MULTICOLLINEARITY AND THE MEAN SQUARE ERROR CRITERION IN MULTIPLE REGRESSION: A TEST AND SOME SEQUENTIAL ESTIMATOR COMPARISONS

by

Carlos Enrique Toro-Vizcarrondo

Institute of Statistics
Mimeograph Series No. 588
August 1968
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>List Item</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES.</td>
<td>vii</td>
</tr>
<tr>
<td>INTRODUCTION.</td>
<td>1</td>
</tr>
<tr>
<td>The Model</td>
<td>1</td>
</tr>
<tr>
<td>Stepwise Regression</td>
<td>2</td>
</tr>
<tr>
<td>Specification Error</td>
<td>3</td>
</tr>
<tr>
<td>Multicollinearity</td>
<td>4</td>
</tr>
<tr>
<td>Preliminary Test of Significance and the Mean Square Error.</td>
<td>5</td>
</tr>
<tr>
<td>TWO ESTIMATORS AND A CRITERION.</td>
<td>6</td>
</tr>
<tr>
<td>Purpose</td>
<td>6</td>
</tr>
<tr>
<td>Multicollinearity and the Mean Square Error Criterion--Two Regressors.</td>
<td>6</td>
</tr>
<tr>
<td>The &quot;t-ratio&quot; Test and the Mean Square Error Criterion--Two Regressors.</td>
<td>8</td>
</tr>
<tr>
<td>The Restricted and Full Least Squares Estimators.</td>
<td>9</td>
</tr>
<tr>
<td>VARIANCE (MEAN SQUARE) OF A LINEAR FUNCTION, GENERALIZED VARIANCE (MEAN SQUARE), AND THE MEAN SQUARE CRITERION</td>
<td>14</td>
</tr>
<tr>
<td>Variance of a Linear Function and the Generalized Variance.</td>
<td>14</td>
</tr>
<tr>
<td>Mean Square of a Linear Function and the Generalized Mean Square.</td>
<td>15</td>
</tr>
<tr>
<td>THE NONCENTRAL F DISTRIBUTION AND THE MEAN SQUARE ERROR CRITERION.</td>
<td>22</td>
</tr>
<tr>
<td>Preliminaries</td>
<td>22</td>
</tr>
<tr>
<td>A Uniformly More Powerful Test for Multicollinearity.</td>
<td>22</td>
</tr>
<tr>
<td>Comparison of the Usual F Test and the Mean Square Error Criterion Test</td>
<td>30</td>
</tr>
<tr>
<td>Minimum Variance Unbiased Estimator of ( \lambda ).</td>
<td>32</td>
</tr>
<tr>
<td>CRITICAL POINTS AND VALUES OF THE POWER OF THE &quot;MULTICOLLINEARITY&quot; TEST FOR ONE DEGREE OF FREEDOM IN THE NUMERATOR</td>
<td>34</td>
</tr>
<tr>
<td>Critical Points</td>
<td>34</td>
</tr>
<tr>
<td>Power Values.</td>
<td>38</td>
</tr>
<tr>
<td>Some Comments about the Critical Points and Power Tables.</td>
<td>45</td>
</tr>
<tr>
<td>ESTIMATION AFTER PRELIMINARY TESTS OF SIGNIFICANCE.</td>
<td>49</td>
</tr>
<tr>
<td>Sequential Estimation</td>
<td>49</td>
</tr>
<tr>
<td>Sequential Estimator Bias</td>
<td>51</td>
</tr>
<tr>
<td>Sequential Estimator Mean Square.</td>
<td>61</td>
</tr>
<tr>
<td>TABLE OF CONTENTS (continued)</td>
<td>Page</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH</td>
<td>68</td>
</tr>
<tr>
<td>Summary</td>
<td>68</td>
</tr>
<tr>
<td>Conclusions</td>
<td>70</td>
</tr>
<tr>
<td>Suggestions for Further Research</td>
<td>72</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>74</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>76</td>
</tr>
<tr>
<td>APPENDIX A. Computer Program for the &quot;Multicollinearity&quot; Test Critical Points</td>
<td>77</td>
</tr>
<tr>
<td>APPENDIX B. Computer Program for the Evaluation of Power Values</td>
<td>79</td>
</tr>
<tr>
<td>APPENDIX C. Fortran Statement for the Evaluation of Sequential Estimators Bias and Mean Squares</td>
<td>80</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Noncentral beta and noncentral F &quot;multicollinearity&quot; test of critical points for one degree of freedom in the numerator.</td>
<td>37</td>
</tr>
<tr>
<td>2</td>
<td>Largest values of i used in computing the power for $\lambda \neq 0$.</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>Power values for the &quot;multicollinearity&quot; test at 5 percent significance level</td>
<td>41</td>
</tr>
<tr>
<td>4</td>
<td>Power values for the &quot;multicollinearity&quot; test at 10 percent significance level</td>
<td>42</td>
</tr>
<tr>
<td>5</td>
<td>Power values for the &quot;multicollinearity&quot; test at 25 percent significance level</td>
<td>43</td>
</tr>
<tr>
<td>6</td>
<td>Power values for the &quot;multicollinearity&quot; test at 50 percent significance level</td>
<td>44</td>
</tr>
<tr>
<td>7</td>
<td>Comparison of Tang's 1 percent and &quot;multicollinearity&quot; 5 percent critical points.</td>
<td>46</td>
</tr>
<tr>
<td>8</td>
<td>Comparison of critical points for selected denominator degrees of freedom when $\beta(0)$ is approximately .01.</td>
<td>47</td>
</tr>
<tr>
<td>9</td>
<td>Comparison of F critical points for selected denominator degrees of freedom when $\beta(0)$ is approximately .025</td>
<td>47</td>
</tr>
<tr>
<td>10</td>
<td>Comparison of F critical points for selected denominator degrees of freedom when $\beta(0)$ is approximately .10.</td>
<td>48</td>
</tr>
<tr>
<td>11</td>
<td>Critical points used in the evaluation of the sequential estimator's bias</td>
<td>57</td>
</tr>
<tr>
<td>12</td>
<td>Powers for the &quot;50 percent&quot; test for selected values of the noncentrality parameter for one degree of freedom in the numerator.</td>
<td>57</td>
</tr>
<tr>
<td>13</td>
<td>Values of $\lambda$ and range i (in parenthesis) used in computing the bias and mean square error of the sequential estimator of $\beta_1$.</td>
<td>59</td>
</tr>
<tr>
<td>14</td>
<td>The bias in estimating $\beta_1$ via the &quot;mean square,&quot; &quot;usual,&quot; and &quot;50 percent&quot; sequential estimators.</td>
<td>60</td>
</tr>
<tr>
<td>15</td>
<td>The mean squares of the &quot;mean square,&quot; &quot;usual,&quot; and &quot;50 percent&quot; sequential estimators.</td>
<td>67</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

The Model

The motivation of this thesis can be found in the following interrelated topics: stepwise regression, specification error, multicollinearity, minimum mean square estimator and estimators after preliminary tests of significance. Consider the linear model of full rank

\[(1) \quad Y = X\beta + \epsilon\]

where \(Y\) is an nx1 column vector of random variables whose variation is to be explained by the \(p\) nonstochastic regression vectors appearing as columns in the \(n\times p\) matrix \(X\).

The \(X\) matrix and \(\beta\) vector may be partitioned conformably so that (1) can be written as

\[(2) \quad Y = X_1\beta_1 + X_2\beta_2 + \epsilon\]

where \(X_1\) is the \(n\times k\) matrix of observations on the first \(k\) regressors: \(x_i \quad (i=1, 2, \ldots, k)\) with rank \(k\), \(\beta_1\) is the \(k\times 1\) vector of coefficients of the first \(k\) regressors, \(X_2\) is the \(n\times m\) matrix of observations of the second \(m\) regressors: \(x_j \quad (j=k+1, k+2, \ldots, k+m = p<n)\) with rank \(m\), \(\beta_2\) is the \(m\times 1\) vector of coefficients of the second \(m\) regressors, and \(\epsilon\) is the \(n\times 1\) vector of disturbances distributed as a multivariate normal distribution with mean vector \(0\) and covariance matrix \(\sigma^2 I\). For convenience both the \(X\)'s and \(Y\) variables are assumed to have zero sample means.
Stepwise Regression

The stepwise least squares procedure first estimates $\beta_1$ as

$$\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'Y$$

and then regresses the residuals

$$\hat{\tilde{Y}} = Y - X_1\hat{\beta}_1 = [I - X_1(X_1'X_1)^{-1}X_1']Y$$

upon $X_2$ to estimate $\beta_2$ as

$$\hat{\beta}_2 = (X_2'X_2)^{-1}X_2'\hat{\tilde{Y}} = (X_2'X_2)^{-1}X_2'[I - X_1(X_1'X_1)^{-1}X_1']Y.$$  

Freund et al. (1961) examined the bias, estimates and associated tests obtained by subjecting the residuals to a regression analysis. Goldberger and Jockems (1961) showed that the stepwise procedure underestimates both the regression coefficient of $x_2$ and the marginal contribution of $x_2$ to the explanation of $y$ in the simple model of two regressors. The results were generalized for the estimate of $\beta_{k+1}$ in the case of $k+1$ explanatory variables. Goldberger (1961) presented the connection between the stepwise and the direct (classical estimator of $\beta_1$ including $X_2$) estimator of $\beta_1$ and concluded that the stepwise estimator ($\hat{\beta}_1$) can be written as a function of the direct estimators ($\hat{\beta}_1, \hat{\beta}_2$), i.e.,

$$\hat{\beta}_1 = \hat{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\hat{\beta}_2.$$  

From (5) it is easily seen that $\hat{\beta}_1$ is biased by the factor $(X_1'X_1)^{-1}X_1'X_2\hat{\beta}_2$ unless $X_1$ and $X_2$ are orthogonal or $\hat{\beta}_2 = 0$. Zyskind (1963) established exact expressions for the difference between "true" additional regression sum of squares due to $X_2$ obtained by direct regression analysis and the
regression sum of squares due to additional variables by residual analysis. He also showed that a residual analysis test statistic considered by Freund et al. (1961) is subject to certain serious defects. Kabe (1963) gave the results obtained by Freund et al. (1961), Goldberger and Jockems (1961) and Goldberger (1961) for a multivariate regression model. Wallace (1964) showed that for the two-regressors case the stepwise estimator has a smaller variance than the direct estimators and that using the criterion of relative mean square errors the stepwise estimators are "better" than direct least squares estimators under certain circumstances. This thesis generalizes the results obtained by Wallace (1964). First, it is shown that when the mean square criterion is used, the selection between a stepwise estimator and a direct estimator as an estimate of $\beta_1$ is determined by a critical value of the noncentrality parameter of the noncentral F distribution. Second, a uniformly most powerful test is derived as a method for determining when to use the stepwise estimator. Third, a minimum variance unbiased estimator is obtained for the noncentrality parameter. Fourth, for various degrees of freedom and various significance levels, the noncentral F is tabulated for the critical value of the noncentrality parameter. Fifth, the power of the test is tabulated over selected ranges of the relevant parameters.

**Specification Error**

The term, specification error, is used for the situation where the true hypothesis is (1) but we mistakenly work with

\[(6) \quad Y = X_1 \beta_1 + \nu.\]

That is, we leave out variables which should be included. The stepwise
estimator $\hat{\beta}_1$ may be viewed as the direct estimator of $\beta_1$ in (6) taking $\mathbf{V} = X_2 \beta_2 + \mathbf{\epsilon}$ as the "disturbance" vector. The stepwise estimator bias is attributable to the fact that there is a "specification error;" the classical model is not appropriate to (6) because the "disturbance" $\mathbf{V}$ does not have zero expectation. Griliches (1957) analyzed some common specification errors in Cobb-Douglas production function estimation; Theil (1961) presented some results on the problems of specification analysis, and Goldberger (1961) spelled out the connection between the stepwise and the direct estimator of $\beta_1$ as a special case of Theil's analysis of specification error. The problem of estimating $\beta_1$ by the stepwise estimator is analogous to misspecifying the model as (6) when the true model is (2). If our interest is to find the "best" estimate of $\beta_1$ between the stepwise and the direct estimator, the mean square error criterion may, under certain circumstances, lead us to specify the model as (6) and estimate $\hat{\beta}_1$ as $(X_1'X_1)^{-1} X_1' \mathbf{V}$.

**Multicollinearity**

When linear relations exist among the observed values of the regressors we say multicollinearity exists. In Johnston (1963) and Goldberger (1964) multicollinearity is discussed in the limiting case where the correlation of the regressors in a two-regressors model is one. Also an intuitive explanation is given when the correlation is close to one. Wallace (1964) suggested that multicollinearity becomes a precise concept by allowing the mean square error criterion to tell us whether we are "better off" in estimating $\beta_1$ by ignoring $x_2$. In the more general model (2), the multicollinearity problem is whether to include $X_2$ or not in the estimation of $\beta_1$. Again the problem is analogous to decide whether $\hat{\beta}_1$ or $\hat{\beta}_1$ is to be used.
Preliminary Test of Significance and the Mean Square Error

Stepwise regression, specification error and multicollinearity are branches of one tree whose trunk is the mean square error concept. The purpose of this thesis is to give the researcher a practical procedure to decide when to use the stepwise estimator, or analogously when to misspecify the model, or delete variables from the model because they are collinear. Usually researchers make the selection between $\hat{b}_1$ and $\tilde{b}_1$ using a preliminary test of significance. The most common is testing the null hypothesis $\beta_2 = 0$; when the hypothesis is accepted $\tilde{b}_1$ is used, when rejected $\hat{b}_1$ is used. The bias involved in the estimation of $\beta_1$ after a preliminary test in a two-regressors model was found by Bancroft (1944). We propose an alternative preliminary test of significance (mean square error test). A comparison will be given between the usual estimator and the mean square estimator in terms of bias and mean square in a two-regressors model.
CHAPTER II
TWO ESTIMATORS AND A CRITERION

Purpose

The purpose of this thesis is to find a "good" estimate of \( \beta_1 \) in the partitioned model (2). Two estimators are to be studied: the restricted least squares estimators (R.L.S.E.) of \( \beta_1 \) (generally known as stepwise estimators) and the full least squares estimators (F.L.S.E.) of \( \beta_1 \) (sometimes called direct least squares estimators). The R.L.S.E. ignores \( X_2 \) in the estimation of \( \beta_1 \) while the F.L.S.E. estimates \( \beta_1 \) including \( X_2 \) in the model. In the first section of this chapter the mean square error criterion is used in order to select between the R.L.S.E. and the F.L.S.E. when the model is simplified to a two-independent variables case. An attempt to explain multicollinearity and comments about the usual procedure of deleting variables by using the t-test are presented. The restricted and full estimators are discussed in the subsequent sections.

Multicollinearity and the Mean Square Error Criterion--Two Regressors

Suppose the "true" model is

\[
y_i = x_{1i}\beta_1 + x_{2i}\beta_2 + \varepsilon_i, \quad i=1, \ldots, n.
\]

(7)

The restricted estimator of \( \beta_1 \) in this simple model is \( \frac{\sum x_1y}{\sum x_1^2} \)

with mean \( \beta_1 + \beta_2 \frac{\sum x_1x_2}{\sum x_1^2} \) and variance \( \sigma^2 \left( \frac{\sum x_1^2}{\sum x_1} \right) \) and the full estimator is

\[
\frac{\sum x_1y}{\sum x_1^2} + \frac{r_{12}^2}{1-r_{12}^2} \left( \frac{\sum x_1y}{\sum x_1^2} - \frac{\sum x_2y}{\sum x_1x_2} \right)
\]

with mean \( \beta_1 \) and variance \( \frac{\sigma^2}{\sum x_1^2 (1-r_{12}^2)} \) where \( r_{12} \) is the "simple correlation" between \( x_1 \) and \( x_2 \);
\[ r_{12}^2 = \frac{(\Sigma x_1 x_2)^2}{\Sigma x_1^2 \Sigma x_2^2}. \]

The variance of \( \hat{b}_1 \) is less than or equal to variance of \( \hat{b}_1 \) but since \( \hat{b}_1 \) is biased, a more meaningful comparison can be made using mean square errors (M.S.E.). By making the comparison, Wallace (1964) showed that

\[ \text{Mean Square Error (} \hat{b}_1 \text{)} \leq \text{Mean Square Error (} \hat{b}_1 \text{)} \]

\[ (8) \quad \frac{\beta_2^2}{V(\hat{b}_1)} = \frac{\beta_2^2}{\sigma^2 / \Sigma x_2^2 (1 - r_{12}^2)} = \frac{\beta_2^2}{\sigma^2 / \Sigma x_2^2} - \frac{(\Sigma x_1 x_2)^2}{\Sigma x_1^2} < 1. \]

The mean square criterion tells the investigator when he is "better off" in estimating \( \beta_1 \) by ignoring \( x_2 \) but even in the simple case of only two regressors, the criterion (8) is difficult to implement since it is unlikely to be known beforehand, i.e., it involves two parameters. The criterion also depends on the correlation between \( x_1 \) and \( x_2 \). When \( x_1 \) and \( x_2 \) are linearly dependent, "perfect multicollinearity"--the simple correlation coefficient is one. As \( r_{12}^2 \) gets very close to one, the variance of the F.L.S.E. for \( \beta_1 \) explodes. Any estimator with finite mean square will be better than \( \hat{b}_1 \) as the correlation between \( x_1 \) and \( x_2 \) gets very close to one. If the two regressors are orthogonal--"zero multicollinearity"--the simple correlation coefficient is zero and the two estimators under consideration are equivalent (Goldberger, 1964, pp. 200-201).

In some areas of research, say econometrics, the above situations rarely, if ever, exist.\(^1\) Econometricians have to accept the experiments

---

\(^1\) In controlled laboratory experiments, orthogonality of the sample observations on x's can be achieved.
performed by nature, and these experiments are likely to feature collinearity rather than orthogonality. In practice an exact linear relationship is highly improbable, but the general interdependence of economic phenomena may result in the appearance of approximate linear relationships in the x's. This phenomenon is known as multicollinearity or intercorrelation (Goldberger, 1964; Johnston, 1963). Excluding the limiting cases, use of the mean square error criterion makes multicollinearity a precise concept.

\[
\frac{\beta_2^2}{\text{V}(b_2)} \leq 1,
\]

If we should, according to the M.S.E. criterion, use the restricted estimator \(b_1\). Rearranging the inequality, we have

\[
r_{12}^2 > 1 - \frac{\sigma^2}{\beta_2^2 \Sigma x_2^2}.
\]

Thus multicollinearity is a problem when \(r_{12}^2\) is greater than or equal to a quantity that depends on the parameters \(\beta_2\) and \(\sigma^2\) (\(\Sigma x_2^2\) is a known constant).

Since the square of the correlation coefficient is bounded by zero and one, then we should throw away \(x_2\) whenever

\[
1 - \frac{\sigma^2}{\beta_2^2 \Sigma x_2^2} \leq r_{12}^2 \leq 1.
\]

(9)

So a precise statement of "multicollinearity" is obtained but it is a rather inconclusive one—it requires knowledge about two unknown parameters, i.e., \(\beta_2\) and \(\sigma^2\).

The "t-ratio" Test and the Mean Square Error Criterion—Two Regressors

Researchers, generally, will approach the problem of including or excluding the variable \(x_2\) by using a "t-ratio." The question the
researcher asks, formally described, is the verbal hypothesis that "y does not vary with \(x_2\)" which in turn is the hypothesis \(\beta_2 = 0\). The so-called "t-ratio" \(\frac{\hat{b}_2}{\hat{S}_{b_2}}\) is the test statistic appropriate to this hypothesis \((\hat{S}_{b_2} is the sample standard deviation of \(\hat{b}_2\)). The use of the t-ratio implicitly assumes that the researcher is interested in the prediction of \(y\), not in the estimation of \(\beta_1\). If the hypothesis \(\beta_2 = 0\) is accepted, the dependent variable \(y\) is estimated by \(\hat{y} = \hat{b}_1 x_1\), and when \(\beta_2 = 0\) is rejected, \(y\) is estimated by \(\hat{y} = \hat{b}_1 x_1 + \hat{b}_2 x_2\). However, if the researcher is interested in finding the minimum mean square estimator between \(\hat{b}_1\) and \(\hat{b}_2\), the following mean square rule should be followed:

Use \(\hat{b}_1\) if \(-\sqrt{\hat{V}(\hat{b}_2)} \leq \beta_2 \leq \sqrt{\hat{V}(\hat{b}_2)}\), use \(\hat{b}_2\) otherwise.

Therefore, \(\beta_2 = 0\) is included in the mean square rule but \(x_2\) might be deleted for other values of \(\beta_2\) as long as \(|\beta_2| \leq \sqrt{\hat{V}(\hat{b}_2)}\).

The Restricted and Full Least Squares Estimators

In the previous sections we considered the mean square error criterion in the simple case of two regressors. In the next chapter the mean square error criterion is extended to the estimation of \(\beta_1\) in the general partitioned model described in (2). This section presents properties of the restricted and full estimators that are needed in the generalization of the mean square error criterion.

If \(X_2\) is ignored, \(\beta_1\) is estimated from the estimation model

\[
\hat{y} = X_1\hat{\beta}_1 + \hat{\nu}
\]

where
\[ V = X_2 \hat{\beta}_2 + \varepsilon. \]

The restricted estimator for \( \hat{\beta}_1 \) is obtained by applying the usual least squares procedure to equation (10). This yields the estimator

\[ \hat{\beta}_1 = (X_1'X_1)^{-1} X_1'Y. \]  

In (5) we saw that the restricted estimator of \( \hat{\beta}_1 \) may be a biased estimator. Equation (12) is obtained substituting (2) into (11).

\[ \hat{\beta}_1 = \hat{\beta}_1 + A\hat{\beta}_2 + (X_1'X_1)^{-1} X_1' \varepsilon \]

where

\[ A = (X_1'X_1)^{-1} X_1'X_2. \]

The columns of \( A \) can be taken to be coefficients from auxiliary regressions of each variable in \( X_2 \) upon the set of variables contained in \( X_1 \). It is evident from (12) that \( \hat{\beta}_1 \) is a biased estimator with corresponding bias equal to \( A\hat{\beta}_2 \) unless \( X_1 \) and \( X_2 \) are orthogonal, i.e., \( X_1'X_2 = 0 \) (kxm) or \( \hat{\beta}_2 = 0 \) (Freund et al., 1961; Goldberger and Jockems, 1961; and Goldberger, 1961).

Now let us turn to examining the second order moments of the estimators and denote the covariance matrix of \( \hat{\beta}_1 \) as \( \sigma^2_{\hat{\beta}_1} \) and its mean square error matrix as \( \Omega^2_{\hat{\beta}_1}. \)

\[ ^2 \text{A mean square error matrix for a vector of estimators, say } \hat{\Theta}, \text{ is defined as } E(\hat{\Theta} - \Theta)(\hat{\Theta} - \Theta)', \text{ covariance matrix } + \text{ (bias vector)(bias vector)}, \]

where \( E \) stands for expected value. If \( \Theta \) is unbiased, the mean square error matrix is equivalent to the covariance matrix.
The covariance matrix of $\hat{b}_1$ is:

$$
\Sigma_{\hat{b}_1} = E[\hat{b}_1 - E(\hat{b}_1)][\hat{b}_1 - E(\hat{b}_1)]' = E(X_1'X_1)^{-1}X_1'eX_1(X_1'X_1)^{-1}
$$

$$
= \sigma^2(X_1'X_1)^{-1}.
$$

And the mean square error matrix of $\hat{b}_1$ is:

$$
\Omega_{\hat{b}_1} = E[\hat{b}_1 - \hat{b}][\hat{b}_1 - \hat{b}]' = \Sigma_{\hat{b}_1} + [\text{bias}(\hat{b}_1)][\text{bias}(\hat{b}_1)]' = \sigma^2(X_1'X_1)^{-1} + A\sigma_2^2\xi^2A'.
$$

As is well-known, the full least squares (F.L.S.E.) estimator for $\beta$ is $\hat{b} = (X'X)^{-1}X'Y$ or, in the partitioned form that is used in this thesis,

$$
\hat{b} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1'Y \\ X_2'Y \end{bmatrix}.
$$

The above can be written as:

$$
\hat{b} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} = \frac{(X_1'X_1)^{-1}[I + X_1'X_2F_2^{-1}(X_2'X_1)(X_1'X_1)^{-1}] - (X_1'X_1)^{-1}X_1'X_2F_2^{-1}}{1 - F_2^{-1}X_2'X_1(X_1'X_1)^{-1}} \cdot \begin{bmatrix} X_1'Y \\ X_2'Y \end{bmatrix}
$$

where $F_2 = X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2$, employing a partitioned inversion procedure (Goldberger, 1964, p. 174).

The properties of $\begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix}$ are summed up in the following well-known theorem (Graybill, 1961, p. 113).
Theorem: If \( Y = X \beta + \varepsilon \) is a general linear hypothesis model of full rank and if \( \varepsilon \) is distributed \( N(0, \sigma^2 I) \), the estimators

\[
\hat{\beta} = (X'X)^{-1}X'Y, \quad \hat{\sigma}^2 = \frac{Y'[I-X(X'X)^{-1}X']Y}{n-p}
\]

have the following properties:

1. Consistent
2. Efficient
3. Unbiased
4. Sufficient
5. \( \hat{\beta} \) is distributed \( N(\beta, \sigma^2(X'X)^{-1}) \)
6. Complete
7. Minimum Variance Unbiased
8. \( \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \) is distributed as \( \chi^2(n-p) \)
9. \( \hat{\beta} \) and \( \hat{\sigma}^2 \) are independent

The F.L.S.E. of \( \beta_1 \) reduces to

\[
\hat{\beta}_1 = (X'X_1)^{-1}X_1'Y - (X'X_1)^{-1}X_1'X_2F_2^{-1}X_2'[I - X_1(X_1'X_1)^{-1}X_1']Y
\]

which in view of the definition of the R.L.S.E. \( \hat{\beta}_1' \) in (11) and the definition of the F.L.S.E. \( \hat{\beta}_2 \) in (16) gives equation (5), i.e.,

\[
\hat{\beta}_1 = \hat{\beta}_1' - A\hat{\beta}_2
\]

where \( A \) has previously been defined. From (17) we can see that \( \hat{\beta}_1' = \hat{\beta}_1 \) when \( X_1 \) and \( X_2 \) are orthogonal.

The F.L.S.E. are unbiased, therefore the covariance and generalized mean squares matrices of \( \hat{\beta}_1 \) are identical.
Using similar notation for $\hat{b}_1$ as for $\hat{b}_1$, the following result is obtained directly from (16).

\begin{equation}
\hat{\Sigma}_{\hat{b}_1} = \Omega_{\hat{b}_1} = \sigma^2 (X_1'X_1)^{-1} [I + X_1'X_2F_2^{-1}X_2'X_1(X_1'X_1)^{-1}] \\
= \sigma^2 (X_1'X_1)^{-1} + \sigma^2 A F_2^{-1} A'.
\end{equation}
CHAPTER III
VARIANCE (MEAN SQUARE) OF A LINEAR FUNCTION,
GENERALIZED VARIANCE (MEAN SQUARE), AND THE
MEAN SQUARE CRITERION

Variance of a Linear Function and the Generalized Variance

In the simple model of two regressors the variance of the restricted estimator is less than or equal to the variance of the full estimator. In looking at the general partitioned model (2), the following criterion is used to compare the restricted estimator, \( \hat{b}'_1 = (\hat{b}_1', \hat{b}_2', \ldots, \hat{b}_k') \), with the full estimator, \( b'_1 = (b_1', b_2', \ldots, b_k') \). The restricted estimator is defined to be preferable from the viewpoint of efficiency if the variance of any arbitrary linear function of the full estimator is greater than or equal to the variance of that same arbitrary linear function of the restricted estimator. Denoting the variance by \( \text{Var} \) the criterion can be written as: the restricted estimator is preferable if,

\[
\text{Var}(c' \hat{b}_1) - \text{Var}(c' b'_1) \geq 0
\]

for all \( c \) except \( c = 0 \), where \( c \) is a kxl vector. From (19) the following inequality is obtained:

\[
c' [\Sigma_{\hat{b}_1} - \Sigma_{b'_1}] c \geq 0
\]

for all \( c \) (kxl), i.e., \( \Sigma_{\hat{b}_1} - \Sigma_{b'_1} \) is positive semidefinite. So the variance of an arbitrary linear function criterion asks whether the

\[
\text{Var} (c' \hat{\theta}) = E[c' \hat{\theta} - E(c' \hat{\theta})][c' \hat{\theta} - E(c' \hat{\theta})']
\]

\[
= c' E[\hat{\theta} - E\hat{\theta}][\hat{\theta} - E\hat{\theta}]' c
\]

\[
= c' \Sigma_{\hat{\theta}} c
\]
covariance matrix of the full estimator minus the covariance matrix of
the restricted estimator is positive semidefinite. Using (13) and (18)
the difference of the covariance matrices can be written as:

\[(21) \quad \Sigma_{\widehat{\beta}_1} - \Sigma_{\widehat{\gamma}_1} = \sigma^2 A F_2^{-1} A'\]

where A and F_2 are previously defined. Since F_2^{-1} is positive definite,
(\sigma^2 F_2^{-1} is the covariance matrix of the full estimator \widehat{\beta}_2) A F_2^{-1}A' is
positive semidefinite (Goldberger, 1964, p. 37). Therefore the restricted
estimator is preferable in terms of variance of a linear function.

Furthermore, the variance of a linear function criterion implies the widely-
used method of comparing two sets of estimators known as the Generalized
Variance Ratio. Since \Sigma_{\widehat{\beta}_1}, \Sigma_{\widehat{\gamma}_1} are positive definite matrix and
\Sigma_{\widehat{\beta}_1} - \Sigma_{\widehat{\gamma}_1} is positive semidefinite, we can conclude by using a well-known
theorem (Graybill, 1961, p. 6) that \(|\Sigma_{\widehat{\beta}_1}| \geq |\Sigma_{\widehat{\gamma}_1}|\). Further, the ratio
\(|\Sigma_{\widehat{\beta}_1}| / |\Sigma_{\widehat{\gamma}_1}|\) is bounded by zero and one, hence the ratio provides a
convenient percentage comparison of efficiency.

**Mean Square of a Linear Function and the Generalized Mean Square**

A comparison of the restricted estimator and the full estimator based
both on bias and variance must be then based on the personal preferences
of the researcher. If bias is considered to be extremely important, the

---

4 If \(\widehat{\Theta}\) is a random vector with covariance matrix \(\Sigma\), then the
determinant of \(\Sigma(|\Theta|)\) is called the generalized variance of \(\widehat{\Theta}\). The
estimator \(\widehat{\Theta}\) is preferable over \(\widehat{\Theta}\) if \(|\Sigma_{\widehat{\Theta}}| \geq |\Sigma_{\widehat{\Theta}}|\). A geometric interpretation
of the generalized variance is given in T. W. Anderson (1964).
full estimator would be used, while if the variance criterion is given great weight the restricted estimator would be used. One criterion, encompassing both bias and variance, is the mean square error criterion. Since we are dealing with vectors of estimators the criterion is reduced to a scalar by looking at the mean square of an arbitrary linear function of the estimators.

The criterion is the following:

\[(22) \quad \text{Whenever M.S.E. } (\mathbf{c}^\prime \mathbf{\hat{b}}_1) \geq \text{M.S.E. } (\mathbf{c}^\prime \mathbf{\hat{b}}_1)\]

for all \(\mathbf{c}\) (kxl) except \(\mathbf{c} = 0\), the restricted estimator would be preferable.

The above inequality can be written as

\[(23) \quad \mathbf{c}^\prime [\mathbf{\Omega}_1^\wedge - \mathbf{\Omega}_1^\nu] \mathbf{c} \geq 0\]

for all \(\mathbf{c}\) which implies that the mean square matrix of the full estimator minus the mean square matrix of the restricted estimator must be positive semidefinite. Using (14) and (18) this difference is written as

\[(24) \quad \sigma^2 \mathbf{A} \mathbf{F}_2^{-1} \mathbf{A}^\prime - \mathbf{A} \mathbf{B}_2 \mathbf{B}_2^{-1} \mathbf{A}^\prime .\]

Since \(\mathbf{\Omega}_1^\wedge\) and \(\mathbf{\Omega}_1^\nu\) are symmetric, and \(\mathbf{\Omega}_1^\nu\) is positive definite\(^5\) and if

\(\mathbf{\Omega}_1^\wedge - \mathbf{\Omega}_1^\nu\) is positive semidefinite, then we can conclude by a well-known theorem (Graybill, 1961, p. 6) that \(|\mathbf{\Omega}_1^\wedge| \geq |\mathbf{\Omega}_1^\nu|\). So the mean square

\(^5\)The mean square matrix of the restricted estimator is \(\sigma^2 (\mathbf{X}_1^\prime \mathbf{X}_1)^{-1} + \mathbf{A} \mathbf{B}_2 \mathbf{B}_2^{-1} \mathbf{A}^\prime\) where the first term is positive definite (Graybill, 1961, p. 4) and the second term is positive semidefinite (Graybill, 1961, p. 4; Goldberger, 1964, p. 37).
error of a linear function criterion implies what can be called the Generalized Mean Square Ratio Criterion.

Using the more general M.S.E. criterion, whether to include $X_2$ to estimate $\beta_1$ depends on the properties of $A F_2^{-1} A' - \frac{1}{\sigma^2} A \beta_2 \beta_2' A'$ since the positive semidefinite property is not changed when (24) is multiplied by the positive scalar $\frac{1}{\sigma^2}$. When it is positive semidefinite we will estimate $\beta_1$ using only $X_1$ by means of the restricted estimator. We shall present two cases for which the difference in the mean square error matrices is positive semidefinite.

I. A sufficient condition for $A (F_2^{-1} - \frac{1}{\sigma^2} \beta_2 \beta_2') A'$ to be positive semidefinite is that $F_2^{-1} - \frac{1}{\sigma^2} \beta_2 \beta_2'$ be positive semidefinite (Goldberger, 1964, p. 37).

II. If $m \leq k$ and rank of $A$ is equal to $m$, a necessary and sufficient condition for $A (F_2^{-1} - \frac{1}{\sigma^2} \beta_2 \beta_2') A'$ be positive semidefinite is that $F_2^{-1} - \frac{1}{\sigma^2} \beta_2 \beta_2'$ be positive semidefinite.

The sufficiency part of II is proven immediately by usage of I. To show necessity we assume $A (F_2^{-1} - \frac{1}{\sigma^2} \beta_2 \beta_2') A'$ is positive semidefinite and rank of $A = (X_1 X_1)^{-1} X_1' X_2$ is $m$ where $A$ is a $k \times m$ matrix and $m \leq k$. The $A'$ matrix consists of $m$ linearly independent vectors (say, e.g., the first $m$ columns) and $k-m$ column vectors that can be written as linear combinations of the first $m$ column vectors since their rank is $m$. Since any $m$ linearly independent subset of vectors from $m$-space spans

---

6 Rank of $A \leq (\text{minimum } m, k)$. Here we assume it is exactly equal to $m$. 


the entire space, we find all m-vectors by changing the values of the k-vector \( c \) in the equation \( A'c = d \). So we can conclude that under the assumptions of II:

\[
(25) \quad \mathbf{c} \mathbf{A}(F_2^{-1} - \frac{1}{\sigma^2} \mathbf{2} \mathbf{2})' \mathbf{A} \mathbf{c} \geq 0
\]

for all \( \mathbf{c} \) (kx1) implies that

\[
(26) \quad \mathbf{d}'(F_2^{-1} - \frac{1}{\sigma^2} \mathbf{2} \mathbf{2})' \mathbf{d} \geq 0
\]

for all \( \mathbf{d} \) (mx1), i.e., \( F_2^{-1} - \frac{1}{\sigma^2} \mathbf{2} \mathbf{2} \)' is positive semidefinite.

It is not too optimistic to expect that in most econometrics problems the matrix \( A \) will be of rank \( m \). Therefore, the condition that \( F_2^{-1} - \frac{1}{\sigma^2} \mathbf{2} \mathbf{2} \)' be positive semidefinite will be a sufficient and necessary condition for deleting \( X_2 \) from the "true model" if our criterion is to select between the full and the restricted estimator that one that minimizes the mean square of an arbitrary linear function of the estimates of \( \mathbf{2}_1 \).

We can arrange (26) as

\[
(27) \quad \frac{\mathbf{d}' \frac{1}{\sigma^2} \mathbf{2} \mathbf{2} \mathbf{d}}{\mathbf{d}' F_2^{-1} \mathbf{d}} \leq 1
\]

for all \( \mathbf{d} \) (mx1). Since \( F_2^{-1} \) is positive definite and \( \mathbf{2}_2 \) is an m-vector, then

\[
(28) \quad \sup_{\mathbf{d}} \frac{\mathbf{d}' \frac{1}{\sigma^2} \mathbf{2} \mathbf{2} \mathbf{d}}{\mathbf{d}' F_2^{-1} \mathbf{d}} = \frac{\mathbf{2}_2 F_2 \mathbf{2}}{\sigma^2}
\]

where the supremum is attained at \( \mathbf{d}^* = F_2 \mathbf{2}_2 \) (Rao, 1965, p. 48). Using
equations (27) and (28) we can show that $F_2^{-1} - \frac{1}{\sigma^2 B_2 B_2'}$ is positive semidefinite if and only if

$$\frac{\beta_2' F_2 B_2}{\sigma^2} \leq 1.$$  \hspace{1cm} (29)

\textbf{Sufficiency:}

Let $d' \left( F_2^{-1} - \frac{1}{\sigma^2 B_2 B_2'} \right) d \geq 0$ for all $d$. Since $\frac{d' \frac{1}{\sigma^2 B_2 B_2'} d}{d' F_2^{-1} d} \leq 1$ for all $d$, the inequality will be satisfied for the particular value

$$d^* = F_2 B_2, \text{ i.e., } \frac{\beta_2' F_2 B_2}{\sigma^2} \leq 1.$$

\textbf{Necessity:}

Let $\frac{\beta_2' F_2 B_2}{\sigma^2} \leq 1$.

From (28) $\frac{\beta_2' F_2 B_2}{\sigma^2} = \sup_d \frac{d' \frac{1}{\sigma^2 B_2 B_2'} d}{d' F_2^{-1} d}$. Since $\frac{\beta_2' F_2 B_2}{\sigma^2} \leq 1$

then $\frac{d' \frac{1}{\sigma^2 B_2 B_2'} d}{d' F_2^{-1} d} \leq 1$ for all $d$, i.e., $F_2^{-1} - \frac{1}{\sigma^2 B_2 B_2'}$ is a positive semi-

\text{define matrix.}

Since $F_2$ is a positive definite, the expression $\frac{\beta_2' F_2 B_2}{\sigma^2}$ is greater than zero for all nonnull $B_2$ vectors, it is zero only if $B_2 = 0$. Thus, equation (29) can be written as

$$0 \leq \frac{\beta_2' F_2 B_2}{\sigma^2} \leq 1.$$  \hspace{1cm} (30)

In the next chapter we present the relationship of $\frac{\beta_2' F_2 B_2}{\sigma^2}$ with the noncentral $F$ distribution that arises in multiple regression problems and develop a test of hypothesis for (30). We conclude this chapter with two simple examples.
I. Using only two regressors the inequality (30) simplifies to

\[
0 \leq \frac{1}{\sigma^2} \beta_2^2 \left[ \frac{\Sigma x_2^2 - (\Sigma x_1 x_2)^2}{\Sigma x_1^2} \right] \leq 1.
\]

Rearranging terms and using the definition of the variance of \( \hat{b}_2 \),
the inequality can be written as

\[
0 \leq \frac{\beta_2^2}{\text{Var}(b_2)} \leq 1 \quad \text{(Wallace, 1964, p. 1182)}.
\]

II. When we are interested in deciding to include or exclude only
one regressor, say \( X_2 = X_{k+1} \), the mean square error criterion (30) simplifies to

\[
0 \leq \frac{\beta_{k+1}^2}{\sigma^2} \Sigma x_{k+1}^2 (1 - R_{k+1,1,2,...,k}^2) \leq 1,
\]

where \( R_{k+1,1,2,...,k}^2 = \frac{X_{k+1}'X_{k+1}(X_{k+1}'X_{k+1})^{-1}X_{k+1}'X_{k+1}}{\Sigma x_{k+1}^2} \), i.e., the "coefficient of
determination." Since \( \text{Var}(\hat{b}_{k+1}) = \frac{\sigma^2}{\Sigma x_{k+1}^2(1 - R_{k+1,1,2,...,k}^2)} \), the above
inequality can be written as

\[
0 \leq \frac{\beta_{k+1}^2}{\text{Var}(\hat{b}_{k+1})} \leq 1.
\]

In the two situations described above the mean square error criterion
(32) and (34) are necessary and sufficient conditions without making the
assumption found in footnote 6. The mean square error criterion, (22),
implies that if \( \hat{b}_1 \) is to be chosen, \( \sigma^2 A F_2^{-1} A' - A \hat{b}_2 \hat{b}_2 A' \) must be positive
semidefinite. In the above situations the mean square error criterion simplifies to require that a scalar must be nonnegative.

For two regressors the M.S.E. criterion simplifies to

\[ [\text{Var}(\hat{b}_2) - \beta_2^2] \left[ \frac{\Sigma x_1 x_2}{\Sigma x_1^2} \right]^2 \geq 0. \]

Since \[ \left[ \frac{\Sigma x_1 x_2}{\Sigma x_1^2} \right]^2 \geq 0, \] (35) is equivalent to (32).

For including or excluding one regressor the M.S.E. criterion simplifies to

\[ [\text{Var}(\hat{b}_{k+1}) - \beta_{k+1}^2] (X_1'X_1)^{-1}X_1'(X_{k+1}2^2 X_1(X_1'X_1)^{-1} \]

Since the matrix in (36) is positive semidefinite (Graybill, 1961, p. 4), the M.S.E. criterion as stated in (36) is equivalent to (34).
CHAPTER IV
THE NONCENTRAL F DISTRIBUTION AND THE MEAN SQUARE ERROR CRITERION

Preliminaries

We found in the previous chapter that the mean square error criterion to include or exclude $X_2$ reduced to finding out if (30) is satisfied.

We showed that, under the linear model of full rank, (30) is a necessary and sufficient condition when we are dealing with a two-regressors model or $m=1$. When $m>2$, we must further assume that rank of $A$ is $m$ in order to have (30) as necessary and sufficient. Using the mean square error criterion the sets of variables $X_1$ (n×k) and $X_2$ (n×m) are defined to be "multicollinear" whenever $\beta_2' F_2 \beta_2 / \sigma^2$ is bounded by 0 and 1. When "multicollinearity" exists, we proceed to estimate $\beta_1$ by the restricted estimator deleting $X_2$ from the model.

Since the mean square error criterion is a function of the unknown parameters $\beta_2$ and $\sigma^2$, it is impossible to compute the value of $\beta_2' F_2 \beta_2 / \sigma^2$ for a specific situation. However, the situation is not hopeless since, as to be shown in the next section, $\frac{1}{2\sigma^2} (\beta_2' F_2 \beta_2)$ is the noncentrality parameter of a widely used "statistic" and we can turn to the uniformly most powerful test of $\frac{1}{2\sigma^2} (\beta_2' F_2 \beta_2)$ being less than or equal to one-half against the alternative of being greater than one-half.

A Uniformly More Powerful Test for Multicollinearity

Usually researchers delete or include $X_2$ by testing the null hypothesis $\beta_2 = 0$. The hypothesis $\beta_2 = 0$ is equivalent to $\frac{1}{2\sigma^2} (\beta_2' F_2 \beta_2)$ is zero since $F_2$ is positive definite. The usual (likelihood ratio or
the conditional error) test statistic for the null hypothesis, \( \hat{\beta}_2 = 0 \), under the model described in Chapter I is

\[
(37) \quad u = \frac{(n-p)Q_1}{mQ_o}
\]

where \( Q_1 \) is the quadratic form \( \hat{\beta}_2' F_2 \hat{\beta}_2 \) and \( Q_o \) is the quadratic form \( Y' [I - X(X'X)^{-1}X'] Y \). The quadratic form \( Q_1 \) is distributed as a noncentral chi square with \( m \) degrees of freedom and noncentrality parameter \( \lambda = \frac{1}{2G^2} (\hat{\beta}_2' F_2 \hat{\beta}_2) \). The quadratic form \( Q_o \) is distributed as a central chi square with \( n-p \) degrees of freedom. Since \( Q_1 \) and \( Q_o \) are independent, the random variable \( u \) is distributed as a noncentral \( F \) with \( m \) and \( n-p \) degrees of freedom, and noncentrality parameter \( \lambda \) (Graybill, 1961, p. 135).

The following theorem is extremely helpful to show that the test statistic \( u \) distributed as noncentral \( F \) distribution provides us with a uniformly most powerful (U.M.P.) test for "multicollinearity," i.e., \( 0 \leq \lambda \leq \frac{1}{2} \). Since \( \lambda \) is nonnegative, the "multicollinearity" criterion is simply \( \lambda \leq \frac{1}{2} \).

**Theorem** (Lehmann, 1964, p. 68):

Let \( \lambda \) be a real parameter, and let the random variable \( u \) have probability density \( f_\lambda(u) \) with monotone likelihood ratio in \( T(u) \). The real parameter family of densities \( f_\lambda(u) \) is said to have monotone likelihood ratio if there exists a real valued function \( T(u) \) such that for any \( \lambda_0 < \lambda_1 \) the distributions are distinct, and the ratio

\[
\frac{f_{\lambda_1}(u)}{f_{\lambda_0}(u)}
\]

is a nondecreasing function of \( T(u) \).
I. For testing $H_0: \lambda \leq \lambda_0$ against $K: \lambda > \lambda_0$, there exists a uniformly most powerful test, which is given by:

\[
\phi(u) = \begin{cases} 
1 \text{ when } T(u) > c; \text{ reject } H_0 \text{ with probability } 1 \\
T(u) > c, \\
\delta \text{ when } T(u) = c; \text{ reject } H_0 \text{ with probability } \delta \\
\text{if } T(u) = c, \text{ and} \\
0 \text{ when } T(u) < c; \text{ reject } H_0 \text{ with probability } 0 \\
\text{if } T(u) < c,
\end{cases}
\]

(38)

where $c$ and $\delta$ are determined by the following expectation:

\[
E_{\lambda_0} \phi(u) = \alpha.
\]

(39)

II. The power function $\beta(\lambda) = E_{\lambda} \phi(u)$ of this test is strictly increasing for all points $\lambda$ for which $\beta(\lambda) < 1$.

III. For all $\lambda'$, the test determined by (38) and (39) is U.M.P. for testing $H': \lambda \leq \lambda'$ against $K': \lambda > \lambda'$ at level $\alpha' = \beta(\lambda')$.

IV. For any $\lambda \leq \lambda_0$, the test minimizes $\beta(\lambda)$ [the probability of an error of the first kind] among all tests satisfying (39).

Let $f(u,m,q,\lambda)$, for short $f_{\lambda}(u)$, stand for the noncentral $F$ probability density of the random variable $u$ with $m$ and $q$ degrees of freedom and noncentrality parameter $\lambda$,

\[
f_{\lambda}(u) = \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{2i+m+q}{2}\right)}{\Gamma\left(\frac{q}{2}\right) \Gamma\left(\frac{2i+m}{2}\right) i!} \left(\frac{u}{q}\right)^{\frac{q}{2}} \lambda^i e^{-\lambda} \frac{1}{2} (2i+2m-2) (1 + \frac{mu}{q})^\frac{1}{2} (2i+m+q)
\]

(40)

where $q = n-p$ and $\Gamma$ stands for the gamma function.
Using the hint given by Lehmann (1964, p. 313, problem 4(i)), we proceed to show that \( f_{\lambda_1} (u) / f_{\lambda_0} (u) \) is a nondecreasing function of the real valued function \( \frac{mu}{q+mu} \) (0 \( \leq \) \( \frac{mu}{q+mu} \) \( \leq \) 1) for all \( \lambda_0 < \lambda_1 \).

After performing some of the operations indicated in (40), we obtained (41).

\[
(41) \quad f_{\lambda} (u) = \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{2i+m+q}{2}\right) \left(\frac{m}{q}\right)^i \left(\frac{m}{q}\right)}{\Gamma\left(\frac{m+q}{2}\right) i!} e^{-\lambda} u^i \left(1 + \frac{mu}{q}\right)^{\frac{m+q}{2}}
\]

Collecting all terms not involving \( i \) outside the summation sign, the following expression is obtained:

\[
(42) \quad f_{\lambda} (u) = \frac{e^{-\lambda} u^{\frac{m-2}{2} \left(\frac{m}{q}\right)}}{\Gamma\left(\frac{m+i}{2}\right) \left(1 + \frac{mu}{q}\right)^{\frac{m+q}{2}}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{2i+m+q}{2}\right) \lambda^i}{\Gamma\left(\frac{2i+m}{2}\right) i!} w^i
\]

where \( w = \frac{mu}{q + mu} \) is a monotone increasing function of \( u \), and \textit{vice versa}.

Substituting the corresponding values for \( f_{\lambda_1} (u) \) and \( f_{\lambda_0} (u) \), their ratio can be written as in (43).

\[
(43) \quad \frac{f_{\lambda_1} (u)}{f_{\lambda_0} (u)} = \frac{b_i}{a_i} \sum_{i=0}^{\infty} \frac{b_i}{a_i} \frac{w^i}{w^i}
\]

where

\[
b_i = \frac{\Gamma\left(\frac{2i+m+q}{2}\right) \lambda^i}{\Gamma\left(\frac{2i+m}{2}\right) i!} > 0, \quad a_i = \frac{\Gamma\left(\frac{2i+m+q}{2}\right) \lambda^i}{\Gamma\left(\frac{2i+m}{2}\right) i!} > 0
\]

and \( \sum_{i=0}^{\infty} b_i w^i \) and \( \sum_{i=0}^{\infty} a_i w^i \) converge for all \( w \) (Kaplan, 1959, p. 350).
For \( f_1(u)/f_0(u) \) to be a nondecreasing function of \( w \), its first derivative must be nonnegative for all values of \( w \), i.e.,

\[
\frac{-\lambda_1}{e^{-\lambda_0}} \sum_{k=0}^{\infty} a_k w^k \sum_{n=0}^{\infty} b_n w^{n-1} - \sum_{n=0}^{\infty} b_n w^n \sum_{k=0}^{\infty} ka_k w^{k-1}
\]

\[
\geq 0
\]

for all \( w \).

The above inequality is obtained by differentiating (43) with respect to \( w \). To avoid confusion, two indexes of summation are used. The differentiation process is valid since we are dealing with convergent power series and continuous derivatives (Kaplan, 1959, pp. 148-150).

Since \( e^{-\lambda_1}/e^{-\lambda_0} \) is always greater than zero and \( \left( \sum_{k=0}^{\infty} a_k w^k \right)^2 \) is a finite positive number, our attention shall be focused on

\[
\sum_{k=0}^{\infty} a_k w^k \sum_{n=0}^{\infty} b_n w^{n-1} - \sum_{n=0}^{\infty} b_n w^n \sum_{k=0}^{\infty} ka_k w^{k-1}
\]

Our task is to show (45) to be nonnegative for all \( w \). We can rewrite (45) as follows:

\[
\sum_{n} \left( \sum_{k=0}^{n-k} a_k b_n w^{n+k-1} \right)
\]

Partitioning the sum over \( k \) into values of \( k \) less than \( n \) and values of \( k \) greater than \( n \) (when \( k=n \) the term \( n-k \) in (46) vanishes), we rewrite (46) as

\[
\sum_{n} \left( \sum_{k=0}^{n-k} a_k b_n w^{n+k-1} + \sum_{n} \left( \sum_{k=n}^{n+k-1} a_k b_n w^{n+k-1} \right) \right)
\]
Adding and subtracting $\sum_{n} \sum_{k<n} (n-k) a_n b_k w^{n+k-1}$ we obtained

$$\Sigma \Sigma (n-k) (a_k b_n - a_n b_k) w^{n+k-1} + \Sigma \Sigma (n-k) a_n b_k w^{n+k-1}$$

$$+ \Sigma \Sigma (n-k) a_k b_n w^{n+k-1}$$

Changing the order of summation in the third term of (48) we obtained

$$\Sigma \Sigma (n-k) a_k b_n w^{n+k-1}$$

Since $n$ and $k$ are indexes of summation we can substitute them by $i$ and $j$. The second term in (48) is written as

$$\Sigma \Sigma (i-j) a_i b_j w^{i+j-1}$$

and the third term in (48) as

$$\Sigma \Sigma (j-i) a_i b_j w^{i+j-1}$$

Then the sum of (50) and (51), i.e., the sum of the second and third term in (48), is zero, and (45) simplifies to the scalar

$$\Sigma \Sigma (n-k) (a_k b_n - a_n b_k) w^{n+k-1}$$

Since $k<n$ and $0<w<1$ it follows that $n-k$ and $w^{n+k-1}$ are nonnegative, but what about $a_k b_n - a_n b_k$?

Taking the ratios of $b_i$ to $a_i$ for the $i^{th}$ and $i+1^{th}$ term, expressions (53) and (54) are obtained.
(53) \[ \frac{b_i}{a_i} = \left( \frac{\lambda_i}{\lambda} \right)^i \]

(54) \[ \frac{b_{i+1}}{a_{i+1}} = \left( \frac{\lambda_i}{\lambda} \right)^{i+1} \]

Since \( \lambda_o < \lambda_1 \), \[ \frac{b_i}{a_i} < \frac{b_{i+1}}{a_{i+1}} \text{ or } a_i b_{i+1} - a_{i+1} b_i > 0 \text{ for all } i, \] more generally \( a_k b_n - a_n b_k > 0 \text{ for } k < n \). Therefore, (44) is nonnegative and hence \( f_{\lambda_1}(u)/f_{\lambda_o}(u) \) is a nondecreasing function of \( w(0 < w < 1) \).

The density function of \( w \); \( h(w; m, q, \lambda) \), is easily derived by using transformation of variables techniques in (40).

(55) \[ h(w; m, q, \lambda) = \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} \frac{\Gamma\left(\frac{2i+m+q}{2}\right)}{\Gamma\left(\frac{q}{2}\right) \Gamma\left(\frac{2i+m}{2}\right)} \frac{1}{2}(2i+m-2) \frac{1}{(1-w)^2}(q-2) \]

where \( 0 < w < 1 \).

If \( \lambda = 0 \), \( h(w; m, q, \lambda = 0) \) is the beta distribution. Similarly when \( \lambda > 0 \), \( h(w; m, q, \lambda) \) is termed the noncentral beta distribution (Graybill, 1961, p. 79).

We have shown that the family of noncentral \( F \) densities have the monotone likelihood ratio property in \( w \). Using the above theorem (Lehmann, 1964, p. 68) we conclude that the U.M.P. test for "multicollinearity," i.e., \( \lambda < \frac{1}{2} \), at an \( \alpha \) level of significance is given by

\[ H_0: \lambda < \frac{1}{2} \text{ against } H: \lambda > \frac{1}{2} \]

Accept \( H_0 \): if \( w^* < w' \)

Reject \( H_0 \): if \( w^* \geq w' \)
where the critical point $w_*'$ is determined by

\[ \int_0^w h(w; m, q, \frac{1}{2}) \, dw = 1 - \alpha \]

and $w^*$ stands for the computed value of $w$.

Carrying out term by term integration indicated in (56) for the
distribution presented in (55), we have the following:

\[ \sum_{i=0}^{\infty} \frac{\left(\frac{1}{2}\right)^i e^{-\frac{1}{2}}}{i!} I_{w_*'}\left(\frac{1}{2}m+i, \frac{1}{2}q\right) = 1 - \alpha . \]

where

\[ I_{w_*'}\left(\frac{1}{2}m+i, \frac{1}{2}q\right) = \int_0^w \frac{\Gamma\left(\frac{m+q}{2} + i\right)}{\Gamma\left(\frac{q}{2}\right)\Gamma\left(\frac{1}{2}m+i\right)} \frac{1}{w^2} \frac{d}{dw} \left(1-w\right)^{\frac{q-1}{2}} \cdot \]

Hence, (57) is an infinite series whose terms are the product of a
term from a Poisson series with an Incomplete Beta Function Ratio. As
the values of the Poisson term and Incomplete Beta Function Ratio lie
between 0 and 1, the calculation of $w_*'$ will not be too heavy if the
number of terms giving either $\sum_{i=0}^{\infty} \frac{\left(\frac{1}{2}\right)^i e^{-\frac{1}{2}}}{i!}$ or $\sum_{i=0}^{\infty} I_{w_*'}\left(\frac{1}{2}m+i, \frac{1}{2}q\right)$ to
a desired accuracy is not too large. In Chapter V critical points, $w_*'$
are given for several $\alpha$ levels of significance for $m=1$ and selected values
of $n-p=q$. Also the power

\[ \beta(\lambda) = \int_{w_*'}^1 h(w; 1, q, \lambda) \, dw = 1 - \int_0^{w_*'} h(w; 1, q, \lambda) \, dw \]

is presented for several $\lambda$'s.

If we transform form $w$ to $u$ by the monotone increasing function

\[ u = \frac{qw}{n-\lambda w} \]

we get
\[
\alpha = \int_{w_\alpha}^{1} h(w; m, q, \lambda) \, dw = \int_{u_\alpha}^{\infty} f(u; m, q, \lambda) \, du
\]

where \( u_\alpha = \frac{q \, w_\alpha}{m - mw_\alpha} \).\(^7\) Hence we can perform the uniform most powerful test for the null hypothesis, \( \lambda \leq \lambda_0 \), by using the beta distribution or the F distribution. The test statistic for the F distribution is

\[
u^* = \frac{qQ_1}{mQ_0} \quad \text{and} \quad w^* = \frac{mu^*}{q + mu^*} \quad \text{for the beta distribution.}
\]

For the null hypothesis, \( \lambda = 0 \), the critical point will be denoted as \( w_\alpha^0 \) for the beta distribution and \( u_\alpha^0 \) for the F distribution. Under the null hypothesis, \( \lambda \leq \frac{1}{2} \), the critical point will be denoted as \( w_\alpha^0 \) for the beta distribution and \( u_\alpha^0 \) for the F distribution.

**Comparison of the Usual F Test and the Mean Square Error Criterion Test**

In most instances researchers interested in estimating \( \beta_1 \) decide to include \( X_2 \) or not by performing the central F test for the null hypothesis \( \beta_2 = 0 \), i.e., \( \lambda = 0 \).

If the computed \( u \) (or \( u^* \)) is nonsignificant at some assigned significance level, we omit the terms containing \( X_2 \) and use \( \hat{\nu} \) as the estimate of \( \beta_1 \). If the computed \( u \) (or \( u^* \)) is significant, we retain the terms containing \( X_2 \) and use \( \hat{\beta}_1 \) as the estimate of \( \beta_1 \). Transforming the variable \( u \) into \( w \), the test for \( \beta_2 = 0 \) can be carried out by using the central beta distribution for the null hypothesis \( \lambda = 0 \). The testing of \( \beta_2 = 0 \), or alternatively, \( \lambda = 0 \), weights heavily the interest in an unbiased estimator of \( \beta_1 \) without

\(^7\)In Chapter V we present critical points and powers for the case of deleting one variable, i.e., \( m = 1 \).
paying due consideration to the variance of the estimator. If \( \lambda > 0 \) the unbiased estimator \( \hat{b}_1 \) (F.L.S.E.) shall be used according to the usual test. This procedure disregards the following: (1) The linear function of the restricted estimators \( \begin{pmatrix} c^t \hat{b}_1 \end{pmatrix} \), though biased, has smaller variance than \( c^t \hat{b}_1 \). (2) If the noncentrality parameter, \( \lambda \), is between zero and one-half the mean square error of \( c^t \hat{b}_1 \) is smaller than the mean square error of \( c^t \hat{b}_1 \).

In the mean square error criterion test we consider the mean square of linear function, hence both bias and variance are considered. We delete \( X_2 \) from our model when we considered \( X_2 \) to be "multicollinear" with \( X_1 \), i.e., when the null hypothesis, \( \lambda < \frac{1}{2} \), is accepted.

The following results are immediate from the properties of the uniformly more powerful test used to test the hypothesis \( \lambda < \frac{1}{2} \).

I. Testing \( H_0: \lambda = 0 \) at \( \alpha_0 \) level is the same as testing \( H_0: \lambda < \frac{1}{2} \) at \( \alpha_1 \) level where \( \alpha_1 > \alpha_0 \) since the power function is strictly increasing for all points \( \lambda \) for which \( \beta(\lambda) < 1 \).

II. For the same \( \alpha \) level the critical point for the M.S.E. test is greater than the critical point for the usual test, i.e., \( \omega_\alpha > \omega_\alpha^0 \) in terms of the beta distribution or \( u_\alpha > u_\alpha^0 \) in terms of the F distribution.

Therefore if \( X_2 \) is deleted in the usual test at a level, i.e., \( H_0: \lambda = 0 \) is accepted then \( X_2 \) is deleted under the M.S.E. criterion, i.e., \( H_0: \lambda < \frac{1}{2} \) is accepted at a level. Analogously, if \( X_2 \) is kept using the M.S.E. test at a level then \( X_2 \) is kept at a level under the usual test. If \( \omega^* \) lies between \( \omega_\alpha^0 \) and \( \omega_\alpha \), we will delete \( X_2 \) under the M.S.E. test but keep \( X_2 \) under the usual test.
Minimum Variance Unbiased Estimator of $\lambda$

From a very well-known theorem (Graybill, 1961, p. 113) we have that $\hat{\sigma}^2, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_p$ form a set of estimators that are jointly sufficient and complete for $\sigma^2, \beta_1, \beta_2, \ldots, \beta_p$. Therefore, if a function $h(\hat{\sigma}^2, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_p)$ can be found such that $E[h(\hat{\sigma}^2, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_p)] = g(\sigma^2, \beta_1, \beta_2, \ldots, \beta_p)$, then $h(\hat{\sigma}^2, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_p)$ is an unbiased estimator of $g(\sigma^2, \beta_1, \beta_2, \ldots, \beta_p)$ and has smaller variance for a given sample size than any other unbiased estimator of $g(\sigma^2, \beta_1, \beta_2, \ldots, \beta_p)$, Rao-Blackwell theorem.

The random variable previously defined as

$$ u = \frac{(n-p)Q_1}{m Q_0} \quad (60) $$

can be written as

$$ u = \frac{\hat{b}_2 F_2 \hat{b}_2}{m \hat{\sigma}^2} \quad (61) $$

and $u$ is distributed as a noncentral $F$ with $m$ and $n-p$ degrees of freedom with noncentrality parameter

$$ \lambda = \frac{\hat{b}_2 F_2 \hat{b}_2}{2 \sigma^2} $$

Now

$$ E(u) = \frac{q}{q-2} + \frac{2\lambda q}{m(q-2)} \quad (62) $$

where $q = n-p > 2$ (Patnaik, 1949, p. 221).
Rearranging (62) we find that

$$E\left[ \frac{um(q-2)}{2q} - \frac{m}{2} \right] = \lambda. $$

Since $\frac{um(q-2)}{2q} - \frac{m}{2}$ is an unbiased estimator of $\lambda$ and is a function of $\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_p, \hat{c}^2$, therefore $\frac{um(q-2)}{2q} - \frac{m}{2}$ is the minimum variance unbiased estimator of $\lambda$. 
CHAPTER V
CRITICAL POINTS AND VALUES OF THE POWER OF THE
"MULTICOLLINEARITY" TEST FOR ONE DEGREE OF
FREEDOM IN THE NUMERATOR

Critical Points

When we are interested in deciding whether to exclude only one regressor in the linear model of full rank, the necessary and sufficient condition under the "multicollinearity" test simplifies to

\[ \beta_{k+1}^2 \frac{1}{\text{Var}(\hat{\beta}_{k+1})} \leq \frac{1}{2} \] (see page 20). Hence using the results from Chapter IV we can perform the uniform most powerful test for the null hypothesis \( \lambda \leq \frac{1}{2} \) by using the Beta or F distribution. The test statistic for the F distribution simplifies to \( u^* = \frac{\beta_{k+1}^2}{\text{Var}(\hat{\beta}_{k+1})} \), the well-known test statistic for the null hypothesis \( \beta_{k+1} = 0 \), when only one regressor is under study for exclusion, i.e., \( m=1 \).

In Chapter IV the theoretical framework of the U.M.P. test for "multicollinearity" and the comparison with the usual test (\( \beta_{k+1} = 0 \)) are presented. We proceed now to describe the procedure by which the critical points were obtained for the "multicollinearity" test for \( m=1 \). The computations were made using the noncentral beta distribution and via the monotonic inverse transformation, \( u^* = \frac{q^\alpha}{1 - q^\alpha} \), the noncentral F critical points were obtained for purpose of ready comparison with critical points of the central F distribution which is so widely used to make classical regression tests.

In order to estimate the critical points in the noncentral beta distribution, we used the approximation

\[
\sum_{i=0}^{7} \frac{(\frac{1}{2})^i}{i!} \frac{1}{I_\alpha'(\frac{1}{2} + i, \frac{1}{2})} = 1 - \alpha
\] (63)
for $\alpha = .05, .10, .25, .50$, where $I_{\frac{1}{2}}$ is the Incomplete Beta Function Ratio, and $q = n - p$.

We truncated the series to the first eight terms with error less than $10^{-6}$ for the following reasons:

(i) The sum of the Poisson weights for $\lambda = \frac{1}{2}$ beyond the first seven terms is less than $10^{-6}$ (Molina, 1947).

(ii) The Incomplete Beta Function Ratio is bounded by $[0,1]$ and for $0 < w_{\alpha} < 1$ is a monotonically decreasing function of $\frac{1}{2} + i$ (Gun, 1965).

A searching process is carried out for $w_{\alpha}$ until equation (63) is satisfied. The searching process is simplified by the fact that equation (63) is a monotonically increasing function of $w_{\alpha}$ with domain and range $[0,1]$.

The searching process for a particular $\alpha$ and a given T-K was:

(i) Try $w_{\alpha} = \frac{1}{2}$ in equation (63).

(ii) If $w_{\alpha} = \frac{1}{2}$ yields an evaluation of equation (63) $= 1 - \alpha$ within the absolute error $5 \times 10^{-4}$, we stop.

(iii) If the evaluation of (63) lies outside the range $1 - \alpha \pm (5 \times 10^{-4})$, we step $w_{\alpha}$ by one-half of the remaining interval either to the right or left depending on whether the error is negative or positive.

(iv) The process continues until the evaluation is within $5 \times 10^{-4}$ of $1 - \alpha$.

A 1410 IBM Computer was used, employing the Incomplete Beta Ratio subroutine written by Gautschi (1964). The values of $w_{\alpha}$ were carried to eight decimals and then were used to compute $u_{\alpha}$ by means of the monotonic transformation $u_{\alpha} = qw_{\alpha}/(1 - w_{\alpha})$. 
A Fortran statement of the computation of the critical points for selected values of \( n - (k+1) = q \) is presented in Appendix A. In Table 1 the critical points for the noncentral beta and F distribution were rounded to three and two decimal points respectively. Due to the great sensitivity of transforming the noncentral beta critical points to the noncentral F critical points through \( u_\alpha = \frac{q\alpha}{1-\alpha} \), an inconsistent result in the second decimal point was obtained for the critical point corresponding to \( q = 28, \alpha = .05 \). For that reason the critical points for \( q = 26, 27, 28, 29 \) were deleted from Table 1. Intuitively, we recognized the sensitivity of mapping the domain \([0,1]\) of the noncentral beta distribution into the noncentral F distribution domain \([0,\infty]\). Furthermore we truncated an infinite series into an eight terms sum and approximated the \( \alpha \) level with an absolute error of \( 5(10^{-4}) \). On the other hand we do claim that for the three decimals in the noncentral beta and for the two decimals in the noncentral F critical points, the results presented in Table 1 are accurate. The supporting evidence for this claim is postponed until some values of the power of the "multicollinearity" test are presented. We have provided four choices for the level of significance to emphasize the fact that the researcher should make his choice of \( \alpha \) by considering the specific problem under study and his personal "utility function" by which he weights the Type I and II errors "costs." If a large \( \alpha \) level is chosen we implicitly regard as more "costly" smaller variance but biased estimates for \( \hat{\beta}_1 \) than the B.L.U.E. of \( \hat{\beta}_1 \) (\( \hat{\beta}_1 \) versus \( \hat{\beta}_1 \)). On the other hand, if a small \( \alpha \) level is chosen we implicitly regard as more "costly" the B.L.U.E. of \( \hat{\beta}_1 \) (\( \hat{\beta}_1 \)) than the smaller variance but biased estimator (\( \hat{\beta}_1 \)).
Table 1. Noncentral beta and noncentral F "multicollinearity" test of critical points for one degree of freedom in the numerator

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( w_{\alpha} )</th>
<th>( u_{\alpha} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>q</td>
<td>.05</td>
<td>.10</td>
</tr>
<tr>
<td>1</td>
<td>.997</td>
<td>.988</td>
</tr>
<tr>
<td>2</td>
<td>.949</td>
<td>.897</td>
</tr>
<tr>
<td>3</td>
<td>.869</td>
<td>.788</td>
</tr>
<tr>
<td>4</td>
<td>.789</td>
<td>.693</td>
</tr>
<tr>
<td>5</td>
<td>.717</td>
<td>.617</td>
</tr>
<tr>
<td>6</td>
<td>.654</td>
<td>.554</td>
</tr>
<tr>
<td>7</td>
<td>.602</td>
<td>.502</td>
</tr>
<tr>
<td>8</td>
<td>.555</td>
<td>.459</td>
</tr>
<tr>
<td>9</td>
<td>.516</td>
<td>.422</td>
</tr>
<tr>
<td>10</td>
<td>.480</td>
<td>.391</td>
</tr>
<tr>
<td>12</td>
<td>.424</td>
<td>.341</td>
</tr>
<tr>
<td>13</td>
<td>.400</td>
<td>.320</td>
</tr>
<tr>
<td>14</td>
<td>.379</td>
<td>.302</td>
</tr>
<tr>
<td>15</td>
<td>.359</td>
<td>.285</td>
</tr>
<tr>
<td>16</td>
<td>.342</td>
<td>.270</td>
</tr>
<tr>
<td>17</td>
<td>.326</td>
<td>.257</td>
</tr>
<tr>
<td>18</td>
<td>.312</td>
<td>.245</td>
</tr>
<tr>
<td>19</td>
<td>.299</td>
<td>.234</td>
</tr>
<tr>
<td>20</td>
<td>.287</td>
<td>.224</td>
</tr>
<tr>
<td>21</td>
<td>.275</td>
<td>.215</td>
</tr>
<tr>
<td>22</td>
<td>.266</td>
<td>.206</td>
</tr>
<tr>
<td>23</td>
<td>.256</td>
<td>.198</td>
</tr>
<tr>
<td>24</td>
<td>.246</td>
<td>.191</td>
</tr>
<tr>
<td>25</td>
<td>.238</td>
<td>.184</td>
</tr>
<tr>
<td>30</td>
<td>.203</td>
<td>.157</td>
</tr>
<tr>
<td>40</td>
<td>.158</td>
<td>.121</td>
</tr>
<tr>
<td>60</td>
<td>.109</td>
<td>.082</td>
</tr>
<tr>
<td>120</td>
<td>.056</td>
<td>.042</td>
</tr>
</tbody>
</table>
Furthermore, as it will be discussed in Chapter VI, the choice of the significance level has implications in the bias and mean square of sequential estimators. The usage of a critical point of one in the noncentral $F$, which corresponds approximately to a 50 percent level as Table 1 indicates, will be of special interest in Chapter VI.

The context for using Table 1 is a multiple regression model of a dependent variable regressed on $K+1$ nonstochastic regressors. It is assumed that the regression error is normally and independently distributed with zero mean and constant variance. The researcher considers the exclusion of $x_{k+1}$ in order to provide a "best" estimator for $\theta_1(kx1)$. The null hypothesis is $\lambda \leq \frac{1}{2}$, i.e., the R.L.S.E. $(\hat{b}_1)$ is to be preferred over the F.L.S.E. $(\hat{b}_1)$.

The procedure is

(i) Compute $u^* = \frac{\text{SSE}(\hat{b}_1) - \text{SSE}(\hat{b})}{\text{SSE}(\hat{b})/n-(k+1)}$

where $\text{SSE}(\hat{b}_1)$ is the error sum of squares when $x_{k+1}$ is excluded from the model and $\text{SSE}(\hat{b})$ is the error sum of squares when $x_{k+1}$ is included in the model.

(ii) For a chosen level of $\alpha$, compare $u^*$ in (i) with critical point $u'_\alpha$ in Table 1 that corresponds to the degrees of freedom $n-(k+1)$.

(iii) If $u^* \geq u'_\alpha$, reject the hypothesis that the restricted estimators are better in MSE. If $u^* < u'_\alpha$, accept the hypothesis that the restricted estimators are better in MSE.

**Power Values**

In this section power values of the U.M.P. test for the null hypothesis $\lambda \leq \frac{1}{2}$ are presented. The values of the power were evaluated
for the following noncentrality parameters: 0, .25, .3333, .5, 1.0, 
2.25, 4.0, 6.25, 7.5625 and 9.0. When $\lambda = 0$, the power is evaluated 
directly by the Incomplete Beta Ratio subroutine since in this case the 
noncentral beta distribution reduces to the central beta distribution, i.e., 

$$\beta(0) = 1 - I^{-1} _{\frac{1}{2}, \frac{a}{2}} \left( \frac{1}{2}, \frac{a}{2} \right),$$ 

where $\beta(\lambda)$ stands for the power corresponding to the noncentrality 
parameter $\lambda$. For the other values of the noncentrality parameter the 
power evaluation is made by the following approximation, 

$$\beta(\lambda) = 1 - \sum_{i=0}^{N(\lambda)} \frac{\lambda^i e^{-\lambda}}{i!} I^{-1} _{\frac{1}{2} + i, \frac{a}{2}} \left( \frac{1}{2} + i, \frac{a}{2} \right)$$ 

where the truncation $N(\lambda)$ depends on $\lambda$. Table 2 indicates the relationship between the truncation and the parameter $\lambda$ and also indicates maximum truncation error for the evaluation of the power presented in 
Tables 3, 4, 5, and 6.

In Appendix B the Fortran statement used in the evaluation of the power is presented. As for the search of the critical points, a 1410 IBM Computer and the Incomplete Beta Ratio subroutine written by Gautschi (1964) were used. The input of the Fortran statement consists of:

(i) The truncation term, $N(\lambda)$.

(ii) The noncentral beta critical point rounded to four decimals for the corresponding $\alpha$ level and denominator degrees of freedom.

From the power tables we can obtain the corresponding significance level for the usual test by looking at the power for $\lambda = 0$. The power for $\lambda = \frac{1}{2}$ was evaluated in order to see how close we come with four
Table 2. Largest values of $i$ used in computing the power for $\lambda \neq 0^a$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$N(\lambda)$</th>
<th>Maximum truncation error</th>
</tr>
</thead>
<tbody>
<tr>
<td>.2500</td>
<td>6</td>
<td>$3 \times 10^{-7}$</td>
</tr>
<tr>
<td>.3333</td>
<td>6</td>
<td>$4 \times 10^{-6}$</td>
</tr>
<tr>
<td>1.0000</td>
<td>9</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>2.2500</td>
<td>13</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>4.0000</td>
<td>17</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>6.2500</td>
<td>22</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>7.5625</td>
<td>24</td>
<td>$2 \times 10^{-6}$</td>
</tr>
<tr>
<td>9.0000</td>
<td>27</td>
<td>$10^{-6}$</td>
</tr>
</tbody>
</table>

$^a$See Molina (1947, Table 2).

decimal points in the critical point to the intended $\alpha$ level since we rounded to three decimals the critical points in Table 1. For the 5 percent level we found one case in which we had an absolute error in the level of significance greater than $5(10^{-4})$. This happened for $q = 1$ where the value of the power was .0509. For the 10 percent level all values of $\beta(\frac{1}{2})$ were within the allowed absolute error. For the 25 percent level the only case for which $\beta(\frac{1}{2})$ was beyond the absolute error was for $q = 120$ where the value of the power was .2506. For the 50 percent level two $\beta(\frac{1}{2})$ were not of the allowed absolute error. They were .5008 for $q = 16$ and .5006 for $q = 22$. We must realize that in practice we are satisfied to be reasonably close to the intended $\alpha$ level and that extreme accuracy in the evaluations of critical points is partially lost when we only use two or three decimals critical points, i.e.,
Table 3. Power values for the "multicollinearity" test at 5 percent significance level

<table>
<thead>
<tr>
<th>q</th>
<th>0</th>
<th>.2500</th>
<th>.3333</th>
<th>1.0000</th>
<th>2.2500</th>
<th>4.0000</th>
<th>6.2500</th>
<th>7.5625</th>
<th>9.0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.035</td>
<td>.043</td>
<td>.046</td>
<td>.065</td>
<td>.093</td>
<td>.123</td>
<td>.154</td>
<td>.169</td>
<td>.184</td>
</tr>
<tr>
<td>2</td>
<td>.026</td>
<td>.038</td>
<td>.042</td>
<td>.074</td>
<td>.131</td>
<td>.205</td>
<td>.291</td>
<td>.336</td>
<td>.383</td>
</tr>
<tr>
<td>3</td>
<td>.021</td>
<td>.035</td>
<td>.040</td>
<td>.081</td>
<td>.161</td>
<td>.273</td>
<td>.408</td>
<td>.478</td>
<td>.548</td>
</tr>
<tr>
<td>4</td>
<td>.018</td>
<td>.033</td>
<td>.039</td>
<td>.085</td>
<td>.182</td>
<td>.323</td>
<td>.490</td>
<td>.574</td>
<td>.654</td>
</tr>
<tr>
<td>5</td>
<td>.016</td>
<td>.032</td>
<td>.038</td>
<td>.089</td>
<td>.199</td>
<td>.362</td>
<td>.550</td>
<td>.641</td>
<td>.724</td>
</tr>
<tr>
<td>6</td>
<td>.015</td>
<td>.032</td>
<td>.038</td>
<td>.092</td>
<td>.212</td>
<td>.391</td>
<td>.593</td>
<td>.688</td>
<td>.770</td>
</tr>
<tr>
<td>7</td>
<td>.014</td>
<td>.031</td>
<td>.037</td>
<td>.093</td>
<td>.221</td>
<td>.412</td>
<td>.624</td>
<td>.719</td>
<td>.801</td>
</tr>
<tr>
<td>8</td>
<td>.013</td>
<td>.031</td>
<td>.037</td>
<td>.096</td>
<td>.231</td>
<td>.432</td>
<td>.650</td>
<td>.746</td>
<td>.825</td>
</tr>
<tr>
<td>9</td>
<td>.013</td>
<td>.030</td>
<td>.037</td>
<td>.096</td>
<td>.236</td>
<td>.444</td>
<td>.668</td>
<td>.763</td>
<td>.841</td>
</tr>
<tr>
<td>10</td>
<td>.012</td>
<td>.030</td>
<td>.036</td>
<td>.098</td>
<td>.243</td>
<td>.457</td>
<td>.684</td>
<td>.779</td>
<td>.855</td>
</tr>
<tr>
<td>11</td>
<td>.012</td>
<td>.030</td>
<td>.036</td>
<td>.098</td>
<td>.246</td>
<td>.465</td>
<td>.694</td>
<td>.789</td>
<td>.864</td>
</tr>
<tr>
<td>12</td>
<td>.012</td>
<td>.029</td>
<td>.036</td>
<td>.099</td>
<td>.250</td>
<td>.474</td>
<td>.705</td>
<td>.800</td>
<td>.873</td>
</tr>
<tr>
<td>13</td>
<td>.011</td>
<td>.029</td>
<td>.036</td>
<td>.100</td>
<td>.253</td>
<td>.480</td>
<td>.713</td>
<td>.807</td>
<td>.880</td>
</tr>
<tr>
<td>14</td>
<td>.011</td>
<td>.029</td>
<td>.036</td>
<td>.100</td>
<td>.256</td>
<td>.487</td>
<td>.721</td>
<td>.814</td>
<td>.885</td>
</tr>
<tr>
<td>15</td>
<td>.011</td>
<td>.029</td>
<td>.036</td>
<td>.101</td>
<td>.259</td>
<td>.493</td>
<td>.728</td>
<td>.821</td>
<td>.890</td>
</tr>
<tr>
<td>16</td>
<td>.011</td>
<td>.029</td>
<td>.036</td>
<td>.102</td>
<td>.262</td>
<td>.498</td>
<td>.734</td>
<td>.826</td>
<td>.895</td>
</tr>
<tr>
<td>17</td>
<td>.011</td>
<td>.029</td>
<td>.035</td>
<td>.102</td>
<td>.264</td>
<td>.502</td>
<td>.738</td>
<td>.830</td>
<td>.898</td>
</tr>
<tr>
<td>18</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.102</td>
<td>.264</td>
<td>.504</td>
<td>.741</td>
<td>.833</td>
<td>.900</td>
</tr>
<tr>
<td>19</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.102</td>
<td>.266</td>
<td>.508</td>
<td>.746</td>
<td>.837</td>
<td>.904</td>
</tr>
<tr>
<td>20</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.102</td>
<td>.267</td>
<td>.510</td>
<td>.749</td>
<td>.839</td>
<td>.906</td>
</tr>
<tr>
<td>21</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.103</td>
<td>.270</td>
<td>.514</td>
<td>.753</td>
<td>.843</td>
<td>.908</td>
</tr>
<tr>
<td>22</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.103</td>
<td>.270</td>
<td>.515</td>
<td>.754</td>
<td>.844</td>
<td>.910</td>
</tr>
<tr>
<td>23</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.103</td>
<td>.271</td>
<td>.518</td>
<td>.757</td>
<td>.847</td>
<td>.911</td>
</tr>
<tr>
<td>24</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.104</td>
<td>.274</td>
<td>.522</td>
<td>.762</td>
<td>.850</td>
<td>.914</td>
</tr>
<tr>
<td>25</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.104</td>
<td>.274</td>
<td>.523</td>
<td>.763</td>
<td>.852</td>
<td>.915</td>
</tr>
<tr>
<td>30</td>
<td>.010</td>
<td>.028</td>
<td>.035</td>
<td>.105</td>
<td>.279</td>
<td>.533</td>
<td>.773</td>
<td>.860</td>
<td>.921</td>
</tr>
<tr>
<td>40</td>
<td>.009</td>
<td>.027</td>
<td>.034</td>
<td>.105</td>
<td>.282</td>
<td>.540</td>
<td>.781</td>
<td>.867</td>
<td>.927</td>
</tr>
<tr>
<td>60</td>
<td>.009</td>
<td>.027</td>
<td>.034</td>
<td>.107</td>
<td>.289</td>
<td>.552</td>
<td>.793</td>
<td>.877</td>
<td>.933</td>
</tr>
<tr>
<td>120</td>
<td>.009</td>
<td>.027</td>
<td>.034</td>
<td>.109</td>
<td>.296</td>
<td>.564</td>
<td>.804</td>
<td>.886</td>
<td>.940</td>
</tr>
</tbody>
</table>
Table 4. Power values for the "multicollinearity" test at 10 percent significance level

<table>
<thead>
<tr>
<th>$q$</th>
<th>0</th>
<th>.2500</th>
<th>.3333</th>
<th>1.0000</th>
<th>2.2500</th>
<th>4.0000</th>
<th>6.2500</th>
<th>7.5625</th>
<th>9.0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.068</td>
<td>.085</td>
<td>.089</td>
<td>.127</td>
<td>.181</td>
<td>.238</td>
<td>.295</td>
<td>.323</td>
<td>.351</td>
</tr>
<tr>
<td>2</td>
<td>.053</td>
<td>.077</td>
<td>.084</td>
<td>.145</td>
<td>.248</td>
<td>.372</td>
<td>.501</td>
<td>.564</td>
<td>.624</td>
</tr>
<tr>
<td>3</td>
<td>.044</td>
<td>.072</td>
<td>.081</td>
<td>.156</td>
<td>.290</td>
<td>.456</td>
<td>.624</td>
<td>.700</td>
<td>.767</td>
</tr>
<tr>
<td>4</td>
<td>.040</td>
<td>.070</td>
<td>.080</td>
<td>.163</td>
<td>.318</td>
<td>.511</td>
<td>.696</td>
<td>.774</td>
<td>.838</td>
</tr>
<tr>
<td>5</td>
<td>.036</td>
<td>.067</td>
<td>.078</td>
<td>.167</td>
<td>.336</td>
<td>.545</td>
<td>.738</td>
<td>.815</td>
<td>.876</td>
</tr>
<tr>
<td>6</td>
<td>.034</td>
<td>.066</td>
<td>.077</td>
<td>.171</td>
<td>.351</td>
<td>.571</td>
<td>.768</td>
<td>.843</td>
<td>.900</td>
</tr>
<tr>
<td>7</td>
<td>.033</td>
<td>.065</td>
<td>.077</td>
<td>.174</td>
<td>.362</td>
<td>.590</td>
<td>.788</td>
<td>.861</td>
<td>.915</td>
</tr>
<tr>
<td>8</td>
<td>.031</td>
<td>.064</td>
<td>.076</td>
<td>.175</td>
<td>.369</td>
<td>.603</td>
<td>.803</td>
<td>.874</td>
<td>.925</td>
</tr>
<tr>
<td>9</td>
<td>.030</td>
<td>.064</td>
<td>.076</td>
<td>.178</td>
<td>.377</td>
<td>.616</td>
<td>.815</td>
<td>.884</td>
<td>.933</td>
</tr>
<tr>
<td>10</td>
<td>.030</td>
<td>.064</td>
<td>.076</td>
<td>.179</td>
<td>.382</td>
<td>.625</td>
<td>.824</td>
<td>.892</td>
<td>.938</td>
</tr>
<tr>
<td>11</td>
<td>.029</td>
<td>.063</td>
<td>.075</td>
<td>.179</td>
<td>.386</td>
<td>.631</td>
<td>.830</td>
<td>.897</td>
<td>.942</td>
</tr>
<tr>
<td>12</td>
<td>.028</td>
<td>.062</td>
<td>.075</td>
<td>.180</td>
<td>.389</td>
<td>.637</td>
<td>.836</td>
<td>.902</td>
<td>.946</td>
</tr>
<tr>
<td>13</td>
<td>.028</td>
<td>.062</td>
<td>.074</td>
<td>.181</td>
<td>.392</td>
<td>.642</td>
<td>.841</td>
<td>.906</td>
<td>.948</td>
</tr>
<tr>
<td>14</td>
<td>.028</td>
<td>.062</td>
<td>.074</td>
<td>.182</td>
<td>.395</td>
<td>.647</td>
<td>.845</td>
<td>.909</td>
<td>.951</td>
</tr>
<tr>
<td>15</td>
<td>.027</td>
<td>.062</td>
<td>.074</td>
<td>.183</td>
<td>.398</td>
<td>.651</td>
<td>.849</td>
<td>.912</td>
<td>.953</td>
</tr>
<tr>
<td>16</td>
<td>.027</td>
<td>.062</td>
<td>.074</td>
<td>.183</td>
<td>.400</td>
<td>.654</td>
<td>.852</td>
<td>.914</td>
<td>.955</td>
</tr>
<tr>
<td>17</td>
<td>.027</td>
<td>.062</td>
<td>.074</td>
<td>.184</td>
<td>.403</td>
<td>.658</td>
<td>.856</td>
<td>.917</td>
<td>.956</td>
</tr>
<tr>
<td>18</td>
<td>.026</td>
<td>.061</td>
<td>.074</td>
<td>.184</td>
<td>.404</td>
<td>.660</td>
<td>.857</td>
<td>.919</td>
<td>.958</td>
</tr>
<tr>
<td>19</td>
<td>.026</td>
<td>.061</td>
<td>.074</td>
<td>.184</td>
<td>.405</td>
<td>.662</td>
<td>.859</td>
<td>.920</td>
<td>.958</td>
</tr>
<tr>
<td>20</td>
<td>.026</td>
<td>.061</td>
<td>.074</td>
<td>.185</td>
<td>.407</td>
<td>.665</td>
<td>.861</td>
<td>.922</td>
<td>.960</td>
</tr>
<tr>
<td>21</td>
<td>.026</td>
<td>.061</td>
<td>.074</td>
<td>.185</td>
<td>.409</td>
<td>.667</td>
<td>.863</td>
<td>.923</td>
<td>.960</td>
</tr>
<tr>
<td>22</td>
<td>.026</td>
<td>.061</td>
<td>.074</td>
<td>.186</td>
<td>.411</td>
<td>.670</td>
<td>.865</td>
<td>.924</td>
<td>.961</td>
</tr>
<tr>
<td>23</td>
<td>.026</td>
<td>.061</td>
<td>.074</td>
<td>.186</td>
<td>.412</td>
<td>.672</td>
<td>.867</td>
<td>.926</td>
<td>.962</td>
</tr>
<tr>
<td>24</td>
<td>.025</td>
<td>.061</td>
<td>.074</td>
<td>.186</td>
<td>.412</td>
<td>.672</td>
<td>.867</td>
<td>.926</td>
<td>.962</td>
</tr>
<tr>
<td>25</td>
<td>.025</td>
<td>.061</td>
<td>.074</td>
<td>.186</td>
<td>.413</td>
<td>.674</td>
<td>.868</td>
<td>.927</td>
<td>.963</td>
</tr>
<tr>
<td>30</td>
<td>.025</td>
<td>.060</td>
<td>.073</td>
<td>.186</td>
<td>.415</td>
<td>.678</td>
<td>.872</td>
<td>.930</td>
<td>.965</td>
</tr>
<tr>
<td>40</td>
<td>.024</td>
<td>.060</td>
<td>.073</td>
<td>.188</td>
<td>.421</td>
<td>.686</td>
<td>.879</td>
<td>.934</td>
<td>.968</td>
</tr>
<tr>
<td>60</td>
<td>.024</td>
<td>.060</td>
<td>.073</td>
<td>.190</td>
<td>.426</td>
<td>.693</td>
<td>.884</td>
<td>.938</td>
<td>.970</td>
</tr>
<tr>
<td>120</td>
<td>.023</td>
<td>.059</td>
<td>.072</td>
<td>.191</td>
<td>.431</td>
<td>.700</td>
<td>.890</td>
<td>.942</td>
<td>.973</td>
</tr>
</tbody>
</table>
Table 5. Power values for the "multicollinearity" test at 25 percent significance level

<table>
<thead>
<tr>
<th>q</th>
<th>0</th>
<th>.2500</th>
<th>.3333</th>
<th>1.0000</th>
<th>2.2500</th>
<th>4.0000</th>
<th>6.2500</th>
<th>7.5625</th>
<th>9.0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.174</td>
<td>.214</td>
<td>.226</td>
<td>.314</td>
<td>.437</td>
<td>.556</td>
<td>.662</td>
<td>.708</td>
<td>.749</td>
</tr>
<tr>
<td>2</td>
<td>.143</td>
<td>.198</td>
<td>.216</td>
<td>.343</td>
<td>.529</td>
<td>.704</td>
<td>.837</td>
<td>.885</td>
<td>.922</td>
</tr>
<tr>
<td>3</td>
<td>.129</td>
<td>.191</td>
<td>.211</td>
<td>.358</td>
<td>.571</td>
<td>.764</td>
<td>.894</td>
<td>.934</td>
<td>.961</td>
</tr>
<tr>
<td>4</td>
<td>.120</td>
<td>.187</td>
<td>.208</td>
<td>.365</td>
<td>.594</td>
<td>.793</td>
<td>.917</td>
<td>.953</td>
<td>.974</td>
</tr>
<tr>
<td>5</td>
<td>.116</td>
<td>.184</td>
<td>.207</td>
<td>.371</td>
<td>.609</td>
<td>.811</td>
<td>.930</td>
<td>.962</td>
<td>.981</td>
</tr>
<tr>
<td>6</td>
<td>.112</td>
<td>.183</td>
<td>.206</td>
<td>.375</td>
<td>.619</td>
<td>.822</td>
<td>.938</td>
<td>.968</td>
<td>.984</td>
</tr>
<tr>
<td>7</td>
<td>.109</td>
<td>.181</td>
<td>.204</td>
<td>.377</td>
<td>.625</td>
<td>.830</td>
<td>.943</td>
<td>.971</td>
<td>.986</td>
</tr>
<tr>
<td>8</td>
<td>.107</td>
<td>.180</td>
<td>.204</td>
<td>.379</td>
<td>.630</td>
<td>.835</td>
<td>.947</td>
<td>.974</td>
<td>.988</td>
</tr>
<tr>
<td>9</td>
<td>.105</td>
<td>.179</td>
<td>.203</td>
<td>.381</td>
<td>.635</td>
<td>.840</td>
<td>.950</td>
<td>.975</td>
<td>.989</td>
</tr>
<tr>
<td>10</td>
<td>.104</td>
<td>.179</td>
<td>.203</td>
<td>.382</td>
<td>.638</td>
<td>.844</td>
<td>.952</td>
<td>.977</td>
<td>.990</td>
</tr>
<tr>
<td>15</td>
<td>.100</td>
<td>.176</td>
<td>.201</td>
<td>.386</td>
<td>.648</td>
<td>.854</td>
<td>.958</td>
<td>.980</td>
<td>.992</td>
</tr>
<tr>
<td>16</td>
<td>.100</td>
<td>.176</td>
<td>.201</td>
<td>.387</td>
<td>.650</td>
<td>.856</td>
<td>.959</td>
<td>.981</td>
<td>.992</td>
</tr>
<tr>
<td>17</td>
<td>.099</td>
<td>.176</td>
<td>.201</td>
<td>.387</td>
<td>.650</td>
<td>.856</td>
<td>.959</td>
<td>.981</td>
<td>.992</td>
</tr>
<tr>
<td>18</td>
<td>.099</td>
<td>.176</td>
<td>.201</td>
<td>.387</td>
<td>.652</td>
<td>.857</td>
<td>.960</td>
<td>.982</td>
<td>.992</td>
</tr>
<tr>
<td>19</td>
<td>.098</td>
<td>.176</td>
<td>.201</td>
<td>.387</td>
<td>.652</td>
<td>.858</td>
<td>.960</td>
<td>.982</td>
<td>.992</td>
</tr>
<tr>
<td>20</td>
<td>.098</td>
<td>.176</td>
<td>.201</td>
<td>.388</td>
<td>.653</td>
<td>.859</td>
<td>.960</td>
<td>.982</td>
<td>.993</td>
</tr>
<tr>
<td>21</td>
<td>.098</td>
<td>.175</td>
<td>.201</td>
<td>.388</td>
<td>.654</td>
<td>.860</td>
<td>.961</td>
<td>.982</td>
<td>.993</td>
</tr>
<tr>
<td>22</td>
<td>.098</td>
<td>.175</td>
<td>.200</td>
<td>.388</td>
<td>.655</td>
<td>.860</td>
<td>.961</td>
<td>.982</td>
<td>.993</td>
</tr>
<tr>
<td>23</td>
<td>.097</td>
<td>.175</td>
<td>.200</td>
<td>.389</td>
<td>.655</td>
<td>.861</td>
<td>.962</td>
<td>.983</td>
<td>.993</td>
</tr>
<tr>
<td>24</td>
<td>.097</td>
<td>.175</td>
<td>.200</td>
<td>.389</td>
<td>.656</td>
<td>.862</td>
<td>.962</td>
<td>.983</td>
<td>.993</td>
</tr>
<tr>
<td>25</td>
<td>.097</td>
<td>.175</td>
<td>.200</td>
<td>.389</td>
<td>.656</td>
<td>.862</td>
<td>.962</td>
<td>.983</td>
<td>.993</td>
</tr>
<tr>
<td>30</td>
<td>.096</td>
<td>.175</td>
<td>.200</td>
<td>.390</td>
<td>.659</td>
<td>.864</td>
<td>.963</td>
<td>.984</td>
<td>.993</td>
</tr>
<tr>
<td>40</td>
<td>.095</td>
<td>.174</td>
<td>.200</td>
<td>.391</td>
<td>.661</td>
<td>.866</td>
<td>.964</td>
<td>.984</td>
<td>.994</td>
</tr>
<tr>
<td>60</td>
<td>.094</td>
<td>.174</td>
<td>.200</td>
<td>.392</td>
<td>.664</td>
<td>.869</td>
<td>.966</td>
<td>.985</td>
<td>.994</td>
</tr>
<tr>
<td>120</td>
<td>.093</td>
<td>.173</td>
<td>.199</td>
<td>.394</td>
<td>.667</td>
<td>.871</td>
<td>.967</td>
<td>.986</td>
<td>.994</td>
</tr>
</tbody>
</table>
Table 6. Power values for the "multicollinearity" test at 50 percent significance level

<table>
<thead>
<tr>
<th>q</th>
<th>0.2500</th>
<th>.3333</th>
<th>1.0000</th>
<th>2.2500</th>
<th>4.0000</th>
<th>6.2500</th>
<th>7.5625</th>
<th>9.0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.372</td>
<td>.440</td>
<td>.461</td>
<td>.600</td>
<td>.764</td>
<td>.882</td>
<td>.949</td>
<td>.968</td>
</tr>
<tr>
<td>2</td>
<td>.338</td>
<td>.425</td>
<td>.451</td>
<td>.622</td>
<td>.813</td>
<td>.930</td>
<td>.980</td>
<td>.990</td>
</tr>
<tr>
<td>3</td>
<td>.324</td>
<td>.419</td>
<td>.447</td>
<td>.631</td>
<td>.829</td>
<td>.943</td>
<td>.986</td>
<td>.994</td>
</tr>
<tr>
<td>4</td>
<td>.317</td>
<td>.415</td>
<td>.445</td>
<td>.635</td>
<td>.837</td>
<td>.948</td>
<td>.989</td>
<td>.995</td>
</tr>
<tr>
<td>5</td>
<td>.313</td>
<td>.414</td>
<td>.440</td>
<td>.638</td>
<td>.842</td>
<td>.952</td>
<td>.990</td>
<td>.996</td>
</tr>
<tr>
<td>6</td>
<td>.309</td>
<td>.412</td>
<td>.443</td>
<td>.640</td>
<td>.844</td>
<td>.954</td>
<td>.991</td>
<td>.996</td>
</tr>
<tr>
<td>7</td>
<td>.308</td>
<td>.411</td>
<td>.443</td>
<td>.642</td>
<td>.847</td>
<td>.955</td>
<td>.991</td>
<td>.997</td>
</tr>
<tr>
<td>8</td>
<td>.305</td>
<td>.410</td>
<td>.442</td>
<td>.642</td>
<td>.848</td>
<td>.956</td>
<td>.991</td>
<td>.997</td>
</tr>
<tr>
<td>9</td>
<td>.304</td>
<td>.410</td>
<td>.441</td>
<td>.643</td>
<td>.849</td>
<td>.957</td>
<td>.992</td>
<td>.997</td>
</tr>
<tr>
<td>10</td>
<td>.303</td>
<td>.409</td>
<td>.441</td>
<td>.643</td>
<td>.850</td>
<td>.957</td>
<td>.992</td>
<td>.997</td>
</tr>
<tr>
<td>11</td>
<td>.303</td>
<td>.409</td>
<td>.441</td>
<td>.644</td>
<td>.851</td>
<td>.958</td>
<td>.992</td>
<td>.997</td>
</tr>
<tr>
<td>12</td>
<td>.302</td>
<td>.409</td>
<td>.441</td>
<td>.644</td>
<td>.852</td>
<td>.958</td>
<td>.992</td>
<td>.997</td>
</tr>
<tr>
<td>13</td>
<td>.301</td>
<td>.408</td>
<td>.440</td>
<td>.644</td>
<td>.852</td>
<td>.958</td>
<td>.992</td>
<td>.997</td>
</tr>
<tr>
<td>14</td>
<td>.301</td>
<td>.408</td>
<td>.440</td>
<td>.645</td>
<td>.853</td>
<td>.959</td>
<td>.992</td>
<td>.997</td>
</tr>
<tr>
<td>15</td>
<td>.300</td>
<td>.408</td>
<td>.440</td>
<td>.645</td>
<td>.853</td>
<td>.959</td>
<td>.992</td>
<td>.997</td>
</tr>
<tr>
<td>16</td>
<td>.300</td>
<td>.408</td>
<td>.440</td>
<td>.646</td>
<td>.854</td>
<td>.959</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>17</td>
<td>.299</td>
<td>.407</td>
<td>.440</td>
<td>.646</td>
<td>.854</td>
<td>.959</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>18</td>
<td>.299</td>
<td>.407</td>
<td>.440</td>
<td>.646</td>
<td>.854</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>19</td>
<td>.299</td>
<td>.407</td>
<td>.440</td>
<td>.646</td>
<td>.854</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>20</td>
<td>.298</td>
<td>.407</td>
<td>.440</td>
<td>.646</td>
<td>.854</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>21</td>
<td>.298</td>
<td>.407</td>
<td>.440</td>
<td>.646</td>
<td>.854</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>22</td>
<td>.298</td>
<td>.407</td>
<td>.440</td>
<td>.646</td>
<td>.855</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>23</td>
<td>.298</td>
<td>.407</td>
<td>.440</td>
<td>.646</td>
<td>.855</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>24</td>
<td>.298</td>
<td>.407</td>
<td>.440</td>
<td>.647</td>
<td>.855</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
<tr>
<td>25</td>
<td>.298</td>
<td>.407</td>
<td>.440</td>
<td>.647</td>
<td>.855</td>
<td>.960</td>
<td>.993</td>
<td>.997</td>
</tr>
</tbody>
</table>
we might be accurate to eight decimals but by rounding to two decimals, we are not better off than having accuracy to, say, five decimals.

Some Comments about the Critical Points and Power Tables

A comparison of Table 1, Noncentral beta and noncentral F "multicollinearity" test critical points for one degree of freedom in the numerator, with a standard central F table indicates that for the same $\alpha$ level the critical point for the "multicollinearity" test is greater than the critical point for the usual test. Tables 3, 4, 5, and 6 indicate that the power function is strictly increasing for all points $\lambda$ for which $\beta(\lambda) < 1$. They also provide the level of significance at which the usual test is carried out ($\lambda=0$). It is also evident from the power value tables that for a given $\alpha$ level and values of $\lambda$ smaller than one-half, as $q$ increases the power decreases, while for values of $\lambda$ larger than one-half, the power increases.

From Table 3 we find that for $q = 18(1)25$ and $q = 30$ the corresponding $\alpha$ level for the usual test is 1 percent. Since Tang (1938) presented critical points in the noncentral beta for the usual test at 1 percent level, we can compare the corresponding values of Table 1 with Tang's values. The differences in Table 1 and Tang's critical points can be attributed to the fact that the values of the power for $\lambda=0$ are rounded to the third decimal point. For that reason in Table 7 we incorporate the value of $\beta(0)$ at four decimals in order to provide a justification for the discrepancy in the critical points.

To indicate the sensitivity of transforming the noncentral beta critical point into the noncentral F critical point we present a comparison in the critical points of the noncentral F distribution and the
Table 7. Comparison of Tang's 1 percent and "multicollinearity" 5 percent critical points

<table>
<thead>
<tr>
<th>q</th>
<th>( \beta(0) )</th>
<th>( w^{0.01} )</th>
<th>( w^{0.05} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>.0103</td>
<td>.315</td>
<td>.312</td>
</tr>
<tr>
<td>19</td>
<td>.0103</td>
<td>.301</td>
<td>.299</td>
</tr>
<tr>
<td>20</td>
<td>.0101</td>
<td>.288</td>
<td>.287</td>
</tr>
<tr>
<td>21</td>
<td>.0101</td>
<td>.276</td>
<td>.275</td>
</tr>
<tr>
<td>22</td>
<td>.0099</td>
<td>.265</td>
<td>.266</td>
</tr>
<tr>
<td>23</td>
<td>.0098</td>
<td>.255</td>
<td>.256</td>
</tr>
<tr>
<td>24</td>
<td>.0099</td>
<td>.246</td>
<td>.246</td>
</tr>
<tr>
<td>25</td>
<td>.0098</td>
<td>.237</td>
<td>.238</td>
</tr>
<tr>
<td>30</td>
<td>.0096</td>
<td>.201</td>
<td>.203</td>
</tr>
</tbody>
</table>

\(^{a}\)See Tang (1938).

central F distribution. The comparison is presented in Table 8 for q = 20, 24, and 30 using Table 1 and a standard central F table. In Table 8 we include the corresponding critical points for the beta distribution.

The difference for q = 24 in the F critical points is due to the fact that even though for three decimals \( w^{0.01} = w^{0.05} = .246 \) the computations for \( u^{0.01} \) and \( u^{0.05} \) were carried out by different procedures and accuracy.

From Table 4 we find that for q = 24, 25, and 30 the corresponding \( \alpha \) level for the usual test is \( 2\frac{1}{2} \) percent. In Table 9 we include the power for \( \lambda = 0 \) using four decimals and compare the corresponding usual and
"multicollinearity" critical points via Table 1 and a standard central F table.

Table 8. Comparison of critical points for selected denominator degrees of freedom when \( \beta(0) \) is approximately .01

<table>
<thead>
<tr>
<th>q</th>
<th>( u^0 .01 )</th>
<th>( u^0 .05 )</th>
<th>( w^0 .01 )</th>
<th>( w^0 .05 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>8.10</td>
<td>8.05</td>
<td>.288</td>
<td>.287</td>
</tr>
<tr>
<td>24</td>
<td>7.82</td>
<td>7.83</td>
<td>.246</td>
<td>.246</td>
</tr>
<tr>
<td>30</td>
<td>7.56</td>
<td>7.65</td>
<td>.201</td>
<td>.203</td>
</tr>
</tbody>
</table>

Table 9. Comparison of F critical points for selected denominator degrees of freedom when \( \beta(0) \) is approximately .025

<table>
<thead>
<tr>
<th>q</th>
<th>( \beta(0) )</th>
<th>( u^0 .025 )</th>
<th>( u^0 .10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>.0254</td>
<td>5.72</td>
<td>5.68</td>
</tr>
<tr>
<td>25</td>
<td>.0253</td>
<td>5.69</td>
<td>5.66</td>
</tr>
<tr>
<td>30</td>
<td>.0246</td>
<td>5.57</td>
<td>5.60</td>
</tr>
</tbody>
</table>

From Table 5 we find that for \( q = 14, 15 \), the corresponding \( \alpha \) level for the usual test is 10 percent. In Table 10 we include the power for \( \lambda = 0 \) using four decimals and compare the corresponding usual and "multicollinearity" critical points via Table 1 and a standard central F table.

This section provides neither rigorous nor exhaustive accuracy checks of the critical points table for the "multicollinearity" test. There are two reasons for this: first, there are not to my knowledge available tables for the central F tables with the wide range of \( \alpha \) levels
Table 10. Comparison of F critical points for selected denominator degrees of freedom when \( \beta(0) \) is approximately .10

<table>
<thead>
<tr>
<th>q</th>
<th>( \beta(0) )</th>
<th>( u_{.10} )</th>
<th>( u_{.25} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>.1001</td>
<td>3.07</td>
<td>3.07</td>
</tr>
<tr>
<td>16</td>
<td>.0999</td>
<td>3.05</td>
<td>3.05</td>
</tr>
</tbody>
</table>

implied in Tables 3, 4, 5, and 6 for the usual test; second, our main purpose in this section is to emphasize the relationship in the levels of significance for the usual and "multicollinearity" tests.
CHAPTER VI
ESTIMATION AFTER PRELIMINARY TESTS OF SIGNIFICANCE

Sequential Estimation

The context in which we considered the mean square error test in Chapter IV assumes a well-specified model where the error term is normally independently distributed with mean zero and constant variance. But the nature of economic data being what it is, a restriction on the parameter space may need to be considered. In considering the restriction, the researcher may be more interested in the properties of his structural estimates than in the "truth" or "falsity" of the restriction. In this thesis we decide between the F.L.S.E. of $\hat{\beta}_1$ and the R.L.S.E. of $\hat{\beta}_1$, i.e., $\hat{b}_1$ vs. $\hat{b}_1$, via the mean square error criterion. The underlying restriction used is $\beta_2 = 0$ with the purpose of finding the "best" estimator of $\beta_1$ where the decision is based upon the existence of "multicollinearity."

In Chapter IV a U.M.P. test for "multicollinearity" is obtained and in Chapter V the critical points for this test given $m=1$, i.e., $\beta_{k+1} = 0$, are presented. For this simple case the criterion simplifies to test if

$$|\hat{\beta}_{k+1}| < \sqrt{V(\hat{b}_{k+1})} \quad \text{or} \quad \frac{\hat{\beta}_{k+1}^2}{2V(\hat{b}_{k+1})} < \frac{1}{2}$$

in order to decide if $x_{k+1}$ should be excluded from the well-specified regression model of full rank in order to obtain the "best" estimator of $\beta_1$.

Traditionally, researchers have approached the problem of excluding variables from a different context. They have started with an uncertain specification in the linear regression model of full rank. After working the F.L.S.E. of $\hat{\beta}$, they have decided about the inclusion of a nonstochastic $x_{k+1}$ by testing the hypothesis that $\beta_{k+1} = 0$, i.e., $\frac{\hat{\beta}_{k+1}^2}{2V(\hat{b}_{k+1})} = 0$. 
A comparison of the M.S.E. test (test for multicollinearity) and the usual test is found in Chapter IV. As we noticed in Chapter V, the basic difference is the level of significance chosen for the test discounting the motivation issue.

The analogy in the above approaches also lies in the fact that in both cases a preliminary test of significance is used as an aid in deciding the estimator to be used for $\hat{\beta}_1$. The sequential estimator of $\hat{\beta}_1$ would be the R.L.S.E. $\left(\hat{b}_1\right)$ if the $H_O: \frac{\hat{\beta}_{k+1}^2}{2V(\hat{b}_{k+1})} \leq \frac{1}{2}$ is accepted, and the F.L.S.E. $\left(\hat{b}_1\right)$ otherwise when the "multicollinearity" test is used. When the usual test is used the estimator of $\hat{\beta}_1$ would be the R.L.S.E. $\left(\hat{b}_1\right)$ if the $H_O: \hat{\beta}_{k+1} = 0$ is accepted, and the F.L.S.E. $\left(\hat{b}_1\right)$ otherwise.

Bancroft (1944) pointed out that sequential estimators of regression coefficients on a two-regressors model based on a t-test are biased, the degree of bias depending on several parameters including the level of significance at which the test is made. In the following sections a comparison of the bias for the sequential estimators of $\hat{\beta}_1$ under the M.S.E. test, the usual test (both at 5 percent significance level) and a test where the critical point is one is provided. From Chapter V we noticed that a critical point of one in the F distribution is approximately equivalent to the M.S.E. test at 50 percent level of significance. An intuitive explanation for such a critical point would be that computed $t$ is the "estimating analogy" of $\frac{\hat{\beta}_2}{\sqrt{V(\hat{b}_2)}}$, i.e., computed $t$ is equal to $\frac{\hat{b}_2}{\sqrt{V(\hat{b}_2)}}$.

Since $\hat{b}_1$ is a biased estimator and $\hat{b}_1$ is an unbiased estimator and realizing that the basic difference among the above mentioned tests is the selection of significance level, or viz. the selection of the
critical point, the bias of the sequential estimators disappears as \( \alpha \) goes to one, i.e., the critical point in the F distribution goes to zero. If the critical point is zero, the null hypothesis is rejected with probability one and the unbiased estimator \( \hat{b}_1 \) is used. Henceforth an analysis is carried out to determine the effect of the chosen level of significance in the mean square of the sequential estimator in the two-regressors model as specified by Bancroft (1944). The analysis would seem to indicate that even though the bias decreases as the critical point goes to zero, holding constant the parameters of interest, the mean square behavior depends on the value of the noncentrality parameter of the noncentral F distribution.

**Sequential Estimator Bias**

The regression model of two regressors to be used is the same as Bancroft (1944). Let

\[
y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \quad i=1, \ldots, n
\]

where \( y_i, x_{1i} \) and \( x_{2i} \) are deviations from their respective means; \( \beta_1 \) and \( \beta_2 \) are the respective population regression coefficients and \( \epsilon_i \) is normally independently distributed with zero mean and unit variance. The \( x \)'s are nonstochastic with unit variances and "correlation coefficient"

\[
r_{12}
\]

so that

\[
\sum_{i=1}^{n} x_{1i}^2 = n-1, \quad \sum_{i=1}^{n} x_{2i}^2 = n-1, \quad \sum_{i=1}^{n} x_{1i} x_{2i} = r_{12}(n-1)
\]

i.e., the nonstochastic variables \( x_{1i} \) and \( x_{2i} \) are standardized variables (Goldberger, 1964, p. 198).
For purpose of easing the calculation of the sequential estimator bias, the following orthogonal transformation is carried out:

\[ \xi_{1i} = x_{1i}, \quad \xi_{2i} = x_{2i} - r_{12}x_{1i} \]

Then the transformed model is

\[ y_i = (\beta_1 + r_{12} \beta_2) \xi_{1i} + \beta_2 \xi_{2i} + \epsilon_i \]  

(67)

The above transformation provides us, after applying least squares estimating procedures, with the following results:

(i) an unadjusted estimator for \( \hat{\beta}_1 \) \((\hat{\beta}_1)\)

(ii) an adjusted estimator for \( \hat{\beta}_2 \) \((\hat{b}_2)\)

(iii) mean square due to \( x_1 \) ignoring \( x_2 \) \( (S_1^2) \)

(iv) mean square due to \( x_2 \) after fitting \( x_1 \) \( (S_2^2) \)

(v) residual mean square \( (S_3^2) \)

(vi) \( \hat{\beta}_1 \) is independent of \( \hat{\beta}_2 \)

(vii) \( S_1^2, S_2^2, \) and \( S_3^2 \) are independent sum of squares

In Goldberger (1964, pp. 176-196), a general presentation of the above results can be found.

Since \( \hat{\beta}_1 = \hat{\beta}_1 - r_{12} \hat{b}_2 \) and \( E(\hat{\beta}_1) = \beta_1 + r_{12} \beta_2 \), we obtained that the \( E(\hat{\beta}_1) = \beta_1 + r_{12} \beta_2 - r_{12} E(\hat{b}_2) \). Hence the unconditional expectation of \( \hat{\beta}_1 \) is \( \beta_1 \) and any expectation (conditional or unconditional) is \( \beta_1 \) if \( r_{12} = 0 \).

The bias and mean square comparisons will be carried out for the following sequential estimators:

(i) The "usual" sequential estimator \( (\hat{B}_1) \) where we test if \( \beta_2 = 0 \) at a level, i.e., \( \lambda = \frac{\beta_2^2}{\text{var}(\hat{b}_2)} = 0 \), where \( \lambda \) has been previously defined
as the noncentrality parameter in the noncentral F distribution and
\[ V(b_2) = \frac{1}{(n-1)(1-\tau_{12}^2)}, \]

\[ B_1 = \begin{cases} \hat{b}_1 & \text{if } u^* < u_\alpha^0, \ i.e., \ w^* < w_\alpha^0 \\ \hat{b}_1 & \text{if } u^* \geq u_\alpha^0, \ i.e., \ w^* \geq w_\alpha^0 \end{cases} \]

(ii) The "mean square" sequential estimator \( B_2 \) where we test if
\[ \beta_2 \leq \frac{\hat{\beta}_2^2}{V(b_2)}, \ i.e., \ \lambda = \frac{\hat{\beta}_2^2}{2V(b_2)} \leq \frac{1}{2} \text{ at } \alpha \text{ level}, \]

\[ B_2 = \begin{cases} \hat{b}_1 & \text{if } u^* < u_\alpha^1, \ i.e., \ w^* < w_\alpha^1 \\ \hat{b}_1 & \text{if } u^* \geq u_\alpha^1, \ i.e., \ w^* \geq w_\alpha^1 \end{cases} \]

(iii) The "50 percent" sequential estimator \( B_3 \) where we test if
\[ \lambda \leq \frac{1}{2} \text{ at approximately } .50 \text{ level}, \]

\[ B_3 = \begin{cases} \hat{b}_1 & \text{if } u^* < 1, \ i.e., \ w^* < \frac{1}{1+q} \\ \hat{b}_1 & \text{if } u^* \geq 1, \ i.e., \ w^* \geq \frac{1}{1+q} \end{cases} \]

We notice that the basic difference in the above sequential estimators is the chosen level of significance where \( u^*, w^*, u_\alpha^0, w_\alpha^0, u_\alpha^1, w_\alpha^1 \) and \( w_\alpha \) have been previously defined in Chapter IV. Therefore we will proceed to obtain the bias of a sequential estimator for a general \( \alpha \) level and subsequently make the proper substitutions for each special case. Let us call the "general" sequential estimator \( B \) where a preliminary test of significance of \( \beta_2 = 0 \) at \( \alpha \) has been used.

Hence,
\[ B = \begin{cases} \hat{b}_1 & \text{if } u^* < u_\alpha, \ i.e., \ w^* > w_\alpha \\ \hat{b}_1 & \text{if } u^* \geq u_\alpha, \ i.e., \ w^* \leq w_\alpha \end{cases} \]

and
\[ E[B] = E[\hat{b}_1 / \big[ u^* < u_\alpha \big] \Pr(u^* < u_\alpha) + E[\hat{b}_1 / \big[ u^* \geq u_\alpha \big] \Pr(u^* \geq u_\alpha)] \]
where $E[\hat{\theta}/\Omega]$ stands for expectation of $\hat{\theta}$ given $\Omega$, $Pr$ stands for probability
and $u^* = \frac{S_2}{S_3}$.

Since $S_1^2, S_2^2, S_3^2$ are independent and $b_1 = \frac{S_1}{\sqrt{n-1}}$, the expectation
of $\hat{b}_1$ given $\frac{S_2}{S_3} < u_\alpha$ is the unconditional expectation of $\hat{b}_1$, i.e.,

$$E(\frac{b_1}{u^*} < u_\alpha) = E(\hat{b}_1) = \beta_1 + r_{12} \beta_2.$$

Henceforth the expectation of $B$ simplifies to

$$E(B) = (B_1 + r_{12} \beta_2) \Pr(u^* < u_\alpha) + [(B_1 + r_{12} \beta_2) - r_{12} E(\hat{b}_2/u^* \geq u_\alpha)]$$
times $[1 - \Pr(u^* < u_\alpha)]$

$$= \beta_1 + r_{12} \beta_2 - r_{12} E(\hat{b}_2/u^* \geq u_\alpha)[1 - \Pr(u^* < u_\alpha)]$$

$$= \beta_1 + r_{12} \beta_2 - r_{12} E(\hat{b}_2/u^* \geq u_\alpha) \Pr(u^* \geq u_\alpha).$$

From Bancroft (1944) we know that

$$E(\hat{b}_2/u^* \geq u_\alpha) = \frac{\beta_2}{\Pr(u^* \geq u_\alpha)} \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} x_i \left(\frac{n-3}{2}, \frac{3}{2} + 1\right)$$

where

$$\lambda = \frac{\beta_2}{2} (n-1)(1-r_{12}^2),$$

$$x = \frac{n-3}{(n-3)+u^*},$$

$$x_0 = \frac{n-3}{(n-3)+u_\alpha},$$

and

$$I_{x_0}^{x} = \int_{x_0}^{x} \frac{\Gamma(n+2)}{\Gamma(n-3/2)\Gamma(3/2+1)} x^{n-3-1}(1-x)^{3/2+1-1} dx$$

(the incomplete beta ratio).
By the relationship between the F distribution and the beta distribution, we obtained the following:

\[ w = \frac{u}{(n-3)+u} = 1 - x, \]

\[ x = 1 - w \]

using the well-known identity

\[ I_{x}^{(n-3)} \left( \frac{3+2i}{2}, \frac{n-3}{2} \right) = 1 - I_{1-x}^{(3+2i)} \left( \frac{3+2i}{2}, \frac{n-3}{2} \right) \]

we obtained

\[ \sum_{i=0}^{\infty} \frac{\lambda^{i}e^{-\lambda}}{i!} I_{x}^{(n-3)} \left( \frac{3+2i}{2}, \frac{n-3}{2} \right) = 1 - \sum_{i=0}^{\infty} \frac{\lambda^{i}e^{-\lambda}}{i!} I_{w}^{(3+2i)} \left( \frac{3+2i}{2}, \frac{n-3}{2} \right) \]

Substituting (70) into (69) we obtained

\[ E(B) = \beta_1 + r_{12} \beta_2 \left[ 1 - \sum_{i=0}^{\infty} \frac{\lambda^{i}e^{-\lambda}}{i!} I_{x}^{(n-3)} \left( \frac{n-3}{2}, \frac{3+2i}{2} \right) \right] \]

Therefore the bias of B is

\[ Bias = r_{12} \beta_2 \left[ 1 - \sum_{i=0}^{\infty} \frac{\lambda^{i}e^{-\lambda}}{i!} I_{x}^{(n-3)} \left( \frac{n-3}{2}, \frac{3+2i}{2} \right) \right] \]

The following deductions follow immediately:

(i) There is no bias in estimating \(\beta_1\), if \(r_{12}\) or \(\beta_2\) is zero,

(ii) the absolute value of the coefficient of \(\beta_2\) is at most one,

(iii) the sign of the bias depends on the signs of \(r_{12}\) and \(\beta_2\),

(iv) the bias is independent of \(\beta_1\),
(v) the absolute bias decreases as the significance level approaches to one, i.e., as $u_{\alpha}$ approaches zero, and

(vi) the absolute bias decreases as $n$ increases for $u_{\alpha} = 1$ since then $x_{o}$ approaches one.

The "mean square" sequential estimator bias was evaluated for sample sizes 5, 11, 21, for $r_{12} = .2, .4, .6, .8$, and $\beta_2 = 0.1, 0.4, 1.0, 2.0,$ and 4.0. These are the same special cases analyzed by Bancroft (1944) under the "usual" test and the "50 percent" level test.

In Table 11 the critical points in terms of $u$'s (F distribution), $w$'s (beta distribution), and the corresponding $x_{o}$'s are summarized. The critical points for the "mean square" test at 5 percent level can be found in Table 1 in Chapter V. The critical points for the "usual" test at 5 percent level can be found for $u$ in any standard F table and in Tang (1938) for $w$.

In Tang (1938) and Table 3 in Chapter V, an evaluation of the "mean square" and "usual" tests of power function for selected values of the noncentrality parameter ($\lambda$) can be obtained. In Table 12 the "50 percent" test power function for selected values of $\lambda$ is presented.

The implications of the behavior of the power for the tests considered will be shown to be more important in the next section when the mean square errors for the sequential estimators are derived. As it will be shown, the mean square error is a function of the significance level and the noncentrality parameter.
Table 11. Critical points used in the evaluation of the sequential estimator's bias

<table>
<thead>
<tr>
<th>n sample size</th>
<th>n-3 degrees of freedom</th>
<th>&quot;Mean square&quot; sequential estimators</th>
<th>&quot;Usual&quot; sequential estimators</th>
<th>&quot;50 percent&quot; sequential estimators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>u ( .05 )</td>
<td>w ( .05 )</td>
<td>x(_o)</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>37.38</td>
<td>.949</td>
<td>.051</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>9.96</td>
<td>.555</td>
<td>.445</td>
</tr>
<tr>
<td>21</td>
<td>18</td>
<td>8.18</td>
<td>.312</td>
<td>.688</td>
</tr>
</tbody>
</table>

Table 12. Powers for the "50 percent" test for selected values of the noncentrality parameter for one degree of freedom in the numerator

<table>
<thead>
<tr>
<th>Denominator degrees of freedom</th>
<th>n-3 = q</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.423</td>
</tr>
<tr>
<td>8</td>
<td>.347</td>
</tr>
<tr>
<td>18</td>
<td>.331</td>
</tr>
</tbody>
</table>
A 1410 IBM Computer and the Incomplete Beta Ratio subroutine written by Gautschi (1964) were used in the evaluation of the sequential estimator biases. For the "usual" and "50 percent" sequential estimators, analogous results to Bancroft (1944) were obtained. A Fortran statement of the program used to evaluate the bias can be found in Appendix C. In Table 13 the corresponding value of the noncentrality parameter ($\lambda$) is obtained for the different combinations of $n$, $r_{12}$, and $\beta_2$. We deleted from the infinite series of the weighted Incomplete Beta Ratios those terms for which the Poisson weights were less than $10^{-7}$ (see Molina, 1947). A slight modification was made in the computation of the bias for some $\lambda$'s (represented with asterisks in Table 14) since the computing time required for the evaluation of the Incomplete Beta Ratio was relatively long. The modification was to set $I_{X_0}$ equal to one for the terms $I_{X_0}(\frac{N-3}{2}, \frac{3}{2} + i' + k), k = 1, 2, \ldots$, when $I_{X_0}(\frac{N-3}{2}, \frac{3}{2} + i') > 0.99999$ since $I_{X_0}$ is an increasing function of the second argument (see Gun, 1965).

In Table 14 the corresponding biases for the "mean square", "usual" and "50 percent" sequential estimators are presented for the special cases considered. The results seem to indicate:

(i) The bias decreases as we fix $r_{12}$, $\beta_2$, and the significance level and increase the sample size,

(ii) the bias decreases considerably as we fix $r_{12}$, $\beta_2$, and $n$ and increase the level of significance,

(iii) the bias increases without exception as we fix $n$, $\beta_2$, and the significance level and increases $r_{12}$ (when the collinearity between $x_{11}$ and $x_{21}$ increases or when $\lambda$ decreases) and
Table 13. Values of $\lambda$ and range $i$ (in parenthesis) used in computing the bias and mean square error of the sequential estimator of $\beta_1$

<table>
<thead>
<tr>
<th>$n = 5$</th>
<th>( r_{12} )</th>
<th>.2</th>
<th>.4</th>
<th>.6</th>
<th>.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>.0192</td>
<td>.0168</td>
<td>.0128</td>
<td>.0072</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-3)</td>
<td>(0-3)</td>
<td>(0-3)</td>
<td>(0-3)</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>.3072</td>
<td>.2688</td>
<td>.2048</td>
<td>.1152</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-7)</td>
<td>(0-6)</td>
<td>(0-6)</td>
<td>(0-5)</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1.9200</td>
<td>1.6800</td>
<td>1.2800</td>
<td>.7200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-11)</td>
<td>(0-11)</td>
<td>(0-10)</td>
<td>(0-8)</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>7.6800</td>
<td>6.7200</td>
<td>5.1200</td>
<td>2.8800</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-24)</td>
<td>(0-23)</td>
<td>(0-19)</td>
<td>(0-14)</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>30.7200</td>
<td>26.8800</td>
<td>20.4800</td>
<td>11,5200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8-61)</td>
<td>(5-55)</td>
<td>(0-46)</td>
<td>(0-31)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n = 11$</th>
<th>( r_{12} )</th>
<th>.2</th>
<th>.4</th>
<th>.6</th>
<th>.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>.0480</td>
<td>.0420</td>
<td>.0320</td>
<td>.0180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-4)</td>
<td>(0-4)</td>
<td>(0-4)</td>
<td>(0-3)</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>.7680</td>
<td>.6720</td>
<td>.5120</td>
<td>.2880</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-8)</td>
<td>(0-8)</td>
<td>(0-7)</td>
<td>(0-6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-19)</td>
<td>(0-18)</td>
<td>(0-16)</td>
<td>(0-12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-45)</td>
<td>(0-40)</td>
<td>(0-33)</td>
<td>(0-24)</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>76.8000</td>
<td>67.2000</td>
<td>51.2000</td>
<td>28.8000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(38-122)</td>
<td>(32-110)</td>
<td>(21-89)</td>
<td>(7-58)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n = 21$</th>
<th>( r_{12} )</th>
<th>.2</th>
<th>.4</th>
<th>.6</th>
<th>.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>.0960</td>
<td>.0840</td>
<td>.0640</td>
<td>.0360</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-5)</td>
<td>(0-5)</td>
<td>(0-5)</td>
<td>(0-4)</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>1.5360</td>
<td>1.3440</td>
<td>1.0240</td>
<td>.5760</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-11)</td>
<td>(0-10)</td>
<td>(0-10)</td>
<td>(0-8)</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>9.6000</td>
<td>8.4000</td>
<td>6.4000</td>
<td>3.6000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0-28)</td>
<td>(0-26)</td>
<td>(0-22)</td>
<td>(0-16)</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>38.4000</td>
<td>33.6000</td>
<td>25.6000</td>
<td>14.4000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(12-72)</td>
<td>(9-65)</td>
<td>(5-54)</td>
<td>(0-36)</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>153.6000</td>
<td>134.4000</td>
<td>102.4000</td>
<td>57.6000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(92-216)</td>
<td>(76-192)</td>
<td>(57-152)</td>
<td>(25-97)</td>
<td></td>
</tr>
</tbody>
</table>
Table 14. The bias in estimating $\beta_1$ via the "mean square," "usual," and "50 percent" sequential estimators

<table>
<thead>
<tr>
<th>$\beta_2$</th>
<th>$r_{12}$</th>
<th>( u_{.05} = 37.38 )</th>
<th>( u_{.05} = 9.96 )</th>
<th>( u_{.05} = 8.18 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.2</td>
<td>0.18</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>0.073</td>
<td>0.067</td>
<td>0.056</td>
</tr>
<tr>
<td>1.0</td>
<td>0.2</td>
<td>0.168</td>
<td>0.076</td>
<td>0.056</td>
</tr>
<tr>
<td>2.0</td>
<td>0.2</td>
<td>0.250</td>
<td>0.022</td>
<td>0.000</td>
</tr>
<tr>
<td>4.0</td>
<td>0.2</td>
<td>0.155</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\beta_2$</th>
<th>$r_{12}$</th>
<th>( u_{.05} = 18.51 )</th>
<th>( u_{.05} = 5.32 )</th>
<th>( u_{.05} = 4.41 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.2</td>
<td>0.017</td>
<td>0.015</td>
<td>0.014</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>0.067</td>
<td>0.049</td>
<td>0.033</td>
</tr>
<tr>
<td>1.0</td>
<td>0.2</td>
<td>0.142</td>
<td>0.028</td>
<td>0.001</td>
</tr>
<tr>
<td>2.0</td>
<td>0.2</td>
<td>0.162</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>4.0</td>
<td>0.2</td>
<td>0.035</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\beta_2$</th>
<th>$r_{12}$</th>
<th>( u = 1 )</th>
<th>( u = 1 )</th>
<th>( u = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.2</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>0.012</td>
<td>0.009</td>
<td>0.005</td>
</tr>
<tr>
<td>1.0</td>
<td>0.2</td>
<td>0.011</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>2.0</td>
<td>0.2</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>4.0</td>
<td>0.2</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
(iv) the bias increases and then decreases as we fix $r_{12}$, $n$, and the level of significance and increase $\beta_2$.

**Sequential Estimator Mean Squares**

Following an analogous procedure to the derivation of the sequential estimator bias, the expression of the sequential estimator mean square is obtained. The sequential estimator mean square for a general significance level is derived and subsequently the appropriate changes are made to calculate the corresponding mean squares for the "mean square", "usual," and "50 percent" sequential estimators. As for the bias, the mean square is dichotomized:

$$
\text{MSE}(B) = \text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{<\alpha}} \Pr(u_{<\alpha}) + \text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{>\alpha}} \Pr(u_{>\alpha}) )
$$

Since the variance and bias of the R.L.S.E. ($\hat{b}_1$) are $\frac{1}{n-1}$ and $r_{12} \beta_2$, and $S_1^2$ is independent of $\frac{S_2^2}{S_3} = U$, the $\text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{<\alpha}} \Pr(u_{<\alpha}) )$ is equal to the unconditional mean square error, i.e., $\text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{>\alpha}} \Pr(u_{>\alpha}) ) = \frac{1}{n-1} + r_{12}^2 \beta_2^2$.

Since $\hat{b}_1 = \hat{b}_{1/\alpha} - r_{12} \hat{b}_2$, the $\text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{>\alpha}} \Pr(u_{>\alpha}) )$ is equal to

$$
\text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{>\alpha}} \Pr(u_{>\alpha}) ) = \frac{1}{n-1} + r_{12}^2 \beta_2^2
$$

By denoting a conditional expectation by $E/$, we can write $\text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{>\alpha}} \Pr(u_{>\alpha}) )$ as

$$
E/[(\hat{b}_{1/\alpha} - \hat{b}_{1/\alpha}) - r_{12} \hat{b}_2]^2
$$

Since $E/(\hat{b}_{1/\alpha} - \hat{\beta}_1)^2 = \text{MSE}(\hat{b}_1)$ and using the independence between $\hat{b}_1$ and $\hat{b}_2$, the $\text{MSE}(\frac{\hat{b}_{1/\alpha}}{u_{>\alpha}} \Pr(u_{>\alpha}) )$ can be written as
\[
\text{MSE(}\hat{b}_{1/2} \mid u_{\geq u_{a}}) = \frac{1}{n-1} + r^2_{12} \rho^2 - \frac{2r^2_{12} \beta^2}{\Pr(u_{\geq u_{a}})} \sum_{i=0}^{\infty} \lambda^i - \lambda^i \frac{1}{i!} \frac{e^{-\lambda} \lambda^i}{x_{o}^{(n-3/2,3/2)+1}} + r^2_{12} E(\hat{b}_{2/2} \mid u_{\leq u_{a}})
\]

where we have substituted the previous result for the \(E(\hat{b}_{2/2} \mid u_{\geq u_{a}})\).

Following the same technique used by Bancroft (1944) to obtain the expectation of \(\hat{b}_{2}\) given \(u_{\geq u_{a}}\), we obtained the expectation of \(\hat{b}_{2}^2\) given \(u_{\geq u_{a}}\).

We wish the expectation of \(\hat{b}_{2}^2\) when \(\frac{\hat{s}_{2}^2}{\hat{s}_{3}^2} = u_{\geq u_{a}}\) or \(\hat{b}_{2}^2 / \hat{s}_{3}^2 \leq u_{a} c\)

since \(\hat{s}_{3}^2 = \hat{b}_{2}^2 / c\) where \(c = \frac{1}{(n-1)(1-r^2_{12})}\).

Since \(\hat{b}_{2}\) and \(\hat{s}_{3}^2\) are independent, their joint distribution is

\[
\begin{align*}
&-\frac{1}{2}(\hat{b}_{2} - \hat{\beta}_{2})^2 / c \quad \frac{1}{2}(n-3) \quad -\frac{1}{2}(n-3) V_3 \quad e^{V_3^2/2} \quad e^{\frac{1}{2}(n-3) V_3} \quad dV_3 \quad db_2, \\
&K = \frac{(n-3)^{n-3/2}}{\sqrt{2\pi c}} \quad \Gamma(n/2) \quad \text{and} \quad V_3 = \hat{s}_{3}^2.
\end{align*}
\]

Making the following transformation of variables

\[
z = \frac{V_3}{\hat{s}_{2}}, \quad \frac{1}{b_2^2} \quad dV_3 = \frac{\hat{b}_{2}^2}{b_2^2} \quad dq,
\]

the joint distribution of \(\hat{b}_{2}\) and \(z\) is

\[
-\frac{(\hat{b}_{2} - \hat{\beta}_{2})^2}{c} \quad \frac{1}{2}(n-5) \quad \frac{1}{2}(n-3) \quad \hat{b}_{2}^2 \quad \hat{s}_{2}^2 \quad d\hat{b}_{2} \quad dz \quad db_2
\]

Taking the expected value of \(\hat{b}_{2}^2\) when \(u_{\geq u_{a}}\), i.e., \(z \leq \frac{1}{u_{a} c}\), we have
\( \text{E}(\hat{b}_2^2) = \frac{K e^{-\frac{\beta^2}{2c}}}{Pr(z \leq \frac{1}{u_a c})} \int_{0}^{\infty} \frac{1}{u_a c} \hat{b}_2^{n-3} \frac{1}{z^\frac{1}{2}(n-5)} - \frac{\hat{b}_2^2[1+(n-3)cz]}{c} \int_{0}^{\infty} \frac{b_2^i \beta_2}{\Sigma i!} \sum_{i=0}^{\infty} \frac{dz}{db_2} \) 

where \( \Pr(z \leq \frac{1}{u_a c}) = \Pr(u^*_u \geq u_a) \), \(-\infty < b_2 < \infty\), and \( 0 < z < \frac{1}{u_a c} \).

The odd terms of the series will vanish whether \( n \) is odd or even when \( \hat{b}_2 \) is integrated out since for those terms we have an odd function on \( \hat{b}_2 \), i.e., \( f(-\hat{b}_2) = -f(\hat{b}_2) \).

After integrating out with respect to \( \hat{b}_2 \) for the even terms of the series using the properties of integrating an even function of the gamma variate-type, we have an infinite series whose typical term is of the form,

\[
\frac{K e^{-\frac{\beta^2}{2c}}}{\Pr(u^*_u \geq u_a)} \int_{0}^{\infty} \frac{1}{u_a c} \left( \frac{\beta^2}{2c} \right)^{n+2i} \frac{1}{2^{i+1}} \frac{1}{(2c)^i} \frac{1}{c^i} \frac{1}{z^\frac{1}{2}(n-5)} \left[ 1+(n-3)cz \right]^{n+2i} \frac{dz}{2} \]

By making the transformation of variable,

\[
x = \frac{c(n-3)z}{1+(n-3)cz} \text{, i.e., } 1 - x = \frac{1}{1+(n-3)cz} \text{,}
\]

\[
dz = \frac{1}{c(n-3)(1-x)^2} \ dx \text{,}
\]

and since \( z = \frac{1}{uc} \text{, } x_o = \frac{n-3}{(n-3)+u_a} \text{,} \)
the expectation of $b^2_2$ given $u^*_\alpha u_\alpha$ can be written as

$$
\frac{c}{\Pr(u^*_\alpha u_\alpha)} \sum_{i=0}^{\infty} \frac{(2i+1)}{(2i)!} \frac{\lambda^{i} e^{-\lambda}}{i!} I_{x_\alpha} \left(\frac{n-3}{2}, \frac{3}{2} + i\right),
$$

using the following identity,

$$
\frac{\Gamma\left(\frac{3}{2} + i\right)}{\Gamma\left(\frac{1}{2}\right) (2i)!} 2^{2i+1} = \frac{2i+1}{i!},
$$

and letting $\lambda = \frac{\beta_2^{2}}{2c}$. Since

$$
\text{MSE} \left( \hat{b}_1^{1/2} \right) \Pr(u^*_u u_\alpha) = \left(\frac{1}{n-1} + \frac{r^2_2 \beta_2^2}{12}\right) \Pr(u^*_u u_\alpha),
$$

and

$$
\text{MSE} \left( \hat{b}_1^{1/2} \right) \Pr(u^*_u u_\alpha) = \left(\frac{1}{n-1} + \frac{r^2_2 \beta_2^2}{12}\right) \Pr(u^*_u u_\alpha) - 2r^2_2 \beta_2^2 \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} I_{x_\alpha} \left(\frac{n-3}{2}, \frac{3}{2} + i\right)
$$

$$
+ r^2_2 c \sum_{i=0}^{\infty} \frac{(2i+1)}{(2i)!} \frac{\lambda^i e^{-\lambda}}{i!} I_{x_\alpha} \left(\frac{n-3}{2}, \frac{3}{2} + i\right),
$$

we finally obtain,

$$
\text{MSE}(B) = \frac{1}{n-1} + \frac{r^2_2 \beta_2^2}{12} - 2r^2_2 \beta_2^2 \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} I_{x_\alpha} \left(\frac{n-3}{2}, \frac{3}{2} + i\right)
$$

$$
+ r^2_2 c \sum_{i=0}^{\infty} \frac{(2i+1)}{(2i)!} \frac{\lambda^i e^{-\lambda}}{i!} I_{x_\alpha} \left(\frac{n-3}{2}, \frac{3}{2} + i\right).
$$

As a check for our algebra, we could let $u_\alpha$ be zero. This value of the critical point implies that $\hat{b}_1$ is always used. If $u_\alpha = 0$ then

$$
I_{x_\alpha} \left(\frac{n-3}{2}, \frac{3}{2} + i\right) = 1 \text{ for all } i \text{ since } x_\alpha = 1,
$$

and the mean square error of $B$ simplifies to
\[
\text{MSE}(B) = \frac{1}{n-1} + r_{12}^2 \beta_2^2 - 2r_{12}^2 \beta_2^2 \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} + 2r_{12}^2 c \sum_{i=0}^{\infty} \frac{i\lambda^i e^{-\lambda}}{i!} + r_{12}^2 c \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!}.
\]

But since \( \sum_{i=0}^{\infty} \frac{i\lambda^i e^{-\lambda}}{i!} = \lambda \), \( \sum_{i=0}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!} = 1 \), \( c\lambda = \frac{\beta_2^2}{2} \) the mean square of \( b_1^* \) simplifies to

\[
\frac{1}{n-1} + r_{12}^2 c = \frac{1}{(n-1)(1-r_{12}^2)}
\]

which is the mean square of \( \hat{b}_1 \).

If \( \hat{b}_1 \) is always used, this implies \( u_\alpha = \infty \), i.e.,

\[
I \left( \frac{n-3}{2}, \frac{3}{2} + 1 \right) = 0 \text{ for all } i \text{ since } \chi^2_0 = 0, \text{ and the mean square of } B
\]
simplifies to

\[
\frac{1}{n-1} + r_{12}^2 \beta_2^2
\]

which is the mean square of \( \hat{b}_1 \).

In this section the simplified model of Bancroft (1944) was used for the purpose of comparison, but we could use the same approach for a regression model with two regressors. The two basic modifications will be that the variance of \( \varepsilon \) will be \( \sigma^2 \) instead of 1 and \( c = \frac{1}{\sum x_i^2 (1-r_{12}^2)} \) instead of \( \frac{1}{(n-1)(1-r_{12}^2)} \). The orthogonalization process will be carried out with a new transformation matrix but the properties of independent sum of squares will remain invariant.

Substituting the correspondent critical points the mean square of the sequential estimators were calculated for the special cases under
consideration. The results are presented in Table 15 in which the single (double) asterisks mean that \( I_{X_0} \) was set equal to one for the terms \( I_{X_0} (\frac{n-3}{2}, \frac{3}{2} + i' + K), K = 1, 2, \ldots \), when \( I_{X_0} (\frac{n-3}{2}, \frac{3}{2} + i') > .99999, (.9999) \), since \( I_{X_0} \) is an increasing function of the second argument.

The results in Table 15 seem to indicate:

(i) The mean square decreases as we fix \( r_{12}, \beta_2 \), and the significance level and increase the sample size.

(ii) The mean square increases as we fix \( n, \beta_2 \), and the significance level and increase \( r_{12} \).

(iii) The mean square increases and then decreases as we fix \( r_{12}, n \), and the level of significance and increase \( \beta_2 \).

(iv) For \( \lambda < \frac{1}{2} \), the mean square gets smaller as we reduce the significance level and fix \( n, r_{12} \) and \( \beta_2 \). For \( \lambda > \frac{1}{2} \), the mean square gets smaller as we increase the significance level and fix \( n, r_{12} \) and \( \beta_2 \) (see Table 13 for the values of \( \lambda \)).

These are the following three exceptions for the behavior of the mean square for \( \lambda > \frac{1}{2} \). The mean square increases and then decreases as we increase the significance level for:

(i) Sample size 5 and \( \lambda = .720 \).

(ii) Sample size 11 and \( \lambda = .512 \).

(iii) Sample size 21 and \( \lambda = .576 \).
Table 15. The mean squares of the "mean square," "usual," and "50 percent" sequential estimators

<table>
<thead>
<tr>
<th></th>
<th>$n = 5$</th>
<th></th>
<th>$n = 11$</th>
<th></th>
<th>$n = 21$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$u_{0.05} = 37.38$</td>
<td></td>
<td>$u_{0.05} = 9.96$</td>
<td></td>
<td>$u_{0.05} = 8.18$</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>$r_{12}$</td>
<td></td>
<td></td>
<td>$r_{12}$</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>.251</td>
<td>.255</td>
<td>.264</td>
<td>.290</td>
<td>.101</td>
</tr>
<tr>
<td>0.4</td>
<td>.257</td>
<td>.279</td>
<td>.317</td>
<td>.385</td>
<td>.107</td>
</tr>
<tr>
<td>1.0</td>
<td>.287</td>
<td>.400</td>
<td>.596</td>
<td>.898</td>
<td>.121</td>
</tr>
<tr>
<td>2.0</td>
<td>.359</td>
<td>.708</td>
<td>1.369</td>
<td>2.488</td>
<td>.105</td>
</tr>
<tr>
<td>4.0</td>
<td>.389</td>
<td>.921</td>
<td>2.322</td>
<td>6.008</td>
<td>.104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$u_{0.05} = 18.51$</th>
<th></th>
<th>$u_{0.05} = 5.32$</th>
<th></th>
<th>$u_{0.05} = 4.41$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_2$</td>
<td>$r_{12}$</td>
<td></td>
<td></td>
<td>$r_{12}$</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>.252</td>
<td>.258</td>
<td>.273</td>
<td>.319</td>
<td>.101</td>
</tr>
<tr>
<td>0.4</td>
<td>.258</td>
<td>.281</td>
<td>.325</td>
<td>.413</td>
<td>.107</td>
</tr>
<tr>
<td>1.0</td>
<td>.284</td>
<td>.391</td>
<td>.583</td>
<td>.901</td>
<td>.112</td>
</tr>
<tr>
<td>2.0</td>
<td>.328</td>
<td>.590</td>
<td>1.142</td>
<td>2.228</td>
<td>.104</td>
</tr>
<tr>
<td>4.0</td>
<td>.290</td>
<td>.472</td>
<td>1.118</td>
<td>3.723</td>
<td>.104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$u = 1$</th>
<th></th>
<th>$u = 1$</th>
<th></th>
<th>$u = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_2$</td>
<td>$r_{12}$</td>
<td></td>
<td></td>
<td>$r_{12}$</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>.258</td>
<td>.289</td>
<td>.365</td>
<td>.611</td>
<td>.104</td>
</tr>
<tr>
<td>0.4</td>
<td>.260</td>
<td>.297*</td>
<td>.383</td>
<td>.646</td>
<td>.105</td>
</tr>
<tr>
<td>1.0</td>
<td>.263</td>
<td>.311</td>
<td>.428</td>
<td>.768</td>
<td>.104</td>
</tr>
<tr>
<td>2.0</td>
<td>.261</td>
<td>.300</td>
<td>.405</td>
<td>.802</td>
<td>.104</td>
</tr>
<tr>
<td>4.0</td>
<td>.260</td>
<td>.298</td>
<td>.391</td>
<td>.696</td>
<td>.104*</td>
</tr>
</tbody>
</table>
CHAPTER VII
SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Summary

It has been our purpose in this thesis to find the "best" estimate of $\beta_1$ between the R.L.S.E. ($b_1$) and the F.L.S.E. ($b_1^*$) in a linear model of full rank. In so doing we have defined "multicollinearity" by minimizing the mean square error of any linear combination of the estimators. This procedure provided us with a criterion to exclude regressors from the model in terms of the noncentrality parameter of the F distribution.

In Chapter I the linear model of full rank with $\epsilon_i \sim NID(0, \sigma^2)$ is presented. A partitioning of the regression coefficients vector is carried out in order to isolate $\beta_1$. The relationships among stepwise regression, specification error, multicollinearity, preliminary tests of significance and mean square error is discussed. Related literature was also cited in Chapter I.

The selection between the R.L.S.E. and the F.L.S.E. of $\beta_1$ in the two-regressor model via the mean square error is discussed in Chapter II. An explanation of "multicollinearity" and comments about the usual procedure of deleting variables by using the t-test are given. In Chapter II properties of the restricted and full estimators that are needed in the generalization of the mean square error criterion are presented.

In Chapter III a comparison of the R.L.S.E. ($b_1$) and F.L.S.E. ($b_1^*$) in terms of the variance of any arbitrary linear function is given. Also the criterion of "betterness" in terms of the mean square error of any arbitrary linear function of the estimators is derived.
The theoretical framework of the U.M.P. test for "multicollinearity" is discussed in Chapter IV. Also a critical points and level of significance comparison for the usual and the "multicollinearity" test is presented. The minimum variance unbiased estimator of the noncentrality parameter is derived in this chapter.

In Chapter V the procedure to obtain the "multicollinearity" critical points and power values is described. A table of critical points in terms of the noncentral beta and the noncentral F distributions is presented for one degree of freedom in the numerator and four levels of significance, with an explanation of its usage. "Multicollinearity" test power value tables for selected noncentrality parameters under four levels of significance are presented in this chapter. A comparison with standard F tables and Tang's power tables is given in order to emphasize the basic difference between the usual test and the "multicollinearity" test. In Appendices A and B the computer programs for the evaluation of critical points and power values are presented.

The bias and mean square error of a sequential estimator of $\beta_1$ in a special two-regressor model are presented in Chapter VI. The evaluation of bias and mean square error is carried out for selected sample sizes, $r_{12}$ and $\beta_2$. Also a comparison of bias and mean square error of the sequential estimators corresponding to three significance levels is provided in Chapter VI. Tables summarizing the results for bias and mean square error are presented in this chapter with the corresponding computer program for evaluating bias and mean square error appearing in Appendix C.
Conclusions

Using the criterion of minimizing the mean square error of any linear combination of the estimators of \( \beta_1 \) in a linear model of full rank, we were able to integrate the topics of stepwise regression, specification error, "multicollinearity" and preliminary tests of significance under a single framework. The decision of selecting the R.L.S.E. \( \hat{\beta}_{21} \) over the F.L.S.E. \( \hat{\beta}_{11} \) under the criterion of \( \text{MSE}(c' \hat{\beta}_{21}) \leq \text{MSE}(c' \hat{\beta}_{11}) \) for all nonnull \( c \) can be articulated as preferring the stepwise estimator, or analogously by model misspecification, or deletion of variables from the model because they are "multicollinear." Since the mean square error criterion of selection requires testing a hypothesis about the noncentrality parameter \( \lambda \) in the noncentral F distribution our estimator is a sequential estimator, i.e., the selection of the "best" estimator is based on a test of significance.

For the two-regressor model we were able to derive a precise but inconclusive statement of "multicollinearity" since we should delete \( x_2 \) whenever \( r_{12}^2 \) is bounded by \( [1 - \frac{\sigma^2}{\hat{\beta}_2^2 \Sigma x_2^2}, 1] \). Also the usual test for deleting \( x_2 \) is found to be more restrictive than required (under the MSE criterion) since \( \hat{\beta}_{12} \) should be used when \( |\beta_2| \leq \frac{\sqrt{V(b_2)}}{\hat{\beta}_2} \). Analogous statements in terms of \( R_k^2 \) and \( V(\hat{b}_{k+1}) \) are derived when we are interested in deciding to exclude only one regressor in the general partitioned linear model of full rank.

The biased restricted estimator is \( (X_1'X_1)^{-1}X_1Y \) with mean \( \hat{\beta}_1 + A\hat{\beta}_2 \), covariance matrix of \( \sigma^2(X_1'X_1)^{-1} \) and mean square error matrix \( \sigma^2(X_1'X_1)^{-1} + A\hat{\beta}_2\hat{\beta}_2' \). The full estimator is \( \hat{\beta}_{12} - A\hat{\beta}_2 \) with mean \( \hat{\beta}_{12} \).
covariance matrix and mean square error matrix \( \sigma^2 (X_1'X_1)^{-1} + \sigma^2 \mathbf{A} \mathbf{F}^{-1} \mathbf{A}' \) since it is unbiased.

Under the criterion of minimizing the variance of any arbitrary linear function, the restricted estimator is better than the full estimator. Consequently the restricted estimator is better in terms of the generalized variance ratio. Under the criterion of minimizing the mean square error of any arbitrary linear function, the restricted estimator of \( \hat{\beta}_1 \) is better than the full estimator if the noncentrality parameter is less or equal to one-half and the mean square error of a linear function criterion implies what can be called the generalized mean square ratio criterion. The condition that the noncentrality parameter is less than or equal to one-half is a sufficient and necessary condition for deleting \( X_2 \) from the true model if \( m \leq K \) and rank of \( \mathbf{A} \) is \( m \). This is not too optimistic to expect in most econometrics problems. When we are interested in deciding whether to exclude only one regressor the mean square error criterion is necessary and sufficient since the criterion simplifies to a scalar.

Since the family of noncentral F densities have the monotone likelihood ratio property in \( T(u) = \frac{\mu u}{q + \mu u} = w \) a U.M.P. test for "multicollinearity," i.e., \( \lambda \leq \frac{1}{2} \), is provided by using the noncentral F distribution or equivalently the noncentral beta distribution. The basic difference between the usual test and the "multicollinearity" test is the level of significance at which the corresponding hypotheses are tested. In Chapter V, "multicollinearity" test critical points for one degree of freedom in the numerator in terms of the beta and F distributions are presented. Also the evaluation of the power for selected values of the noncentrality parameter are given. From this
table we can readily see the relationship between the levels of
significance for the $H_0: \lambda = 0$ versus the $H_0: \lambda \leq \frac{1}{2}$. The minimum variance
unbiased estimator of the noncentrality parameter is $\frac{\psi m(q-2)}{2q} + \frac{m}{2}$ in
terms of the $F$ variable and are $\frac{w(q-2)}{2(1-w)} + \frac{m}{2}$ in terms of the beta variable.

In the regression model of two regressors used by Bancroft (1944)
the bias and mean square error of the sequential estimators of $\beta_1$ implied
by a usual test at 5 percent, a "multicollinearity" test at 5 percent, and
a "50 percent" test for selected sample sizes, $r_{12}$ and $\beta_2$ were evaluated.
The results seem to indicate that the bias decreases for increases in the
sample size and level of significance with corresponding *ceteris paribus.*
The bias increases as the "collinearity" increases, *ceteris paribus,* and
the bias increases and then decreases as $\beta_2$ increases, *ceteris paribus.*
For increases in sample size, "collinearity" and $\beta_2$ with corresponding
*ceteris paribus* the behavior of the mean square error of the sequential
estimator for $\beta_1$ is similar to the behavior of the bias. For increases
in the level of significance, *ceteris paribus,* the behavior of the mean
square error depends on the value of the noncentrality parameter, i.e.,
if $\lambda < \frac{1}{2}$ the mean square increases and for $\lambda > \frac{1}{2}$ the mean square
decreases, eventually, attaining maximums for noncentrality parameters
close to one-half.

**Suggestions for Further Research**

In this thesis we have used a very special restriction on the
parameter space. Since our interest was to decide between $\hat{\beta}_1$ and $\hat{\beta}_1$
the implicit restriction was $\beta_2 = 0$. An investigation about the
properties of the R.L.S.E. and F.L.S.E. under the mean square error
criterion for a more general set of restrictions seems natural.
Analogously to experimental designs where an optimum design matrix is of great interest, a search for optimal restrictions should be carried out especially for the case where only one linear combination of the regression coefficients is studied.

The properties of the residual estimator of $\hat{\beta}_1$ in terms of the mean square error criterion should be analyzed. The residual estimator of $\hat{\beta}_1$ is obtained by regressing $Y$ on $X_2$ and then regressing the residual on $X_1$. Following the mean square error criterion a "best" estimator for $\beta_1$ should be investigated without restricting ourselves to least squares estimators. Also it would be interesting to apply the mean square error criterion in estimating linear combinations of regression coefficients in error in variables models as well as in simultaneous equation models where a linear combination of the structural parameters is of interest.

A Bayesian approach for estimating the noncentrality parameter and an investigation of the distributional properties of the minimum variance unbiased estimator of the noncentrality parameter seem relevant.

"Multicollinearity" critical points tables should be provided for more degrees of freedom in the numerator. Due to the great sensitivity of critical points evaluation, greater precision should be required in their computation.

The empirical results obtained on the mean square error of sequential estimators for $\beta_1$ in the two-regressor model raise the possibility of an optimum choice of $\alpha$ for a given objective function.
LIST OF REFERENCES


LIST OF REFERENCES (continued)


Appendix A. Computer Program for the "Multicollinearity" Test Critical Points

MESSAGE

FORTRAN LISTING 1410-FO-970

DIMENSIONA (19), B(19), BETA(61), U(4), V(4), W(4)
00100 FORMAT(6X, 4F9.2, 4X, 4F9.2)
00200 FORMAT(32HSSUBROUTINE ERROR, MACHINE STOPPED)
00300 FORMAT(F6.0, 4F9.4, 8X, 4F9.4)
00400 FORMAT(I3)
00500 FORMAT(1H1, 7HF1 IS 1)
00600 FORMAT(3I4, F6.4)

M = 1
WRITE(3, 500)
EPS = .5E-6
GM = M
V(1) = .05
V(2) = .1
V(3) = .25
V(4) = .5
WRITE(3, 100) V(1), V(2), V(3), V(4), V(1), V(2), V(3), V(4)
00025 READ(1, 400) N
GN = N
Q = .5*GN
AZERO = EXP(-.5)
D010 J = 1, 4
SLIMIT = 1 - V(J)
XA = 1.
Y = 0.
X = .5

00004 P = .5*GM
CALLDRIVER(X, P, Q, EPS, NMAX, BETA, KEY)
IF(KEY) 14, 15, 14
00014 WRITE(3, 200)
STOP
00015 BZERO = BETA(NMAX+1)
A(1) = .5
S = AZERO*BZERO
P = P + 1.
CALLDRIVER(X, P, Q, EPS, NMAX, BETA, KEY)
IF(KEY) 16, 17, 16
00016 WRITE(3, 200)
STOP
00017 B(1) = BETA(NMAX+1)
S = S + AZERO*A(1)*B(1)
D081 = 2, 7
DI = 1
A(I) = A(I-1) * .5 / DI
P = P + 1.
CALLDRIVER(X, P, Q, EPS, NMAX, BETA, KEY)
IF(KEY) 6, 7, 6
00006 WRITE(3, 200)
STOP
MESSAGE

FORTRAN LISTING

1410-FO-970

00007  B(I)=BETA(NMAX+1)
00008  S=S+A(I)*B(I)*AZERO
       P=S-SLIMIT
       IF(ABS(P)-.0005)2,2,5
00005  IF(P)1,2,3
00001  Y=X
       X=(XA+X)/2.
       GOT04
00003  XA=X
       X=(Y+X)/2.
       GOT04
00002  W(J)=X
       WRITE(2,600)M,N,J,X
00010  U(J)=GN*X/((1.-X)*GM)
       WRITE(3,300)GN,W(1),W(2),W(3),W(4),U(1),U(2),U(3),U(4)
       GOT025

STOP
END
Appendix B. Computer Program for the Evaluation of Power Values

MESSAGE FORTRAN LISTING 1410-PO-970
DIMENSIONA (27),B(27),C(9),BETA(61),ID(9)
00100 FORMAT(33H F1 F2 J LAMBDAPower)
00200 FORMAT(32HSSUBROUTINEERROR,MACHINESTOPPED)
00300 FORMAT(315,2F10.4)
00500 FORMAT(F7.4,12)
00600 FORMAT(314,F6.4)
EPS=.5E-6
WRITE(3,100)
DO2I=1,9
00002 READ(1,500)C(I),ID(I)
00001 READ(1,600)M,N,K,X
DO13J=1,9
ALAMBA=C(J)
JJ=ID(J)
GM=M
GN=N
Q=.5*GN
P=.5*GM
AZERO=EXP(-ALAMBA)
calla driver(x,p,q,eps,nmax,beta,key)
if(key)18,19,18
00018 WRITE(3,200)
STOP
00019 BZERO=BETA(NMAX+1)
S=AZERO*BZERO
P=P+1.
A(1)=ALAMBA
calla driver(x,p,q,eps,nmax,beta,key)
if(key)20,21,20
00020 WRITE(3,200)
STOP
00021 B(1)=BETA(NMAX+1)
S=S+AZERO*B(1)
DO12III=2,JJ
DI=III
A(III)=A(III-1)*ALAMBA/DI
P=P+1.
calla driver(x,p,q,eps,nmax,beta,key)
if(key)22,23,22
00022 WRITE(3,200)
STOP
00023 S=S+AZERO*A(III)*B(III)
T=1.-S
WRITE(3,300)M,N,K,ALAMBA,T
00013 CONTINUE
GOTO1
STOP
END
Appendix C. Fortran Statement for the Evaluation of Sequential Estimators Bias and Mean Squares

MESSAGE

FORTRAN LISTING

DIMENSION BETA(220), A(216), B(216), U(4), V(4)

00100 FORMAT(12, F2.1, F4.4, 213, F2.2)

00200 FORMAT(32H SUBROUTINE ERROR, MACHINE STOPPED)

00300 FORMAT(40H1 BIAS MEAN SQUARE ERROR)

00400 FORMAT(22H .2 .4 .6 .8)

00700 FORMAT(F4.1, F8.4, 18X, F9.4)

00800 FORMAT(F4.1, 6X, F8.4, 18X, F9.4)

00900 FORMAT(F4.1, 12X, F8.4, 18X, F9.4)

00950 FORMAT(F4.1, 18X, F8.4, 18X, F9.4)

EPS = .5E-6

V(1) = .2

V(2) = .4

V(3) = .6

V(4) = .8

WRITE(3, 300)

WRITE(3, 400)

0013 READ(1, 100) N, BE, X, IA, IO, RHO

S = 0.

T = 0.

AN = N

P = (AN - 3.)/2.

Q = 1.5

K = RHO*5.

CA = RHO*BE

CB = 1. / (AN - 1.)

CC = CA*CA

CD = -2. * CC

CE = RHO*RHO / ((AN - 1.)*(1. - RHO*RHO))

ALPHA = (1. - RHO*RHO)*((AN - 1.)*BE*BE/2.

A(1) = ALPHA

AZERO = EXP(-ALPHA)

IF(IA-3) 3, 3, 2

00003 CALLDIVER(X, P, Q, EPS, NMAX, BETA, KEY)

IF(KEY) 4, 5, 4

00004 WRITE(3, 200)

STOP

00005 BZERO = BETA(NMAX+1)

Q = Q+1.

CALLDIVER(X, P, Q, EPS, NMAX, BETA, KEY)

IF(KEY) 6, 7, 6

00006 WRITE(3, 200)

STOP

00007 B(1) = BETA(NMAX+1)

S = AZERO*(BZERO + A(1)*B(1))

T = AZERO*A(1)*B(1)

IA = 2

GOTO 8

00002 IAM = IA-1
MESSAGE  FORTRAN LISTING

DO91 = 2, IAM
DI = I

00009  A(I) = A(I-1) * ALPHA / DI
00008  AI = IA
        Q = 1.5 + AI
DO10I = 1A, I0
        DI = I
        A(I) = A(I-1) * ALPHA / DI
CALLDRIVER(X, P, Q, EPS, NMAX, BETA, KEY)
        IF (KEY) 11, 12, 11

00011  WRITE(3, 200)
        STOP

00012  B(I) = BETA(NMAX+1)
        Z = A(I) * B(I) * AZERO
        Q = Q + 1.
        S = S + Z

00010  T = T + Z * DI
        BIAS = CA *(1. - S)
        SSQE = CB + CC + (CD + CE) * S + 2. * CE * T
GOTO(15, 16, 17, 18), K

00015  WRITE(3, 700) BE, BIAS, SSQE
GOTO13

00016  WRITE(3, 800) BE, BIAS, SSQE
GOTO13

00017  WRITE(3, 900) BE, BIAS, SSQE
GOTO13

00018  WRITE(3, 950) BE, BIAS, SSQE
GOTO13

00014  CONTINUE
        STOP
        END