1973

The Finite Sampling Properties Of Several Estimators Of The Linear Regression Models With Random Coefficients

Baldev Raj

Follow this and additional works at: https://ir.lib.uwo.ca/digitizedtheses

THE FINITE SAMPLING PROPERTIES OF

SEVERAL ESTIMATORS OF THE LINEAR

REGRESSION MODELS WITH RANDOM COEFFICIENTS

by

Baldev Raj

Department of Economics

Submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Faculty of Graduate Studies

The University of Western Ontario

London, Canada

September, 1972

ABSTRACT

Interest in the linear random regression models has
been increasing in the last decade or so. The regression
coefficient in such models is assumed to possess a probability
distribution, and in that sense it is a random variable. Then
the mean and variance of this probability distribution are
estimated from given set of observations. These models pro-
vide an elegant technique for handling inter-unit heterogeneity
found especially in cross-section behaviour. The model is
appropriate in all situations where relevant variables (eco-
nomic and non-economic) cannot be explicitly included in the
regression - either becuase they cannot be measured or data
is not available etc.

Under simplifying assumptions the model reduces to a
classical model with the means as regression coefficients
whose disturbances are heterskedastic. The least squares
estimators for the means are inefficient. The minimum variance
Aitken's estimates for means are not directly obtainable because
the diagonal elements in the variance-covariance matrix of
disturbances are functions of variances, the unknown parameters.
Thus we require estimates of the variances not only for their
own sake but as an aid to obtain efficient estimators for means.
This thesis discusses several estimators some of which have

been developed and/or generalized by us.

The study provides a basis for selecting a suitable estimator among several methods now available. We compute by Monte Carlo experimentation the quality of alternative estimators in terms of bias and dispersion of each estimator for parameters, and of forecasts made from the estimated equations. In addition some indication of the relative cost is provided. The biases of standard errors obtained from asymptotic formulas are determined.

The samples sizes studied are 10, 20 and 50. Three models are analyzed at two different points (chosen well apart) in the parameter space. To mention a few among several conclusions that emerge from our study. we find that (i) all estimators of means and variances are unbiased (ii) rankings of alternative estimators do change both across sample sizes and parameter values (iii) iterative estimators analyzed generally converge within five iterations and (iv) iterative procedures show rather small gain in efficiency in comparison to increased cost of computation.

The study also makes several theoretical contributions. The problem common to all estimators, that sometimes negative estimates of variance may appear, is discussed both analytically and geometrically. It is argued that in seeking more efficient estimators of variances, we reduce the probability of obtaining negative estimates of variances. This theoretical conjecture does hold in our empirical investigations. Further,

it is argued that in seeking efficient estimators of variances we reduce but do not eliminate the possibility of obtaining negative estimates of variance. The solution therefore lies in search for an estimator which provides guaranteed positive estimates and also has other desirable properties. It turns out that the simple proportional model with random coefficient always yields a guaranteed positive maximum likelihood estimator of variance, without any non-linearity problems found with this method when applied in general. This property is used to develop a stepwise least square estimator analogous to the maximum likelihood estimator which is consistent and generally positive. Hopefully, this is an improvement over other methods. Here we have developed only theory.

In addition a theoretical proof of unbiasedness of mean response coefficients estimators is developed. The formulas for calculating standard error of forecasts and coefficient of multiple correlation corresponding to alternative estimators are proposed. A more general assumption for the mean of random coefficients wherein the mean is functionally related to excluded and/or included variable/s is proposed. The thesis concludes with several suggestions for further research.

ACKNOWLEDGEMENT

Thanks are also due to Professors R. G. Bodkin,
R.A.L. Carter, S.C. Clark, T.J. Courchene, P.G. Kirkham,
K.L. Krishna and Balvir Singh. The author had the benefit
of discussion with them on one or the other aspect of the
thesis. Professor Carter was particularly helpful on the
programming part of the research.

Last but not the least my sincere appreciation is
due to my wife and my parents who provided the necessary
affection and moral support which helped a great deal in
speedier completion of the project.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

Page

TABLE OF CONTENTS (cont'd)

LIST OF TABLES

LIST OF TABLES (cont'd)

LIST OF FIGURES

NOMICLATURE

CMR	:	Constant Mean Response
GLS	:	Generalized Least Squares
HH	:	Hildreth and Houck
IHH	:	Iterative Hildreth and Houck Estimator
ISALS	:	Iterative Step-wise Aitken's Least Squares
ITWLS	:	Iterative Theil's Weighted Least Squares
Model I	:	The Proportional CMR Model
Model II	:	The Bivariate CMR model with an Intercept
Model III	:	The Trivariate CMR model without an Intercept
OLS	:	Ordinary Least Squares
PERBIA	:	Percentage Bias
PERT QRT DIV	:	Percentage Quartile Deviation
PMEDBI	:	Percentage Median Bias
PRMSE	:	Percentage Root Mean Square Error
PSD	:	Percentage Standard Deviation
Q1	:	First Quartile
Q3	:	Third Quartile
QUART DIV	:	Quartile Deviation
RMSE	:	Root Mean Square Error
SALS	:	Step-wise Aitken's Least Squares
SD or Sd	:	Standard Deviation
SE	:	Standard Error
SET I	:	First Set of Values chosen for the Sampling Experiments ('Small' values)
SET II	:	Second Set of Values chosen for the Sampling Experiments ('Large' values)
STAD DIV	:	Standard Deviation (Corrected for degrees of freedom)
t	:	t-statistic
TWLS	:	Theil's Weighted Least Squares
Var or VAR	:	Variance
VMR	:	Variable Mean Response
XMEDBI	:	Median Bias
XMSE	:	Mean Square Error
% Bias	:	Percentage Bias
% E	:	Percentage Efficiency
% M Bias	:	Percentage Median Bias
% RMSE	:	Percentage Root Mean Square Error
% QD	:	Percentage Quartile Deviation
% -ve δ_1^2	:	Percentage Negative Estimates of Variance of Random Coefficients β_1 (t)
% -ve δ_2^2	:	Percentage Negative Estimates of Variance of Random Coefficient β_2 (t)

LIST OF APPENDICES

CHAPTER I

PURPOSE AND PLAN OF THE STUDY

I.1   INTRODUCTION

In classical linear regression analysis it is assumed
that the regression co-efficients are constant over the entire
sample period in time series or that they do not change from
observation to observation in cross sections. It can be argued
that this assumption is rather restrictive. For example, sup-
pose in demand analysis we are regressing the quantity demanded
of a certain commodity on, say; the price of that commodity and
some other explanatory variables. If we are using time series
data on all relevant variables, then it is quite likely that
the price elasticity of demand (or the regression coefficient)
does not remain the same over the sample period. The change
may occur because of changes in tastes over time. Similarly
in a cross section study of the production function of firms,
the labour and capital elasticities might vary from small
firms to large firms due to economies of scale or managerial
abilities etc. Therefore, it seems desirable to postulate
linear regression models which permit regression coefficients
to change from observation to observation.

Interest in the linear regression models with random
coefficients has been increasing in the last decade or so.

In such models we assume that the regression coefficient possesses a probability distribution and in that sense, it is a random variable. It is a common practice to estimate the mean and variance of this probability distribution from the given set of observations. In fact, the classical linear regression model with constant coefficients can be interpreted as a particular case of the regression model with random coefficients, where only the intercept term is postulated to be a random variable.

If we are using a regression model with random coefficients and an estimate of the variance of random coefficients turns out to be statistically insignificant, this would then, support the hypothesis that the regression coefficient being considered is, in fact, constant. Thus under this framework of random coefficient regressions we have a statistical test for the non-constancy of a regression coefficient.

The linear regression model with random coefficients has been put to various uses in econometric analyses in the past. Zellner (1966) used the model in the context of the aggregation problem. He showed that if we aggregate micro relations,we end up with a macro relation which has coefficients changing over time. The only way to handle such macro relations is to consider them as random coefficient regressions; then there will be no aggregation bias.[1]

---

[1]See also Swamy (1968), pp. 9-13.

Nerlove (1965) proposed the use of this model in the specification and estimation of demand for output and supply of input functions for an industry when variables like managerial ability, technological progress etc. cannot be properly accounted for in the specification.

In a different context Swamy (1970) specified and analyzed the linear regression model with random coefficients in an aggregate consumption study of twenty-four countries using panel data.[2] Some early attempts to specify models with a random intercept using time series of cross-section data were made by Kuh (1959), Mundlak (1963) and Hock (1962). The analysis of covariance was applied to such linear regression models with a random intercept.

The linear regression model with random coefficients may be used for regional, urban or quantitative economic analyses where data on some of the explanatory variables are practically non-existent and some economic and non-economic forces are not clearly specified.

In a recent study Singh, Nagar and Raj (1972) used the random coefficients model to analyze parameter shifts. The authors have shown that the parameter shifts can be built into the specification and analyzed statistically. In a separate study on the Econometric Model for the Indian economy it has been proposed by Professor A. L. Nagar to put the random coefficients model to extensive use of

_____

[2]Time series of cross-section data.

various sectors of the economy. In the blueprints of the Indian Econometric Model, it is proposed to carry out a large number of micro studies (in cross-section and time series alike) within the frame of the linear regression model with random coefficients to collect basic information about the size and direction of economic forces. This basic information then, shall be used in identifying important parameters in the macro equation and in making meaningful disaggregation for the macro equations of the econometric model.

With the recognition of the importance of the random coefficients model, it would now be helpful if users have some basis for selecting a most suitable method of estimation from among the several methods now available. To assist in this selection is the purpose and the objective of the present project. To achieve this goal we shall study the quality of the various methods in terms of bias and dispersion of both estimates of the parameters, and of forecasts made from the estimated equations. In addition we shall give some indication of the relative costs of the methods. Given qualities and relative costs, the user or consumer has some basis for selecting the best method for handling his particular problem.

I.2   SCOPE OF THE STUDY

1.2.1   Specification and Estimation in the Linear
          Regression Model with Random Coefficients

If we let the regression coefficients change from observation to observation, we will have to estimate too many parameters. It is ideal but hopeless to estimate and run tests on all parameters in such situations. Therefore, one has to make some simplifying assumptions about these coefficients. In the present study, we assume that they are random variables and we will estimate their means and variances only. In Chapter II, we discuss three different specifications of such random coefficient regression models. In the first, we simply assume that the mean of the probability distribution of a particular coefficient remains constant over the sample period or over various observation units. We call this the Constant Mean Response (CMR) model. Our second specification permits the mean response coefficient to change in a systematic fashion over sample units. Therefore, we call it the Variable Mean Response (VMR) model. Last of all, we let the mean response be functionally related to excluded and/or included explanatory variables. For example, the marginal propensity to consume may be interpreted as a function of an included variable like income, or of excluded variables like rate of interest, ratio of nonhuman to human wealth etc. as proposed by Friedman (1957).

The estimation of these models is discussed in Chapter III. It has been noted that for the purposes of

estimation all three specifications can be reduced to the same format as the CMR Model. The only difference is that in the latter two specifications the number of parameters to be estimated is large.

The methods of estimation employed are the following:

    i) Ordinary Least Squares Estimation (OLS)
   ii) Hildreth and Houck Estimation (HH)
  iii) Step-Wise Aitken's Least Squares Estimation (SALS)
   iv) Theil's Weighted Least Squares Estimation (TWLS)
    v) Maximum Likelihood Estimation (MLE).

In the estimation of mean and/or standard errors of mean response coefficients, we need to estimate the variance of random coefficients first.[3] A major drawback common to all the methods is that the estimate of variance sometimes turns out to be negative. Some ad-hoc methods of overcoming this difficulty have been discussed in Chapter III. In Chapter VI we propose a stepwise estimation method analogous to the maximum likelihood estimation of the linear regression model with random coefficients as a means of getting positive consistent estimators for the variances. As a digression we note that the Maximum Likelihood Method applied to the bivariate regression model with random coeffi-

---

[3]The Ordinary Least Squares (OLS) method of estimation gives estimates of mean response coefficients only. Therefore, we need extraneous estimates of variances to obtain consistent OLS estimates of the standard errors of the mean response coefficients.

cients always provides positive estimates of the variances.
We also note that the Step-Wise Least Squares concept devel-
oped by Golderger (1961) can be employed in the present
context. Therefore we consider step-wise regressions and
propose a consistent estimation of variances obtained
analogous to the Maximum Likelihood Estimators. We cannot
call this estimator a Maximum Likelihood estimator.
However, the estimator possesses two desirable properties:

      (a)   the estimate of variance will be positive, and
      (b)   the estimator is consistent.

This is hopefully an improvement over the methods mentioned
above.

I.2.2   <u>Monte Carlo Study of Several Methods of</u>
              <u>Estimators</u>

      The exact sample properties of the estimators listed
above are not known and need a great deal of theoretical
analysis. In the present study we seek to examine the
small/^{sample}properties of several estimators through sampling
experiments.[4] In our study we have done bias analysis and
efficiency ranking of several estimators for the mean
response and variances of random coefficients. The descrip-

----

[4]The Maximum Likelihood Estimator requires solution of non-
linear equations through an iterative procedure. This
estimator has not been included in sampling experiments
due to limited computational funds and research time.
For the same reasons the new estimator developed in
Chapter VI is not studied by Monte Carlo experiments in
this project.

tive statistics used are a combination of statistics based on first two moments (i.e. bias, standard deviation and root mean square error) of the sampling distribution and the corresponding non-parametric statistics (i.e. median bias and quartile deviation). These descriptive statistics have generally been expressed in percentage form to make them scale free. Various descriptive statistics have been defined in Chapter IV.

In addition to making efficiency comparison of several estimators, we have compared the estimators on the basis of their forecasting ability and the coefficient of multiple correlation. The biases of estimators of the asymptotic standard errors have also been computed to determine the least biased estimator. The sampling experiments extend to three models. For each of the three models the analysis has been carried out for more than one point in the parameter space in order to broaden and strengthen the conclusions. The results of the sampling experiments are discussed in Chapter V.

The generation of random numbers is an important aspect of a Monte Carlo study. The standard pseudo-random numbers commonly used in sampling studies suffer from two handicaps - (i) the hypothesis that they come from fixed (uniform) population does not hold in many instances and (ii) they are serially correlated. The pseudo-random numbers used for this study do not suffer from these

deficiencies. These pseudo-standard normal deviates have been generated by a computer routine developed by Carter (1972). This routine has built in test for the hypothesis that they are from a uniform population with no serial correlation. The technique of random number generation is briefly discussed in Chapter IV.

1.2.3 Efficiency and Convergence of Alternative Iterative Estimators

In the Hildreth and Houck, the Step-Wise Aitken's Least Squares and the Theil's Weighted Least Squares methods of estimation, we must estimate the variance first and then the mean response of random coefficients. These methods of estimation obtain the estimates for the variances of random coefficients from the Ordinary Least Squares residuals. Intuitively it seems that we can obtain more efficient estimators for the variances of random coefficients (hence the mean estimator for mean response coefficients) if we use the generalized least squares residuals. Thus, the Hildreth and Houck, the Step-Wise Aitken's Least Squares and Theil's Weighted Least Squares estimators of variances and mean response coefficients may be regarded respectively, as the initial step for the following iterative estimators:

(vi)   Iterative Hildreth and Houck Estimator (IHH)
(vii) Iterative Step-Wise Aitken's Least Squares Estimator (ISALS)
(viii) Iterative Theil's Weighted Least Squares Estimator. (ITWLS).

These iterative estimators are discussed in Chapter III. The properties of these iterative estimators are examined through Monte Carlo experiments in Chapter V. For instance, we shall examine if there is a gain in efficiency in terms of smaller standard error of estimates; if the proportion of negative estimate of the variance of the random coefficients is reduced, and finally, whether the iterative schemes really converge.

The conclusions of our investigation of the sampling properties of alternative estimators of the linear regression models with random coefficients are given in Chapter VII.

CHAPTER II

SPECIFICATION OF THE LINEAR

REGRESSION MODEL WITH RANDOM COEFFICIENTS

II.1   MODEL SPECIFICATION

Consider a general linear regression model of the following type:

$$y(t) = \sum_{j=1}^{K} \beta_j(t) \, x_j(t) + u(t), \quad t = 1, 2, \ldots, T; \quad j = 1, \ldots, K \tag{II.1}$$

where $y(t)$ is the t-th observation on the dependent variable and $x_j(t)$ is the t-th observation on the j-th explanatory variable. There are K explanatory variables in the model and T observations on all variables. $\beta_j(t)$ is the regression coefficient of the j-th explanatory variable corresponding to the t-th observation. It represents the partial response rate at which the dependent variable changes for a unit change in the j-th explanatory variable for the t-th observation, when influence of all other explanatory variables is held constant. In terms of partial derivatives, it can be represented as

$$\beta_j(t) = \frac{\partial \, y(t)}{\partial \, x_j(t)}, \quad \begin{array}{l} t = 1, \ldots, T, \\ j = 1, \ldots, K. \end{array} \tag{II.2}$$

It is clear that there are KT partial response rates which

are unknown in the model (II.1) and need investigation. We will come to this point later. The u(t)'s are unobservable disturbances in the model which are assumed to be independently and identically distributed with zero mean and constant finite variance, σ_0^2.

In the specification of the model (II.1), we have formulated the regression coefficients as function of t. This feature of the general linear model needs explanation. It implies that the structural coefficients relating to the x's differ for each observation.[1] For example, consider a simple bivariate regression relation between savings and income of households. Suppose we have sample on savings from a cross-section of households at various income levels. Select any two households who have the same level of income but different wealth holdings. These two households will very likely have different savings response to a unit change in income. The reason for this is simple; wealth holdings also affect the savings behaviour of households. This shows that the structural coefficients will change from observation to observation if explanatory variables (like wealth) have been omitted from the model. Even if we include wealth in our model, there are scores of other factors such as age of households, number of persons in households, marital

---

[1] We are assuming that the form of the function between y and x's does not change with observation. Further, we are concerned with linear structures only.

status, social status, price expectations, income expectations and many more, which affect the savings behaviour of households and all cannot be simultaneously included in the savings income relationship; therefore, the regression model will have different partial savings response coefficients from observation to observation.

Consider another example. We have a cross-section sample of firms to explain variation in output by level of inputs, labour and capital. The change of output per unit change in capital/labour is likely to differ from firm to firm if the firms differ in size, managerial ability, technology etc. Thus in a simple regression relation of output and two factor inputs, capital and labour, the model such as (II.1) appears appropriate.

The objective in specifying the general linear model, which allows the regression coefficients to change from observation to observation, is to relax the restrictive assumption of the classical general linear model wherein the regression coefficients are assumed constant over the entire sample period. In the latter model, it is assumed that excluded variables effect only the intercept term, not other response coefficients; while, in the former model we generalize this principle by assuming that excluded variables effect all response coefficients. This point will become clear in subsequent discussion.

We have rationalized the model (II.1) for cross-

section studies by considering two examples. However, the model is equally plausible for time series also. For example, in a simple bivariate regression of inventory investment on the sales of a firm, the response of inventory investment to a change in sales will differ from year to year if business conditions, expectations, governmental policy etc., which also affect the inventory investment, change over time. Similarly, if we were to explain variation in output of a firm by level of capital and labour inputs, changes in output per unit change in each input will differ from year to year if technical progress takes place, better working conditions are instituted in the firm etc., as these factors also affect the output of a firm.

The specification of the model (II.1) is equally justified for macro economic behaviour. In a study of aggregate consumption, the changes in income of equal amount will produce different effects on aggregate consumption if consumer attitudes, income expectations, price expectation etc., change from year to year.

In the above discussion we provided certain illustrations (in micro and macro economics) where the use of the model (II.1), with changing coefficients, might be appropriate. We emphasized that the response coefficients change from sample unit to sample unit due to the fact that not all variables (economic and non-economic) can be explicitly introduced in the model. The variables are usually omitted from

the model for various reasons--because they cannot be measured quantitatively or data may not be available or because there are not enough degrees of freedom or simply they cannot be identified. In case of a homogeneous sample, if the model is well specified in terms of the explanatory variables, one may be able to use the regression model with constant coefficients. In other words, if the omitted variables could be held constant (as is possible only in laboratory experimentation in physical sciences) then regression model with constant coefficients could still be applicable.

The model (II.1), as noted earlier, has KT unknown regression coefficients to be estimated. We can however, only estimate and run tests on less than T parameters in the model because we have only T observations in the sample. It is, therefore, necessary that we impose certain simplifying restrictions on the structural coefficients of the model. These restrictions may take one of the following forms.

A.   Grouping of Observations

The sample can be grouped into small homogeneous strata. For example, in a study of savings behaviour of households in the cross-section, the single and married households can be put in two separate groups and savings behaviour of each analysed separately. In this way the effect of marital status on savings behaviour can be controlled within each group. For each of these two groups further grouping of households can be done with respect to social

status, making the effect of socio-status on savings behaviour constant within each group and so on. The stratification reduces the number of parameters to be estimated since the number of explanatory variables required in the regression is reduced.[2] Further, the structural coefficients can be assumed to be independent of the observations in the sample and we can use the familiar regression model with constant coefficients because the effect of omitted variables is then considerably reduced. However, this approach is conditional to the availability of data in suitable form and the number of observations required will be rather large.

B. Use of Prior Information

If we have prior information about the structural coefficients it can be utilized in their estimation along with the sample data. Either one may then employ mixed estimation as suggested by Theil and Goldberger (1961) or alternatively follow the Bayesian approach. This subject is beyond the scope of the present study.

C. Random Coefficients Regression Approach

Another approach to the problem is to make some simplifying assumptions about the structural coefficients.

---

[2]We are assuming that we have sufficient data and we don't run into the degrees of freedom problem. There may be a methodology point here - why not introduce dummies? It may be pointed out that use of dummies is a discrete case of the regression model with random coefficients discussed as Assumption C.

We may, for example, assume that they are random variables with some specified properties. In the following analysis we shall elaborate this approach in detail.

II.2 Constant Mean Response Approach

This approach requires that, for any specified j, the regression coefficient $\beta_j(t)$, $t = 12 \ldots, T$, is a random variable distributed independently and identically with constant mean and finite variance.

Algebraically we can write

$$E\left[\beta_j(t)\right] = \bar{\beta}_j \qquad (II.3)$$

where $\bar{\beta}_j$ is the mean response which is independent of t.

$$E\left[\beta_j(t) - \bar{\beta}_j\right]^2 = \text{Var } \beta_j(t) = \sigma_j^2 \qquad (II.4)$$

where σ_j^2 is a finite constant,

$$E\left[\bar{\beta}_j(t) - \bar{\beta}_j\right]\left[\beta_{j'}(t') - \bar{\beta}_{j'}\right] = 0 ; \qquad (II.5)$$

$$\text{if } j \neq j' \text{ or } t \neq t' ,$$
$$j, j' = 1, \ldots, K$$
$$t, t' = 1, \ldots, T,$$

i.e. $\beta_j(t)$'s are uncorrelated among themselves and across the sample units. We may then write

$$\beta_j(t) = \bar{\beta}_j + \varepsilon_j(t) ; \quad \begin{array}{l} t = 1, \ldots, T \\ j = 1, \ldots, K \end{array} \qquad (II.6)$$

where

$$E \ \varepsilon_j(t) = 0 \qquad \text{for all } j \text{ and } t \qquad \text{(II.7)}$$

$$\text{Var } \varepsilon_j(t) = \sigma_j^2 \qquad \text{(II.8)}$$

$$\text{Cov} \left[\varepsilon_j(t) \ \varepsilon_{j'}(t') \right] = 0 \quad \text{if } j \neq j' \text{ or} \qquad \text{(II.9)}$$
$$t \neq t'$$

This assumption implies that $\beta_j(t)$, $t = 1, 2, \ldots$
.., T, fluctuates around the constant mean $\bar{\beta}_j$ with constant
variance σ_j^2. The situation can be pictured as in Fig. II.1,
for any specified j. The disturbances centered around $\bar{\beta}_j$
are the assumed distribution for $\varepsilon_j(t)$.
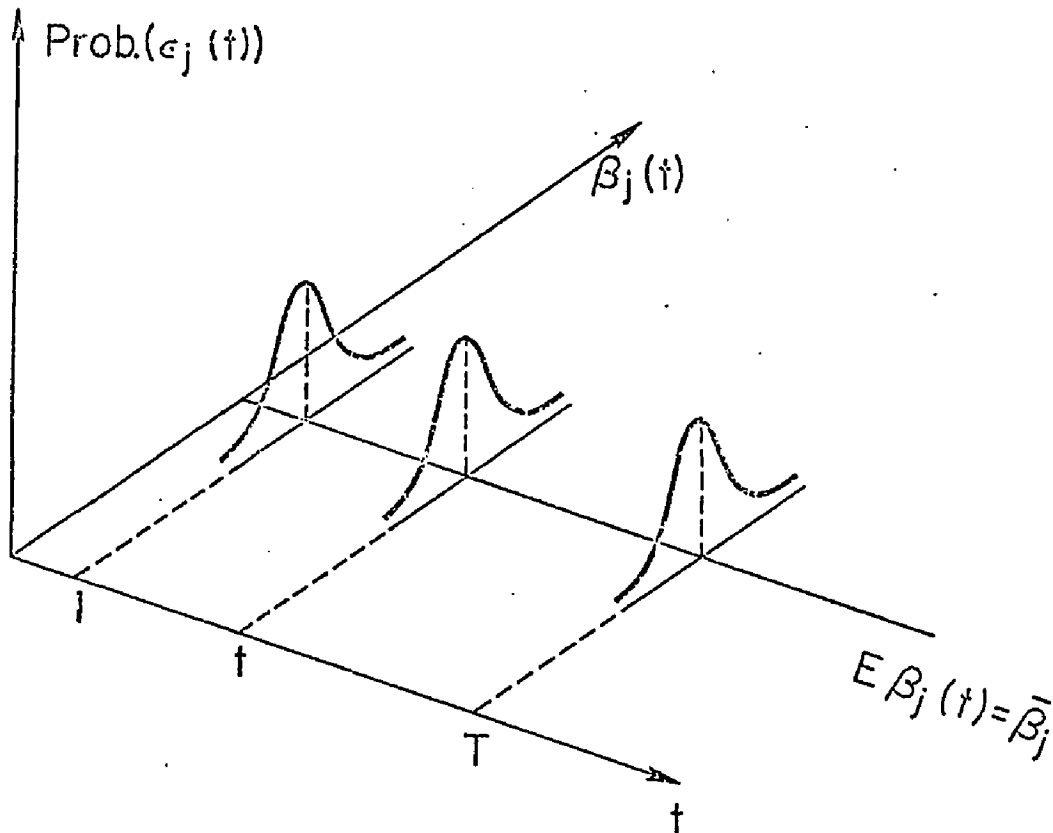
Fig. II.1

Since the mean response $\bar{\beta}_j$ does not depend on t we call this approach the constant mean response approach (CMR).

Substituting (II.6) in (II.1), we have

$$y(t) = \sum_j (\bar{\beta}_j + \varepsilon_j(t)) \, x_j(t) + u(t) ; \qquad (II.10)$$

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

or

$$y(t) = \sum_j \bar{\beta}_j \, x_j(t) + u^*(t) \qquad (II.11)$$

where

$$u^*(t) = \sum_j \varepsilon_j(t) \, x_j(t) + u(t) \qquad (II.12)$$

If we make an additional assumption that x's are non stochastic, then if follows from (II.7) - (II.9) that

$$Eu^*(t) = 0 \text{ because } E\,\varepsilon_j(t) = Eu(t) = 0 \text{ for all j and t (II.13)}$$

Finally

$$\text{Var } u^*(t) = \sum_j x_j^2(t) \, \sigma_j^2 + \sigma_o^2 \qquad (II.14)$$

where

$$\text{Var } u(t) = \sigma_o^2$$

Thus in the model (II.11), we have to estimate $2K + 1$ unknown parameters ($\bar{\beta}_j(t)$'s, $\sigma_j^2(t)$'s, and σ_0^2). So long as this number of parameters is less than T (number of

observations) they can be estimated.

Here we digress a little to discuss the role of the disturbance term u(t) in the model (II.1). It may be pointed out that the disturbance term in any regression model is included to account for omitted variables in the regression. One can also postulate that there is an inherent element of randomness in individual behaviour even after all systematic causes have been included. Since the concept of random coefficients in the model (II.1) has been introduced for similar reasons, one may do away with the disturbance term from the model, at least for the sake of simplicity. Retaining the disturbance term would add one more parameter to be estimated in the model (II.1) (viz., σ_0^2). This additional term cannot be estimated separately if the model (II.1) has an intercept, for in this case

$$\text{variance of } u^*(t) = \left[\sum_{j=1}^{k} x_j^2(t) \ \sigma_j^2 \right] + \sigma_0^2 = (\sigma_0^2 + \sigma_1^2) +$$

$\sum_{j=2}^{k} x_j^2(t) \ \sigma_j^2$ (assuming $x_1(t) = 1$ for all t) and it is diffi-

cult to unscramble the estimates of σ_0^2 and σ_1^2 (cf. Klein (1953)). Following the convention in the literature, we drop u(t) from the model (II.1) and (II.11).

To sum up

$$y(t) = \sum_j \beta_j(t) \ x_j(t) \ ; \ t = 1, 2, \ldots, T, \qquad (II.15)$$

where $\beta_j(t)$'s are random variables. This may be written as

$$y(t) = \sum_j \bar{\beta}_j \, x_j(t) + \eta(t) \; ; \; t = 1, \ldots\ldots T \qquad \text{(II.16)}$$

where

$$\eta(t) = \sum_j \varepsilon_j(t) \, x_j(t) \qquad ; \; t = 1, \ldots\ldots T \qquad \text{(II.17)}$$

It follows that

$$E\eta(t) = 0 \text{ and } \text{Var } \eta(t) = \phi^2(t) = \sum_j x_j^2(t) \, \sigma_j^2 \qquad \text{(II.18)}$$

$$\text{for } t = 1, \ldots\ldots, \; T.$$

A variety of mixed models (where some regressions coefficient are random and others constant) can be constructed as special cases of (II.16) to suit particular need of the researcher. In particular, the familiar general linear regression model with constant coefficients is obtained if we postulate that $x_1(t) = 1$, for all t and the regression coefficient attached to it in the model (II.15) is random, while other regression coefficients are constant (i.e. all other random coefficients have zero variances). Thus we have

$$y(t) = \beta_1(t) + \sum_{j=2}^{k} \bar{\beta}_j \; x_j(t) \qquad \text{(II.19)}$$

$$= \bar{\beta}_1 + \sum_{j=2}^{k} \bar{\beta}_j \; x_j(t) + \varepsilon_1(t) \; .$$

where $\bar{\beta}_1$ and $\bar{\beta}_j$ are all constants.

II.3   Variable Mean Response Approach

The argument for constant mean response assumption could be extended one step further; suppose $\bar{\beta}_j$'s for any specified j is not constant but shifts (again shifts must be systematic or estimation will be hopeless). For example, it is likely that in time series analysis the marginal propensity to consume (or any other such parameter) shows a systematic shift. This systematic shift might take place due to various reasons, for example, due to change in the attitude of the people, changes in institutions, technological change etc. Similarly in cross-section analysis that the mean response may have a systematic shift due to occupational, sociological, geographical factors etc. Suppose that in such situations we write

$$E \, \beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j \, f_j(t) \qquad\qquad (II.20)$$

where $f_j(t)$ is a continuous function of t (the sample unit) which possesses partial derivatives at least up to second order.[3]

In the simplest form $f_j(t) = t$, then,

$$E \, \beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j t \qquad\qquad (II.21)$$

---

[3]This idea has been explored in a separate paper by Singh, Nagar and Raj (1972).

i.e. $\beta_j(t)$'s fluctuate around a linear trend.

In other instances, it may be suitable to assume that

$$f_j(t) = \log t \tag{II.22}$$

then

$$E\,\beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j \log t \tag{II.23}$$

i.e. $\beta_j(t)$ fluctuates around a semi-log linear trend.

Or suppose,

$$f_j(t) = e^t \tag{II.24}$$

in which case,

$$E\,\beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j e^t \tag{II.25}$$

implying that $\beta_j(t)$ fluctuates around an exponential trend.

In fact, in the upswing and downswing of business cycles the parameter shifts may be such that they systematically increase during the upswing and fall during the downswing or vice-versa. In such situations $f_j(t)$ may take the following form,

$$f_j(t) = \delta_1 + \delta_2 t + \delta_3 t^2 \tag{II.26}$$

then,

$$E\, \beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j \left[\delta_1 + \delta_2 t + \delta_3 t^2 \right] \qquad \text{(II.27)}$$

$$= (\bar{\beta}_j + \bar{\alpha}_j\, \delta_2) + (\bar{\alpha}_j\, \delta_2)t + (\bar{\alpha}_j\, \delta_3)\, t^2$$

i.e. $\beta_j(t)$ fluctuates around a second degree polynomial trend.[4]

The major difficulty that might arise in this approach is the specification of the form of $f_j(t)$, j=1, ..., K. However, given computing facilities it is possible to try out different forms of $f_j(t)$, j=1, 2, ... K and choose the one which yeilds maximum explained variation (i.e. multiple correlation). Further it is difficult to unscramble the individual coefficients $\bar{\beta}_j$ and $\bar{\alpha}_j$ from δ_1, δ_2 and δ_3 for formulation (II.27).

We may, in a general case, write

$$\beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j\, f_j(t) + \varepsilon_j(t) \; ; \; \begin{array}{l} t=1, \ldots\ldots, \; T \\ j=1,2, \ldots\ldots, \; K. \end{array}$$

$$\text{(II.28)}$$

and assume that

$$E\, \varepsilon_j(t) = 0 \qquad \text{for all t and j} \qquad \text{(II.29)}$$

$$\text{Var}\, \varepsilon_j(t) = \sigma_j^2 \; , \; \text{a finite constant} \qquad \text{(II.30)}$$

---

[4]Perhaps here, $\beta_j(t) = \delta_1 \cos (\delta_2 t + \delta_3)$ would be more appropriate. However, if we are considering only one cycle, quardratic form will serve the purpose.

$$\text{Cov} \left[\varepsilon_j(t) \; \varepsilon_{j'}(t') \right] = 0 \;, \; \text{if } j \neq j' \text{ or } t \neq t' \tag{II.31}$$

The hypothesis (II.28) can be pictured as in Fig. II.2 for any specified j and $f_j(t) \equiv t$. The disturbances around the line $\bar{\beta}_j + \bar{\alpha}_j t$ are the assumed distribution of $\varepsilon_j(t)$.
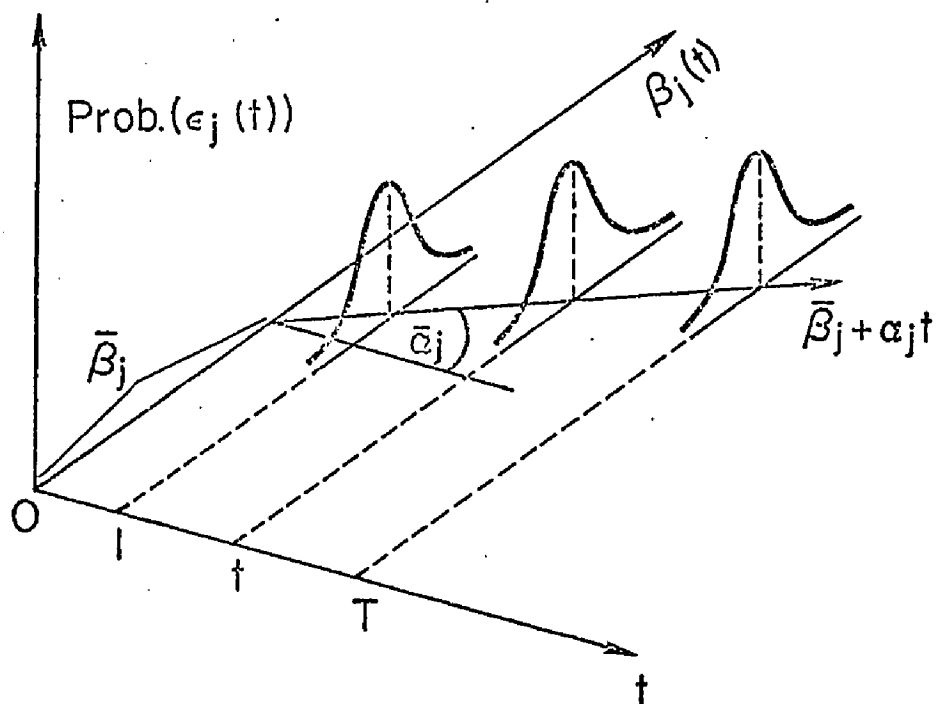
Fig. II.2

Since mean response varies with observation, we may call this approach the *variable mean response* approach (VMR). It may be pointed out that the CMR approach is a particular case of VMR approach when $\bar{\alpha}_j = 0$ for all j.

Substituting (II.28) in (II.15) we have

$$y(t) = \sum_j (\bar{\beta}_j + \bar{\alpha}_j f_j(t) + \varepsilon_j(t)) x_j(t) \qquad \text{(II.32)}$$

$$= \sum_j \bar{\beta}_j x_j(t) + \sum_j \bar{\alpha}_j x_j^*(t) + \eta(t)$$

where

$$x_j^*(t) = f_j(t) x_j(t) \; ; \quad j=1,2,\ldots\ldots, K \qquad \text{(II.33)}$$

$$\eta(t) = \sum_j x_j(t) \, \varepsilon_j(t), \quad t=1,2,\ldots\ldots, T \qquad \text{(II.34)}$$

Assuming that x's are non stochastic, it follows from (II.29) - (II.31) that

$$E \, \eta(t) = 0 \; , \quad \text{for all } t \qquad \text{(II.35)}$$

$$\text{Var } \eta(t) = \sum_j x_j^2(t) \, \sigma_j^2 \quad ; \; t=1,\ldots\ldots, T \qquad \text{(II.36)}$$

In the model (II.32), if $f_j(t)$ is prespecified, there are 3K unknown parameters $\bar{\beta}_j$'s, $\bar{\alpha}_j$'s and σ_j^2's. So long as the number of parameters is less than T they can be estimated.

II.4   Mean Response as a Function of Some Explanatory and/or Excluded Exogenous Variables.

In the preceding sections we have discussed the cases where mean response either fluctuates around a constant

or about some specified trend. However, one cannot rule out
the possibility that the mean response is a function of some
included explanatory variables or even excluded exogenous
variables. For example, Friedman (1957) assumed that the
marginal propensity to consume is a function of rate of
interest, ratio of nonhuman to human wealth etc. Similarly
in a separate study Agrawala and Drinkwater (1972) assumed
that the marginal propensity to consume is a function of
socio-economic variables, such as female participation, which
make the consumption function to shift over time. By way of
illustration we may consider a cross-section sample of house-
holds to explain the savings behaviour by their levels of
income. Assume further that the sample consists of house-
holds with varying amounts of wealth holdings. In this situ-
ation, it is likely that marginal propensity to save be
functionally related to wealth holdings of households, even
when wealth is explicity introduced in the regression.

Thus it may be reasonable to assume that

$$E \, \beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j \, g_j(\mathbf{z}(t)) \, , \, j = 1, \ldots\ldots, K \qquad \text{(II.37)}$$

where $g_j(\mathbf{z})$ is a continuous function of exogenous (excluded
or included) variable $\mathbf{z}(t)$, which possess partial deriva-
tive at least up to second order.[5]

_____

[5]Here we have included one explanatory variable in the

The simplest form $g_j(z)$ might take is

$$g_j(z) = z \tag{II.38}$$

then

$$E \beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j z \tag{II.39}$$

i.e. $\beta_j(t)$ fluctuates around the linear function in z.

The other simple possibilities are:

$$g_j(z) = \log z \tag{II.40}$$

and

$$g_j(z) = e^z \tag{II.41}$$

More complicated form of $g_j(z)$ are possible. For example,

$$g_j(z) = \gamma_1 + \gamma_2 z + \gamma_3 z^2 \tag{II.42}$$

then

$$E \bar{\beta}_j(t) = \bar{\beta}_j + \bar{\alpha}_j \left[\gamma_1 + \gamma_2 z + \gamma_3 z^2\right] \tag{II.43}$$

$$= (\bar{\beta}_j + \bar{\alpha}_j \gamma_1) + (\bar{\alpha}_j \gamma_2) z + (\bar{\alpha}_j \gamma_3) z^2$$

i.e. $\beta_j(t)$ fluctuates around a second degree polynominal function in z.

---

function g_j, for simplicity reasons. The assumption can be generalized to include more explanatory variables.

The optimal form of $g_j(\beta)$ can be selected on the basis of the criterion of maximum explanation of variation in terms of multiple correlation. However, for complicated formulations of the functional form of $g_j(\beta)$ such as (II.42), it seems difficult to separate individual coefficients $\bar{\beta}_j$ and $\bar{\alpha}_j$ from γ_1, γ_2 and γ_3.

Now, let us write

$$\beta_j(t) = \bar{\beta}_j + \bar{\alpha}_j \, g_j(\beta) + \epsilon_j(t) \ , \quad t=1, \ldots\ldots, T \quad (II.44)$$

and assume

$$E \, \epsilon_j(t) = 0 \qquad \text{for all } j \text{ and } t, \qquad (II.45)$$

$$Var \, \epsilon_j(t) = \sigma_j^2 \ , \text{ a finite constant, and} \qquad (II.46)$$

$$Cov\left[\epsilon_j(t) \, \epsilon_{j'}(t')\right] = 0 \ , \text{ if } j \neq j' \text{ or } t \neq t' \qquad (II.47)$$

Substituting (II.44) in (II.15), we have

$$y(t) = \sum_j (\bar{\beta}_j + \bar{\alpha}_j \, g_j(\beta) + \epsilon_j(t)) \, x_j(t) \qquad (II.48)$$

or

$$y(t) = \sum_j \bar{\beta}_j \, x_j(t) + \sum_j \bar{\alpha}_j \, x_j^{**}(t) + \eta(t) \qquad (II.49)$$

where

$$x_j^{**}(t) = g_j(\beta(t)\, x_j(t)) \; ; \quad \begin{matrix} t=1,2, \ldots\ldots, T \\ j=1,2, \ldots\ldots, K \end{matrix} \qquad (II.50)$$

and

$$\eta(t) = \sum_j x_j(t)\, \varepsilon_j(t) \qquad\qquad (II.51)$$

Under the additional assumption that x^{**}'s are non stochastic, it follows from (II.45) - (II.47) that

$$E\,\eta(t) = 0 \quad \text{for all } t \;\Big/\; \overset{\text{and}}{\text{Var}}\,\eta(t) = \sum_j x_j^2(t)\, \sigma_j^2\,, \qquad (II.52)$$

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

The parameters of the model to be estimated are: $\bar{\beta}_j$'s, $\bar{\alpha}_j$'s and σ_j^2's,

It is easy to see that this approach embraces the CMR and the VMR models as particular cases.

CHAPTER III

ESTIMATION OF THE LINEAR REGRESSION

MODELS WITH RANDOM COEFFICIENTS

III.1   Matrix Formulation of the General Linear Regression
         Models with Random Coefficients.

For estimation purposes it is useful to write the

general linear models with random coefficients, discussed

in Sections II.2 - II.4, in matrix notation, as follows.

The CMR model (II.16) can be written as

$$y = X \bar{\beta} + \eta \qquad \text{(III.1)}$$

where

$$y_{T \times 1} = \begin{bmatrix} y(1) \\ \cdot \\ \cdot \\ \cdot \\ y(T) \end{bmatrix} \quad ; \quad X_{T \times K} = \begin{bmatrix} x_1(1) & \cdots & x_K(1) \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ x_1(T) & \cdots & x_K(T) \end{bmatrix} \quad ;$$

$$\bar{\beta}_{K \times 1} = \begin{bmatrix} \bar{\beta}_1 \\ \cdot \\ \cdot \\ \cdot \\ \bar{\beta}_K \end{bmatrix} \quad \text{and} \quad \eta_{T \times 1} = \begin{bmatrix} \eta(1) \\ \cdot \\ \cdot \\ \cdot \\ \eta(T) \end{bmatrix}$$

such that the elements of X are fixed constants and

$$
E\eta = \begin{bmatrix} E\eta(1) \\ \cdot \\ \cdot \\ \cdot \\ E\eta(T) \end{bmatrix} = 0_{T \times 1} \; , \tag{III.2}
$$

$$
E\eta\eta' = \Phi_{T \times T} = \begin{bmatrix} E\eta^2(1) & \cdots & E\eta(1)\eta(T) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ E\eta(T)\eta(1) & \cdots & E\eta^2(T) \end{bmatrix} = \begin{bmatrix} \phi^2(1) & & \\ & \cdot & 0 \\ & \cdot & \\ & \cdot & \\ 0 & & \phi^2(T) \end{bmatrix} \tag{III.3}
$$

where[1]

$$
\phi^2(t) = \sum_j x_j^2(t) \, \sigma_j^2 \; , \quad t = 1, 2 \ldots, T, \tag{III.4}
$$

as defined in (II.18).

Similarly, the VMR model (II.32) in matrix notation may be written as

$$
y = X^* \, \bar{\beta}^* + \eta. \tag{III.5}
$$

where y and η are column vectors as defined in (III.1) and

$$
X^*_{T \times 2K} = \begin{bmatrix} x_1(1) & \cdots & x_K(1) & x_1^*(1) & \cdots & x_K^*(1) \\ \cdot & & \cdot & \cdot & & \\ \cdot & & \cdot & \cdot & & \\ \cdot & & \cdot & \cdot & & \\ x_1(T) & \cdots & x_K(T) & x_1^*(T) & \cdots & x_K^*(T) \end{bmatrix} \tag{III.6}
$$

---

[1] The summation over j always runs from 1 through K and over t from 1 through T.

such that $x_j^*(t)$, for all j and t, has been defined in (II.33). Thus the elements of x* are nonstochastic and fixed in repeated samples

$$\bar{\beta}^*_{2K \times 1} = \begin{bmatrix} \bar{\beta}_1 \\ \cdot \\ \cdot \\ \cdot \\ \bar{\beta}_K \\ \bar{\alpha}_1 \\ \cdot \\ \cdot \\ \cdot \\ \bar{\alpha}_K \end{bmatrix} \qquad (III.7)$$

where $\bar{\beta}_j$'s and $\bar{\alpha}_j$'s are defined in (II.28) above.

Also the disturbance vector η is the same as in model (III.1) with Eη = 0 and Eηη' = Φ defined in (III.2) and (III.3) respectively.

Finally, the model (II.49) may be written in matrix notation as

$$y = X^{**}\bar{\beta}^* + \eta \qquad (III.8)$$

where the vectors y and η are as defined in (III.1) and the vector $\bar{\beta}^*$ in (III.7) above.

The matrix X** of order Tx2K is defined as

$$
X^{**}_{T \times 2K} = \begin{bmatrix} x_1(1) & \cdots & x_K(1) & x_1^{**}(1) & \cdots & x_K^{**}(1) \\ \vdots & & \vdots & \vdots & & \\ \vdots & & \vdots & \vdots & & \\ x_1(T) & \cdots & x_K(T) & x_1^{**}(T) & \cdots & x_K^{**}(T) \end{bmatrix} \qquad \text{(III.9)}
$$

where $x_j^{**}(t)$, for all j and t has been defined in (II.50). The elements of the matrix x^{**} are non-stochastic fixed in repeated samples. Once again, the assumptions of the disturbance vector η are the same as in the CMR model (III.1) and are given in (III.2) – (III.4) above.

In the models (III.7), (III.5) and (III.8) the disturbances are independently and identically distributed with zero mean and heteroskedastic variances $\phi^2(t)$'s. The elements of the matrices of explanatory variables, X, X* and X** (defined in (III.1), (III.6) and (III.9), respectively) are non-stochastic/fixed in repeated samples. Thus the models differ only in the number of parameters to be estimated. We may therefore restrict our discussion, which follows in subsequent sections, to the CMR model (III.1) only, without any loss of generality.

III.2   <u>Ordinary Least Squares Estimator</u>[2]

For estimation, we assume further that

$$\rho(X) = K \leq T \qquad \text{(III.10}$$

i.e. the columns of X are assumed to be linearly independent.

The ordinary least squares (OLS) estimate of $\bar{\beta}$ in the CMR model (III.1) is given as

$$\bar{b} = (X'X)^{-1} X'y \qquad \text{(III.11)}$$

$$= Ay$$

where

$$A = (X'X)^{-1} X' .$$

Under the assumptions stated above, it is easy to show the OLS estimator $\bar{b}$ is linear, unbiased and consistent estimator of $\bar{\beta}$.

The linearity of $\bar{b}$ is obvious from (III.11), for any jth element of $\bar{b}$ is given by

$$\bar{b}_j = \sum_j A_{ij} y_j \qquad \text{(III.12)}$$

---

[2] It is well known that this estimator is not appropriate in the present context because it yields an inefficient estimator for mean response. It is included primarily in the study for its historical importance. Besides, since the covariance matrix Φ is not known a priori, it is not entirely clear if other estimators which use the estimated covariance matrix will be more efficient than the OLS estimator (cf. Rao (1970). Therefore, the OLS estimator has some comparative value also. We will come to this point later while discussing the results of the sampling experiments in Chapter V.

where A_{ij} is an element in the ith row and jth column of matrix A, defined in (III.11) above.

In order to prove unbiasedness, we must show that the expected value of sampling error of $\bar{b}$ is zero

$$E(\bar{b} - \bar{\beta}) = 0 \qquad\qquad (III.13)$$

Substituting the value of y from (III.1) in (III.11), we have

$$\bar{b} = (x'x)^{-1} x'[x\bar{\beta} + \eta] \qquad\qquad (III.14)$$

or

$$\bar{b} - \bar{\beta} = (x'x)^{-1} x'\eta$$

Taking expectation on both sides, we have

$$E(\bar{b}) = \bar{\beta} \qquad\qquad (III.15)$$

($\because$ $E\eta = 0$, from (III.2) and elements of X are nonstochastic and fixed in repeated samples). Hence $\bar{b}$ is an unbiased estimator.

To prove consistency we must show that $\operatorname{Plim} (\bar{b} - \bar{\beta})$ = 0. Taking the probability limit of the sampling error given in (III.14), we have

$$\operatorname*{Plim}_{T \to \infty} (\bar{b} - \bar{\beta}) = \operatorname*{Plim}_{T \to \infty} (x'x)^{-1} x'\eta \qquad\qquad (III.16)$$

$$= \operatorname*{Plim}_{T \to \infty} (\frac{1}{T} x'x)^{-1} . \operatorname*{Plim}_{T \to \infty} (\frac{1}{T} x'\eta)$$

Now assuming that

$$\text{Plim} \quad (\tfrac{1}{T} x'x)^{-1} \qquad\qquad\qquad \text{(III.17)}$$

is bounded, and noting that

$$\underset{T \to \infty}{\text{Plim}} \tfrac{1}{T} x'\eta = \underset{T \to \infty}{\lim} E \ (\tfrac{1}{T} x'\eta) = 0 \qquad\qquad \text{(III.18)}$$

(because $E\eta = 0$ and elements of x's are non-stochastic and fixed in repeated samples), we have

$$\text{plim } \bar{b} = \bar{\beta} \ , \qquad\qquad\qquad \text{(III.19)}$$

i.e. $\bar{b}$ is a consistent estimator of $\bar{\beta}$.

However, the estimator $\bar{b}$ will not be efficient because $E\eta\eta' \neq \sigma^2 I$, i.e. the covariance matrix of the disturbances is not a scalar times an identity matrix as required under the classical Gauss-Markov assumptions. In particular, the disturbances are heteroskedastic.

Given the unbiasedness of $\bar{b}$, the covariance matrix of the estimator $\bar{b}$ can be obtained easily as

$$\begin{aligned}
\text{Var } (\bar{b}) &= E[\bar{b}-\bar{\beta}) \ (\bar{b}-\bar{\beta})'] \qquad\qquad \text{(III.20)}\\
&= (x'x)^{-1} x' (E\eta\eta') \ x(x'x)^{-1}\\
&= (x'x)^{-1} x' \ \Phi \ x(x'x)^{-1}
\end{aligned}$$

because

$$E\eta\eta' = \Phi \qquad\qquad\qquad \text{(III.21)}$$

The elements Φ are not completely determined because they involve unknown parameters σ_j^2's.

The traditional formula $\hat{\sigma}_0^2 \, (x'x)^{-1}$ for obtaining the variance of $\bar{b}$ would yield biased results and hence is not suitable.

The co-variance matrix Φ of the disturbance vector η in the CMR model (III.1) is diagonal, a typical element of which is, $\phi^2(t) = \sum_j x_j^2(t) \, \sigma_j^2$. Since $\phi^2(t)$ is a function of observations on the explanatory variables, it will change over time. In other words the disturbances are hetroskedastic.

We need to estimate the unknowns σ_j^2's for several reasons.

(i) If we have these estimates, we may obtain variance of $\bar{b}$ in (III.20).

(ii) We may obtain the Aitken estimate of $\bar{\beta}$.

(iii) The hypothesis regarding the randomness of regression coefficients can be tested only if we have these estimates and knowledge of their sampling distribution.

III.3   Hildreth and Houck Estimator

Hildreth and Houck (1968) have proposed an estimation procedure to estimate $\bar{\beta}_j$'s and σ_j^2's in the model (III.1), which may be described briefly in steps as

1) Fit the model (III.1) by OLS yielding the fitted residuals;

2) regress the squares of the OLS residuals as obtained in step 1 above on a given function of x's yielding the estimators of variances;

3) obtain an estimator of the covariance matrix Φ using the estimates of the variances as obtained in step 2 above; and finally,

4) obtain an Aitken estimator of the mean response coefficients using the estimated covariance matrix Φ as obtained in step 3 above.

· Having described various steps involved for the Hildreth and Houck (HH) method of estimation, we now discuss their theoretical foundations. Consider the OLS fitted residuals in (III.1)

$$\hat{\eta} = y - x\bar{b} \tag{III.22}$$

where y, X and η are defined in (III.1) and $\bar{b}$ is defined in (III.11). We may write (III.22) as

$$\hat{\eta} = My \tag{III.23}$$

where

$$M = I - X(X'X)^{-1} X' \tag{III.24}$$

is a symmetric idempotent matrix of order TxT. Substituting the value of y from (III.1) in (III.23) we obtain

$$\hat{\eta} = M(X\bar{\beta} + \eta) \qquad\qquad (III.25)$$

$$= M\eta \qquad (\because MX = 0)$$

Let the elements of the matrix M be given as

$$M_{TxT} = \begin{bmatrix} m_{11} & \cdots & m_{1T} \\ \vdots & & \vdots \\ m_{T1} & \cdots & m_{TT} \end{bmatrix} \qquad\qquad (III.26)$$

and since these elements are functions of X's only they are known and fixed in repeated samples.

Using (III.26) we can write the elements of the vector $\hat{\eta}$, defined in (III.25), as

$$\hat{\eta} = \begin{bmatrix} \hat{\eta}(1) \\ \vdots \\ \hat{\eta}(T) \end{bmatrix} = \begin{bmatrix} \sum_t m_{1t}\, \eta(t) \\ \vdots \\ \sum_t m_{Tt}\, \eta(t) \end{bmatrix} \qquad\qquad (III.27)$$

Taking expectations on both sides of (III.27) we have

$$E\hat{\eta} = \begin{bmatrix} E \sum_t m_{1t}\, \eta(t) \\ \vdots \\ E \sum_t m_{Tt}\, \eta(t) \end{bmatrix} = 0 \qquad\qquad (III.28)$$

Define a column vector $\overset{\circ}{\overset{\wedge}{\eta}}$ whose elements are the squares of elements of the column vector $\overset{\wedge}{\eta}$ defined in (III.27), i.e.

$$\overset{\circ}{\overset{\wedge}{\eta}} = \begin{bmatrix} \overset{\wedge}{\eta}^2(1) \\ \vdots \\ \overset{\wedge}{\eta}^2(T) \end{bmatrix} = \begin{bmatrix} \left(\sum_t m_{1t}\, \eta(t)\right)^2 \\ \vdots \\ \left(\sum_t m_{Tt}\, \eta(t)\right)^2 \end{bmatrix} \quad . \qquad (III.29)$$

The expected value of the vector $\overset{\circ}{\overset{\wedge}{\eta}}$ can be obtained from (III.29), using assumptions (III.2) $-$ (III.4), as

$$E\overset{\circ}{\overset{\wedge}{\eta}} = \begin{bmatrix} E\left[\sum_t m_{1t}\, \eta(t)\right]^2 \\ \vdots \\ E\left[\sum_t m_{Tt}\, \eta(t)\right]^2 \end{bmatrix} = \begin{bmatrix} \sum_t m_{1t}^2\, \phi^2(t) \\ \vdots \\ \sum_t m_{Tt}^2\, \phi^2(t) \end{bmatrix} \qquad (III.30)$$

where $\phi^2(t)$ for $t = 1, 2, \ldots, T$, is defined in (III.4).
In matrix notation, (III.30) can be written as

$$E\,\overset{\circ}{\overset{\wedge}{\eta}} = \overset{\bullet\bullet}{M}\,\overset{\bullet}{X}\,\overset{\bullet}{\sigma} \qquad (III.31)$$

where

$$\overset{\bullet\bullet}{M}_{TxT} = \begin{bmatrix} m_{11}^2 & \cdots & m_{1T}^2 \\ \vdots & & \vdots \\ m_{T1}^2 & \cdots & m_{TT}^2 \end{bmatrix} \quad ; \quad \overset{\bullet}{X}_{TxK} = \begin{bmatrix} x_1^2(1) & \cdots & x_K^2(1) \\ \vdots & & \vdots \\ x_K^2(T) & \cdots & x_K^2(T) \end{bmatrix}$$

$$(III.32)$$

and

$$\dot{\sigma}_{Kx1} = \begin{bmatrix} \sigma_1^2 \\ \cdot \\ \cdot \\ \cdot \\ \sigma_K^2 \end{bmatrix} \cdot \tag{III.33}$$

Let us write

$$\hat{\eta} = \ddot{M X} \dot{\sigma} + \xi \tag{III.34}$$

where ξ is a disturbance vector such that $E\xi = 0$.

The elements of the matrix $\ddot{M X}$ are functions of X's and hence independent of the elements of the vector ξ. We may therefore, use the regression relation (III.34) to estimate $\dot{\sigma}$ as:

$$\dot{s} = (\dot{z}' \ \dot{z})^{-1} \ \dot{z}' \ \hat{\eta} \tag{III.35}$$

where

$$\dot{z}_{TxK} = \ddot{M X} \tag{III.36}$$

and

$$\dot{s} = \begin{bmatrix} s_1^2 \\ \cdot \\ \cdot \\ \cdot \\ s_K^2 \end{bmatrix} = \begin{bmatrix} \text{est } \sigma_1^2 \\ \cdot \\ \cdot \\ \cdot \\ \text{est } \sigma_K^2 \end{bmatrix} \cdot \tag{III.37}$$

It is easy to show that the estimator $\dot{s}$ is unbiased and consistent. Substituting the value of $\hat{\eta}$ from (III.34) in (III.35), the sampling error of $\dot{s}$ can be obtained, i.e.

$$\dot{s} - \dot{\sigma} = (\dot{Z}'\dot{Z})^{-1} \dot{Z}'\xi \quad . \tag{III.38}$$

It follows that

$$E(\dot{s} - \dot{\sigma}) = 0 , \quad (\because E\dot{\xi} = 0), \tag{III.39}$$

i.e. $\dot{s}$ is an unbiased estimator of $\dot{\sigma}$.

Further $\dot{s}$ is a consistent estimator, if

$$\underset{T \to \infty}{\text{Plim}} \ \dot{s} = \dot{\sigma}. \tag{III.40}$$

It follows from (III.38) that

$$\underset{T \to \infty}{\text{Plim}} (\dot{s} - \dot{\sigma}) = \underset{T \to \infty}{\text{Plim}} (\frac{1}{T} \dot{Z}'\dot{Z})^{-1} \underset{T \to \infty}{\text{Plim}} (\frac{1}{T} \dot{Z}'\xi) \tag{III.41}$$

Now assume that

$$\underset{T \to \infty}{\text{Plim}} (\frac{1}{T} \dot{Z}'\dot{Z})^{-1} \tag{III.42}$$

is bounded, and noting that[2a]

$$\underset{T \to \infty}{\text{Plim}} \frac{1}{T} \dot{Z}'\xi = \underset{T \to \infty}{\text{Lim}} E(\frac{1}{T} \dot{Z}'\xi) = 0 \tag{III.43}$$

(because $E\dot{\xi} = 0$ and the elements of $\dot{Z}$ are non-stochastic and fixed in repeated samples),

---

[2a]Hussain and Wallace (1967) have shown that $\underset{T \to \infty}{\text{Plim}} \ g_T = \theta$ does not necessarily imply $\underset{T \to \infty}{\text{Lmt}} E(g_T) = \theta$, as is erroneously believed sometimes. Of course if g_T is uniformally bounded (i.e., $|g| \leqslant M < \infty$ for all T), then $g_T \overset{p}{\to} \theta \implies Eg_T \to \theta$. (Also see original reference [5] in Hussain and Wallace). Throughout this study we will continue to assume that g_T is continuously bounded.

we have

$$\text{Plim } \dot{s} = \dot{\sigma} \quad , \qquad \text{(III.44)}$$

i.e. $\dot{s}$ is a consistent estimator of $\dot{\sigma}$.

The estimator $\dot{s}$ will be efficient if the disturbances in ξ are uncorrleated and homoskedastic, i.e. $E(\xi\xi') = \sigma^2 I$. This is unfortunately not true, for

$$E(\xi\xi') = E[\hat{\dot{\eta}} - E\hat{\dot{\eta}}) \; (\hat{\dot{\eta}} - E\hat{\dot{\eta}})'] \qquad \text{(III.45)}$$

$$= E(\hat{\dot{\eta}}\hat{\dot{\eta}}') - E(\hat{\dot{\eta}}) \; E(\hat{\dot{\eta}}') \; .$$

It is shown in Appendix 1 that

$$E(\hat{\dot{\eta}}\hat{\dot{\eta}}') = E(\hat{\dot{\eta}}) \; E(\hat{\dot{\eta}}') + 2 \; \dot{\ddot{\Psi}} \qquad \text{(III.46)}$$

where $\dot{\ddot{\Psi}}$ is a matrix of order TXT whose elements are the squares of the elements of the matrix Ψ of order TXT and

$$\dot{\Psi} = M \; \Phi \; M \qquad \text{(III.47)}$$

where M and Φ have been defined in (III.24) and (III.4) respectively.

Thus,

$$E(\xi\xi') = 2\dot{\ddot{\Psi}} \; .^3 \qquad \text{(III.48)}$$

The elements of $\dot{\ddot{\Psi}}$ are all non zero and hence the disturbances

---

[3]Also see the Appendix A in Singh, Nagar and Raj (1972).

ξ's are correlated and heteroskedastic. Hence, the estimator $\dot{s}$ is not efficient.

The covariance matrix of the estimator $\dot{s}$ is given by

$$\text{Var}(\dot{s}) = E(\dot{s} - \dot{\sigma})(\dot{s} - \dot{\sigma})' \qquad \text{(III.49)}$$

$$= (\dot{z}'\dot{z})^{-1} \dot{z}' \, E(\xi\xi') \, \dot{z}(\dot{z}'\dot{z})^{-1} \, .$$

Substituting for $E(\xi\xi')$ from (III.48) in (III.49), we have

$$\text{Var}(\dot{s}) = (\dot{z}'\dot{z})^{-1} \dot{z}' (2\dot{\Psi}) \, \dot{z}(\dot{z}'\dot{z})^{-1} \, . \qquad \text{(III.50)}$$

The variances in (III.50) are not fully determined because Ψ is a function of unknown σ_j^2's. However it is appropriate to replace $2\dot{\Psi}$ by its consistent estimator $2\hat{\dot{\Psi}}$ obtained by replacing the unknown σ_j^2's by their consistent estimates s_j^2's.

It should be noted that the estimator of σ_j^2 obtained according to (III.35) will have a range of $-\infty$ to ∞. Thus it is always possible to arrive at negative estimate of σ_j^2 -- with, howsoever, small probability.

III.3.1 Problem of Negative Variances: A Digression[4]

Here we digress a little to discuss the problem of negative estimates of σ_j^2's, which might arise according to

---

[4]The arguments developed by us in this section are helpful in interpretations of the sampling results in Chapter V.

(III.35). No doubt, in the limit, the probability of getting a negative estimate will approach its lower limit zero because the estimator $\overset{\centerdot}{s}$ is consistent. The probability of getting negative values of s_j^2's according to (III.35) may be substantial in small samples especially so because the estimator in (III.35) is inefficient (its standard errors will be large) and it can be reduced in seeking a more efficient estimator. This point will become clearer by the following diagrametrical representation.
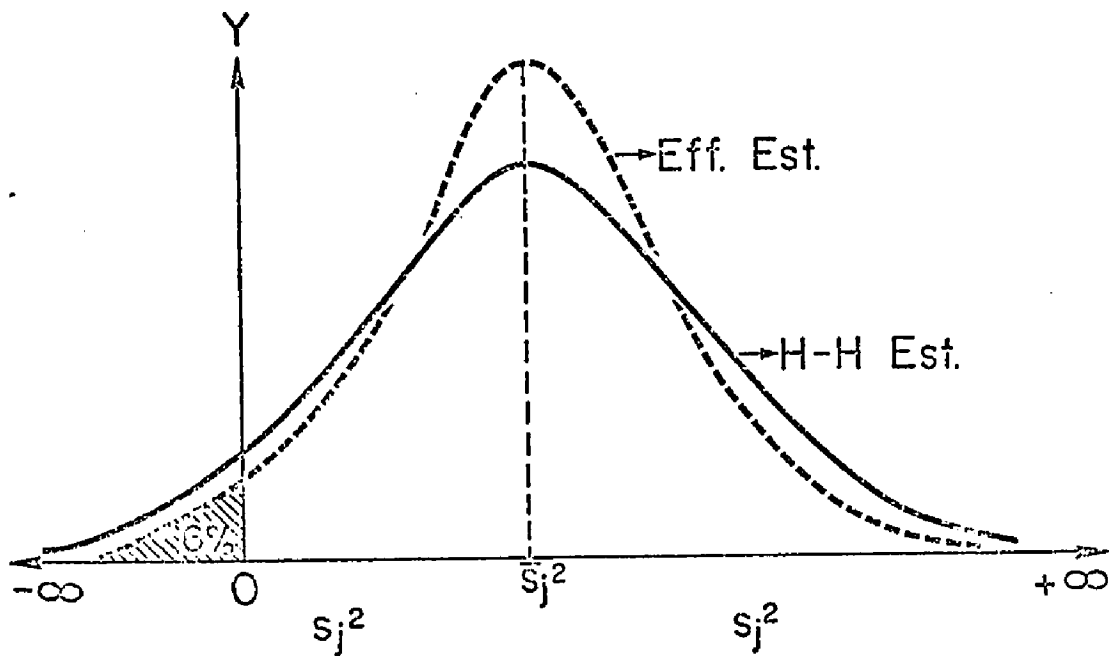
Fig. III.1

In Fig. III.1 we have drawn a hypothetical probability density function for any representative s_j^2 obtained by

the H-H method in free hand line.[5] The shaded area under the solid lined probability density function to the left of OY line gives the percentage chance of getting negative estimate (say 10%) for a given sample size (say a size 20). In the same diagram we have drawn the probability density function in broken line for an efficient estimator, which has smaller spread than the H-H estimator by definition, (the shaded area under the broken line to the left of line OY is less than that under the solid line curve). The percentage chance of getting negative estimate is reduced (say it is now 6%). It is intuitively clear from the above argument that by seeking more efficient estimates of σ_j^2's we reduce the probability of getting negative values. We will discuss a few efficient estimators in the sections that follow.

Hildreth and Houck (1968), have suggested two alternative remedies to the problem of negative estimates of σ_j^2's.

(i) We may replace negative estimates by zeros. The estimator so obtained is biased but has lower mean square error than Hildreth and Houck estimate obtained in (III.35) above.

(ii) Employ constrained minimization of errors subject to non-negativity constraints on σ_j^2's, in obtaining their

---

[5]We are implicitly assuming that the distributions of estimators are symmetric which may not hold in small samples.

estimators, i.e. use quadratic programming. The suggestion in (i) is rather arbitrary and that in (ii) has not been applied by anybody so far--perhaps due to its complex nature.[6] Further, in the latter approach there is no way to measure the precision of the estimates.

III.3.2 Hildreth and Houck's Estimator of Mean Response Coefficients.

It is well known that, if Φ is known, the BLU estimator of $\hat{\bar{\beta}}$ is Aitken's estimator, given as

$$\hat{\bar{\beta}} = (X' \Phi^{-1} X)^{-1} X' \Phi^{-1} Y \qquad (III.52)$$

and the covariance matrix of $\hat{\bar{\beta}}$ is given by

$$Var \ \hat{\bar{\beta}} = (X' \Phi^{-1} X)^{-1} \qquad (III.53)$$

However as observed earlier Φ is not known a priori because its elements are linear functions of σ_j^2's, which are unknown. In (III.35) above, we obtained s_j^2's, which are consistent estimates of σ_j^2's. We may use these to obtain consistent estimate of Φ and use it in place of Φ in (III.52) and (III.53). Thus the H-H estimator $\hat{\bar{\beta}}$ is given by

$$\hat{\bar{\beta}} = (X' \hat{\Phi}^{-1} X)^{-1} X' \hat{\Phi}^{-1} Y \qquad (III.54)$$

---

[6]In a separate problem Judge and Takayama (1966) have dealt with the inequality restrictions in regression analyses.

$$\hat{\Phi} = \begin{bmatrix} \sum_j x_j^2(1) s_j^2 & & & \\ & \cdot & & 0 \\ & & \cdot & \\ 0 & & & \cdot \\ & & & \sum_j x_j^2(T) s_j^2 \end{bmatrix} \qquad (III.55)$$

The asymptotic covariance matrix of $\hat{\bar{\beta}}$ is

$$\text{Var} (\hat{\bar{\beta}}) = (X' \hat{\Phi}^{-1} X)^{-1} \simeq (X' \Phi^{-1} X)^{-1} \qquad (III.56)$$

where

$$\hat{\Phi}^{-1} = \begin{bmatrix} \dfrac{1}{\sum_j x^2(1) s_j^2} & & & & & 0 \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & \dfrac{1}{\sum_{j=1}^{k} x_j^2(t) s_j^2} & & \\ & & & & \cdot & \\ & & & & & \cdot \\ 0 & & & & & \dfrac{1}{\sum_j x_j^2(T) s_j^2} \end{bmatrix} \qquad (III.57)$$

Hildreth and Houck (1968) have shown that $\hat{\bar{\beta}}$ as given in (III.54) is a consistent estimator. However it is not known if the estimator (III.54) is unbiased. We shall prove below that $\hat{\bar{\beta}}$ is unbiased also.

III.3.3 Proof of Unbiasedness of $\hat{\bar{\beta}}$.[7]

Assume that the elements of the vector η in (III.1) follow a continuous symmetric probability law, viz.

$$h[\eta(t)] = h[-\eta(t)] \qquad (III.58)$$

where $\eta(t)$ is the t-th element of the vector η, defined in (III.1) above.

Now consider the t-th element of $\hat{\eta}$ in (III.27)

$$\hat{\eta}(t) = \sum_t m_{tt}\, \eta(t) \qquad (III.59)$$

and the j-th element of $\dot{s}$ in (III.35).

$$s_j^2 = \sum_{t=1}^{T} P_{jt} \left[\sum_{t=1}^{T} m_{tt}\, \eta(t) \right]^2 \qquad (III.60)$$

where

$$P = ((P_{jk})) = (\dot{z}'\dot{z})^{-1}\, \dot{z}' \qquad (III.61)$$

and

$$\left[\sum_{t=1}^{T} m_{kt}\, \eta(t) \right]^2 = \hat{\eta}^2(k) \quad , \quad k = 1,\, 2,\, \ldots,\, T \qquad (III.62)$$

---

[7] The proof is based on arguments similar to the arguments in the paper by Kakwani (1967).

finally, the t-th element of matrix $\hat{\Phi}$ in (III.55) is

$$\hat{\Phi}^2(t) = \sum_{j=1}^{K} x_j^2(t) \sum_{t=1}^{T} P_{jt} \left[\sum_{t=1}^{T} m_{tt} \, \eta(t) \right]^2 \qquad \text{(III.63)}$$

Since the elements of η follow a continuous probability law, it follows from (III.59) - (III.63) that the elements of $\hat{\eta}$, $\dot{s}$ and $\hat{\Phi}$ also follow a continuous symmetric probability law, i.e. their values do not change with a change of the sign of the $\eta(t)$'s. Substituting (III.1) in (III.54), we can obtain the sampling error of $\hat{\bar{\beta}}$ as

$$\hat{\bar{\beta}} - \bar{\beta} = (X' \, \hat{\Phi}^{-1} \, X)^{-1} \, X' \, \hat{\Phi}^{-1} \, \eta \qquad \text{(III.64)}$$

Our objective is to show that $E(\hat{\bar{\beta}} - \bar{\beta}) = 0$ given that the expectation exists. Writing $(\hat{\bar{\beta}} - \bar{\beta})$ as

$$\hat{\bar{\beta}} - \bar{\beta} = H(\eta) \, \eta \qquad \text{(III.65)}$$

where

$$H(\eta) = (X' \, \hat{\Phi}^{-1} \, X)^{-1} \, X' \, \hat{\Phi}^{-1} \qquad \text{(III.66)}$$

we observe that

$$H(\eta) = H(-\eta) \, , \qquad \text{(III.67)}$$

i.e. $H(\eta)$ is an even function of the vector η in the sense that a change in sign of η does not change the value of the function.

On the other hand the sampling error given in (III.65) is an odd function of η. However, η and $-\eta$ have the same probability density function by assumption. Therefore the sampling error $(\hat{\beta} - \bar{\beta})$ and $(\bar{\beta} - \hat{\beta})$ will have same density functions also.

Thus $\hat{\beta}$ is symmetrically distributed about $\bar{\beta}$. This proves that $\hat{\beta}$ is unbiased, if its mean exists.[8]

However, the estimator $\hat{\beta}$ will not be efficient because $\hat{\Phi}^{-1}$ is obtained by using the inefficient estimates of σ_j^2's.

---

[8] The following illustration provides a theoretical background for the above proof. Consider a random variable x (say the sampling error) with density function

$$f(x)\ dx \quad ; \quad -\infty < x < \infty$$

Suppose $f(x)$ is an even function of x i.e. $f(x) = f(-x)$. Then it follows that

(1) $\int_{-\infty}^{\infty} f(x)\ dx = 2 \int_{0}^{\infty} f(x)\ dx$

because

(2) $\int_{-\infty}^{+\infty} f(x)\ dx = \int_{-\infty}^{0} f(x)\ dx + \int_{0}^{\infty} f(x)\ dx.$

Now substituting $x = -x$ in (2) above, we have

(3) $-\int_{\infty}^{0} f(-x)\ dx + \int_{0}^{\infty} f(x)\ dx = 2 \int_{0}^{\infty} f(x)\ dx.$

Now, since $f(x)$ is an even function of x (assuming that the mean exists), it can be shown that

(4) $E(x) = \int_{-\infty}^{\infty} x\ f(x)\ dx = 0$

because, (4) can be written as

(5) $E(x) = \int_{-\infty}^{0} x\ f(x)\ dx + \int_{0}^{\infty} x\ f(x)\ dx.$

Substituting $x = -x$ in (5) above, we have

(6) $E(x) = \int_{\infty}^{0} x\ f(-x)\ dx + \int_{0}^{\infty} x\ f(x)\ dx = 0$

III.3.4 Coefficient of Multiple Correlation in the CMR Model.[9]

Let us for a moment assume that the covariance matrix Φ is known and is positive definite. Then it is always possible to transform models such as (III.1) so that the Gauss-Markov assumption hold, i.e. that the transformed disturbances possess a covariance matrix which is a scalar times the identity matrix. Let S be a matrix of order TXT, whose elements are known, such that

$$S'S = \Phi^{-1} \qquad \text{and} \qquad S\Phi^{-1}S' = I \text{ .}[10] \qquad (III.68)$$

Then

$$SY = SX\bar{\beta} + S\eta \qquad (III.69)$$

is the transformed model, which satisfies all the Gauss-Markov assumptions. In the model (III.1) the disturbances $\eta(t)$'s are heteroskedastic, thus S is easy to obtain. In fact, in that case, the transformation matrix S is given by

$$S = \begin{bmatrix} \dfrac{1}{\phi(1)} & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \dfrac{1}{\phi(T)} \end{bmatrix} \qquad (III.70)$$

---

[9] The material in Sections III.3.4 and III.3.5 is specially developed by us for use as criterion in choosing among alternative estimators.

[10] Goldberger (1964), p. 36.

where $\phi(t)$ is the square root of $\phi^2(t)$, defined in (III.4).
It is easy to verify that (III.68) is true. It follows
that

$$E(S\eta) = 0 \quad (\because \text{ S is known and} \quad \text{(III.71)}$$
$$E(\eta) = 0)$$

and

$$E(S\eta\eta'S') = I \quad (\because \text{ (III.4) is true}). \text{(III.72)}$$

Writing the transformed model as

$$Y^\bullet = X^\bullet \bar{\beta} + \eta^\bullet \qquad \text{(III.73)}$$

where

$$Y^\bullet = SY \quad ; \quad X^\bullet = SX \text{ and } \eta^\bullet = S\eta. \qquad \text{(III.74)}$$

The <u>coefficient</u> of <u>multiple</u> <u>correlation</u> is the
square root of R^2 defined as

$$R^2 = 1 - \frac{\hat{\eta}^{\bullet\prime}\,\hat{\eta}^\bullet}{Y^{\bullet\prime}\,Y^\bullet} \qquad \text{(III.75)}$$

where $\hat{\eta}^\bullet = Y^\bullet - X^\bullet\hat{\bar{\beta}}$ and $\hat{\bar{\beta}} = (X^{\bullet\prime}X^\bullet)^{-1}X^{\bullet\prime}Y^\bullet$

Substituting the value of $Y^\bullet$ and $\hat{\eta}^\bullet$ defined in (III.74), we
have

$$R^2 = 1 - \frac{\hat{\eta}'\phi^{-1}\hat{\eta}}{Y'\phi^{-1}Y} \qquad \text{(III.76)}$$

where

$$\hat{\eta} = Y - X\hat{\bar{\beta}} \qquad \text{(III.77)}$$

and $\hat{\bar{\beta}}$ is defined in (III.52).

However, R^2 can not be computed because it is a function of unknown σ_j^2's.

We obtained a consistent estimator for δ in (III.35) which may be used to obtain a consistent estimator for Φ. Thus, $\hat{R}^2$ may be defined as

$$\hat{R}^2 = 1 - \frac{\hat{\eta}' \; \hat{\Phi}^{-1} \; \hat{\eta}}{Y' \; \hat{\Phi}^{-1} \; Y} \qquad (III.78)$$

where

$$\hat{\eta} = Y - X\hat{\bar{\beta}} \qquad (III.79)$$

and $\hat{\bar{\beta}}$ is the Hildreth and Houck estimator given in (III.54).

III.3.5 Prediction with the CMR Model

Assume that the CMR model (III.1) $\hat{\bar{\beta}}$ has been esti-mated by the generalized least squares procedure defined in (III.52). Suppose we are given values of the explanatory variables for m-periods of prediction and we want to predict the future values (in m-periods) of the dependent variable. Let X_f be a mxK matrix whose rows consist of observations on the explanatory variables outside the sample period.

Assume that the model (III.1) still holds for future observations, then the true values of the dependent variable will be given as

$$Y_f = X_f\bar{\beta} + \eta_f \qquad\qquad \text{(III.80)}$$

where subscript f refers to the forecasts or future values and Y_f is a column vector of size m consisting of the true unknown values of the dependent variable and η_f is an m-element disturbance vector distributed with zero mean and covariance matrix Φ_f, such that

$$E(\eta_f \eta_f') = \Phi_f = \begin{bmatrix} \phi_f^2(1) & & 0 \\ & \ddots & \\ 0 & & \phi_f^2(m) \end{bmatrix} \qquad \text{(III.81)}$$

where

$$\phi_f^2(t) = \sum_{j=1}^{K} x_{fj}^2(t)\, \sigma_j^2 \ ; \ t = 1, \ldots m \ . \qquad \text{(III.82)}$$

Let $X_f\hat{\bar{\beta}}$ be the predictor, where $\hat{\bar{\beta}}$ is defined in (III.52). Then the prediction error is given by

$$X_f\hat{\bar{\beta}} - Y_f = X_f\hat{\bar{\beta}} - (X_f\bar{\beta} + \eta_f) \qquad \text{(III.83)}$$

$$= X_f(\hat{\bar{\beta}} - \bar{\beta}) - \eta_f$$

The predictor is unbiased in the sense that the expected value of the prediction error is zero, i.e.

$$E(X_f\hat{\bar{\beta}} - Y_f) = 0 \qquad\qquad \text{(III.84)}$$

because

$$E\hat{\bar{\beta}} = \bar{\beta} \qquad \text{and} \qquad E\eta_f = 0 \qquad \qquad (III.85)$$

The sampling covariance matrix of predication error is given as

$$E\left[(X_f\hat{\bar{\beta}} - Y_f)(X_f\hat{\bar{\beta}} - Y_f)'\right] \qquad \qquad (III.86)$$

$$= E\left[(X_f(\hat{\bar{\beta}}-\bar{\beta}) - \eta_f)(X_f(\hat{\bar{\beta}}-\bar{\beta}) - \eta_f)'\right]$$

$$= X_f E(\hat{\bar{\beta}}-\bar{\beta})(\hat{\bar{\beta}}-\bar{\beta})' X_f' - X_f E(\hat{\bar{\beta}}-\bar{\beta})\eta_f'$$

$$- E\eta_f(\hat{\bar{\beta}}-\bar{\beta})' X_f' + E(\eta_f \eta_f')$$

If we assume that the elements of η are uncorrelated with those of η_f, then it is easy to verify that $E\,\eta_f(\hat{\bar{\beta}}-\bar{\beta})'X_f' = 0$ and $E(\hat{\bar{\beta}}-\bar{\beta})\eta_f' = 0$. Therefore, we can write (III.86) as

$$E\left[(X_f\hat{\bar{\beta}}-Y_f)(X_f\hat{\bar{\beta}}-Y_f)'\right] = X_f (X'\Phi^{-1}X)^{-1} X_f' + \Phi_f \qquad (III.87)$$

It can be shown that predictor $X_f\hat{\bar{\beta}}$ is the best in the class of linear unbiased predictors.

Now since Φ, Φ_f and $\hat{\bar{\beta}}$ are functions of the unknown σ_j^2, we may replace σ_j^2's by their consistent estimators s_j^2's to obtain the variance of the predictor given in (III.87) as

$$\text{Estimated Var (Prictor)} = X_f(X' \hat{\Phi}^{-1}X)^{-1} X_f' + \hat{\Phi}_f \qquad (III.88)$$

where $\hat{\Phi}^{-1}$ is defined in (III.57)

and

$$\hat{\phi}_f = \begin{bmatrix} \sum_{j=1}^{K} x_{fj}^2(1)s_j^2 & & & 0 \\ & \bullet & & \\ & & \bullet & \\ & & \bullet & \\ 0 & & & \sum_{j=1}^{K} x_{fj}^2(T)s_j^2 \end{bmatrix} \quad . \tag{III.89}$$

III.4   Iterative Hildreth and Houck Estimator

Hildreth and Houck (1968) have proposed an iterative scheme for the estimation procedure described in the preceding Section III.3, which may be briefly described in a few steps as follows:

1) Carry out Step 1 to Step 4 of the HH method of estimation discussed in Section III.3 and obtain the initial estimates of the variances and the mean response coefficients;

2) Obtain Aitken's fitted residuals in (III.1) using the initial estimates of the mean response co-efficients, obtained in Step 1 above;

3) Carry out Step 2 to Step 4 of the HH method of estimation discussed in Section III.3 to obtain fresh estimates of the mean response coefficients;

4) Obtain fresh Aitken's fitted residuals in (III.1) using the fresh estimates of the mean response coefficients, obtained in Step 3 above; and then

5) Go back to Step 3 above and keep iterating between Step 4 and Step 3 until convergence.

Although Hildreth and Houck did not give any specific reasons for iterating between different stages of their method discussed in Section III.3, intuitively it seems that we can obtain more efficient estimates of parameters of model (III.1) by the above iterative procedure.

It may be recalled that in deriving the estimator $\dot{s}$ in (III.35) we made use of the ordinary least squares residuals, defined in (III.22). We may use Aitken's fitted residuals instead and obtain fresh estimates of σ_j^2's analogous to (III.35). Then the estimators s_j^2's obtained in (III.35) and $\hat{\beta}_j$'s in (III.54) may be regarded as the initial estimates of the Iterative Hildreth and Houck (IHH) Estimator.

Given the initial estimates σ_j^2's and $\hat{\beta}_j$'s we may obtain fitted generalized least squares residuals in the CMR model (III.1) as

$$\hat{\eta}_2 = Y - X\hat{\beta}_1 \qquad\qquad (III.90)$$

where

$$\hat{\beta}_1 \equiv \hat{\beta} \quad , \text{ defined in (III.54)}, \qquad (III.91)$$

while Y and X have been defined in (III.1).

Subscripts 1, 2, etc. indicate the round of iteration considered.

Define

$$
\hat{\eta}_2 = \begin{bmatrix} \hat{\eta}_2^2(1) \\ \cdot \\ \cdot \\ \cdot \\ \hat{\eta}_2^2(T) \end{bmatrix} \qquad , \qquad \text{(III.92)}
$$

which may be used to obtain a fresh estimator of $\dot{\sigma}$

analogous to (III.35) as

$$
\dot{s}_2 = (\dot{z}'\dot{z})^{-1} \dot{z}' \hat{\eta}_2 \qquad \text{(III.93)}
$$

where

$$
\dot{s}_2 = \begin{bmatrix} s_2^2(1) \\ \cdot \\ \cdot \\ \cdot \\ s_2^2(K) \end{bmatrix} \qquad \text{(III.94)}
$$

Then the estimates $s_2^2(j)$'s may be used to obtain a fresh

estimator of covariance matrix Φ analogous to (III.55) as

$$
\hat{\phi}_2 = \begin{bmatrix} \sum_j x_j^2(1) \, s_2^2(j) & & & \\ & \cdot & & 0 \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sum_j x_j^2(T) s_2^2(j) \end{bmatrix} \qquad \text{(III.95)}
$$

The second round of iteration can be completed by obtaining a fresh estimate of $\bar{\beta}$ analogous to (III.54) as

$$\hat{\beta}_2 = (X' \hat{\Phi}_2^{-1} X)^{-1} X' \hat{\Phi}_2^{-1} Y \qquad (III.96)$$

where matrix $\hat{\Phi}_2^{-1}$ is the inverse of the matrix $\hat{\Phi}_2$ in (III.95). If we continue iterating the values of $\dot{s}$ and $\hat{\beta}$, then, at the T-th iteration, we have

$$\dot{s}_T = (\dot{s}'\dot{s})^{-1} \dot{s}' \hat{\eta}_T \qquad (III.97)$$

where

$$\hat{\eta}_T = \begin{bmatrix} \hat{\eta}_T^2(1) \\ \cdot \\ \cdot \\ \cdot \\ \hat{\eta}_T^2(T) \end{bmatrix} \quad , \qquad \hat{\eta}_T = Y - X \hat{\beta}_{T-1} \qquad (III.98)$$

and

$$\hat{\beta}_T = (X' \hat{\Phi}_T^{-1} X)^{-1} X' \hat{\Phi}_T^{-1} Y \quad . \qquad (III.99)$$

However, it is not entirely clear if the iterative scheme presented above will result in gain in efficiency in terms of smaller standard errors of the parameter estimates. Moreover, whether the proportion of negative estimates of variances σ_j^2's will be reduced. Also, whether the process will really converge. In fact, these questions require a

great deal of theoretical analysis. We shall, in the present work, look into these problems through Monte Carlo experiments.

III.5   Step-wise Aitken's Least Squares Estimator

The Step-wise Aitken's Least Squares (SALS) estimation procedure goes a step further than Hildreth and Houck method in estimating the variances of random coefficients. The steps involved in the SALS method of estimation are

1) Carry out Step 1 and Step 2 of the HH method of estimation discussed in section III.3 and obtain the initial estimators of variances;

2) Obtain an estimator of the covariance matrix $2\overset{\bullet}{\Psi}$ using the initial HH estimates of the variance as obtained in Step 1 above;

3) Obtain an Aitken estimator of the variances using the estimated covariance matrix obtained in Step 2 above;

4) Obtain an estimator of the covariance matrix Φ using Aitken's estimates of the variances obtained in Step 3 above;

5) Obtain an Aitken estimator of the mean response coefficients using the estimated covariance matrix obtained in Step 4 above.

The theoretical considerations behind above Steps are explained below:[13]

---

[13]The material in Section III.5 and III.7 was developed jointly by Professor A. L. Nagar and the author. The

The disturbances ξ's in the regression relation (III.34) were shown to be correlated and heteroskedastic and hence the estimator $\dot{s}$ as obtained in (III.35), though unbiased and consistent, is not efficient. Intuition suggests that an estimator that takes into account the covariance structure of the disturbances ξ_j's in the regression relation (III.34) will be efficient. Aitken's procedure cannot be applied directly to (III.34) because the covariance matrix of the disturbances 2Ψ defined in (III.48), is a function of the unknown σ_j^2's. We may however, apply a step-wise procedure to estimate $\dot{\sigma}$ efficiently.

Obtain $\dot{s}$ according to ordinary least squares applied to (III.34). Then, the estimators s_j^2's so obtained can be used to estimate Φ from (III.3) and finally Ψ and $\dot{\Psi}$ from (III.45) - (III.48). Next use the estimated covariance matrix $2\hat{\Psi}$, obtained above, to get generalized least squares estimator for $\dot{\sigma}$ as

$$\hat{\sigma} = (\dot{s}'(2\hat{\Psi})^{-1}\dot{s})^{-1}\dot{s}'(2\hat{\Psi})^{-1}\hat{\eta} \qquad (III.100)$$

$$= (\dot{s}'\hat{\Psi}^{-1}\dot{s})^{-1}\dot{s}'\hat{\Psi}^{-1}\hat{\eta}$$

where $\dot{s} = \dot{M}X$ is defined in (III.36).

The estimate so obtained will have the same aymptotic properties as the Aitken estimator.[11]

The asymptotic covariance matrix of $\hat{\hat{\sigma}}$ is given by

$$\text{Var}(\hat{\hat{\sigma}}) = 2(\hat{\dot{z}}' \; \hat{\Psi}^{-1} \; \dot{z})^{-1} \approx 2(\dot{z}' \; \Psi^{-1} \; \dot{z})^{-1} \qquad \text{(III.101)}$$

Using the estimate $\hat{\dot{\sigma}}$, we can obtain an estimator $\hat{\hat{\Phi}}$ from (III.3). This new estimate may then be used to obtain an efficient estimator of $\bar{\beta}$ as .

$$\hat{\hat{\beta}} = (X' \; \hat{\hat{\Phi}}^{-1} \; X)^{-1} \; X' \; \hat{\hat{\Phi}}^{-1} \; Y \; . \qquad \text{(III.102)}$$

Further, the asymptotic covariance matrix of $\hat{\hat{\beta}}$ is given by

$$\text{Var}(\hat{\hat{\beta}}) = E(\hat{\hat{\beta}}-\bar{\beta})(\hat{\hat{\beta}}-\bar{\beta})' = (X'\hat{\hat{\Phi}}^{-1}X)^{-1} \approx (X'\Phi^{-1}X)^{-1}.$$

$$\text{(III.103)}$$

The estimator $\hat{\hat{\beta}}$ will have the same aymptotic properties as Aitken's estimator obtained by using the true Φ.[12]

The proof of unbiasedness and the concepts of Multiple Correlation Coefficient and Prediction developed for the HH estimator can be easily extended for use with this estimator.

---

[11]   Theil (1971).

[12]Op. cit.

III.6   Iterative Step-wise Aitken's Least
         Squares Estimator

Analogous to the IHH estimation procedure discussed
in Section III.4, we may define an iterative procedure based
on the SALS estimator discussed in the preceding section.
We may first describe the steps involved in the Iterative
SALS (ISALS) estimation procedure as

1) Carry out Step 1 to Step 5 of the SALS method of
   estimation discussed in Section III.5 and obtain
   the initial estimates of the variances and the
   mean response coefficients;

2) Obtain Aitken's fitted residuals in (III.1) using
   the initial estimates of the mean response co-
   efficients obtained in Step 1 above;

3) Carry out Step 2 of the HH method of estimation
   discussed in Section III.3 to obtain fresh HH
   estimators of the variances.

4) Carry out Step 2 to Step 5 of the SALS method of
   estimation to obtain fresh estimators of the
   variances and the mean response coefficients;

5) Obtain fresh Aitken's fitted residuals in (III.1)
   using the fresh estimates of the mean response
   coefficients as obtained in Step 4 above;

6) Go back to Step 3 and keep iterating among Step
   3 - Step 6 until convergence.

It may be recalled that in deriving the HH esti-
mator of $\dot{\sigma}$ we used ordinary least squares estimates to
obtain the mean response coefficients defined in (III.22).
The HH estimates s_j^2's were used to obtain an estimator of

covariance matrix $2\dot{\Psi}$. Then, the estimated covariance matrix $2\hat{\dot{\Psi}}$ was used to obtain the SALS estimator $\hat{\dot{\sigma}}$ in (III.100). The SALS estimated variances are used to obtain an estimator of the covariance matrix Φ. Finally, the estimated co-variance matrix $\hat{\Phi}$ is used to obtain the SALS estimator, $\hat{\hat{\beta}}$ in (III.103). Therefore, the estimates $\hat{\sigma}_j^2$'s and $\hat{\hat{\beta}}_j^2$'s may be regarded as the first round of Iterative Step-wise Aitken's Least Squares Estimator. To begin the second round of iterations we may use $\hat{\hat{\beta}}$ in place of $\bar{b}$ to get $\dot{s}_2$ analogous to (III.35) as

$$\dot{s}_2 = (\dot{\hat{z}}'\dot{\hat{z}})^{-1} \dot{\hat{z}}'\hat{\dot{\eta}}_1 \qquad \text{(III.104)}$$

where $\dot{\hat{z}}$ is defined in (III.36) and $\hat{\dot{\eta}}_1$ is a vector of the squared elements of the vector $\hat{\eta}_1$ such that

$$\hat{\eta}_1 = Y - X\hat{\hat{\beta}} \quad . \qquad \text{(III.105)}$$

The estimator $\dot{s}_2$ obtained in (III.104) gives us a fresh estimator $2\hat{\dot{\Psi}}_2$, where $\hat{\dot{\Psi}}_2$ is the matrix of squared elements of

$$\Psi_2 = M\hat{\Phi}_2 M \qquad \text{(III.106)}$$

and

$$\hat{\Phi}_2 = \begin{bmatrix} \sum_j x_j^2(1) \, s_2^2(j) & & \\ & \ddots & 0 \\ & \ddots & \\ 0 & \ddots & \\ & & \sum_j x_j^2(T) s_2^2(j) \end{bmatrix} \qquad \text{(III.107)}$$

It follows that

$$\hat{\sigma}_2 = (\dot{z}' (2\hat{\Psi}_2)^{-1} \dot{z})^{-1} \dot{z}' (2\hat{\Psi}_2)^{-1} \hat{n}_2 \qquad \text{(III.108)}$$

This yields a fresh estimator $\hat{\hat{\Phi}}_2$ from (III.3) which may be used to obtain a fresh Aitken estimator of $\bar{\beta}$ as

$$\hat{\hat{\beta}}_2 = (X' \, \hat{\hat{\Phi}}_2^{-1} \, X)^{-1} \, X' \, \hat{\hat{\Phi}}_2^{-1} \, Y \qquad \text{(III.109)}$$

The process could be repeated until convergences in $\hat{\hat{\beta}}_j$'s are obtained.

Once again, intuitively it seems that the iterative procedure would lead to gain in efficiency and would reduce the proportion of negative estimates of σ_j^2's. The properties of this iterative estimator are not known and need investigation either analytically or by Monte Carlo experiments. We would follow the latter approach.

III.7   Theil's Weighted Least Squares Estimator -
         Special Case of the Step-Wise Aitken's
         Estimator.

For the bivariate regression model, Theil and

Mennes (1959) have shown that the off-diagonal elements

of the covariance matrix $2\dot{\Psi}$, defined in (III.48), are of

lower order of magnitude than the diagonal elements. Con-

sequently they proposed replacing the off-diagonal elements

of $2\dot{\Psi}$ by zero and then applying weighted least squares

procedure to (III.34).

We shall show that this procedure can be extended

to the multiple linear regression case in a straight for-

ward manner.

Consider the Ψ matrix defined in (III.47), viz.,

$$\Psi = M\Phi M \qquad\qquad\qquad (III.110)$$

where Φ is defined in (III.3) and M in (III.24).
Let us define

$$M^* = X(X'X)^{-1} X' \qquad\qquad (III.111)$$

and

$$M = I - M^* \qquad (\therefore\ M = I - X(X'X)^{-1} X') \qquad (III.112)$$

Substituting the value of M from (III.112) in (III.110) we

have

$$\Psi = (I - M^*)\ \Phi\ (I - M^*) \qquad\qquad (III.113)$$

$$= \Phi - \Phi M^* - M^*\Phi + M^*\Phi M^* .$$

Now Φ is of order 1, or O(1), because its elements do not increase or decrease with sample size and M* is of order 1/T because $(X'X)^{-1}$ is O(1/T). Therefore, the diagonal elements of Ψ are of order 1 and the off-diagonal ones are O(1/T).

Consequently, the covariance matrix $2\dot{\Psi}$ whose elements are 2 times the square of the elements of $\dot{\Psi}$ will have the leading diagonal terms of O(1) and the off-diagonal terms of $O(\frac{1}{T^2})$.

Thus for large samples (at least) one may ignore the off-diagonal elements of the covariance matrix $2\dot{\Psi}$. We may then proceed exactly as in the step-wise Aitken procedure discussed in Section III.6.

III.8   Iterative Theil's Weighted Least Squares
          Estimator - Special Case of Iterative
          Aitken's Step-Wise Least Squares Estimator.

Once again, treating the off-diagonal terms zero in the covariance matrix $2\dot{\Psi}$ we may proceed as in the Iterative Aitken's Step-Wise Least Squares to obtain the Iterative Theil's Weighted Least Squares (ITWLS) Estimator. The convergence properties of this special case are also unknown and need investigation.

III.9   Maximum Likelihood Estimator

So far we have discussed estimators which use step-wise procedure to estimate parameters of the model

(III.1). One may attempt joint estimation of parameters (elements of $\bar{\beta}$ and $\dot{\sigma}$) using the Maximum Likelihood procedure, as proposed by Rubin (1950). Thus, we may assume that the $\eta(t)$'s are normally distributed with means and variances specified in (III.2) - (III.4).

The joint distribution of $\eta(1)$, $\eta(2)$, ..., $\eta(T)$, in (III.1) can be written as

$$L = \frac{1}{(2\pi)^{\frac{T}{2}} \prod_t \phi(t)} \; e^{-\frac{1}{2} \sum_t \frac{\eta^2(t)}{\phi^2(t)}} \qquad \text{(III.114)}$$

where $\phi^2(t)$ is defined in (III.4) above and $\eta(t)$ is the t-th element of the vector η defined in (III.1).

The Log likelihood function of parameters in the CMR model (III.1), given Y and X's can be written as

$$L = -\frac{T}{2} \log 2\pi -\frac{1}{2} \sum_t \log \phi^2(t) -\frac{1}{2} \sum_t \frac{\eta^2(t)}{\phi^2(t)} \qquad \text{(III.115)}$$

where

$$\eta(t) = Y(t) - \sum_j x_j(t) \, \bar{\beta}_j \qquad \text{(III.116)}$$

because, Jacobian, $|J| = \left| \frac{\partial(\eta(1), \ldots, \eta(T))}{\partial(Y(1), \ldots, (T))} \right| = 1$

The normal equations of $\tilde{\tilde{\beta}}_j$'s and $\tilde{\sigma}_j^2$'s for all j, can then be written as

$$\frac{\partial L}{\partial \tilde{\tilde{\beta}}_j} = \sum_t \frac{\tilde{\eta}(t) \; x_j(t)}{\tilde{\phi}^2(t)} = 0 \quad , \quad \text{for } j = 1, 2, \ldots K$$

(III.117)

where

$$\tilde{\phi}^2(t) = \sum_j x_j^2(t) \; \tilde{\sigma}_j^2 \quad , \quad \text{for } t = 1, 2, \ldots, T \quad \text{(III.118)}$$

and

$$\tilde{\eta}(t) = Y(t) - \sum_j x_j(t) \; \tilde{\tilde{\beta}}_j \; . \tag{III.119}$$

We may write (III.117) as

$$(X' \; \tilde{\phi}^{-1} \; X) \; \tilde{\tilde{\beta}} = X' \; \tilde{\phi}^{-1} \; Y \tag{III.120}$$

The normal equation of $\tilde{\sigma}_j^2$ are

$$\frac{\partial L}{\partial \tilde{\sigma}_j^2} = 0 = -\frac{1}{2} \sum_{t=1}^{T} \frac{1}{\tilde{\phi}^2(t)} \cdot x_j^2(t) -$$

(III.121)

$$\frac{1}{2} \sum_{t=1}^{T} \frac{(\tilde{\eta}^2(t) \; (-1) \; (x_j^2(t))}{(\tilde{\phi}^2(t))^2}$$

or

$$\sum_t \frac{x_j^2(t)}{\tilde{\phi}^2(t)} - \sum_t \frac{\tilde{n}^2(t)\ x_j^2(t)}{(\tilde{\phi}^2(t))^2} = 0, \qquad \text{(III.122)}$$

for $j = 1, 2, \ldots, K$.

These normal equations are highly non-linear in the parameters. In principle, solution of these non-linear equations can be obtained by some suitable method of solving non-linear equations. These methods seek to find the roots of a linearised version of the non-linear equations through successive iterations starting with some initial estimates. The selection of the initial values of the parameters is an important aspect of these iterative procedures as convergence depends on them. The initial values could be intelligent guesses or estimates obtained by less efficient methods. It is generally recommended that consistent estimators be used as the initial values because they are found to be good approximations. Further, the final round estimates will be consistent because they are functions of initial consistent estimates.

If the likelihood equations have multiple roots and corresponding local maxima,it is possible that if iterative procedure converges at all, it will converge to a local maxima rather than the global maximum. To ensure a global maximum we may have to find all the roots and choose

the one for which L is maximum. In addition, if the likelihood function has saddle points, or region of inflection oscilations may occur and thus convergence may not be achieved. A variety of numerical methods varying in computational complexity, their capabilities to overcome the problems discussed above and speed for convergence are available to suit the need of researcher.[14] The solution of the non-linear system may thus come to remain as an exercise in Numerical Analysis.[15]

In the present study, we have not attempted to study the sampling properties of this estimator for paucity of computing funds and research time, and this has been left to be tackeld as post doctoral research.

---

[14]A few references are included in the Bibliography.

[15]In a separate study Singh, Nagar and Raj (1972) employed the Modified Gauss-Newton Procedure to solve the non-linear equations, and found that almost half of the times as many as 22 to 36 iterations were required for convergence.

CHAPTER IV

THE DESCRIPTION OF THE SAMPLING EXPERIMENTS

IV.1   Purpose of the Sampling Experiments

We disucssed several methods of estimating the mean and variance in the linear regression model with random coefficients in Chapter III.   Intuitively these estimators may be ranked as follows, because they incorporate successively less information regarding the structure of the covariance matrix $2\overset{*}{\Psi}$ defined in (III.48).[1]

    (i)   SALS

   (ii)   TWLS

 (iii)   HH

  (iv)   OLS

The purpose of the sampling experiments is to determine if the above intuitive ranking holds good in large and small samples.   In the preceding chapter we also derived the estimates of the asymptotic standard errors of the mean and variance of the regression coefficients.[2]   The

---

[1] We have excluded the Maximum Likelihood Estimator (MLE) from the ranking because it is not subjected to sampling experiments in the present study. Further, the iterative procedures have not been included in the ranking because they will be dealt with separately.

[2] The OLS provides estimator of the means only.

exact standard errors for finite samples of the mean
response and the variance of regression coefficients are
not known. The sampling experiments are designed to
determine if the conventionally calculated estimates of
the asymptotic standard errors are biased upward or
downward in small samples.

In Chapter III, we also discussed three iterative
estimators for estimating the parameters of the model,
viz.

    (a) Iterative SALS (ISALS),

    (b) Iterative TWLS (ITWLS), and

    (c) Iterative HH (IHH).

However, it is not clear if these iterative methods will
result in a gain in efficiency in terms of smaller standard
errors of the parameter estimates. Moreover, we do not
know whether the iterative process will really converge.
In fact these questions require a great deal of theoretical
analysis. We shall look into these questions through
sampling experiments.

All the methods have the common difficulty that
the estimate of the variance may sometimes turn out to be
negative. As pointed out in Chapter III, we expect a
method yielding a smaller number of negative estimates of
variance to be more efficient and rank them on this basis.
Further, since we expect the Iterative estimators to be
more efficient than their counterpart, we propose to

undertake sampling experiments to determine if the iterative schemes yield a smaller proportion of negative estimates of variances.

IV.2   Design of the Sampling Experiments

IV.2.1   Specification of the Models

We shall analyse the following three models in the present study.

The Proportional CMR Model I.

$$Y(t) = \beta_1(t) \, x_1(t) \quad , \quad t = 1, \, 2, \, \ldots, \, T, \qquad \text{(IV.1)}$$

and

$$\beta_1(t) = \bar{\beta}_1 + \varepsilon_1(t) \qquad \qquad \text{(IV.2)}$$

$$E\varepsilon_1(t) = 0 \, , \quad \text{for all } t,$$

$$\text{Var } \varepsilon_1(t) = \sigma_1^2 \text{ for all } t$$

we assume that $\varepsilon_1(t)$'s are independently normally distributed.

Combining the two equations (IV.1) and (IV.2), we may write

$$Y(t) = \bar{\beta}_1 \, x_1(t) + \eta_1(t) \, , \quad t = 1, \, 2, \, \ldots, \, T \qquad \text{(IV.3)}$$

where

$$\eta_1(t) = x_1(t) \; \varepsilon_1(t) \;\; , \; t = 1, \, 2, \, \ldots, \, T. \qquad \text{(IV.4)}$$

For the Monte Carlo experimentation we specify a priori the following sets of parametric values.[3]

		Set I	Set II
Mean Response	$= \bar{\beta}_1$:	0.6	10.1
Var $\beta_1(t)$ = Var $\varepsilon_1(t) = \sigma_1^2$:		0.03	0.5

The alternative sample sizes chosen are 10, 20 and 50. The values of x_1 in these experiments are

---

[3]It may be pointed out that conclusions to be drawn from the sampling experiments for each of the three models, depend on the values of the parameters and the values of independent variables selected by us. In view of this, Thornber (1967) suggested that the whole range of parameter space be investigated to make more definite conclusions. Consequently, we may select an estimator among the host of other estimators with smallest loss function. The suggestion however, entails enormous computation cost even when a small range of parameter space is considered in conjunction with variation in other parameters. Besides, there is the difficulty that selection of the various parameters values may not always give meaningful results. In our case larger values of the variances of random parameters generally lead to a set of dependent variables, some of which are negative. We have chosen two sets of values of parameters well apart in the hope that if the selection of an estimator is dependent on parameter space it will show up. We have limited the variation of parameter values to two only, mainly to meet the limitation of computation cost and research time.

$$x_1$$

T = 10	3.452	3.895	3.953	4.157	4.575
	5.555	6.998	7.344	8.027	8.311

T = 20	3.452	3.895	3.953	4.157	4.575
	5.555	6.998	7.344	8.027	8.311
	8.923	9.584	11.079	11.849	12.688
	14.794	16.072	16.904	16.984	18.239

T = 50	3.452	3.895	3.953	4.157	4.575
	5.555	6.998	7.344	8.027	8.311
	8.923	9.584	11.079	11.849	12.688
	14.794	16.072	16.904	16.984	18.239
	3.452	3.895	3.953	4.157	4.575
	5.555	6.998	7.344	8.027	8.311
	8.923	9.584	11.079	11.849	12.688
	14.794	16.072	16.904	16.984	18.239
	3.452	3.895	3.953	4.157	4.575
	5.555	6.998	7.344	8.027	8.311

The Bivariate CMR Model II.

$$Y(t) = \beta_1(t)\ \bar{x}_1 + \beta_2(t)\ x_2(t)\ , \quad t = 1,\ 2,\ \ldots,\ T \qquad (IV.5)$$

where $\bar{x}_1$ is a constant; and

$$\beta_j(t) = \bar{\beta}_j + \varepsilon_j(t)\ , \quad j = 1,\ 2, \qquad (IV.6)$$

$$E\ \varepsilon_j(t) = 0\ , \quad j = 1,\ 2,$$

$$\text{Var}\ \varepsilon_j(t) = \sigma_j^2\ , \quad j = 1,\ 2\ .$$

We assume that the $\varepsilon_1(t)$'s and $\varepsilon_2(t)$'s are independently normally distributed.

Combining (IV.5) and (IV.6), we may write the

model II as

$$Y(t) = \bar{\beta}_1 \bar{x}_1 + \bar{\beta}_2 x_2(t) + \eta_2(t) \ , \ t = 1, \ 2, \ \ldots, \ T \qquad (IV.7)$$

where

$$\eta_2(t) = \bar{x}_1 \varepsilon_1(t) + x_2(t) \varepsilon_2(t) \ , \qquad t = 1, \ 2, \ \ldots, \ T \qquad (IV.8)$$

This gives us a case more like the conventional linear relation where we associate the disturbance or error with the whole relation rather than just the parameters. Indeed the error in equation can be assumed to be an error associated with the constant term or intercept. In the CMR model II we have in effect the composite case of error in parameter and error in equation.

We specify a priori two sets of values of the parameters for the Monte Carlo experiments as

		Set I	Set II
Mean Response	$= \bar{\beta}_1$:	0.8	30.2
Mean Response	$= \bar{\beta}2$:	0.6	10.1
$\text{Var } \beta_1(t) = \text{Var } \varepsilon_1(t) = \sigma_1^2$:		0.04	1.5
$\text{Var } \beta_2(t) = \text{Var } \varepsilon_2(t) = \sigma_2^2$:		0.03	0.5

The alternative samples sizes chosen are the same as in the CMR model I and the values of x_2 in these experiments

are the same as values of x_1 in the CMR model I.

The values of $\bar{x}_1$ (a constant) for sample size 10 are as follows:

$$\bar{x}_1$$

| T = 10 | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 |
| | 10.0 | 10.0 | 10.0 | 10.0 | 10.0 |

The values of explanatory variable $\bar{x}_1$ for sample size 20 and 50 are obtained by repeating the values of size 10.[5]

The Trivariate CMR Model III.

$$Y(t) = \beta_1(t) \, x_1(t) + \beta_2(t) \, x_2(t) \ , \ t = 1, \ 2, \ \ldots, \ T \tag{IV.9}$$

and

$$\beta_j(t) = \bar{\beta}_j + \varepsilon_j(t) \ , \ j = 1, \ 2, \tag{IV.10}$$

$$E \, \varepsilon_j(t) = 0 \ , \ \text{for all } j; \ t = 1, \ 2, \ \ldots, \ T,$$

$$\text{Var } \varepsilon_j(t) = \sigma_j^2 \ , \ j = 1, \ 2.$$

---

[5] In order to avoid any possible rounding of errors in calculation of the type mentioned by Johnston (1963), pp. 131, we shift the origin to make the numbers in x_1 and x_2 of the same size.

We assume that the $\varepsilon_1(t)$'s and $\varepsilon_2(t)$'s are independently normally distributed.

Combining (IV.9) and (IV.10), we may write the CMR model III as

$$Y(t) = \bar{\beta}_1 \, x_1(t) + \bar{\beta}_2 \, x_2(t) + \eta_3(t) \ , \ t = 1, \, 2, \, \ldots, \, T$$

$$(IV.11)$$

where

$$\eta_3(t) = x_1(t) \, \varepsilon_1(t) + x_2(t) \, \varepsilon_2(t) \qquad , \ t = 1, \, 2, \, \ldots, \, T.$$

$$(IV.12)$$

The a priori specification of the parameter values for the sampling experiments is the same as in the CMR model II.

The values of the independent variable x_1 in these experiments for sample sizes 10, 20 and 50 are the same as values of the independent variable x_1 in the Monte Carlo experiments with the CMR model I. The values of x_2 in the sampling experiments in the CMR model III are:

x_2

T = 10	10.100	10.150	10.200	10.250	10.200
	10.100	10.300	10.350	10.400	10.400
T = 20	10.200	10.600	10.700	10.800	10.880
	11.555	11.660	11.886	11.886	11.995
	10.998	10.987	11.567	11.554	11.990
	11.997	12.000	12.332	12.775	12.887

$$x_2$$

T = 50				
10.200	10.600	10.700	10.800	10.880
11.555	11.660	11.886	11.886	11.995
10.998	10.987	11.567	11.554	11.990
11.997	12.000	12.332	12.775	12.887
10.200	10.600	10.700	10.800	10.880
11.555	11.660	11.886	11.886	11.995
10.998	10.987	11.567	11.554	11.990
11.997	12.000	12.332	12.775	12.887
10.200	10.600	10.700	10.800	10.880
11.555	11.660	11.886	11.886	11.995

The correlation matrices indicating the simple correlation between two independent variables for sample sizes 10, 20 and 50 are given as

Sample Size

Correlation Matrix

$$
\begin{array}{ccc}
10 & 20 & 50 \\
\begin{bmatrix} 1 & 0.8569 \\ 0.8569 & 1 \end{bmatrix} &
\begin{bmatrix} 1 & 0.8635 \\ 0.8635 & 1 \end{bmatrix} &
\begin{bmatrix} 1 & 0.8516 \\ 0.8516 & 1 \end{bmatrix}
\end{array}
$$

It is clear from the simple correlation matrices that there is high degree of intercorrelation between two variables. The objective is to study the performance of alternative estimators when regressors are highly correlated.

IV.2.2 Generation of Random Numbers

Having specified a priori the parameter values, observations on the independent variables and the distribution of the $\varepsilon_j(t)$'s we now seek to generate a set of independently normally distributed pseudo-random numbers

corresponding to the disturbances $\eta_j(t)$'s in the three models. The procedure to generate a series of pseudo-random numbers $\eta_j(t)$'s may be explained as follows:

The first step is to generate a series of pseudo-random numbers $u(t)$'s as random drawings from a uniform population. There are many standard routines available to accomplish this task.[6] The shortcoming with the pseudo-random drawings produced by one such computer routine are firstly that they fail the statistical test of the hypothesis that they come from some known (uniform) population; and secondly they turn out to be serially correlated.[7] Carter (1972) used the shuffling technique proposed by MacLaren and Marsaglie (1965) to generate pseudo-random numbers free from these two deficiencies.[8]

---

[6] One such routine used for the present study has been fully described by K. Roberts (1968).

[7] R.A. L. Carter (1972).

[8] Op. cit., Carter (1972):- The shuffling technique was found to have no effect on the statistical test of hypothesis that the pseudo-random numbers come from a uniform population. Therefore, if a set of random drawings did not pass the null hypothesis that the sample is from a uniform population over the range 0 to 1, that set is discarded from the set of random drawings, otherwise it is put to further tests for correlations (up to order five and even odd items). The set is retained if the hypotheses that sample drawings come from a population in which each of the above correlation is zero. If any of correlation is not zero, the set is shuffled and the tests for serial correlations are applied. The shuffled sample is retained only if the above hypothesis is accepted for the shuffled sample. In this way the pseudo-random drawings can be obtained

The second step is to transform the pseudo-random drawings from a uniform population to pseudo-random standard normal deviates. This is accomplished by the Box-Muller (1958) transformation:

$$w_1(t) = (-2 \log u_1(t))^{\frac{1}{2}} \cos (2\pi u_2(t)) \qquad \text{(IV.13)}$$

$$w_2(t) = (-2 \log u_1(t))^{\frac{1}{2}} \sin (2\pi u_2(t))$$

where the $u_1(t)$'s and $u_2(t)$'s, $t = 1, 2, \ldots, T$ are pseudo-random drawings from a uniform distribution obtained in the first step; and the $w_1(t)$'s and $w_2(t)$'s are the independent standard normal deviates.[9]

IV.2.3 Conversion of Standard Normal Deviates to $\eta_j(t)$'s

The final step is to transform the series of independent standard normal deviates to independent normal deviates $\eta(t)$'s for each of the three models selected by us for Monte Carlo study. By way of illustration consider

---

such that the desired assumptions hold 100 per cent within the domain of above tests.

[9] The routine for generating standard normal deviates was generously made available to me by Professor R.A.L. Carter.

the problem of finding a suitable transformation in the context of model III.[10] Consequently,

$$\eta(t) \equiv \eta_3(t) \quad , \quad t = 1, 2, \ldots, T. \qquad \text{(IV.14)}$$

Further,

$$E \, \eta_3(t) = E(x_1(t) \, \varepsilon_1(t) + x_2(t) \, \varepsilon_2(t)) \qquad \text{(IV.15)}$$

$$= 0 \ , \text{ for all } t \text{ (because } E \, \varepsilon_1(t) =$$

$$E \, \varepsilon_2(t) = 0))$$

and

$$\text{Var } \eta_3(t) = x_1^2(t) \, \text{Var } \varepsilon_1(t) + x_2^2(t) \, \text{Var } \varepsilon_2(t)$$

$$\text{(IV.16)}$$

because the $\varepsilon_1(t)$'s and $\varepsilon_2(t)$'s are assumed to be indepently normally distributed.

Therefore, we may write (IV.16) as

$$\text{Var } \eta_3(t) = x_1^2(t) \, \sigma_1^2 + x_2^2(t) \, \sigma_2^2 \ . \qquad \text{(IV.17)}$$

---

[10]It can be easily varified that Model I and Model II are particular cases of Model III when $\sigma_2^2 = 0$ (Model I); $x_1(t) = \bar{x}_1$ for all t, while $x_2(t)$ becomes $x_1(t)$ (Model II).

Let the transformation be given as

$$\eta_3(t) = \phi_3(t) \, w(t) \quad , \quad t = 1, 2, \ldots, T \qquad \text{(IV.18)}$$

where $w(t)$'s are the Standard Normal Variates obtained in (IV.13).

Then, it follows that

$$E \, \eta_3(t) = E \, (\phi_3(t) \, w(t)) \qquad \text{(IV.19)}$$

$$= \phi_3(t) \, E \, w(t)$$

$$= 0 \quad , \quad \text{for all } t \quad \text{because}$$

the $\phi_3(t)$'s are known and fixed in repeated samples, and the $w(t)$'s are standard normal variates, therefore

$$\text{Var } \eta_3(t) = \phi_3^2(t) \, \text{Var } w(t) \qquad \text{(IV.20)}$$

$$= \phi_3^2(t)$$

$$= x_1^2(t) \, \sigma_1^2 + x_2^2(t) \, \sigma_2^2 \quad .$$

Further, the $\eta_3(t)$'s will be independently normally distributed because they are linear transformation of independently normally distributed variates $w(t)$'s.

Thus the transformation (IV.18) is the desired transformation to transform the pseudo independent standard normal variates, obtained in the second step of the trans-

formation, to independent normal variates with mean 0 and variance $\phi_3^2(t)$.[11]

In the matrix notation (IV.18) it may be written as

$$\eta_3 = P_3 w_3 \tag{IV.21}$$

where

$$\eta_{3_{Tx1}} = \begin{bmatrix} \eta_3(1) \\ \cdot \\ \cdot \\ \cdot \\ \eta_3(T) \end{bmatrix} , \quad P_{3_{TxT}} = \begin{bmatrix} \phi_3(1) & & & \\ & \cdot & & 0 \\ & & \cdot & \\ 0 & & \cdot & \\ & & & \phi_3(T) \end{bmatrix}$$

and

$$w_{3_{Tx1}} = \begin{bmatrix} w(1) \\ \cdot \\ \cdot \\ \cdot \\ w(T) \end{bmatrix}$$

The derivation of the transformation for disturbances in the general linear model may be obtained analogously as

$$\eta = Pw \tag{IV.22}$$

___

[11]See footnote 10 regarding the derivation of transformation for Model I and Model II.

where

$$\eta_{Tx1} = \begin{bmatrix} \eta(1) \\ \cdot \\ \cdot \\ \cdot \\ \eta(T) \end{bmatrix} \quad , \quad W_{Tx1} = \begin{bmatrix} w(1) \\ \cdot \\ \cdot \\ \cdot \\ w(T) \end{bmatrix} \qquad \text{(IV.23)}$$

and

$$P_{TxT} = \begin{bmatrix} \phi(1) & & & \\ & \cdot & & 0 \\ & & \cdot & \\ 0 & & \cdot & \\ & & & \phi(T) \end{bmatrix} \qquad \text{(IV.24)}$$

such that

$$\phi(t) = \sqrt{\phi^2(t)} = \sqrt{\sum_{j=1}^{K} \sigma_j^2 \, x_j^2(t)} \quad \text{for } t = 1, 2, \ldots, T.$$

Thus, the peseudo-random numbers $\eta(t)$'s for given values of σ_j^2's and x's may be obtained by pre-multiplying the vector w of the Standard Normal Variates by the transformation matrix P defined in (IV.24).

IV.2.4 Descriptive Statistics

The description of an empirical distribution is commonly studied in terms of first two moments, mean and

the variance (or the standard deviation), defined as[12]

$$\text{Mean } (\hat{\theta}) = \frac{1}{N} \sum_{i=1}^{N} \theta_i = \bar{\hat{\theta}} \qquad (IV.25)$$

where θ_i is an estimate of θ.[13]

The first moment of empirical distribution is used to estimate the bias of the parameter, defined as

$$\text{Estimated Bias } (\hat{\theta}) = \bar{\hat{\theta}} - \theta \qquad (IV.26)$$

where θ is the true value of the parameter.[14]

The standard deviation of the sampling distribution (a measure of dispersion) is the square root of the second moment about the sample mean and is defined as

---

[12]This is under the assumption that small sample moments in population are defined. As pointed out by Basmann (1961) however, these moments may not always exist. The problem of verifying the existence of small samples moments in the population is a rather complicated one. In the absence of this verification it is instructive to compute the corresponding non-parametric statistics for the sampling distribution (the existence of the population counterpart of these is without doubt), which has been done in the present study. These non-parametric statistics are defined later in this section.

[13]The sample size N throughout the sampling experiments is equal to one hundred unless otherwise noted.

[14]The true value of parameter is fixed (a priori) in sampling experiments.

$$S\ d(\hat{\theta}) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{\theta}_i - \bar{\hat{\theta}})^2} \qquad\qquad\text{(IV.27)}$$

The standard deviation of different sampling distribution are computed to compare the relative efficiency of alternative estimators.

Sometimes a percentage relative efficiency concept based on the second moment about the sample mean of sampling distribution is used to compare the relative efficiency of two alternative estimators.

$$\text{Percentage Efficiency (%E)} = \frac{\text{Var}(\hat{\theta})_I}{\text{Var}(\hat{\theta})_{II}} \text{ x } 100.0 \qquad\qquad\text{(IV.28)}$$

$$= \frac{[Sd(\hat{\theta})_I]^2}{[Sd(\hat{\theta})_{II}]^2} \text{ x } 100.0$$

where the subscript I and II indicate the estimator under consideration.

The estimator II is said to be relatively more efficient than estimator I if E > 100. The descriptic statistics (IV.27) and (IV.28) are particularly useful to compare two unbiased estimators. If the estimators are biased a more useful descriptic statistic is the root mean square error based on the second moment of the empirical distribution about the true value of para-

meters.[15]

$$\text{Root Mean Square Error } (RMSE(\hat{\theta})) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{\theta}_i - \theta)^2} \qquad (IV.29)$$

or

$$RMSE(\hat{\theta}) = \sqrt{Var(\hat{\theta}) + Bias^2(\hat{\theta})} \qquad (IV.30)$$

The descriptive statistics defined above have the common shortcoming that their values are dependent on the size of true parameter value. The scale free descriptive statistics which are analogous to the above concepts are obtained by dividing each statistic by the true value of the parameter and representing the ratio in percentage form.

$$\text{Percentage Bias } (\% \text{ Bias}) = \frac{Bias(\hat{\theta})}{\theta} \times 100.0 \qquad (IV.31)$$

$$\text{Percentage Standard Deviation } (\% Sd) = \frac{Sd(\hat{\theta})}{\theta} \times 100.0 \qquad (IV.32)$$

The descriptic statistic (IV.32) is analogous to the coefficient of variation.

---

[15]The empirical bias of unbiased estimator will generally be zero, yet there may be instances when this is not so. In situations where empirical bias is not zero descriptive statistic (IV.29) rather than (IV.27) is useful.

Finally,

Percentage Root Mean Square Error (% RMSE) $= \frac{RMSE(\hat{\theta})}{\theta}$ x 100.0

$$(IV.33)$$

In addition to computing the first two moments of the sampling distribution we may compute a number of non-parametric statistics.[16]

As a measure of central tendency, we may compute the Median to obtain,

Median Bias (M Bias$(\hat{\theta})$) = Median$(\hat{\theta})$ - θ (IV.34)

and

Percentage Median Bias (% M Bias) $= \frac{M\ Bias(\hat{\theta})}{\theta}$ x 100.0 (IV.35)

The non-parametric measure of dispersion computed is the quartile deviation, defined as

Quartile Deviation (QD$(\hat{\theta})$) $= \frac{Q_3 - Q_1}{2}$ (IV.36)

where Q_1 and Q_3 are the first and third quartiles for the data.[17]

We may define scale free dispersion concept for (IV.36) which is analogous to the percentage standard deviation.

Percentage Quartile Deviation (%QD) $= \frac{Q_3 - Q_1}{2\theta}$ (IV.37)

---

[16]See footnote 12.

[17]Let a set of data be arranged in order of magnitude and divided into four equal parts. Further, let Q_1 Q_2 and Q_3 be the values which divides the arranged set, then, Q_1 and Q_3 are the first and third quartiles respectively while Q_2 is th median. The probability of an observa-

IV.2.5   Data for the Sampling Experiments

The values of independent variables for each of the three CMR models and sample sizes 10, 20 and 50 have been specified in Section IV.2.1. These values of the independent variables for any given model and sample size T remain constant in repeated samples as required by the assumption of the CMR models. We now consider generation of sample of data on the dependent variable for a given model and sample size T.

Suppose, we are interested in generating the sample of data on the CMR model III for sample size T = 10; then

$$Y = X\bar{\beta} + \eta \qquad\qquad \text{(IV.37)}$$

where

$$Y_{10x1} = \begin{bmatrix} Y(1) \\ \cdot \\ \cdot \\ \cdot \\ Y(10) \end{bmatrix}, \; X_{10x2} = \begin{bmatrix} x_1(1) & x_2(1) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ x_1(10) & x_2(10) \end{bmatrix}, \; \bar{\beta}_{2x1} = \begin{bmatrix} \bar{\beta}_1 \\ \bar{\beta}_2 \end{bmatrix} \qquad \text{(IV.38)}$$

and

$$\eta_{10x1} = \begin{bmatrix} \eta(1) \\ \cdot \\ \cdot \\ \cdot \\ \eta(10) \end{bmatrix}$$

tion falling within the interval Med $\pm$ QD is .50.

such that

$$E\eta = 0 \qquad\qquad (IV.39)$$

and

$$E\eta\eta' = Var~\eta = \begin{bmatrix} (x_1^2(1)~\sigma_1^2 + x_2^2(1)~\sigma_2^2 & & 0 \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ 0 & (x_1^2(10)~\sigma_1^2 + x_2^2(10)~\sigma_2^2 \end{bmatrix} \qquad (IV.40)$$

The values of $\bar{\beta}_j$ and σ_j^2 were also specified for the CMR Model III in Section IV.2.1. Therefore, the values in $X\bar{\beta}$ can be obtained directly while the values of η can be obtained analogous to transformation (IV.18). Hence, a sample of data on the dependent variable Y are obtained from (IV.37). The process of obtaining sample of data on Y's is repeated 100 times. The set of observations in $X\bar{\beta}$ remains constant in each of 100 replications in the sampling experiment. The set of observations in Y differ from replication to replication because the set of pseudo-random numbers in w for the transformation (IV.18) change with replication. In a similar manner, the sample of data may be obtained for other sample sizes 20 and 50 in the CMR Model III, and for the other two models and samples sizes 10, 20 and 50.

Having obtained 100 samples of data for each of the three models and sample sizes, the estimates of the $\bar{\beta}_j$'s and σ_j^2's, and their standard errors are obtained for alternative estimators. For each of these estimators the frequency distributions of individual estimates are assembled and then, the descriptive statistics of their frequency distributions are calculated.

CHAPTER V

RESULTS OF SAMPLING EXPERIMENTS[1]

In the preceding chapter we described the design of the sampling experiments. In this chapter we present the results of the sampling experiments. For discussion purposes we have divided this chapter into two sections. In the first section, we shall discuss the sampling properties of the Ordinary Least Squares (OLS), the Hildreth and Houck (HH), the Theil's Weighted Least Squares (TWLS) and the Step-wise Aitken's Least Squares (SALS) estimators for sample sizes 10, 20 and 50 (only for the CMR Model III-Set I).[2] In the second section, we shall present the results of the sampling experiments on the Iterative HH (IHH), the Iterative TWLS (ITWLS) and the Iterative SALS (ISALS) estimators for sample size

---

[1]The computer programming for computations in this chapter was done by the author. The Department of Economics, University of Western Ontario provided the financial assistance for the computations.

[2]Originally, we planned to undertake the sampling experiments for sample size 50 for all the models. However, after finding out the cost of computations for alternative estimators in the CMR Model III-Set I we decided to drop the experiments for other models for reasons similar to given in footnote 3.

20 only.[3]

V.1   Sampling Properties of the OLS, the HH, the
               TWLS and the SALS Estimators.

V.1.1   Estimation of the Mean Response Coefficients.

V.1.1.1   Bias Analyses of the Mean Response Coefficients
               of Alternative Estimators.

In Chapter III we proved that the OLS and the HH

Estimators of the mean response coefficients are unbiased.

We also noted that the proof of unbiasedness for the HH

estimator can be extended for the TWLS and the SALS estimators

of the mean response coefficients. Now, we will investigate

if the empirical bias is zero. The results of the sampling

experiments relating to the bias of mean response coefficients

are summarized in Table 1 to Table 3. The summary statistics

used are the percentage bias (% Bias) and the percentage

median bias (% M Bias).[4] The hypothesis that each of the

_____

[3]Ideally sampling experiments for the iterative estimators
   should also be done for sample sizes 10 and 50. However,
   iterative properties are rather costly to investigate. In
   view of the limitation of funds, we decided to study the
   properties of the estimators for sample size 20 only. For
   the same reason we had to limit our investigation to six
   iterations only. This limitation was imposed after seeing
   the results of sampling experiments for sample size 20 and
   other sample runs. Further investigation will follow as
   post doctoral research.

[4]The % Bias, the % M Bias are defined respectively in
   (IV.31) and (IV.35).

estimators for the mean response coefficients is unbiased

is tested by the t-test.[5]

Bias Analysis in the Proportional CMR Model I.

Table 1 shows that the empirical % Bias, the % M

Bias and the t-statistic values for the two sets of values

of the parameters in the CMR Model I for sample sizes 10

and 20. It may be noticed that for each sample size the

numbers within each descriptive statistic in this table for

the HH, the TWLS and the SALS estimators are the same. This

is to be expected theoretically because of the special nature

of the CMR Model I. It is shown in Chapter VI (see footnote

TABLE 1 Percentage Bias, Percentage Median Bias and t-statistic of Alternative Estimators of the Mean Response Coefficient in the CMR Model I

Sample Size	Estimator	MODEL I — SET I			MODEL I — SET II		
		Coefficient: $\bar{\beta}_1 = 0.6$			Coefficient: $\bar{\beta}_1 = 10.1$		
		% Bias	% M Bias	t	% Bias	% M Bias	t
10	SALS	0.2094	0.2952	0.2266	0.0508	0.0716	0.2266
	TWLS	0.2094	0.2952	0.2266	0.0508	0.0716	0.2266
	HH	0.2094	0.2952	0.2266	0.0508	0.0716	0.2266
	OLS	1.3166	0.6342	1.1605	0.3193	0.1538	1.1605
20	SALS	0.0136	0.0489	−0.0210	0.0033	0.0119	−0.0210
	TWLS	0.0136	0.0489	−0.0210	0.0033	0.0119	−0.0210
	HH	0.0136	0.0489	−0.0210	0.0033	0.0119	−0.0210
	OLS	0.1047	0.3358	0.4115	0.0254	0.0814	0.1223

[5]The t-test is defined in footnote 6.

5, Ch. VI) that in the proportional CMR model I the HH, the TWLS, and the SALS estimators are identical to the Aitken estimator of the mean response coefficient.

It is clear from table 1 that the % Bias and the % M Bias figures for sample sizes 10 and 20 in each of the two sets of values of parameters are less than one and a half percent for all the four estimators of the mean response coefficient suggesting that the estimators are likely unbiased. More importantly, in each case the t-statistic value lies inside the interval $-t_{.975}$ to $t_{.975}$, which for large samples is the interval -1.96 to 1.96 at .05 level of significance.[6] Thus empirically also we find the estimators to be unbiased in the CMR model I. Note however, the results indicate that the OLS has greater biases than the other methods. Thus if we

---

[6]The empirical bias is obtained from a sample of size 100. We can therefore, apply the large sample theory here. The hypothesis is

$$H_o: \quad \hat{\bar{\theta}} = \theta$$

where $\hat{\theta}_i$ is the estimate of θ obtained from one sample of data by a particular method and $\hat{\bar{\theta}} = \frac{1}{N} \sum_{i=1}^{N} \hat{\theta}_i$, $N = 100$.

The Alternative hypothesis is

$$H_1: \quad \hat{\bar{\theta}} \neq \theta$$

The t-statistic is defined as:

$$t = \frac{\hat{\bar{\theta}} - \theta}{Sd\hat{\theta}|\sqrt{N}}$$

increase the significance level (SL) the OLS becomes biased
before the other methods do.[7]

Bias Analyses in the Bivariate CMR Model II with an Intercept

Table 2 gives the % Bias, the % M Bias, and the
t-statistic for each of the four estimators of the mean
response in the CMR Model II for sample sizes 10 and 20.

It may be seen that for sample size 10 the % Bias and
the % M Bias figures are always more than three percent. On
the other hand for sample size 20 the % Bias and the % M Bias
figures are always less than two percent except in the case of
the TWLS estimator of the mean response coefficient $\bar{\beta}_2$ in the
set I. However, the % Bias figures are only a rough guideline
on the unbiasedness of an estimator and not a test of hypothesis.
In order to test the hypothesis that the estimators are un-
biased we must verify if the t-statistic values lie inside the
interval $-t_{.975}$ to $t_{.975}$ at .05 level of significance (see

---

[7]The Median Bias is tested using the binomial test.
The null hypothesis is

H_0: Median $(\hat{\theta}_i) - \theta = 0$

An equivalent method of testing this hypothesis is

Prob $(\theta_i \leq$ Median $(\hat{\theta}_i)) =$ Prob $(\theta_i >$ Median $(\hat{\theta}_i)) = \frac{1}{2}$

For large samples the standard normal variate

$(|L-\frac{N}{2}| - \frac{1}{2}) \frac{1}{2} \sqrt{N}$ is calculated, where L is the number of
estimates to the left of Median $(\hat{\theta})$. The binomial test is
described in Kendall and Stuart (1961) pp. 514-517. For our
study we did not use this test.

TABLE 2 Percentage Bias, Percentage Median Bias and t-statistic of Alternative Estimators of the Mean Response Coefficients in the CMR Model II

Sample Size	Estimator	MODEL II — SET I						MODEL II — SET II					
		Coefficient: $\bar{\beta}_1 = .8$			Coefficient: $\bar{\beta}_2 = .6$			Coefficient: $\bar{\beta}_1 = 30.2$			Coefficient: $\bar{\beta}_2 = 10.1$		
		% Bias	% M Bias	t	% Bias	% M Bias	t	% Bias	% M Bias	t	% Bias	% M Bias	t
10	SALS	3.0882	6.2073	0.5295	7.4748	19.2282	-0.5286	0.5718	0.8446	-1.2340	3.1662	5.7876	1.2506
	TWLS	3.1324	3.2233	-1.0504	7.8510	17.7230	1.0639	0.6815	0.9214	-1.4386	3.7899	5.3460	1.4605
	HH	4.7060	7.7696	-0.9663	11.8492	20.0371	0.9724	0.8356	1.2112	-0.8724	4.7147	6.1296	0.9007
	OLS	5.3121	8.5167	-2.0330	13.3287	16.4630	2.0605	0.7997	0.9840	-1.9916	4.4596	5.0769	2.0422
20	SALS	0.0286	0.6380	-0.0185	0.0044	0.0044	-0.0019	0.0449	0.1377	0.1612	0.3989	0.3989	-0.1855
	TWLS	1.4591	0.6168	0.8022	2.3500	2.3500	-0.8572	0.0300	0.1427	-0.1327	0.0949	0.4911	0.1347
	HH	1.1865	0.6859	-0.6544	1.9580	1.9580	0.7017	0.2336	0.0953	0.1472	0.1107	0.5188	-0.1492
	OLS	0.2530	0.3511	-0.1692	0.3696	0.3696	0.1652	0.0317	0.0510	-0.1463	0.0995	0.5698	0.1476

footnote 6 of this chapter).

It may be seen from Table 2 that in all cases except
the OLS estimator for sample size 10, the t-statistic value
does lie inside the interval -1.96 to 1.96. Hence, all
estimators except the OLS estimator for sample size 10 are
unbiased in the CMR Model II at .05 level of significance.

Bias Analyses in the Tri-variate CMR Model III without
an Intercept.

The results in table 3 for sample size 10 and 20
are similar to the CMR Model II. Once again the hypothesis
of unbiasedness is substantiated on the t-test for all esti-
mators except the OLS estimator for sample size 10.

In Model III - Set I, the sampling experiments were
extended to sample size 50. The results for sample size 50
are in accordance with sample size 20. The t-statistic values
lie inside the interval $-t_{.975}$ to $t_{.975}$ at .05 level of
significance. Hence each estimator of mean response is
unbiased at .05 level of significance.

V.1.1.2   Efficiency Ranking of Alternative Estimators
               of the Mean Response Coefficients.

In Chapter IV we provided an intuitive efficiency
ranking of the SALS, the TWLS, the HH and the OLS estimators
on the basis of the information regarding the structure of
estimated disturbances used by each estimator. In this

TABLE 3 Percentage Bias, Percentage Median Bias and t-statistic of Alternative Estimators of the Mean Response Coefficients in the CMR Model III

Sample Size	Estimator	MODEL III — SET I						MODEL III — SET II					
		Coefficient: $\bar{\beta}_1 = 0.8$			Coefficient: $\bar{\beta}_2 = 0.6$			Coefficient: $\bar{\beta}_1 = 30.2$			Coefficient: $\bar{\beta}_2 = 10.1$		
		% Bias	% M Bias	t	% Bias	% M Bias	t	% Bias	% M Bias	t	% Bias	% M Bias	t
10	SALS	6.7176	6.4115	-1.8402	17.1893	21.7080	1.8106	1.0986	2.2744	-1.0969	2.5905	4.2457	1.1301
	TWLS	3.4557	4.0883	-1.1420	8.8736	19.2917	1.1572	0.7061	0.9516	-1.5575	4.0290	5.4976	1.5849
	HH	4.8839	8.3343	-1.2519	12.5937	20.9394	1.2598	2.6359	1.2793	-1.6687	15.0196	7.1554	1.6511
	OLS	5.4587	7.9950	-2.0349	14.0269	16.5260	2.0626	0.8250	1.1752	-1.9968	4.7160	5.3525	2.0456
20	SALS	0.1747	0.0311	-0.1066	0.2824	0.3635	0.1026	0.2435	0.0434	-0.7145	0.9650	0.1551	0.7589
	TWLS	0.6979	0.0581	0.3961	1.3185	1.5682	-0.4353	0.0762	0.0453	-0.3228	0.2924	0.5001	0.3428
	HH	1.3775	0.5923	-0.7218	2.3777	0.3933	0.6957	0.0058	0.1324	-0.0235	0.0032	0.5936	-0.0035
	OLS	0.4620	0.4377	-0.2911	0.6925	0.6107	0.2575	0.0659	0.1334	-0.2825	0.2144	0.3211	0.2554
50	SALS	1.6505	2.7459	-1.6274	2.6496	2.1092	1.3490						
	TWLS	1.6511	2.7706	-1.6279	2.6512	2.3716	1.3494		Not Available*				
	HH	1.6011	2.3240	-1.5610	2.5564	2.3250	1.2749						
	OLS	1.8106	1.5146	-1.7518	2.9600	3.8589	1.4935						

*See footnote 2 of this chapter.

section, we shall investigate if this intuitive efficiency ranking holds in for sample sizes 10, 20 and 50 (only for the CMR Model III - Set I) in the monte carlo experiments. The descriptive statistics computed for the efficiency analysis in the sampling experiments are the percentage root mean square error (% RMSE), the percentage quartile deviation (% QD) and the percentage efficiency (% E).[8] The results of the sampling experiments on the three models are given in Tables 4 to 8.

Efficiency ranking in the Proportional CMR Model I.

For a given sample size the numbers within each descriptive statistics in Table 4 for the HH, the TWLS, the SALS estimators are identical. In fact theoretically we expect this to happen. As noted earlier (see discussion on Table 1), the HH, the TWLS and the SALS estimators in the proportional CMR model I are identical to the Aitken estimator of the mean response. Further, the Aitken estimator of the mean response is independent of the variance of the random coefficient σ_1^2 or its estimator $\hat{\sigma}_1^2$ (see footnote 5, Chapter VI). Thus, in view of the special nature of the CMR model I,

---

[8]The % RMSE, the % QD and % E are defined in (IV.33), (IV.37) and (IV.28) respectively. Unless otherwise noted throughout the section V.1.1, the % E statistics has been calculated as

$$\frac{Var(\hat{\Theta})I}{Var(\hat{\Theta})OLS} \times 100.00, \quad \text{where } I = SALS, TWLS, \text{ or } HH \text{ estimator.}$$

we expect that for a given sample size the HH, the TWLS and the SALS estimators of the mean response coefficient to be identical and each to be more efficient than the OLS estimator.

Table 4   Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage Efficiency of Alternative Estimators of the Mean Response Coefficient in the CMR Model I

Sample Size	Estimator	MODEL I — SET I			MODEL I — SET II		
		Coefficient: $\bar{\beta}_1 = 0.6$			Coefficient: $\bar{\beta}_1 = 10.1$		
		% RMSE	% QD	% E	% RMSE	% QD	% E
10	SALS	9.2417	4.8692	150.7882	2.2413	1.1809	150.7882
	TWLS	9.2417	4.8692	150.7882	2.2413	1.1809	150.7882
	HH	9.2417	4.8692	150.7882	2.2413	1.1809	150.7882
	OLS	11.4216	6.9521	100.0000	2.7700	1.6861	100.0000
20	SALS	6.4639	4.7152	175.4184	1.5676	1.1435	175.4184
	TWLS	6.4639	4.7152	175.4184	1.5676	1.1435	175.4184
	HH	6.4639	4.7152	175.4184	1.5676	1.1435	175.4184
	OLS	8.5617	6.0464	100.0000	2.0764	1.4664	100.0000

It can easily be verified from Table 4 that for a given sample size the SALS, the TWLS and the HH estimators rank equal in efficiency while each rank higher than the OLS estimator on both the % RMSE and the % QD criteria. Since the estimators under consideration have been proved unbiased we can say, using the % E criterion, that for sample size 10 and 20 respectively each of the SALS, the TWLS and the HH estimators is 50 and 75 percent more efficient than the OLS estimator in both sets of values of parameters.

Efficiency Ranking in the Bivariate CMR Model II with an Intercept.

The sampling results on the bivariate CMR Model II with an Intercept are given in Table 5. These results are particularly interesting in different ways, which we discuss below.

It may be noticed that none of the SALS, the TWLS and the HH estimators is more efficient than the OLS estimator on the % RMSE and the % E criteria. On the contrary we find that the OLS estimator ranks higher in estimated efficiency among the four estimators under study. This result may appear surprising. However, reader is reminded of the fact that each of the SALS, the TWLS and the HH estimators of the mean response coefficients is obtained by replacing the covariance matrix Φ in the Aitken formula (III.52) by its consistent estimator because Φ is not known a priori. The efficiency comparison of an 'estimated' Aitken estimator, say HH estimator defined in (III.54) to the OLS estimator defined in (III.11) amounts to comparing the asymptotic covariance matrices of the estimators given in (III.20) (with Φ replaced by its consistent estimator) and (III.56). Theoretically it is difficult to prove that the covariance matrix in (III.56) will be equal to the covariance matrix in (III.20) (with Φ replaced by its consistent estimator) plus a positive definite matrix or vice versa. Rao (1970) has pointed out the difficulty in the theoretical analysis and suggested that Monte Carlo experimentation be done to determine the relative efficiency

Table 5   Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage
Efficiency of Alternative Estimators of the Mean Response Coefficient in the CMR
Model II

| Sample Size | Estimator | MODEL II – SET I | | | | | | MODEL II – SET II | | | | | |
| | | Coefficient: $\bar{\beta}_1 = 0.8$ | | | Coefficient: $\bar{\beta}_2 = 0.6$ | | | Coefficient: $\bar{\beta}_1 = 30.2$ | | | Coefficient: $\bar{\beta}_2 = 10.1$ | | |
		% RMSE	% QD	% E	% RMSE	% QD	% E	% RMSE	% QD	% E	% RMSE	% QD	% E
10	SALS	58.4027	18.2459	20.0725	141.5963	43.3969	20.9293	4.6685	2.9108	75.0976	25.5152	13.6294	74.3894
	TWLS	29.9849	17.9464	76.7739	74.2083	43.4241	76.8476	4.7864	2.7820	71.8298	26.2257	13.5072	70.8075
	HH	48.9307	17.8493	28.7824	122.4257	40.6608	28.1831	9.6144	2.8892	17.5737	13.4204	13.4204	17.4016
	OLS	26.6637	17.3903	100.0000	66.0468	40.6150	100.0000	4.0941	2.7364	100.0000	22.2873	13.3831	100.0000
20	SALS	15.4365	10.9365	93.8210	22.8499	16.5213	95.9173	2.5112	1.5447	56.5854	7.8209	5.2280	74.2170
	TWLS	18.2484	10.9064	67.5723	27.5144	16.8950	66.6095	2.2616	1.6014	91.8423	7.0477	5.2455	91.3803
	HH	18.1422	11.4560	68.2213	27.9744	16.8337	64.2824	2.3371	1.6681	86.0079	7.4232	5.3423	83.3725
	OLS	14.9547	10.8340	100.0000	22.3767	17.9968	100.0000	2.1674	1.5257	100.0000	6.7372	5.3622	100.0000

of the two estimators.

It may also be pointed out that the bivariate CMR Model II has the problem that sometimes negative estimates of the variances are obtained (refer to the discussion in Section III.3.1).[9] The fact that the OLS estimator in our sampling experimentation has turned out to be more efficient than the other three estimators may well be due to the problem of negative estimates of the variances and may not hold in general for all heteroskedastic models. However, the study of all heteroskedastic models is beyond the scope of this thesis.

In Chapter IV we noted that the results of the sampling experiments depend on the values of the parameter chosen by the researcher (see footnote 3, Chapter IV). As pointed out by Thornber (1967) the ranking of alternative estimators may or may not remain the same from one point to another point in the parameter space. For example, the ranking of alternative estimators remained the same for two points in the parameter space in the proportional CMR Model I. However, we will show that the ranking of some of the estimators reverses itself from one point to the other point in the parameter space in the CMR Model II.[10] We may call

---

[9]A theoretical solution to the problem of negative estimates of variances is proposed in Chapter VI.

[10]The sampling experiments, as noted in footnote 3, Chapter IV were restricted to two points in the parameter space for limitation of computer and research time. The two sets of

this phenomenon 'ranking reversal'.

The CMR Model II has two regression coefficients each of which may have different rankings for alternative estimators. In order to determine the overall ranking of alternative estimators, we may calculate the overall % RMSE and overall % QD ranking of alternative estimators and rank them accordingly. The Overall % RMSE and Overall % QD Rankings of alternative estimators for the CMR Model II are given in table 6.[11]

It can be seen from table 6 that on the overall % RMSE the OLS performs better than all other estimators both

---

values were chosen well apart to detect possible ranking reversal. To determine the point in parameter space at which the reversal takes place and to determine if there is more than one reversals would require a study of whole parameter space, which may not be always possible.

[11] In order to obtain the overall % RMSE and the overall % QD rankings in table 6, we first ranked alternative estimators for each of the two regression coefficients separately on the basis of the values of % RMSE and % QD in table 5 - an estimator for each regression coefficient with lowest % RMSE (or % QD) is ranked first while the estimator with highest % RMSE (or % QD) is ranked last. Having ranked alternative estimators for % RMSE and % QD for individual coefficients the overall % RMSE and overall % QD ranking for alternative estimators in table 6 were calculated respectively as the sum of the % RMSE and % QD for the individual regression coefficients as obtained from table 5.

The choice criterion among alternative estimators is that an estimator with lower overall % RMSE and overall % QD rank is always preferred.

Sometimes Kendall's coefficient of concordance (1956) is used for measuring the strength of rankings. However in our case this was not considered necessary in view of clear ranking of estimators. Besides, forecasting precision criterion has been applied later, which does similar job in some sense.

in sample size 10 and 20 since it has the lowest rank. The ranking of other estimators relative to each other changes both across the points in the parameter space and across the sample sizes. Thus, we do see the phenomenon of ranking reversal. In Set I the TWLS estimator is the second best for sample size 10 while the SALS estimator is the second best for sample size 20 on the overall % RMSE ranking criterion. On the same criterion in Set II, on the other hand, the SALS estimator is the second best for sample size 10 while the TWLS estimator is the second best for sample size 20.

Table 6 Overall Percentage Root Mean Square Error and Overall Percentage Quartile Deviation Ranking of Alternative Estimators of the Mean Response Coefficients in the CMR Model II

Sample Size	Estimator	MODEL II – SET I		MODEL II – SET II	
		Overall % RMSE Rank	Overall % QD Rank	Overall % RMSE Rank	Overall % QD Rank
10	SALS	8	7	4	8
	TWLS	4	7	6	5
	HH	6	4	8	5
	OLS	2	2	2	2
20	SALS	4	·4	8	3
	TWLS	7	5	4	5
	HH	7	6	6	7
	OLS	2	.5	2	5

The reader may be reminded that the estimates of variances of random coefficients are used to estimate the mean response coefficients for all estimators except the OLS estimator. It seems that negative estimates of variances which do occur in a fairly large number of times is responsible for the violent behaviour of the estimators. In the Model I since there is no problem of negative estimates there is a definite pattern about the performance of estimators

both across the sample size and the points in parameter space.

On the overall % QD ranking criterion for sample size 10 the OLS estimators ranks first, while for sample size 20 the SALS estimators ranks first in Set I and Set II.

For sample size 10 the OLS estimator is empirically biased and hence the relative efficiency comparison of alternative estimators in terms of the most efficient estimator (the OLS estimator) is not really meaningful. However, for sample size 20 we can make a relative efficiency comparison of alternative estimators on the basis of the % E criterion. Thus in the Set I, the SALS estimator is only 4 to 6 percent less efficient than the OLS estimator (Table 5). The TWLS estimator is 32 to 33 percent less efficient than the OLS estimator. In the Set II the SALS estimator is 26 to 43 percent less efficient, the TWLS estimator 8 to 9 percent less efficient and the HH estimator is 24 to 27 percent less efficient than the OLS estimator.

Before we go on with the discussion of efficiency in CMR Model III we may say a few words in regards to choice between two estimators ranked first and second. The OLS estimator appears better choice for both sample sizes 10 and 20 not only on the overall % RMSE ranking criterion but also on the cost of computation criterion.[12] However, the OLS

---

[12]The cost of computation may be, however, not a very important consideration in single equation computation. The cost of computing a single run in a sample of size 20 for the SALS or the TWLS is less than $2.00 (including the cost of compilation).

does not give estimates of variances and variances of mean response coefficients. We need to use some other estimator to estimate variances.[13] Furthermore, the properties of the standard error (in terms of bias) of the OLS estimator need to be taken into consideration in contrast to the SALS and the TWLS estimators (the second ranked estimators). Also, if the overall % QD is the relevant criterion the choice of the OLS estimator remains as only the second best at least for sample size 20 in the CMR Model II.

Efficiency Ranking in the Trivariate CMR Model III.

The Monte Carlo experiments on the CMR model III for sample size 10, 20 and 50 (for Set I only) are summarized in Table 7. The results are (for sample sizes 10 and 20) generally analogous to Table 5 and many of the comments and conclusions on the CMR model II hold much the same qualitatively. For example, the OLS once again ranks above the other three estimators on the % RMSE and the % E criteria. The explanations for this are the same as set out in discussion on the sampling experiments on Model II. Further, once again we find that the phenomenon of ranking reversal takes place on all criteria. Since there are more than one regression coefficients in the

---

[13]If we take into consideration the cost of computing the estimates of $\bar{\beta}_j$ and the variances of $\bar{\beta}_j$'s the OLS estimator is roughly 80 percent of the cost of computation for the SWLS or the TWLS estimators.

Table 7 Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage Efficiency of Alternative Estimators of the Mean Response Coefficients in the CMR Model III

Sample Size	Estimator	MODEL III — SET I						MODEL III — SET II					
		Coefficient: $\bar{\beta}_1 = 0.8$			Coefficient: $\bar{\beta}_2 = 0.6$			Coefficient: $\bar{\beta}_1 = 30.2$			Coefficient: $\bar{\beta}_2 = 10.1$		
		% RMSE	% QD	% E	% RMSE	% QD	% E	% RMSE	% QD	% E	% RMSE	% QD	% E
10	SALS	37.1174	18.3308	53.9984	44.8290	51.3128	10.0756	5.4946	17.0195	23.5318	11.5512	100.0121
	TWLS	30.4566	17.9510	78.5895	41.0049	77.1915	96.4793	2.7689	83.0603	25.7384	13.9765	82.2453
	HH	39.3181	17.8980	47.2763	100.7546	42.4525	16.0143	2.9636	6.8424	92.2014	14.6159	6.4223
	OLS	27.3746	17.8505	100.0000	69.4368	44.0547	4.2134	2.8308	100.0000	23.5312	14.4244	100.0000
20	SALS	16.3824	10.9341	93.8458	27.5383	20.3698	3.4163	1.6837	46.8856	12.7522	6.4502	43.3581
	TWLS	17.6301	10.9960	81.1516	30.3186	20.2310	2.3609	1.6751	97.7778	8.5349	6.2224	96.8479
	HH	19.1364	10.8048	60.1284	34.2580	20.9528	2.4656	1.8303	89.5547	9.1060	6.5069	84.9825
	OLS	15.8761	10.7241	100.0000	26.9006	22.0441	2.3342	1.6249	100.0000	8.3972	6.2809	100.0000
50	SALS	10.2756	6.3318	103.8578	19.8187	13.5619	101.8308					
	TWLS	10.2759	6.3508	103.8528	19.8258	13.6009	101.7587		Not Available			
	HH	10.3914	7.0933	101.5409	20.2141	13.7211	97.6988					
	OLS	10.4934	5.7521	100.0000	20.0396	11.8374	100.0000					

model and the rankings for each sometimes differs, we have computed the overall % RMSE and overall % QD ranking (analogous to Table 6), which is given in Table 8.

On the overall % RMSE ranking the OLS estimator continues to out-perform other estimators both in sample 10 and 20 as in the CMR model II. Furthermore, with one exception the OLS does very well on the overall % QD ranking also (as it ranks first or share first rank with other estimators). The one exception is for sample size 10 - Set II where it ranks second to the TWLS estimator. Once again in the overall % RMSE ranking the TWLS estimator ranks second in Set I for sample size 10. The second ranked estimator in Set II is shared by the SALS and the TWLS estimators. For sample size 20 the second ranked estimators are the same as in the CMR Model II for Set I and Set II.

Table 8   Overall Percentage Root Mean Square Error and Overall Percentage Quartile Deviation Ranking of Alternative Estimators of the Mean Response Coefficients in the CMR Model III

Sample Size	Estimator	MODEL III — SET I		MODEL III — SET II	
		Overall % RMSE Rank	Overall % QD Rank	Overall % RMSE Rank	Overall % QD Rank
10	SALS	6	8	5	5
	TWLS	4	4	5	3
	HH	8	4	8	7
	OLS	2	4	2	5
20	SALS	4	5	8	6
	TWLS	6	5	4	3
	HH	8	5	6	8
	OLS	2	5	2	3
50	SALS	2	4	Not Available	
	TWLS	4	6		
	HH	7	8		
	OLS.	7	2		

As pointed out earlier, we performed the sampling experiments for sample size 50 for the CMR Model III - Set I. The results are summarized in Tables 7 and 8 along with the results on sample sizes 10 and 20. It may be seen that on the overall % RMSE ranking the OLS estimator does marginally worse than the SALS and the TWLS estimators. Furthermore, the ranking of alternative estimators is close to the intuitive ranking of estimator set forth in Chapter IV with the exception that the HH estimator does not rank above the OLS estimator but shares the rank with the OLS estimator. On the overall % QD ranking, however, the OLS continues to be superior over other estimators.

The OLS estimator was shown to have empirical bias for sample size 10 and therefore efficiency comparison relative to the OLS estimator as computed by descriptive statistics % E is not very meaningful. However for other estimators the relative efficiency have been computed as percentage of the OLS estimator and they can be read from Table 7 under the column for % E.

In conclusion, on the % RMSE criterion, the choice once again is between the OLS and the SALS or the TWLS estimators for all sample sizes. Therefore, all the arguments given above in the concluding remarks on the CMR Model II regarding choice between alternative estimators hold for the CMR Model III also.

V.1.1.3   Forecasting Efficiency Ranking of Alternative
           Estimators in the CMR Models

The choice of a relevant criterion depends on the
purpose of the sampling experiments. Insofar as the purpose
is to have a precise knowledge of the structural parameters
the criterion of efficiency outlined in Section V.1.1.2
seems satisfactory. The interest, however, does not have to
be limited to the precise knowledge of the structural para-
meters only; for example, the interest may be in the con-
ditional forecasting precision of alternative estimators.
In the latter case the criterion of conditional forecasting
(forecasting for short) ability for alternative estimators
will be more relevant. In Chapter III we developed the
theoretical framework for this purpose in the context of the
HH estimator. We further noted that the analysis can be
extended to other estimators. In this section we shall
present the results of the sampling experiments on the fore-
casting ability of alternative estimators. The descriptive
statistics used for the comparison are the % RMSE, the % QD and
the % E. The results from the Monte Carlo experiments for
the three models and samples sizes 10, 20 and 50 (for the CMR
Model III-Set I only) are given in Table 9 to Table 11.

Forecasting Efficiency of Alternative Estimators in
the Proportional CMR Model I.

It can be easily verified from Table 9 that for a
given sample size the numbers within each descriptive

statistic are the same for the HH, the TWLS and the SALS estimators.

Table 9   Percentage Root Mean Square Error, Percentage Quartile Deviation, and Percentage Efficiency of Forecast of Alternative Estimators in the CMR Model I

Sample Size	Estimator	MODEL I — SET I			MODEL I — SET II		
		Forecast			Forecast		
		% RMSE	% QD	% E	% RMSE	% QD	% E
10	SALS	11.1827	4.5092	150.7882	2.2697	1.1753	150.7892
	TWLS	11.1827	4.5092	150.7882	2.2697	1.1753	150.7882
	HH	11.1827	4.5092	150.7882	2.2697	1.1753	150.7882
	OLS	12.1868	6.4381	100.0000	2.7429	1.6781	100.0000
20	SALS	7.0768	4.5593	175.4184	1.5780	1.1412	175.4184
	TWLS	7.0768	4.5593	175.4184	1.5780	1.1412	175.4184
	HH	7.0768	4.5593	175.4184	1.5780	1.1412	175.4184
	OLS	8.8768	5.8465	100.0000	2.0796	1.4634	100.0000

As explained in the discussion on the bias and efficiency in Sections V.1.1.1 and V.1.1.2, this is due to special nature of the proportional CMR Model I. It can be easily verified from the Table 9 that the ranking of alternative estimators is the same as efficiency ranking obtained in the Section V.1.1.2. On the % RMSE and the % QD criteria, the SALS, the TWLS and the HH rank equal and each ranks higher than the OLS estimator both in Set I and Set II within each sample sizes of 10 and 20.

On the % E criteria we can say that for sample sizes 10 and 20 respectively the SALS, the TWLS and the HH estimators of forecast are 50 and 75 percent more efficient than

the OLS estimator of forecast both in Set I and Set II.[14]

Forecasting Efficiency of Alternative Estimators in the Bivariate CMR Model II with an Intercept

Table 10 gives the % RMSE, the % QD and the % E of alternative estimators in the Bivariate Model II with an intercept.

The forecasting efficiency ranking on the % RMSE criterion is similar in many respects to the efficiency ranking in Table 5 - Table 6. For instance, the OLS estimators ranks first for sample sizes 10 and 20 on the % RMSE

Table 10 Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage Quartile Deviation and Percentage Efficiency of Forecast of Alternative Estimators in the CMR Model II

Sample Size	Estimator	MODEL II — SET I			MODEL II — SET II		
		Forecast			Forecast		
		% RMSE	% QD	% E	% RMSE	% QD	% E
10	SALS	21.4109	9.0860	30.0381	2.5073	1.6649	80.6443
	TWLS	12.7647	8.4772	83.2153	2.5636	1.6555	77.3822
	HH	19.3158	8.3533	35.7032	4.5932	1.5802	24.0363
	OLS	11.5656	7.7797	100.0000	2.2641	1.5155	100.0000
20	SALS	8.7993	6.2654	97.8588	1.8664	1.2832	84.1409
	TWLS	10.3685	6.7087	71.5711	1.7686	1.2657	93.4837
	HH	10.5764	7.0110	67.3175	1.9108	1.2656	80.2887
	OLS	8.7139	6.0060	100.0000	1.7101	1.3041	100.0000

[14]We ran the t-test for bias of forecast in the CMR Model I (sample sizes 10 and 20) and found that the forecast is unbiased. The figures are however not included to conserve space. This finding is in concurrence with theoretical expectation.

criterion. For sample size 10, the TWLS estimator is second best and is 17 percent less efficient than the OLS estimator in the Set I, while the SALS estimator ranks second and is 19 percent less efficient than the OLS estimator. For sample size 20 the SALS estimator ranks second in Set I while the TWLS estimator ranks second in Set II respectively. Thus we find that phenomenon of ranking reversal takes place both across the sample sizes and across points in the parameter space. The viotile behaviour of the HH, the TWLS and the SALS estimators can be attributed to their dependence on negative estimates of variances of random coefficients.

The ranking of alternative estimator on the % QD criterion sometimes differs in the two sets of values of parameters from that in Table 6. For example in Set I for sample size 20 the OLS estimator ranks first on the fore-casting efficiency criteria in Table 9 while the SALS esti-mators ranks first according to efficiency criteria in Table 6. Similarly the ranking on the % QD differs in Set II for sample size 20. The ranking for sample size 10 in both Set I and Set II on the % QD criterion is the same as in Table 6.

<u>Forecasting Efficiency of Alternative Estimators in the Trivariate CMR Model III</u>

The results of sampling experiments on alternative estimators in the trivariate CMR Model III are contained in Table 11.

Table 11 Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage Efficiency of Forecast of Alternative Estimators in the CMR Model III

Sample Size	Estimator	MODEL III – SET I Forecast			MODEL III – SET II Forecast		
		% RMSE	% QD	% E	% RMSE	% QD	% E
10	SALS	15.1410	8.4513	57.3053	3.0353	1.7041	54.7567
	TWLS	12.5322	8.3540	84.7828	2.4024	1.6014	86.6744
	HH	15.5322	8.2045	54.6943	7.6635	1.5387	8.6237
	OLS	11.4960	7.7391	100.0000	2.2443	1.5027	100.0000
20	SALS	8.4729	6.0597	98.2725	2.2718	1.1644	50.2188
	TWLS	9.3278	6.1528	82.1631	1.6306	1.1410	97.3126
	HH	10.3693	6.6565	63.6995	1.7420	1.1328	85.4035
	OLS	8.3851	5.8044	100.0000	1.6089	1.1727	100.0000
50	SALS	4.1822	2.0944	100.4199			
	TWLS	4.1823	2.0961	100.3975		Not Available	
	HH	4.1835	2.1095	99.7658			
	OLS	4.1835	2.1095	100.0000			

The results on the forecasting efficiency of the CMR Model III are similar to the results of the CMR Model II for sample sizes 10 and 20 as can be verified from Table 10 and Table 11. The only exception being that in the Set II for sample size 10 the TWLS estimator does better than the SALS estimator which is contrary to the findings for the CMR Model II.

For sample size 50 we find that the SALS estimator of forecast is the best while the TWLS estimators performs the second best. The HH and the OLS estimators perform equally well. This finding is in concurrence with the findings in Table 8.

For each of three sample sizes and two set of values (except Set II for Model III) the relative efficiency of alternative estimators has been expressed as percentage of the OLS estimator, which may be read from Table 1.[15]

In Set I, for sample sizes 20 the forecasting efficiency ranking on the % QD criterion is the same as that on the % RMSE. In Set II, however, the forecasting efficiency ranking on the % QD criterion differs from that on the % RMSE. The ranking is identical to the ranking in Set II of Model II on the % QD criterion (see Table 8).

In conclusion, the first and second rank positions for alternative estimators is the same for efficiency and forecasting efficiency criteria.

V.1.1.4   Coefficient of Multiple Determination Ranking of Alternative Estimators.[16]

Having ranked the SALS, the TWLS, the HH and the OLS estimators in terms of efficiency in Section V.1.1.2 and in terms of forecasting efficiency in Section V.1.1.3, we go on to rank the estimators on the basis of the mean and the median

---

[15]The percentage bias figures in the CMR Model III for sample sizes 10 and 20 were generally less than 2 percent which suggest that estimators are unbiased. We did not however, specifically run t-test for the unbiasedness of alternative estimators of forecasting.

[16]The Coefficient of Multiple Determination, R^2, is the square of the Multiple Correlation Coefficient, R.

values of the coefficients of multiple Determination (R^2) as obtained in the sampling experiments. The formula for calculating R^2 was developed in Chapter III for the HH estimator. Further, it was noted that the formulae for other estimators can easily be obtained on the lines of formula for the HH estimator. The criteria for ranking alternative estimators is that the higher the sample mean and the median value of R^2 the better the estimator.

The mean and median values of R^2, for alternative estimators for the three models and sample sizes 10, 20 and 50 (for the CMR Model III-Set I only) were computed. The mean and median values of R^2 were generally very close. Most of the times the difference were in the third and fourth decimal places only.[17] In many instances the SALS and TWLS marginally perform better than the other two estimators but no systematic trend seemed to emerge. In conclusion the ranking of alternative estimators on the basis of coefficients of Multiple Determination was rather ambiguous.

---

[17]The numbers for R^2 for alternative estimators were suppressed to conserve space.

V.1.2   Estimation of the Standard Errors of the Mean
         Response Coefficients.

V.1.2.1 Bias Analysis of Alternative Estimators of The
         Standard Errors of the Mean Response Coefficients.

The standard error of an estimator has special
importance in testing of hyoptheses. A biased standard
error is undesirable because it makes the testing of hypo-
theses unreliable. For example, if the standard error of
an estimator is biased upward, we will commit more Type I
errors. On the other hand if the standard error of an
estimator is downward biased, we will commit more errors of
Type II.[18] As pointed out in Chapter IV, the exact standard
errors of the SALS, the TWLS, the HH and the OLS estimators
are not known. Therefore, in small samples also we compute
the standard errors of alternative estimators using the asymp-
totic formulae given in Chapter III. It is clear from the
above discussion that we need to study the bias properties
of asymptotic standard errors of alternative estimators. In
this section we shall study the bias properties of alterna-
tive asymptotic standard errors in the three models for
sample sizes 10, 20 and 50 (for the CMR III - Set I only).

---

[18]Type I Error is committed if a hypothesis is
rejected when it should be accepted, while Type II Error
is committed if a hypothesis is accepted when it should be
rejected.

The bias of an asymptotic standard error has been computed as a difference between the mean of asymptotic standard error (obtained from sampling experiments) and the standard deviation of the mean response coefficients estimates (taken from Section V.1.1.2). The descriptive statistics used for bias analysis are the % Bias and the % M Bias and the t-statistic.[19]

Bias Analysis of Alternative Estimators of the Standard Error of the Mean Response Coefficient in the Proportional CMR Model I.

The results of sampling experiments on the bias of alternative estimators of standard error of the mean response coefficient for sample Sizes 10 and 20 in the CMR Model I are given in Table 12.

The values of the t-statistic for the bias of alternative estimators of standard error for sample Size 10 lie outside the interval - 1.96 to 1.96 and hence all estimators of standard error are biased. However, the SALS is the least biased estimator. The ranking of alternative estimators of standard on the basis of their % Bias and % M Bias performance is in accordance with the intuitive ranking of estimators of mean response given by us in Chapter V.

---

[19]It is interesting and important to test the small sample distributions of the standardised ratio t = $\hat{\theta}-\theta/se(\hat{\theta})$ for normality. This is usually tested by the Kolmogrov-Smirnov test. The test is fully described by Sidney Siegel, <u>Non parametric Statistics for Behavioral Sciences</u>, New York: MCGraw-Hill (1956) pp. 229-238. This test was not applied for limitation of computer and research time.

Table 12 Percentage Bias, Percentage Median Bias and t-statistic of Asymptotic Standard Error of Alternative Estimator of the Mean Response Coefficient in the CMR Model I

Sample Size	Estimator	MODEL I – SET I			MODEL I – SET II		
		$SE(\hat{\bar{\beta}}_1)$			$SE(\hat{\bar{\beta}}_1)$		
		% Bias	% M Bias	t	% Bias	% M Bias	t
10	SALS	6.3965	8.4707	−2.5601	6.3968	8.4711	−2.5602
	TWLS	6.6320	9.7026	−2.6763	6.6319	9.7026	−2.6763
	HH	10.0893	14.6567	−3.5784	10.0893	14.6569	−3.5784
	OLS	14.4929	18.8366	−5.4050	14.4917	18.8355	−5.4044
20	SALS	0.1755	0.0166	−0.1144	0.1755	0.0166	−0.1116
	TWLS	0.2615	0.4337	−0.2405	0.2615	0.4337	−0.1710
	HH	2.7331	7.5680	−1.2104	2.7332	7.5680	−1.2104
	OLS	2.0324	6.9021	−19.2566	2.0324	6.9021	−0.8936

The values of the t-statistic for the bias of alternative estimators for sample Size 20 lie between -1.96 and 1.96 at .05 level of significance both in Set I and Set II, except for the OLS estimator in Set I.[20] Therefore, the standard error are unbiased in all estimators except for the OLS estimator of the standard error for sample Size 20. Once again the SALS estimator of the standard error is the least biased estimator in the CMR Model I on all criteria for sample Size 20.

---

[20]The standard errors of the OLS estimates of the mean response coefficients have been obtained as square root of the variance given in (III.20) replacing Φ by its consistent estimator obtained by HH Method, throughout in this section. It is well known that the conventional formula of variance, viz. $\hat{\sigma}^2(x'x)^{-1}$ yields biased and inconsistent estimates of standard error if disturbances in the regression are heteroskedasic. The bias of standard errors computed for sample Size 20 by the conventional formula were mostly higher than standard errors obtained from consistent formula. More importantly the behaviour of bias of standard errors obtained from standard formula was

Bias Analysis of Alternative Estimator of the Standard
Errors of the Mean Response Coefficients in the Bivariate
CMR Model II.

The results of the sampling experiments for sample
Sizes 10 and 20 in the CMR Model II are given in Table 13.
An examination of the sampling results will show that the
OLS estimator of the standard error of the mean response
coefficient is unbiased for sample Size 10 at .05 level of
significance. On the other hand the SALS and the TWLS (for
β_1 only) estimators of the standard error of the mean re-
sponse coefficients in Set I and Set II respectively have
the t-statistic values lying inside the interval -1.96 to
1.96 and hence these estimators are unbiased at .05 level
of significance for sample Size 20. In all other cases the
t-statistic values lie outside the interval -1.96 to 1.96
and hence all other estimators of the standard error of the
mean response are biased.

Since ranking of alternative estimators for the two
coefficients sometimes differ in terms of the % Bias and the
% M Bias we have computed the overall % Bias and overall % M
Bias rankings analogons to the overall % RMSE and the over-
all % OD rankings in Table 6 (see footnote 11) and are
given in Table 14.

It can easily be seen from Table 14 that the OLS
estimator is the least biased estimator on both the overall

rather viotile due to inconsistency property of the estimator.

Table 13 Percentage Bias, Percentage Medium Bias and t-statistic of Asymptotic Standard Errors of Alternative Estimators of the Mean Response Coefficients in the CMR Model II

Sample Size	Estimates	MODEL II — SET I						MODEL II — SET II					
		SE ($\hat{\beta}_1$)			SE ($\hat{\beta}_2$)			SE ($\hat{\beta}_1$)			SE ($\hat{\beta}_2$)		
		% Bias	% M Bias	t	% Bias	%M Bias	t	% Bias	%M Bias	t	% Bias	%M Bias	t
10	SALS	53.6348	56.0759	-32.7098	55.1352	57.5408	-33.0176	12.1206	14.7607	-4.4305	18.6492	20.3260	-7.2108
	TWLS	12.0495	13.4857	-4.2990	18.7794	20.3252	-6.9049	15.0521	19.7801	-4.9280	19.9287	21.5896	6.9844
	HH	43.3747	47.8443	-18.6870	48.0832	51.1012	-19.7176	56.1067	59.2357	-29.9960	59.2216	61.5255	-32.2944
	OLS	8.1662	4.3128	1.6869	2.1423	4.7327	-0.7533	5.4254	3.0920	1.7486	1.5665	3.0198	-0.5533
20	SALS	0.0953	3.8881	0.0359	2.6589	1.9651	-0.8244	11.5764	13.0182	-7.2987	15.2229	15.5189	-9.6382
	TWLS	17.3004	18.7135	-10.7534	22.4486	22.3368	-15.2767	0.7856	3.1720	-0.4324	5.1611	5.9519	-3.0690
	HH	18.6584	17.7933	-9.9790	22.7713	23.8273	-17.4494	5.6321	6.6180	-3.0469	10.1287	10.8002	-6.1696
	OLS	6.9671	5.2120	4.1824	3.1566	3.9382	1.8703	6.7144	5.1641	3.9833	4.4860	5.3099	2.3228

Table 14   Overall Percentage Bias, Overall Percentage Median Bias Ranking of the Asymptotic Standard Errors of Alternative Estimators of the Mean Response Coefficients in CMR Model II

Sample Size	Estimator	MODEL II – SET I		MODEL II – SET II	
		Overall % Bias Rank	Overall % M Bias Rank	Overall % Bias Rank	Overall % M Bias Rank
10	SALS	8	8	4	4
	TWLS	4	4	6	6
	HH	6	6	8	8
	OLS	2	2	2	2
20	SALS	2	2	8	8
	TWLS	6	7	3	3
	HH	8	7	5	6
	OLS	4	4	4	3

% Bias and the overall % M Bias ranking criteria for sample Size 10. the SALS and the TWLS estimators on the other hand are respectively least biased estimators in Set I and Set II on both the overall criteria of the % Bias and the % M Bias.

If we combine the results in Table 7, Table 10 and Table 14 we can conclude that the OLS estimator is the best choice for sample Size 10, while the SALS and the TWLS estimators are respectively the best choices when parameters are small and large for sample Size 20.

Bias Analysis of Alternative Estimators of the
Standard Error of the Mean Response Coefficients
in the Trivariate CMR Model III.

The summary of the Monte Carlo experiments on the bias of alternative estimators of the standard error in the CMR Model III is given in Table 15.

Table 15   Percentage Bias, Percentage Median Bias and t-statistic of Asymptotic Standard Errors of Alternative Estimators of Mean Response Coefficients in the CMR Model III

Sample Size	Estimates	MODEL III — SET I						MODEL III — SET II					
		$SE(\bar{\hat{\beta}}_1)$			$SE(\bar{\hat{\beta}}_1)$			$SE(\bar{\hat{\beta}}_1)$			$SE(\bar{\hat{\beta}}_1)$		
		% Bias	%M Bias	t	% Bias	%M Bias	t	% Bias	%M Bias	t	% Bias	%M Bias	t
10	SALS	24.9556	28.1560	-10.1311	31.4233	33.7793	-12.5907	33.4248	37.5267	-9.7712	34.9706	39.9961	-12.1536
	TWLS	11.2661	14.6061	-4.0435	17.9479	18.8608	-6.5731	9.7176	14.0742	-3.2650	15.4806	16.4384	-5.4265
	HH	28.6305	33.1471	-11.3623	34.0141	37.0915	-13.1909	72.2282	74.4633	-50.0352	74.8703	76.7547	-52.2618
	OLS	4.7719	4.4410	1.5629	2.1888	4.3186	-0.7740	5.0163	2.9410	1.6225	1.6553	2.7152	-0.5875
20	SALS	2.4645	4.6586	-1.4307	5.1573	6.4510	-2.9001	30.3437	31.4567	-25.1542	34.8976	35.3571	-31.0259
	TWLS	7.5468	12.3326	-2.1765	12.5549	16.2613	-3.4725	0.7236	0.1625	0.4179	3.1243	4.9406	-1.8637
	HH	19.9264	18.6542	-10.3482	24.6172	25.4554	-18.8092	4.8503	5.0733	-2.4656	8.8961	9.6847	-5.6118
	OLS	6.6056	6.2693	3.9747	4.0653	4.2249	2.0229	6.5030	5.2398	3.8893	4.6282	0.9631	2.3972
50	SALS	6.8513	7.8302	-6.7427	15.4180	15.0910	-15.8525						
	TWLS	6.8527	8.3649	-6.7560	15.3497	15.8903	-15.8142			Not Available			
	HH	9.8538	11.5019	-8.2088	16.8958	17.3086	-18.3476						
	OLS	3.1022	3.5394	-3.1344	9.2798	10.8604	-6.5634						

An examination of Table 15 shows that in most cases on the % Bias and the % M Bias criteria the least biased estimators of the standard error are different for the two regression coefficient estimates (i.e. $\hat{\beta}_1$ and $\hat{\beta}_2$). Therefore, we will choose the least biased estimator among alternative estimators on the overall % Bias and overall % M Bias ranking criteria, tabulated below in Table 16.

Table 16   Overall Percentage Bias, Overall Percentage Median Bias Ranking of Alternative Estimators of Standard Errors in the CMR Model III

Sample Size	Estimator	MODEL III — SET I		MODEL III — SET II	
		Overall % Bias Rank	Overall % M Bias Rank	Overall % Bias Rank	Overall % M Bias Rank
10	SALS	6	6	6	6
	TWLS	4	4	4	4
	HH	8	8	8	8
	OLS	2	2	2	2
20	SALS	3	3	8	8
	TWLS	6	6	2	3
	HH	8	8	5	5
	OLS	4	3	5	4
50	SALS	5	4		
	TWLS	5	6		
	HH	8	8	Not Available	
	OLS	2	2		

An examination of Table 15 and 16 will show that the conclusions on the CMR Model II regarding ranking of the alternative estimators of the standard errors of the mean response coefficients are upheld again for sample Size 10

and 20. Furthermore if we combine the results in Table 8, Table 11 and Table 16 we can once again conclude that the OLS estimator is the best choice for sample Size 10, while for the sample Size 20 the SALS is the best choice if the parameter values are 'small' and the TWLS is the best choice if the parameter values are 'large'.

On the basis of the t-test for sample Size 50, the alternative estimators of the standard error of the mean response coefficients are all biased at .05 level of significance. However the OLS estimator of the standard error is the least biased estimator on the basis of the overall % Bias and overall % M Bias ranking in Table 16. Further, if we combine the results in Table 9, Table 11 and Table 16, the OLS may not be a bad choice considering the cost of computation. However, if we could evaluate the bias of standard error of mean response and correct it, the SALS or the TWLS would be preferable to the OLS estimator.

V.1.3   Estimation of the Variance of Random Coefficients

Estimation of the variance of random coefficients (variance for short) is an important aspect of the estimation and the hypothesis testing in the linear regression model with random coefficients. The variances are required to estimate the mean response coefficients efficiently (except in the porportional CMR Model I and OLS estimator) and to estimate the standard errors of the mean response coeffi-

cients. Besides, the values of the variances influence our faith in the specification of the model and the quality of data.

In this section we shall study the sampling properties of the three estimators of variances. We are interested in studying the bias and efficiency properties of the variance, and the bias of the standard error of variance for alternative estimators. In the CMR Model I, such an analysis seems straight forward on the lines of analysis of the sampling properties of alternative estimators of the mean response coefficients. This is because the CMR Model I has no problem of negative estimates of variances. The possibility that some of the estimates of the variances be negative in an estimator makes various descriptive statistics meaningless and hence an investigation of sample properties in terms of these statistics is ambiguous.[21] Since the problem of negative estimates of variance exists for the CMR Model II and the CMR Model III, we will not do the traditional monte carlo analyses on the alternative estimators of variances for the CMR Models II and III. Instead, we shall measure the efficiency of alternative estimators by the proportion of negative estimates of variances obtained

---

[21]For example, if negative estimates are obtained fifty percent of the time (which may well be) then, there is a possibility that the sample mean will be zero and so on. Some of the descriptic statistics such as median bias, fourth quartile and ranges may still be meaningful.

in each estimator in the sampling experiments (see Fig. III.1 Chapter III).

V.1.3.1 Bias Analysis of Alternative Estimators of Variances of the Random Coefficients in the CMR Models.

Bias Analysis of Alternative Estimators of Variances in the Proportional CMR Model I.

Table 17 contains the results of sampling experiments on alternative estimators of variances for two sets of values of parameters in the CMR Model I for sample Sizes 10 and 20. It may be seen that the t-statistic values for all the estimators of variances lie inside the interval -1.96 and 1.96 at 5 percent level of significance for sample Sizes 10 and 20 in both sets. Therefore, all the estimators of variance are unbiased emperically.

Table 17 Percentage Bias, Percentage Median Bias and t-statistic of Alternative Estimators of Variances of the Random Coefficient in the CMR Model I

Sample Size	Estimator	MODEL I — SET I			MODEL I — SET II		
		Coefficient: $\sigma_1^2 = 0.03$			Coefficient: $\sigma_1^2 = 0.5$		
		% Bias	%M Bias	t	% Bias	%M Bias	t
10	SALS	9.0467	25.3898	−1.4344	9.0467	25.3898	−1.4344
	TWLS	4.4089	16.4763	−0.7865	4.4089	16.4763	−0.7865
	HII	3.8531	14.1817	−0.6845	3.8531	14.1817	−0.6845
20	SALS	2.2837	0.2421	0.7484	2.2837	0.2421	0.7484
	TWLS	2.0968	0.5926	0.6894	2.0968	0.5926	0.6894
	HII	1.8201	14.3280	−0.0009	1.8201	14.3280	−0.0039

The unbiasedness of variances for the HH estimator is in concurrence with theoretical results of Chapter III. However, we did not give a theoretical proof of unbiasedness for the TWLS and the SALS estimators of variance. Hence the emperical results here are rather important.[22]

As pointed out earlier, such an analysis is not feasible for the CMR Model II and the CMR Model III and therefore, we now go on to discuss the efficiency property of alternative estimators.

V.1.3.2 Efficiency Analysis of Alternative Estimators of Variances of the Random Coefficients.

Efficiency Analysis of Alternative Estimators of Variances in the Bivariate CMR Model I.

The % RMSE, the % QD and the % E statistics have been computed for the efficiency analysis of alternative estimators in the CMR Model I and are given in Table 18.[23] The SALS and the TWLS estimators are more efficient than the HH estimator on the % RMSE, the QD and the % E criteria as conjectured on theoretical grounds in Chapter III for sample sizes 10 and 20.[24]

---

[22]It is our belief that the proof of unbiasedness for the TWLS and the SALS estimators of variances can be constructed on the lines of Kakwani (1967) arguments.

[23]The % E statistic throughout section V.1.3.2 has been computed as $\dfrac{\text{Var } (\hat{\theta}) \text{ SALS or TWLS}}{\text{Var } (\hat{\theta}) \text{ HH}} \times 100.0$.

[24]The HH estimator is in fact the Ordinary Least squares estimator of variances.

Table 18   Percentage Root Mean Square Error and the Percentage Quartile
Deviation of Alternative Estimators of the Variances of Random
Coefficients in the CMR Model I

Sample Size	Estimator	MODEL I – SET I			MODEL 1 – SET II		
		Coefficient: σ_1^2 = 0.03			Coefficient: σ_1^2 = 0.5		
		% RMSE	% QD	% E	% RMSE	% QD	% E
10	SALS	56.4253	24.5700	125.5321	56.4253	24.5700	125.5321
	TWLS	56.2303	27.9571	126.5933	56.2303	27.9571	126.5933
	HH	63.7175	28.2943	100.0000	63.7175	28.2943	100.0000
20	SALS	30.5996	21.8376	222.8216	30.5996	21.8376	222.8216
	TWLS	30.4855	21.5068	224.3038	30.6855	21.5068	224.3038
	HH	45.5493	31.0804	100.0000	45.5493	31.0804	100.0000

However, our conjecture that the SALS estimator
is more efficient than the TWLS estimator does not seem to
hold.

The efficiency analysis of the above type is not
meaningful for the CMR Model II and III as pointed out
earlier and for these Models we will carry out efficiency
analysis in terms of proportion of negative variance esti-
mates (See Fig. III.1).

Efficiency Analysis as Measured by the Percentage
Negative Estimates of the Variances of the Random
Coefficients in the Bivariate CMR Model II.[25]

The percentage negative estimators of the variance
of the random coefficients are tabulated in Table 19 for

---

[25]Since the sampling experiments for each estimator
is repeated 100 times, the actual and percentage numbers of
negative estimates for each parameter are identical.

alternative estimators for sample Sizes 10 and 20 in the CMR Model II.

An examination of Table 19 brings out a few interesting points. The percentage of negative estimates of variance does not necessarily decrease, as may be expected intuitively, when parameters values are increased from Set I to Set II. A priori value of σ_2^2 in Set II is 17 times its value in Set I but the percentage of negative estimates

Table 19 Percentage Negative Estimates of Variances in Alternative Estimators in the CMR Model II

Sample Size	Estimator	MODEL II — SET I		MODEL II — SET II	
		Coefficient: $\sigma_1^2 = 0.04$	Coefficient: $\sigma_2^2 = 0.03$	Coefficient: $\sigma_1^2 = 1.5$	Coefficient: $\sigma_2^2 = 0.5$
10	SALS	10	48	9	49
	TWLS	11	46	9	49
	HH	14	45	10	49
20	SALS	3	16	0	30
	TWLS	1	17	0	29
	HH	7	15	3	30

of variances is increased one and a half times for sample Size 20, while it remained constant for sample Size 10. However, when a priori value of σ_1^2 in Set II is nearly 40 times its value in Set I, the percentage of negative estimates of variance is decreased for sample Size 20. Therefore no definite relationship seems to exist between the values of parameters to be estimated and the percentage of negative estimates obtained for an estimator.

The rationale for efficiency analysis of alternative estimators in terms of the number of negative estimates of variance was given in Fig. III.1 of Chapter III. It can be seen that in Set I and Set II, the overall percentage negative estimates of variances (the sum of negative estimates of variances of σ_1^2 and σ_2^2) for each of the SALS and the TWLS estimators is less than the overall percentage negative estimates of variances for the HH estimator for all sample Sizes.[26] In this sense the SALS and the TWLS estimators are more efficient than the HH estimator. Further, the TWLS estimator does marginally better than the SALS estimator except in the case of Set II for sample Size 10 where it does as well as the SALS estimator.[27]

One representive Histogram for estimates of σ_2^2 showing the negative estimates is given in Appendix 2.

---

[26]Strictly speaking the percentage of negative estimates of variance is to be compared for alternative estimators for each of two parameters σ_1^2 and σ_2^2 separately. If this be the case then, only the TWLS estimator in Set II will be more efficient than the TWLS and the HH estimators. Further, it may be pointed out that in the proposed comparative efficiency analysis in Fig. III.1, Chapter III holds for the smmetric distribution only. (And perhaps between distributions with a given skewness.)

[27]In the CMR Model III, the reverse is true.

[28]The histogram has been plotted through the use of HIST routine available in the SSP.

Efficiency Analysis as Measured by the Percentage
Negative Estimates of Variances of Random Coefficients
in the Trivariate CMR Model III.

In table 20 we have given the percentage negative estimates of variances of random coefficients as obtained from the sampling experiments on the CMR Model III for sample Sizes 10, 20 and 50 (for the Set I only).

An examination of Table 20 reveals that many of the conclusions on the CMR Model II for sample sizes 10 and 20 (refer to the discussion in Table 19) continue to hold.

Table 20   Percentage Negative Estimates of Variances of Alternative Estimators in the CMR Model III

Sample Size	Estimator	MODEL III — SET I		MODEL III — SET II	
		Coefficient: $\sigma_1^2 = 0.04$	Coefficient: $\sigma_3^2 = 0.03$	Coefficient: $\sigma_1^2 = 1.5$	Coefficient: $\sigma_3^2 = 0.5$
		No. of -ve Est.	No. of -ve Est.	No. of -ve Est.	No. of -ve Est.
10	SALS	10	48	29	41
	TWLS	11	47	9	49
	HH	16	45	12	48
20	SALS	1	26	0	39
	TWLS	2	27	0	40
	HH	7	23	5	39
50	SALS	0	15		
	TWLS	0	15		
	HH	3	15		

For example, we find that there is no definite relationship between the priori values of variances and the percentage of negative estimates of variances in alternative estimator for sample Sizes 10 and 20. Further, on the criterion of over-all percentage negative estimates of variances, the SALS (except for sample Size 10 - Set II) and the TWLS estimators are more efficient than the HH estimator for sample Sizes 10 and 20.[29]

In the Set II the SALS estimator of variance does worse than the TWLS and the HH estimators for sample Size 10, while the SALS estimator of variance does marginally better than the TWLS estimator for sample Size 20.

For sample Size 50, the SALS and the TWLS estimators of variances perform equal while each perform better than the HH estimator.

It may be seen that as we increase the sample Size the proportion of negative estimates of variances is reduced, which we would expect theoretically.

V.1.4 Estimation of the Standard Errors of Variances

In Section V.1.3 we pointed out the difficulty in obtaining meaningful descriptive statistics such as the root

---

[29]The comments in footnote 26 are again relevant here.

mean square error or the standard deviation in the CMR Model

II and Model III and to carry out the traditional sampling

analysis in terms of moments. It is obvious then that the

bias analysis of standard error of variance can not be done

in the CMR Model II and III because we need the standard

deviations of estimates of variances. In view of this

difficulty we shall do bias analysis for only the CMR

Model I.

V.1.4.1 Bias Analysis of Alternative Estimators of
         the Standard Errors of Variances of the
         Random Coefficients.

Bias Analysis of Alternative Estimators of the Standard
Errors of Variances of the Random Coefficients in the
Proportion CMR Model I.

The results of the sampling experiments on the bias

of alternative estimators of the standard error of variances

of the random coefficients are given in Table 21.

It can be verified from the Table 21 that

the HH estimator of standard error is the least biased

estimator on both the % Bias and % M Bias criteria for sample

Size 10.[30] Furthermore, the HH estimator is also unbiased

on the t-test at .05 level of significance for sample Size

10. For sample Size 20, the HH estimator is the only biased

estimator on the t-test at .05 level of significance, while

---

[30]refer footnote 20, Chapter V.

Table 21   Percentage Bias, Percentage Median Bias and t-statistic of Alternative Estimators
of Standard Error of the Variance of Random Coefficients in the CMR Model I

Sample Size	Estimator	MODEL I — SET I			MODEL I — SET II		
		SE $(\hat{\sigma}_1^2)$			SE $(\hat{\sigma}_1^2)$		
		% Bias	% M Bias	t	% Bias	% M Bias	t
10	SALS	23.4162	37.1773	−4.4092	23.4162	37.1773	−4.4092
	TWLS	27.4392	40.4774	−5.4532	27.4392	40.4774	−5.4532
	HH	6.9804	23.6948	−1.0822	6.9804	23.6948	−1.0822
20	SALS	7.6878	7.7249	1.5670	7.6888	7.7247	1.5636
	TWLS	3.9578	10.9212	0.8357	3.9578	10.9210	0.8357
	HH	19.0348	1.9979	3.5100	18.9478	1.9234	3.4960

the SALS and TWLS estimator of standard error of variances
are unbiased.

V.2   Efficiency and Convergence of the ISALS, the
ITWLS, the IHH Estimators.

In this section we shall discuss the results of
sampling experiments for sample Size 20 on the three iter-
ative estimators discussed in Chapter III. We are interested
in investigating if there is a possibility of improving the
efficiency of the SALS, the TWLS and the HH estimators by
iterating them.[31] Furthermore, we want to know whether the

[31]The SALS, the TWLS and the HH estimators are the
first step in iteration of the ISALS, the ITWLS and IHH es-
timators respectively (see Chapter III).

alternative estimators really converge. We shall study the efficiency properties of alternative iterative estimators of the mean response coefficients in the three models and the variance of random coefficient in the CMR Model I in terms of the % RMSE, the % QD and % E criteria. While the efficiency properties of alternative iterative estimators of the variances of random coefficients in the CMR Models II and III are studied in terms of percentage negative estimates of variances at each iteration. The convergence properties of alternative estimators are studied in terms of the convergence of the mean response coefficients.[32]

Convergence is defined as follows:

An estimator, $\hat{\beta}_j$, is said to converge at iteration τ, if

$$|\bar{\hat{\beta}}_j (\tau+1) - \bar{\hat{\beta}}_j (\tau)| < \delta \quad \text{for all } j \quad (V.1)$$

where δ is any preassigned small number.

In the study of efficiency and convergence of alternative iterative estimators in the CMR Models I, II and III we performed respectively three, six and six iterations.[33]

---

[32]The estimates of variance of the random coefficients are used to estimate the mean response coefficients efficiently and therefore the convergence of the mean response coefficients is important for estimation.

[33]The number of iterations required for convergence in each of the CMR Models were decided on the basis of the performance of alternative estimators in sample runs. The cost consideration was also kept in mind (see footnote 3, Chapter V.)

V.2.1. <u>Efficiency Analysis of Alternative Iterative</u>
<u>Estimators of the Mean Response Coefficients</u>
<u>in the CMR Models for Sample Size 20.</u>

The results of sampling experiments on the efficiency analyses of the alternative iterative estimators of the mean response coefficients in the CMR Models are given in Table 22 to Table 24.[34]

<u>Efficiency Analysis of Alternative Iterative Estimator</u>
<u>of the Mean Response Coefficients in the Proportional</u>
<u>CMR Model I for Sample Size 20.</u>

Table 22 gives the % RMSE, the %QD and the % E statistics for each of the three iterations of alternative iterative estimators of the mean response coefficients.

An examination of Table 22 shows that the ISALS, the ITWLS and the IHH estimators of the mean response coefficient are identical and there is no gain in efficiency at any stage of the three iterations. As pointed out earlier, the SALS, the TWLS and the HH estimators of the mean response coefficients are identical to the Aitken estimator in the proportional CMR Model I. Furthermore, the Aitken estimator does

---

[34]The % E statistics have been computed as:

$$\frac{Var \ (\hat{\theta}) \ ITER = j}{Var \ (\theta) \ ITER = 1} \ X \ 100.0 \ , \ j = 2, 3, \ldots 6$$

throughout this section.

Table 22 Percentage Root Mean Square Error, Percentage Quartile Deviation
and Percentage Efficiency for each of the three iterations of
Alternative Iterative Estimators of the Mean Response Coefficient
in the CMR Model I for Sample Size 20

| Iteration | Estimator | MODEL I — SET I | | | MODEL I — SET I | | |
| | | Coefficient: $\bar{\beta}_1$ = 0.6 | | | Coefficient: $\bar{\beta}_1$ = 10.1 | | |
		% RMSE	% QD	% E	% RMSE	% QD	% E
1	SALS	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
2	ISAS	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
3	ISALS	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
1	TWLS	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
2	ITWLS	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
3	ITWLS	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
1	HH	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
2	IHH	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000
3	IHH	6.4638	4.7152	100.0000	1.5676	1.1435	100.0000

not depend on the knowledge of variance at all and it has
BLU Properties.[35] Therefore, it is expected that there
will be no gain in efficiency in iterating alternative esti-
mators of the mean response coefficients.

---

[35]See footnote 5, Chapter VI.

Efficiency Analysis of Alternative Iterative Estimators of the Mean Response Coefficients in the Bivariate CMR Model II for Sample Size 20.

The results of sampling experiments on alternative iterative estimators in the Bivariate CMR Model II for sample Size 20 are summarized in Table 23. An examination of results will show that except for the IHH estimators in Set II there is an improvement in efficiency on the % RMSE and the % E criteria at one or more stages of iterations. But the fluctuations in efficiency between iterations are often large and there is no systematic trend for convergence (we will come to convergence later).[36]

It may be recalled that the SALS estimator in Set I and the TWLS estimator in Set II were ranked second in the CMR Model II. The gain in efficiency on the % RMSE and the % E criteria in iterating these two estimators is rather small and it shows up only at the sixth iteration. Further, the ISALS (I=6) and the ITWLS (I=6) estimators are respectively less efficient than the OLS estimators on the % RMSE criterion in the CMR Model II (see Table 5 for the efficiency of the OLS estimator).

-----

[36]It will be shown that convergence is achieved in 80 to 90 percent of the cases for δ = 0.0001. The fluctuations in efficiency results from the fact that in 10 to 20 percent cases the convergence is not achieved, which is possibly due to the problem of negative estimates of variances in the CMR Models II and III.

Table 23   Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage
Efficiency for each of the six iterations of Alternative Iterative Estimators of the
Mean Response Coefficients in the CMR Model II for Sample Size 20

Iteration	Estimator	MODEL II – SET I						MODEL II – SET II					
		Coefficient $\bar{\beta}_1 = 0.8$			Coefficient $\bar{\beta}_2 = 0.6$			Coefficient $\bar{\beta}_1 = 30.2$			Coefficient $\bar{\beta}_2 = 10.1$		
		%RMSE	%QD	%E	%RMSE	%QD	%E	%RMSE	%QD	%E	%RMSE	%QD	%E
1	SALS	15.4365	10.9365	100.0000	22.8449	16.5231	100.0000	2.5112	1.5447	100.0000	7.8209	5.2280	100.0000
2	ISALS	24.6246	11.0499	39.3502	39.5809	16.9666	33.3625	2.2289	1.5158	127.0538	6.9840	5.1085	125.5168
3	ISALS	16.6644	10.0905	85.8879	24.4564	17.0783	87.3412	2.5312	1.5525	98.4277	7.9124	5.2609	97.6971
4	ISALS	16.5915	10.2369	86.9937	25.1899	16.6433	82.6755	2.2648	1.5440	123.2378	7.0929	5.1372	121.9227
5	ISALS	18.8500	10.2361	67.2083	27.3687	16.9323	69.8314	2.5317	1.5568	98.3864	7.9136	5.2646	97.6638
6	ISALS	15.1773	10.7778	103.5358	22.5245	15.8527	102.9510	3.8497	1.5774	42.6996	13.0278	5.2608	36.1703
1	TWLS	18.2484	10.9064	100.0000	27.5144	16.9850	100.0000	2.2616	1.6814	100.0000	7.0477	5.2455	100.0000
2	ITWLS	33.4486	11.0791	29.6612	51.7106	17.5103	28.1957	24.2893	15.4815	0.8759	86.7298	51.8655	0.6671
3	ITWLS	19.5917	11.0266	86.6299	29.4463	18.0304	87.1324	2.26747	1.5957	99.5900	7.1359	5.2707	97.7173
4	ITWLS	21.6156	11.0211	71.4544	32.6979	17.5516	70.9581	2.2702	1.5393	99.2738	7.0926	5.2594	99.7743
5	ITWLS	16.2231	11.0151	125.7203	24.0061	17.4788	130.4107	2.2633	1.5957	99.9666	7.1196	5.2983	98.1349
6	ITWLS	15.9686	10.8766	129.9606	24.2028	17.0000	128.4943	2.2340	1.5471	102.5887	7.0010	5.1570	101.1633
1	HH	18.1422	11.4560	100.0000	27.9744	16.8337	100.0000	2.3371	1.6681	100.0000	7.4232	5.3423	100.0000
2	IHH	18.9635	11.7593	91.4078	29.7067	18.2080	88.5096	2.3812	1.6872	96.3505	7.6465	5.4014	94.2883
3	IHH	19.0046	12.4351	91.2667	28.4600	18.1612	96.7819	2.5711	1.7181	83.0226	8.3676	5.2747	79.9517
4	IHH	16.4626	12.4673	121.3393	24.0980	18.1281	125.0769	2.9990	1.7479	61.7962	9.9231	5.4055	56.2338
5	IHH	17.1222	12.8597	111.7866	26.5341	18.1093	110.6046	2.6884	1.6968	75.6210	8.4525	5.3528	77.1964
6	IHH	16.9194	12.8597	114.486	26.2050	18.1453	113.4009	2.8178	1.7181	68.967	9.8039	5.3674	56.3661

The gain in iterating the SALS estimator in Set I
and the TWLS estimator in Set II is negligible both on the
% RMSE, the % QD and the % E criteria in comparison to in-
crease in cost of computation. There is a gain in efficiency
in iterating the TWLS and the HH estimators in Set I on the
% RMSE and the % E criteria. However, the ITWLS and the
IHH estimators rank lower in efficiency than the OLS and the
SALS estimators. If the % QD is the criterian of efficiency,
then there is loss in efficiency iterating the TWLS and the
HH estimators in Set I.

In Set II, there is gain in efficiency in iterating
the SALS estimator on the % RMSE and the % E criteria. But,
the ISALS estimator ranks lower than the TWLS and the OLS
estimators on the % RMSE criterion. There is a loss in
efficiency in iterating the HH estimator on all the three
criteria.

Efficiency Analysis of Alternative Iterative Estimators of the Mean Response Coefficients in the Trivariate CMR Model III for Sample Size 20.

In Table 24 we have summarized the results of samp-
ling experiments on alternative iterative estimators of the
mean response coefficients in the Trivariate CMR Model III
for sample Size 20. The results are similar to the results
in Table 23. For instance, in Set I, there is a slight gain
in efficiency in iterating the SALS estimator on all the
three criteria. The ISALS is less efficient than the OLS

Table 24 Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage Efficiency for each of the Six iterations of Alternative Iterative Estimators of the Mean Response Coefficients in the CMR Model III for Sample Size 20

| | | MODEL III — SET I | | | | | | MODEL III — SET II | | | | | |
| | | Coefficient: $\bar{\beta}_1 = 0.8$ | | | Coefficient: $\bar{\beta}_2 = 0.6$ | | | Coefficient: $\bar{\beta}_1 = 30.2$ | | | Coefficient: $\bar{\beta}_2 = 10.1$ | | |
Iteration	Estimator	% RMSE	% QD	% E	% RMSE	% QD	% E	% RMSE	% QD	% E	% RMSE	% QD	% E
1	SALS	16.3824	10.9341	100.0000	20.3698	20.3698	100.0000	3.4163	1.6837	100.0000	12.7522	6.4502	100.0000
2	ISALS	15.9065	10.0865	106.1829	18.6421	18.6421	105.2876	3.0850	1.6469	121.1528	11.82075	6.2316	115.8732
3	ISALS	16.5178	10.6463	98.3592	19.6057	19.6057	98.3961	2.4367	1.6159	195.7374	8.7730	6.2307	210.3016
4	ISALS	15.9794	10.6462	105.1705	19.3480	19.3480	103.6658	2.3824	1.6261	204.7769	8.6439	6.2306	216.6292
5	ISALS	16.3586	10.6462	100.2889	19.6045	19.6045	100.2448	11.7864	1.6570	8.4529	47.1078	6.4386	7.3600
6	ISALS	15.9943	10.6462	104.9686	19.3822	19.3822	100.8604	2.4000	1.6241	201.7045	8.7162	6.3226	212.9501
1	TWLS	17.6301	10.9960	100.0000	30.3186	20.2310	100.0000	2.3609	1.6751	100.0000	8.5349	6.2224	100.0000
2	ITWLS	16.5223	10.7414	113.7136	28.2039	19.3412	115.3680	97.6504	0.2697	297.3020	91.79802	1.0925	309.7570
3	ITWLS	18.2687	10.9267	93.0926	30.7556	20.3232	97.1327	2.3735	1.6309	98.9872	8.6023	6.2011	98.4853
4	ITWLS	17.9807	10.6480	96.3215	31.2053	19.6962	94.3872	2.4161	1.6335	95.4033	8.8562	6.3019	92.7765
5	ITWLS	16.6294	10.9316	112.225	28.4703	20.0469	113.1998	2.5600	1.6682	85.4514	9.3173	6.3403	84.3659
6	ITWLS	22.5059	10.6415	61.4510	38.0232	20.0469	63.6753	2.3768	1.6240	98.7288	8.6160	6.2011	98.1890
1	HH	19.1364	12.8048	100.0000	34.2580	20.9528	100.0000	2.4656	1.8303	100.0000	9.1060	6.5069	100.0000
2	IHH	16.6583	11.6105	131.5914	28.8747	20.4939	140.4315	3.1658	1.8208	36.6955	12.6731	6.6765	51.8189
3	IHH	20.3966	13.1206	88.1994	36.2054	20.9090	89.7814	2.6766	1.8078	84.8593	9.8989	6.6429	84.6304
4	IHH	17.6340	12.8764	117.1883	30.4077	21.2685	126.3636	2.4967	1.8543	97.9048	9.3532	6.5141	95.2078
5	IHH	116.650	129.9350	2.7106	219.3228	219.5590	2.4522	2.3893	1.7592	106.5202	8.8186	6.1726	106.6455
6	IHH	21.3602	12.7882	80.5101	37.6139	21.8833	83.3126	2.4941	1.8351	47.7284	9.1459	6.6685	99.1293

estimator on the % RMSE criterion. On the % QD criterion the ISALS is more efficient than the OLS estimator. However, cost of computation far outweighs the gain in efficiency in iterating the SALS estimators on all criteria. But, the gain in efficiency still leaves the ITWLS and the IHH estimators less efficient than the OLS estimator on the % RMSE criterion. On the % QD criterion, the ITWLS (I=6) is more efficient than the ITWLS (I=4), the IHH (I=2) estimators.

In Set II, there is no gain in efficiency in iterating the TWLS estimator on the % RMSE criterion. It may be seen that the ITWLS (I=2) estimator does badly on the % RMSE criterion while it performs well on the % E criterion. This is due to the large bias in the ITWLS (I=2) estimator.[37] The IHH estimator perform badly in comparison to the HH estimator on all criteria.

In conclusion, iterating the SALS, the TWLS and the HH estimators may improve the efficiency of the estimator, but the gain in efficiency does not affect the ranking relation among alternative estimators. The gain in efficiency in some cases is often marginal, and hence we cannot justify iterating the alternative estimators when the increase in computing cost is taken into consideration.

---

[37]We plotted a histrogram for the ITWLS (I=2) estimator. The histrogram had large numbers of observations in one string while a few observations appeared as outliars.

V.2.2. Efficiency Analysis of Alternative Iterative Estimators of Variances of Random Coefficients in the CMR Models for Sample Size 20.

The results of sampling experiments on alternative iterative estimators of variances in the CMR Models are given in Table 25 to Table 27.

Efficiency Analysis of Alternative Iterative Estimators of Variances of Random Coefficients in the Proportional CMR Model I for Sample Size 20.

The % RMSE, the % QD and the % E statistics have been computed for each of the three iterations in the proportional CMR Model I and are given in Table 25. Both in

Table 25 Percentage Root Mean Square Error, Percentage Quartile Deviation and Percentage Efficiency for each of the three Iterative of Alternative Iterative Estimator of the Variance of Random Coefficients in the CMR Model I for Sample Size 20

Iteration	Estimator	MODEL I — SET I			MODEL I — SET II		
		Coefficient: $\sigma_1^2 = 0.03$			Coefficient: $\sigma_1^2 = 0.5$		
		% RMSE	% QD	% E	% RMSE	% QD	% E
I	SALS	30.5996	21.8376	100.0000	30.5996	21.8376	100.0000
II	I SALS	29.2217	18.8600	109.1700	29.2217	18.8660	109.1800
III	I SALS	29.2217	18.8600	109.1700	29.2217	18.8660	109.1800
I	TWLS	30.4855	21.5068	100.0000	30.4855	21.5068	100.0000
II	I TWLS	29.3468	19.4001	107.4600	29.3468	19.4001	107.4000
III	I TWLS	29.3468	19.4001	107.4600	29.3468	19.400ʻ	107.4600
I	HH	45.5493	31.0804	100.0000	45.5493	31.0804	100.0000
II	I HH	47.7797	33.8213	93.2600	47.7797	33.8231	93.2600
III	I HH	47.7797	33.8231	93.2600	47.7797	33.8213	93.260000

Set I and Set II there is a gain in efficiency in iterating
the SALS and the TWLS estimators of variances but there is
a loss in efficiency in iterating the HH estimator.   In the
SALS and the TWLS all possible gain in efficiency is
realized at the second iteration.   There is a systematic
trend for convergence in all of the three estimators.   The
results in the second and third iterations are identical for
each of the ISALS, the ITWLS and the IHH estimators on all
criteria.

Table 26 . Percentage Negative Estimates in Alternative Estimators of the
Variances of Random Coefficients in the CMR Model I for
Sample Size 20

Iteration	Estimator	MODEL II — SET I		MODEL II — SET II	
		Coefficient: $\sigma_1^2 = 0.04$	Coefficient: $\sigma_2^2 = 0.03$	Coefficient: $\sigma_1^2 = 1.5$	Coefficient: $\sigma_2^2 = 0.5$
		%-ve $\hat{\sigma}_1^2$	%-ve $\hat{\sigma}_1^2$	%-ve $\hat{\sigma}_1^2$	%-ve $\hat{\sigma}_2^2$
1	SALS	3	16	0	30
2	I SALS	3	16	0	30
3	I SALS	5	18	0	30
4	I SALS	3	17	0	30
5	I SALS	4	16	0	30
6	I SALS	3	16	0	30
1	TWLS	1	17	0	29
2	I TWLS	3	17	0	31
3	I TWLS	2	17	0	30
4	I TWLS	3	16	0	30
5	I TWLS	2	18	1	29
6	I TWLS	2	16	0	30
1	HH	7	15	0	30
2	I HH	10	15	0	30
3	I HH	9	15	2	30
4	I HH	8	15	3	30
5	I HH	10	15	4	30
6	I HH	10	15	2	29

Efficiency as Measured by the Percentage Negative Estimates of Variances of Random Coefficients in the Bivariate Model II for Sample Size 20.

The results of the sampling experiments are summarized on the iterative estimators of variances in Table 26. There is no gain in efficiency in iterating the SALS, the TWLS and the HH estimators, on the criterion of overall percentage negative estimates of variance.

Thus, on the basis of sampling properties of the ISALS, the ITWLS and the IHH estimators of variances in the CMR Model II, the iterative estimators are not recommended.

Efficiency as Measured by the Percentage Negative Estimates of Variance of Random Coefficients in the Trivariate Model III for Sample Size 20.

The percentage negative estimates of variance (% -ve $\hat{o}_1^2$ or % -ve $\hat{o}_2^2$) are tabulated for the alternative iterative estimators in the Bivariate CMR Model III in Table 27. On the criterion of overall percentage negative estimates of variances, the ITWLS and the IHH estimators in Set I are more efficient at one or more stages of iterations than the TWLS and the HH estimators respectively. Similarly, in Set II, the HH, the TWLS and the SALS estimators do better when iterated. But the decrease in percentage negative estimates of variance in all cases is less than two percent. In many situations, this small gain in efficiency may not be justifiable in terms of increased cost of iterations. It may be noted that in Set I, the overall percentage

Table 27   Percentage Negative Estimates in Alternative Iterative Estimates
of the Variances of Random Coefficients in the CMR Model
III for Sample Size 20

| Iteration | Estimator | MODEL III — SET I | | MODEL III — SET II | |
| | | Coefficient: $\sigma_1^2 = 0.04$ | Coefficient: $\sigma_2^2 = 0.03$ | Coefficient: $\sigma_1^2 = 1.5$ | Coefficient: $\sigma_2^2 = 0.5$ |
		%-ve $\hat{\sigma}_1^2$	%-ve $\hat{\sigma}_2^2$	%-ve $\hat{\sigma}_1^2$	%-ve $\hat{\sigma}_2^2$
1	SALS	1	26	0	39
2	I SALS	3	26	0	38
3	I SALS	4	25	0	38
4	I SALS	5	25	0	37
5	I SALS	5	25	0	37
6	I SALS	5	25	0	37
1	TWLS	2	27	0	40
2	I TWLS	3	26	0	38
3	I TWLS	3	26	0	38
4	I TWLS	2	27	0	38
5	I TWLS	2	26	0	38
6	I TWLS	3	26	0	38
1	HH	7	23	5	39
2	I HH	9	20	5	38
3	I HH	11	21	6	38
4	I HH	10	21	6	38
5	I HH	11	21	6	38
6	I HH	9	21	5	38

negative estimates criterion, the ITWLS (I=5) and the IHH
(I=2) estimators have least number of negative estimates
among six iterations and yet each rank lower than in the SALS
estimator.

In Set II the ISALS (I=4) is more efficient than
the ITWLS (I=2) estimator and both estimators are more
efficient than the IHH estimator.

V.2.3   Convergence of Alternative Iterative Estimators
         of Mean Response Coefficients in the CMR Models.

The test of convergence was described in Section

V.2. We may choose the value of δ equal to 0.0001.[38]   The

results of the convergence test on alternative estimators

in the three models for sample Size 20 are described below:

Convergence of Alternative Iterative Estimators of the
Mean Response Coefficients in the CMR Model I for Sample
Size 20.

Alternative iterative estimators in the proportional

CMR Model I converge (both in Set I and Set II) at the first

iteration in 100 percent cases.[39]

Convergence of Alternative Iterative Estimators of the
Mean Response Coefficients in the CMR Model II for
Sample Size 20.

The convergence test was run at the fifth iteration

for alternative iterative estimators in the CMR Model II

for sample Size 20. Convergence was achieved in 80 to 90

replications out of 100. Percentage convergence in the ITWLS

and the ISALS estimators was always higher than in the IHH

estimator. The ISALS and ITWLS estimators performed equal in

terms of percentage convergence.

---

[38]The value of δ = 0.0001 seems appropriate for most
economic experiments. In sample cases another value of δ
(i.e. δ = 0.00001) was tried and convergence was achieved in
70 to 80 replications out of 100.

[39]This result is true for all values of δ up to 7
decimal points.

Convergence of Alternative Iterative Estimators of the
Mean Response Coefficients in the CMR Model III for
Sample Size 20.

The percentage convergences in the two sets in the

CMR Model III are similar to the CMR Model II.

In conclusion, the alternative estimators of mean

response coefficients generally converge within five itera-

tions.[40]

---

[40]In Tables 23 and 24 the iterations seem to get
worse and then get better and then sometimes get worse
again. In short there seems to be a cycling process. This
cycling process in efficiency arises for the non covergence
of 10 to 20 percent cases. A more stringent test for con-
vergence would be to verify

$$| \hat{\bar{\beta}}_j (t + 2) - \hat{\bar{\beta}}_j (t + 1)| \leq | \hat{\bar{\beta}}_j (t + 1) - \hat{\bar{\beta}}_j(t) |$$

for all t after say t = 5. This was brought to my notice by
Professor T. M. Brown. Unfortunately, we did not apply this test.
We therefore caution the reader about the conclusions re-
garding convergence which hold only within the domain of the
test of convergence used for the present study.

CHAPTER VI

STEPWISE LEAST SQUARES APPROACH TO ESTIMATING PARAMETERS IN THE REGRESSION MODEL WITH RANDOM COEFFICIENTS ANALOGOUS TO THE MAXIMUM LIKELIHOOD ESTIMATOR[1]

VI.1. INTRODUCTION

In Chapter III we discussed several methods of estimating regression models with random coefficients. All these methods provide estimates of the variances and the mean response coefficients. However, there is the difficulty common to all the methods that the estimates of variances may sometimes turn out to be negative. Some ad hoc methods of overcoming this shortcoming were also pointed out. For example, it has been suggested that negative estimates be replaced by zeros. Hildreth and Houck (1968) have shown that the restricted estimator will be biased but will have lower mean square error than the corresponding estimator without zero restrictions.[2] Alternatively it has been suggested that a programming approach be applied to the

---

[1] The idea emerged in the course of discussion with Professor A.L. Nagar on the problem of negative estimates of variances. I am thankful for his guidance in preparation of this Chapter.

[2] This procedure was applied in a few sample runs and the estimates of the mean response coefficients were obtained. These estimates were not always meaningful. We did not, however, carry out extensive test of this procedure.

problem using the restriction that $\sigma_j^2 \geq 0$, for all j. This approach has the shortcoming that the precision of the estimator cannot be measured. Further, we argued that in seeking more efficient estimators we reduce the probability of getting negative estimate of variance and we discussed a few efficient estimators of variances. However, as has been said earlier, we cannot eliminate the possibility of getting negative estimates through this approach. In the present Chapter we will propose a 'step-wise least squares approach' to the problem, an approach which does in fact produce a guaranteed positive estimate of variances.

VI.2. MAXIMUM LIKELIHOOD ESTIMATORS OF THE MEAN RESPONSE COEFFICIENT AND THE VARIANCE IN THE BI-VARIATE RE-GRESSION MODEL WITH A RANDOM COEFFICIENT

The bivariate regression model with a random coefficient has some very interesting properties. The model gives a guaranteed positive estimate of the variance by all methods discussed in Chapter III. Further, the problem of non-linearity in the parameters of the normal equations does not arise in this special case.[3]

We may write the bivariate regression model with random coefficients, on the lines of model (II.15) as

---

[3]In this section we have included the discussion on the MLE only because other estimators were found not suit-able for the solution to the problem of negative variances as discussed in this Chapter.

$$y(t) = \beta_1(t) \, x_1(t) \quad , \, t = 1, \, 2, \, \ldots, \, T \, . \, . \quad (VI.1)$$

where $y(t)$ is the t-th observation on the dependent variable and $x_1(t)$ is the t-th observation on the explanatory variable. $\beta_1(t)$ is a random regression coefficient such that[4]

$$\beta_1(t) = \bar{\beta}_1 + \epsilon_1(t) \quad \ldots \ldots \ldots \quad (VI.2)$$

where the assumptions (II.7) - (II.9) hold, for $j = 1$. Substituting (VI.2) into (VI.1), we have

$$y(t) = \bar{\beta}_1 \, x_1(t) + \eta_1(t) \quad , \, t = 1, \, 2, \, \ldots, \, T \quad (VI.3)$$

where

$$\eta_1(t) = \epsilon_1(t) \, x_1(t) \quad , \, t = 1, \, 2, \, \ldots, \, T \quad (VI.4)$$

such that

$$E\eta_1(t) = o \text{ for all } t, \text{ and Var } \eta_1(t) = \phi_1^2(t) = \sigma_1^2 x_1^2(t),$$

$$\text{for } t = 1, \, 2, \, \ldots, \, T \quad \ldots \quad (VI.5)$$

Under the assumption that the $\eta_1(t)$'s are independently distributed with the means and variances given in (VI.5), the joint distribution of $\eta_1(1)$, $\eta_1(2)$, $\ldots$, $\eta_1(T)$ is:

$$\mathcal{L} = \frac{1}{(2\pi)^{T/2} \, \prod_t \phi_1(t)} \, e^{-1/2 \sum_{t=1}^{T} \frac{\eta_1^2(t)}{\phi_1^2(t)}} \quad \ldots \ldots \quad (VI.6)$$

_____

[4] Once, again, we will restrict the discussion to the CMR model only without any loss of generality.

where $\phi_1(t) = \sqrt{\phi_1^2(t)}$, for $t = 1, 2, \ldots, T$.

Further, the log likelihood function in the parameters $\bar{\beta}_1$ and σ_1^2 can be written as

$$L(\bar{\beta}_1, \sigma_1^2 \mid y, x_1) = -\frac{T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \log x_1^2(t)\, \sigma_1^2$$

$$-\frac{1}{2} \sum_{t=1}^{T} \frac{[y(t) - \bar{\beta}_1 x_1(t)]^2}{x_1^2(t)\, \sigma_1^2} \quad \ldots \ldots \quad (VI.7)$$

because the Jacobian of the transformation $|J| = 1$

or

$$L = -\frac{T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \log x_1^2(t) - \frac{1}{2} \sum_{t=1}^{T} \log \sigma_1^2$$

$$-\frac{1}{2} \sum_{t=1}^{T} \frac{[y(t) - \bar{\beta}_1 x_1(t)]^2}{x_1^2(t)\, \sigma_1^2} \quad \ldots \ldots \quad (VI.8)$$

The normal equations for parameter estimator $\tilde{\beta}_1$ may then be obtained as

$$\frac{\partial L}{\partial \tilde{\beta}_1} = -\frac{1}{2} \sum_{t=1}^{T} \frac{[y(t) - \tilde{\beta}_1 x_1(t)].\, 2.(-x_1(t)}{x_1^2(t)\, \tilde{\sigma}_1^2} = 0 \quad \ldots \quad (VI.9)$$

or

$$\frac{1}{\tilde{\sigma}_1^2} \sum_{t=1}^{T} \left(\frac{y(t) - \tilde{\tilde{\beta}}_1 x_1(t)}{x_1(t)} \right) = 0 \quad \ldots \ldots \quad (VI.10)$$

or

$$\tilde{\beta}_1 = \frac{1}{T} \sum_{t=1}^{T} \frac{y(t)}{x_1(t)} \quad \ldots \ldots \ldots \ldots \quad (VI.11)$$

because $\sigma_1{}^2$ is finite by assumption.

It is clear from (VI.11) that $\tilde{\beta}_1$ is independent of the estimate of variance for the simple proportional CMR model (VI.1).[5]

Similarly the normal equation for $\tilde{\sigma}_1{}^2$ is given as

$$\frac{\partial L}{\partial \tilde{\sigma}_1{}^2} = \frac{-T}{2\tilde{\sigma}_1{}^2} - \frac{1}{2} \sum_{t=1}^{T} \frac{[y(t) - \tilde{\beta}_1 x_1(t)]^2}{x_1{}^2(t)} \quad (-\frac{1}{\tilde{\sigma}_1{}^4}) = 0 \quad \text{(VI.12)}$$

or

$$\frac{1}{2\tilde{\sigma}_1{}^2} \left[-T + \frac{1}{\tilde{\sigma}_1{}^2} \sum_{t=1}^{T} \frac{[y(t) - \tilde{\beta}_1 x_1(t)]^2}{x_1{}^2(t)} \right] = 0 \quad \ldots \quad \text{(VI.13)}$$

or

$$\tilde{\sigma}_1{}^2 = \frac{1}{T} \sum_{t=1}^{T} \frac{[y(t) - \tilde{\beta}_1 x_1(t)]^2}{x_1{}^2(t)} \quad \ldots \ldots \ldots \quad \text{(VI.14)}$$

because $\sigma_1{}^2$ is finite by assumption.

---

[5]In fact the estimator (VI.11) is also the Aitken estimator of $\tilde{\beta}_1$ for the model (VI.3). This can be seen easily from the fact that the variances of the disturbances $\eta_1(t)$'s as given in (VI.5) are proportional to $x_1{}^2(t)$'s. Thus the Aitken estimator of $\bar{\beta}_1$ can be obtained by deflating the model (VI.3) by $x_1(t)$'s and then applying the OLS, i.e.

$$y(t)/x_1(t) = \bar{\beta}_1 + \eta_1^* \qquad , t = 1, 2, \ldots, T$$

where $\eta_1^*(t) = \eta_1(t) / x_1(t)$.

It is clear that the OLS estimator of $\bar{\beta}_1$ in above is (VI.11) because $E \eta_1^*(t) = o$ for all t and $Var \ \eta_1^*(t) = \sigma_1^2$ for all t. Now the estimator is independent of value of $\sigma_1{}^2$. Further, the HH, the TWLS and the SWLS estimators of $\bar{\beta}_1$ are identical with (VI.11).

It is clear from (VI.11) and (VI.14) that explicit solution of the normal equations for the simple bivariate model (VI.3) is obtainable and the problem of non-linearity in parameters, as in the multivariate case, does not arise. Further, the estimator (VI.14) of variance is guaranteed to be positive. It is this interesting feature that we will exploit to define an estimator with desirable properties.

VI.3. STEP-WISE LEAST SQUARES ESTIMATOR

The step-wise least squares technique is well known.[6] We will discuss below some aspects of this technique which are relevant for our purpose . Consider a simple model

$$y = x_1 \beta_1 + x_2 \beta_2 + u, \quad \ldots \ldots \ldots \text{(VI.15)}$$

where

$$y = \begin{bmatrix} y(1) \\ \vdots \\ y(T) \end{bmatrix}, \quad x_1 = \begin{bmatrix} x_1(1) \\ \vdots \\ x_1(T) \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_2(1) \\ \vdots \\ x_2(T) \end{bmatrix} \quad \ldots \text{(VI.16)}$$

β_1 and β_2 are scalars. Further

$$u = \begin{bmatrix} u(1) \\ \vdots \\ u(T) \end{bmatrix} \quad \ldots \ldots \ldots \ldots \ldots \text{(VI.17)}$$

---

[6]Goldberger (1961).

such that

$$Eu = o \quad \ldots \ldots \ldots \ldots \ldots \text{(VI.18)}$$

and

$$Euu' = \sigma^2 \, I_{TxT} \quad \ldots \ldots \ldots \ldots \text{(VI.19)}$$

In the first step of the stepwise least squares procedure we regress y on x_1

$$Y = \beta_1 x_1 + u_1 \quad \ldots \ldots \ldots \ldots \text{(VI.20)}$$

It follows that

$$u_1 = x_2 \beta_2 + u \quad \text{and} \quad Eu_1 = x_2 \beta_2 \neq o \quad \ldots \ldots \text{(VI.21)}$$

and therefore, the OLS estimator applied to (VI.20) will yield an estimator

$$\hat{\beta}_1 = (x_1 \, 'x_1)^{-1} \, x_1 \, y \quad \ldots \ldots \ldots \text{(VI.22)}$$

of β_1, which is biased and inconsistent.

The next step requires that we regress the OLS residuals,

$$y^o = y - x_1 \, \hat{\beta}_1 \quad \ldots \ldots \ldots \ldots \text{(VI.23)}$$

on x_2

$$y^o = x_2 \beta_2 + \eta_2 \quad \ldots \ldots \ldots \ldots \text{(VI.24)}$$

where η_2 is the disturbance term defined in (VI.26) below.

Since the true model is given in (VI.15), subtracting $x_1 \hat{\beta}_1$ from both sides of (VI.15), we get

$$y - x_1 \hat{\beta}_1 = x_2 \hat{\beta}_2 + (x_1 \beta_1 + u - x_1 \hat{\beta}_1) \quad \ldots \ldots \text{(VI.25}$$

Therefore,

$$\eta_2 = x_1(\beta_1 - \hat{\beta}_1) + u \quad \ldots \ldots \ldots \text{(VI.26)}$$

The OLS estimator of β_2 in (VI.24) is given as

$$\hat{\beta}_2 = (x'_2 \, x_2)^{-1} \, x'_2 \, y^o \quad \ldots \ldots \ldots \text{(VI.27)}$$

It is easy to see that the estimator $\hat{\beta}_2$ is biased and inconsistent (as $\hat{\beta}_1$ is biased and inconsistent) because

$$E(\eta_2) \neq o \quad \ldots \ldots \ldots \ldots \text{(VI.28)}$$

Further, since

$$Ey^o = E(y - x_1\hat{\beta}_1) \quad \ldots \ldots \ldots \ldots \ldots \text{(VI.29)}$$

$$\text{(because (VI.23) holds)}$$

$$= Ey - x_1 \, E \, \hat{\beta}_1$$

$$= x_1\beta_1 + x_2\beta_2 - x_1(x_1 \, 'x_1)^{-1} \, x_1'(x_1\beta_1 + x_2\beta_2)$$

$$\text{(because (VI.22) and (VI.15) hold)}$$

$$= x_2\beta_2 - x_1(x_1'x_1)^{-1} \, x_1'x_2\beta_2$$

$$= [x_2 - x_1(x_1'x_1)^{-1} \, x_1'x_2]\beta_2$$

or

$$Ey^o = x_2^o \, \beta_2 \quad \ldots \ldots \ldots \ldots \ldots \text{(VI.30)}$$

where

$$x_2^o = x_2 - x_1 \, \hat{P} \quad \ldots \ldots \ldots \ldots \text{(VI.31)}$$

and $\hat{P}$ is the OLS estimator of the auxilliary regression coefficient in $x_2 = Px_1 + \text{disturbance}$, i.e.

$$\hat{P} = (x_1'x_1)^{-1} x_1'x_2 \quad \ldots \ldots \ldots \text{(VI.32)}$$

It follows from (VI.30) that in order to obtain a consistent estimator of β_2 we should regress y^o on x_2^o (rather than y^o on x_2 as in (VI.24)) because the disturbance term in the regression

$$y^o = x_2^o \beta_2 + \text{disturbance}, \quad \ldots \ldots \text{(VI.33)}$$

will have zero expected value. The vector x_2^o may be interpreted as the vector of the observations on the explanatory variable x_2, after it has been corrected for relationship between the two explanatory variables.

The OLS estimator of β_2 in (VI.33) is given as

$$\hat{\beta}_2^o = (x_2^{o'} x_2^o)^{-1} x_2^{o'} y^o \quad \ldots \ldots \text{(VI.34)}$$

The estimator $\hat{\beta}_2^o$ is unbiased and consistent, for

$$E \hat{\beta}_2^o = (x_2^{o'} x_2)^{-1} x_2^{o'} E y^o \quad \ldots \ldots \text{(VI.35)}$$

$$(\because x_2^o \text{ is non stochastic})$$

It follows from (VI.30) then,

$$E \hat{\beta}_2^o = (x_2^{o'} x_2^o)^{-1} x_2^{o'} x_2^o \ \beta_2$$

$$= \beta_2$$

i.e. $\hat{\beta}_2^o$ is an unbiased estimator of β_2. Further

$$\underset{T \to \infty}{\text{Plim}} \ \hat{\beta}_2^o = \underset{T \to \infty}{\text{Plim}} \ [(x_2^{o'} x_2^o)^{-1} (x_2^{o'} y^o)] \quad \ldots \ldots \text{(VI.37)}$$

$$= \Plim_{T \to \infty} \left[\left(\frac{1}{T} x_2^{o'} x_2^o \right)^{-1} \left(\frac{1}{T} x_2^{o'} y^o \right) \right]$$

$$= \Plim_{T \to \infty} \left[\left(\frac{1}{T} x_2^{o'} x_2^o \right)^{-1} \right] \Plim_{T \to \infty} \left[\frac{1}{T} x_2^{o'} y^o \right]$$

Let $\displaystyle \Plim_{T \to \infty} \left(\frac{1}{T} x_2^{o'} x_2^o \right)^{-1} = \sum_{x_2^o x_2^o}^{-1}$ be bounded, then it follows

that $\displaystyle \Plim_{T \to \infty} \hat{\beta}_2^o = \beta_2$

because

$$\Plim_{T \to \infty} \left(\frac{1}{T} x_2^o y^o \right) = \Lmt_{T \to \infty} \frac{1}{T} x_2^{o'} E\, y^o \quad \ldots \ldots \ldots \text{(VI.38)}$$

$$= \Lmt_{T \to \infty} \frac{1}{T} x_2^{o'} \quad (x_2^o \beta_2)$$

$$(\because \text{(VI.30) holds})$$

$$= \Lmt_{T \to \infty} \left(\frac{1}{T} x_2^{o'} x_2^o \right) \beta_2$$

Now, since $\displaystyle \Plim_{T \to \infty} \left(\frac{1}{T} x_2^{o'} x_2^o \right)^{-1} = \sum_{x_2^o x_2^o}^{-1}$ by assumption, it

follows then that $\displaystyle \Plim_{T \to \infty} \left(\frac{1}{T} x_2^o y^o \right) = \Lmt_{T \to \infty} \left(\frac{1}{T} x_2^{o'} x_2^o \right) \beta_2$

$= \sum_{x_2^o x_2^o}$.

Thus, from (VI.37) and (VI.38), we have $\displaystyle \Plim_{T \to \infty} \hat{\beta}_2^o = \beta_2$

i.e. estimator $\hat{\beta}_2^o$ is consistent.

Further since,

$$E\, \hat{\beta}_1 = (x_1'x_1)^{-1} x_1'\, Ey \quad \ldots \ldots \ldots \ldots \text{(VI.40)}$$

$$= (x_1'x_1)^{-1} x_1'\, [x_1 \beta_1 + x_2 \beta_2]$$

$$=\beta_1 + (x_1'x_1)^{-1} x_1'x_2 \, \beta_2$$

$$=\beta_1 + \hat{P} \, \beta_2 \quad [\because (VI.32) \text{ holds}],$$

we may obtain an unbiased and consistent estimator $\hat{\beta}_1^{\,o}$ for

β_1 as

$$\hat{\beta}_1^{\,o} = \hat{\beta}_1 - \hat{P} \, \hat{\beta}_2^{\,o} \quad \dots \dots \dots \dots \quad (VI.41)$$

where $\hat{\beta}_1$, $\hat{P}$ and $\hat{\beta}_2^{\,o}$ are defined in (VI.22), (VI.32) and

(VI.34) respectively.[7]

VI.4. STEP-WISE ESTIMATOR OF THE MEANS AND VARIANCES IN THE REGRESSION MODEL WITH RANDOM COEFFICIENTS ANALO-GOUS TO THE MAXIMUM LIKELIHOOD ESTIMATOR

In this section we shall investigate the possibility

of applying the maximum likelihood procedure in step-wise

fashion to a linear regression model with random coefficients.

Consider a simple regression model with random

coefficients

$$y(t) = \beta_1(t) \, x_1(t) + \beta_2(t) \, x_2(t) \quad , \, t = 1, \, 2, \, \dots \, ., \, T$$

$$(VI.42)$$

such that

$$\beta_j(t) = \bar{\beta}_j + \varepsilon_j(t) \quad , \, t = 1, \, 2, \, \dots \, ., \, T; \quad j = 1, \, 2 \quad (VI.43)$$

$$E \, \beta_j(t) = \bar{\beta}_j \quad , \, \text{for all } t; \quad j = 1, \, 2, \quad (VI.44)$$

---

[7]Goldberger (1964) has shown that $\hat{\beta}_1^{\,o}$ and $\hat{\beta}_2^{\,o}$ are in fact the true OLS estimators of β_1 and β_2 in combined (True) model (VI.15).

$$\text{Var } \beta_j(t) = \sigma_j{}^2 \quad , \; j = 1, \; 2, \qquad \ldots \ldots \ldots \ldots \qquad \text{(VI.45)}$$

$$E \; (\beta_j(t) - \bar{\beta}_j) \; (\beta_j{}'(t') - \bar{\beta}_j{}') = 0 \quad \ldots \ldots \ldots \ldots \qquad \text{(VI.46)}$$

$$\text{if } j \neq j' \text{ or } t \neq t'$$

Substituting (VI.43) into (VI.42), we have

$$y(t) = \bar{\beta}_1 \; x_1(t) + \bar{\beta}_2 \; x_2(t) + \eta(t), \quad t = 1, \; 2, \; \ldots ., \; T \quad \text{(VI.47)}$$

where

$$\eta(t) = \varepsilon_1(t) \; x_1(t) + \varepsilon_2(t) \; x_2(t) \quad , \; t = 1, \; 2, \; \ldots ., \; T \quad \text{(VI.48)}$$

such that

$$E\eta(t) = 0 \quad , \text{ for all } t \; \ldots \ldots \; \text{(VI.49)}$$

and

$$\text{Var } (\eta(t)) = \phi^2(t) = x_1{}^2(t) \; \sigma_1{}^2 + x_2{}^2(t) \; \sigma_2{}^2, \; \ldots \; \text{(VI.50)}$$

Now suppose, in the first step of the step-wise procedure, we regress y (t) on $x_1(t)$ as

$$y(t) = \bar{\beta}_1 \; x_1(t) + \eta_1(t) \quad , \; t = 1, \; 2, \; \ldots ., \; T \quad \ldots \; \text{(VI.51)}$$

where

$$\eta_1(t) = \bar{\beta}_2 \; x_2(t) + \eta(t) \quad , \; t = 1, \; 2, \; \ldots ., \; T \quad \ldots \; \text{(VI.52)}$$

because (VI.42) is the true model. It follows that $E \; \eta_1(t)$ $\neq$ o. However, analogous to the MLE (VI.11) and (VI.14) we may define estimator for $\bar{\beta}_1$ and $\sigma_1{}^2$ as

$$\tilde{\bar{\beta}}_1 = \frac{1}{T} \sum_{t=1}^{T} y(t)/x_1(t) \quad \ldots \ldots \ldots \ldots \; \text{(VI.53}$$

$$\tilde{\sigma}_1{}^2 = \frac{1}{T} \sum_{t=1}^{T} \frac{(y(t) - \tilde{\bar{\beta}}_1 \; x_1(t))^2}{x_1{}^2(t)} \quad \ldots \ldots \ldots \; \text{(VI.54)}$$

It is easy to establish that $\tilde{\tilde{\beta}}_1$ and $\tilde{\sigma}_1{}^2$ defined in (VI.53) and (VI.54) are biased and inconsistent estimators of $\bar{\beta}_1$ and $\sigma_1{}^2$ respectively because of the specification bias in regression (VI.51). In particular, the bias of $\tilde{\tilde{\beta}}_1$ is given by

$$E \,\tilde{\tilde{\beta}}_1 = \frac{1}{T} \sum_{t=1}^{T} \frac{E\, y(t)}{x_1(t)} \qquad \ldots \ldots \ldots \ldots \ldots \quad (VI.55)$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{\bar{\beta}_1\, x_1(t) + \bar{\beta}_2\, x_2(t)}{x_1(t)}$$

$$= \bar{\beta}_1 + \beta_2 \,\frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)}$$

The second step in the step-wise approach is to regress the residuals from the regression of y on x_1, i.e.

$$y^o(t) = y(t) - \tilde{\tilde{\beta}}_1\, x_1(t) \qquad , \quad t = 1, 2, \ldots, T, \qquad (VI.56)$$

on x_2

$$y^o(t) = \bar{\beta}_2\, x_2(t) + \eta_2(t) \qquad \ldots \ldots \ldots \ldots \ldots \quad (VI.57)$$

where

$$\eta_2(t) = (\bar{\beta}_1 - \tilde{\tilde{\beta}}_1)\, x_1(t) + \eta(t) \ldots \ldots \ldots \ldots \quad (VI.58)$$

(on the lines of (VI.25) and (VI.26).)

Again, we may define estimators of $\bar{\beta}_2$ and $\sigma_2{}^2$ analogous to the MLE estimators (VI.11) and (VI.14)

$$\tilde{\tilde{\beta}}_2 = \frac{1}{T} \sum_{t=1}^{T} \frac{y^o(t)}{x_2(t)} \qquad \ldots \ldots \ldots \ldots \ldots \quad (VI.59)$$

$$\tilde{\sigma}_2^{\,2} = \frac{1}{T} \sum_{t=1}^{T} \frac{[y^O(t) - \tilde{\bar{\beta}}_2 \, x_2(t)]^2}{x_2^{\,2}(t)} \qquad \ldots \ldots \text{(VI.60)}$$

It is clear from (VI.58) that

$$E\eta_2(t) = E(\bar{\beta}_1 - \tilde{\bar{\beta}}_1) \, x_1(t) + E\eta(t) \ldots \ldots \ldots \ldots \text{(VI.61)}$$

$$= - \left(\bar{\beta}_2 \, \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right) x_1(t) \neq 0, \quad t = 1, 2, \ldots, T,$$

because (VI.49) and (VI.55) hold. Therefore, the estimators $\tilde{\bar{\beta}}_2$ and $\tilde{\sigma}_2^{\,2}$ defined in (VI.59) and (VI.60) are biased and inconsistent.

In particular, the bias of $\tilde{\bar{\beta}}_2$ is given by

$$E \, \tilde{\bar{\beta}}_2 = \frac{1}{T} \sum_{t=1}^{T} \frac{E \, [y^O(t)]}{x_2(t)} \qquad \ldots \ldots \ldots \ldots \ldots \text{(VI.62)}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{E[y(t) - \tilde{\bar{\beta}}_1 \, x_1(t)]}{x_2(t)}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{E \, y(t) - E \, \tilde{\bar{\beta}}_1 \, x_1(t)}{x_2(t)}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{\left[\bar{\beta}_1 \, x_1(t) + \bar{\beta}_2 \, x_2(t) \right] - \left[\bar{\beta}_1 + \bar{\beta}_2 \, \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right] x_1(t)}{x_2(t)}$$

$$(\because \text{(VI.55) holds})$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{\bar{\beta}_2 \left[x_2(t) - x_1(t) \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right]}{x_2(t)}$$

$$= \bar{\beta}_2 - \bar{\beta}_2 \left[\frac{1}{T} \sum_{t=1}^{T} \frac{x_1(t)}{x_2(t)} \right] \left[\frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right]$$

It follows from (VI.61) that the regression relation (VI.57) has specification bias which may be corrected as

$$y^o(t) = \bar{\beta}_2 \left[x_2(t) - \left(\frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right) x_1(t) \right] \quad \ldots \ldots \text{(VI.63)}$$

$$+ \left[\eta_2(t) + \left(\bar{\beta}_2 \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right) x_1(t) \right]$$

$$= \bar{\beta}_2 \left[x_2(t) - x_1(t) \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right]$$

$$+ \left[\eta_2(t) + \bar{\beta}_2 x_1(t) \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right]$$

or

$$y^o(t) = \bar{\beta}_2 x_2^o(t) + \eta_2^o(t) \quad , \ t = 1, 2, \ldots, T \quad \text{(VI.64)}$$

where

$$x_2^o(t) = x_2(t) - x_1(t) \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \ , \ t = 1, 2, \ldots, T \quad \text{(VI.65)}$$

and

$$\eta_2^o(t) = \eta_2(t) + \bar{\beta}_2 x_1(t) \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)}, \ t = 1, 2, \ldots, T.$$

$$\ldots \ \text{(VI.66)}$$

It follows from (VI.61) that

$$E \; \eta_2^{\,o}(t) = o \quad , \text{ for all } t \; . \; . \; . \; . \; . \quad (VI.67)$$

Thus if we define the estimator $\tilde{\beta}_2^{\,o}$ in regression relation (VI.64) analogous to (VI.53) as

$$\tilde{\beta}_2^{\,o} = \frac{1}{T} \sum_{t=1}^{T} \frac{y^o(t)}{x_2^{\,o}(t)} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (VI.68)$$

then $\tilde{\beta}_2$ will be unbiased and consistent because the disturbances $\eta_2^{\,o}(t)$'s have expected value zero. Viz.

$$E \; \tilde{\beta}_2^{\,o} = \frac{1}{T} \sum_{t=1}^{T} \frac{E \; y^o(t)}{x_2^{\,o}(t)} \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (VI.69)$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{E[y(t) - \tilde{\beta}_1 \; x_1(t)]}{x_2^{\,o}(t)}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{\bar{\beta}_1 \; x_1(t) + \bar{\beta}_2 \; x_2(t) - \left[\bar{\beta}_1 + \frac{1}{T} \bar{\beta}_2 \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right] x_1(t)}{x_2^{\,o}(t)}$$

[because (VI.55) holds]

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{\bar{\beta}_2 \left[x_2(t) - x_1(t) \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right]}{x_2^{\,o}(t)}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \frac{\bar{\beta}_2 \; x_2^{\,o}(t)}{x_2^{\,o}(t)}$$

$$= \bar{\beta}_2$$

Similarly, it can be shown that $\tilde{\beta}_2^{\,o}$ is consistent.

In (VI.65) $x_2^O(t)$ is defined as

$$x_2^O(t) = x_2(t) - x_1(t) \; \tilde{\tilde{P}} \quad \ldots \ldots \quad \text{(VI.70)}$$

where

$$\tilde{\tilde{P}} = \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \quad \ldots \ldots \quad \text{(VI.71)}$$

i.e. $\tilde{\tilde{P}}$ is the maximum likelihood estimator of the mean response $\bar{P}$ of the random coefficient $P(t)$ analogous to the MLE estimator (VI.11) in the auxilliary regression.

$$x_2(t) = P(t) \; x_1(t) \quad , \; t = 1, \; 2, \; \ldots, \; T \quad \text{(VI.72)}$$

where

$$P(t) = \bar{P} + \xi(t) \quad \ldots \ldots \quad \text{(VI.73)}$$

such that

$$E \; \xi(t) = o \quad , \text{ for all } t \; ; \text{ Var } \; \xi(t) = \sigma_\xi^2 \text{ and} \quad \text{(VI.74)}$$

$$\text{Cov } (\xi(t) \quad \xi(t')) = o, \text{ if } t \neq t'$$

In other words $x_2^O(t)$ is the corrected $x_2(t)$ after adjusting for the regression relationship between two explanatory variables with random coefficient defined in (VI.72).

We may, then, define the estimator $\tilde{\sigma}_2^{2O}$ for the model (VI.64) analogous to (VI.14) as

$$\tilde{\sigma}_2^{2O} = \frac{1}{T} \sum_{t=1}^{T} \frac{[y^O(t) - \tilde{\tilde{\beta}}_2^O \; x_2^O(t)]^2}{x_2^{2O}(t)} \quad \ldots \quad \text{(VI.75)}$$

In order to show that the estimator $\tilde{\sigma}_2^{O^2}$ is consis-

tent, we must prove that

$$\underset{T\to\infty}{\text{Plim}}\ \tilde{\sigma}_2{}^{o2} = \sigma_2{}^2 \quad \ldots \ldots \ldots \ldots \quad \text{(VI.76)}$$

Taking the probability limit of $\tilde{\sigma}_2{}^{o2}$ in (VI.75) we have

$$\underset{T\to\infty}{\text{Plim}}\ \tilde{\sigma}_2^2{}^o = \underset{T\to\infty}{\text{Plim}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{[y^o(t) - x_2{}^o(t)\ \tilde{\tilde{\beta}}_2{}^o]^2}{x_2^{2o}(t)} \quad \ldots\ \text{(VI.77)}$$

$$= \underset{T\to\infty}{\text{Plim}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{[y^o(t) - x_2{}^o(t)\ \bar{\beta}_2]^2}{x_2^{2o}(t)}$$

because $\qquad \underset{T\to\infty}{\text{Plim}}\ \tilde{\tilde{\beta}}_2{}^o = \bar{\beta}_2$

Furthermore,

$$\underset{T\to\infty}{\text{Plim}}\ \tilde{\sigma}_2^2 = \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{E[y^o(t) - x_2{}^o(t)\ \bar{\beta}_2]^2}{x_2^{2o}(t)} \quad \ldots \ldots \quad \text{(VI.78)}$$

$$= \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{\text{Var } y^o(t)}{x_2^{2o}(t)}$$

But

$$y^o(t) = y(t) - \frac{1}{T}x_1(t)\sum_{t=1}^{T}\frac{y(t)}{x_1(t)} \quad \ldots \ldots \ldots \quad \text{(VI.79)}$$

$$[\because (\text{VI.56})\text{ holds}]$$

$$= [x_1(t)\ \bar{\beta}_1 + x_2(t)\ \bar{\beta}_2 + \eta(t)] - \frac{1}{T}x_1(t)$$

$$\sum_{t=1}^{T}\frac{[x_1(t)\ \bar{\beta}_1 + x_2(t)\ \bar{\beta}_2 + \eta(t)]}{x_1(t)}$$

$$= x_1(t) \ \bar{\beta}_1 + x_2(t) \ \bar{\beta}_2 + \eta(t) - x_1(t) \ \bar{\beta}_1$$

$$- \frac{1}{T} \ \bar{\beta}_2 \ x_1(t) \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} - \frac{1}{T} \ x_1(t) \sum_{t=1}^{T} \frac{\eta(t)}{x_1(t)}$$

$$= \left[x_2(t) - x_1(t) \ \frac{1}{T} \sum_{t=1}^{T} \frac{x_2(t)}{x_1(t)} \right] \bar{\beta}_2$$

$$+ \left[\eta(t) - x_1(t) \ \frac{1}{T} \sum_{t=1}^{T} \frac{\eta(t)}{x_1(t)} \right]$$

$$= x_2^{\,O}(t) \ \bar{\beta}_2 + \left[\eta(t) - x_1(t) \ \frac{1}{T} \sum_{t=1}^{T} \frac{\eta(t)}{x_1(t)} \right]$$

Therefore

$$\text{Var } y^O(t) = \text{Var } (x_2^{\,O}(t) \ \bar{\beta}_2) + \text{Var } \left[\eta(t) - x_1(t) \right.$$

$$\left. \frac{1}{T} \sum_{t=1}^{T} \frac{\eta(t)}{x_1(t)} \right]$$

$$+ 2 \ \text{Cov } \left[x_2^{\,O}(t) \ \bar{\beta}_2, \ \eta(t) - x_1(t) \right.$$

$$\left. \frac{1}{T} \sum_{t=1}^{T} \frac{\eta(t)}{x_1(t)} \right] \qquad \cdots \cdots \text{(VI.80)}$$

or

$$\text{Var } y^O(t) = \text{Var } \left[\eta(t) - x_1(t) \ \frac{1}{T} \sum_{t=1}^{T} \frac{\eta(t)}{x_1(t)} \right] \qquad \cdots \cdots \text{(VI.81)}$$

because Var $x_2^o(t)$ $\bar{\beta}_2$ = 0 and

$$\text{Cov}\left[x_2^o(t)\ \bar{\beta}_2\quad,\ \eta(t)\ -\ x_1(t)\ \frac{1}{T}\sum_{t=1}^{T}\frac{\eta(t)}{x_1(t)}\right]\ =\ 0$$

Further, (VI.81) can be written as

$$\text{Var}\ y^o(t)\ =\ \text{Var}\left[\eta(t)\ -\ x_1(t)\left(\frac{1}{T}\frac{\eta(1)}{x_1(t)}\ +\ \ldots\ +\ \frac{\eta(t)}{x_1(t)}\right.\right.$$

$$\left.\left.+\ \frac{\eta(T)}{x_1(T)}\right)\right] \tag{VI.82}$$

$$=\ \text{Var}\left[-\ x_1(t)\ \frac{1}{T}\frac{\eta(1)}{x_1(1)}-\ \ldots\ -\left(x_1(t)\ \frac{1}{T}\ \frac{1}{x_1(t)}-1\right)\ \eta(t)-\right.$$

$$\left.\ldots\ -\ x_1(t)\ \frac{1}{T}\frac{\eta(T)}{x_1(T)}\ \right]$$

$$=\left[x_1^2(t)\ \frac{1}{T^2}\ \frac{\phi^2(1)}{x_1^2(1)}\ +\ \ldots\ +\ (x_1(t)\ \frac{1}{T}\ \frac{1}{x_1(t)}\ -\ 1)^2\right.$$

$$\left.\phi^2(t)\ +\ \ldots\ +\ x_1^2(t)\ \frac{1}{T^2}\frac{\phi^2(T)}{x_1^2(T)}\right]$$

because Cov $[\eta(t),\ \eta(t')] = o$, if $t \neq t'$

and   Var $\eta(t) = \phi^2(t)$, $t = 1, 2, \ldots, T$.

Rearranging the terms in (VI.82), we get

$$\text{Var}\ y^o(t)\ =\ x_1^2(t)\ \frac{1}{T^2}\sum_{t=1}^{T}\ \frac{\phi^2(t)}{x_1^2(t)}\ +\ \left(\frac{T\ -\ 2}{T}\right)\ \phi^2(t) \tag{VI.83}$$

Substituting the value of var $[y^o(t)]$ in (VI.78), we get

$$\underset{T\to\infty}{\text{Plim}}\ \tilde{\sigma}_2^2{}^o = \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{\left[x_1^2(t)\ \frac{1}{T^2}\sum_{t=1}^{T}\frac{\phi^2(t)}{x_1^2(t)}\ +(\frac{T-2}{T})\phi^2(t)\right]}{x_2^{2^o}(t)} \qquad (VI.84)$$

$$= \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\phi^2(t)}{x_1^2(t)}\ \frac{1}{T}\sum_{t=1}^{T}\frac{x_1^2(t)}{x_2^{2o}(t)}\right.$$

$$\left. + (\frac{T-2}{T})\sum_{t=1}^{T}\frac{\phi^2(t)}{x_2^{2o}(t)}\right]$$

$$= \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\phi^2(t)}{x_1^2(t)}\ \frac{1}{T}\sum_{t=1}^{T}\frac{x_1^2(t)}{x_2^{2o}(t)}\right]$$

$$+\ \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{\phi^2(t)}{x_2^{2o}(t)}$$

$$-\ \underset{T\to\infty}{\text{Lmt}}\ \frac{2}{T}\ \frac{1}{T}\sum_{t=1}^{T}\frac{\phi^2(t)}{x_2^{2o}(t)}$$

$$= \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{\phi^2(t)}{x_2^{2o}(t)} \qquad \text{(because other terms approach}$$
$$\text{zero limit)}$$

$$= \underset{T\to\infty}{\text{Lmt}}\ \frac{1}{T}\sum_{t=1}^{T}\frac{x_1^2(t)\ \sigma_1^2 + x_2^2(t)\ \sigma_2^2}{x_2^{2o}(t)}$$

Therefore,

$$\underset{T\to\infty}{\text{Plim}}\ \tilde{\sigma}_2^2{}^o = \sigma_1^2\ \underset{T\to\infty}{\text{Lmt}}\ \left[\frac{1}{T}\sum_{t=1}^{T}\frac{x_1^2(t)}{x_2^{2o}(t)}\right]\ +\ \sigma_2^2\ \underset{T\to\infty}{\text{Lmt}} \qquad (VI.85)$$

$$\left[\frac{1}{T}\sum_{t=1}^{T}\frac{x_2^2(t)}{x_2^{2o}(t)}\right]\ \neq\ \sigma_2^2$$

i.e. $\tilde{\sigma}_2^{o^2}$ is inconsistent estimator for σ_2^2.

The result in (VI.84) suggests that if we define an estimator $\hat{\sigma}_2^{o^2}$ as

$$\hat{\sigma}_2^{o^2} = \frac{1}{T} \sum_{t=1}^{T} \frac{[y^o(t) - \tilde{\beta}_2^o x_2^o(t)]^2}{x_2^2(t)} - s_1^2 \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)} \qquad \text{(VI.86)}$$

where s_1^2 is a consistent estimator for σ_1^2 obtained by the HH method.[8]

It is easy to show that $\hat{\sigma}_2^{o^2}$ will be consistent.

$$\Plim_{T\to\infty} \hat{\hat{\sigma}}_2^2 = \Plim_{T\to\infty} \sum_{t=1}^{T} \frac{1}{T} \frac{[y^o(t) - \bar{\beta}_2 x_2^o(t)]^2}{x_2^2(t)}$$

$$- \Plim_{T\to\infty} s_1^2 \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)} \qquad \text{(VI.87)}$$

because $\Plim_{T\to\infty} \tilde{\beta}_2^o = \bar{\beta}_2$

or

$$\Plim_{T\to\infty} \hat{\hat{\sigma}}_2^2 = \Lmt_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{\text{Var } [y^o(t)]}{x_2^2(t)} - \sigma_1^2 \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)} \qquad \text{(VI.88)}$$

_____

[8] If the s_1^2 obtained by the HH method is negative we could reverse the process by choosing a positive estimate of, say, σ_2^2.

Substituting the value of var $[y^o(t)]$ from (VI.83) (VI.88) we have,

$$\text{Plim}_{T\to\infty} \hat{\sigma}_2^2 = \text{Lmt}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left[\frac{x_1^2(t) \frac{1}{T^2} \sum_{t=1}^{T} \frac{\phi^2(t)}{x_1^2(t)} + \frac{T-2}{T} \phi^2(t)}{x_2^2(t)} \right] \quad (VI.89)$$

$$- \sigma_1^2 \text{Lmt}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)}$$

$$= \text{Lmt}_{T\to\infty} \frac{1}{T} \left[\left(\frac{1}{T} \sum_{t=1}^{T} \frac{\phi^2(t)}{x_1^2(t)} \right) \left(\frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)} \right) \right]$$

$$+ \text{Lmt}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{\phi^2(t)}{x_2^2(t)} - \text{Lmt}_{T\to\infty} \frac{1}{T} \left[\frac{2}{T} \sum_{t=1}^{T} \frac{\phi^2(t)}{x_2^2(t)} \right]$$

$$- \sigma_1^2 \text{Lmt}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)}$$

$$= \text{Lmt}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{\phi^2(t)}{x_2^2(t)} - \sigma_1^2 \text{Lmt}_{T} \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)}$$

Substituting the value of $\phi^2(t)$, and rearranging, we have

$$\text{Plim}_{T\to\infty} \hat{\sigma}_2^2 = \sigma_2^2 + \sigma_1^2 \text{Lmt}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)} \quad (VI.90)$$

$$- \sigma_1^2 \text{Lmt}_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{x_1^2(t)}{x_2^2(t)}$$

$$= \sigma_2^2$$

i.e. $\hat{\hat{\sigma}}_2^{\,2}$ is a consistent estimator.[9]

This estimator cannot be called the maximum likelihood estimator but it has two desirable properties:

(a) the estimate of variance is hopefully positive and

(b) the estimator is consistent.

This is hopefully an improvement over the methods discussed in Chapter III.[10] However, a detailed comparative study in this direction is required.

Having obtained consistent positive estimates of $\sigma_j^{\,2}$, $j = 1, 2.$, we may estimate, $\hat{\hat{\phi}}^2(t) = s_1^{\,2} x_1^{\,2}(t) + \hat{\hat{\sigma}}_2^{\,2}(t)$ and $\hat{\hat{\phi}}$. This may be used to estimate the mean response vector as

$$\hat{\bar{\beta}} = (x'\hat{\hat{\phi}}^{-1} x)^{-1} x'\hat{\hat{\phi}}^{-1}y, \quad \ldots \ldots \quad (VI.91)$$

where

$$y = \begin{bmatrix} y(1) \\ \cdot \\ \cdot \\ \cdot \\ y(T) \end{bmatrix} \quad \hat{\bar{\beta}} = \begin{bmatrix} \hat{\bar{\beta}}_1 \\ \hat{\bar{\beta}}_2 \end{bmatrix} \quad , \quad x = [x_1 : x_2] = \begin{bmatrix} x_1(1) & x_2(1) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ x_1(T) & x_2(T) \end{bmatrix} (VI.92)$$

---

[9] The analysis in this chapter can also be taken as an analysis of the specification error (of omitting a relevant variable) in the regression model with random coefficients.

[10] It cannot be ruled out that estimator (VI.87) may be negative sometimes (though in couple of tests we did not find it to be so). However, we have not worked out the conditions under which the estimator (VI.87) will yield guaranteed positive estimates. These conditions need to be spelled out.

$$\text{and } \hat{\hat{\Phi}} = \begin{bmatrix} x_1^{\,2}(1)s_1^{\,2} + x_2^{\,2}(1)\hat{\hat{\sigma}}_2^{\,2} & & & & & 0 \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ 0 & & & & & x_1^{\,2}(T)s_1^{\,2} + x_2^{\,2}(T)\hat{\hat{\sigma}}_2^{\,2} \end{bmatrix} \qquad \text{(VI.93)}$$

CHAPTER VII

SUMMARY AND EPILOGUE

In this thesis we discussed several aspects of the specification and the estimation in the linear regression models with random coefficients. In the course of development of the thesis we made several theoretical and emperical contributions, which are as follows:

1) We proposed an extention in the specification of the linear regression model random coefficient wherein we let the mean response be functionally related to the excluded and/or included variables. For example, the marginal propensity to consume may be interpreted as function of included variable like income or of excluded variable like rate of interest, ratio of non-human to human wealth like rate of interest as proposed by Friedman (1957). This idea is being explored further seperately.

2) A new estimator and an extention of Theil's Weighted Least Squares (TWLS) estimator were proposed theoretically. The new estimator is called the Stepwise Aitken's Weighted Least Squares estimator (SALS). This work was jointly done with Professor A. L. Nagar.

3) A theoretical proof of the unbiasedness was developed for the Hildreth and Houch (HH) Estimator of the mean response coefficients. Similar proofs could be constructed

181

for the SALS and the TWLS estimators of the mean response coefficients.

4) The formula for the standard error of the conditional forecast was developed for the alternative estimators of the random coefficients regression models. Further, we used the forecasting efficiency as a criterion in choosing among alternative estimators in the sampling study.

5) The formula for the coefficient of multiple correlation (R) was developed for the alternative estimators of the random coefficients regression models. Further, we used the mean and the median values of R as a criterion in choosing among alternative estimators in the sampling study.

6) The problem of negative estimates of variances was discussed geometrically along with some ad-hoc solutions proposed in the literature. We proposed a Stepwise Least Squares approach to the estimation of the means and variances of the random coefficients regression model analogous to the Maximum Likelihood estimator. This approach enables us to obtain estimates of the variances, which are consistent and positive. Hopefully this is an improvement over other methods. Here we have developed only the theory.

7) The finite sample properties of four estimators of the mean response coefficients and three estimators of the variances of random coefficients were studied using the Monte Carlo technique. The study extends to three models (one of which has the property that it guarantees to give positive

estimates of variance i.e. Model I - a simple proportional constant Mean Response Model) for the sample sizes 10, 20 and 50 (only for Model III-Set I). In the Monte Carlo study we studied the performances of estimators both in terms of parametric and nonparametric statistics and the analysis has been carried out for more than one point in the parameter space.

The results of the sampling experimentation are:

Estimation of the Mean Response Coefficients

(i) All estimators of the mean response coefficients are emperically unbiased on the t-test at .05 level of significance but for one exception - the Ordinary Least Squares (OLS) estimator for sample size 10 in Models II and III is biased.

(ii) In the Model I with guaranteed positive estimates of variance the SALS estimator of mean response coefficient is the best choice.

(iii) In the models, II and III, which have the problem of negative estimates of variances, there are several conclusions.

iiia. For sample size 10 the OLS estimator of the mean response is the best choice.

iiib. For sample size 20 the SALS estimator is the best choice when parameters values are 'small' and the TWLS estimator is the best choice when parameter values are 'large'.

iiic. For sample size 50 the OLS estimator of the mean

response coefficients is once again the best

choice (even though it does worse marginally

than the SALS and the TWLS estimators of the mean

response coefficients in terms of the efficiency

of estimation and forecasting precision).

In proclaiming an estimator the 'best choice' we have
taken into consideration the performance of the estimator for
its estimation efficiency (bias conbined with dispersion - root
mean square error), forecasting precision, mean/median value of
coefficient of multiple correlation, the bias of the standard
errors of the estimates derived from asymptotic formulas and
cost of the computation. Since there is no definite way to
assign weights for each criterion our conclusions may appear
subjective to some readers. Further, the conclusions are
generally based on the parametric statistics. In some instances
these results were different from the results of the non-
parametric statistics. We decided however to suppress the
conclusions on the non-parametric statistics in order to avoid
the burden on the reader of a variety of conclusions. The
interested reader may refer to Chapter V for the results of
the sampling experiments on the non-parametric statistics. The
results on biases of standard errors of estimates derived from
the asymptotic formulas may also be read in the text.

Estimation of the Variances of the Random Coefficients

(i) All estimators of the variances of the random coeffic-
ients are emperically unbiased (conclusion based on the
CMR Model I only as similar analysis on other model was
not practical).

(ii) The SALS and the TWLS estimators of the variances of the
random coefficients (with one exception) are more efficient
than the HH estimator of variances in all models and all
sample sizes. The efficiency for the CMR Model II and
III have been measured in terms of the proportion of
negative estimate of variances (see Fig. III.1).

8) The efficiency and convergence properties for three iter-
ative estimators were investigated for sample size 20
using the Monte Carlo technique. The convergence of the
mean response co-efficient was studied at the third
iteration for the proportional CMR Model I, while for the
CMR Models II and III convergence was studied at the fifth
iteration. The results of samplings experiments are as
follows:

Efficiency of the Iterative estimators of the Mean Response Coefficients

(i) In the CMR Model I there is no gain in iterating any of
the HH, the TWLS and the SALS estimators of the mean
response coefficients as can be expected theoretically.

(ii) In the CMR Models II and III iterating the SALS, the HH
and the TWLS estimators does increase efficiency in several

instances but the gain is often marginal and hence we cannot justify iterating the estimators when increase in computing cost is taken into consideration. Sometimes there is loss in iterating the HH and the TWLS estimators.

Efficiency of the Iterative Estimators of the Variances of the Random Coefficients

(i) In the CMR Model I there is a definite gain in iterating the SALS and the TWLS estimators of variances. The gain is realised at the second stage of iterations. Further, there is loss in iterating the HH estimator of variance.

(ii) In the CMR Models II there is no gain in iterating the SALS, the TWLS and the HH estimators of variances of random coefficients. In Model III there is slight gain but cost consideration may outweigh this slight gain.

Convergence of the Mean Response Coefficients

(i) In the proportional CMR Model I with guaranteed positive estimates of variances, the convergence is achieved in 100 percent cases. In the CMR Models II and III wherein we have the problem of negative estimates of variances convergence is achieved in 80 to 90 percent cases within the domain of the convergence test applied by us (see footnote 40 Chapter V).

As is well known, the Monte Carlo results are only sugggestive and not conclusive. We therefore caution the reader of the limitations of our conclusion. Further, we have presented the results as they were obtained, often not knowing why they occurred. With full theoretical explanation, there would, of course, be no need of the sampling experiments.

We conclude this chapter with a few suggestions for further research. In the text and in the footnotes at several places we pointed out the need for further research. We may again mention here a few areas requiring attention. There is a great deal of theoretical analyses required to be done in the estimation of random coefficient regression model. It is our hope that our work will stimulate research in this direction. Besides, we did not undertake the sampling study on the Maximum Likelihood Estimator, and the new estimator proposed by us in Chapter VI. Both of these need to be studied with the empirical tools. In addition several inconsistent but positive estimators of variances can be defined (refer discussion on Chapter VI) using stepwise regression technique whose properties need be studied for small samples. Further, there is need to study the properties of alternative estimators when the disturbances are not normal and may for example, follow a Cauchy or a lognormal distribution.

Another area which may prove interesting is to study estimation problems on dynamic equations, where, for example, theory suggests the lagged dependent variable in the regression as a causal variable.

BIBLIOGRAPHY

Agarwala, R., and I. Drinkwater (1972). "Consumption Functions with Shifting Parameters Due to Socio-Economic Factors." The Review of Economics and Statistics, Vol. LIV, No. 1, pp. 89-93.

Basmann, R.L. (1961). "A Note on the Exact Sample Frequency Functions of Generalized Classical Linear Estimators in Two Leading Overidentified Cases." Journal of American Statistical Association, 50, pp. 619-636.

Beltrami, Edward J. (1970). An Algorithmic Approach to Non-Linear Analysis and Optimization. Academic Press, New York.

Box, G.E.P., and M.E. Muller (1958). "A Note On the Generation of Normal Deviates." Annals of Mathematical Statistics, 28, pp. 610-611.

Carter, R.A.L. (1972). The Generation and Testing of Random Numbers, Research Report 7121, Sept. 1971, Department of Economics, The University of Western Ontario, London, Canada.

Chernoff, Herman (1956). "Large Sample Theory: Parametric Case." Annals of Mathematical Statistics, 27, pp. 1-22.

Chernoff, Herman, and Leon E. Moses (1959). Elementary Decision Theory. New York: John Wiley & Sons Inc., 1959.

Chetty, V.K. (1968). "Pooling of Times Series and Cross Section Data." Econometrica, 36, pp. 279-290.

Christ, Carl F. (1966). Econometric Models and Methods. New York: John Wiley & Sons, Inc., 1966.

Chew, Victor (1970). "Covariance Matrix Estimation in Linear Models." Journal of the American Statistical Association, Vol. 65, No. 329, pp. 173-181.

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

Friedman, Milton (1957). A Theory of Consumption Function. Princeton University Press for NBER, Princeton, N.J.

Glejser, H. (1969). "A New Test for Heteroscedasticity." Journal of the American Statistical Association, 64, pp. 316-323.

Goldberger, A.S. (1961). "Stepwise Least Squares: Residual Analysis and Specification Error." Journal of the American Statistical Association, 56, pp. 998-1000.

Goldberger, Arthur S. (1964). Econometric Theory. New York: John Wiley & Sons, Inc.

Goldberger, A.S., and D.B. Jochems (1961). "Note on Stepwise Least Squares." Journal of the American Statistical Association, 56, pp. 105-110.

Hadley, G. (1964). Nonlinear and Dynamic Programming. Reading, Mass.: Addison-Wesley Publishing Company, Inc.

Hartley, H.O. (1961). "The Modified Gauss-Newton Method for Fitting Non-Linear Regression Functions by Least Squares." Technometrics, Vol. 3, No. 2, pp. 269-280.

Hildreth, C. and J.P. Houck (1968). "Some Estimators for a Linear Model with Random Coefficients." Journal of the American Statistical Association, 63, pp. 584-595.

Hoch, I. (1962). "Estimation of Production Function Parameters Combining Time Series and Cross-Section Data." Econometrica, 30, pp. 34-53.

Hussain, A., and T.D. Wallace (1967). A Note on Some Misconceptions of Asymptotic Distribution Theory in Econometrics, Econometric Workshop, Department of Economics and Experimental Statistics, North Carolina State University, Raleigh, U.S.A.

Huzurbazar, V.S. (1948). "The Likelihood Equations, Consistency and the Maxima of the Likelihood Function." Annals of Eugenics, Vol. 14, pp. 185-200.

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

Judge, G.G., and T. Takayama (1966). "Inequality Restrictions in Regression Analysis." Journal of the American Statistical Association, 61, pp. 168-181.

Kakwani, N.C. (1967). "The Unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators." Journal of the American Statistical Association, 62, pp. 141-142.

Kane, Edward J. (1968). Economic Statistics and Econometrics: An Introduction to Quantitative Economics. New York: Harper & Row.

Kendall, M.G., and A. Stuart (1963, 1967, 1966). The Advanced Theory of Statistics. Three-Volume edition. Volume 1, Second edition 1963 (first edition, 1958). Volume 2, Second edition 1967 (first edition, 1961). Volume 3, 1966. New York: Hafner Publishing Company.

Klein, L.R. (1953). A Textbook of Econometrics. Evanston: Row, Peterson and Co., pp. 211-215.

Kmenta, J. (1964). Elements of Econometrics. Ontario: The Collier Macmillan Canada Ltd.

Kmenta, J., and R.F. Gilbert (1968). "Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions." Journal of the American Statistical Association, 63, pp. 1180-1200.

Koopmans, Tjalling C., ed. (1950). Statistical Inference in Dynamic Economic Models. Cowles Commission Monograph, No. 10, John Wiley & Sons, Inc.

Kuh, E. (1959). "The Validity of Cross-Sectionally Estimated Behavior Equations in Time-Series Applications." Econometrica, 27, pp. 197-214.

MacLaren, M.D. and G. Marsaglia (1965). "Uniform Random Number Generators." Journal of the Association for Computing Machinery, 12, pp. 83-89.

Malinvaud, E. (1966). Statistical Methods of Econometrics. (English translation). Chicago: Rand McNally & Company. Amsterdam: North Holland Publishing Company.

Mood, A.M., and F.A. Graybill (1963). Introduction to the Theory of Statistics. Second Edition (first edition 1950). New York: McGraw-Hill Book Company, Inc.

Mundlak, Y. (1963). "Estimation of Production and Behavioural Functions from a Combination of Cross-section and Time Series Data; in C.F. Christ (ed). Measurement in Economics. Stanford: University Press.

Nerlove, M. (1965). Estimation and Identification of Cobb-Douglass Production Functions. Chicago: Rand McNally and Co. Amsterdam: North Holland Publishing Co., pp. 18-38.

Rao, C.R. (1965a). Linear Statistical Inference and its Applications. New York: John Wiley and Sons.

Rao, C.R. (1965b). "The Theory of Least-Squares When the Parameters are Stochastic and its Application to the Analysis of Growth Curves." Biometrika, 52, pp. 447-458.

Rao, C.R. (1967). "Least Squares Theory Using an Estimated Dispersion Matrix and its Application to Measurement of Signals." Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press, pp. 355-71.

Rao, C.R. (1970). "Estimation of Heteroscedastic Variances in Linear Models." Journal of the American Statistical Association, 65, pp. 161-166.

Rao, P., and Z. Griliches (1969). "Small-Sample Properties of Several Two-Stage Regression Methods in the Context of Auto-Correlated Errors." Journal of American Statistical Association, 64, pp. 253-272.

Roberts, K. (1968). Investigation of UC RANDOM. University of Chicago Computation Center Technical Note Number 5.

Rubin, H. (1950). "Note on Random Coefficients," in the Statistical Inference in Dynamic Economic Models. T.C. Koopmans (Editor), Cowles Commission Monograph No. 10, pp. 419-421.

Singh, B., A.L. Nagar, and Baldev Raj. (1972). Measurement of Structural Change: An Application of Random Coefficients Regression Model, Research Report 7203, Department of Economics, The University of Western Ontario, London, Canada.

Srinivasan, T.N. (1970). "Approximations to Finite Sample Moments of Estimators Whose Exact Sampling Distributions are Unknown." Econometrica, Vol. 38, No. 3, pp. 1-10.

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

Swamy, P.A.V.B. (1967). "Statistical Inference in Random Coefficient Regression Models using Panel Data." Systems Formulation, Methodology and Policy Workshop Paper 6701, University of Wisconsin.

Swamy, P.A.V.B. (1968). "Statistical Inference in Random Coefficient Regression Models." Unpublished Ph.D. thesis, University of Wisconsin.

Swamy, P.A.V.B. (1970). "Efficient Inference in a Random Coefficient Regression Model." Econometrica, Vol. 38, No. 2.

Swamy, P.A.V.B., and G.S. Maddala (1968). Tests of Random Coefficient vs. Fixed Coefficient Models Based on the Likelihood Ratio Principle." Department of Economics Workshop Paper No. 35, State University of New York at Buffalo.

Telser, L.G. (1964). "Iterative Estimation of a Set of Linear Regression Equations." Journal of the American Statistical Association, 59, pp. 845-862.

Theil, H. (1954). Linear Aggregation of Economic Relations. Amsterdam: North-Holland Publishing Company.

Theil, H. (1968). "Consistent Aggregation of Micromodels with Random Coefficients." Report 6816, Center for Mathematical Studies in Business and Economics, The University of Chicago.

Theil, H. (1971). Principles of Econometrics. New York: John Wiley and Sons, Inc.

Theil, H. and A.S. Goldberger (1961). "On Pure and Mixed Statistical Estimation in Economics." International Economic Review, 2, pp. 65-78.

Theil, H., and L.B.M. Mennes (1959). "Conception Stochastique de Coefficients Multiplicateurs dans l'Adjustment Linéarire des Series Temparilles." Publications de l'Institude de Statistique de l'Université de Paris, Vol. 8, pp. 211-227.

Theil, H. and L.B.M. Mennes (1959). "Multiplicative Randomness in Time Series Regression Analysis." Mimeographed Report No. 5901, Econometric Institute of the Netherlands School of Economics.

Tiao, G.C. and A. Zellner (1964a). "On the Bayesian Estimation of Multivariate Regression." Journal of the Royal Statistical Society, Series B, 26, pp. 277-285.

Tiao, G.C. and A. Zellner (1964b). "Bayes's Theorem and the Use of Prior Knowledge in Regression Analysis." Biometrika, 51, pp. 219-230.

The Rand Corporation (1955). One Million Random Digits and One Hundred Thousand Deviates. Glencoe, Ill.: The Free Press.

Thornber, H. (1966). "Application of Decision Theory of Econometrics." Unpublished Ph.D. Thesis, Department of Economics, The University of Chicago.

Thornber, H. (1967). "Finite Sample Monte Carlo Studies: An Autoregressive Illustration." Journal of the American Statistical Association, 62, pp. 801-818.

Rosenberg, B. (1967). "Regression in the Presence of Stochastically Varying Parameters." Mimeographed report, School of Business Administration, University of California at Berkeley.

Wonnacott, Thomas H., and Ronald J. Wonnacott (1969). Introductory Statistics. New York: John Wiley & Sons, Inc., 1969.

Wonnacott, Ronald J., and Thomas H. Wonnacott (1970). Econometrics. New York: John Wiley & Sons, Inc.

Zellner, A. (1962). "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." Journal of the American Statistical Association, 57, 348-368.

Zellner, A. (1963). "Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results." Journal of the American Statistical Association, 58, 977-992.

Zellner, A. (1966). "On the Aggregation Problem: A New Approach to a Troublesome Problem." Report # 6628, Center for Mathematical Studies in Business and Economics, University of Chicago.

Zellner, A. (1970). Introduction to Bayesian Inference in Econometrics. New York: Wiley and Sons.

APPENDIX 1

THE COVARIANCE MATRIX OF THE DISTURBANCES IN THE REGRESSION (III.34).

In (III.45) we showed that

$$E(\xi\xi') = E(\hat{\eta}\ \hat{\eta}') - E(\hat{\eta})\ E(\hat{\eta}') \quad \ldots \ldots \quad \text{(A1.1)}$$

where

$$E\hat{\eta} = \dot{M}\dot{X}\dot{\sigma} \quad \ldots \ldots \ldots \ldots \quad \text{(A1.2)}$$

as shown in (III.31) and thus to evaluate the covariance matrix of ξ, we must evaluate $E(\hat{\eta}\ \hat{\eta}')$

From (III.29) we have

$$\hat{\eta}_{Tx1} = \begin{bmatrix} \hat{\eta}^2(1) \\ \cdot \\ \cdot \\ \cdot \\ \hat{\eta}^2(T) \end{bmatrix} = \begin{bmatrix} \left(\sum_{t=1}^{T} m_{1t}\ \eta(t)\right)^2 \\ \cdot \\ \cdot \\ \cdot \\ \left(\sum_{t=1}^{T} m_{Tt}\ \eta(t)\right)^2 \end{bmatrix} \ldots \ldots \quad \text{(A1.3)}$$

which may be written as

$$\hat{\eta}_{Tx1} = \begin{bmatrix} \sum_{t,t'=1}^{T} m_{1t}\ m_{1t'}\ \eta(t)\ \eta(t') \\ \cdot \\ \cdot \\ \sum_{t,t'=1}^{T} m_{Tt} m_{Tt'}\ \eta(t)\ \eta(t') \end{bmatrix} \ldots \quad \text{(A1.4)}$$

where $m_{tt'}$ is the element in the t-th row and t'-th column of the matrix M of order TxT defined in (III.24) and $\eta(t)$ is

194

the t-th element of η defined in (III.1).

The element in the k-th row and k'-th column (k, k' = 1,2 ..., T) of $E\hat{\eta}\hat{\eta}'$ may then be written as

$$E\left[\sum_{t,t'=1}^{T} m_{kt}m_{kt'}\eta(t)w(t')\right]\left[\sum_{t'',t'''=1}^{T} m_{k't''}m_{k't'''}\eta(t'')\eta(t''')\right]$$

.... (A1.5)

$$= \sum_{t,t',t'',t'''=1}^{T} m_{kt}m_{kt'}m_{k't''}m_{k't'''}E[\eta(t)\eta(t')\eta(t'')\eta(t''')]$$

In order to evaluate the mathematical expectation in (A1.5) we may assume that $\eta(t)$'s are independently, normally distributed with means and variances defined in (III.2) – (III.4). This assumption simplifies the results drastically.

The terms in the summation in (A1.5) will be non zero only if either t,t',t'',t''' are pairwise equal or $t=t'=t''=t'''$. In a typical former case the value of the summation is equal to

$$\left[\sum_{t,t'=1}^{T} m_{kt}^2 m_{k't'}^2 \phi^2(t) \phi^2(t')\right] - \sum_{t=1}^{T} m_{kt}^2 m_{k't}^2 \phi^4(t) \quad . . \text{(A1.6)}$$

In the latter case, however, the value of the summation is equal to

$$3 \sum_{t} m_{kt}^2 m_{k't}^2 \phi^4(t) \quad .... \text{(A1.7)}$$

If we combine and rearrange the terms such as (A1.6) and (A1.7) we may write the element in k-th row and k'-th

column of $E\hat{\eta}\hat{\eta}'$ as

$$\left[\sum_{t=1}^{T} m_{kt}^2 \ \phi^2(t)\right] \ \left[\sum_{t=1}^{T} m_{k't'}^2 \ \phi^2(t')\right] \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ (A1.8)$$

$$+ \ 2\left[\sum_{t=1}^{T} m_{kt} \ m_{k't} \ \phi^2(t)\right]^2$$

Therefore,

$$E\hat{\eta}\hat{\eta}' = E(\hat{\eta}) \ E(\hat{\eta}') + 2\dot{\psi} \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ (A1.9)$$

where $E(\hat{\eta})$ is defined in (III.30) and $\dot{\psi}$ is the matrix of squared elements of TXT matrix

$$\psi = M\Phi M, \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ (A1.10)$$

Φ is the covariance matrix of disturbances η's defined in (III.3).

It follows from (A1.1) and (A1.9) that

$$E\xi\xi' = 2\dot{\psi} \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ (A1.11)$$

APPENDIX 2

A REPRESENTATIVE HISTOGRAM OF THE ESTIMATES OF σ_1^2 SHOWING THE POSSIBILITY OF NEGATIVE ESTIMATES (MODEL II-SET I; ESTIMATOR: TWLS, SAMPLE SIZE (10))

HISTOGRAM