D 05
19	0.7364947D 04	0.6848356D 05
20	0.7601420D 04	0.5999381D 05
21	0.7912994D 04	0.6968877D 05

TABLE C.12

2SLS: UNIFORM DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1045132D 06	0.6122312D 02
2	0.1083412D 06	0.5332050D 02
3	0.1147944D 06	0.3525800D 03
4	0.1208978D 06	0.8477562D 03
5	0.1508932D 06	0.2477818D 05
6	0.1555624D 06	0.1837042D 05
7	0.1644955D 06	0.1678845D 05
8	0.1718678D 06	0.1418083D 05
9	0.1890711D 06	0.2154314D 05
10	0.2014958D 06	0.2247082D 05
11	0.2151107D 06	0.2502172D 05
12	0.2269562D 06	0.2106821D 05
13	0.2480780D 06	0.3083596D 05
14	0.2639409D 06	0.3161688D 05
15	0.2798641D 06	0.2980708D 05
16	0.2997423D 06	0.4809226D 05
17	0.3154716D 06	0.5022964D 05
18	0.3362969D 06	0.5695386D 05
19	0.3594251D 06	0.6692512D 05
20	0.3686352D 06	0.5838520D 05
21	0.3877518D 06	0.6805176D 05

TABLE C.13

2SLS: UNIFORM DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.5696397D 03	0.5729203D 03
2	0.5858485D 03	0.7921654D 03
3	0.5075883D 03	0.6180280D 03
4	0.5933602D 03	0.1034779D 04
5	0.2397194D 04	0.2489387D 05
6	0.2411690D 04	0.1852805D 05
7	0.2421780D 04	0.1696223D 05
8	0.2728985D 04	0.1444103D 05
9	0.5481737D 04	0.2222963D 05
10	0.5661086D 04	0.2317295D 05
11	0.5952359D 04	0.2571997D 05
12	0.6234214D 04	0.2197123D 05
13	0.3441353D 04	0.3102739D 05
14	0.3930239D 04	0.3186023D 05
15	0.4170283D 04	0.3009740D 05
16	0.4623891D 04	0.4831403D 05
17	0.4918141D 04	0.5046984D 05
18	0.5065809D 04	0.5717871D 05
19	0.5282159D 04	0.6713325D 05
20	0.5418625D 04	0.5863611D 05
21	0.5604012D 04	0.6828211D 05

TABLE C.14

2SLS: LOGNORMAL DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1044495D 06	-0.2501250D 01
2	0.1082788D 06	0.4708250D 03
3	0.1147459D 06	-0.4010650D 03
4	0.1208021D 06	0.7521169D 03
5	0.1513825D 06	0.2526752D 05
6	0.1560572D 06	0.1886515D 05
7	0.1650108D 06	0.1730380D 05
8	0.1724021D 06	0.1471511D 05
9	0.1893364D 06	0.2180842D 05
10	0.2018066D 06	0.2278156D 05
11	0.2154422D 06	0.2535319D 05
12	0.2273183D 06	0.2143030D 05
13	0.2484732D 06	0.3123125D 05
14	0.2643420D 06	0.3201799D 05
15	0.2802723D 06	0.3021533D 05
16	0.3001872D 06	0.4853717D 05
17	0.3159288D 06	0.5068682D 05
18	0.3367795D 06	0.5743647D 05
19	0.3599505D 06	0.6745059D 05
20	0.3690853D 06	0.5883533D 05
21	0.3882194D 06	0.6851937D 05

TABLE C.15

2SLS: LOGNORMAL DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.3750589D 03	0.3750672D 03
2	0.3834020D 03	0.6071847D 03
3	0.3232044D 03	0.5150866D 03
4	0.3713465D 03	0.8387956D 03
5	0.1479492D 04	0.2531080D 05
6	0.1489618D 04	0.1892387D 05
7	0.1510452D 04	0.1736960D 05
8	0.1657486D 04	0.1480816D 05
9	0.3301301D 04	0.2205688D 05
10	0.3416347D 04	0.2303629D 05
11	0.3588175D 04	0.2560585D 05
12	0.3758345D 04	0.2175737D 05
13	0.2140713D 04	0.3130453D 05
14	0.2421030D 04	0.3210939D 05
15	0.2557654D 04	0.3032339D 05
16	0.2827157D 04	0.4861944D 05
17	0.2997306D 04	0.5077537D 05
18	0.3085377D 04	0.5751928D 05
19	0.3218535D 04	0.6752734D 05
20	0.3270387D 04	0.5892615D 05
21	0.3378574D 04	0.6860261D 05

TABLE C.16

UBK: NORMAL DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1053009D 06	0.8489269D 03
2	0.1091605D 06	0.1352457D 04
3	0.1156475D 06	0.5004775D 03
4	0.1218156D 06	0.1765631D 04
5	0.1520347D 06	0.2591974D 05
6	0.1567415D 06	0.1954949D 05
7	0.1657514D 06	0.1804438D 05
8	0.1731655D 06	0.1547845D 05
9	0.1906977D 06	0.2316974D 05
10	0.2032268D 06	0.2420181D 05
11	0.2169584D 06	0.2686938D 05
12	0.2289066D 06	0.2301860D 05
13	0.2500977D 06	0.3285568D 05
14	0.2661026D 06	0.3377862D 05
15	0.2821588D 06	0.3210185D 05
16	0.3022084D 06	0.5055839D 05
17	0.3180711D 06	0.5282910D 05
18	0.3390641D 06	0.5972113D 05
19	0.3623805D 06	0.6988049D 05
20	0.3716675D 06	0.6141748D 05
21	0.3909399D 06	0.7123985D 05

TABLE C.17

UBK: NORMAL DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.5046822D 04	0.5117723D 04
2	0.5189847D 04	0.5363176D 04
3	0.5553418D 04	0.5575924D 04
4	0.5873942D 04	0.6133567D 04
5	0.5214511D 04	0.2643906D 05
6	0.5366251D 04	0.2027262D 05
7	0.5567169D 04	0.1888367D 05
8	0.6057942D 04	0.1662171D 05
9	0.8223085D 04	0.2458569D 05
10	0.8542655D 04	0.2566524D 05
11	0.9040673D 04	0.2834956D 05
12	0.9455109D 04	0.2488483D 05
13	0.8846421D 04	0.3402580D 05
14	0.9517345D 04	0.3509380D 05
15	0.1013773D 05	0.3366455D 05
16	0.1083893D 05	0.5170718D 05
17	0.1144551D 05	0.5405472D 05
18	0.1215545D 05	0.6094561D 05
19	0.1290143D 05	0.7106145D 05
20	0.1357880D 05	0.6290064D 05
21	0.1427080D 05	0.7265516D 05

TABLE C.18

UBK: UNIFORM DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1046872D 06	0.2351850D 03
2	0.1085259D 06	0.7178875D 03
3	0.1149043D 06	-0.2426550D 03
4	0.1211089D 06	0.1058873D 04
5	0.1508760D 06	0.2476097D 05
6	0.1555526D 06	0.1836056D 05
7	0.1645192D 06	0.1681221D 05
8	0.1718192D 06	0.1413220D 05
9	0.1900263D 06	0.2249826D 05
10	0.2024803D 06	0.2345528D 05
11	0.2161482D 06	0.2605919D 05
12	0.2280409D 06	0.2215286D 05
13	0.2487331D 06	0.3149107D 05
14	0.2646921D 06	0.3236810D 05
15	0.2806677D 06	0.3061073D 05
16	0.3006257D 06	0.4897571D 05
17	0.3164137D 06	0.5117173D 05
18	0.3372735D 06	0.5793052D 05
19	0.3604429D 06	0.6794288D 05
20	0.3697115D 06	0.5946147D 05
21	0.3888699D 06	0.6916992D 05

TABLE C.19

UBK: UNIFORM DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.1055344D 04	0.1081232D 04
2	0.1095591D 04	0.1309841D 04
3	0.9214751D 03	0.9528892D 03
4	0.1234774D 04	0.1626616D 04
5	0.2278085D 04	0.2486554D 05
6	0.2290030D 04	0.1850283D 05
7	0.2301340D 04	0.1696899D 05
8	0.2594263D 04	0.1436834D 05
9	0.6319509D 04	0.2336896D 05
10	0.6521961D 04	0.2434515D 05
11	0.6863312D 04	0.2694784D 05
12	0.7184421D 04	0.2328873D 05
13	0.4103004D 04	0.3175724D 05
14	0.4693498D 04	0.3270662D 05
15	0.4998612D 04	0.3101617D 05
16	0.5523443D 04	0.4928619D 05
17	0.5882448D 04	0.5150873D 05
18	0.6078144D 04	0.5824851D 05
19	0.6338805D 04	0.6823793D 05
20	0.6592344D 04	0.5982579D 05
21	0.6834524D 04	0.6950675D 05

TABLE C.20

UBK: LOGNORMAL DISTRIBUTION

	$\bar{\hat{C}}_{21}$	$b(\hat{C}_{21})$
1	0.1044847D 06	0.3268187D 02
2	0.1083188D 06	0.5107887D 03
3	0.1147252D 06	-0.4217975D 03
4	0.1208461D 06	0.7961012D 03
5	0.1513041D 06	0.2518908D 05
6	0.1559807D 06	0.1878870D 05
7	0.1649543D 06	0.1724732D 05
8	0.1722865D 06	0.1459954D 05
9	0.1899292D 06	0.2240123D 05
10	0.2024162D 06	0.2339121D 05
11	0.2160822D 06	0.2599322D 05
12	0.2279873D 06	0.2209925D 05
13	0.2488019D 06	0.3155991D 05
14	0.2647296D 06	0.3240560D 05
15	0.2806859D 06	0.3062894D 05
16	0.3006478D 06	0.4899782D 05
17	0.3164210D 06	0.5117902D 05
18	0.3372821D 06	0.5793909D 05
19	0.3604688D 06	0.6796876D 05
20	0.3696293D 06	0.5937927D 05
21	0.3887798D 06	0.6907977D 05

TABLE C.21

UBK: LOGNORMAL DISTRIBUTION

	$SD(\hat{C}_j)$	$RMSE(\hat{C}_j)$
1	0.3943855D 03	0.3957373D 03
2	0.4058945D 03	0.6524228D 03
3	0.3282625D 03	0.5344805D 03
4	0.3971216D 03	0.8896532D 03
5	0.1506562D 04	0.2523409D 05
6	0.1516102D 04	0.1884977D 05
7	0.1531692D 04	0.1731520D 05
8	0.1698578D 04	0.1469802D 05
9	0.3669271D 04	0.2269976D 05
10	0.3792383D 04	0.2369664D 05
11	0.3982158D 04	0.2629648D 05
12	0.4168869D 04	0.2248903D 05
13	0.2322854D 04	0.3164528D 05
14	0.2641471D 04	0.3251308D 05
15	0.2794013D 04	0.3075612D 05
16	0.3090982D 04	0.4909522D 05
17	0.3280297D 04	0.5128404D 05
18	0.3372056D 04	0.5803713D 05
19	0.3510864D 04	0.6805937D 05
20	0.3584634D 04	0.5948737D 05
21	0.3701154D 04	0.6917885D 05

TABLE C.22

(a) $\lvert Cov(\hat{\delta}_{21})\rvert$ (b) $\lvert \overline{Cov}(\hat{\delta}_{21})\rvert$		ESTIMATORS: OLS	2SLS	UBK
NORMAL DISTRIBUTION	(a)	.37328160 09	.10309090 12	.30492400 12
	(b)	.42358760 09	.45998370 12	.15125140 13
UNIFORM DISTRIBUTION	(a)	.73495110 08	.39304930 11	.10606480 12
	(b)	.43397370 08	.24406980 12	.78068810 12
LOGNORMAL DISTRIBUTION	(a)	.46918280 08	.10878720 10	.31355810 10
	(b)	.53111580 08	.54224660 11	.15187500 12

TABLE C.23

$e = |\hat{Cov}(\hat{\delta}_{21})_{k_a}| \, / \, |\hat{Cov}(\hat{\delta}_{21})_{k_a}|$

DENOMINATOR \ NUMERATOR		NORMAL DISTRIBUTION OLS	NORMAL DISTRIBUTION 2SLS	NORMAL DISTRIBUTION UBK
NORMAL DISTRIBUTION	OLS	1.0	.27617D +03	.81687D +03
	2SLS	3.62089D -03	1.0	2.95781
	UBK	1.22417D -03	.33808	1.0
UNIFORM DISTRIBUTION	OLS	.50789D +01	.14026D +04	.41489D +04
	2SLS	.94970D -02	.26228D +01	.77579D +01
	UBK	3.51937D -03	.97196	2.87488
LOGNORMAL DISTRIBUTION	OLS	.79559D +01	.21972D +04	.64990D +04
	2SLS	3.43130D -01	.94763D +02	2.80294D +02
	UBK	1.19047D -01	.32877D +02	.97246D +02

TABLE C.24

$$e = |Cov(\hat{\delta}_{21})_{k_a}| / |Cov(\hat{\delta}_{21})_{k_a'}|$$

DENOMINATOR \ NUMERATOR		UNIFORM DISTRIBUTION OLS	UNIFORM DISTRIBUTION 2SLS	UNIFORM DISTRIBUTION UBK
NORMAL DISTRIBUTION	OLS	1.96889D -01	1.05295D +02	.28414D +03
	2SLS	7.12915D -04	3.81264D -01	1.02884
	UBK	2.41027D -04	1.28900D -01	.34784
UNIFORM DISTRIBUTION	OLS	1.0	.53479D +03	.14431D +04
	2SLS	1.86987D -03	1.0	.26985D +01
	UBK	6.92926D -04	3.70574D -01	1.0
LOGNORMAL DISTRIBUTION	OLS	1.56644	.83773D +03	.22606D +04
	2SLS	6.75586D -02	3.61301D +01	.97497D +02
	UBK	2.34390D -02	1.25351D +01	.33827D +02

TABLE C.25

$$e = |\mathrm{Cov}(\hat{\delta}_{21})_{k_a}| / |\mathrm{Cov}(\hat{\delta}_{21})_{k_{a'}}|$$

DENOMINATOR \ NUMERATOR		LOGNORMAL DISTRIBUTION		
		OLS	2SLS	UBK
NORMAL DISTRIBUTION	OLS	1.25691D -01	.29143D +01	.84000D +01
	2SLS	4.55115D -04	1.05525D -02	3.04156D -02
	UBK	1.53868D -04	.35676D -02	1.02831D -02
UNIFORM DISTRIBUTION	OLS	.63838	.14801D +02	.42663D +02
	2SLS	1.19369D -03	.27677D -01	.79775D -01
	UBK	4.42354D -04	1.02566D -02	2.95628D -02
LOGNORMAL DISTRIBUTION	OLS	1.0	.23186D +02	.66830D +02
	2SLS	4.31284D -02	1.0	2.88230
	UBK	1.49631D -02	.34694	1.0

TABLE C.26

$$e = \left|\overline{Cov}(\hat{\delta}_{21})_{k_a}\right| / \left|\overline{Cov}(\hat{\delta}_{21})_{k_a}\right|$$

DENOMINATOR \ NUMERATOR		NORMAL DISTRIBUTION		
		OLS	2SLS	UBK
NORMAL DISTRIBUTION	OLS	1.0	1.08592D +03	.35707D +04
	2SLS	.92087D -03	1.0	.32881D +01
	UBK	2.80055D -04	3.04118D -01	1.0
UNIFORM DISTRIBUTION	OLS	.97606D +01	1.05993D +04	.34852D +05
	2SLS	1.73551D -03	1.88463	.61970D +01
	UBK	.54258D -03	.58920	.19374D +01
LOGNORMAL DISTRIBUTION	OLS	.79754D +01	.86607D +04	.28478D +05
	2SLS	.78117D -02	.84829D +01	.27893D +02
	UBK	2.78905D -03	3.02869	.99589D +01

TABLE C.27

$e_o = |\overline{Cov}(\hat{\delta}_{21})_{k_a}| / |\overline{Cov}(\hat{\delta}_{21})_{k_a}|$

DENOMINATOR \ NUMERATOR		UNIFORM DISTRIBUTION		
		OLS	2SLS	UBK
NORMAL DISTRIBUTION	OLS	1.0245 D -01	.57619D +03	1.84303D +03
	2SLS	.94345D -04	.53060	1.69720
	UBK	2.86922D -05	1.61366D -01	5.16157D -01
UNIFORM DISTRIBUTION	OLS	1.0	.56240D +04	1.79892D +04
	2SLS	1.77807D -04	1.0	3.19862
	UBK	.55588D -04	.31263	1.0
LOGNORMAL DISTRIBUTION	OLS	.81709	.45954D +04	1.46990D +04
	2SLS	.80032D -03	.45010D +01	1.43972D +01
	UBK	2.85744D -04	1.60704	5.14033

TABLE C.28

$$e = |\widehat{Cov}(\hat{\delta}_{21})_{k_a}| / |\widehat{Cov}(\hat{\delta}_{21})_{k_a}|$$

DENOMINATOR \ NUMERATOR		LOGNORMAL DISTRIBUTION OLS	LOGNORMAL DISTRIBUTION 2SLS	LOGNORMAL DISTRIBUTION UBK
NORMAL DISTRIBUTION	OLS	1.25385D -01	1.28012D +02	.35854D +03
	2SLS	1.15464D -04	1.17883D -01	.33017
	UBK	3.51147D -05	3.58506D -02	1.00412D -01
UNIFORM DISTRIBUTION	OLS	1.22384	1.24949D +03	.34996D +04
	2SLS	2.17608D -04	2.22168D -01	.62226
	UBK	.68031D -04	.69457D -01	.19453
LOGNORMAL DISTRIBUTION	OLS	1.0	1.02095D +03	.28595D +04
	2SLS	.97947D -03	1.0	.28008D +01
	UBK	3.49705D -04	3.57034D -01	1.0

TABLE C.29

$e = \lvert Cov(\hat{\delta}_{21})_{k_a} \rvert / \lvert Cov(\hat{\delta}_{21})_{k_{a'}} \rvert$		NORMAL DISTRIBUTION			UNIFORM DISTRIBUTION			LOGNORMAL DISTRIBUTION		
DENOMINATOR \ NUMERATOR		OLS	2SLS	UBK	OLS	2SLS	UBK	OLS	2SLS	UBK
NORMAL DISTRIBUTION	OLS	1	+	+	−	+	+	−	+	+
	2SLS	−	1	+	−	−	+	−	−	−
	UBK	−	−	1	−	−	−	−	−	−
UNIFORM DISTRIBUTION	OLS	+	+	+	1	+	+	−	+	+
	2SLS	−	+	+	−	1	+	−	−	−
	UBK	−	−	+	−	−	1	−	−	−
LOGNORMAL DISTRIBUTION	OLS	+	+	+	+	+	+	1	+	+
	2SLS	−	+	+	−	+	+	−	1	+
	UBK	−	+	+	−	+	+	−	−	1

TABLE C.30

$e = |\overline{Cov}(\hat{\delta}_{21})_{k_s}| / |\overline{Cov}(\hat{\delta}_{21})_{k_s}|$

DENOMINATOR \ NUMERATOR		NORMAL DISTRIBUTION OLS	NORMAL DISTRIBUTION 2SLS	NORMAL DISTRIBUTION UBK	UNIFORM DISTRIBUTION OLS	UNIFORM DISTRIBUTION 2SLS	UNIFORM DISTRIBUTION UBK	LOGNORMAL DISTRIBUTION OLS	LOGNORMAL DISTRIBUTION 2SLS	LOGNORMAL DISTRIBUTION UBK
NORMAL DISTRIBUTION	OLS	1	+	+	-	+	+	-	+	+
	2SLS	-	1	+	-	-	+	-	-	-
	UBK	-	-	1	-	-	-	-	-	-
UNIFORM DISTRIBUTION	OLS	+	+	+	1	+	+	+	+	+
	2SLS	-	+	+	-	1	+	-	-	-
	UBK	-	-	+	-	-	1	-	-	-
LOGNORMAL DISTRIBUTION	OLS	+	+	+	-	+	+	1	+	+
	2SLS	-	+	+	-	+	+	-	1	+
	UBK	-	+	+	-	+	+	-	-	1