INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in “sectioning” the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
ON THE RELATIVE PROPERTIES OF ORDINARY
LEAST SQUARES ESTIMATION FOR THE
PREDICTION PROBLEM WITH ERRORS IN VARIABLES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Bong Jin Yum, B.S., M.S.

* * * * *

The Ohio State University

1981

Reading Committee: Approved By
Albert B. Bishop
Clark A. Mount-Campbell
John B. Neuhardt

Adviser
Department of Industrial and Systems Engineering
ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to his adviser, Dr. J. B. Neuhardt for his assistance and guidance which made the author's graduate career in the Ohio State University extremely rewarding and unforgettable. The author also wishes to thank Dr. A. B. Bishop and Dr. C. A. Mount-Campbell for their constructive criticisms and suggestions in the preparation of this dissertation.

Particular thanks are due to Dr. E. A. Whitehurst for supporting this research through the Ohio State University Engineering Experiment Station.

The author wishes to thank his parents for their encouragement, and his wife, Tong Kyung for her patience and typing this manuscript.
VITA

August 13, 1948 . . . Born - Seoul, Korea
February, 1971 . . . B.S.E.E., Seoul National University, Seoul, Korea
1977-1978 . . . . . Teaching Assistant, Department of Industrial Engineering, Oregon State University, Corvallis, Oregon
June, 1977 . . . . . M.S.I.E., Oregon State University, Corvallis, Oregon
1979-1981 . . . . . Research Associate, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELD OF STUDY

Major Field: Industrial Engineering

Studies in Applied Linear Models. Professor John B. Neuhardt
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF SYMBOLS AND ABBREVIATIONS</td>
<td>viii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td><strong>I.</strong> INTRODUCTION: LINEAR STATISTICAL MODELS</td>
<td></td>
</tr>
<tr>
<td>Errors-in-variables Models</td>
<td>4</td>
</tr>
<tr>
<td>Estimation Techniques</td>
<td>8</td>
</tr>
<tr>
<td>The Prediction Model</td>
<td>11</td>
</tr>
<tr>
<td>The Linear Calibration Problem and Constraints on Parameters</td>
<td>13</td>
</tr>
<tr>
<td>Optimal Design</td>
<td>16</td>
</tr>
<tr>
<td>Summary of Objectives and Outline of Research</td>
<td>17</td>
</tr>
<tr>
<td><strong>II.</strong> LARGE-SAMPLE APPROXIMATION TO THE INTEGRATED MEAN SQUARED ERROR</td>
<td></td>
</tr>
<tr>
<td>Derivation of IMS&lt;sub&gt;1&lt;/sub&gt;</td>
<td>19</td>
</tr>
<tr>
<td>Derivation of IMS&lt;sub&gt;1&lt;/sub&gt;</td>
<td>19</td>
</tr>
<tr>
<td>Properties of the Integrated Mean Squared Error</td>
<td>22</td>
</tr>
<tr>
<td>Comparisons of IMS&lt;sub&gt;1&lt;/sub&gt; and IMS&lt;sub&gt;m&lt;/sub&gt;</td>
<td>23</td>
</tr>
<tr>
<td><strong>III.</strong> MONTE CARLO SIMULATION STUDY OF THE INTEGRATED MEAN SQUARED ERROR</td>
<td></td>
</tr>
<tr>
<td>Design of the Monte Carlo Simulation</td>
<td>34</td>
</tr>
<tr>
<td>Computer Program</td>
<td>37</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                      Page
1. Components of Asymptotic Integrated Mean Squared Error          24
2. Factor Levels for the Simulation Experiment          34
3. Design of the Monte Carlo Simulation Experiment          36
4. Functions of Main and Subprograms          39
5. Results of the Monte Carlo Simulation Experiment          42
6. Results of Additional Runs for N=30 to Compare Simulation and Large-Sample Approximation          43
7. Results of Additional Runs for N=300 to Compare Simulation and Large-Sample Approximation          45
8. Summary of Analysis of Variance (Dependent Variable = DF)          49
9. Summary of Analysis of Variance (Dependent Variable = DELTA)          50
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. DF versus $\tau^2$ when $N=300$ and $\sigma^2=0.0001$</td>
<td>29</td>
</tr>
<tr>
<td>2. DF versus $\tau^2$ when $N=30$ and $\sigma^2=0.0001$</td>
<td>30</td>
</tr>
<tr>
<td>3. DF versus $\tau^2$ when $N=300$ and $\sigma^2=0.01$</td>
<td>31</td>
</tr>
<tr>
<td>4. DF versus $\tau^2$ for varying $\beta_1$ ($N=300$, $\sigma^2=0.0001$)</td>
<td>32</td>
</tr>
<tr>
<td>5. Interactions between Main and Subprograms</td>
<td>40</td>
</tr>
<tr>
<td>6. Asymptotic versus Simulated DF Values for $N=30$ ($\beta_1=1.0$ and $\sigma^2=0.0001$)</td>
<td>44</td>
</tr>
<tr>
<td>7. Asymptotic versus Simulated DF Values for $N=300$ ($\beta_1=1.0$ and $\sigma^2=0.0001$)</td>
<td>46</td>
</tr>
<tr>
<td>8. The Effect of $\beta_1$ and $\tau^2$ on DELTA</td>
<td>51</td>
</tr>
<tr>
<td>9. IMS$<em>{10}$ versus $\Delta</em>{10}$</td>
<td>58</td>
</tr>
</tbody>
</table>
LIST OF SYMBOLS AND ABBREVIATIONS

\( \beta \) vector of unknown coefficients of linear models.
\( \beta_0 \) intercept of the two-variable linear model.
\( \beta_1 \) slope of the two-variable linear model.
b an estimate of \( \beta \).
\( \hat{b} \) least squares estimate of \( \beta \).
\( \tilde{b} \) maximum likelihood estimate of \( \beta \).

\( \text{BVN}(A, B) \) bivariate normal distribution with mean \( A \) and variance \( B \).
\( \text{COV}(\cdot, \cdot) \) covariance

\( D \) factor representing the design of the experiment for estimating \( \beta \) in the Monte Carlo simulation.

\( \Delta_1 \) (observed DF) - (calculated DF).

\( \text{diag} A \) matrix with only diagonal elements of \( A \).

\( \text{DF} \) IMS\(_1\) - IMS\(_m\).

\( \delta \) vector of bias in \( b \).

\( \Delta_1 = \sum x_i^2 / N + \gamma^2 - \bar{x}^2 \).

\( \Delta_2 = \sum x_i^2 / N - \bar{x}^2 \).

\( e \) measurement error in the independent variables for the general linear model. \( e' = (0, e_1, \ldots, e_p) \).

\( \epsilon \) measurement error in the dependent variable for the general linear model. \( \epsilon' = (\epsilon_1, \ldots, \epsilon_N) \).

\( \text{E}(\cdot) \) expectation operation.

\( \text{EVM} \) errors-in-variables model.
\( I_k \) identity matrix of order \( k \).

\( \text{IMS} \) integrated mean squared error of prediction.

\( \text{IMS}_l \) IMS using the least squares estimation.

\( \text{IMS}_m \) IMS using the maximum likelihood estimation.

\( \text{LS} \) least squares estimation.

\( m \) total number of future measurements of the independent variable \( x \).

\( \text{ML} \) maximum likelihood estimation.

\( \text{MSE} \) mean squared error.

\( n \) total number of design points in the experiment for estimating \( \beta \).

\( N \) total number of observations in the experiment for estimating \( \beta \).

\( \text{N}(A,B) \) univariate normal distribution with mean \( A \) and variance \( B \).

\( p \) total number of independent variables in a general linear model.

\( R \) volume of \( \mathcal{X} \).

\( r \) total number of repeated measurements at each design point.

\( S \) pooled sample covariance matrix.

\( \text{s.t.} \) subject to.

\( \sigma^2 \) variance of \( v \) in the two-variable model.

\( \Sigma \) covariance matrix of \( e \) in the general linear model.

\( t \)

\[ \sigma^2 + \beta_1^2 \gamma^2. \]

\( \text{tr} \) trace of a matrix.

\( \gamma^2 \) variance of \( u \) in the two-variable model.

\( V \) covariance matrix of \( b \).
Var(•) variance operation.

w vector of the error-contaminated independent variables in the general linear model.

W observed design matrix in the general linear model.

x vector of the independent variables in the general linear model, or the independent variable in the two-variable model.

\( \hat{x} \) estimated value of the independent variable.

X design matrix in the general linear model.

\( \xi \) error-contaminated independent variable.

\( \xi^* \) extreme-point design.

\( \chi \) domain of x, or design space.

y vector of the dependent variable in the general linear model, or the dependent variable in the two-variable model.

\( \hat{y} \) estimated value of the dependent variable.

\( \eta \) observed values of the dependent variable.

u measurement error in the independent variable for the two-variable model.

v measurement error in the dependent variable for the two-variable model.

\( \Omega \) covariance matrix of (u, v).

\( \varphi \) correlation coefficient of u and v.

\( \Lambda \) covariance matrix of \( \varepsilon \).

\( \prime \) transpose of.

\( \sim \) distributed as.

\( \tilde{\sim} \) asymptotically equal to.
\( \approx \) approximately equal to.
\( \Rightarrow \) implies that.
\( \propto \) proportional to.
\( \partial \) partial derivative operation.
CHAPTER I

INTRODUCTION:

LINEAR STATISTICAL MODELS

The linear statistical model is perhaps one of the most widely accepted statistical tools in scientific investigation due to its mathematical tractability and practical value. A linear model may be represented by:

\[ \eta = X\beta + \varepsilon, \]

where \( \eta \) is a vector of the observed dependent variable, \( X \) is a matrix of the values of the independent variables, \( \beta \) is a vector of unknown parameters, and \( \varepsilon \) is a vector of random errors with mean 0 and covariance matrix, \( \Lambda \). Then, the statistical problems related with the model may include the determination of the relationship (i.e., estimation of, and hypothesis testing related to \( \beta \)) and the prediction of \( E(\eta) \) in the future based upon the estimated relationship.

Under certain assumptions (i.e., known and controlled \( X \), uncorrelated error terms, and constant variance of \( \varepsilon \)), elegant theories are available for solving the above statistical problems. On the other hand, there are many instances when some of the assumptions are violated, and
therefore, the existing theories may not be directly applicable. One important departure from the basic assumptions arises when X is not known exactly due to measurement errors in the independent variables. In fact, in virtually all applications of the linear model, one may face the problem of errors or uncertainties not only in the dependent variable but also in the independent variables as well (simple round-off errors in computation, for example). Typically, consider a calibration problem between a non-standard and a standard instrument in which one wishes to predict the standard measurement given a non-standard observation. In the classical theory of calibration, the standard measurements are assumed to be exact, and therefore the non-standard measurements are regressed on the standard measurements to estimate the relationship between the two (e.g., Williams, 1969). However, the standard measurements are frequently contaminated by errors, especially when a large scale calibration experiment is performed in the 'field' (under less than the highly controlled conditions of a laboratory experiment).

When the experimenter feels that measurement errors are present in the independent variables, the usual least squares (LS) estimation technique may be applied to the 'observed' values, or one of the more elaborate analysis techniques may be selected which reflect the uncertainties
explicitly. At that point, the experimenter's obvious question is "what alternative analysis procedure should be chosen based upon the objectives?". In other words, should one use the usual LS estimation technique and obtain a hopefully satisfying result, or choose a more elaborate technique (e.g., maximum likelihood) at the expense of more complicated computation and additional assumptions? Unfortunately, there have been no definite answers to these questions of choice, although a number of theories are available for given techniques. It is particularly true that no comparison of analysis procedures is known when the objective of the experimenter is to use the estimated relationship in the future to predict $E(\eta)$ at certain values of the independent variables.

In the literature, the model which explicitly considers the uncertainties in the independent variables is commonly called an 'errors-in-variables' model (EVM)\(^1\). Most contributions in this area are more concerned with a specific estimator of the unknown $\beta$ and its properties than the relative performance of the competitive estimators. Besides, the current theories in the areas of EVM are restricted in the sense that little work has been done with regard to the prediction problem.

\(^1\) see Malinvaud (1970, p. 374)
Considering the above shortcomings, there is a need for providing guidelines which experimenters can follow when their objective is to predict $E(\eta)$ in the future based upon a current estimated relationship. For this purpose, the usual LS estimation procedure will be compared to a more elaborate technique, the maximum likelihood (ML) estimation procedure based upon a certain criterion which will be defined later. These two estimation techniques are selected since the usual LS estimation procedure is simple and widely accepted in almost all areas of scientific disciplines (as well as having computer programs readily available). The ML estimation procedure is the most typically suggested alternative technique in the areas of EVM because of its known asymptotic properties.

In the following sections, several types of EVM will be introduced together with the corresponding LS and ML estimation procedures. Then, a model will be formulated considering the prediction of future $E(\eta)$, and the approach taken in this research will be explained. Finally, the statistical calibration problem will be introduced as a typical application area of this research as well as the optimal design of the calibration experiment.

Errors-in-variables Models

There is an extensive literature on the topic of EVM. For a general survey, one may refer to Lindley (1947),

During the past thirty years, the two-variable case of EVM has been considered somewhat extensively in the statistical literature. Among them, Lindley (1947) first formalized the problem and provided significant theoretical results. Since then, numerous models under different assumptions have emerged together with corresponding different estimation techniques. However, it is the author's opinion that the structure of the model and the corresponding estimation techniques considered in this research are most typical ones.

Consider the following mathematical relationship between two variables, $x$ and $y$:

$$y = \beta_0 + \beta_1 x$$

(1-1)

Assume that the true values, $x$ and $y$ cannot be observed exactly. Then, depending upon the nature of the observational process there exist two different types of EVM. First, consider the case when the experimenter has no control on the observed values, but simply observes $\xi$ and $\eta$ such that

$$\xi = x + u \quad \text{and}$$

$$\eta = y + v$$

(1-2)
where \(u\) and \(v\) are random errors such that \(\mathbb{E}(u) = \mathbb{E}(v) = 0\), \(\text{Var}(u) = \gamma^2\), and \(\text{Var}(v) = \sigma^2\). For the simplicity of presentation assume that \(u\) and \(v\) are independent. In the above model, the experimenter has no prior information about \(\xi\) nor \(\eta\) before he actually observes them. This first type of EVM can be further classified into two classes based upon the nature of the variables, \(x\) and \(y\). If \(x\), and hence \(y\), is a random variable, the model is called a 'structural relationship' model, and if \(x\) and \(y\) are simply mathematical variables, then a 'functional relationship' is defined (Kendall and Stuart, 1973). In this research only the latter class, more typically found in engineering literature, will be considered.

Then, from Eq. (1-1) and (1-2),

\[
\eta = \beta_0 + \beta_1 \xi + (v - \beta_1 u). 
\]

Let \(v' = v - \beta_1 u\). Then,

\[
\eta = \beta_0 + \beta_1 \xi + v'
\]

where \(v' \sim (0, \sigma^2 + \beta_1^2 \gamma^2)\).

Although the model in Eq. (1-4) appears to be the usual two-variable regression model, one significant difference is that \(\xi\) and \(v'\) are now correlated since

\[
\text{COV}(\xi, v') = \text{COV}(x+u, v - \beta_1 u) = -\beta_1 \gamma^2 \neq 0
\]
unless $\beta_1 = 0$, which is a trivial case. Hence, the classical theory of regression cannot be directly applied in this case.

Consider now the other type of EVM where the experimenter has a known target, or nominal value of $\xi$ before the experiment is performed. For instance, in an experiment for determining the relationship between the extension($y$) of a spring and weights($x$), the nominal values($\xi$) of the weights may be known to the experimenter in advance. Since these nominal values may be inaccurate, when the experimenter hangs the weight to the spring, what is actually applied might be the true weight, $x = \xi - u$. Then $x$ determines $y$ through the physical law, Eq.(1-1), and $\eta = y + v$ is measured. Hence, in this model, $\xi$ is a known fixed value, but $x$ is considered as a random variable which is perfectly, negatively correlated with $u$. Then, Eq.(1-4) for this model is exactly a regression model since $\xi$ is not now a random variable, and is not correlated with $v$. Hence, the usual LS estimation technique can be applied to Eq.(1-4) without any difficulty.

The above model is frequently called 'Berkson's Model' subsequent to his work (Berkson, 1950), and is applicable in many experimental situations. However, in this research only

2. see Kendall and Stuart (1973, p.424)
the functional relationship model in the first type will be examined.

**Estimation Techniques**

In an experiment to estimate the relationship in Eq. (1-1), assume that \( n \) experimental units are available, and \( r \) pairs of repeated measurements are taken at each experimental unit. That is,

\[
\begin{align*}
    y_i &= \beta_0 + \beta_1 x_i \\
    \xi_{ij} &= x_i + u_{ij} \\
    \eta_{ij} &= y_i + v_{ij}
\end{align*}
\]

for all \( i \) and \( j \), where \( \xi_{ij} \) is unknown and \( \rho \) is the correlation coefficient of \( u \) and \( v \).

Further, assume that an error vector,

\[
\begin{pmatrix} u_{ij} \\ v_{ij} \end{pmatrix} \sim \text{BIVN} \left\{ 0, \Omega = \begin{bmatrix} \gamma^2 & \rho \sigma \gamma \\ \rho \sigma \gamma & \sigma^2 \end{bmatrix} \right\}
\]

for all \( i \) and \( j \), where \( \Omega \) is unknown and \( \rho \) is the correlation coefficient of \( u \) and \( v \). Error vectors corresponding to different combinations of \( i \) and \( j \) are assumed to be independent. Now Eq. (1-4) becomes

\[
\eta_{ij} = \beta_0 + \beta_1 \xi_{ij} + v_{ij}
\]

for \( i = 1, 2, \ldots, n \) and \( j = 1, 2, \ldots, r \). Let \( N = rn \), the total number of observations. Then, the usual LS estimate,
\( \hat{b}_0, \hat{b}_1 \) of \((\beta_0, \beta_1)\) is given by:

\[
\hat{b}_1 = \frac{\sum_{i=1}^{n} \sum_{j=1}^{r} (\xi_{ij} - \bar{\xi})(\eta_{ij} - \bar{\eta})}{\sum_{i=1}^{n} \sum_{j=1}^{r} (\xi_{ij} - \bar{\xi})^2}, \quad \text{and} \\
\hat{b}_0 = \bar{\eta} - \hat{b}_1 \bar{\xi}
\]

where

\[
\bar{\eta} = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{r} \eta_{ij}, \quad \text{and} \\
\bar{\xi} = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{r} \xi_{ij}.
\]

As mentioned earlier, the above LS estimator is biased due to the non-zero correlation between \( \xi \) and \( v^t \). For details of the above property, refer to Richardson and Wu(1970).

For the same model in Eq. (1-6) and (1-7), the ML estimates of \( \beta_0 \) and \( \beta_1 \) are explicitly given by Villegas (1961). One may refer to Appendix D for details. First, define

\[
\bar{\xi}_i = \frac{1}{r} \sum_{j=1}^{r} \xi_{ij} , \\
\bar{\eta}_i = \frac{1}{r} \sum_{j=1}^{r} \eta_{ij} , \\
F = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}
\]

(1-11)
\[
\begin{bmatrix}
\sum_{i=1}^{n} (\bar{\xi}_i - \bar{\xi})^2 & \sum_{i=1}^{n} (\bar{\xi}_i - \bar{\xi})(\bar{\eta}_i - \bar{\eta}) \\
\sum_{i=1}^{n} (\bar{\eta}_i - \bar{\eta})^2 & 0
\end{bmatrix}, \text{ and}
\]

\[S = \begin{bmatrix}
s_{11} & s_{12} \\
s_{21} & s_{22}
\end{bmatrix}\]

\[
= \frac{1}{N-n} \begin{bmatrix}
\sum_{i=1}^{n} \sum_{j=1}^{r} (\xi_{ij} - \bar{\xi})^2 & \sum_{i=1}^{n} \sum_{j=1}^{r} (\xi_{ij} - \bar{\xi})(\eta_{ij} - \bar{\eta}) \\
\sum_{i=1}^{n} \sum_{j=1}^{r} (\eta_{ij} - \bar{\eta})^2 & 0
\end{bmatrix}.
\]

That is,

\[\bar{\xi}_i, \bar{\eta}_i = \text{sample means at each design point},\]

\[F = \text{matrix of the sum of squares based upon sample means at each design point, and}\]

\[S = \text{pooled sample covariance}.\]

Then, based upon Villegas,

\[\tilde{b}_1 = \frac{f_{12} - \lambda^*s_{12}}{f_{11} - \lambda^*s_{11}}, \text{ and}\]

\[\tilde{b}_0 = \bar{\eta} - \tilde{b}_1 \bar{\xi} \]

where \(\lambda^*\) is the minimum eigenvalue of the system:

\[F\tilde{b} = \lambda S\tilde{b}.\]
The above ML estimators are also biased, although Bradley and Gart (1962) show that they are consistent for finite n.

Hereafter, the model described by Eq. (1-6) and (1-7) will be called the 'two-variable model'.

The Prediction Model

The important feature, prediction, is now incorporated into the basic two-variable model. Suppose that an estimate, $(b_0, b_1)$ of $(\beta_0, \beta_1)$ has been obtained either by LS or ML estimation procedure, that is, an estimated relationship of Eq. (1-6) is available. Assume that to predict a future $y = E(\eta)$ corresponding to an unknown $x$, m repeated measurements (also in error) are taken as follows:

$$\xi_j = x + u_j ; \; j = 1, 2, \ldots, m. \quad (1-14)$$

Then, an estimate of the unknown $x$ is obtained by:

$$\hat{x} = \frac{1}{m} \sum_{j=1}^{m} \xi_j. \quad (1-15)$$

Inserting $x$ into the estimated relationship yields an estimate of $y$ as

$$\hat{y} = b_0 + b_1 \hat{x} \quad (1-16)$$
Eq.(1-6), (1-14), and (1-16) will be called the prediction model. This, in fact, describes the use of the 'calibration curve', which will be explained later.

Note that depending upon the method used for estimating \((f_0, f_1)\), different values of \(y\) are obtained. In addition, \(y\) may be biased due to the bias in \(b_0\) and \(b_1\).

The statistical behavior of \(\hat{y}\) is represented by its mean squared error(MSE), which is the sum of the variance and squared bias of \(\hat{y}\). However, a specific MSE may be meaningless when many future \(y\)'s are to be predicted over the range of \(x\). Hence, the integrated mean squared error(IMS) is proposed as a measure of the behavior of \(\hat{y}\)'s. Let \(X\) and \(R\) be the domain of \(x\) and the length of \(X\), respectively. Then, the IMS is defined by:

\[
 IMS = \frac{1}{R} \int_X \text{MSE}(\hat{y})dx. \tag{1-17}
\]

That is, the IMS is the average of \(\text{MSE}(\hat{y})\) over \(X\). In the above definition of the IMS, all the points in \(X\) are considered equally important. The case when different weights are given to different points in \(X\) is not considered in this research(refer to Helms, 1974 for the weighted IMS criterion for selecting a good set of the independent variables in linear models). Since different estimates of \((\beta_0, \beta_1)\) would result in different values of IMS, the relative performance of estimation procedures will be assessed by comparing the corresponding IMS values.
In Chapter II, expressions for the IMS based upon large-sample approximation will be developed and compared for the LS as well as ML estimation procedure. The small-sample behavior of the IMS's will be investigated in Chapter III by the Monte Carlo simulation.

The Linear Calibration Problem and Constraints on Parameters

Although the theory and results obtained in this research are fairly general, the calibration problem is taken as an illustration of the prediction model discussed in the previous section. This problem will allow reasonable constraints to be placed on the values of the parameters in the simulation study.

The statistical calibration of a non-standard instrument involves the following two distinct, but closely related activities:

1. determination of the relationship between the non-standard and standard (commensurate) measurements by performing a calibration experiment, and
2. prediction of the standard measurements in the future based upon the estimated relationship and future non-standard observations.

In the classical model, it is assumed that the standard measurements \((x)\) are exact while the non-standard measurements \((\eta)\) are subject to measurement errors or random fluctuations.
That is,
\[ \eta = \beta_0 + \beta_1 x + v \]  \hspace{1cm} (1-18)
where \( v \sim N(0, \sigma^2) \).

To estimate the above relationship, in the calibration experiment a finite number of quantities are measured by the non-standard as well as the standard instrument to obtain a series of \((x_i, \eta_i)\)'s. Then, based upon the LS estimation technique, the calibration equation is determined by:
\[ \hat{E}(\eta) = b_0 + b_1 x. \]  \hspace{1cm} (1-19)

In the future, a quantity is measured \(m\) times by the non-standard instrument and an estimate of \(E(\eta)\) is obtained by the sample mean, \(\bar{\eta}\) of the \(m\) measurements. Then, the corresponding \(x\) is estimated by:
\[ \hat{x} = (\bar{\eta} - b_0)/b_1. \]  \hspace{1cm} (1-20)

That is, the calibration equation is used in reverse to predict future \(x\). However, this mathematically logical procedure results in a statistical problem, especially when a confidence interval for \(x\) is to be constructed. That is, \(\hat{x}\) in Eq.(1-20) has undefined expectation and infinite variance due to the term, \(\frac{1}{b_1}\) as pointed out by Williams (1969). Although Scheffe(1973) solves this problem, the inverse estimation step in the classical calibration problem is generally considered as a computationally demanding,
cumbersome procedure. In passing, it is worth mentioning that Krutchkoff (1967) tried to eliminate this undesirable step by regressing \( x \) on \( \eta \) (inverse regression) rather than \( \eta \) on \( x \), and claimed that his procedure gives a smaller variance of the predicted \( x \). However, this approach has been criticized by Williams (1969), Berkson (1969), and several other authors due to the misuse of the regression relationship and the unjustifiable criterion used for comparison.

Note that the above statistical problem arises because of the asymmetry between \( \eta \) and \( x \) in Eq. (1-18). That is, the relation:

\[
x = -\frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} \eta - \frac{v}{\beta_1}
\]  

(1-21)

is not meaningful in the regression context (Eisenhart, 1939). However, in the prediction model proposed in the earlier sections, this inverse estimation procedure can be completely eliminated without any difficulties. Note that from Eq. (1-1), the role of \( x \) and \( y \) are symmetric, and hence,

\[
x = -\frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} y
\]  

(1-22)

is an equally meaningful relation. Since \( x \) and \( y \) are both measured with errors, the relation:

\[
\xi = -\frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} \eta - \frac{v'}{\beta_1}
\]  

(1-23)

is also equally meaningful as Eq. (1-4).
The above argument implies that the experimenter may eliminate the inverse estimation procedure in the future by putting the standard measurements on the left side of the equation, which is not allowed in classical sense.

Finally, calibration of the instruments will allow the assumption that typical values of $\beta_1$ are unity. Further, the variation in the errors will be kept within what might be considered a 'reasonable' range in the engineering context. In particular, the standard deviation of the error terms will always be assumed to be no greater than 5% of the range of the variable.

**Optimal Design**

Thus far, the two-variable model with measurement errors in both variables has been considered from both an estimation and prediction point of view. Another important aspect of the problem is the optimal design of the experiment for estimating the relationship considering prediction as a future activity. For the classical model, an extensive amount of literature is available on this topic. For instance, one may refer to Ott (1968) and Naszodi (1978) for the optimal design of the classical calibration experiment. Unfortunately, however, little research has been done in the area of EVM. In this research, the optimal design is defined as an optimal selection of the levels of the
independent variable, which minimizes the IMS for a given estimation procedure.

Summary of Objectives and Outline of Research

The two-variable linear model with measurement errors in both variables has been discussed in terms of estimation, prediction, and optimal design of the experiment. Although the proposed approach is applicable to fairly general situations, investigations are made in the context of the statistical calibration problem for convenience.

The purpose of this research is to investigate the problem of choosing an analysis technique (in particular, LS or ML estimation procedure) for the calibration experiment utilizing the IMS as the criterion of selection. Given either analysis procedure, implications in the design of the calibration experiment are also considered.

Chapter II investigates the large-sample performance of the LS and ML estimation procedure by comparing the corresponding IMS's analytically. In Chapter III, the relative performance of the two estimation techniques is investigated for small samples through the use of Monte Carlo simulation. The simulation results are also compared to the large-sample findings to evaluate possible discrepancies between the two. The optimal design of the calibration experiment is considered in Chapter IV based upon the large-
sample expressions developed in Chapter II. Chapter V includes summary, conclusions, and recommendations for future research.
CHAPTER II

LARGE-SAMPLE APPROXIMATION TO
THE INTEGRATED MEAN SQUARED ERROR

The purpose of this chapter is two-fold: (i) to develop expressions for large-sample integrated mean squared error (IMS) using the least squares (LS) as well as using the maximum likelihood (ML) estimation procedure, and (ii) to identify the model parameters which affect the relative performance of the two estimation procedures in terms of the IMS. In addition, the results of this chapter will be compared to those of Monte Carlo simulation in Chapter III to determine whether these large-sample expressions are valid for small samples as well.

Hereafter, IMS₁ will denote the IMS using LS estimation procedure, and IMSₘ using ML estimation procedure.

**Derivation of IMS₁**

For the general linear model described in Appendix A, the LS estimator \( \hat{b} \) of the true value \( \beta \) is given by:

\[
\hat{b} = (W'W)^{-1}W'\eta
\]  

(2-1)
where \( \eta \) is an \((N \times 1)\) vector of the observed dependent variables, and \( W \) is the observed design matrix. For a more precise definition of \( \eta \) or \( W \), refer to Eq. (A-1) and (A-3) in Appendix A. The case when errors in the independent and the dependent variables are independent (i.e., \( \rho = 0 \) for the two-variable model) is of interest.

Let \( \delta \) and \( V \) be the bias and variance of \( \hat{b} \), respectively. That is,

\[
\delta = E(\hat{b}) - \beta, \quad \text{and} \quad V = E\left\{ (\hat{b} - E(\hat{b}))^2 \right\}.
\]

Then, for large \( N \), Davies and Hutton (1975) show that

\[
\delta \sim - \left( \frac{X'X}{N} + \sum \right)^{-1} \sum \beta, \quad (2-2)
\]

and

\[
V \sim \frac{\sigma^2 + \beta' \sum \beta}{N} \left( \frac{X'X}{N} + \sum \right)^{-1}. \quad (2-3)
\]

For the two-variable model considered,

\[
X'X = \begin{bmatrix}
N & \sum_{i=1}^{N} x_i \\
\sum_{i=1}^{N} x_i & N \sum_{i=1}^{N} x_i^2
\end{bmatrix}, \quad (2-4)
\]

\[
\sum = \begin{bmatrix}
0 & 0 \\
0 & \tau^2
\end{bmatrix}, \quad (2-5)
\]

and

\[
\beta' = (\beta_0, \beta_1). \quad (2-6)
\]

As in Appendix A, \( x \) is assumed to be scaled such that

\[-1 \leq x \leq 1.\]
Inserting Eq. (2-4), (2-5), and (2-6) into Eq. (2-2) and (2-3),

\[ \delta \sim \frac{\beta_1 \gamma^2}{\Delta_1} \left( \frac{\bar{x}}{-1} \right), \]  \hspace{1cm} (2-7)

and

\[ V \sim \frac{t}{N \Delta_1} \begin{bmatrix} \gamma^2 + \frac{\sum x_i^2}{N} & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix} \]  \hspace{1cm} (2-8)

where

\[ \Delta_1 = \frac{\sum x_i^2}{N} + \gamma^2 - \bar{x}^2 \]  \hspace{1cm} (2-9)

\[ t = \sigma^2 + \beta_1^2 \gamma^2. \]  \hspace{1cm} (2-10)

Inserting Eq. (2-7) and (2-8) into Eq. (A-14) in Appendix A,

\[ \text{MSE}(\hat{y}) \sim \frac{t \gamma^2}{mN \Delta_1} + \frac{\beta_1^2 \gamma^2}{m} \frac{(1 - \frac{T^2}{\Delta_1})^2}{\Delta_1} \]

\[ + \frac{t}{N \Delta_1} \left( \Delta_1 + \bar{x}^2 - 2\bar{x}x + x^2 \right) \frac{\beta_1^2 \gamma^4}{\Delta_1^2} \left( \bar{x}^2 - 2\bar{x}x + x^2 \right). \]  \hspace{1cm} (2-11)

Similarly, inserting Eq. (2-7) and (2-8) into the expression for IMS, Eq. (A-18) in Appendix A,
Derivation of $\text{IMS}_m$

Since the ML estimator $\tilde{\beta}$ of $\beta$ is consistent, the bias becomes negligible as $N$ becomes large. Let $V = \text{Var}(\tilde{\beta})$.

Then, when $N$ is large, Dolby and Lipton (1972) show that

$$V \sim \frac{t \tau^2}{mN \Delta_1} + \frac{\beta_1^2 \tau^2}{m} \left(1 - \frac{\tau^2}{\Delta_1}\right)$$

$$+ \frac{t}{N} + \frac{t}{N \Delta_1} \left(\bar{x}^2 + \frac{1}{3}\right)$$

$$+ \frac{\beta_1^2 \tau^4}{\Delta_1^2} \left(\bar{x}^2 + \frac{1}{3}\right).$$

(2-12)

Inserting Eq. (2-13) and $\delta = 0$ into Eq. (A-14) in Appendix A,

$$\text{MSE}(\hat{y}) \sim \frac{t \tau^2}{mN \Delta_2} + \frac{\beta_1^2 \tau^2}{m} + \frac{t(\Delta_2 + \bar{x}^2)}{N \Delta_2}.$$
\[- \frac{2t\bar{x}}{N \Delta_2} x + \frac{t}{N \Delta_2} x^2 \quad (2-14)\]

for the two-variable model. Similarly, from Eq.(A-18) in Appendix A,\[\text{IMS}_m \sim \frac{t \gamma^2}{mN \Delta_2} + \frac{\beta_1^2 \gamma^2}{m} + \frac{t}{N} \]
\[+ \frac{t(\bar{x}^2 + \frac{1}{3})}{N \Delta_2}. \quad (2-15)\]

**Properties of the Integrated Mean Squared Error**

Either IMS\(_1\) or IMS\(_m\) may be decomposed into several components as shown in Table 1. Define\[LF = \text{component of IMS}_1 \text{ due to the error in future measurements},\]
\[LV = \text{component of IMS}_1 \text{ due to the variance of } \hat{b},\]
\[LB = \text{component of IMS}_1 \text{ due to the bias of } \hat{b},\]
\[MF = \text{component of IMS}_m \text{ due to the error in future measurements}, \text{ and}\]
\[MV = \text{component of IMS}_m \text{ due to the variance of } b.\]

Then,\[\text{IMS}_1 = LF + LV + LB, \quad \text{and} \quad (2-16)\]
\[\text{IMS}_m = MF + MV.\]

In addition, the parameters involved in the IMS may be classified into the following two classes:
Table 1. Components of Asymptotic Integrated Mean Squared Error.

<table>
<thead>
<tr>
<th>Source of Error</th>
<th>IMS₁</th>
<th>IMSₘ</th>
</tr>
</thead>
<tbody>
<tr>
<td>due to future measurement error</td>
<td>(\frac{t \tau^2}{mN \Delta_1} + \frac{\beta_1^2 \tau^2}{m} \left(1 - \frac{\tau^2}{\Delta_1}\right)^2)</td>
<td>(\frac{t \tau^2}{mN \Delta_2} + \frac{\beta_1^2 \tau^2}{m})</td>
</tr>
<tr>
<td>due to variance</td>
<td>(\frac{t}{N} + \frac{t}{N \Delta_1} (\bar{x}^2 + \frac{1}{3}))</td>
<td>(\frac{t}{N} + \frac{t(\bar{x}^2 + \frac{1}{3})}{N \Delta_2})</td>
</tr>
<tr>
<td>due to bias</td>
<td>(\frac{\beta_1^2 \tau^4 (\bar{x}^2 + \frac{1}{3})}{\Delta_1^2})</td>
<td>-</td>
</tr>
</tbody>
</table>
uncontrollable: $\beta_1, \sigma^2, \tau^2$

controllable: $m, N, \Delta_1, \Delta_2$.

Then, based upon Table 1, it is observed that

1. all components depend upon $\beta_1, \tau^2, \Delta_1$, or $\Delta_2$, where the $\{\Delta_i\}$ correspond to the design of the experiment for estimating $\beta$,

2. all but LB depends upon $\sigma^2$,

3. for $\sigma^2, \tau^2, \beta_1$, and $\{\Delta_i\}$ given,
   a. LF and MF largely depend upon $m$, the number of repeated measurements in the future,
   b. LV and MV depend upon $N$, and
   c. LB depends upon neither $m$ nor $N$.

Since $m, N$, and $\Delta'$s are controllable, those values may be selected for a proper magnitude of IMS. The effect of $\Delta$ on IMS will be examined in Chapter IV.

Comparisons of $IMS_{1}$ and $IMS_{m}$

One of the objectives of this research is to examine the relative performance of LS and ML estimation procedures with respect to the IMS.

From the previous section,

\[
IMS_{1} - IMS_{m} = (LF - MF) + (LV - MV) + LB. \quad (2-17)
\]
However, for any given model parameters,

\[ LF - MF < 0, \text{ and} \]
\[ LV - MV < 0, \]

since

\[ \Delta_1 = \frac{\sum x_i^2}{N} + \tau^2 - \bar{x}^2 > \Delta_2 = \frac{\sum x_i^2}{N} - \bar{x}^2. \]

This implies that there exists a tradeoff between the first two terms and the last in Eq.(2-17).

More rigorously, the difference between IMS\(_1\) and IMS\(_m\) may be examined as follows. From Eq.(2-12) and (2-15),

\[
DF = IMS_1 - IMS_m
= \frac{t \tau^2 (\Delta_2 - \Delta_1)}{mN \Delta_1 \Delta_2} + \frac{\beta_1^2 \tau^2}{m} \left\{ \left( 1 - \frac{\tau^2}{\Delta_1} \right)^2 \right\}
+ \frac{t(\bar{x}^2 + \frac{1}{3})(\Delta_2 - \Delta_1)}{N \Delta_1 \Delta_2}
+ \frac{\beta_1^2 \tau^4 (\bar{x}^2 + \frac{1}{3})}{\Delta_1^2}.
\]

Since \[ \Delta_2 - \Delta_1 = -\tau^2, \]
DF = \frac{\tau^2}{\Delta_1} \left[ - \left\{ - \frac{t \tau^2}{mN \Delta_2} + \frac{\beta_1^2 \tau^2 (2 \Delta_1 - \tau^2)}{m \Delta_1} \right\} + \frac{t(\bar{x}^2 + \frac{1}{3})}{N \Delta_1} \right] + \frac{\beta_1^2 \tau^2 (\bar{x}^2 + \frac{1}{3})}{\Delta_1} \right] \right] (2-19)

For fairly large \( N \),

\[
DF \approx \frac{\tau^2}{\Delta_1} \left\{ - \frac{\beta_1^2 \tau^2 (2 \Delta_1 - \tau^2)}{m \Delta_1} + \frac{\beta_1^2 \tau^2 (\bar{x}^2 + \frac{1}{3})}{\Delta_1} \right\}
= \frac{\beta_1^2 \tau^4}{\Delta_1} \left\{ \frac{-2}{m} + \frac{\tau^2}{m \Delta_1} + \frac{\bar{x}^2 + \frac{1}{3}}{\Delta_1} \right\} \] (2-20)

Hence, for any given design \( \Delta_1 \), the sign of DF largely depends upon \( m \), and the magnitude upon \( \beta_1 \) and \( \tau^2 \). For a better illustration of the above dependency, the following specific design is considered.

\[
\Delta_1^* = 1 + \tau^2 \quad \text{and} \quad \Delta_2^* = 1. \] (2-21)

The above \( \Delta \)'s are obtained when one-half of the observations \( \left( \frac{N}{2} \right) \) are allocated at \( x = -1 \) and the other half at \( x = +1 \).

Inserting Eq. (2-21) into Eq. (2-19),

\[
DF^* = \frac{\tau^2}{1 + \tau^2} \left\{ - \frac{t \tau^2}{mN} - \frac{\beta_1^2 \tau^2 (2 + \tau^2)}{m(1 + \tau^2)} \right\} + \frac{t}{3N} + \frac{\beta_1^2 \tau^2}{3(1 + \tau^2)} \right\} \] (2-22)
Figure 1 shows the relationships between $D F^*$ and $\gamma^2$ for $m = 1, 3, 5, 7,$ and $10 (\beta_1 = 1.0, \sigma^2 = 0.0001,$ and $N = 300)$. Figure 2 and 3 are the same as Figure 1 except that $N = 30$ in Figure 2, and $\sigma^2 = 0.01$ in Figure 3. It is interesting to note that $D F^*$ is not sensitive to $N$ when $N$ is large. Similar weak dependency of $D F^*$ on $\sigma^2$ may be realized by comparing Figures 1 and 3. Figure 4 shows the effect of $\beta_1$.

Based upon the figures, it is observed that:

1. for smaller $m$ (future sample size), the LS estimation procedure results in a smaller IMS for the range of $\gamma^2$ (variance of the errors in the independent variable) considered. As $m$ increases the ML estimation procedure becomes slightly better.

2. the magnitude of the rate of change of $D F = IMS_1 - IMS_m$ with respect to $\gamma^2$ increases as $\gamma^2$ increases.

In summary, the relative performance of $IMS_1$ and $IMS_m$ was examined in terms of $D F$ when $N$ is large. It is observed that $D F$ largely depends upon $m$ and $\gamma^2$, but is essentially insensitive to $N$ and $\sigma^2$. The LS estimation procedure is superior for small $m$ while the ML estimation procedure is slightly better for large $m$.

In Chapter III, the behavior of $D F$ will be re-examined by Monte Carlo simulation to see if the above properties still hold for small samples.
Figure 1. DF versus \( \gamma^2 \) when \( N=300 \) and \( \sigma^2=0.0001 \).
Figure 2. DF versus $\gamma^2$ when $N=30$ and $\sigma^2=0.0001$. 
Figure 3. DF versus $\gamma^2$ when $N=300$ and $\sigma^2=0.01$. 
Figure 4. DF versus $\tau^2$ for varying $\beta_1$ ($N = 300, \sigma^2 = 0.0001$).
CHAPTER III

MONTE CARLO SIMULATION STUDY OF
THE INTEGRATED MEAN SQUARED ERROR

The purpose of the Monte Carlo simulation experiment is to evaluate the relative performance of the LS and ML estimation procedure in terms of the IMS for small samples, and to compare the results with the large-sample expressions in Chapter II.

In view of the results of Chapter II, the hypotheses to be examined are (for small N and over the range of the parameters in Chapter II):

(H1) DF is unaffected by changes in $\sigma^2$,

(H2) the most significant factors affecting DF are m (future sample size in prediction), $\tau^2$ (variance of the errors in the independent variable), $\beta_1$ and their interactions,

(H3) a moderate change in the design of experiment for estimating $\beta$ does not affect DF,

(H4) over the sample size of $N \leq 30$, none of the effects have significant interactions with N.

Given failure to reject H4,

(H5) $N$, as a main effect, does not affect DF for $N \leq 30$. 

33
The terminology 'main effect' and 'interaction' refer to standard definitions in a factorial linear model. Since computation becomes rather extensive in the estimation of DF for a single set of values of the parameters, a coarse grid is structured for parameter values, along with a factorial experimental design. A linear model is assumed, and main effects of the factors and their interactions are assessed relative to their effect on DF.

**Design of the Monte Carlo Simulation**

Parameters, or factors considered in the design of simulation experiment are $N, m, \sigma^2, \tau^2, \beta_1$, and $D$, where $D$ is the design of the experiment for estimating $\beta'=(\beta_0, \beta_1)$. Table 2 shows the selected levels of each factor.

**Table 2. Factor Levels for the Simulation Experiment.**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>6, 18, 30</td>
</tr>
<tr>
<td>$m$</td>
<td>1, 5, 10</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.0001, 0.01</td>
</tr>
<tr>
<td>$\tau^2$</td>
<td>0.0001, 0.01</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.5, 1.0, 1.5</td>
</tr>
<tr>
<td>$D$</td>
<td>(-1, +1), (-1, 0, +1)</td>
</tr>
</tbody>
</table>
The intercept, $\beta_0$ is not included as a factor since the large sample expression for $DF = IMS_I - IMS_m$ does not involve $\beta_0$ as shown in Eq. (2-18) in Chapter II. Hence, throughout the simulation experiment $\beta_0$ is fixed to zero. The correlation coefficient $\rho$ between errors in $x$ and $y$ is set to zero, the case of interest. The two levels of $\tau^2$ are selected such that the square root of each (standard deviation) is equal to 0.5%, and 5% of the range of $x$. The levels of $\sigma^2$ considered are the same as $\tau^2$, as a percent of the range. For factor $D$, the first level $D_1 = (-1, +1)$, consists of two extreme points, and the second level, $D_2 = (-1, 0, +1)$, consists of two extreme points and the center of the range of $x$. For both designs, $N$ is equally allocated to each design point. Three levels of $m$ and $\beta_1$ were chosen because of the nonlinear effect of those variables on $DF$ as observed in Chapter II. Since the experimental question is fundamentally related to small samples, two levels of $N$ were considered insufficient, and three were chosen.

From Table 2, it is seen that a complete factorial design requires $3^3 \times 2^3$ runs. In order to economize, a one-sixth fraction of the complete factorial was selected as illustrated in Table 3. This experiment will result in an inability to estimate interactive effects of $m$, $\beta_1$, and $N$ as well as those of $\sigma^2$, $\tau^2$, and $D$. However, this seemed the most effective experiment, given that the total number of
Table 3. Design of the Monte Carlo Simulation Experiment

<table>
<thead>
<tr>
<th></th>
<th>$\beta_1 = 0.5$</th>
<th>1.0</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N=6$</td>
<td>$m=1$</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Legend: T1

Legend: T2

Legend: T

Legend: S

$D$: $D_1 = -1, +1$
$D_2 = -1, 0, +1$

$\sigma^2$: $S_1 = 0.0001$
$S_2 = 0.01$

$\gamma^2$: $T_1 = 0.0001$
$T_2 = 0.01$

Legend: $\square$: not run
combinations was not to exceed 50(an external constraint). Further, since the objective is a comparison of IMS\(_1\) and IMS\(_m\), the same randomly generated data(for a given set of parameter values) is used to estimate both IMS\(_1\) and IMS\(_m\). This correlation between the quantities reduces the variance of DF, and hence, increases the power of the tests. Further, it substantially reduces the cost of simulation.

**Computer Program**

Appendix B includes the listing of the simulation program, definitions of scalar and array variables, and flowcharts.

The FORTRAN language is used, and all computation is done on an Amdahl 470 system. For certain statistical computations, the IMSL(International Mathematical and Statistical Libraries) subroutines are used.

Briefly the entire program consists of the following three major steps:

1. generation of N sets of multivariate normal deviates,
2. calculation of the LS and ML estimate of \( \beta \)(for the same data generated in 1), and
3. Repetition of step 1 and 2 'NSMPL' times to estimate the biases and variances of the LS and ML estimates. Calculation of IMS\(_1\), IMS\(_m\), and \( DF = IMS\(_1\) - IMS\(_m\) \).

The above procedures are repeated 'NEXP' times for a given
'run' (a given set of independent variable values or factors), and then the mean and standard deviation of DF are calculated in order to construct a confidence interval for the mean DF for that set of values of the factors. Table 4 shows more detailed functions of the main and subprograms while Figure 5 illustrates their interactions.

The IMSL subroutine GGNSM generates k-variate normal deviates, \( u' = (u_1, \ldots, u_k) \) with mean \( \mu \) and covariance \( \Sigma \) based upon the following formula:

\[
  u = Tz + \mu
\]

where \( z \) is a \( k \)-variate normal deviate with mean 0 and covariance \( I_k \), and

\[
  T = \text{a unique lower triangular matrix such that } \Sigma = TT'.
\]

To compute the IMS for either estimation procedure, Eq. (A-18) is used together with the estimated bias and variance of either estimator in step 3.

In the simulation, the values of NSMPL and NEXP are set to 20 and 10, respectively.

Analysis of the Results

All the simulation results are summarized in Appendix C.1. A 99% confidence interval, \((L, U)\) for the mean DF was constructed based upon the t-distribution with 9 degrees of
Table 4. Functions of Main and Subprograms.

Main Program

(a) read in input data
(b) calculation of: auxiliary quantities; IMS₁, IMSₘ, and DF; the mean and standard deviations of DF; and large-sample IMS₁, IMSₘ, and DF.
(c) report the results

Subprograms

LSQ: calculation of the LS estimate of \( \beta \).
MLE: calculation of the ML estimate of \( \beta \).
CALS: calculation of the pooled sample covariance.
EGV: identification of the minimum eigenvalue and scaling of the corresponding eigenvector.
MV: calculation of mean and variance.
CIMS: calculation of the IMS.

IMSL Subroutines

EIGZF: calculation of eigenvalues and eigenvectors of the system, \( Ax = Bx \).
LINV1F: matrix inversion.
GGNSM: generation of a set of multivariate normal deviates.
VMULFP, VMULFM, VMULFF: matrix multiplication.
Figure 5. Interactions between Main and Subprograms.
freedom. Then, the results of each of the entire 36 runs were classified into the following three categories:

- : \( U < 0 \),
0 : \((L, U)\) includes 0, and
+ : \( L > 0 \).

Based upon the above classification, Table 5 is constructed. It is observed from the table that:

1. No '+' is observed, indicating that, for the values of the parameters considered, the ML estimation procedure is nowhere better (99% confidence) than the LS estimation procedure.

2. For \( m=1 \), the effect of \( \gamma^2 \) on DF is revealing. That is, the two estimation procedures are not practically different when \( \gamma^2=0.0001 \). However, for \( \gamma^2=0.01 \) the LS estimation procedure is definitely superior.

3. For \( m=10 \), all 12 runs correspond to '0', implying that one estimation procedure may not be preferred to the other in this case. A similar phenomenon is observed for \( m=5 \).

4. The above phenomena do not depend upon \( \sigma^2 \) nor the design of experiment considered.
Table 5. Results of the Monte Carlo Simulation Experiment

<table>
<thead>
<tr>
<th></th>
<th>$\beta_1 = 0.5$</th>
<th>1.0</th>
<th>1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=6</td>
<td>m=1</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Legend:
- $D_1$: D1 = -1, +1
- $D_2$: D2 = -1, 0, +1
- $S_1$: S1 = 0.0001
- $S_2$: S2 = 0.01
- $r^2$: T1 = 0.0001
- $T^2$: T2 = 0.01

$: not run
Furthermore, to compare the results of Monte Carlo simulation for small samples with large-sample expressions, additional runs, as shown in Table 6, were made for N=30, D=(-1, +1), $\beta_1=1.0$, and $\sigma^2=0.0001$. The DF values obtained are plotted on Figure 6, where continuous curves represent large-sample approximation. From Figure 6, it is clear that the large-sample approximation underestimates the superiority of the LS estimation procedure.

Table 6. Results of Additional Runs for N=30 to Compare Simulation and Large-Sample Approximation.

<table>
<thead>
<tr>
<th>m</th>
<th>$\gamma^2$</th>
<th>DFx10^{-4}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.01</td>
<td>-1.74451</td>
</tr>
<tr>
<td>5</td>
<td>.0001</td>
<td>-.00003</td>
</tr>
<tr>
<td>5</td>
<td>.01</td>
<td>-.20572</td>
</tr>
<tr>
<td>10</td>
<td>.0001</td>
<td>.00002</td>
</tr>
<tr>
<td>10</td>
<td>.01</td>
<td>-.10723</td>
</tr>
</tbody>
</table>

To validate the behavior of the simulated DF for larger N, another set of runs were made for N=300(D, $\beta_1$, and $\sigma^2$ were the same as in Table 6). The results are summarized
Figure 6. Asymptotic versus Simulated DF Values for N=30 
($\beta_1 = 1.0$ and $\sigma^2 = 0.0001$).
in Table 7, and plotted in Figure 7. As expected, the large-sample expressions and the simulation results are closer than those previously obtained for N=30.

Table 7. Results of Additional Runs for N=300 to Compare Simulation and Large-Sample Approximation.

<table>
<thead>
<tr>
<th>m</th>
<th>$\gamma^2$</th>
<th>DFx10^-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0001</td>
<td>-0.00017</td>
</tr>
<tr>
<td>1</td>
<td>0.01</td>
<td>-1.60489</td>
</tr>
<tr>
<td>5</td>
<td>0.0001</td>
<td>0.00001</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>-0.10666</td>
</tr>
<tr>
<td>10</td>
<td>0.0001</td>
<td>0.00001</td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>0.06571</td>
</tr>
</tbody>
</table>

In summary, the simulation results for small samples generally agree with large-sample expressions in the sense that m and $\gamma^2$ are again important factors to evaluate the relative performance of LS and ML estimation procedures for small samples. However, one important difference between the two is that the superiority of the ML estimation procedure for large m in large-sample expressions is not observed in the simulation experiment.

Based upon the general analysis of the simulation results, for small m and large $\gamma^2$, the LS estimation
Figure 7. Asymptotic versus Simulated DF Values for N=300 ($\beta_1=1.0$ and $\sigma^2=0.0001$).
procedure is definitely preferred, and for the other combinations of parameter values, the LS and ML estimation procedures are not statistically different. This is undoubtedly due to a bias in small-sample ML estimator (such a small-sample bias was not calculable analytically). Furthermore, the comparison made between the large-sample expressions and the simulation results indicates possible discrepancies between the two.

These questions and the original hypotheses of this chapter are now addressed in a more detailed manner using analysis of variance.

**Linear Model Analysis**

Since a one-sixth fraction of the complete $3^3 \times 2^3$ factorial experiment was run, the following set of (orthogonal polynomial) variables ($NL, NQ, ML, MQ, BL, and BQ$) are defined for a more detailed analysis of variance:

<table>
<thead>
<tr>
<th></th>
<th>$NL$</th>
<th>$NQ$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N=6$ :</td>
<td>$-1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$=18$ :</td>
<td>$0$</td>
<td>$-2$</td>
</tr>
<tr>
<td>$=30$ :</td>
<td>$1$</td>
<td>$1$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$ML$</th>
<th>$MQ$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m=1$ :</td>
<td>$-1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$=5$ :</td>
<td>$0$</td>
<td>$-2$</td>
</tr>
<tr>
<td>$=10$ :</td>
<td>$1$</td>
<td>$1$</td>
</tr>
</tbody>
</table>
The 'L', and 'Q' refer to the 'linear' and 'quadratic' effects of each of the equi-spaced variables (e.g., refer to Hicks, 1973).

The linear model used for the analysis of variance includes: the main effect terms of NL, NQ, ML, MQ, BL, BQ, \( \gamma^2 \), \( \sigma^2 \), and D; and their estimable second-order interaction terms. Three or higher order interactions are considered negligible, and pooled as remainder, or error term. The dependent variable, DF is multiplied by \( 10^4 \) for convenience.

The results of analysis of variance indicated that none of the effects considered were significant at the 1% level. However, at the 5% level, several effects were significant, and are listed in Table 8. Note that the five hypotheses stated on page 33 are not rejected at the 5% level.

In general, the result of analysis of variance conforms to the findings in the previous section. However, the model may not be adequate to predict DF since the variance of error (0.95) seems somewhat large. Hence, an attempt is made to fit DF to the large-sample expression in Eq. (2-19) so that
Table 8. Summary of Analysis of Variance (Dependent Variable = DF).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>p &gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau^2$</td>
<td>1</td>
<td>10.21</td>
<td>.011</td>
</tr>
<tr>
<td>ML $\times \tau^2$</td>
<td>1</td>
<td>6.46</td>
<td>.031</td>
</tr>
<tr>
<td>ML</td>
<td>1</td>
<td>6.38</td>
<td>.032</td>
</tr>
<tr>
<td>BL</td>
<td>1</td>
<td>5.26</td>
<td>.047</td>
</tr>
<tr>
<td>BL $\times \tau^2$</td>
<td>1</td>
<td>5.14</td>
<td>.049</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>7.63</td>
<td></td>
</tr>
</tbody>
</table>

$\text{MSE} = .95$
the discrepancies between the two may be examined. For this purpose, define

\[ \text{DELTA} = \text{DF} - \text{DF}_c \]

where

\[ \text{DF} = \text{observed IMS}_1 - \text{IMS}_m \text{ by the simulation, and} \]
\[ \text{DF}_c = \text{calculated IMS}_1 - \text{IMS}_m \text{ by the large-sample expression.} \]

Then, the same linear model was fit to DELTA, and the effects which are significant at 1% level are listed in Table 9.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>p &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma^2 )</td>
<td>1</td>
<td>1.06</td>
<td>0.003</td>
</tr>
<tr>
<td>BL</td>
<td>1</td>
<td>0.76</td>
<td>0.007</td>
</tr>
<tr>
<td>BL ( \times \gamma^2 )</td>
<td>1</td>
<td>0.72</td>
<td>0.008</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>0.48</td>
<td></td>
</tr>
</tbody>
</table>

\( \text{MSE} = 0.06 \)

Note that the variance of error is now 0.06 as compared to 0.95 when DF is directly fitted to the linear model. It may be interesting to note that the unexplained portion of DF
(DELTA) is affected by factor BL, $\tau^2$ and their interaction. Graphically, this is illustrated in Figure 8.

![Graph illustrating the effect of $\beta_1$ and $\tau^2$ on DELTA.]

In other words, the large-sample expression may be fairly accurate for predicting small-sample DF when $\tau^2$ is small. However, if $\tau^2$ is large the large-sample expression underestimates the superiority of the LS estimation procedure depending upon the magnitude of $\beta_1$. This was also illustrated in Figure 6 for certain values of parameters.
In conclusion, $m$, $\gamma^2$, and $\beta_1$ are found to be the most important factors to determine the relative performance of the two estimation procedures for small samples. For small $m$, the LS estimation procedure is preferred for a modest change in the design of experiment as well as changes in the magnitude of $\sigma^2$. As $m$ increases, the two estimation procedures are not substantially different. The factor $\beta_1$ affects the magnitude of DF, but not the superiority of the LS estimation procedure. For the purpose of predicting DF for small sample sizes, the large-sample expression is fairly accurate when $\gamma^2$ is small, however, when $\gamma^2$ and $\beta_1$ are large one must realize that the 'superiority' of IMS$_1$ may be underestimated in this case. Indeed, for the range of values of the parameters, including small samples, least squares procedures are nowhere inferior to those of maximum likelihood.

In passing, it is worth mentioning that the case when $\Omega$ is known was also simulated, and similar results were obtained (refer to Appendix C.2 for the simulation results, and Appendix D for the ML solution when $\Omega$ is known). This implies that the ML estimation procedure does not become better given knowledge of $\Omega$. 
CHAPTER IV

OPTIMAL DESIGN USING ASYMPTOTIC EQUATIONS

In Chapter II and III the LS and ML estimation procedures were compared with respect to the problem of choosing an appropriate analysis technique. That is, given an experimental design for estimating $\beta$, the relative performance of the two estimation procedures were compared in terms of the IMS. In this chapter, the optimal design of the experiment in the estimation of $\beta$ is considered for a given estimation procedure (either LS or ML estimation).

The 'optimal' design of the experiment is defined as the selection of the values of the independent variables so that the (asymptotic) IMS may be minimized. The large-sample expressions developed in Chapter II will be used for this purpose.

First, the two-variable model will be considered for both LS and ML estimation, and then the general linear model for LS estimation only (asymptotic expressions for the IMS in the case of ML estimation are not available for the general linear model).
The Two-Variable Model Using LS Estimation

From Eq. (2-12),

\[\text{IMS}_1 \sim \frac{t \tau^2}{mN \Delta_1} + \frac{\beta_1^2 \tau^2}{m} \left(1 - \frac{\tau^2}{\Delta_1}\right)^2\]

\[+ \frac{t}{N} + \frac{t}{N \Delta_1} \left(\bar{x}^2 + \frac{1}{3}\right)\]

\[+ \frac{\beta_1^2 \tau^4}{\Delta_1^2} \left(\bar{x}^2 + \frac{1}{3}\right)\]

where

\[\Delta_1 = \frac{\sum x_i^2}{N} + \tau^2 - \bar{x}^2, \text{ and}\]

\[t = \sigma^2 + \beta_1^2 \tau^2.\]

Then, the problem of the optimal design of the experiment is to select \(\{x_i, i = 1, 2, \ldots, N\}\), or equivalently, to determine \(\Delta_1\) such that IMS\(_1\) is minimized.

Since \(x\) is scaled such that \(-1 \leq x \leq 1\),

\[\tau^2 \leq \Delta_1 \leq 1 + \tau^2\]  \hspace{1cm} (4-1)

Note that for any given value of \(\Delta_1\) within the range in Eq. (4-1), there exist many different sets of \(\{x_i\}_{i=1}^N\) (and hence, many different values of \(\bar{x}\)) which result in that
given value. Especially, there always exists a set \( \{ x_i \}_{i=1}^N \) for which \( x = 0 \) for a given value of \( \Delta_1 \). Then, partial differentiation of IMS\(_1\) with respect to \( x^2 \) for any given value of \( \Delta_1 \) yields:

\[
\frac{\partial \text{IMS}_{1\text{g}}}{\partial (x^2)} \bigg|_{\text{given } \Delta_1} = \frac{t}{N \Delta_1} + \frac{\beta_1^2 \tau^4}{\Delta_1^2} > 0 \quad (4-2)
\]

Eq. (4-2) suggests that, for any given \( \Delta_1 \), the IMS\(_1\) is decreased as \( x^2 \) decreases. Hence, \( x^2 \) can be eliminated from the minimization procedure by setting \( x^2 \) to zero.

Inserting \( x^2 = 0 \) to IMS\(_1\),

\[
\text{IMS}_{10} = \text{IMS}_{1\text{g}}(x^2 = 0)
\]

\[
\sim \frac{t \tau^2}{mN \Delta_{10}} + \frac{\beta_1^2 \tau^2}{m} \left(1 - \frac{\tau^2}{\Delta_{10}}\right)^2
\]

\[
+ \frac{t}{N} + \frac{t}{3N \Delta_{10}} + \frac{\beta_1^2 \tau^4}{3 \Delta_{10}^2} \quad (4-3)
\]

where

\[
\Delta_{10} = \frac{\sum x_i^2}{N} + \tau^2 \quad (4-4)
\]
Considering only the terms in $\Delta_{10}$ in Eq. (4-3),

$$\text{IMS}_{10} \propto \frac{t \tau^2}{mN \Delta_{10}} + \frac{\beta_1^2 \tau^2}{m} \left( \frac{2 \tau^2}{\Delta_{10}} + \frac{\tau^4}{\Delta_{10}^2} \right)$$

$$+ \frac{t}{3N \Delta_{10}} + \frac{\beta_1^2 \tau^4}{3 \Delta_{10}^2}$$

$$= \frac{1}{\Delta_{10}} \left( \frac{t \tau^2}{mN} - \frac{2 \beta_1^2 \tau^4}{m} + \frac{t}{3N} \right)$$

$$+ \frac{1}{\Delta_{10}^2} \left( \frac{\beta_1^2 \tau^6}{m} + \frac{\beta_1^2 \tau^4}{3} \right). \quad (4-5)$$

Let

$$f = \frac{t \tau^2}{mN} - \frac{2 \beta_1^2 \tau^4}{m} + \frac{t}{3N},$$

and

$$g = \frac{\beta_1^2 \tau^6}{m} + \frac{\beta_1^2 \tau^4}{3}.$$ 

Then,

$$\text{IMS}_{10} \propto \frac{f}{\Delta_{10}} + \frac{g}{\Delta_{10}^2}. \quad (4-6)$$

Note that $g>0$. Hence, if $f \geq 0$, then the $\text{IMS}_{10}$ is minimized when $\Delta_{10}$ is maximized. However, $f \geq 0$ implies that:

$$\frac{t \tau^2}{mN} - \frac{2 \beta_1^2 \tau^4}{m} + \frac{t}{3N} \geq 0$$
\[
\frac{t(3 \tau^2 + m)}{3mN} \geq \frac{2 \beta_1^2 \tau^4}{m}
\]

\[N \leq \frac{t(3 \tau^2 + m)}{6 \beta_1^2 \tau^4}\]

Since \( t = \sigma^2 + \beta_1^2 \tau^2 \),

\[N \leq \frac{(\sigma^2 + \beta_1^2 \tau^2)(3 \tau^2 + m)}{6 \beta_1^2 \tau^4}\]

\[N \leq N_0 = \frac{1}{\tau^2} \left( 1 + \frac{\sigma^2}{\beta_1^2 \tau^2} \right) \left( \frac{m}{6} + \frac{\tau^2}{2} \right) \quad (4-7)\]

From Eq. (4-4), \( \Delta_{10} \) attains its maximum when one half (\( \frac{N}{2} \)) of \( x_i \)'s are allocated at \(-1\), and the other half at \(+1\) so that \( \sum x_i^2 / N = 1 \). Hereafter, this design will be called as an extreme-point design, and denoted by:

\[\xi^* = \begin{pmatrix} -1 & +1 \\ \frac{N}{2} & \frac{N}{2} \end{pmatrix} \quad (4-8)\]

If \( f < 0 \), that is, if \( N > N_0 \), the graph of IMS_{10} looks like Figure 9. Note that IMS_{10} has a unique minimum at \( \Delta_{10} = -2g/f \), which is always greater than \( \tau^2 \). Hence, it is important to check if
If the inequality (4-9) holds, then the $\text{IMS}_{10}$ is minimized when $\Delta_{10}$ is maximized, otherwise, $\Delta_{10} = -\frac{2g}{f}$ is optimal.

\[
-\frac{2g}{f} > 1 + \gamma^2.
\] (4-9)

Note that

\[
-\frac{2g}{f} - (1 + \gamma^2)
\]

\[
= \frac{-2N \beta^2 \gamma^4 (m-3) - t(1 + \gamma^2)(m+3\gamma^2)}{3t\gamma^2 - 6N \beta_1^2 \gamma^4 + mt}.
\] (4-10)
The denominator of Eq.(4-10) is negative since \( f < 0 \).
Hence, if \( m \geq 3 \), the inequality (4-9) holds, and therefore, the extreme-point design, \( \xi^* \) is optimal. Consider the case when \( m < 3 \). If the numerator in Eq.(4-10) is negative, that is, if

\[
N \leq N_1 = \frac{3}{3-m} \left( 1 + \frac{1}{\gamma^2} \right) \left( 1 + \frac{\sigma^2}{\beta_1 \gamma^2} \right) \left( \frac{m}{6} + \frac{\gamma^2}{2} \right) \quad (4-11)
\]

then, \( \xi^* \) is again optimal. However, if \( N > N_1 \), then the \( \text{IMS}_{10} \) is minimized at \( \Delta_{10} = -2g/f \), which is less than \( 1 + \gamma^2 \).

Note that for small \( \gamma^2 \), \( N_0 < N_1 \). Hence, the whole results in this section can be summarized as follows:

1. if \( m \geq 3 \), then \( \xi^* \) if optimal for all \( N \).
2. if \( m < 3 \), and \( N \leq N_1 \), then \( \xi^* \) is optimal.
3. if \( m < 3 \) and if \( N > N_1 \), then there exists a better design than \( \xi^* \).

For the case of \( m < 3 \) and \( N > N_1 \), it is interesting to note that

\[
\lim_{N \to \infty} \left( \Delta_{10} = -\frac{2g}{f} \right) = \frac{m}{3} + \gamma^2. \quad (4-12)
\]

Eq.(4-12) implies that, as \( N \) becomes arbitrarily large, the optimal allocation of \( x_i \)'s must satisfy:

\[
\frac{\sum x_i^2}{N} = \frac{m}{3}.
\]
Two-Variable Model Using ML Estimation

From Eq. (2-15)

\[ IMS_m \sim \frac{t \tau^2}{mN \Delta_2} + \frac{\beta_1^2 \tau^2}{m} + \frac{t}{N} + \frac{t(\bar{x}^2 + \frac{1}{3})}{N \Delta_2} \]

where

\[ \Delta_2 = \frac{\sum x_i^2}{N} - \bar{x}^2. \]

Then,

\[ \frac{\partial IMS_m}{\partial (\bar{x}^2)} \bigg|_{\text{given } \Delta_2} = \frac{2t}{N \Delta_2} > 0. \]

By the similar argument as in the LS estimation procedure, \( \bar{x}^2 \) can be eliminated from IMS to yield:

\[ IMS_{m0} = IMS_m(\bar{x}^2 = 0) \]

\[ \sim \frac{t \tau^2}{mN \Delta_{20}} + \frac{\beta_1^2 \tau^2}{m} + \frac{t}{N} + \frac{t}{3N \Delta_{20}} \]  

(4-13)

where

\[ \Delta_{20} = \frac{\sum x_i^2}{N}. \]  

(4-14)

From Eq. (4-13), it is obvious that IMS\(_{m0}\) is minimized when \( \Delta_{20} \) is maximized. From Eq. (4-14), \( \Delta_{20} \) attains its maximum when the end-point design, \( \xi^* \) is selected.
The General Linear Model Using LS Estimation

Since the large-sample properties of the ML estimator for the general linear model is not known, only the LS estimation procedure is considered.

For the linear model described in Appendix A, assume that the first k independent variables are measured exactly, and the remaining l measured with errors (p = k + l). Then, the covariance matrix $\Sigma$ has the following structure:

$$
\Sigma = \begin{bmatrix}
0 & \cdots & 0 \\
\cdots & \gamma_{k+1}^2 & \cdots \\
0 & \cdots & \gamma_{k+1}^2 \\
\end{bmatrix}^{k+1} \begin{bmatrix}
\gamma_{k+1}^2 \\
\cdots \\
\gamma_{k+1}^2 \\
\end{bmatrix}^{l}
$$

Consider a class of designs which result in a diagonal $XX'(e.g.,$ balanced designs) such that

$$
\frac{X'X}{N} = \begin{bmatrix}
1 \\
\sum_{j=1}^{N} x_{1j}^2 / N \\
\vdots \\
\sum_{j=1}^{N} x_{k+1,j}^2 / N \\
\end{bmatrix}
$$

Let $z_i = \sum_{j=1}^{N} x_{ij}^2$ for $i = 1, 2, \ldots, k+1$. 
Then,

\[ \frac{X'X}{N} + \sum = \begin{bmatrix} 1 & \vdots & \vdots & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & 1 & \vdots \\ \vdots & \vdots & \vdots & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_k \\ z_{k+1} + \gamma_{k+1}^2 \\ \vdots \\ \vdots \\ z_{k+1} + \gamma_{k+1}^2 \end{bmatrix} \]

Inserting these results to Eq. (2-2), (2-3), and (A-18) yields:

\[
\text{IMS}_1 \propto \frac{t}{mN} \left[ \sum_{i=k+1}^{k+1} \frac{\gamma_i^2}{z_i + \gamma_i^2} \right. \\
+ \frac{1}{m} \left\{ -2 \sum_{k+1}^{k+1} \frac{\beta_i^2 \gamma_i^4}{z_i + \gamma_i^2} + \sum_{k+1}^{k+1} \frac{\beta_i^2 \gamma_i^6}{(z_i + \gamma_i^2)^2} \right\} \\
+ \frac{1}{3} \sum_{k+1}^{k+1} \frac{\beta_i^2 \gamma_i^4}{(z_i + \gamma_i^2)^2} \\
+ \frac{t}{3N} \left\{ \sum_{1}^{k} \frac{1}{z_i} + \sum_{k+1}^{k+1} \frac{1}{z_i + \gamma_i^2} \right\} \right] \\

Note that IMS$_1$ is separable with respect to the index i.
Let IMS$_{1i}$ be the ith component of IMS$_1$. Then,
\[ \text{IMSL}_i \propto \frac{1}{z_i} \quad \text{for } i = 1, 2, \ldots, k, \text{ and} \]

\[ \propto \frac{t}{mN} \frac{\tau_i^2}{\Delta_i} + \frac{\beta_i^2 \tau_i^2}{m} \frac{2 \tau_i^2}{\Delta_i} + \frac{\tau_i^4}{\Delta_i^2} \]

\[ + \frac{t}{3N\Delta_i} + \frac{\beta_i^2 \tau_i^4}{3\Delta_i^2} \]

for \( i = k + 1, \ldots, k + l \), where \( \Delta_i = z_i + \tau_i^2 \).

For \( i = 1, 2, \ldots, k \), the IMSL is minimized by maximizing \( z_i \), which corresponds to the extreme point design, \( \xi^* \).

For \( i = k + 1, \ldots, k + l \), IMSL has exactly the same form as Eq.(4-5), indicating that all of the results for the two-variable model are directly applicable.

In summary, the optimal design of experiment for estimating \( \beta \) has been considered based upon large-sample expressions for the IMS. Unless \( m \) is small and \( N \) is very large, the extreme-point design is preferred.

Although there exist uncertainties in determining \( \{ x_i \} \), the extreme-point design may be easily implemented when \( \{ x_i \} \) can be at least rank-ordered based upon their expected values.
CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The relative performance of the least squares (LS) and the maximum likelihood (ML) estimation procedure has been investigated for the two-variable linear model with measurement errors in both variables. The integrated mean squared error (IMS) related to the calibration prediction problem was utilized as a criterion of comparison. For the large-sample case, analytic expressions for the IMS were developed in Chapter II, and the relative performance of the two estimation procedures was evaluated in terms of the model parameters. In addition, it was shown that the IMS for a given estimation procedure can be decomposed into several meaningful components, which could be utilized for analysis purposes. For the small-sample case, the two estimation techniques were compared using Monte Carlo simulation in Chapter III. A reasonable set of the parameter levels was considered based upon the calibration problem of two commensurate measurements, and a fraction of a complete factorial was run to economize the simulation experiment,
along with a variance reduction technique adopted by applying the two estimation techniques to the same, simulated data. Further, the simulation results and the large-sample expressions were compared with each other from a qualitative as well as quantitative point of view. For a qualitative analysis, the simulation results were classified into three categories based upon the confidence interval of \( DF = IMS_1 - IMS_m \), and were compared to the general behavior of the large-sample \( DF \). For a quantitative analysis, the large-sample expression of \( DF \) was fit to the simulated \( DF \) values by analysis of variance technique to observe conformance or discrepancies between the two.

The optimal design of the experiment for the purpose of prediction was also considered for a given estimation technique. The IMS was used as a criterion of the design. In addition, the principle of optimal design was extended to the general linear model when the LS estimation technique is used.

Finally, it is worth mentioning that the large-sample expression of the IMS for any given estimation technique has been developed for the general linear models, with the hope that it may be utilized by any future research in this area.
Research Findings and Conclusions

With regard to the relative performance of the LS and ML estimation procedure, it was found that:

1. For large samples, the sign(quality) of $\text{DF} = \text{IMS}_1 - \text{IMS}_m$ largely depends upon $m$ (the number of future measurements). For small $m$, the LS is superior to the ML estimation procedure (i.e., negative DF), and as $m$ increases, the ML estimation procedure becomes slightly better (i.e., positive DF). In addition, the magnitude (quantity) of DF largely depends upon $\beta_1$ (the slope of the linear relationship) and $\gamma^2$ (variance of errors in the independent variable). As $\beta_1$ or $\gamma^2$ increases, the DF increases in absolute value. These relationships do not depend upon $N$ (the total number of observations) nor $\sigma^2$ (variance of errors in the dependent variable).

2. For small samples, the simulation results generally conform to the large-sample expression. In particular, the superiority of the LS estimation procedure is observed for small $m$ and large $\gamma^2$ regardless of $\beta_1$, $N$, $\sigma^2$, or the design. As $m$ is increased, the two estimation techniques are not statistically different (type I error of .01).

3. The large-sample expression for DF is a fair approximation to the small-sample behavior of DF. However, the large-sample value tends to underestimate the superiority of the LS estimation procedure for large $\beta_1$ and $\gamma^2$. 
With regard to the optimal design of the experiment, it was shown that:

1. Based upon the large-sample expression of the IMS, the extreme-point design is preferred for the LS estimation procedure over a reasonable range of m and N. When the ML estimation technique is used, the extreme-point design is always optimal.

2. The findings in (1) can be extended to the optimal design of the general linear model with measurement errors in the independent variables.

Based upon the above findings, two major conclusions may be drawn with regard to the 'analysis' and 'design' strategy of the problem. For the analysis strategy, the LS procedure is preferred to the ML over a reasonable range of the parameter values. Considering the simplicity in addition, one may safely use the LS procedure for the purpose of prediction, especially when m is small. When DF is to be predicted by the large-sample expression for large $\beta_1$ and $\gamma^2$, it must be realized that the superiority of the LS procedure could even be underestimated. Once the estimation technique is selected, an 'average' tolerance interval for a predicted value may be constructed based upon the IMS components in Table 1.

For the design strategy, the extreme-point design is recommended for both estimation techniques when N is fairly
large and \( m \) is not small. For a proper implementation of the above, the true values of the experimental quantities must be at least rank-ordered since there exist uncertainties in the independent variable. This could be achieved if the past history exists. Otherwise, a preliminary experiment could be performed for this purpose.

**Recommendations for Future Research**

It is recommended that future research be pursued in the following directions:

1. The criterion, IMS adopted in this research could be modified to the 'weighted' IMS to reflect the importance of points within the region of interest. For this purpose, a weight function could be defined as Helms(1974) did in a different regression problem.

2. Other estimation techniques such as method of grouping (e.g., Wald, 1940; Bartlett, 1949) could be compared to the LS or ML procedure.

3. The relative performance of estimation techniques could be evaluated in a different context. For instance, the power of the hypothesis tests related to the unknown \( \beta_0 \) and \( \beta_1 \) may be compared either by simulation or, possibly, by the large-sample approximation.
Appendix A

Integrated Mean Squared Error for General Linear Model

Model and Estimation

Consider the following linear model:

\[ \eta = X\beta + \epsilon \]  

(A-1)

where \( \eta' = (\eta_1, \eta_2, \ldots, \eta_N) \),

\[
X = \begin{bmatrix} x' \\ x_2' \\ \vdots \\ x_N' \end{bmatrix} = \begin{bmatrix} 1, x_{11}, \ldots, x_{p1} \\ 1, x_{12}, \ldots, x_{p2} \\ \vdots \\ 1, x_{1N}, \ldots, x_{pN} \end{bmatrix},
\]

\( \beta' = (\beta_0, \beta_1, \ldots, \beta_p) \),

\( \epsilon' = (\epsilon_1, \epsilon_2, \ldots, \epsilon_N) \),

\[ \text{E}(\epsilon) = 0 \text{ and } \text{Var}(\epsilon) = \sigma^2 I. \]

In the classical theory of linear models, it is assumed that all elements of \( X \) are measured exactly. Then, the least squares estimator \( \hat{b} \) of \( \beta \) is given by:
\( b = (X'X)^{-1}X'\eta, \)

which is an unbiased, minimum-variance estimator.

However, in many instances, \( X \) cannot be measured precisely, but only observations \( W \) made, where

\[
W = X + \varepsilon
\]

and

\[
\varepsilon = \begin{bmatrix}
e_1' \\
e_2' \\
\vdots \\
e_N'
\end{bmatrix} = \begin{bmatrix}
0, e_{11}, \ldots, e_{p1} \\
0, e_{12}, \ldots, e_{p2} \\
\vdots & \ddots & \vdots \\
0, e_{1N}, \ldots, e_{pN}
\end{bmatrix},
\]

\( E(\varepsilon_i) = 0, \)

\( \text{Var}(\varepsilon_i) = \sum = \text{diag}(0, \tau_1^2, \ldots, \tau_p^2) \) with the

\( \varepsilon_i \)'s pairwise independent.

Hence,

\[
W = \begin{bmatrix}
w_1' \\
w_2' \\
\vdots \\
w_N'
\end{bmatrix} = \begin{bmatrix}
x_1' + e_1' \\
x_2' + e_2' \\
\vdots & \ddots & \vdots \\
x_N' + e_N'
\end{bmatrix} = \begin{bmatrix}
1, x_{11} + e_{11}, \ldots, x_{p1} + e_{p1} \\
1, x_{12} + e_{12}, \ldots, x_{p2} + e_{p2} \\
\vdots & \ddots & \vdots \\
1, x_{1N} + e_{1N}, \ldots, x_{pN} + e_{pN}
\end{bmatrix}
\]
Then, based upon \( W \), an estimator \( \hat{b} \) of \( \beta \) may be found by a chosen estimation procedure.

**Prediction**

In the future, one may wish to predict \( y = E(\mathcal{Y}) \) at an unknown \( x' = (1, x_1, \ldots, x_p) \in \mathcal{X} \), where \( \mathcal{X} \) is the region of interest. Again, it is assumed that the future \( x \) cannot be measured precisely, but one only observes a sequence of \( 1 \times (p+1) \) vectors, \( \{w_j\} \) where

\[
w_j = x_j' + e_j ; \quad j = 1, 2, \ldots, m.
\]

Then, the predicted value of \( E(y) \) is given by:

\[
\hat{y} = \bar{w}'b \tag{A-5}
\]

where

\[
\bar{w} = \frac{1}{m} \sum_{j=1}^{m} w_j = x + \bar{e}.
\]

In the following sections, the property of \( \hat{y} \) over \( \mathcal{X} \) is examined in terms of the integrated mean squared error. The following results hold for any estimation procedure as long as the procedure generates an estimator \( b \) which has properly defined expectation and variance.
Mean Squared Error of $\hat{y}$

Since the true value of $\hat{y}$ is $x'\beta$, the mean squared error (MSE) of $\hat{y}$ is, by definition,

$$MSE(y) = E(\hat{y} - x'\beta)^2$$

$$= E \left( \{ \hat{y} - E(\hat{y}) \} + \{ E(\hat{y}) - x'\beta \} \right)^2$$

$$= E \{ \hat{y} - E(\hat{y}) \}^2 + E \{ E(\hat{y}) - x'\beta \}^2$$

$$+ 2 E \left[ \{ \hat{y} - E(\hat{y}) \} \{ E(\hat{y}) - x'\beta \} \right]$$

$$= Var(\hat{y}) + \{ E(\hat{y}) - x'\beta \}^2 \quad (A-7)$$

From Eq. (A-6),

$$E(\hat{y}) = E(\bar{w}'b)$$

$$= E \left\{ (x+\bar{e})'b \right\}$$

$$= x'E(b) \quad (A-8)$$

assuming that the future measurement error, $\bar{e}$ and $b$ are independent. Define the bias introduced in $b$ as

$$\delta = E(b) - \beta. \quad (A-9)$$

Then, Eq. (A-7) and (A-8) yield

$$MSE(\hat{y}) = Var(\hat{y}) + (x'\delta)^2. \quad (A-10)$$
However, by definition and from Eq. (A-8),

\[
\text{Var}(\hat{y}) = E \{ \hat{y} - E(\hat{y}) \}^2
\]

\[
= E(\hat{y}^2) - \{ E(\hat{y}) \}^2
\]

\[
= E \{ (x+\bar{e})'b \}^2 - \{ x'E(b) \}^2 \quad (A-11)
\]

To evaluate \( E \{ (x+\bar{e})'b \}^2 \), the following theorem is needed.

**Theorem.** Let \( u \) and \( v \) be \( q \)-variate, independent random vectors such that

\[
u \sim (\bar{U},R), \quad R = (r_{ij}), \text{ and} \\
v \sim (\bar{V},S), \quad S = (s_{ij}).
\]

Then,

\[
E(u'v)^2 = \text{tr } RS + \bar{v}'R\bar{v} + \bar{u}'S\bar{u} + (\bar{u}'\bar{v})^2.
\]

**Proof.**

\[
E(u'v)^2 = E(\sum_{i=1}^{q} \sum_{j=1}^{q} u_i v_i u_j v_j)
\]

\[
= \sum_{i=1}^{q} \sum_{j=1}^{q} E(u_i u_j) E(v_i v_j)
\]

\[
= \sum_{i=1}^{q} \sum_{j=1}^{q} (r_{ij} + \bar{u}_i \bar{u}_j)(s_{ij} + \bar{v}_i \bar{v}_j)
\]

\[
= \sum_{i=1}^{q} \sum_{j=1}^{q} r_{ij} s_{ij} + \sum_{i=1}^{q} \sum_{j=1}^{q} r_{ij} \bar{v}_i \bar{v}_j
\]

\[+ \sum_{i=1}^{q} \sum_{j=1}^{q} s_{ij} \bar{u}_i \bar{u}_j + \sum_{i=1}^{q} \sum_{j=1}^{q} \bar{u}_i \bar{u}_j \bar{v}_i \bar{v}_j
\]

\[= \text{tr } RS + \bar{v}'R\bar{v} + \bar{u}'S\bar{u} + (\bar{u}'\bar{v})^2.
\]
Let $V = \text{Var}(b)$. Applying the theorem to $E \left\{ (x + \bar{e})' b \right\}^2$ with

$E(x + \bar{e}) = x$ and $\text{Var}(x + \bar{e}) = \frac{1}{m} \Sigma$,

$$E \left\{ (x + \bar{e})' b \right\}^2 = \frac{1}{m} \text{tr} \Sigma V + \frac{1}{m} E(b)' \Sigma E(b)$$

$$+ x'Vx + \{x'E(b)\}^2. \quad (A-12)$$

From Eq. (A-10), (A-11), and (A-12),

$$\text{MSE}(\hat{y}) = \frac{1}{m} \text{tr} \Sigma V + \frac{1}{m} E(b)' \Sigma E(b)$$

$$+ x'Vx + (x' \delta)'^2. \quad (A-13)$$

Since $E(b) = \beta + \delta$ and $(x' \delta)'^2 = x' (\delta \delta')x$,

$$\text{MSE}(\hat{y}) = \frac{1}{m} \text{tr} \Sigma V + \frac{1}{m} (\beta + \delta)' \Sigma (\beta + \delta)$$

$$+ x'Vx + x' (\delta \delta')x. \quad (A-14)$$

Note that $\text{MSE}(\hat{y})$ consists of three components. The first two terms are due to the future measurement errors, the third term due to the variance of $b$, and the last term due to the bias of $b$. If $x$ is measured exactly (no errors), then the usual least squares procedure results in $\text{MSE}(\hat{y})$ with only the variance component remaining.
The Integrated Mean Squared Error

The elements of $x$, $d$, $\Sigma$, and $V$ are first partitioned as follows:

$$x = \begin{pmatrix} d \\ x \end{pmatrix}, \quad d = \begin{pmatrix} d_0 \\ \delta \end{pmatrix},$$

$$\Sigma = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma \end{bmatrix}, \quad \text{and} \quad V = \begin{bmatrix} v_{00} & v_{01}' \\ v_{01} & \nu \end{bmatrix}.$$  

Inserting the above expressions into Eq.(A-14) results in:

$$\text{MSE}(\hat{y}) = \frac{1}{m} \text{tr} \Sigma V + \frac{1}{m} (\beta + \delta)' \Sigma (\beta + \delta)$$

$$+ v_{00} + 2\hat{x}'v_{01} + \hat{x}'\nu \hat{x}$$

$$+ \delta_0^2 + 2\delta_0 \delta' \hat{x} + \hat{x}'(\delta \delta') \hat{x}.$$  \hspace{1cm} (A-15)

Next, assume that each element of $x$ is scaled such that:

$$\chi = \{-1 \leq x_i \leq 1 ; \ i = 1,2, \ldots, p\}.$$  \hspace{1cm} (A-16)

Then, by definition,

$$\text{IMS} = \frac{1}{R} \int_{\chi} \text{MSE}(\hat{y}) \, d\hat{x}$$  \hspace{1cm} (A-17)

where $R$ is the volume of $\chi$. Integrating Eq.(A-15) according to Eq.(A-17) yields:
\[ \text{IMS} = \frac{1}{m} \text{tr} \sum V + \frac{1}{m} (\beta + \delta)' \sum (\beta + \delta) + v_{00} + \delta_0^2 \]
\[ + \frac{1}{R} \int_{\mathcal{X}} \dot{x}' (\tilde{\delta} \tilde{\delta}' + \tilde{V}) \dot{x} \, dx \]
\[ = \frac{1}{m} \text{tr} \sum V + \frac{1}{m} (\beta + \delta)' \sum (\beta + \delta) + v_{00} \]
\[ + \delta_0^2 + \frac{1}{3} \text{tr} (\tilde{\delta} \tilde{\delta}' + \tilde{V}) \quad (A-18) \]

Again, it is noted that the IMS consists of the following three components:

(a) component due to future measurement errors
\[ = \frac{1}{m} \text{tr} \sum V + \frac{1}{m} (\beta + \delta)' \sum (\beta + \delta) \]

(b) component due to the variance of \( b \)
\[ = v_{00} + \frac{1}{3} \text{tr} \dot{V} \quad \text{and} \]

(c) component due to the bias of \( b \)
\[ = \delta_0^2 + \frac{1}{3} \text{tr} \tilde{\delta} \tilde{\delta}' \]

In case of no measurement errors in \( x \), the results are:
\[ \text{IMS} = v_{00} + \frac{1}{3} \text{tr} \dot{V} \]
\[ = v_{00} + \frac{1}{3} \sum_{i=1}^{p} v_{ii} . \]
Appendix B

Details of Monte Carlo Simulation Program

B.1. Definitions of Variables and Arrays

**Variables**

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABLO</td>
<td>average value of $\hat{b}_0$ for a given run.</td>
</tr>
<tr>
<td>ABL1</td>
<td>average value of $\hat{b}_1$ for a given run.</td>
</tr>
<tr>
<td>ABMO</td>
<td>average value of $\tilde{b}_0$ for a given run ($\Omega$ unknown).</td>
</tr>
<tr>
<td>ABM1</td>
<td>average value of $\tilde{b}_1$ for a given run ($\Omega$ unknown).</td>
</tr>
<tr>
<td>ABMKO</td>
<td>average value of $\tilde{b}_0$ for a given run ($\Omega$ known).</td>
</tr>
<tr>
<td>ABMK1</td>
<td>average value of $\tilde{b}_1$ for a given run ($\Omega$ known).</td>
</tr>
<tr>
<td>ADF0</td>
<td>$\text{AMSL - AMSM}_1$</td>
</tr>
<tr>
<td>ADF1</td>
<td>$\text{AMSL - AMSL}_2$</td>
</tr>
<tr>
<td>ADF2</td>
<td>$\text{AMSM}_1 - \text{AMSM}_2$</td>
</tr>
<tr>
<td>AMSL</td>
<td>average $\text{IMS}_1$</td>
</tr>
<tr>
<td>AMSM1</td>
<td>average $\text{IMS}_m$ ($\Omega$ unknown).</td>
</tr>
<tr>
<td>AMSM2</td>
<td>average $\text{IMS}_m$ ($\Omega$ known).</td>
</tr>
<tr>
<td>Name</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DLSQ</td>
<td>$\Delta_1 = \sum x_i^2 / N + \gamma^2 - \bar{x}^2$.</td>
</tr>
<tr>
<td>DMLE</td>
<td>$\Delta_2 = \sum x_i^2 / N - \bar{x}^2$.</td>
</tr>
<tr>
<td>DSEED</td>
<td>seed value for multivariate normal deviate generator, GGNSM.</td>
</tr>
<tr>
<td>DXBAR</td>
<td>$\bar{x}$ (mean of the independent variable).</td>
</tr>
<tr>
<td>MF</td>
<td>$m$ (number of future measurements).</td>
</tr>
<tr>
<td>ND</td>
<td>$n$ (number of design points).</td>
</tr>
<tr>
<td>NEXP</td>
<td>number of runs.</td>
</tr>
<tr>
<td>NR</td>
<td>$r$ (number of repeated measurements at each design point).</td>
</tr>
<tr>
<td>NSMPL</td>
<td>number of simulated experiments in a given run.</td>
</tr>
<tr>
<td>NV</td>
<td>number of the independent variables.</td>
</tr>
<tr>
<td>STERO</td>
<td>standard error of ADFO.</td>
</tr>
<tr>
<td>STER1</td>
<td>standard error of ADF1.</td>
</tr>
<tr>
<td>STER2</td>
<td>standard error of ADF2.</td>
</tr>
<tr>
<td>VBLO</td>
<td>sample variance of $\hat{b}_0$.</td>
</tr>
<tr>
<td>VBL1</td>
<td>sample variance of $\hat{b}_1$.</td>
</tr>
<tr>
<td>VBM1</td>
<td>sample variance of $\tilde{b}_0(\Omega$ unknown).</td>
</tr>
<tr>
<td>VBMK1</td>
<td>sample variance of $\tilde{b}_1(\Omega$ known).</td>
</tr>
<tr>
<td>VX</td>
<td>$\gamma^2$</td>
</tr>
<tr>
<td>VY</td>
<td>$\sigma^2$</td>
</tr>
</tbody>
</table>
Arrays

<table>
<thead>
<tr>
<th>Name</th>
<th>Dimension</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLO</td>
<td>(NSMPL)</td>
<td>least squares estimate of $\beta_0$.</td>
</tr>
<tr>
<td>BL1</td>
<td>(NSMPL)</td>
<td>least squares estimate of $\beta_1$.</td>
</tr>
<tr>
<td>BLS</td>
<td>(NV+1)</td>
<td>least squares estimate of $\beta$.</td>
</tr>
<tr>
<td>BMO</td>
<td>(NSMPL)</td>
<td>maximum likelihood estimate of $\beta_0(\Omega$ unknown).</td>
</tr>
<tr>
<td>BM1</td>
<td>(NSMPL)</td>
<td>maximum likelihood estimate of $\beta_1(\Omega$ unknown).</td>
</tr>
<tr>
<td>BMK0</td>
<td>(NV+1)</td>
<td>maximum likelihood estimate of $\beta$.</td>
</tr>
<tr>
<td>BMK1</td>
<td>(NSMPL)</td>
<td>maximum likelihood estimate of $\beta_0(\Omega$ known).</td>
</tr>
<tr>
<td>BT</td>
<td>(NV+1)</td>
<td>true coefficients $\beta$.</td>
</tr>
<tr>
<td>D</td>
<td>(ND, NV+2)</td>
<td>matrix of true values of the independent and dependent variable at design points.</td>
</tr>
<tr>
<td>F</td>
<td>(NV+1, NV+1)</td>
<td>matrix calculated from ZBAR and ZDOT.</td>
</tr>
<tr>
<td>RAHDA</td>
<td>(NV+1)</td>
<td>eigenvalues of the system, $Fb = \lambda Sb$.</td>
</tr>
<tr>
<td>S</td>
<td>(NV+1, NV+1)</td>
<td>sample covariance matrix.</td>
</tr>
<tr>
<td>SIGMA</td>
<td>(NV+1)(NV+2)/2</td>
<td>vector of error covariance.</td>
</tr>
<tr>
<td>SKN</td>
<td>(NV+1, NV+1)</td>
<td>matrix of error covariance.</td>
</tr>
<tr>
<td>W</td>
<td>(NDxNR, NV+1)</td>
<td>observed design matrix.</td>
</tr>
<tr>
<td>Name</td>
<td>Dimension</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>WD</td>
<td>(NDxNR, NV+2)</td>
<td>matrix of the observed dependent variable and the observed design matrix.</td>
</tr>
<tr>
<td>YOB</td>
<td>(NDxNR)</td>
<td>vector of the observed dependent variable.</td>
</tr>
<tr>
<td>Z</td>
<td>(NV+1, NV+1)</td>
<td>complex matrix containing eigenvectors of the system, $Fb = \lambda Sb$.</td>
</tr>
<tr>
<td>ZBAR</td>
<td>(NV+1, ND)</td>
<td>matrix containing ZBI's.</td>
</tr>
<tr>
<td>ZBI</td>
<td>(NV+1)</td>
<td>a vector containing sample means of independent and dependent variable at the ith design point.</td>
</tr>
<tr>
<td>ZDOT</td>
<td>(NV+1)</td>
<td>a vector of grand means of the independent and dependent variable.</td>
</tr>
<tr>
<td>ZREAL</td>
<td>(NV+1)</td>
<td>eigenvector corresponding to the minimum eigenvalue.</td>
</tr>
</tbody>
</table>
B.2. Flowcharts

**Main Program**

1. **Start**

   **Preliminary Steps:**
   (a) read in factor levels and simulation parameters.
   (b) initiate variables and arrays.
   (c) calculate auxiliary quantities.

2. Generate multivariate normal deviates by calling subroutine GGNSM.

3. Calculate necessary statistics.

4. Construct the vector of the observed dependent variable and the observed design matrix.
Calculate the LS and ML estimate of \( \beta \) by calling subroutine LSQ, and MLE, respectively.

If the number of experiments is greater than NSMPL, calculate the mean and variance of estimates. Calculate IMS\(_1\), IMS\(_m\), and DF.

If the number of runs is greater than NEXP, calculate the average and standard error of \( (\text{IMS}_1 - \text{IMS}_m) \). Calculate large-sample approximation. Print out the results.

Stop
Subroutine LSQ

Start

Calculate $W'$

Calculate $(W'W)^{-1}$

Calculate $W'y$

Calculate the least squares estimate of $\beta$ by

$$\hat{b} = (W'W)^{-1}W'y$$

Return
Subroutine MLE

Start

Construct ZDOT and F

Solve the system:
\[ Fb = \lambda Sb \]

Find the minimum eigenvalue and scale the corresponding eigenvector by calling subroutine EGV.

Find the maximum likelihood estimate of \( \beta \).

Return
Main Program

DIMENSION BT(4), D(10,5), SIGMA(6), SIG(6), SKN(3,3), WD(1000,5),
1(3,3), RVEC(300,3), WKVEC(3), ZB1(3), ZBAR(3,10), ZDEL(3),
ZOB(1000), N(1000,4), BLS(4), BML(4), BL0(200), BL1(200),
3BM0(200), BM1(200), BMK0(200), BMK1(200), DF0(200), DF1(200),
4DF2(200)

C DOUBLE PRECISION DSEED

C *****************************************************
C READ IN INPUT DATA.
C *****************************************************
C
C READ(5,100) NP, NQ, ND, NR, MF, NSMPL, NEXP
NV=NP+NQ
K1=NV+1
K2=NV+2
K3=(NQ+1)*(NQ+2)/2
K4=NQ+1
K5=ND*NR

C READ(5,101) (BT(I), I=1,K1)
C READ(5,101) ((D(I,J), J=1,K1), I=1,ND)
C READ(5,101) (SIGMA(I), I=1,K3)

C INITIATE VARIABLES AND ARRAYS. CALCULATE NECESSARY
C QUANTITIES.
C
C IEXP=1
SL=0.
SM1=0.
SM2=0.
VX=SIGMA(1)
VY=SIGMA(3)
DXBAR=0.
DSS=0.

C DO 9 1=1, ND
DXBAR=DXBAR+D(I,2)*NR
DSS=DSS+D(I,2)*D(I,2)*NR
9 CONTINUE

C SKN(1,1)=SIGMA(1)
C SKN(1,2)=SIGMA(2)
C SKN(2,1)=SIGMA(2)
C SKN(2,2)=SIGMA(3)

C DO 11 I=1, ND
Y=0.
DO 10 J=1, K1
Y=Y+D(I,J)*BT(J)
10 CONTINUE
D(I,K2)=Y
11 CONTINUE

C 750 DO 20 I=1, K3
(Main Program continued)

\[ \text{SIG}(I) = \text{SIGMA}(I) \]

\[ \text{CONTINUE} \]

\[ \text{ISM} = 1 \]

-----------------------------------
READ IN SEED VALUE AND INITIALIZE SUBROUTINE GGNSM.
-----------------------------------

\[ \text{READ}(5,102) \ DSEED \]
\[ \text{WKVEC}(1) = 0. \]
\[ \text{CALL GGNSM} \ (\ DSEED, NR, K4, \text{SIG}, 300, \text{RVEC}, \text{WKVEC}, \text{IER}) \]

-----------------------------------
GENERATE MULTIVARIATE NORMAL DEVIATES. CONSTRUCT A MATRIX(WD) OF OBSERVED DESIGN MATRIX AND Y.
-----------------------------------

\[ \text{DO 30} \ I = 1, \text{ND} \]
\[ \text{L1} = (I-1) \times NR + 1 \]
\[ \text{L2} = I \times NR \]
\[ \text{DO 30} \ J = \text{L1}, \text{L2} \]
\[ \text{DO 30} \ L = 1, K2 \]
\[ \text{WD}(J, L) = D(I, L) \]

\[ \text{CONTINUE} \]

\[ \text{DO 31} \ I = 1, K4 \]
\[ \text{DO 31} \ J = 1, K4 \]
\[ S(I, J) = 0. \]

\[ \text{CONTINUE} \]

\[ \text{WKVEC}(1) = 1. \]

\[ \text{DO 40} \ I = 1, \text{NR} \]
\[ \text{DO 40} \ J = 1, K4 \]
\[ \text{L2} = \text{L1} + 11 \]
\[ \text{J2} = \text{J1} + \text{NP} + 1 \]
\[ \text{WD}(12, J2) = \text{WD}(12, J2) + \text{RVEC}(I1, J1) \]

\[ \text{CONTINUE} \]

-----------------------------------
CALCULATE SAMPLE MEANS AND SAMPLE COVARIANCE FOR THE ML ESTIMATION PROCEDURE.
-----------------------------------

\[ \text{DO 42} \ J3 = 1, K4 \]
\[ \text{ZBI}(J3) = 0. \]
\[ \text{DO 41} \ I3 = 1, \text{NR} \]
\[ \text{ZBI}(J3) = \text{ZBI}(J3) + \text{RVEC}(I3, J3) \]

\[ \text{CONTINUE} \]

\[ \text{ZBI}(J3) = \text{ZBI}(J3) / \text{NR} \]

\[ \text{CONTINUE} \]

\[ \text{DO 43} \ J4 = 1, K4 \]
\[ \text{ZBAR}(J4, 1) = \text{ZBI}(J4) + D(1, J4 + \text{NP} + 1) \]

\[ \text{CONTINUE} \]
DO 45 I5 = 1, NR
DO 44 J5 = 1, K4
ZDEL(J5) = RVEC(15, J5) - ZBI(J5)
44 CONTINUE
CALL CALS(ZDEL, K4, S)
45 CONTINUE

DO 46 I = 1, K4
DO 47 J = 1, K4
S(I, J) = S(I, J) / (K5 - ND)
47 CONTINUE

DO 50 I = 1, K5
YOB(I) = WD(I, K2)
DO 50 J = 1, K1
W(I, J) = WD(I, J)
50 CONTINUE

CALL LSQ(K1, K5, W, YOB, BLS)
CALL MLE(ZBAR, S, D, ND, K4, K1, BML)
BLOC ISM) = BLS(1)
BL1(ISM) = BLS(2)
BM0(ISM) = BML(1)
BM1(ISM) = BML(2)
CALL MLEC(ZBAR, SKN, D, ND, K4, K1, BML)
BM0(ISM) = BML(1)
BM1(ISM) = BML(2)

TEST IF NSMPL EXPERIMENTS HAVE BEEN COMPLETED.

ISM = ISM + 1
IF (ISM .LE. NSMPL) GO TO 800

CALL MV(BL0, ABL0, VBL0, NSMPL)
CALL MV(BL1, ABL1, VBL1, NSMPL)
CALL MV(BM0, ABM0, VBMO, NSMPL)
CALL MV(BM1, ABM1, VBM1, NSMPL)
CALL MV(BMK0, ABMK0, VBMK0, NSMPL)
(Main Program continued)

CALL MV(BMK1, ABMK1, VBMK1, NSMPL)

CALL CIMS(VX, ABL0, VBL0, ABL1, VBL1, BT, MF, AMSL)
CALL CIMS(VX, ABM0, VBM0, ABM1, VBM1, BT, MF, AMSM1)
CALL CIMS(VX, ABMK0, VBMK0, ABMK1, VBMK1, BT, MF, AMSM2)

SL = SL + AMSL
SM1 = SM1 + AMSM1
SM2 = SM2 + AMSM2
DF0(IEXP) = AMSL - AMSM1
DF1(IEXP) = AMSL - AMSM2
DF2(IEXP) = AMSM1 - AMSM2

WRITE(6, 201) IEXP, DSEED, AMSL, AMSM1, AMSM2, DF0(IEXP),
            DF1(IEXP), DF2(IEXP), ABL0, ABL1, ABM0, ABM1, ABMK0, ABMK1,
            VBL0, VBL1, VBM0, VBM1, VBMK0, VBMK1

TEST IF NEXP RUNS HAVE BEEN COMPLETED.

IEXP = IEXP + 1
IF (IEXP.LE.NEXP) GO TO 750

CALCULATE THE MEAN IMS'S AND THEIR AVERAGE DIFFERENCES. CALCULATE THE STANDARD ERROR OF MEAN DIFFERENCES.

CALL MV(DF0, ADF0, VDF0, NEXP)
CALL MV(DF1, ADF1, VDF1, NEXP)
CALL MV(DF2, ADF2, VDF2, NEXP)
STER0 = SQRT(VDF0/NEXP)
STER1 = SQRT(VDF1/NEXP)
STER2 = SQRT(VDF2/NEXP)
AMSL = SL/NEXP
AMSM1 = SM1/NEXP
AMSM2 = SM2/NEXP

WRITE(6, 202) AMSL, AMSM1, AMSM2, ADF0, VDF0, STER0,
            ADF1, VDF1, STER1, ADF2, VDF2, STER2

CALCULATE THE IMS'S BASED UPON LARGE SAMPLE EXPRESSION.

T = VY + BT(2) * BT(2) * VX
F1 = T * VX
F2 = BT(2) * BT(2) * VX
F3 = DXBAR * DXBAR
F4 = F2 * VX
C1 = F1 / (MF * K5 * DLSQ) + F2 * ((1 - VX / DLSQ) ** 2) / MF
1 + T / K5 + T * (F3 + 1 / 3) / (K5 * DLSQ) +
2F4 * (F3 + 1 / 3) / (DLSQ * DLSQ)
C2 = F1 / (MF * K5 * DMLE) + F2 / MF + T * (DML0 + F3) / (K5 * DMLE) +
1T / (3 * K5 * DMLE)
C0 = C1 - C2
(Main Program continued)

```fortran
WRITE(6,203) C1,C2,C0
WRITE(6,204) K5,ND,NR,MF,BT(2),SIGMA(1),SIGMA(3)
C
100 FORMAT(714)
101 FORMAT(F10.2)
102 FORMAT(D10.0)
201 FORMAT(/,5X,15,D15.6,6F15.10,/25X,6F15.10)
202 FORMAT(1H1,/,5X,3F15.10)
203 FORMAT(1H1,/,5X,3F15.10)
204 FORMAT(1H1,415,3F10.5)
STOP
END

Subroutine LSQ

SUBROUTINE LSQ(K1,K5,W,YOB,BLS)
DIMENSION W(1000,4),YOB(1000),BLS(4),WTW(4,4),WTWI(4,4),
WK(4),WT(4,1000),WTY(4)
C
CALCULATE W'W, INVERSE OF W'W, AND W'Y.
CALL VMULFM(W,V,K5,K1,K1,1000,1000,WTW,4,IER)
CALL LINVIF(WTW,K1,4,WTWI,0,WK,IER)
DO 10 I= 1,K5
DO 10 J= 1,K1
WT( J, I) = W(I,J)
10 CONTINUE
DO 20 I= 1,K1
WTY(I) = 0.
DO 20 J= 1,K5
WTY(I) = WTY(I) + WT(I,J) * YOB(J)
20 CONTINUE
CALCULATE THE LS ESTIMATE.
DO 30 I= 1,K1
BLS(I) = 0.
DO 30 J= 1,K1
BLS(I) = BLS(I) + WTW(I,J) * WTY(J)
30 CONTINUE
RETURN
END
```
Subroutine EGV

SUBROUTINE EGV(ALFA, BETA, Z, ZREAL, K4)
COMPLEX ALFA(3), Z(3,3)
DIMENSION BETA(3), ZREAL(3), RAMDA(3)

C
************
IDENTIFY THE MINIMUM EIGENVALUE AND ITS EIGENVECTOR
WHICH IS SCALED.
************

C
IFLAG=1
T=99999.
DO 100 I=1,K4
RAMDA(I)=ALFA(I)/BETA(I)
IF(RAMDA(I).GE.T) GO TO 100
IFLAG=I
T=RAMDA(I)
100 CONTINUE

C
DO 150 I=1, K4
ZREAL(I)=Z(I,IFLAG)
150 CONTINUE

C
T=ZREAL(K4)
DO 200 I=1, K4
ZREAL(I)=-ZREAL(I)/T
200 CONTINUE

C
RETURN
END

Subroutine MV

SUBROUTINE MV(D, A, V, N)
DIMENSION D(200)
SUM=0.
SS=0.
DO 100 I=1,N
SUM=SUM+D(I)
100 CONTINUE

C
A=SUM/N

C
DO 200 I=1,N
SS=SS+(D(I)-A)**2
200 CONTINUE

C
V=SS/(N-1.)

C
RETURN
END

Subroutine CIMS

SUBROUTINE CIMS(VX, A0, V0, A1, V1, BT, M, AMS)
DIMENSION BT(4)
AMS=VX*(V1+A1+A1)/M + V0 + (A0-BT(1))**2 +
8 * ((A1-BT(2))**2 + V1)/3.
RETURN
END
APPENDIX C

C.1. Simulation Results When $\Omega$ is Unknown.

<table>
<thead>
<tr>
<th>N</th>
<th>$\beta_1$</th>
<th>m</th>
<th>ND</th>
<th>$\gamma^2$</th>
<th>$\sigma^2$</th>
<th>Average</th>
<th>Average</th>
<th>Average</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IMS</td>
<td>IMS</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>0.56378</td>
<td>0.56391</td>
<td>-0.00014</td>
<td>0.00016</td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>1</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>28.49232</td>
<td>28.46695</td>
<td>0.02338</td>
<td>0.01346</td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
<td>.0100</td>
<td>.001</td>
<td>51.35754</td>
<td>32.00746</td>
<td>-0.64992</td>
<td>0.6825</td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>2</td>
<td>.001</td>
<td>.0100</td>
<td>0.67793</td>
<td>0.67826</td>
<td>-0.00033</td>
<td>0.00016</td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>23.65450</td>
<td>25.05732</td>
<td>0.01718</td>
<td>0.91641</td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>2</td>
<td>.0100</td>
<td>.001</td>
<td>44.85613</td>
<td>45.66029</td>
<td>-0.81017</td>
<td>0.32364</td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>3</td>
<td>.0100</td>
<td>.0100</td>
<td>73.42796</td>
<td>73.44248</td>
<td>-1.03632</td>
<td>0.26972</td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>.00014</td>
<td>.00016</td>
<td>0.00000</td>
<td>0.000047</td>
<td></td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>27.05310</td>
<td>27.11011</td>
<td>-0.02701</td>
<td>0.02173</td>
<td></td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>31.35754</td>
<td>31.00746</td>
<td>-0.64992</td>
<td>0.06825</td>
<td></td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>3</td>
<td>.0100</td>
<td>.0100</td>
<td>50.69025</td>
<td>51.00732</td>
<td>0.01718</td>
<td>0.91641</td>
<td></td>
</tr>
<tr>
<td>6 1.0</td>
<td>5</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>0.67793</td>
<td>0.67826</td>
<td>-0.00033</td>
<td>0.00016</td>
<td></td>
</tr>
<tr>
<td>6 1.0</td>
<td>5</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>23.65450</td>
<td>25.05732</td>
<td>0.01718</td>
<td>0.91641</td>
<td></td>
</tr>
<tr>
<td>6 1.0</td>
<td>5</td>
<td>2</td>
<td>.0100</td>
<td>.001</td>
<td>44.85613</td>
<td>45.66029</td>
<td>-0.81017</td>
<td>0.32364</td>
<td></td>
</tr>
<tr>
<td>6 1.0</td>
<td>5</td>
<td>3</td>
<td>.0100</td>
<td>.0100</td>
<td>73.42796</td>
<td>73.44248</td>
<td>-1.03632</td>
<td>0.26972</td>
<td></td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>.00014</td>
<td>.00016</td>
<td>0.00000</td>
<td>0.000047</td>
<td></td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>27.05310</td>
<td>27.11011</td>
<td>-0.02701</td>
<td>0.02173</td>
<td></td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>31.35754</td>
<td>31.00746</td>
<td>-0.64992</td>
<td>0.06825</td>
<td></td>
</tr>
<tr>
<td>6 1.5</td>
<td>10</td>
<td>3</td>
<td>.0100</td>
<td>.0100</td>
<td>50.69025</td>
<td>51.00732</td>
<td>0.01718</td>
<td>0.91641</td>
<td></td>
</tr>
<tr>
<td>18 0.5</td>
<td>5</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>0.14460</td>
<td>0.14459</td>
<td>0.00002</td>
<td>0.00005</td>
<td></td>
</tr>
<tr>
<td>18 0.5</td>
<td>5</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>7.75952</td>
<td>7.76273</td>
<td>-0.00499</td>
<td>0.00502</td>
<td></td>
</tr>
<tr>
<td>18 0.5</td>
<td>5</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>7.13723</td>
<td>7.14042</td>
<td>-0.00301</td>
<td>0.00372</td>
<td></td>
</tr>
<tr>
<td>18 0.5</td>
<td>5</td>
<td>2</td>
<td>.0100</td>
<td>.0001</td>
<td>14.62751</td>
<td>14.52481</td>
<td>0.10270</td>
<td>0.05553</td>
<td></td>
</tr>
<tr>
<td>18 1.0</td>
<td>10</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>0.23466</td>
<td>0.23459</td>
<td>0.0007</td>
<td>0.00017</td>
<td></td>
</tr>
<tr>
<td>18 1.0</td>
<td>10</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>7.55542</td>
<td>7.55862</td>
<td>-0.00499</td>
<td>0.00502</td>
<td></td>
</tr>
<tr>
<td>18 1.0</td>
<td>10</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>19.54163</td>
<td>19.53366</td>
<td>0.00797</td>
<td>0.17984</td>
<td></td>
</tr>
<tr>
<td>18 1.0</td>
<td>10</td>
<td>3</td>
<td>.0100</td>
<td>.0100</td>
<td>26.35674</td>
<td>26.46746</td>
<td>-0.0471</td>
<td>0.09571</td>
<td></td>
</tr>
<tr>
<td>18 1.5</td>
<td>1</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>2.47481</td>
<td>2.47537</td>
<td>-0.00565</td>
<td>0.00022</td>
<td></td>
</tr>
<tr>
<td>18 1.5</td>
<td>1</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>11.69897</td>
<td>11.70531</td>
<td>-0.00634</td>
<td>0.000724</td>
<td></td>
</tr>
<tr>
<td>18 1.5</td>
<td>1</td>
<td>3</td>
<td>.0100</td>
<td>.001</td>
<td>241.43530</td>
<td>241.46897</td>
<td>5.63367</td>
<td>0.58587</td>
<td></td>
</tr>
<tr>
<td>18 1.0</td>
<td>1</td>
<td>2</td>
<td>.0100</td>
<td>.0100</td>
<td>243.99475</td>
<td>247.72144</td>
<td>-3.72688</td>
<td>0.29264</td>
<td></td>
</tr>
<tr>
<td>30 0.5</td>
<td>10</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>0.68726</td>
<td>0.68796</td>
<td>-0.00068</td>
<td>0.00084</td>
<td></td>
</tr>
<tr>
<td>30 0.5</td>
<td>10</td>
<td>3</td>
<td>.001</td>
<td>.0100</td>
<td>4.90020</td>
<td>4.89564</td>
<td>0.00456</td>
<td>0.00339</td>
<td></td>
</tr>
<tr>
<td>30 0.5</td>
<td>10</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>3.89399</td>
<td>3.89833</td>
<td>-0.00443</td>
<td>0.00399</td>
<td></td>
</tr>
<tr>
<td>30 0.5</td>
<td>10</td>
<td>2</td>
<td>.0100</td>
<td>.0001</td>
<td>8.66335</td>
<td>8.66133</td>
<td>0.00002</td>
<td>0.00162</td>
<td></td>
</tr>
<tr>
<td>30 1.0</td>
<td>1</td>
<td>2</td>
<td>.001</td>
<td>.001</td>
<td>1.09841</td>
<td>1.09871</td>
<td>-0.00030</td>
<td>0.00010</td>
<td></td>
</tr>
<tr>
<td>30 1.0</td>
<td>1</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>5.89932</td>
<td>5.90233</td>
<td>-0.00301</td>
<td>0.000724</td>
<td></td>
</tr>
<tr>
<td>30 1.0</td>
<td>1</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>104.51308</td>
<td>106.91901</td>
<td>-2.40593</td>
<td>0.14798</td>
<td></td>
</tr>
<tr>
<td>30 1.0</td>
<td>1</td>
<td>2</td>
<td>.0100</td>
<td>.0100</td>
<td>108.63216</td>
<td>110.30309</td>
<td>-1.67093</td>
<td>0.13677</td>
<td></td>
</tr>
<tr>
<td>30 1.5</td>
<td>5</td>
<td>2</td>
<td>.0001</td>
<td>.001</td>
<td>0.68752</td>
<td>0.68723</td>
<td>0.00027</td>
<td>0.00016</td>
<td></td>
</tr>
<tr>
<td>30 1.5</td>
<td>5</td>
<td>3</td>
<td>.0001</td>
<td>.0100</td>
<td>5.81856</td>
<td>5.82295</td>
<td>-0.00436</td>
<td>0.00511</td>
<td></td>
</tr>
<tr>
<td>30 1.5</td>
<td>5</td>
<td>3</td>
<td>.0100</td>
<td>.0001</td>
<td>55.33313</td>
<td>55.72374</td>
<td>-0.39661</td>
<td>0.31858</td>
<td></td>
</tr>
<tr>
<td>30 1.5</td>
<td>5</td>
<td>2</td>
<td>.0100</td>
<td>.0100</td>
<td>58.13612</td>
<td>58.49184</td>
<td>-0.35571</td>
<td>0.16607</td>
<td></td>
</tr>
</tbody>
</table>

* standard error
### Simulation Results When $\Omega$ is Known.

<table>
<thead>
<tr>
<th>N</th>
<th>$\beta_1$</th>
<th>m</th>
<th>ND</th>
<th>$\tau^2$</th>
<th>$\sigma^2$</th>
<th>Average $\text{IMS}_1$</th>
<th>Average $\text{IMS}_m$</th>
<th>Average $\text{DF}$</th>
<th>Average $\text{DF}$</th>
<th>s.e.* of $\text{DF}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.56378</td>
<td>0.56392</td>
<td>-0.00014</td>
<td>0.00016</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>1</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>28.49232</td>
<td>28.47310</td>
<td>0.01923</td>
<td>0.0075</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>1</td>
<td>3</td>
<td>0.0001</td>
<td>0.0001</td>
<td>31.39754</td>
<td>32.09840</td>
<td>-0.64235</td>
<td>0.06468</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>2</td>
<td>0.0001</td>
<td>0.0100</td>
<td>50.64490</td>
<td>51.09529</td>
<td>-0.45639</td>
<td>0.09180</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>3</td>
<td>0.0001</td>
<td>0.0001</td>
<td>3.67799</td>
<td>6.78222</td>
<td>0.67822</td>
<td>0.1190</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>25.05458</td>
<td>25.02077</td>
<td>0.1392</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>5</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>45.67821</td>
<td>45.66231</td>
<td>0.05774</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>10</td>
<td>2</td>
<td>0.0001</td>
<td>0.0001</td>
<td>73.42796</td>
<td>74.46410</td>
<td>-1.03613</td>
<td>0.25976</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>10</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>27.08310</td>
<td>27.16954</td>
<td>-0.09634</td>
<td>0.01440</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>10</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>82.91891</td>
<td>84.62167</td>
<td>-1.70276</td>
<td>0.90725</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>10</td>
<td>2</td>
<td>0.0100</td>
<td>0.0100</td>
<td>96.62962</td>
<td>97.02675</td>
<td>-0.39693</td>
<td>0.04033</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.5</td>
<td>5</td>
<td>3</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.14460</td>
<td>0.14459</td>
<td>0.00000</td>
<td>0.00005</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.5</td>
<td>5</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>7.75952</td>
<td>7.76294</td>
<td>0.00002</td>
<td>0.00712</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.5</td>
<td>5</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>7.13729</td>
<td>7.16428</td>
<td>0.00003</td>
<td>0.00573</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.5</td>
<td>5</td>
<td>2</td>
<td>0.0001</td>
<td>0.0001</td>
<td>14.62751</td>
<td>14.52483</td>
<td>0.00004</td>
<td>0.00553</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.0</td>
<td>10</td>
<td>2</td>
<td>0.0001</td>
<td>0.0001</td>
<td>2.47581</td>
<td>2.47537</td>
<td>0.00005</td>
<td>0.00022</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.0</td>
<td>10</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>7.55542</td>
<td>7.54962</td>
<td>0.00004</td>
<td>0.00479</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.0</td>
<td>10</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>19.54183</td>
<td>19.53445</td>
<td>0.00007</td>
<td>0.01715</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.0</td>
<td>10</td>
<td>2</td>
<td>0.0100</td>
<td>0.0001</td>
<td>26.36674</td>
<td>26.40735</td>
<td>0.00009</td>
<td>0.00571</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.5</td>
<td>1</td>
<td>2</td>
<td>0.0001</td>
<td>0.0001</td>
<td>2.47581</td>
<td>2.47537</td>
<td>0.00005</td>
<td>0.00022</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.5</td>
<td>1</td>
<td>3</td>
<td>0.0001</td>
<td>0.0001</td>
<td>11.69897</td>
<td>11.71288</td>
<td>0.00013</td>
<td>0.00644</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.5</td>
<td>1</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>241.43530</td>
<td>247.07996</td>
<td>-5.64455</td>
<td>0.58692</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.5</td>
<td>1</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>243.99475</td>
<td>247.72136</td>
<td>-3.72681</td>
<td>0.23004</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>10</td>
<td>2</td>
<td>0.0001</td>
<td>0.0001</td>
<td>8.60535</td>
<td>8.60133</td>
<td>0.00030</td>
<td>0.00513</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>10</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>8.60535</td>
<td>8.60133</td>
<td>0.00030</td>
<td>0.00513</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>10</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>4.39059</td>
<td>4.99627</td>
<td>0.60667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>10</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>3.89390</td>
<td>3.89560</td>
<td>0.00460</td>
<td>0.03710</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>10</td>
<td>2</td>
<td>0.0100</td>
<td>0.0001</td>
<td>8.60535</td>
<td>8.60133</td>
<td>0.00030</td>
<td>0.00513</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>10</td>
<td>2</td>
<td>0.0001</td>
<td>0.0100</td>
<td>1.09941</td>
<td>1.09870</td>
<td>0.00029</td>
<td>0.00010</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.0</td>
<td>1</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>5.89393</td>
<td>5.89560</td>
<td>0.00426</td>
<td>0.00364</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.0</td>
<td>1</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>104.51308</td>
<td>106.92360</td>
<td>-2.41652</td>
<td>0.14805</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.0</td>
<td>1</td>
<td>2</td>
<td>0.0100</td>
<td>0.0100</td>
<td>183.63216</td>
<td>110.30302</td>
<td>-1.67087</td>
<td>0.13076</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.0</td>
<td>1</td>
<td>3</td>
<td>0.0001</td>
<td>0.0100</td>
<td>5.89393</td>
<td>5.89560</td>
<td>0.00426</td>
<td>0.00364</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.0</td>
<td>1</td>
<td>3</td>
<td>0.0100</td>
<td>0.0001</td>
<td>55.33313</td>
<td>55.72533</td>
<td>-0.39220</td>
<td>0.31920</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.5</td>
<td>5</td>
<td>2</td>
<td>0.0100</td>
<td>0.0100</td>
<td>58.13612</td>
<td>58.49184</td>
<td>-0.35571</td>
<td>0.16668</td>
<td></td>
</tr>
</tbody>
</table>

* standard error
Appendix D

Maximum Likelihood Solutions to the Functional Relationship Model between Two Variables

General

Suppose that the two mathematical variables, \( x \) and \( y \), are related by:

\[
y = \beta_0 + \beta_1 x
\]  

(\text{D}-1)

where \( \beta_0 \) and \( \beta_1 \) are unknown parameters. In an experiment to estimate the above relationship, \( n \) experimental units may be selected and \( r \) pairs of repeated measurements are then taken at each experimental unit. Assume that due to the measurement errors, \( x \) and \( y \) cannot be measured exactly, but instead, \( \xi \) and \( \eta \) are observed such that

\[
\begin{align*}
\xi_{ij} &= x_i + u_{ij} \\
\eta_{ij} &= y_i + v_{ij}
\end{align*}
\]  

(\text{D}-2)

where

\[
\begin{pmatrix}
u_{ij} \\
v_{ij}
\end{pmatrix} \sim \text{BVN} \quad \left\{ \begin{array}{c}
\mathbf{0}, \\
\Sigma
\end{array} \right\} \quad \text{with} \quad \Sigma = \begin{bmatrix}
\tau^2 & \rho \sigma \tau \\
\rho \sigma \tau & \sigma^2
\end{bmatrix}
\]  

(\text{D}-3)
for all $i$ and $j$. Further, assume that the error vector sequence, $\{(u_{ij}, v_{ij})\}$ is independent for all $i$ and $j$. In the literature, the model described in Eq. (D-1) through (D-3) is called a functional relationship model between two variables.

Among the unknown parameters in the above model, $\beta_0$, and $\beta_1$ are of primary concern. Hence, it would be computationally convenient if the unwanted parameters, $\{x_i\}$ and $\Omega$, could be eliminated in the course of maximum likelihood estimation. In the following sections, this possibility will be examined for three different cases. Namely, unknown $\Omega$, known $\Omega$, and unknown $\Omega$ with $\delta = 0$.

**Case I: Unknown $\Omega$**

The following is a two-variable version of the approach of Villegas (1961) to obtain an explicit maximum likelihood solution to $\beta_0$ and $\beta_1$ by successive elimination of $\Omega$ and $\{x_i\}$. It is repeated here for completeness.

First, define

$$N = rn$$

$$\theta_{ij} = (\xi_{ij})$$

$$\bar{\theta}_i = \frac{1}{r} \sum_{j=1}^{r} \theta_{ij} = \left(\frac{\xi_i}{\eta_i}\right)$$

(D-4)
\begin{align*}
  z_i &= \begin{pmatrix} x_i \\ y_i \end{pmatrix} \\
  \bar{\theta} &= (\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_n) \\
  Z &= (z_1, z_2, \ldots, z_n) \\
  S &= \frac{1}{N-n} \sum_{i=1}^{n} \sum_{j=1}^{r} \left( \theta_{ij} - \bar{\theta}_i \right) \left( \theta_{ij} - \bar{\theta}_i \right)' \\

  \text{Then, the likelihood function, } L, \text{ is given by:}
  \begin{align*}
  L &= (2\pi)^{-N} \Omega^{-N/2} \exp \left\{ -\frac{N-n}{2} \text{tr}\Omega^{-1}S \\
  &\quad - \frac{r}{2} \text{tr}\Omega^{-1}(\bar{\theta} - Z)(\bar{\theta} - Z)' \right\} \\
  \end{align*} \\
  \text{(D-5)}

  \text{Taking the natural logarithm of } L \text{ and ignoring the constant term,}
  \begin{align*}
  \log L &\propto - N \log |\Omega| - \left( N-n \right) \text{tr}\Omega^{-1}S \\
  &\quad - r \text{tr}\Omega^{-1}(\bar{\theta} - Z)(\bar{\theta} - Z)' \\
  \end{align*} \\
  \text{(D-6)}

  \text{Hence, the maximum likelihood estimation problem may be described by:}
  \begin{align*}
  \text{(P1)} \quad \begin{cases} \\
  \max \log L \\
  \text{s.t. } \beta_0 f' + \alpha' Z = 0 \\
  \end{cases}
  \end{align*}

  \text{where}
  \begin{align*}
  f' &= (1, 1, \ldots, 1), \text{ and} \\
  \alpha' &= (\beta_1, -1).
  \end{align*}
Since maximizing \((\log L)\) is equivalent to minimizing \((-\log L)\), (P1) is equivalent to:

\[
\begin{align*}
(P2) \quad & \begin{cases}
\min \ (N-n) \text{tr} \Omega^{-1} S + r \text{tr} \Omega^{-1} (\Theta-Z)(\Theta-Z)' \\
\quad + N \log |\Omega| \\
\text{s.t.} \quad \beta_0 f' + \alpha'Z = 0
\end{cases}
\end{align*}
\]

Then, for any given \(\beta_0, \beta_1, \) and \(Z\), Press(1972) gives the maximum likelihood estimate of \(\Omega\) as:

\[
\hat{\Omega} = \frac{1}{N} \left\{ (N-n)S + r(\Theta-Z)(\Theta-Z)' \right\}
\]

(D-7)

Inserting Eq. (D-7) into the objective function in (P2),

\[
(P3) \quad \begin{cases}
\min \ \text{tr} \ \Omega \Omega^{-1} + N \log |\hat{\Omega}| \\
\text{s.t.} \quad \beta_0 f' + \alpha'Z = 0
\end{cases}
\]

Equivalently,

\[
(P4) \quad \begin{cases}
\min \ \left| (N-n)S + r(\Theta-Z)(\Theta-Z)' \right| \\
\text{s.t.} \quad \beta_0 f' + \alpha'Z = 0
\end{cases}
\]

Note that in (P4), \(\Omega\) is eliminated. To eliminate \(Z\), consider the following change of variables:

\[
H = S^{-\frac{1}{2}} (\Theta-Z) \quad \text{where} \quad S^{-\frac{1}{2}} S^{-\frac{1}{2}} = S^{-1}.
\]

(D-8)

Then,

\[
(N-n)S + r(\Theta-Z)(\Theta-Z)'
= S^{\frac{1}{2}} \left\{ (N-n)I_2 + r S^{-\frac{1}{2}} (\Theta-Z)(\Theta-Z)' S^{-\frac{1}{2}} \right\} S^{\frac{1}{2}}
= S^{\frac{1}{2}} \left\{ (N-n)I_2 + rHH' \right\} S^{\frac{1}{2}}, \quad \text{and}
\]
\[\beta_0 f' + \alpha' z\]
\[= \beta_0 f' + \alpha' (\tilde{\Theta} - S^2 H)\]
\[= \beta_0 f' + \alpha' \tilde{\Theta} - \alpha' S^2 H\]
\[= \beta_0 + \alpha' \tilde{\Theta}_i - \alpha' S^2 h_i \text{ for } i = 1, 2, \ldots, n,\]

where \(h_i\) is the \(i^{th}\) column of \(H\). Let \(\gamma_i = \beta_0 + \alpha' \tilde{\Theta}_i\).

Then, (P4) is equivalent to

\[
\begin{align*}
\text{(P5)} \quad & \min \quad |(N-n)I_2 + rHH'| \\
\text{s.t.} \quad & \gamma_i - \alpha' S^2 h_i = 0 \quad \text{for } i = 1, \ldots, n.
\end{align*}
\]

The determinant in (P5) can be expanded as:

\[
(N-n)^2 + Q_1 (N-n) \cdot r + Q_2 r^2 \tag{D-9}
\]

where \(Q_k\) is the sum of all the principal minors of order \(k\) of the matrix, \(HH' = \sum_{i=1}^{n} h_i h_i'\). Note that \(Q_1\) and \(Q_2\) are non-negative. In particular,

\[
Q_1 = \text{tr} HH' = \sum_{i=1}^{n} h_i' h_i \tag{D-10}
\]

which is the sum of the squared distances from the origin to the points \(\{h_i\}\). This implies that \(Q_1\) has a minimum when \(\{h_i\}\) are on the perpendicular from the origin to the corresponding hyperplanes in (P5). Since all the hyperplanes are parallel, those perpendicular points are on the same
straight line passing through the origin, implying that $Q_2$ vanishes at these points. Hence, the determinant in (p5) is minimized at these perpendicular points for any given $\beta_0$ and $\beta_1$. These perpendicular points are found by:

$$h_i^* = \frac{\gamma_1 S^{\frac{1}{2}} \alpha}{\alpha' S \alpha}; \; i=1, 2, \ldots, n. \quad (D-11)$$

Going back to the original variables with Eq. (D-11),

$$\bar{\omega} - \omega = \frac{S \alpha}{\alpha' S \alpha} (\beta_0 f' + \alpha' \bar{\omega}),$$

and

$$\text{tr } HH' = \text{tr } S^{-1}(\bar{\omega} - \omega)(\bar{\omega} - \omega)'$$

$$= \frac{(\beta_0 f' + \alpha' \bar{\omega})(\beta_0 f' + \alpha' \bar{\omega})'}{\alpha' S \alpha}.$$

Hence, (P5) is equivalent to

$$(P6) \min. \frac{(\beta_0 f' + \alpha' \bar{\omega})(\beta_0 f' + \alpha' \bar{\omega})'}{\alpha' S \alpha}$$

Note that $Z$ is eliminated from the optimization procedure.

To minimize the objective function in (P6) for any given $\alpha$,

$$\frac{\partial (P6)}{\partial \beta_0} \alpha (\beta_0 f' + \alpha' \bar{\omega})f = 0$$

$$\Rightarrow \; n \beta_0 + \alpha' \bar{\omega} f = 0 \quad (D-12)$$
Define
\[ \overline{\Theta} = \frac{1}{n} \bar{\Theta} f = \frac{1}{n} \left( \sum_{i=1}^{n} \frac{\xi_i}{n} \right) = \left( \frac{\bar{\xi}}{n} \right). \]

Then, from Eq. (D-12),
\[ \tilde{b}_0 = -\alpha' \overline{\Theta}. \] (D-13)

Define
\[ \Delta \overline{\Theta} = \overline{\Theta} - \bar{\Theta} f = \left[ (\bar{\xi}_1 - \bar{\xi}), \ldots, (\bar{\xi}_n - \bar{\xi}) \right] \], and
\[ F = (\Delta \overline{\Theta})(\Delta \overline{\Theta})'. \]
\[ = \left[ \sum_{i=1}^{n} (\bar{\xi}_i - \bar{\xi})^2, \ldots, \sum_{i=1}^{n} (\bar{\xi}_i - \bar{\xi})(\bar{\eta}_i - \bar{\eta}) \right] \]
\[ \ldots \left( \sum_{i=1}^{n} (\bar{\eta}_i - \bar{\eta})^2 \right]. \]

Then, for \( \tilde{b}_0 \), (P6) is reduced to:
\[ (P7) \quad \min. \quad \frac{\alpha' F \alpha}{\alpha' S \alpha} \]

Then, the maximum likelihood solution to \( \alpha \) is obtained by solving
\[ \tilde{\alpha}(F - \alpha^* S) = 0 \] (D-14)

where \( \alpha^* \) is the smallest eigenvalue of
\[ |F - \lambda S| = 0. \] (D-15)
Explicit solutions to $\beta_0$ and $\beta_1$ are given in Eq. (1-12) in Chapter I.

**Case II: Known $\Omega$**

When the error covariance matrix, $\Omega$ is known, it can be shown that the maximum likelihood solutions to $\beta_0$ and $\beta_1$ are exactly the same as in Case I except that in Eq. (D-14) and (D-15), the sample covariance matrix, $S$ is replaced by the known $\Omega$.

**Case III: Unknown Diagonal $\Omega$**

Suppose that $\Omega$ is unknown, but diagonal. That is, it is known that $f = 0$. Then, the maximum likelihood estimation of $\Omega$ is given by:

$$
\frac{1}{N} \left[ (N-n) \text{diag } S +
+ r \text{ diag } \left\{ (\bar{\theta} - Z)(\bar{\theta} - Z)\right\} \right] \quad (D-14)
$$

where $\text{diag } A$ represents a matrix with only the diagonal elements of $A$. Then, the problem (P4) in Case I now becomes

$$
\begin{align*}
\text{(P4)'} \quad & \min \left| (N-n) \text{ diag } S + r \text{ diag } \left\{ (\bar{\theta}-Z)(\bar{\theta}-Z)' \right\} \right| \\
\text{s.t.} \quad & \beta_0 f' + \alpha'Z = 0
\end{align*}
$$

For (P4)', such transformation as in Eq. (D-8) does not work due to the diagonalization. Note that (P4)' is equivalent to:
\[
\begin{align*}
\text{(P8)} \quad \min \quad & \log |(N-n) \text{ diag } S \\
& + r \text{ diag } \{(\bar{\Theta} - Z)(\bar{\Theta} - Z)\} \\
= \log & \left| (N-n)s_{11} + r \sum_{i=1}^{n} (\bar{\xi}_i - x_i)^2, 0 \right| \\
& 0, (N-n)s_{22} + r \sum_{i=1}^{n} (\bar{\eta}_i - x_i)^2 \\
\text{s.t.} \quad & \beta_0 + \beta_1 x_i - y_i = 0; i=1, 2, \ldots, n.
\end{align*}
\]

where \( S = \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix} \).

Equivalently,
\[
\text{(P9)} \quad \min \quad \log \left\{ (N-n)s_{11} + r \sum_{i=1}^{n} (\bar{\xi}_i - x_i)^2 \right\}
+ \log \left\{ (N-n)s_{22} + r \sum_{i=1}^{n} (\bar{\eta}_i - \beta_0 - \beta_1 x_i)^2 \right\}
\]

Differentiating the objective function with respect to \( \beta_0 \), \( \beta_1 \), and \( \{x_i\} \), and setting the derivatives to zero, the following are obtained:

\[
\frac{\partial (P9)}{\partial \beta_0} = -2r \frac{\sum_{i=1}^{n} (\bar{\eta}_i - \beta_0 - \beta_1 x_i)}{(N-n)s_{22} + r \sum_{i=1}^{n} (\bar{\eta}_i - \beta_0 - \beta_1 x_i)^2} = 0
\]

\[
\frac{\partial (P9)}{\partial \beta_1} = -2r \frac{\sum_{i=1}^{n} (\bar{\eta}_i - \beta_0 - \beta_1 x_i)x_i}{(N-n)s_{22} + r \sum_{i=1}^{n} (\bar{\eta}_i - \beta_0 - \beta_1 x_i)^2} = 0
\]
\[
\frac{\partial (P9)}{\partial x_i} = \frac{-2r (\xi - x_i)}{(N-n)s_{11} + r \sum_{i=1}^{n} (\xi - x_i)^2} + \frac{-2r (\eta - \beta_0 - \beta_1 x_i) \beta_1}{(N-n)s_{22} + r \sum_{i=1}^{n} (\eta - \beta_0 - \beta_1 x_i)^2} = 0
\]

Currently, the explicit solution to the above system of nonlinear equations cannot be explicitly represented.
APPENDIX E

Glossary

**Two-variable Model**

Prediction model between two variables (dependent and independent variable).

**Errors-in-variables Model**

A linear model with measurement errors in the independent variable(s). Further classified into structural and functional relationship model based upon the nature of the variables.

**Extreme-point Design**

A design of experiment in which one half of the total number of observations are allocated at one extreme point of the design space and the other half at the other extreme point.

**Functional Relationship Model**

A class of the errors-in-variables model in which all the variables involved are mathematical ones.
**Integrated Mean Squared Error**

Mean squared error of prediction averaged over the design space.

**Mathematical Variable (Spratt, 1969)**

An entity with no associated probability distribution.

**Mean Squared Error**

Variance plus squared bias of an estimator.

**Optimal Design**

The design of experiment which optimizes a certain criterion (e.g., minimization of the integrated mean squared error).

**Prediction Model**

Errors-in-variables model including estimation of the relationship between the variables as well as prediction of the expected values of the future dependent variable.

**Standard and Non-Standard Instrument**

A standard instrument is a measuring device which is maintained under the highly controlled environment, and serves as a reference for a non-standard instrument which is used in the working area. A standard instrument is in general more accurate and precise than a non-
standard instrument, but needs more efforts and costs to maintain.

Statistical Calibration Problem

Determination of the relationship between non-standard and standard measurements, and prediction of standard measurements based upon non-standard measurements in the future.

Structural Relationship Model

A class of the errors-in-variables model in which all the variables involved are random.
BIBLIOGRAPHY


Wald, A., "The Fitting of Straight Lines If Both Variables are Subject to Error", The Annals of Mathematical Statistics, v. 11, p. 284-300, 1940.