ESTIMATION OF A LINEAR REGRESSION MODEL WHEN THE VARIANCE IS A LINEAR FUNCTION OF UNKNOWN PARAMETERS

By

HUEI CHU TSAI

Bachelor of Arts
Tunghai University
Taiwan, Republic of China
1962

Master of Arts
University of Missouri-Columbia
Columbia, Missouri
1970

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1976

ESTIMATION OF A LINEAR REGRESSION MODEL WHEN
THE VARIANCE IS A LINEAR FUNCTION
OF UNKNOWN PARAMETERS

Thesis Approved:

Ronald W. McNew
Thesis Adviser

David L. Weeks

W. W. Wade

Lyle Broemeling

G. H. Weaver

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to express sincere appreciation to Dr. Ronald W. McNew, major adviser, for suggesting the problem, for guidance and counsel throughout the course of this study and the preparation of this thesis.

I would like to thank each member of my advisory committee, Dr. David L. Weeks, Dr. Lyle D. Broemeling, Dr. William Warde, and Dr. G. H. Weaver for their interest and helpful suggestions.

Thanks are extended to Mrs. Debbie Williams for the typing of the thesis.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

For the analysis of data in practice, a linear model with homoscedastic error is usually assumed. However, it may be known that homogeneity of errors is not a realistic assumption and it may be suspected that the variance of the dependent variable varies with the mean or with the independent variables. This point can be illustrated by a few examples. In the study of the results of family budget inquiries, Prais and Houthakker (2) found that the residuals have variance increasing with household income. Ezekiel and Fox (1) found that the variance of auto stopping distance is proportionate to the square of the speed. In this thesis, we assume that the variance in the general linear regression model is a linear function of unknown parameters.

Two models will be referred to throughout this thesis.

Model I

Let $y = X\beta + \varepsilon$ denote a linear regression model where

y is an nx1 vector of observable random variables,

X is an nxp matrix of known constants with rank $(X) = p$,

β is a px1 vector of unknown constants,

ε is an nx1 vector of unobservable random variables such that $V(\varepsilon) = V = \text{diagonal } \{\sigma_i^2 : i=1, 2, \ldots, n\}$, where (1) $\sigma_i^2 = \sum_{j=1}^{k} r_j \, d_{ij}$ is positive

(2) r_i's are unknown constants

(3) $d_{ij} \neq d_{i'j}$ for $i \neq i'$ and d_{ij}'s are known positive constants

(4) k satisfies the conditions in Appendix A of this article, which are required for the β_i's to be estimable.

Model II

The same as Model I except $V(\varepsilon) = aI + bD$ where

$a = r_1$

$b = r_2$

$D = \text{diag}\{d_i,\ i=1, 2, \ldots, n \text{ and } 0<d_1<d_2<\ldots<d_n\}$.

We consider the problems of estimating the regression coefficient vector β and the variance components r_i, i=1, 2, ..., k in Model I.

In Chapter II, we propose an estimator of the regression coefficient vector of Model I. The proposed estimator is obtained based on a maximum rank transformation which results in equal variances. Properties of the transformation matrix and the proposed estimator are presented. Finally, the estimator is proved to be consistent, unbiased, and has a smaller variance than the ordinary least squares estimator when the relative size of the largest variance to the smallest variance in the regression model is large.

In Chapter III, we combine the orthonormal basis of the error space" (OBES) technique proposed by Putter (3) and the transformation technique proposed in Chapter II to estimate the variance components in Model II. We develop estimators for the variance components in Model II and our proposed estimators are compared with minimum norm quadratic unbiased estimators (MINQUE). There always exist such proposed unbiased estimators in case the MINQUE do not exist. Furthermore, the proposed method provides a simple calculation procedure where MINQUE may require a generalized inverse procedure..

Chapter II and III will be presented in a form acceptable to JASA. Chapter IV is a general summary of the two studies. Additional algebraic results related to the material in Chapter II, a computer program source list for obtaining the proposed transformation, and a sample output are presented in Appendix B and Appendix C.

REFERENCES

(1) Ezekiel, M., and Fox, K. A. Methods of Correlation and Regression Analysis, New York: John Wiley and Sons, Inc., 1959, p. 40.

(2) Prais, S. J., and Houthakker, H. S. The Analysis of Family Budgets, The University Press, Cambridge, 1955.

(3) Putter, J. "Orthonormal Bases of Error Spaces and Their Use For Investigating the Normality and Variances in Linear Models," JASA, 1970, 65:161-171.

CHAPTER II

ESTIMATION OF REGRESSION COEFFICIENTS IN A LINEAR REGRESSION MODEL WHEN THE VARIANCE IS A LINEAR FUNCTION OF UNKNOWN PARAMETERS

Abstract

We proposed a transformation technique and an estimator for the vector of regression coefficients in a linear regression model when the variance is a linear function of unknown parameters. A maximum rank transformation matrix, which eliminates the variance heterogeneity, is constructed. The proposed estimator is BLUE based on the transformed model. Furthermore, the proposed estimator has a smaller variance than the estimator based on ordinary least squares when the relative size of the largest variance to the smallest variance in the regression model is large.

Introduction

<u>Model I</u>

Let $y = X\beta + \varepsilon$ denote a linear regression model where

y is an nx1 vector of observable random variables,

X is an nxp matrix of known constants with rank $(X) = p$,

β is a px1 vector of unknown constants,

ε is an nx1 vector of unobservable random variables such that $V(\varepsilon) = V = \text{diagonal } \{\sigma_i^2 : i=1, 2, \ldots, n\}$,

where (1) $\sigma_i^2 = \sum_{j=1}^{k} r_j \, d_{ij}$ is positive

(2) r_i's are unknown constants

(3) $d_{ij} \neq d_{i'j}$ for $i \neq i'$ and d_{ij}'s are known positive constants

(4) k satisfies the conditions in Appendix A of this article, which are required for the β_i's to be estimable.

Model II

The same as Model I except $V(\varepsilon) = aI + bD$ where

$a = r_1$

$b = r_2$

$D = \text{diagonal}\{d_i,\ i=1, 2, \ldots, n \text{ and } 0<d_1<d_2<\ldots<d_n\}$.

We consider the problem of estimating β the vector of regression coefficients.

In general, if the errors in the linear regression model are mutually uncorrelated but have differing variances, the ordinary least squares estimates (OLSE) are unbiased and consistent (as proved) in (2)). However, an error will arise if we use the conventional formula for calculating the standard errors of the OLSE (4). As a consequence, tests of hypotheses are affected if we use the biased estimates of the variance-covariance matrix of the OLSE. It is known that the OLSE of the regression coefficients are not BLUE, except when Zyskind's (13) conditions are satisfied, in the case of unequal variances (1). However, Monte Carlo evidence for some simple special cases indicated that OLSE may not be a bad procedure (2,3).

Rao (7) developed the method of MINQUE to estimate the variance and covariance components in a linear model. Rao (8) also suggested estimating the regression coefficients using generalized least squares estimates in which the unknown variance and covariance components are replaced by their MINQUE's. The properties of such estimators remain to be investigated.

Hartley and Jayatillake (1) estimated the regression coefficients and variance components by maximum likelihood under the assumption of a lower bound for the variance components. The asymptotic distributions of these maximum likelihood estimates (MLE) of the regression coefficients are normal. It is known that such estimators may not perform optimally for small sample sizes.

Rutemiller and Bower (9) permitted the variance of the dependent variable to be a function of the independent variables in the linear regression model. Under the normality assumption, MLE and their variance covariance matrix are approximated by the "Method of Scoring" (6) based on large samples. The properties of asymptotic normality are used to construct confidence intervals and to make other statistical inferences.

Takeshi Amemlya (10) considered a regression model where the variance of the dependent variable is proportional to the square of its expectation. He estimated the regression coefficients using generalized least squares estimates (GLSE) in which the unknown variance components, which are functions of the regression coefficients, are replaced by OLSE of the regression coefficients.

It was proved that this weighted least squares estimator is asymptotically normal.

Due to the analytical complexities of solving non-linear equations and of inverting a random matrix, the small sample distributions of the maximum likelihood estimators and the generalized least squares estimators are seldom obtainable. The difficulty of inverting a matrix containing random variables arises in cases such as the following:

If $V^{-1} = (\hat{a}I + \hat{b}D)^{-1}$ in $\hat{\beta}_{gls} = (X'V^{-1}X)^{-1}X'V^{-1}y$, then V^{-1} cannot be rewritten as $f(\hat{a},\hat{b})U$, where f is a real valued function of $\hat{a}$ and $\hat{b}$ and U is a matrix in terms of I and D only. Therefore, the asymptotic approach has become a common practice in inference using these estimators. Although the small sample distribution of the ordinary least squares estimator is easily obtained, this estimator will be inefficient if the relative size of the largest variance to the smallest variance in the regression model is large (12).

In this article we are interested in finding a small sample estimator for β which is free from the inefficiency resulting from the use of OLSE. For the specified Model II, we constructed an $m \times n$ matrix T such that $TVT' = \sigma^2 I_m$ where σ^2 is an unknown constant and m $(< n)$ is a positive integer to be justified in section 3 of this article. Let $z=Ty$, $W=TX$, and $U=T\varepsilon$, it is clear that the ordinary least squares estimator of β based on this new model, $z=W\beta + U$, is BLUE.

In section 3, we are going to construct the transformation matrix and propose an estimator for β. In section 4, we will provide the conditions for the transformation to preserve the rank of X in the specified Model I when p=2. In section 5, we compare the variance of the proposed estimator with the variance of the ordinary least squares estimator when p=2. Finally, in section 6, we summarize the results and state the unsolved problems.

Construction of the Transformation Matrices and a Proposed Estimator for β

Given two points on the real line, any point between them can be represented in terms of the two end points. Based on this statement, we can prove the following three lemmas and a theorem.

Lemma 1: Let $y_i \sim (0,\sigma_i^2)$, i=1, 2, ..., n, be n uncorrelated random variables. Let $\sigma_i^2 = \sigma^2 f(x_i)$, where f is a strictly monotone real valued function of x_i, and σ^2 is a real valued constant. Then for every $x_i < x_j < x_k$, there exists a $\lambda \varepsilon (0,1)$ such that $V(\sqrt{\lambda}\, y_i \pm \sqrt{1-\lambda} y_k) = V(y_j)$.

Proof. Let $\lambda = \dfrac{f(x_k) - f(x_j)}{f(x_k) - f(x_i)}$, then $0 < \lambda < 1$ and

$$V(\sqrt{\lambda} y_i \pm \sqrt{1-\lambda}\, y_k) = V(y_i) + (1-\lambda) V(y_k)$$

$$= \frac{f(x_k) - f(x_j)}{f(x_k) - f(x_i)} \sigma^2 f(x_i) + \frac{f(x_j) - f(x_i)}{f(x_k) - f(x_i)} \sigma^2 F(x_k)$$

$$= \sigma^2 f(x_j)$$

$$= V(y_j). \qquad \text{QED}$$

<u>Lemma 2</u>: Let $D = \text{diagonal } \{d_i,\ i=1,\ 2,\ \ldots,\ n\}$ such that $0 < d_1 < d_2 \ldots < d_n$. Let n be an odd positive integer such that $n = 2m+1$. Let

$$\underset{(m+1)\times(n)}{A} = \begin{bmatrix} & & & 1 & & & \\ & & \sqrt{\lambda_2} & & -\sqrt{1-\lambda_2} & & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots & \\ & \sqrt{\lambda_i} & & & & -\sqrt{1-\lambda_i} & \\ \cdot^{\cdot^{\cdot}} & & & & & & \ddots \\ \sqrt{\lambda_{m+1}} & & & & & & -\sqrt{1-\lambda_{m+1}} \end{bmatrix}$$

where

$$\lambda_2 = \frac{d_{m+2} - d_{m+1}}{d_{m+2} - d_m}$$

$$\vdots$$

$$\lambda_i = \frac{d_{m+i} - d_{m+1}}{d_{m+i} - d_{m+1-i+1}}$$

$$\vdots$$

$$\lambda_{m+1} = \frac{d_{2m+1} - d_{m+1}}{d_{2m+1} - d_1} .$$

Then (1) $AA' = I_{m+1}$.

(2) ADA' is a diagonal matrix for any arbitrary diagonal matrix D.

(3) $ADA' = d_{m+1}I_{m+1}$ for the specified D matrix.

(4) $A'ADA'A = d_{m+1}A'A$.

(5) $(A'A)(A'A) = A'(AA')A = A'A$.

(6) Rank $(A) = m+1$.

(7) Trace (A'A) = m+1

(8) Norm (A'A - V) = $m+1-2(a+bd_{m+1})(m+1)-\sum_{i=1}^{n}(a+bd_i)^2$.

Proof. (1)

$$AA' = \begin{bmatrix} & & & 1 & & & \\ & & \sqrt{\lambda_2} & & -\sqrt{1-\lambda_2} & & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots & \\ & \sqrt{\lambda_i} & & & & -\sqrt{1-\lambda_i} & \\ \cdot^{\cdot^{\cdot}} & & & & & & \ddots \\ \sqrt{\lambda_{m+1}} & & & & & & -\sqrt{1-\lambda_{m+1}} \end{bmatrix}$$

$$\begin{bmatrix} & & & & \sqrt{\lambda_{m+1}} \\ & & \sqrt{\lambda_i} & \cdot^{\cdot^{\cdot}} & \\ \sqrt{\lambda_2} & \cdot^{\cdot^{\cdot}} & & & \\ 1 & & & & \\ -\sqrt{1-\lambda_2} & \ddots & & & \\ & & -\sqrt{1-\lambda_i} & & \\ & & & \ddots & \\ & & & & -\sqrt{1-\lambda_{m+1}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & & & & & \\ & \lambda_2+(1-\lambda_2) & & & & \\ & & \ddots & & & \\ & & & \lambda_i+(1-\lambda_i) & & \\ & & & & \ddots & \\ & & & & & \lambda_{m+1}+(1-\lambda_{m+1}) \end{bmatrix}$$

$$= I_{m+1}.$$

(2) & (3)

$$ADA' = \begin{bmatrix} & & & 1 & & & \\ & & \sqrt{\lambda_2} & & -\sqrt{1-\lambda_2} & & \\ & \iddots & & & & \ddots & \\ & \sqrt{\lambda_i} & & & & -\sqrt{1-\lambda_i} & \\ \iddots & & & & & & \ddots \\ \sqrt{\lambda_{m+1}} & & & & & & -\sqrt{1-\lambda_{m+1}} \end{bmatrix} DA'$$

$$= \begin{bmatrix} & & & d_{m+1} & & & \\ & & \sqrt{\lambda_2}\,d_m & & -\sqrt{1-\lambda_2}\,d_{m+2} & & \\ & \iddots & & & & \ddots & \\ & \sqrt{\lambda_i}\,d_{m-i+2} & & & & -\sqrt{1-\lambda_i}\,d_{m+i} & \\ \iddots & & & & & & \ddots \\ \sqrt{\lambda_{m+1}}\,d_1 & & & & & & -\sqrt{1-\lambda_{m+1}}\,d_n \end{bmatrix}.$$

$$\begin{bmatrix} & & & & \sqrt{\lambda_{m+1}} \\ & & \sqrt{\lambda_i} & \cdots & \\ & \sqrt{\lambda_2} \cdots & & & \\ 1 & & & & \\ & -\sqrt{1-\lambda_2} \cdots & & & \\ & & -\sqrt{1-\lambda_i} & \cdots & \\ & & & & -\sqrt{1-\lambda_{m+1}} \end{bmatrix}$$

$$= \begin{bmatrix} d_{m+1} & & & & \\ & \lambda_2 d_m + (1-\lambda_2) d_{m+2} & & & \\ & & \ddots & & \\ & & & \lambda_i d_{m-i+2} + (1-\lambda_i) d_{m+i} & \\ & & & & \ddots \\ & & & & & \lambda_{m+1} d_1 + (1-\lambda_{m+1}) d_n \end{bmatrix}$$

$$= \quad d_{m+1} I_{m+1}.$$

(4) Follows from (3), (5) follows from (1) and (7) follows from the idempotence of A'A in (5).

(8) $\text{Norm } (A'A - V) = \text{Trace } (A'A - V)^2 = \text{Trace } (A'A - V)(A'A - V)$

$$= \text{Trace } (A'AA'A - VA'A - A'AV + V^2)$$

$$= \text{Trace } (A'A) - \text{Trace } (VA'A) - \text{Trace } (A'AV) + \text{Trace } (V^2)$$

$$= m+1 - \text{Trace } (AVA') - \text{Trace } (AVA') + \text{Trace } (V^2)$$

$$= m+1-2(a+bd_{m+1})(m+1) - \sum_{i=1}^{n} (a+bd_i)^2.$$

Note: For the case $n = 2m$, let $d = (d_m + d_{m+1})/2$ and

$$A = \begin{bmatrix} & & & \sqrt{\lambda_1} & -\sqrt{1-\lambda_1} & & & \\ & & \sqrt{\lambda_2} & & & -\sqrt{1-\lambda_2} & & \\ & & \cdot^{\cdot^{\cdot}} & & & & \ddots & \\ & \sqrt{\lambda_i} & & & & & -\sqrt{1-\lambda_i} & \\ & \cdot^{\cdot^{\cdot}} & & & & & & \ddots \\ \sqrt{\lambda_m} & & & & & & & -\sqrt{1-\lambda_m} \end{bmatrix}$$

where $\lambda_i = \dfrac{d_{m+1} - d}{d_{m+1} - d_{m-i+1}}$, $i=1, 2, \ldots, m$.

Then similar results and proof of Lemma 2 exist.

<u>Lemma 3</u>: If there is an mxn matrix T of rank m such that $T(aI+bD)T' = cI$, where T does not depend on a, b, or c, then the maximum value of m is n/2 if n is even and (n+1)/2 if n is odd.

Proof: $T(aI+bD)T' = aTT' + bTDT' = cI_m$.

Because T does not depend on a, b, or c, we have $TT' = c_1I$ and $TDT' = c_2I$. Let $c_2 = kc_1$, then $TT' = c_1I$ and $TDT' = kc_1I = kTT'$.

Hence $T(D-kI)T' = 0$.

Case 1: If rank $(D - kI) = n$, then by using the Frobenius inequality (5) rank $((D-kI)T')$ - rank $(T(D-kT)T')$ = rank $(D-kI)$ - rank $(T(D-kI)$ or $m = n-m$ or $m = n/2$.

Case 2: If rank $(D-kI) = n-1$, then one may shrink the dimension of T and $(D-kI)$ from mxn and nxn to mx(n-1) and (n-1)x(n-1). Let T*, and D* be the shrunk matrices such that rank (T*) = m or (m-1) and

and rank $(D^*) = n-1$, we have $T^*D^*T^{*\prime} = 0$. Again using the Frobenius inequality, we have $m \leqq (n-1)-m$ or $(m-1) \leqq (n-1)-(m-1)$. Hence, we have $m \leqq (n-1)/2$ or $m \leqq (n+1)/2$. To combine the results from case 1 and case 2, we have the maximum rank of T is $n/2$ if n is even and $(n+1)/2$ if n is odd.

Theorem: Assume the stated Model I, where k satisfies the specification of Appendix A. Then there exists a matrix A such that $z = Ay \sim (AX, \sigma^2 I)$, where σ^2 is a real valued constant to be determined, and $\hat{\beta}_{mt} = (X'A'AX)^{-1}X'A'Ay$ is BLUE based on the transformed model.

Proof. Step 1: To show it is true for $k = 1$, or $\sigma_i^2 = r_1 d_{i1}$ for $i=1, 2, \ldots, n$.

Case 1: (1) $n = 2(m_1) + 1$

(2) $0 < d_{11} < d_{21} < \cdots < d_{n1}$

Define $z_1^{(1)} = y_{m_1+1}$

$z_2^{(1)} = \lambda_2 y_{m_1} + (1-\lambda_2) y_{m_1+2}$

.

.

.

$z_i^{(1)} = \lambda_i y_{m_1+1} + (1-\lambda_i) y_{m_1+1-i+1}$

.

.

.

$$z^{(1)}_{m_1+1} = \lambda_{m_1+1} y_1 + (1-\lambda_{m_1+1}) y_n.$$

where

$$\lambda_2 = \frac{d_{m_1+2} - d_{m_1+1}}{d_{m_1+2} - d_m}$$

.

.

.

$$\lambda_i = \frac{d_{m_1+i} - d_{m_1+1}}{d_{m_1+i} - d_{m_1+1-i+1}}$$

.

.

.

$$\lambda_{m_1+1} = \frac{d_n - d_{m_1+1}}{d_n - d_1}$$

Let

$$A^{(1)} = \begin{bmatrix} & & 1 & & & \\ & & \sqrt{\lambda_2} & & -\sqrt{1-\lambda_2} & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots \\ & \sqrt{\lambda_i} & & & -\sqrt{1-\lambda_i} & \\ \cdot^{\cdot^{\cdot}} & & & & & \ddots \\ \sqrt{\lambda_{m_1+1}} & & & & & -\sqrt{1-\lambda_{m_1+1}} \end{bmatrix}$$

Then

$$\underset{(m_1+1)x1}{z^{(1)}} = \underset{(m_1+1)xn}{A^{(1)}} \underset{nx1}{y}, \quad E(z^{(1)}) = A^{(1)} X \beta, \quad V(z^{(1)}) = D^{(1)}$$

where

$D^{(1)} = r_1 d_{m_1+1} I_{m_1+1}$ by Lemma 1 and Lemma 2.

Case 2: (1) $n = 2(m_1)$

(2) $0 < d_{11} < d_{21} < \dots < d_{n1}$.

Using the same approach as in case 1 and the transformation matrix defined in the note of Lemma 2, we would have a similar result as in case 1.

Step 2: To show it is true for k = 2

$$\sigma_i^2 = r_1 d_{i1} + r_2 d_{i2}$$

Case 1: (1) $n = 2m_1 + 1$

(2) $m_1+1 = 2(m_2) + 1$

(3) $0 < d_{11} < d_{21} < \ldots < d_{n1}$

obtaining $\underset{(m_1+1)x1}{z^{(1)}} = A^{(1)}y$ as step 1, we have now

$z^{(1)} \sim (A^{(1)} X \beta,\ D^{(1)} + D^{(2)})$ by Lemma 1 and Lemma 2

where $D^{(1)} = r_1 d_{m_1+1} I_{m_1+1}$ and

$$D^{(2)} = A^{(1)} \begin{bmatrix} r_2 d_{2,1} & & & \\ & r_2 d_{2,2} & & \\ & & \ddots & \\ & & & r_2 d_{2,n} \end{bmatrix} (A^{(1)})'$$

$$= r_2 \begin{bmatrix} d_{2,m_1+1} & & & \\ & \lambda_2 d_{2,m_1} + (1-\lambda_2) d_{2,m_1+2} & & \\ & & \ddots & \\ & & & \lambda_{m_1+1} d_{2,1} + (1-\lambda_{m_1+1}) d_{2,n} \end{bmatrix}$$

by Lemma 2.

Denote $D^{(2)} = \text{Diag}\ \{d_i^{(2)}\}$, i=1, 2, ..., m_1+1, and assume

(4) $0 < d_1^{(2)} < d_2^{(2)} < \ldots < d_{m+1}^{(2)}$.

Define

$$\lambda_2 = \frac{d_{m_2+2}^{(2)} - d_{m_2+1}^{(2)}}{d_{m_2+2}^{(2)} - d_{m_2}^{(2)}}$$

$$\vdots$$

$$\lambda_i = \frac{d_{m_2+i}^{(2)} - d_{m_2+1}^{(2)}}{d_{m_2+i}^{(2)} - d_{m_2+1-i+1}^{(2)}}$$

$$\vdots$$

$$\lambda_{m_2+1} = \frac{d_{m_1+1}^{(2)} - d_{m_2+1}^{(2)}}{d_{m_1+1}^{(2)} - d_{m}^{(2)}}$$

$$z_1^{(2)} = z_{m_2+1}^{(1)}$$

$$z_2^{(2)} = \sqrt{\lambda_2}\ z_{m_2}^{(1)} + \sqrt{1-\lambda_2}\ z_{m_2+2}^{(1)}$$

$$\vdots$$

$$z^{(2)} = \sqrt{\lambda_i}\ z_{m_2}^{(1)} + \sqrt{1-\lambda_2}\ z_{m_2+2}^{(1)}$$

$$\vdots$$

$$z_{m_2+1}^{(2)} = \sqrt{\lambda_{m_2+1}}\ z_1^{(1)} + \sqrt{1-\lambda_{m_2+1}}\ z_{m_1+1}^{(1)}$$

Now let

$$\underset{(m_2+1)(m_1+1)}{A^{(2)}} = \begin{bmatrix} & & 1 & & \\ & & \sqrt{\lambda_2} & -\sqrt{1-\lambda_2} & \\ & \cdot^{\cdot^{\cdot}} & & & \ddots \\ & \sqrt{\lambda_i} & & -\sqrt{1-\lambda_i} & \\ \cdot^{\cdot^{\cdot}} & & & & \ddots \\ \sqrt{\lambda_{m_2+1}} & & & & -\sqrt{1-\lambda_{m_2+1}} \end{bmatrix}$$

By Lemma 1 and Lemma 2, we have

$$\underset{(m_2+1)x1}{z^{(2)}} = \underset{(m_2+1)x(m_1+1)}{A^{(2)}} \underset{(m_1+1)x1}{z^{(1)}}, \quad E(z^{(2)}) = A^{(2)}A^{(1)} X \beta ,$$

$$V(z^{(2)}) = D^{(1)} + A^{(2)}D^{(2)}(A^{(2)})'$$

where $D^{(1)} = r_1 d_{m_1+1} I_{m_2+1}$, and $A^{(2)}D^{(2)}(A^{(2)})' = r_2 d^{(2)}_{m_2+1} I_{m_2+1}$.

Case 2: $m_1+1 = 2(m_2)$

Using the note of Lemma 2, a similar result as in case 1 can be obtained.

Step 3: Assume $k > 2$ and

$$z^{(k)} = A^{(k)}z^{(k-1)} \sim [\prod_{j=1}^{k} A^{(j)}) X \beta, \sum_{j=1}^{k} r_j d^{(j)}_{m_j+1} I_{m_k+1}] \quad \text{is true.}$$

Step 4: To show it is true for $k + 1$

$z^{(k)} = A^{(k)}z^{(k-1)}$ as in step 3, we have

$$z^{(k)} \sim [(\prod_{j=1}^{k} A^{(j)}) X \beta, \sum_{j=1}^{k} r_j d^{(j)}_{m_j+1} I_{m_k+1} + (\prod_{j=k}^{1} A^{(j)}) D^{(k+1)} (\prod_{j=1}^{k} A^{(j)})']$$

By Lemma 2, we can denote

$$D = \text{Diag}\ \{d_i\} = (\prod_{j=k}^{1} A^{(j)}) D^{(k+1)} (\prod_{j=1}^{k} A^{(j)})'$$

Case 1: (1) $m_k + 1 = 2m + 1$

(2) $0 < d_1 < d_2 < \ldots < d_{m+1}$

Define

$$\lambda_2 = \frac{d_{m+2} - d_{m+1}}{d_{m+2} - d_m}$$

.
.
.

$$\lambda_i = \frac{d_{m+i} - d_{m+1}}{d_{m+i} - d_{m+1-i+1}}$$

.
.
.

$$\lambda_{m+1} = \frac{d_{m_k+1} - d_{m+1}}{d_{m_k+1} - d_1} .$$

Define

$$z_1^{(k+1)} = z_{m+1}^{(k)}$$

$$z_2^{(k+1)} = \sqrt{\lambda_2}\ z_m^{(k)} + \sqrt{1-\lambda_2}\ z_{m+2}^{(k)}$$

.
.
.

$$z_i^{(k+1)} = \sqrt{\lambda_i}\, z_{m+i}^{(k)} + \sqrt{1-\lambda_2}\, z_{m+1-i+1}^{(k)}$$

$$\vdots$$

$$z_{m+1}^{(k+1)} = \sqrt{\lambda_{m+1}}\, z_1^{(k)} + \sqrt{1-\lambda_{m+1}}\, z_{m_k+1}^{(k)}$$

Now let

$$A^{(l+1)} = \begin{bmatrix} & & & 1 & & & \\ & & \sqrt{\lambda_2} & & -\sqrt{1-\lambda_2} & & \\ & \iddots & & & & \ddots & \\ & \sqrt{\lambda_i} & & & & -\sqrt{1-\lambda_i} & \\ \iddots & & & & & & \ddots \\ \sqrt{\lambda_{m2+1}} & & & & & & -\sqrt{1-\lambda_{m2+1}} \end{bmatrix}$$

and we have

$$z^{(k+1)} = A^{(k+1)}\, z^{(k)}, \quad E(z^{(k+1)}) = (\prod_{j=1}^{k+1} A^{(j)})\, X 3,$$

$$V(z^{(k+1)}) = A^{(k+1)}\, D\, (A^{(k+1)})'.$$

By Lemma 2, we have

$$A^{(k+1)} D (A^{(k+1)})' = \sum_{j=1}^{k} r_i d_{m_j+1}^{(j)} I_{m+1} + r_{k+1} d_{m+1} I_{m+1}$$

Denote

$$\sigma^2 = \sum_{j=1}^{k} r_j d_{m_j+1}^{(j)} + r_{k+1} d_{m+1}, \text{ and } A = \prod_{j=1}^{k+1} A^{(j)},$$

then we have

$$\underset{(m+1)x1}{Z^{(k+1)}} \sim (A\,X\,\beta\,, \sigma^2 I_{m+1}).$$

Case 2: $m_k + 1 = 2m$.

Using the note of Lemma 2, a similar result as in case 1 can be obtained. Finally, $\hat{\beta}_{mt} = (X'A'AX)^{-1}X'A'AY$ is BLUE under the transformed model by applying the Gauss-Markov Theorem. QED

The relationship among n, m_i, and k is specified in Appendix A of this thesis. The formula in Appendix A can be illustrated by the following example. Let $n = 7$, then $m_1 = (7+1)/2 = 4$, $m_2 = 4/2 = 2$. Hence, for the simple linear regression case, $k \leq 3$.

Conditions for the Transformation to Preserve the Rank of X in Model I When p = 2

Let A be the defined (m+1)xn transformation matrix. Then rank (AX) = 2 when p = 2 in Model I if at least one of the following equalities does not hold:

$$\lambda_i = \frac{E_i}{E_i + F_i}\,,\ i=2,\ 3,\ \ldots,\ m+1,$$

where $E_i = X_{2,m+1}X_{1,m+i} - X_{1,m+1}X_{2,m+1})$ and

$$F_i = (X_{1,m+1}X_{2,m+2-i} - X_{1,m+2-i}X_{2,m+1}).$$

In practice, it would be easier to calculate the determinant of AX for any positive integer p and it is not necessary to check the conditions provided above.

Comparisons of the Proposed Estimators With OLSE as a Function of d_n in Model II

Since both the proposed estimators and the OLSE of the regression coefficients are unbiased and consistent, it is interesting to compare the elements, the trace, and the determinant of the variance covariance matrix for the estimators of the two different procedures.

In the specified Model II, $\hat{\beta}_{ols} = (X'X)^{-1}X'y$ and $\hat{\beta}_{mt} = (X'A'AX)^{-1}X'A'Ay$, are the estimators for the two procedures, where A is the transformation matrix defined in section 3 of this article. The corresponding variance-covariance matrices are

$$V(\hat{\beta}_{ols}) = (X'X)^{-1}X'VX(X'X)^{-1} \text{ and } V(\hat{\beta}_{mt}) = (X'A'AX)^{-1}(a+bd_{m+1}).$$

Let $X' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix}$ in the stated Model II. For the ordinary least squares case, it can be shown (Appendix B) that

$$V(\hat{\beta}_{1ols}) = \frac{1}{|X'X|^2} \{ \sum_{i=1}^{n-1} (a+bd_i) [(\sum_{j=1}^{n} X_j^2) - X_i(\sum_{j=1}^{n} X_j)]^2$$

$$+ (a+bd_n) [(\sum_{j=1}^{n} X_j^2) - X_n(\sum_{j=1}^{n} X_j)]^2 \quad \text{-----------(5.1)}$$

$$COV(\hat{\beta}_{1ols}, \hat{\beta}_{2ols}) = \frac{1}{|X'X|^2} \{ \sum_{i=1}^{n-1} (a+bd_i)(\sum_{j=1}^{n} X_j - nX_i)(X_i \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_i^2)$$

$$+ (a+bd_n)(\sum_{j=1}^{n} X_j - nX_n)(X_n \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2) \text{-------} (5.2)$$

$$V(\hat{\beta}_{2ols}) = \frac{1}{|X'X|^2} \{ \sum_{i=1}^{n-1} (a+bd_i)[\sum_{j=1}^{n} X_j - nX_i]^2$$

$$+ (a+bd_n)[\sum_{j=1}^{n} X_j - nX_n]^2 \} \text{-------------------} (5.3)$$

$$|V(\hat{\beta}_{ols})| = \frac{1}{|X'X|^2} \{ (\sum_{i=1}^{n-1} (a+bd_i)[\sum_{i=1}^{n-1} (a+bd_i)X_i^2 - \sum_{i=1}^{n-1} (a+bd_i)X_i^2]\}$$

$$+ \frac{1}{|X'X|^2} \{ (\sum_{i=1}^{n-1} (a+bd_i)X_n^2 + [\sum_{i=1}^{n-1} (a+bd_i)X_n^2 - 2X_n \sum_{i=1}^{n-1} (a+bd_i)X_i]\}$$

$$(a+bd_n) \text{--} (5.4)$$

Let $X' = \begin{bmatrix} 1 & \cdots & 1 \\ X_1 & \cdots & X_n \end{bmatrix}$ in Model II. For the proposed transformation procedure, it can be shown (Appendix B)

$$V(\hat{\beta}_{1mt}) = (a+bd_{m+1}) \frac{z^2(X_1^2 - X_n + C) - 2z\sqrt{d_{m+1}-d_1}\, X_1 X_n + C(d_{m+1}-d_1)}{Dz^2 - 2E\sqrt{d_{m+1}-d_1}\, z + F(d_{m+1}-d_1)} \text{----} (5.5)$$

where $z = \sqrt{d_n - d_{m+1}}$

$$A = (m+1)-2[\sqrt{\lambda_2(1-\lambda_2)} + \dots + \sqrt{\lambda_m(1-\lambda_m)}] \quad \text{where } n = 2m+1.$$

$$B = \lambda_2(X_m-X_{m+2}) + \dots + \lambda_{m+1}(X_1-X_n) - \sqrt{\lambda_2(1-\lambda_2)}\,(X_m+X_{m+2}) -$$

$$\dots - \sqrt{\lambda_{m+1}(1-\lambda_{m+1})}\,(X_1+X_n) + (X_{m+1} + \dots + X_n)$$

$$C = \lambda_2(X_m^2 - X_{m+2}^2) + \dots + \lambda_{m+1}(X_1^2 - X_n^2) - 2\sqrt{\lambda_2(1-\lambda_2)}\,X_m X_{m+2} -$$

$$\dots - 2\sqrt{\lambda_{m+1}(1-\lambda_{m+1})}\,X_1 X_n + (X_{m+1}^2 + \dots + X_n^2)$$

$$D = (A-1)X_1^2 + 2X_1X_n - (A+1)X_n^2 - 2B(X_1-X_n) + AC - B^2$$

$$E = 2AX_1X_n + 2C - 2BX_1 - 2BX_n$$

$$F = A.C - B^2$$

$$\lambda_2 = \frac{d_{m+2} - d_{m+1}}{d_{m+2} - d_m}$$

$$\vdots$$

$$\lambda_{m+1} = \frac{d_n - d_{m+1}}{d_n - d_1}$$

$$V(\hat{\beta}_{2mt}) = (a+bd_{m+1})^2 \frac{z^2A - 2z\sqrt{d_{m+1}-d_1} + A(d_{m+1}-d_1)}{Dz^2-2Ez\sqrt{d_{m+1}-d_1} + F\,(d_{m+1}-d_1)} \quad \text{-----------(5.6)}$$

$$COV(\hat{\beta}_{1mt}, \hat{\beta}_{2mt}) = \frac{-z^2(B+X_1-X_n) + z\sqrt{d_{m+1}-d_1}\,(X_1+X_n)-B(d_{m+1}-d_1)}{Dz^2 - 2E\sqrt{d_{m+1}-d_1}\,z + F\,(d_{m+1}-d_1)} \quad \text{----(5.7)}$$

$$|V(\hat{\beta}_{mt})| = (a+bd_{m+1})^2 \; \frac{z^2 + (d_{m+1}-d_1)}{z^2 D - 2z\sqrt{d_{m+1}-d_1}\,E + F\,(d_{m+1}-d_1)} \text{ -----------} (5.8)$$

It is only a matter of algebra to prove (Appendix B) that the diagonal elements and the determinant of $V(\hat{\beta}_{ols})$ are linear functions of d_n with positive slopes and intercepts. The off diagonal element of $V(\hat{\beta}_{ols})$ is a linear function of d_n with negative slope and either a positive or a negative intercept. On the other hand, the elements and trace of $V(\hat{\beta}_{mt})$ are smooth (continuous first derivative) and bounded (both from above and below) functions of d_n with two or fewer critical values. The determinant of $V(\hat{\beta}_{mt})$ is a smooth and bounded function of d_n with one critical value. Hence, analytical solutions of the intersection values of d_n can be obtained by simultaneous solution of the corresponding two equations (Appendix B). For any value of d_n which is greater than the largest value of the intersection values of d_n, the proposed estimator would do better than the OLSE in terms of the variance of the estimator.

In the case that n=3 for Model II, the discussion in the previous paragraph can be demonstrated by the following five figures.

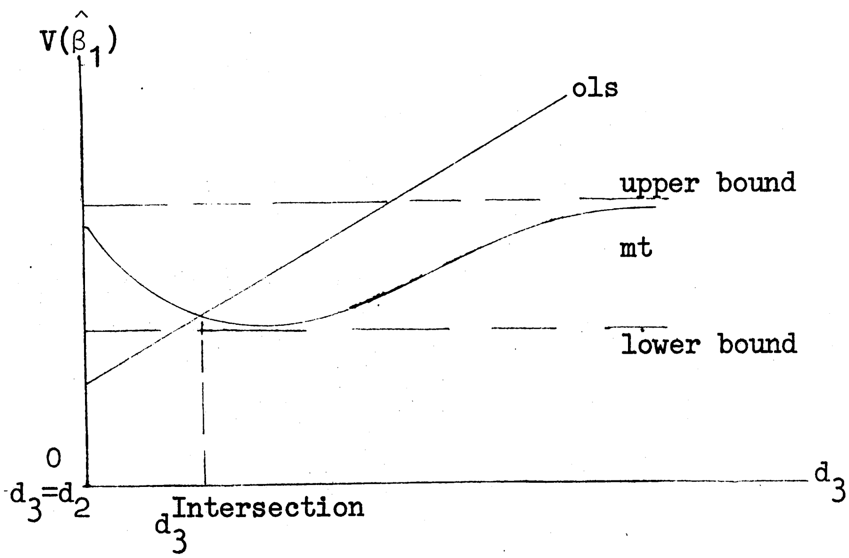

Figure 1. The Intersection of $V(\hat{\beta}_{1ols})$ and $V(\hat{\beta}_{1mt})$ as Functions of d_n

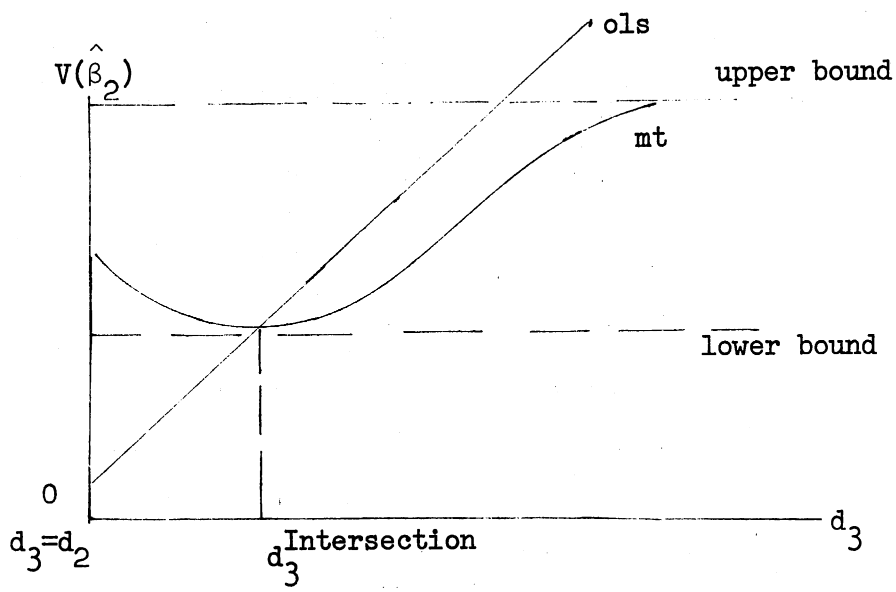

Figure 2. The Intersection of $V(\hat{\beta}_{2ols})$ and $V(\hat{\beta}_{2mt})$ as Functions of d_n

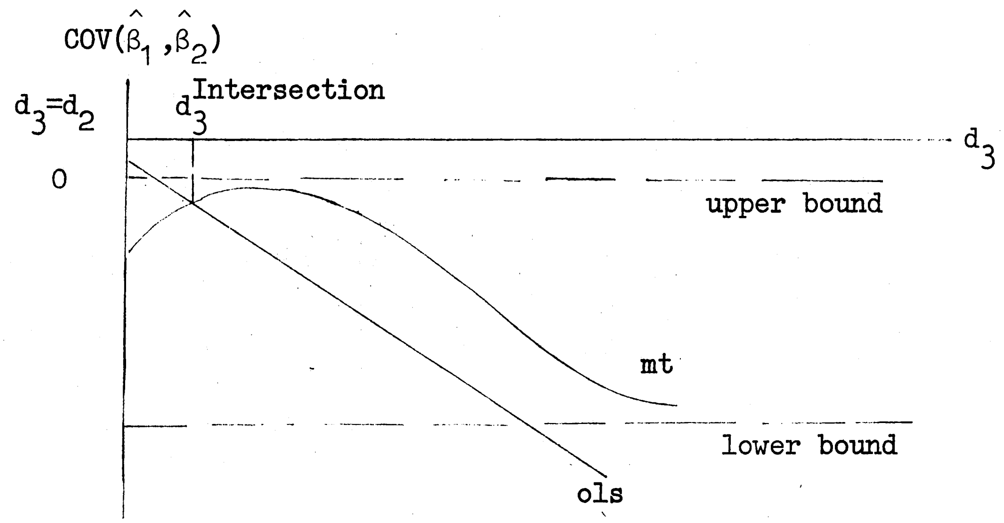

Figure 3. The Intersection of $COV(\hat{\beta}_{1ols}, \hat{\beta}_{2ols})$ and $COV(\hat{\beta}_{1mt}, \hat{\beta}_{2mt})$ as Functions of d_n

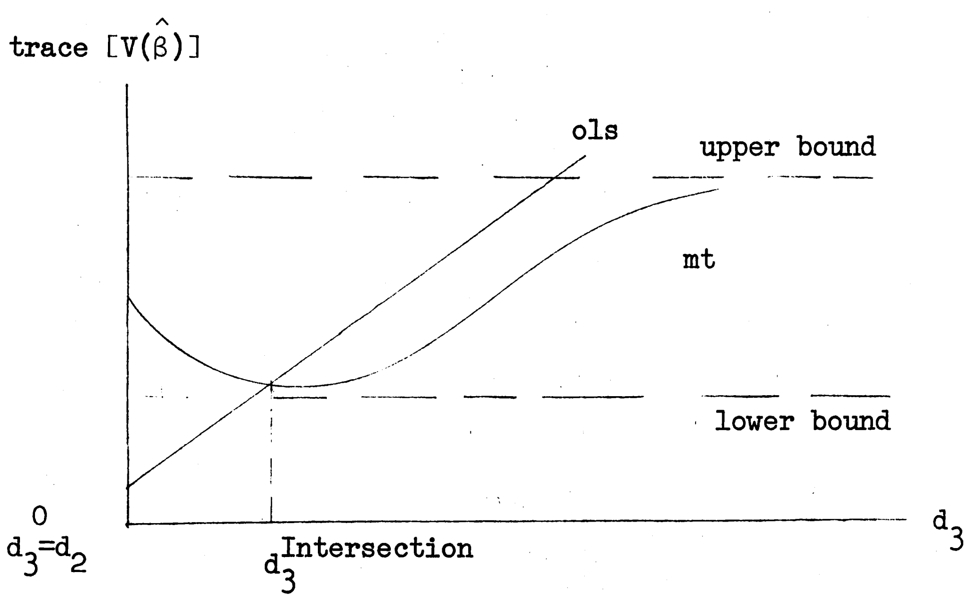

Figure 4. The Intersection of trace $[V(\hat{\beta}_{ols})]$ and trace $[V(\hat{\beta}_{mt})]$ as Functions of d_n

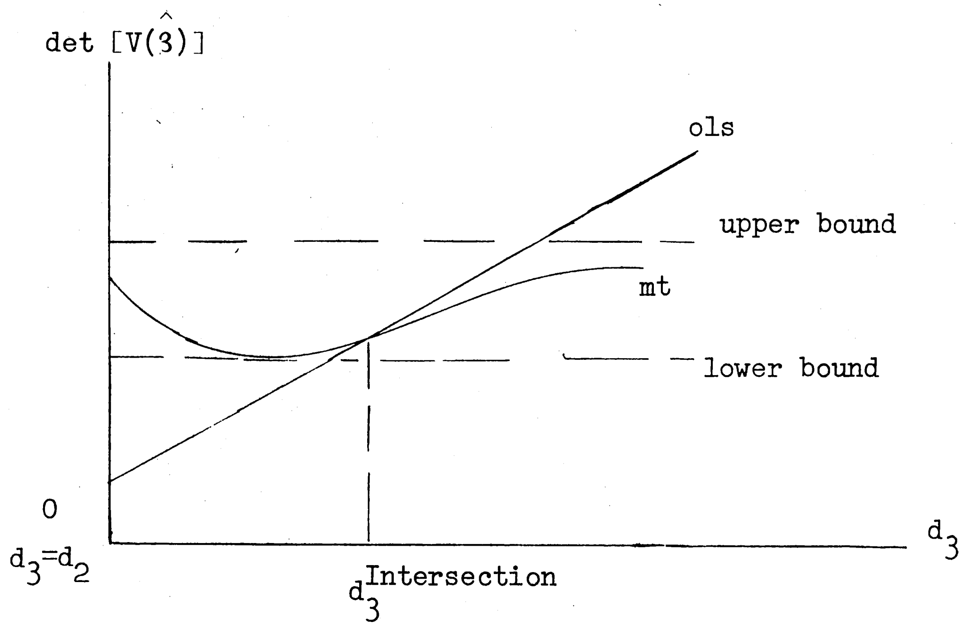

Figure 5. The Intersection of det $[V(\hat{\beta}_{ols})]$ and det $[V(\hat{\beta}_{mt})]$ as Functions of d_n

Summary and Open Problems

We have developed a transformation technique and an estimator of the regression coefficients $\boldsymbol{\beta}$, when the errors in the linear model are heterogeneous of a particular form. The proposed transformation matrix results in homogeneous variances and has the maximum possible rank. The proposed estimator is unbiased and consistent. When the relative size of the largest variance to the smallest variance in the stated regression model is large, the variances of the proposed estimators of intercept and slope will be smaller than the corresponding variances of the OLSE. The determinant of the variance covariance matrix of the proposed estimator is also smaller than the corresponding determinant of the OLSE.

The transformation matrix is not unique. There are m! pairing methods, m = n/2 if n is even and m = (n-1)/2 if n is odd. In addition, the two elements in a row of the matrix can have the same sign or different signs. For example, let n=5 and V = aI + bD, where

$$\underset{5\times5}{D} = \begin{bmatrix} d_1 & & & & \\ & d_2 & & & \\ & & d_3 & & \\ & & & d_4 & \\ & & & & d_5 \end{bmatrix} \quad \text{such that } 0<d_1<d_2<d_3<d_4<d_5.$$

Then there are m = (5-1)/2 = 2 pairing methods. The transformation matrix based on pairing method 1 is

$$\underset{3\times5}{A_1} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & \pm\sqrt{\frac{d_4-d_3}{d_4-d_2}} & 0 & \pm\sqrt{\frac{d_3-d_2}{d_4-d_2}} & 0 \\ \pm\sqrt{\frac{d_5-d_3}{d_5-d_1}} & 0 & 0 & 0 & \pm\sqrt{\frac{d_3-d_1}{d_5-d_1}} \end{bmatrix}.$$

The transformation matrix based on a second pairing method is

$$A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ \pm\sqrt{\frac{d_4-d_3}{d_4-d_1}} & 0 & 0 & \pm\sqrt{\frac{d_3-d_1}{d_4-d_1}} & 0 \\ 0 & \pm\sqrt{\frac{d_5-d_3}{d_5-d_2}} & 0 & 0 & \pm\sqrt{\frac{d_3-d_2}{d_5-d_2}} \end{bmatrix}.$$

Selection of the optimum pairing method needs further investigation.

The transformation was devised to equalize the variances by taking a weighted average of paired d_i's to obtain a predetermined common variance. For each given value of the independent variable in Model II, there is a corresponding value of the dependent variable which is the result of pooling two independent observations using the corresponding weights of the transformation. The transformation results in homoscedastic variances but the number of degrees of freedom is decreased and the range of values of the dependent variable is decreased.

If $k > 2$ or there are more than two variance components in Model I, then the theorem in section 3 guarantees the existence of the homogenizing transformation matrix A. In practice, the process described in this section should be repeated k-1 times to equalize the variances in Model I. For example, if k=3 and $V = aI + bD_1 + cD_2$ in Model I, then the following sequence should be followed:

First stage: Find A_1 such that $A_1(aI+bD_1+cD_2)A_1' = (a+bd_{m+1})I_{m+1} + cA_1D_2A_1'$ where $A_1D_2A_1'$ is an $(m+1)x(m+1)$ diagonal matrix (by Lemma 2).

Second stage: Assume (m+1) is an odd number, find A_2 such that $A_2((a+bd_{m+1})I_{m+1} + cA_1D_2A_1')A_2' = ((a+bd_{m+1}) + ce)I_{(m+2)/2}$ where e is the middle element on the diagonal of $A_1D_2A_1'$.

Hence, $A = A_2A_1$ is the transformation matrix needed to obtain homoscedasticity.

Finally, in the case where $V = aI + bU$ for any known p.s.d. matrix U, it is common knowledge that there exists an nxn orthogonal matrix P such that $PVP' = aI + bD$, where $D = PUP'$ is an nxn diagonal matrix. Consequently, the proposed transformation technique can apply to a general linear regression model with

$$V = a_1I + a_2U_1 + \ldots + a_kU_{k-1},$$

where k satisfies the upper limits provided by the formulas of Appendix A, and V satisfies the assumptions of the theorem after the orthogonal transformation.

REFERENCES

(1) Hartley, H. O., and Jayatillake, K. S. E. "Estimation of Linear Models with Unequal Variances" JASA, 1970, No. 341, p. 189.

(2) Jacquez, J. A. "Sampling Experiments on the Estimation of Parameters in Heteroscedastic Linear Regression" Biometrics, 1973, 29, pp. 771-9.

(3) Maddala, G. S., and Mount, T. D. "A Comparative Study of Alternative Estimators for Variance Components Models Used in Econometric Applications" JASA, 1973, 68, No. 342, pp. 324-8.

(4) Malinvaud, E., "Statistical Method of Econometrics" Rand McNally and Company, 1966, pp. 254-8.

(5) Marcus, M., and Ming, H. "A Survey of Matrix Theory and Matrix Inequalities," Allyn and Bacon, Inc., Boston, 1964, p. 27.

(6) Rao, C. R. "Linear Statistical Inference and Its Applications" John Wiley and Sons, Inc., 1965, pp. 302-9.

(7) Rao, C. R. "Estimation of Heteroscedastic Variances in Linear Models" JASA, 1970, 65, pp. 161-172.

(8) Rao, C. R. "Estimation of Variance and Covariance Components in Linear Models" JASA, 1972, 67, pp. 112-5.

(9) Rutemiller, H. C., and Bower, D. A. "Estimation in a Heteroscedastic Regression Model" JASA, 1968, pp. 552-7.

(10) Takeshi, Amemiya "Regression Analysis When Variance of the Dependent Variable is Proportional to the Square of Its Expectation" JASA, 1973, 68, No. 344, pp. 928-934.

(11) Theil, Henri "Principles of Econometrics" John Wiley and Sons, Inc., 1971, pp. 160, 196-9, 214-6, 244-50, 362-5, 403-5, 624-5.

(12) Watson, G. S., and Bloomfield, P. "The Inefficiency of Least Squares" Biometrika, 1975, 62, No. 1, pp. 121-8.

(13) Zyskind, G. "On Canonical Forms, Non-negative Covariance Matrices and Best and Simple Least Squares Linear Estimators in Linear Models" Ann. Math. Stat., 1967, 38, 1092-1109.

CHAPTER III

ESTIMATION OF VARIANCE COMPONENTS IN A LINEAR REGRESSION MODEL WHEN THE VARIANCE IS A LINEAR FUNCTION OF UNKNOWN PARAMETERS

Abstract

The problem considered in this Chapter involves a linear model in which the variance of the observable random variable is a linear function of unknown variance components. The objective is to estimate the variance components. This is accomplished by combining the "orthonormal basis of error space" technique and a proposed maximum rank orthogonal linear transformation technique to estimate the variance components in a linear regression model when the variance is a linear function of unknown parameters. There always exists a proposed unbiased estimator whereas the estimator based on MINQUE may not exist. Furthermore, the proposed method provides a simple calculation procedure where MINQUE may require a generalized inverse procedure.

Introduction

Model I

Let $y = X\beta + \varepsilon$ denote a linear regression model where

y is an nx1 vector of observable random variables,

X is an nxp matrix of known constants with rank $(X) = p$,

β is a px1 vector of unknown constants,

ε is an nx1 vector of unobservable jointly normal random variables such that $V(\varepsilon) = V = \text{diagonal } \{\sigma_i^2: i=1, 2, \ldots, n\}$,

where (1) $\sigma_i^2 = \sum_{j=1}^{k} r_j d_{ij}$ is positive

(2) r_i's are unknown constants

(3) $d_{ij} \neq d_{i'j}$ for $i \neq i'$ and d_{ij}'s are known positive constants

(4) k satisfies the conditions in Appendix A of this article, which is required for the β_i's to be estimable.

Model II

The same as Model I except $V(\varepsilon) = aI + bD$ where

$a = r_1$

$b = r_2$

$D = \text{diagonal } \{d_i, i=1, 2, \ldots, n \text{ and } 0<d_1<d_2< \ldots <d_n\}$.

Some of the general methods for estimating variance components in a general linear model are "orthonormal basis of error space" (OBES), minimum norm quadratic unbiased estimation (MINQUE), and maximum likelihood estimation (MLE).

Putter (4) proposed methods of constructing OBES and applied his results to estimating variance components. For the case of an unreplicated two-way layout,

$$y_{jk} = u + a_j + b_k + e_{jk} (j=1, 2, \ldots, J;\ k=1, 2, \ldots, K),$$

he showed that when the relative differences among σ_j's, j=1, 2, ..., J are large estimators obtained from OBES would have smaller variances compared with the estimators proposed by Grubbs (2) and Ehrenberg (1).

Rao (5) introduced MINQUE for variance components in a general linear model. Unfortunately, the MINQUE estimators for $\sigma_i^2 = V(e_i)$ in $\underset{nx1}{y} = \underset{nxp}{X}\ \underset{px1}{\beta} + \underset{nx1}{e}$, although unbiased may be negative. Furthermore, for some structures of the X matrix, not all σ_i^2's are estimable.

Hartley and Jayatillake (3) estimated the variance components by MLE under the assumption of a lower bound for the variance components. The estimators are consistent and asympototically efficient. However, it is known that such estimators may not perform optimally for small sample sizes.

A Proposed Method of Variance Components Estimation

In this Chapter, we are going to combine OBES with the transformation technique developed in the previous Chapter in order to estimate the variance components in Model I.

For simplicity, we assume the specified Model II. Either applying the Gram-Schmidt orthonormalization procedure to $(I-X(X'X)^{-1}X')$ or based on methods proposed by Putter (4), we can find an $(n-p)xn$ matrix H in the orthogonal column space of X such that $HX = \phi$ and $HH' = I_{(n-p)}$, premultiplying both sides of the linear model by H leads to

$$Hy = HX\beta + H\varepsilon = H\varepsilon,$$

where

$$E(Hy) = E(H\varepsilon) = \phi$$

$$V(Hy) = H(aI + bD)H'$$

$$= aI + bHDH'.$$

Since HDH' is p.d., there exists an orthogonal $(n-p)x(n-p)$ matrix P such that $P(HDH')P' = D^*$, where $D^* = \text{diag}\{\Delta_i, i=1,\ldots,n-p\}$. Premultiplying both sides of $Hy = H\varepsilon$ by P results in

$$PHy = PH\varepsilon$$

where

$$E(PHy) = \phi$$

$$V(PHy) = P(aI + bHDH')P'$$

$$= aI + bD^*.$$

Define $Z = PHy$; then $E(Z) = \phi$ and the elements in Z, say z_i, $i=1, 2, \ldots, n-p$, are mutually independent since Z is normal and $V(Z) = aI + bD^*$. Let $s_i = z_i^2 = Z'\Lambda_i Z$, $i=1, 2, \ldots, n-p$ and $S = [s_1, s_2, \ldots, s_{n-p}]'$, where Λ_i is a $(n-p)x(n-p)$ diagonal matrix with the i^{th} diagonal element one and all other elements zero.

Hence,

(1) s_i, $i=1, 2, \ldots, n-p$ are mutually independent sinze z_i $i=1, 2, \ldots, n-p$ are mutually independent.

(2) $E(s_i) = E(z_i^2) = V(z_i) = a + b\Delta_i$, i=1, 2, ..., n-p.

(3) $V(s_i) = V(z_i^2) = V(Z'\Lambda_i Z)$

$= 2 \text{ trace } (\Lambda_i(aI + bD^*))^2$

$= 2(a+b\Delta_i)^2$, i=1, 2, ..., n-p.

In order to estimate a and b, based on the previous results, we define the following linear regression model:

$$\begin{bmatrix} s_1 \\ s_2 \\ \cdot \\ \cdot \\ \cdot \\ s_{n-p} \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} + b \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta_{n-p} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_{n-p} \end{bmatrix} .$$

Letting $W = \begin{bmatrix} 1 & \Delta_1 \\ 1 & \Delta_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & \Delta_{n-p} \end{bmatrix}$ and $u = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_{n-p} \end{bmatrix}$

then the previous model can be rewritten as

$$S = W \begin{bmatrix} a \\ b \end{bmatrix} + u$$

where $E(u) = \phi$ and

$$V(u) = \begin{bmatrix} 2(a+b\Delta_1)^2 & & & & \\ & 2(a+b\Delta_2)^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & 2(a+b\Delta_{n-p})^2 \end{bmatrix}$$

This is the case that the variance of the dependent variable is proportional to the square of its expectation. As shown in the previous Chapter, the ordinary least squares estimators of a and b will be inefficient compared with the proposed maximum rank transformation estimators if the relative size of the largest variance to the smallest variance in the regression model is large.

Let $\gamma_1 = 2a^2$, $\gamma_2 = 4ab$, and $\gamma_3 = 2b^2$, then

$$V(U) = \begin{bmatrix} \gamma_1 + \gamma_2\Delta_1 + \gamma_3\Delta_1^2 & & & & \\ & \gamma_1 + \gamma_2\Delta_2 + \gamma_3\Delta_2^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & \gamma_1 + \gamma_2\Delta_{n-p} + \gamma_3\Delta_{n-p}^2 \end{bmatrix} .$$

Without loss of generality, assume (n-p) and (n-p+1)/2 are odd; then using the theorem proved in the previous Chapter, there exist matrices A_1, ((n-p+1)/2) x (n-p), and A_2, ((((n-p+1)/2)+1)/2) x ((n-p+1)/2), such that $A = A_2A_1$ and

$$AS = AW\begin{bmatrix} a \\ b \end{bmatrix} + AU$$

where $E(AU) = \phi$ and $V(AU) = (\gamma_1+\gamma_2\Delta_{\frac{n-p+1}{2}}+\gamma_3\Delta^2_{\frac{n-p+3}{4}})I_{\frac{n-p+3}{4}}$.

Based on Model II, the BLUE of [a b]' is

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (W'A'AW)^{-1}(W'A'AS).$$

The variance-covariance matrix of $\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}$ is

$$V(\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}) = (W'A'AW)^{-1}(\gamma_1 + \gamma_2\Delta_{\frac{n-p+1}{2}} + \gamma_3\Delta_{\frac{n-p+3}{4}}).$$

Discussion

Based on the theorem developed in the previous Chapter, the proposed method on variance components estimation may generalize to Model I. Since there always exists an (n-p)xn matrix H in the orthogonal column space of X, the existence of the proposed estimators is ensured in case the minimum norm quadratic unbiased estimators do not exist. Furthermore, the proposed method provides a simple calculation procedure where MINQUE may require a generalized inverse procedure.

Both the estimators based on MINQUE and the proposed method are unbiased. However, the proposed estimators may have larger variances in comparison with the estimators, if they exist, obtained by MINQUE.

Although an attempt to obtain an "optimum" (uniformly most powerful) test on the variance components has been made, no concrete results have been obtained. Since the generalized likelihood ratio is the ratio of probability density functions of linear combinations of noncentral Chi-squares, the closed form of the distribution of the test statistics is not obtainable. Intuitively, an approximate F-test based on the sum of squares due to full model and restricted model in the model $AS = AW \begin{bmatrix} a \\ b \end{bmatrix} + AU$ may be used yet the properties of this approximate test are unknown.

REFERENCES

(1) Ehrenberg, A. C. S. "The Unbiased Estimation of Heterogenous Error Variances" Biometrika, 1950, 37:347-57.

(2) Grubbs, F. E. "On Estimating Precision of Measuring Instruments and Product Variability," JASA, 1948, 43:243-64.

(3) Hartley, H. O., and Jayatillaile, K. S. E. "Estimation for Linear Models With Unequal Variances," JASA, 1973, 68:189-192.

(4) Putter, J. "Orthonormal Bases of Error Spaces and Their Use For Investigating the Normality and Variances of Residuals," JASA, 1967, 62:1022-36.

(5) Rao, C. R. "Estimation of Heteroscedastic Variances in Linear Models," JASA, 1970, 65:161-171.

CHAPTER IV

SUMMARY

Model I

Let $y = X\beta + \varepsilon$ denote a linear regression model where

y is an nx1 vector of observable random variables,

X is an nxp matrix of known constants with rank $(X) = p$,

β is a px1 vector of unknown constants,

ε is an nx1 vector of unobservable random variables such that $V(\varepsilon) = V = \text{diagonal } \{\sigma_i^2 : i=1, 2, \ldots, n\}$,

where (1) $\sigma_i^2 = \sum_{j=1}^{k} r_j \, d_{ij}$ is positive

(2) r_i's are unknown constants

(3) $d_{ij} \neq d_{i'j}$ for $i \neq i'$ and d_{ij}'s are known positive constants

(4) k satisfies the conditions in Appendix A of this article, which is required for the β_i's to be estimable.

Model II

The same as Model I except $V(\varepsilon) = aI + bD$ where

$a = r_1$

$b = r_2$

$D = \text{diagonal } \{d_i,\ i=1,\ 2,\ \ldots,\ n \text{ and } 0<d_1<d_2< \ldots \leq d_n\}.$

We developed methods of estimation both for the regression coefficients and variance components in a linear regression model when the variance is a linear function of unknown parameters. In order to obtain homoscedasticity in a linear regression model when the variance is a linear function of unknown parameters, a maximum rank transformation was developed. After the transformation, BLUE estimators of the regression coefficients are obtained based on the transformed model.

For Model II, we proved that a full rank transformation which homogenizes the variances does not exist and the maximum rank of the transformation matrix is $m = n/2$ when n is even and $m = (n+1)/2$ when n is odd. So far, only one of the m! possible pairing methods for Model II has been studied in detail. Criteria such as determinant (generalized variance), norm, trace, or individual elements of the variance-covariance matrix of the estimates may be selected for the purpose of comparing any two possible pairing methods.

The proposed procedure led to regression coefficients estimators that are unbiased and consistent. In Model II, it was shown that the variances of the proposed estimators are smaller than the estimators based on OLSE when the relative size of the largest variance to the smallest variance in the regression model is large.

Variance components estimation was tried but with little success. We combined the OBES technique and the proposed transformation technique to obtain unbiased estimators of the variance components in Model II. The proposed method provides a simple calculation procedure where MINQUE may require a generalized inverse procedure. Furthermore, there always exists a proposed estimator whereas the MINQUE may not exist.

In general, most of the effort in this study was directed at finding small sample unbiased estimators of the regression coefficients and variance components in a linear regression model when the variance is a linear function of unknown parameters. We presented a method of estimation based on a maximum rank orthogonal linear transformation when the full rank orthogonal linear transformation does not exist. We proposed a method for constructing such a transformation matrix. Based on this transformation matrix, we derived the proposed estimators. We hope problems of selecting an optimum (minimum variance) transformation matrix among all possible transformation matrices, searching for a small sample best unbiased estimator for the regression coefficient vector, comparing the power of the tests of the regression coefficients based on the proposed procedure, ordinary least square estimation, or maximum likelihood estimation will stimulate further investigation to the problem of heteroscedasticity in a general linear regression model.

APPENDIX A

CONDITIONS OF k IN MODEL I (R_k IS A FUNCTION OF n AND k)

R_i, i = 1, 2, ..., k are the number of observations which remain after the i^{th} transformation. For the p regression coefficients to be estimable, R_k should be greater than or equal to p, where k is the number of variance components in Model I.

$$R_0 = n$$

$$R_1 = \frac{R_0 + \text{Mod}(R_0,2)}{2} = \frac{n + \text{Mod}(n, 2)}{2}$$

.

.

.

$$R_k = \frac{R_{k-1} + \text{Mod}\ (R_{k-1}, 2)}{2} \geq p$$

R_k

n \ k	1	2	3
3	2		
4	2		
5	3	2	
6	3	2	
7	4	2	
8	4	2	
9	5	3	2
10	5	3	2
11	6	3	2
12	6	3	2
13	7	4	2
.			
.			
.			

APPENDIX B

COMPARISON OF $V(\hat{\beta}_{ols})$ AND $V(\hat{\beta}_{mt})$ IN MODEL II

This appendix concerns the algebraic derivation and comparisons of $V(\hat{\beta}_{ols})$ and $V(\hat{\beta}_{mt})$, as a function of d_n, when

$$X' = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ X_1 & X_2 & . & . & . & X_n \end{bmatrix}$$

and $V = aI + bD$ where $D = \text{diagonal } \{d_i: i=1, 2, \ldots, n\}$.

Denote

$$V(\hat{\beta}_{ols}) = \begin{bmatrix} V(\hat{\beta}_{1ols}) & V(\hat{\beta}_{12ols}) \\ V(\hat{\beta}_{12ols}) & V(\hat{\beta}_{2ols}) \end{bmatrix}$$

and

$$V(\hat{\beta}_{mt}) = \begin{bmatrix} V(\hat{\beta}_{1mt}) & V(\hat{\beta}_{12mt}) \\ V(\hat{\beta}_{12mt}) & V(\hat{\beta}_{2mt}) \end{bmatrix}$$

We are going to find the values of d_n for which equality holds between $V(\hat{\beta}_{1ols})$ and $V(\hat{\beta}_{1mt})$; between $V(\hat{\beta}_{12ols})$ and $V(\hat{\beta}_{12mt})$; between $V(\hat{\beta}_{2ols})$ and $V(\hat{\beta}_{2mt})$; and between $|V(\hat{\beta}_{ols})|$ and $|V(\hat{\beta}_{mt})|$.

I. $V(\hat{\beta}_{ols})$

$$V(\hat{\beta}_{ols}) = (X'X)^{-1}X'VX(X'X)^{-1}$$

$$= \frac{1}{|X'X|^2} \begin{bmatrix} \sum_{i=1}^{n} X_i^2 & -\sum_{i=1}^{n} X_i \\ -\sum_{i=1}^{n} X_i & n \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} (a+bd_i) & \sum_{i=1}^{n} (a+bd_i)X_i \\ \sum_{i=1}^{n} (a+bd_i)X_i & \sum_{i=1}^{n} (a+bd_i)X_i^2 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} X_i^2 & -\sum_{i=1}^{n} X_i \\ -\sum_{i=1}^{n} X_i & n \end{bmatrix}$$

A. $\left|V(\hat{\beta}_{ols})\right|$:

$$\left|V(\hat{\beta}_{ols})\right| = \frac{1}{|X'X|^2}\{[\sum_{i=1}^{n}(a+bd_i)][\sum_{i=1}^{n}(a+bd_i)X_i^2] - [\sum_{i=1}^{n}(a+bd_i)X_i]^2\}$$

$$= \frac{1}{|X'X|^2}\{[\sum_{i=1}^{n-1}(a+bd_i)][\sum_{i=1}^{n-1}(a+bd_i)X_i^2] + [\sum_{i=1}^{n-1}(a+bd_i)X_i^2]d_n + [\sum_{i=1}^{n-1}(a+bd_i)X_i^2]d_n + d_n^2X_n^2 - [\sum_{i=1}^{n-1}(a+bd_i)X_i]^2 + 2[\sum_{i=1}^{n-1}(a+bd_i)X_i]d_nX_n + d_n^2X_n^2\}$$

$$= \frac{1}{|X'X|^2}\{[\sum_{i=1}^{n-1}(a+bd_i)][\sum_{i=1}^{n-1}(a+bd_i)X_i^2] - [\sum_{i=1}^{n-1}(a+bd_i)X_i]^2\} + \frac{1}{|X'X|^2}\{[\sum_{i=1}^{n-1}(a+bd_i)X_n^2] + [\sum_{i=1}^{n-1}(a+bd_i)X_i^2] - 2[\sum_{i=1}^{n-1}(a+bd_i)X_i]X_n\}d_n$$

Thus, $|V(\hat{\beta}_{ols})|$ is a linear function of d_n

where

(a) the intercept, the first term of $|V(\hat{\beta}_{ols})|$, is positive since $|X'X|^2 > 0$ and V_{n-1} is p.d. which implies $|X'_{n-1}V_{n-1}X_{n-1}| > 0$, and

(b) the slope, the coefficient of d_n, is positive since $|X'X|^2 > 0$ and

$$\{[\sum_{i=1}^{n-1}(a+bd_i)]X_n^2 + \sum_{i=1}^{n-1}(a+bd_i)X_i^2 - 2[\sum_{i=1}^{n-1}(a+bd_i)X_n]\}$$

$$= \sum_{i=1}^{n-1}(a+bd_i)(X_n - X_i)^2 > 0.$$

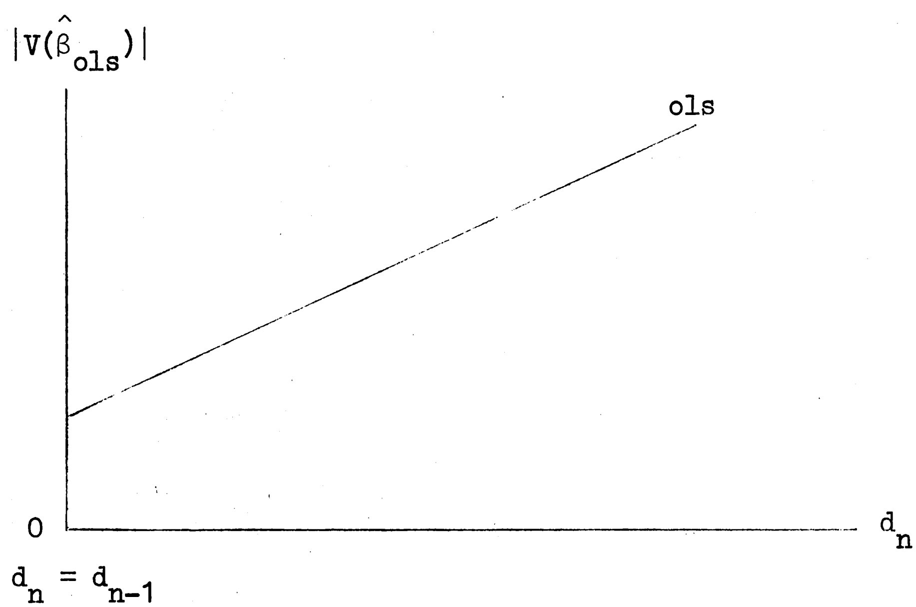

Figure 6. Det $[V(\hat{\beta}_{ols})]$ as a Function of d_n

B. $V(\hat{\beta}_{1ols})$:

$$V(\hat{\beta}_{1ols}) = \frac{1}{|X'X|^2} \{ \sum_{i=1}^{n-1} (a+bd_i) [(\sum_{j=1}^{n} X_j^2) - X_i(\sum_{j=1}^{n} X_j)]^2$$

$$+ (a+bd_n) [(\sum_{j=1}^{n} X_j^2) - X_n(\sum_{j=1}^{n} X_j)]^2 \}$$

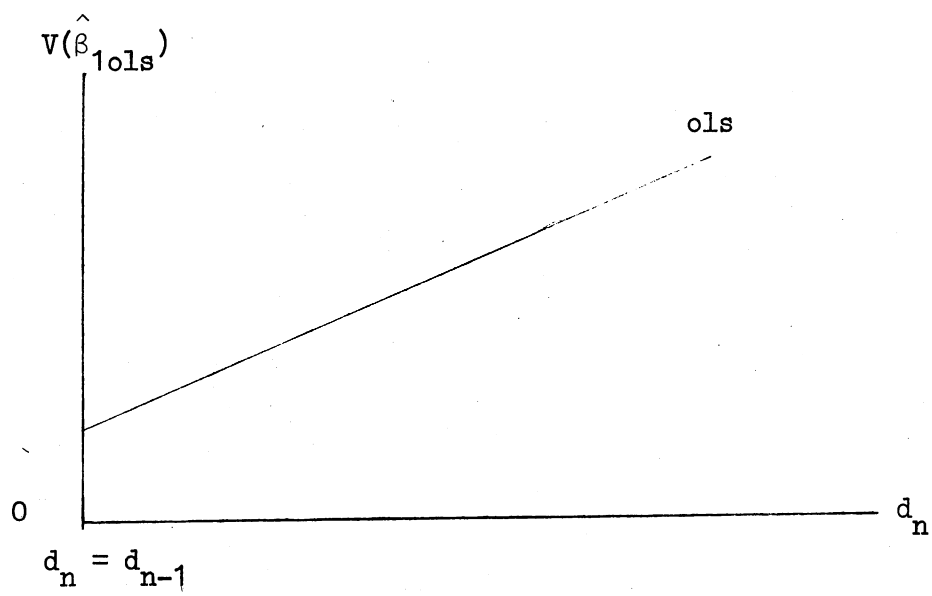

Figure 7. $V(\hat{\beta}_{1ols})$ as a Function of d_n

C. $V(\hat{\beta}_{12ols})$:

$$V(\hat{\beta}_{12ols}) = \frac{1}{|X'X|^2} \{ \sum_{j=1}^{n-1} (a+bd_j)(\sum_{i=1}^{n} X_i - nX_1)$$

$$(X_1 \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} X_i^2) + (a+bd_n)(\sum_{i=1}^{n} X_i - nX_n)$$

$$(X_n \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} X_i^2)\}$$

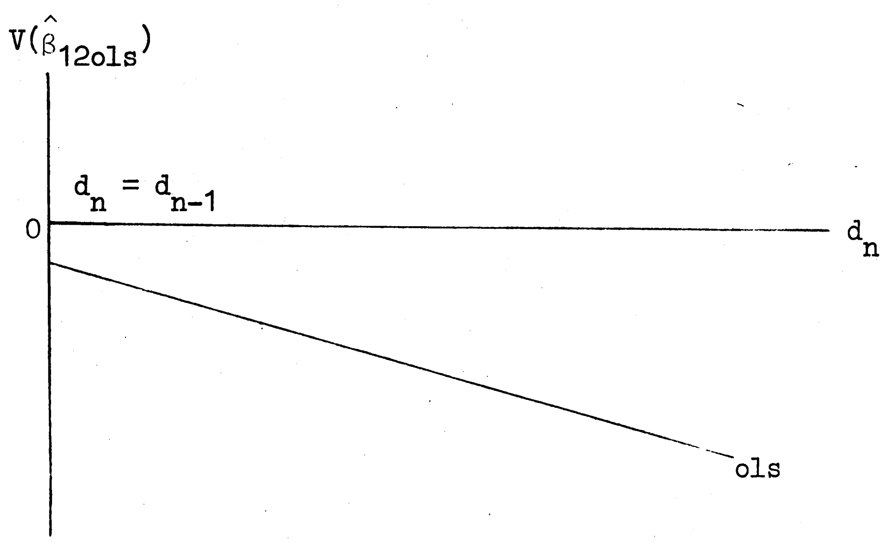

Figure 8. $V(\hat{\beta}_{12ols})$ as a Function of d_n

D. $V(\hat{\beta}_{2ols})$:

$$V(\hat{\beta}_{2ols}) = \frac{1}{|X'X|^2} \{ \sum_{j=1}^{n-1} (a+bd_j) [(\sum_{i=1}^{n} X_i) - nX_j]^2$$

$$+ (a+bd_n) [(\sum_{i=1}^{n} X_i) - nX_n]^2 \}$$

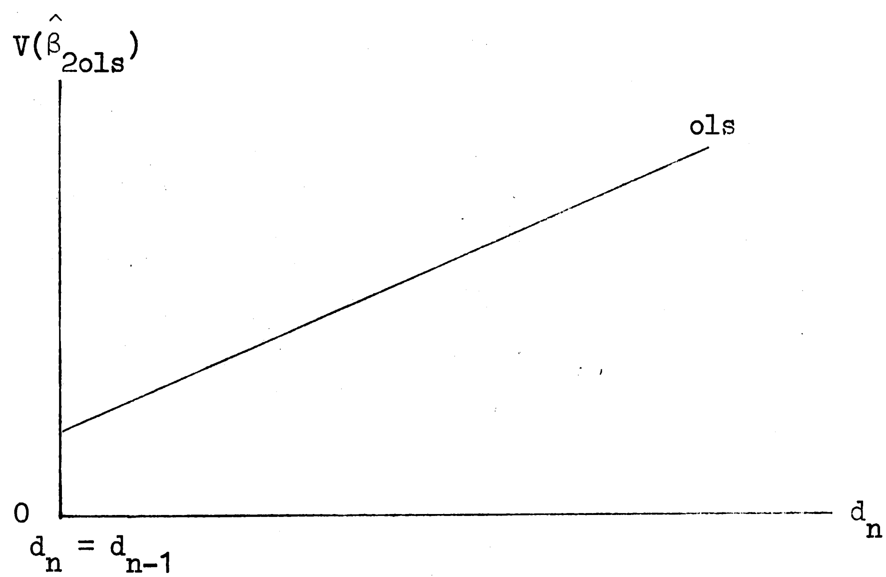

Figure 9. $V(\hat{\beta}_{2ols})$ as a Function of d_n

II. $V(\hat{\beta}_{mt}) = (X'A'AX)^{-1}(a+bd_{m+1})$ and assume $n = 2m + 1$. By definition of matrix A, we have

$$A'A = \begin{bmatrix} \lambda_{m+1} & & & & & & -\sqrt{\lambda_{m+1}(1-\lambda_{m+1})} \\ & \ddots & & & & \cdot^{\cdot^{\cdot}} & \\ & & \lambda_2 & & -\sqrt{\lambda_2(1-\lambda_2)} & & \\ & & & 1 & & & \\ & & -\sqrt{\lambda_2(1-\lambda_2)} & & 1-\lambda_2 & & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots & \\ -\sqrt{\lambda_{m+1}(1-\lambda_{m+1})} & & & & & & 1-\lambda_{m+1} \end{bmatrix}$$

where $\lambda_2 = \dfrac{d_{m+2} - d_{m+1}}{d_{m+2} - d_m}$

$\vdots$

$$\lambda_{m+1} = \frac{d_n - d_{m+1}}{d_n - d_1}$$

$$X'A'AX = \begin{bmatrix} 1 & 1 & \cdot & \cdot & \cdot & 1 \\ X_1 & X_2 & \cdot & \cdot & \cdot & X_n \end{bmatrix} A'A \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

$$XA'AX = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \quad \text{where}$$

$$D_{11} = (m+1) - 2(\sqrt{\lambda_{m+1}(1-\lambda_{m+1})} + \ldots + \sqrt{\lambda_2(1-\lambda_2)}\ )$$

$$D_{12} = \lambda_2(X_m - X_{m+2}) + \ldots + \lambda_{m+1}(X_1 - X_n) - \sqrt{\lambda_2(1-\lambda_2)}(X_m + X_{m+2}) - \ldots$$
$$- \sqrt{\lambda_{m+1}(1-\lambda_{m+1})}\ (X_1 + X_n) + (X_{m+1} + \ldots + X_n)$$

$$D_{21} = D_{12}$$

$$D_{22} = \lambda_2(X_m^2 - X_{m+2}^2) + \ldots + \lambda_{m+1}(X_1^2 - X_n^2) - 2\sqrt{\lambda_2(1-\lambda_2)}\ X_m X_{m+2} - \ldots$$
$$-2\sqrt{\lambda_{m+1}(1-\lambda_{m+1})}\ X_1 X_n + (X_{m+1}^2 + \ldots + X_n^2).$$

Letting

$$A = (m+1) - 2[\sqrt{\lambda_2(1-\lambda_2)} + \ldots + \sqrt{\lambda_m(1-\lambda_m)}\ ]$$

$$B = \lambda_2(X_m - X_{m+2}) + \ldots + \lambda_m(X_2 - X_{n-1}) - \sqrt{\lambda_2(1-\lambda_2)}\ (X_m + X_m + 2) - \ldots$$
$$- \sqrt{x_m(1-\lambda_m)}\ (X_2 + X_{n-1}) + \sum_{i=m+1}^{n} X_i$$

and

$$C = \lambda_2(X_m^2 - X_{m+2}^2) + \ldots + \lambda_m(X_2^2 - X_{m-1}^2) - 2\sqrt{\lambda_2(1-\lambda_2)}\ X_m X_{m+2} - \ldots$$
$$2\sqrt{\lambda_m(1-\lambda_m)}\ X_2 X_{n-1} + \sum_{i=m+1}^{n} X_i^2,$$

then

$$|X'A'AX| = 2\sqrt{\lambda_{m+1}(1-\lambda_{m+1})}\ \ [AX_1X_n + C - B\ (X_1 + X_n)] + [AC - B^2].$$

A. $|V(\hat{\beta}_{mt})|$:

1. $$|V(\hat{\beta}_{mt})| = |(X'A'AX)^{-1}(a+bd_{m+1})| = (a+bd_{m+1})^2|X'A'AX|^{-1}$$

$$= (a+bd_{m+1})^2\{\frac{d_n-d_{m+1}}{d_n-d_1}[(A-1)X_1^2 + 2X_1X_n - (A+1)X_n^2$$

$$- 2B(X_1-X_n)] - \frac{\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}}{d_n - d_1}[2AX_1X_n + 2C$$

$$- 2BX_1 - 2BX_n] + (AC - B^2)\}^{-1}.$$

Letting $D = (A - 1)X_1^2 + 2X_1X_n - (A+1)X_n^2 - 2B(X_1-X_n) + AC - B^2$

$E = 2AX_1X_n + 2C - 2BX_1 - 2BX_n$, and

$F = AC - B^2$,

then

$$|V(\hat{\beta}_{mt})| = (a+bd_{m+1})^2(d_n-d_1)\ \{(d_n+d_{m+1})D - 2\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}$$

$$E + (d_{m+1} - d_1)F\}^{-1}.$$

Letting $Z = \sqrt{d_n - d_{m+1}} \geq 0$, then

$$|V(\hat{\beta}_{mt})| = [\frac{Z^2 + (d_{m+1} - d_1)}{DZ^2 - 2E\sqrt{d_{m+1}-d_1}\ \ Z + F(d_{m+1}-d_1)}]\ (a+bd_{m+1})^2\ .$$

2. The nature of this function is determined by evaluation at $d_n = d_{n-1}$, $d_n \to +\infty$ and at d_n^m, the critical point of this function.

a. $$|V(\hat{\beta}_{mt})|_{d_n=d_{n-1}} = \{(a+bd_{m+1})^2(d_{n-1}-d_1)\}$$
$$/\{[D(d_{n-1}-d_{m+1})-2E\sqrt{(d_{m+1}-d_1)(d_{n-1}-d_{m+1})}+F(d_{m+1}-d_1)]\}$$

b. $$|V(\hat{\beta}_{mt})|_{d_n\to+\infty} = \frac{(a+bd_{m+1})^2}{D}$$

c. $$\frac{\partial}{\partial Z}|V(\hat{\beta}_{mt})| =\{[DZ^2-2E\sqrt{d_{m+1}-d_1}\ Z+F(d_{m+1}-d_1)]\ [2Zd_{m+1}^2]$$
$$-[Z^2+(d_{m+1}-d_1)]\ [2DZ-2E\sqrt{d_{m+1}-d_1}]\ d_{m+1}^2\}/\{[DZ^2$$
$$-2E\sqrt{d_{m+1}-d_1}\ Z+F(d_{m+1}-d_1)]^2\}$$

$$= \frac{2d_{m+1}^2\sqrt{d_{m+1}-d_1}\ [-EZ^2+(F-D)\sqrt{d_{m+1}-d_1}\ Z+E(d_{m+1}-d_1)]}{[DZ^2-2E\sqrt{d_{m+1}-d_1}\ Z+F(d_{m+1}-d_1)]^2}$$

Let $\frac{\partial}{\partial Z}|V(\hat{\beta}_{mt})| = 0$, we have

$$-EZ^2+(F-D)\sqrt{d_{m+1}-d_1}\ Z+E(d_{m+1}-d_1) = 0.$$

Since $Z \geq 0$,

$$Z^m = \tfrac{1}{2}\left[\frac{F-D}{E}\sqrt{d_{m+1}-d_1}+\sqrt{[(\frac{F-D}{E})+4]\ (d_3-d_1)}\ \right]$$

is the only solution, and

$$d_n^m = d_{m+1} + (Z^m)^2.$$

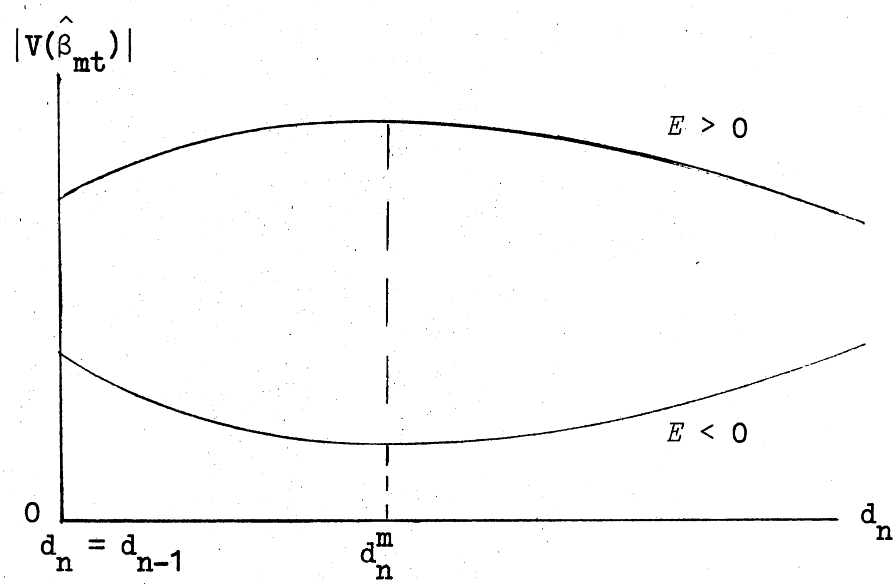

Figure 10. Det $[V(\hat{\beta}_{mt})]$ as a Function of d_n

B. $V(\hat{\beta}_{1mt})$:

1. $$V(\hat{\beta}_{1mt}) = \frac{C + \lambda_{m+1}\ (X_1^2 - X_n^2) - Z\sqrt{\lambda_{m+1}(-\lambda_{m+1})}\ X_1 X_n}{|X'A'AX|}\ (a+bd_{m+1})$$

$$= \frac{(a\mp bd_{m+1})}{|X'A'AX|}\ [C + (\frac{d_n - d_{m+1}}{d_n - d_1})(X_1^2 - X_n^2) -$$

$$2\frac{\sqrt{d_n - d_{m+1}}\sqrt{d_{m+1} - d_1}}{d_n - d_1}\ X_1 X_n]$$

$$= \frac{Z^2(X_1^2 - X_n^2 + C) - 2Z\sqrt{d_{m+1}-d_1}\, X_1X_n + C\,(d_{m+1}-d_1)}{DZ^2 - 2E\sqrt{d_{m+1}-d_1}\, Z + F\,(d_{m+1}-d_1)}\,(a+bd_{m+1})$$

2. The nature of this function is determined by evaluation at $d_n = d_{n-1}$, $d_n \to +\infty$ and at d_n^m, the critical points of this function.

a. $[V(\hat{\beta}_{1mt})]_{d_n = d_{n-1}} = (a+bd_{m+1}) \cdot$

$$\{(d_{n-1}-d_{m+1})(X_1-X_n+C) - 2\sqrt{(d_{n-1}-d_{m+1})(d_{m+1}-d_1)}\; X_1X_n$$

$$+C(d_{m+1}-d_1)\}/\{D(d_n-d_{m+1}) - 2E\sqrt{(d_{m+1}-d_1)(d_{n-1}-d_{m+1})} + F(d_{m+1}-d_1)\}$$

b. $[V(\hat{\beta}_{1mt})]_{d_n \to +\infty} = \dfrac{(X_1^2 - X_n^2 + C)}{D}\,(a+bd_{m+1})$

c. $\dfrac{\partial V(\hat{\beta}_{1mt})}{\partial Z}$ is a continuous function of d_n since $|X'A'AX|^2 > 0$, and $\dfrac{\partial V(\hat{\beta}_{1mt})}{\partial Z} = 0 \to$

$$2(a+bd_{m+1})\ \{DZ^3(X_1^2 - X_n^2 + C) - DZ^2\sqrt{d_{m+1}-d_1}\; X_1X_n$$

$$- 2E\, Z^2\sqrt{d_{m+1}-d_1}\;(X_1^2 - X_n^2 + C) + 2EZ(d_{m+1}-d_1)X_1X_n$$

$$+ FZ(d_{m+1}-d_1)(X_1^2 - X_n^2 + C) - F(d_{m+1}-d_1)\sqrt{d_{m+1}-d_1}$$

$$X_1X_n\} - 2(a+bd_{m+1})\ \{DZ^3(X_1^2 - X_n^2 + C) - 2DZ^2\sqrt{d_{m+1}-d_1}$$

$$X_1X_n + DZC(d_{m+1}-d_1) - EZ^2\sqrt{d_{m+1}-d_1}\;(X_1^2 - X_n^2 + C)$$

$$+ 2EZ(d_{m+1}-d_1)X_1X_n - EC(d_{m+1}-d_1)^{3/2}\} = 0 \to$$

$$Z^2[D\sqrt{d_{m+1}-d_1}\, X_1X_n - E\sqrt{d_{m+1}-d_1}\, (X_1^2 - X_n^2 + C)]$$

$$+ Z\, [F(d_{m+1}-d_1)(X_1^2 - X_n^2 + C) - DC(d_{m+1}-d_1)]$$

$$+ [EC(d_{m+1}-d_1)\sqrt{d_{m+1}-d_1} - C(d_{m+1}-d_1)\sqrt{d_{m+1}-d_1}\, X_1X_n] = 0$$

Since this is a quadratic function of Z, and hence d_n, there are four possible cases for the dependence of $V(\hat{\beta}_{1mt})$ upon d_n. These are illustrated in Figure 11.

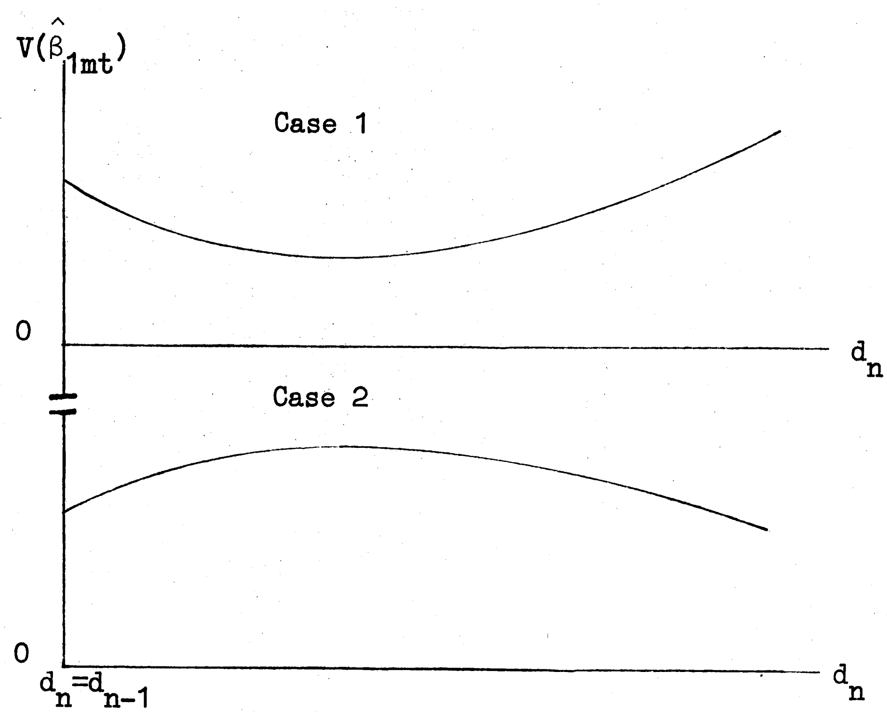

Figure 11. $V(\hat{\beta}_{1mt})$ as a Function of d_n

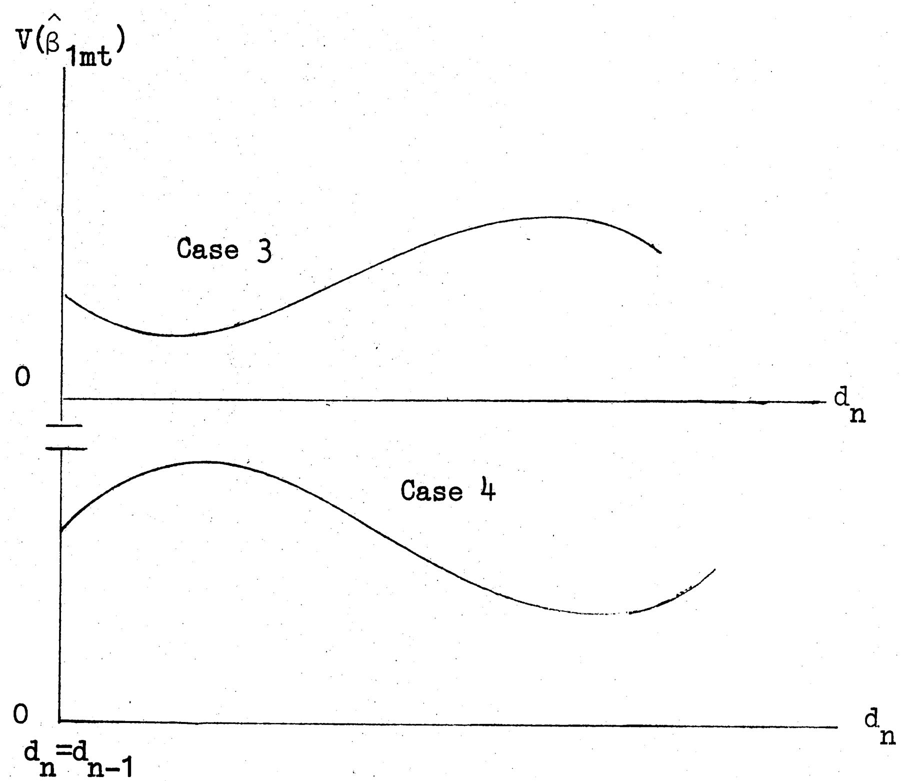

Figure 11. Continued

C. $V(\hat{\beta}_{2mt})$:

1. $$V(\hat{\beta}_{2mt}) = \frac{A - 2\sqrt{\lambda_{m+1}(1-\lambda_{m+1})}}{|X'A'AX|} (a+bd_{m+1})$$

$$= \frac{A[(d_n-d_{m+1}) + (d_{m+1}-d_1)] - 2\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}}{D\ (d_n-d_{m+1})-2E\sqrt{(d_{m+1}-d_1)(d_n-d_{m+1})} + F(d_{m+1}-d_1)} \cdot$$

$$(a+bd_{m+1})$$

$$= \frac{Z^2 A - 2Z\sqrt{d_{m+1}-d_1} + A(d_{m+1}-d_1)}{DZ^2 - 2EZ\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)} \quad (a+bd_{m+1})$$

2. The nature of this function is determined by evaluation at $d_n = d_{n-1}$, $d_n \to +\infty$ and at d_n^m, the critical points of this function.

a. $[V(\hat{\beta}_{2mt})]_{d_n = d_{n-1}} =$

$$\frac{A(d_{n-1}-d_{m+1}) + A(d_{m+1}-d_1) - 2\sqrt{(d_{n-1}-d_{m+1})(d_{m+1}-d_1)}}{D(d_{n-1}-d_{m+1}) - 2E\sqrt{(d_{n-1}-d_{m+1})(d_{m+1}-d_1)} + F(d_{m+1}-d_1)}(a+bd_{m+1})$$

b. $[V(\hat{\beta}_{2mt})]_{d_n \to +\infty} = \frac{A}{D}(a+bd_{m+1})$

c. $\frac{\partial V(\hat{\beta}_{2mt})}{\partial Z}$ is continuous function of d_n since

$|X'A'AX| > 0$, and $\frac{\partial V(\hat{\beta}_{2mt})}{\partial Z} = 0 \to$

$$[DZ^2 - 2EZ\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)]^2[2AZ-2\sqrt{d_{m+1}-d_1}]$$

$$- [Z^2A - 2Z\sqrt{d_{m+1}-d_1} + A(d_{m+1}-d_1)][2DZ-2E\sqrt{d_{m+1}-d_1}] = 0$$

$\to$

$$Z^2\sqrt{d_{m+1}-d_1}\,[-EA + D] + ZA(d_{m+1}-d_1)[F - D]$$

$$+ (d_{m+1}-d_1)\sqrt{d_{m+1}-d_1}\,(EA - F) = 0$$

Since this is a quadratic function of Z, and hence d_n, there are four possible cases for the dependence of $V(\hat{\beta}_{2mt})$ upon d_n. These are illustrated in Figure 12.

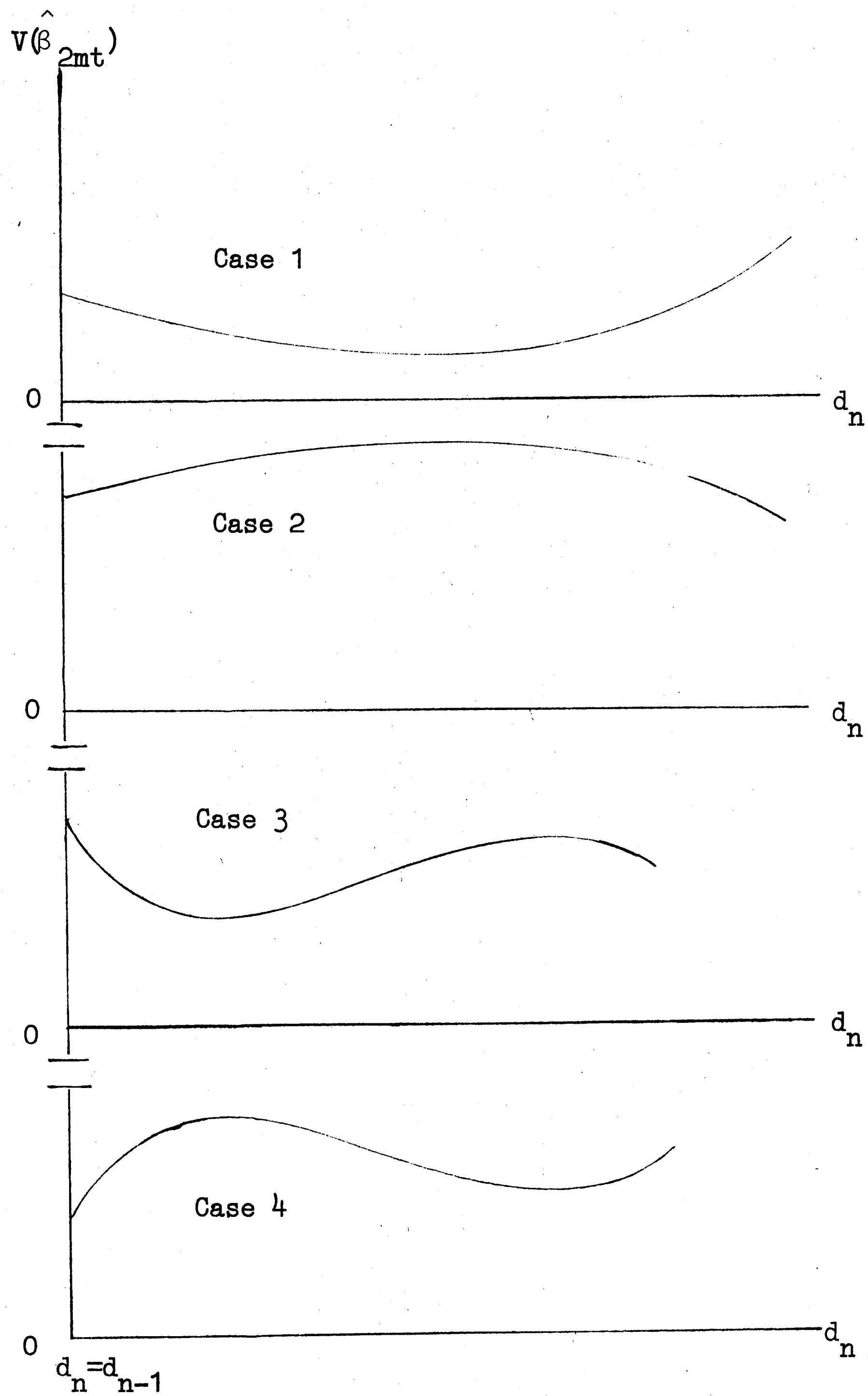

Figure 12. $V(\hat{\beta}_{2mt})$ as a Function of d_n

D. $V(\hat{\beta}_{12mt})$:

1. $$V(\hat{\beta}_{12mt}) = -\frac{[B + \lambda_{m+1}(X_1-X_n)-\sqrt{\lambda_{m+1}(1-\lambda_{m+1})}\ (X_1+X_n)]}{|X'A'AX|}(a+bd_{m+1})$$

$$= (a+bd_{m+1})\{-B[(d_n-d_{m+1})+(d_{m+1}-d_1)]-(d_n-d_{m+1})(X_1-X_n)$$

$$+\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}\ (X_1+X_n)\}/\{(d_n-d_1)|X'A'AX|\}$$

$$= \frac{-Z^2(B + X_1 - X_n)+Z\sqrt{d_{m+1}-d_1}\ (X_1+X_n) - B(d_{m+1}-d_1)}{DZ^2 - 2E\sqrt{d_{m+1}-d_1}\ Z + F(d_{m+1}-d_1)}(a+bd_{m+1})$$

2. The nature of this function is determined by evaluation at $d_n=d_{n-1}$, $d_n\to+\infty$ and at d_n^m, the critical points of this function.

a. $$[V(\hat{\beta}_{12mt})]_{d_n=d_{n-1}} = (a+bd_{m+1})\{-(d_{n-1}-d_{m+1})(B+X_1-X_n)$$

$$+\sqrt{(d_{n-1}-d_{m+1})(d_{m+1}-d_1)}\ (X_1+X_n)-B(d_{m+1}-d_1)\}/$$

$$/\{D(d_{n-1}-d_{m-1})-2E\sqrt{(d_{m+1}-d_1)(d_{n-1}-d_{m-1})}\ +F(d_{m+1}-d_1)\}$$

b. $$[V(\hat{\beta}_{12mt})]_{d_n\to+\infty} = \frac{-B + X_n - X_1}{D}(a+bd_{m+1})$$

c. $\dfrac{\partial V(\hat{\beta}_{12mt})}{\partial Z}$ is a continuous function of d_n since $|X'A'AX| > 0$, and $\dfrac{\partial V(\hat{\beta}_{12mt})}{\partial Z} = 0 \to$

$$[DZ^2 - 2E\sqrt{d_{m+1}-d_1}\; Z + F(d_{m+1}-d_1)][-2Z(B+X_1-X_n+\sqrt{d_{m+1}-d_1}(X_1+X_n)]$$

$$-[-Z^2(B+X_1-X_n)+Z\sqrt{d_{m+1}-d_1}(X_1+X_n)-B(d_{m+1}-d_1)][2DZ-2E\sqrt{d_{m+1}-d_1}] = 0$$

$$\rightarrow$$

$$Z^2[2E\sqrt{d_{m+1}-d_1}\;(B+X_1-X_n) - D\sqrt{d_{m+1}-d_1}\;(X_1+X_n)]$$

$$+ Z[-2F(d_{m+1}-d_1)(B+X_1-X_n) + 2DB(d_{m+1}-d_1)]$$

$$+ (d_{m+1}-d_1)\sqrt{d_{m+1}-d_1}\;[F(X_1+X_n) - 2\,BE] = 0$$

Since this is a quadratic function of Z, and hence d_n, there are four possible cases for the dependence of $V(\hat{\beta}_{12mt})$ upon d_n. These are illustrated in Figure 13.

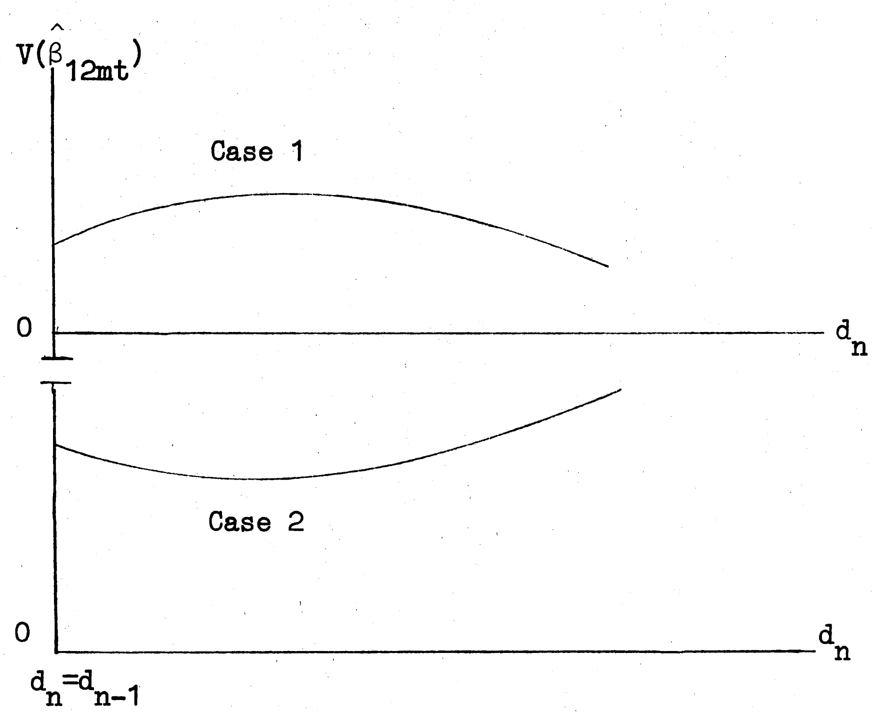

Figure 13. $V(\hat{\beta}_{12mt})$ as a Function of d_n

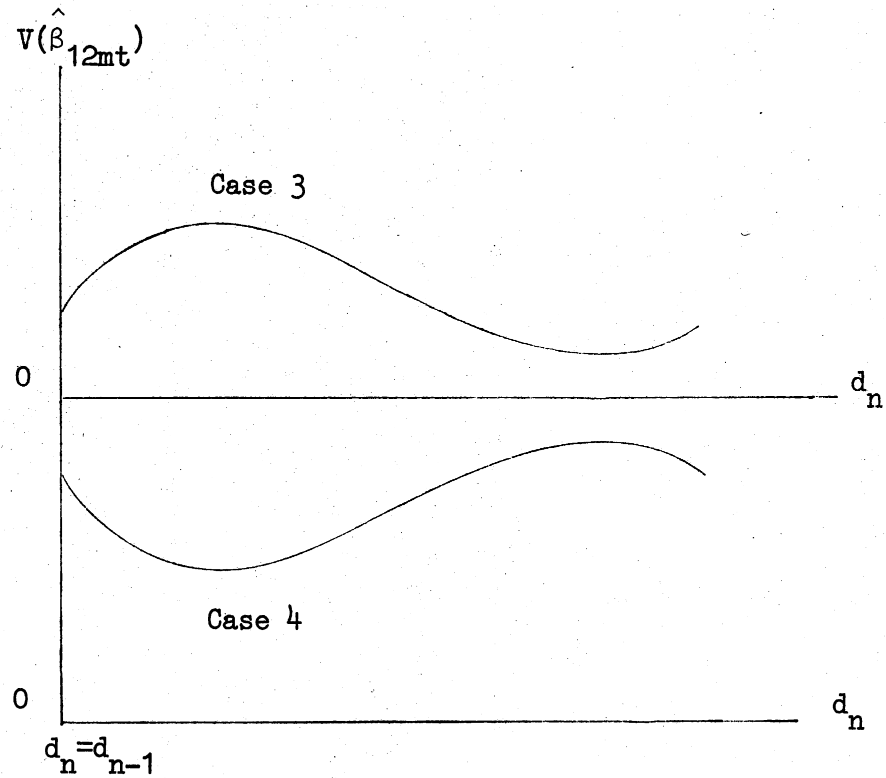

Figure 13. Continued

III. A. We next find d_n such that $V(\hat{\beta}_{1ols}) = V(\hat{\beta}_{1mt})$. The solution will be denoted as d_n^*.

$$V(\hat{\beta}_{1ols}) = \frac{1}{|X'X|^2} \sum_{i=1}^{n} (a+bd_i)(\sum_{j=1}^{n} X_j - X_i \sum_{j=1}^{n} X_j)^2$$

$$V(\hat{\beta}_{1mt}) = (a+bd_{m+1})\{(d_n-d_{m+1})(X_1^2-X_n^2+C)-2\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}\, X_1X_n + C(d_{m+1}-d_1)\}/\{D(d_n-d_{m+1})-2E\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)} + F(d_{m+1}-d_1)\}$$

$V(\hat{\beta}_{1ols}) = V(\hat{\beta}_{1mt})\ \forall\ a,b \ni aI + bD$ is p.d. $\rightarrow$

A.1 $$\frac{1}{|X'X|^2}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{n} X_j - X_i \sum_{j=1}^{n} X_j\right)^2\right]$$

$$= \frac{(d_n-d_{m+1})(X_1^2-X_n^2+C)-2\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1}\ X_1X_n + C(d_{m+1}-d_1)}{D(d_n-d_{m+1}) - 2E\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)}$$

A.2 $$\frac{1}{|X'X|^2}\left[\sum_{i=1}^{n} d_i\left(\sum_{j=1}^{n} X_j - X_i \sum_{j=1}^{n} X_j\right)^2\right]$$

$$= d_{m+1}\frac{(d_n-d_{m+1})(X_1^2-X_n^2+C)-2\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1}\ X_1X_n+C(d_{m+1}-d_1)}{D(d_n-d_{m+1}) - 2E\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)}$$

$$\rightarrow \frac{d_{m+1}}{|X'X|^2}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{n} X_j - X_i \sum_{1}^{n} X_j\right)^2\right] = \frac{1}{|X'X|^2}\left[\sum_{i=1}^{n} d_i\left(\sum_{j=1}^{n} X_j - X_i \sum_{j=1}^{n} X_j\right)^2\right]$$

$$\rightarrow \quad d_n^* = \frac{\sum_{\substack{i=1\\ i\neq m+1}}^{n-1} d_i\left(\sum_{j=1}^{n} X_j - X_i \sum_{j=1}^{n} X_j\right)^2}{\sum_{j=1}^{n} X_j - X_n \sum_{j=1}^{n} X_j)^2}$$

if $d_n \geq d_n^*$ then $V(\hat{\beta}_{1ols}) \geq V(\hat{\beta}_{1mt})$

B. We next find d_n such that $V(\hat{\beta}_{2ols}) = V(\hat{\beta}_{2mt})$. The solution will be denoted as d_n^*.

$$V(\hat{\beta}_{2ols}) = \frac{1}{|X'X|^2}\sum_{i=1}^{n}(a+bd_i)\left(\sum_{j=1}^{n} X_j - nX_i\right)^2$$

$$V(\hat{\beta}_{2mt}) = (a+bd_{m+1}) \frac{(d_n-d_{m+1})A-2\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + A(d_{m+1}-d_1)}{D(d_n-d_{m+1})-2E\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)}$$

$V(\hat{\beta}_{2ols}) = V(\hat{\beta}_{2mt})$ $\forall$ a & b $\ni$ aI + bD is p.d. $\rightarrow$

B.1 $$\frac{1}{|X'X|^2} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} X_j - nX_i \right)^2$$

$$= \frac{(d_n-d_{m+1})A-2\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + A(d_{m+1}-d_1)}{D(d_n-d_{m+1}) - 2E\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)}$$

B.2 $$\frac{1}{|X'X|^2} \sum_{i=1}^{n} d_i \left(\sum_{j=1}^{n} X_j - nX_i \right)^2$$

$$= d_{m+1} \frac{(d_n-d_{m+1})A - 2\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + A(d_{m+1}-d_1)}{D(d_n-d_{m+1})-2E\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)}$$

$$\rightarrow \frac{d_{m+1}}{|X'X|^2} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} X_j - nX_i \right)^2 = \frac{1}{|X'X|^2} \sum_{i=1}^{n} d_i \left(\sum_{j=1}^{n} X_j - nX_i \right)^2$$

$$\rightarrow \quad d_n^* = \frac{\sum_{\substack{i=1 \\ i \neq m+1}}^{n-1} d_i \left(\sum_{j=1}^{n} X_j - nX_i \right)^2}{\left(\sum_{j=1}^{n} X_j - nX_n \right)^2}$$

$\rightarrow$ if $d_n \geq d_n^*$ then $V(\hat{\beta}_{2ols}) \geq V(\hat{\beta}_{2mt})$.

C. We next find d_n such that $|V(\hat{\beta}_{12ols})| = |V(\hat{\beta}_{12mt})|$. The solution will be denoted as d_n^*.

$$V(\hat{\beta}_{12ols}) = \frac{1}{|X'X|^2} \sum_{i=1}^{n} (a+bd_i)(\sum_{j=1}^{n} X_j - nX_i)(X_i \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2)$$

$$V(\hat{\beta}_{12mt}) = (a+bd_{m+1})\{-(d_n-d_{m+1})(B+X_1-X_n)+\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}$$
$$(X_1+X_n) - B(d_{m+1}-d_1)\}/\{D(d_n-d_{m+1})$$
$$-2E\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)} + F(d_{m+1}-d_1)\}$$

$V(\hat{\beta}_{12ols}) = V(\hat{\beta}_{12mt})$ $\forall$ a,b $\ni$ aI + bD is p.d. $\rightarrow$

C.1 $$\frac{1}{|X'X|^2} \sum_{i=1}^{n} (\sum_{j=1}^{n} X_j - nX_i)(X_i \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2)$$

$$= -\frac{-(d_n-d_{m+1})(B+X_1-X_n)+\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1}\,(X_1+X_n)-B(d_{m+1}-d_1)}{D(d_n-d_{m+1})-2E\sqrt{d_n-d_{m+1}}\sqrt{d_{m+1}-d_1} + F(d_{m+1}-d_1)}$$

C.2 $$\frac{1}{|X'X|^2} \sum_{i=1}^{n} d_i (\sum_{j=1}^{n} X_j - nX_i)(X_i \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2)$$

$$= (-d_{m+1})\{-(d_n-d_{m+1})(B+X_1-X_n)+\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}$$
$$(X_1+X_n)-B(d_{m+1}-d_1)\}/\{D(d_n-d_{m+1})-2E\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}$$
$$+ F(d_{m+1}-d_1)\}$$

$$\rightarrow \frac{d_{m+1}}{|X'X|^2} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} X_j - nX_i \right)\left(X_i \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2\right)$$

$$= \frac{1}{|X'X|^2} \sum_{i=1}^{n} d_i \left(\sum_{j=1}^{n} X_j - nX_i \right)\left(X_i \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2\right)$$

$$\rightarrow \quad d_n^* = \frac{\sum_{\substack{i=1 \\ i\neq m+1}}^{n-1} d_i \left(\sum_{j=1}^{n} X_j - nX_n \right)\left(X_n \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2\right)}{\left(\sum_{j=1}^{n} X_j - nX_n \right)\left(X_n \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} X_j^2\right)}$$

$\rightarrow$ if $d_n \geq d_n^*$ then $|V(\hat{\beta}_{12ols})| \geq |V(\hat{\beta}_{12mt})|$.

D. Find d_n^* for $|V(\hat{\beta}_{ols})| = |V(\hat{\beta}_{mt})| \;\forall\; a,b$

$\Rightarrow aI + bD$ to be p.d.

$$|V(\hat{\beta}_{ols})| = \frac{|X'VX|}{|X'X|^2} = \frac{1}{|X'X|^2} \left\{ \sum_{i=1}^{n} [a+bd_i]\right)\left(\sum_{i=1}^{n} [a+bd_i]X_i^2\right)$$

$$- \left(\sum_{i=1}^{n} [a+bd_i]\, X_i \right)^2 \Big\}$$

$$= a^2 \left[(n-1) \sum_{i=1}^{n} X_i^2 - \sum_{i\neq j} X_iX_j \right]$$

$$+ ab\left[\sum_{k=1}^{n} d_k \sum_{i\neq k} X_i^2 + (n-1) \sum_{1}^{n} d_iX_i - \sum_{i\neq j} d_iX_iX_j \right.$$

$$\left. - \sum_{i\neq j} d_jX_iX_j \right] + b^2 \left[\sum_{i}^{n} d_k \sum_{j\neq k} d_iX_i^2 - \sum_{i\neq j} d_id_jX_iX_j \right]$$

$$= a^2 G + abH + b^2 I$$

$$|V(\hat{\beta}_{mt})| = (a+bd_{m+1}) \frac{(d_n - d_1)}{D(d_n-d_{m+1})-2E\sqrt{(d_n-d_{m+1})(d_{m+1}-d_1)}+F(d_{m+1}-d_1)}$$

$|V(\hat{\beta}_{ols})| = |V(\hat{\beta}_{mt})|$ for all $a,b \neq 0 \ni aI + bD$ is p.d.

D.1 $\quad d_{m+1}\, G = H$

D.2 $\quad d_{m+1}\, H = I$

D.3 $\quad d^2_{m+1}\, G = I$

Solve D.1 for d^o_n substitute into D.2 and D.3 if:

1. The equality of D.2 and D.3 hold then

   a $d^*_n \ni |V(\hat{\beta}_{ols})| \geq |V(\hat{\beta}_{mt})|\ \forall\ a,b \ni aI+bD$ is p.d.

2. Otherwise, there does not exist a $d^*_n \ni$

   $|V(\hat{\beta}_{ols})| \geq |V(\hat{\beta}_{mt})|$ for all a and b $\ni aI+bD$ is p.d.

APPENDIX C

LISTING OF PL/1 COMPUTER PROGRAM AND OUTPUT

The following computer program is designed to eliminate the unequal variances in Model I specified in this thesis. The transformed data will be punched on cards (program statement #231 and #243) which may be used as input to any ordinary least squares computer program. In order to find the estimates of the regression coefficients in a single run, the transformed data would be written on a temporary file named TSAI (program statement #232 and #244) which can be used as input to the next job step for an ordinary least squares computer program. Either option can be eliminated by withdrawing the corresponding cards. A sample JCL of using SAS in job step 2 is shown below:

```
//     JOB
//STEP1 EXEC PL1LFCLG
//PL1L.SYSIN DD *
     (Pl1 source program)
/*
//GO.PUNCH DD SYSOUT=B,DCB=BLKSIZE=80
//GO.TSAI DD DSN=&TA,UNIT=SYSDA,
//           SPACE=(CYL,(4,1)),DISP=(NEW,PASS),
//           DCB=(RECFM=FB,BLKSIZE=80,LRECL=80)
//GO.SYSIN DD *
     (input data cards)
/*
//STEP2 EXEC SAS3
//GO.TB DD DSN=*.STEP1.GO.TSAI,DISP=(OLD,DELETE),
//         DCB=(RECFM=FB,BLKSIZE=80)
     DATA;
```

```
        INPUT DDNAME=TB Y 1-8 X1 9-16 X2 17-24;
        . . .
/*
//
```

The input stream to the PL1 program should be ordered as follows:

Number of observations(n)

Number of regression coefficients(p)

Number of variance components(c)

Dependent variable observations: y(1), y(2), ..., y(n)

Independent variable observations(row major)

X(1,1), ..., X(1,p)

X(2,1), ..., X(2,p)

...

X(n,1), ..., X(n,p)

where X(1,i), ..., X(n,i) are the n observations of the i^{th} independent variable.

Variance-Covariance matrix(row major)

D(1,1), ..., D(1,c)

D(2,1), ..., D(2,c)

...

D(n,1), ..., D(n,c)

where D(1,i), ..., D(n,i) are the n diagonal elements of the i^{th} variance component.

The numerical values in the specified order can be punched in a stream on any number of cards with at least one blank between any two values.

A sample output of the PL1 program and SAS are attached at the end of the PL1 source program list.

```
1     (NOUNDERFLOW):
      MT: PROC OPTIONS(MAIN);
      /* CHECK DETERMINANT OF NEW XX BEFORE RUN REGRESSION */
2     DCL ((Y(N),X(N,NP),D(N,NC),T(N)) BIN FLOAT(53)) CTL;
3     DCL ((YY(MM),XX(MM,NP),NUD(MM,NC),LAMB(N,NC)) BIN FLOAT(53)) CTL;
4     DCL A(MM,N) BIN FLOAT(53) CTL;
5     DCL (N,M,M1,MM,NP,NC,IS,MOOD) BIN FLOAT(53);
6     DCL S BIN FLOAT(53);
7     DCL ((T1(9),T2(10)) BIN FLOAT) INIT(0);
8     IS=0;
9         ND=0;
10    IN: GET LIST(N,NP,NC);
11    ALLOCATE Y,X,D,T,LAMB;
12    GET LIST(Y,X,D); CALL OUT;
14    LUP: INDEX=1;
15    PUT PAGE;
16    PUT SKIP(5) EDIT('STAGE =',IS) (COLUMN(5),A(7),F(5));
17    PUT SKIP(5) EDIT('NUMBER OF OBSERVATIONS =',N) (COLUMN(5),A(24),F(10));
18          CALL SORTD;
19    PUT SKIP(5) EDIT('X MATRIX:') (COLUMN(5),A(9));
20    PUT SKIP(2) LIST(X);
21    PUT SKIP(5) EDIT('D MATRIX:') (COLUMN(5),A(9));
22    PUT SKIP(2) LIST(D);
23          IF INDEX=0 THEN GO TO LOP;
25          CALL TRANSF;
26    PUT SKIP(5) EDIT('LAMBAD MATRIX:') (COLUMN(5),A(14));
27    PUT SKIP(2) LIST(LAMB);
28    PUT SKIP(5) EDIT('A MATRIX:') (COLUMN(5),A(9));
29    PUT SKIP(2) LIST(A);
30          CALL NUDATA;
31          N=MM;   FREE A,Y,X,D,T; ALLOCATE Y,X,D,T;
34          DO I=1 TO MM;    Y(I)= YY(I); END;                      FREE YY;
38          DO I=1 TO MM; DO J=1 TO NP;  X(I,J)= XX(I,J); END; END; FREE XX;
44    DO I=1 TU MM; DO J=1 TO NC; D(I,J) = NUD(I,J); END; END; FREE NUD;
50          IF IS<NC THEN GOTU LOP;
52          PUT SKIP(10) EDIT('THE NEW DATA ON THE FOLLOWING PAGE(S)') (COLUMN
      (10),A(37));
53        ND=1;
54    CALL OUT;
55    SORTD: PROC;
56     N1=N-1; IS=IS+1;
58     DO I=1 TO N1;
59        I1=I+1;
60        DO K=I1 TO N;
61           IF D(I,IS) > D(K,IS) THEN DO;
63           DUM = D(I,IS);  D(I,IS)= D(K,IS);  D(K,IS) = DUM;
66           DO J1=1 TO NP;
67              DUM = X(I,J1); X(I,J1) = X(K,J1); X(K,J1) = DUM;
70           END;
71           DUM = Y(I);  Y(I)=Y(K); Y(K)=DUM;
74           END;
75        END;
76     END;
77     IF D(1,IS) = D(N,IS) THEN INDEX=0;
79    RETURN;
80    END SORTD;
81    TRANSF: PROC;
```

(NOUNDERFLOW):

```
82        MOOD = MOD(N,2);
83      IF MOOD=1 THEN DO:
85        M1=(N+1)/2; M=M1-1; MM=M1; ALLOCATE A;
89        DO I=1 TO M1; DO J=1 TO N; A(I,J)=0; END; END;
94        DO I=2 TO M1;
95           MI= M+I;
96           MJ=M-I+2;
97           LAMB(I,IS) = ( D(MI,IS) - D(M1,IS))/( D(MI,IS) - D(MJ,IS));
98        END;
99        DO I=1 TO M1; DO J=1 TO N;  A(I,J) = 0 ;  END; END;
104       A(1,M1) = 1;
105       DO I=2 TO M1; I1 = I-1;  MI1 = M1 - I1;  MI2= M1+I1;
109          A(I,MI1)= SQRT( LAMB(I,IS) );
110          A(I,MI2)= -SQRT( 1 -  LAMB(I,IS));
111       END;
112     END;
113     IF MOOD=0 THEN DO;
115       M = N/2;  MM=M; ALLOCATE A;
118       M1= M+1;  DO I=1 TO M; DO J=1 TO N; A(I,J)=0; END; END;
124       DM= ( D(M,IS)+ D(M1,IS))/2;
125       DO I=1 TO M; MI =M+I; M2 =M-I+1;
128          LAMB(I,IS) = ( D(MI,IS) - DM)/(D(MI,IS) - D(M2,IS));
129       END;
130       DO I=1 TO M:  M3= M+1-I; M4= M+I;
133          A(I,M3) = SQRT(LAMB(I,IS));
134          A(I,M4) =-SQRT(1 - LAMB(I,IS));
135       END;
136   RETURN;
137   END TRANSF;
139   NUDATA: PROC;
140         ALLOCATE YY,XX,NUD;        M1=MM+1;
142   AY:  DO I=1 TO MM;
143        S=0;
144      DO K=1 TO N;
145           S = S + A(I,K)*Y(K);
146        END;
147        YY(I) = S;
148        END;
149   PUT SKIP(5) EDIT('AY') (COLUMN(5),A(2));
150   PUT SKIP(2) LIST(YY);
151   AX: DO J=1 TO NP;
152      DO I=1 TO MM;
153           S=0;
154        DO K=1 TO N;
155        S=S+ A(I,K)*X(K,J);
156        END;
157        XX(I,J)=S;
158   END;
159   END;
160   PUT SKIP(5) EDIT('AX')  (COLUMN(5),A(2));
161   PUT SKIP(2);
162        DO I=1 TO MM;
163        PUT SKIP LIST(XX(I,1),XX(I,2));
164         END;
165   ADA: ID=IS+1;  IF ID<=NC THEN DO;
168   IF MOOD =1 THEN DO;
170      DO K=ID TO NC;  DO K1= 2 TO MM;   MK=M -K1+2; MKK= M+K1;
```

```
            NUD(K1,K) =  LAMB(K1,IS)*D(MK,K) + (1 - LAMB(K1,IS))*D(MKK,K);
          END; NUD(1,K) = D(MM,K);  END;
        END;
        IF MOOD=0 THEN DO;
          DO K=ID TO NC;  DO K1=1 TO MM;  MK= M+1-K1; MKK=M+K1;
            NUD(K1,K) =  LAMB(K1,IS)*D(MK,K) + (1 - LAMB(K1,IS))*D(MKK,K);
          END;            END;
        END;
        PUT SKIP(5) EDIT('ADA') (COLUMN(5),A(3));
        PUT SKIP(2);
            DO I=1 TO MM;  DO J=ID TO NC;
            PUT SKIP LIST(NUD(I,J));
        END; END;
        END;
        RETURN;
        END NUDATA;
        OUT: PROC;
             PUT PAGE;
            PUT SKIP EDIT('NUMBER OF OBSERVATIONS=',N) (COLUMN(5),A(23),F(10));
            PUT SKIP EDIT('NUMBER OF REG. COEFFS. =',NP)(COLUMN(5),A(24),F(9));
        PUT SKIP EDIT('NUMBER OF VAR-COMPONENTS =',NC)(COLUMN(5),A(26),F(7));
            PUT SKIP(5) EDIT('DEPENDENT VARIABLE:')(COLUMN(5),A(19));
            PUT SKIP(2) EDIT(Y)  (10F(12,3));
            DO J=1 TO NP;
            PUT SKIP(5) EDIT('INDEPENDENT VARIABLE',J)(COLUMN(5),A(20),F(5));
            DO I=1 TO N;
            T(I)=X(I,J);
            END;
            PUT SKIP(2) EDIT(T) (10F(12,3));
                END;
            IF ND=0 THEN DO;
            DO J=1 TO NC;
            PUT SKIP(5) EDIT('DIAGONAL ELEMENTS OF THE VAR-COV MATRIX: COMPONEN
        T',J)  (COLUMN(5),A(50),F(5));
            DO I=1 TO N;
            T(I)= D(I,J);
            END;
            PUT SKIP(2) EDIT(T) (10F(12,3));
            END;
            END;
        IF ND=1 THEN DO;
            DO I=1 TO N;
            K=NP;
            M=MIN(NP,9);
            DO J=1 TO M;
            T1(J) = X(I,J);
            END;
            PUT FILE(PUNCH) EDIT(Y(I),T1) (10F(8,3));
        PUT FILE(TSAI) EDIT(Y(I),T1) (10F(8,3));
            K1=0;
            K=K-9;
         LOP: IF K>0 THEN DO;
              M=MIN(K,9);
              L=10+K1*10;
             DO J=L TO M;
             M1 = L - (L-1);
             T2(M1) = X(I,J);
```

(NOUNDERFLOW):

```
             END;
             PUT FILE(PUNCH) EDIT(T2) (10F(8,3));
        PUT FILE(TSAI) EDIT(T2) (10F(8,3));
              K1=K1+1;
              END;
              K=K-10;
              IF K>0 THEN GO TO LOP;
           END;
        END;
        RETURN;
        END OUT;
        END MT;
```

INPUT BEFORE THE TRANSFORMATION

NUMBER OF OBSERVATIONS= 13
NUMBER OF REG. COEFFS. = 2
NUMBER OF VAR-COMPONENTS = 3

DEPENDENT VARIABLE:

1.000	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
11.000	12.000	13.000							

INDEPENDENT VARIABLE 1

1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000							

INDEPENDENT VARIABLE 2

1.000	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
11.000	12.000	13.000							

DIAGONAL ELEMENTS OF THE VAR-COV MATRIX: COMPONENT 1

1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000							

DIAGONAL ELEMENTS OF THE VAR-COV MATRIX: COMPONENT 2

1.000	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000
11.000	12.000	13.000							

DIAGONAL ELEMENTS OF THE VAR-COV MATRIX: COMPONENT 3

1.000	4.000	9.000	16.000	25.000	36.000	49.000	64.000	81.000	100.000
121.000	144.000	169.000							

OUTPUT AFTER THE TRANSFORMATION

```
NUMBER OF OBSERVATIONS=          4
NUMBER OF REG. COEFFS. =         2
NUMBER OF VAR-COMPONENTS =       3


DEPENDENT VARIABLE:

  -4.243       1.491       2.928      10.305


INDEPENDENT VARIABLE     1

   0.000       0.000       0.000       0.866


INDEPENDENT VARIABLE     2

  -4.243       1.491       2.928      10.305
```

S T A T I S T I C A L   A N A L Y S I S   S Y S T E M

ANALYSIS OF VARIANCE TABLE , REGRESSION COEFFICIENTS , AND STATISTICS OF FIT FOR DEPENDENT VARIABLE Y

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB > F	R-SQUARE	C.V.
REGRESSION	2	107.52949875	53.76474937	999999.99999	0.0001	1.00000000	0.00000 %
ERROR	1	0.00000000	0.00000000				
CORRECTED TOTAL	3	107.52949875				STD DEV	Y MEAN
						0.00000004	2.62025

SOURCE	DF	SEQUENTIAL SS	F VALUE	PROB > F	PARTIAL SS	F VALUE	PROB > F
X1	1	78.74051008	999999.99999	0.0006	0.00000000	0.00000	1.0000
X2	1	28.78898867	999999.99999	0.0006	28.78898867	999999.99999	0.0006

SOURCE	B VALUES	T FOR H0:B=0	PROB > |T|	STD ERR B	STD B VALUES
INTERCEPT	0.00000000	0.00000	1.0000	0.00000002	0.0
X1	0.00000000	0.00000	1.0000	0.00000011	0.00000000
X2	1.00000000	999999.99999	0.0001	0.00000001	1.00000000

VITA

Huei Chu Tsai

Candidate for the Degree of

Doctor of Philosophy

Thesis: ESTIMATION OF A LINEAR REGRESSION MODEL WHEN THE VARIANCE IS A LINEAR FUNCTION OF UNKNOWN PARAMETERS

Major Field: Statistics

Biographical:

Personal Data: Born in Kwantung, China, January 23, 1940, the second son of Cheng-tung and Yet-ho Tsai.

Education: Attended primary school in Taiwan, China; graduate from Taichung High School, Taiwan, China, in 1957; attended Tunghai University, Taichung, Taiwan, China, 1958-1962; received the Bachelor of Arts in Economics from Tunghai University, Taiwan, China, in 1962; attended College of Chinese Culture, Taipei, Taiwan, China, 1964-1967; received Master of Arts in Statistics from University of Missouri at Columbia, Missouri, in 1970; attended Washington University, St. Louis, Missouri, 1971-1973; requirements for the Degree of Doctor of Philosophy completed at Oklahoma State University, Stillwater, Oklahoma, July, 1976.

Professional Experience: Graduate Teaching Assistant in the Department of Statistics, University of Missouri at Columbia, August, 1968 to May, 1970; Graduate Research Assistant in the Department of Statistics, University of Missouri at Columbia, June, 1970 to April, 1971; Computer Programmer, Research Department, Federal Reserve Bank of St. Louis, May, 1971 to July, 1973; Graduate Teaching Assistant in the Department of Statistics, Oklahoma State University, Stillwater, Oklahoma, 1973-1976.