Ceylon Journal of Science: Physical Sciences, 2(1), 57-64 (1995)

PRE-TEST ESTIMATION IN THE LINEAR REGRESSION MODEL UNDER STOCHASTIC RESTRICTIONS

P. WIJEKOON
Department of Mathematics, University of Peradeniya, Peradeniya, Sri Lanka

and

G. TRENKLER
FB Statistik, University of Dortmund, 44221, Dortmund, Germany.

ABSTRACT

In this paper, a general method to find stochastic properties of various stochastic pre-test estimators and forecasts is developed. This method is applicable if the covariance matrix Ω of the prior information is nonsingular.

1. INTRODUCTION

Consider the general linear model

$$Y = X\beta + \varepsilon; \quad \varepsilon \sim N(0, \sigma^2 V), \tag{1.1}$$

where **Y** is an $(n \times 1)$ vector of observations on the dependent variable Y, **X** is an $n \times p$ matrix of explanatory variables with full column rank p, β is a vector of p parameters, and the n-dimensional error term ε is multinormally distributed with a mean of zero and a covariance matrix of $\sigma^2 V$. Also assume that **V** is a known positive definite (p.d.) matrix, and that both β and σ^2 are unknown. If no further information than that contained in the observation model (1.1) is available, the generalized least squares estimator (GLSE)

$$\hat{\beta} = \left(X'V^{-1}X\right)^{-1} X'V^{-1}Y \tag{1.2}$$

is the best linear unbiased estimator of the parameter vector β.

When different estimators are available for the same parameter vector β in the linear regression model one must solve the problem of their comparison. Usually as a simultaneous measure of covariance and bias, the mean square error matrix is used, and is defined by

$$M(\hat{\beta},\beta) = E\left[(\hat{\beta}-\beta)(\hat{\beta}-\beta)'\right] = D(\hat{\beta}) + B(\hat{\beta})B(\hat{\beta})', \tag{1.3}$$

where $D(\hat{\beta})$ is the dispersion matrix and $B(\hat{\beta}) = E(\hat{\beta}) - \beta$ denotes the bias vector.

Definition 1 (MSE- Matrix Superiority of Estimators)

Let two alternative estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ of β be given. Then $\hat{\beta}_2$ is said to be superior to $\hat{\beta}_1$ with respect to the MSE- matrix criterion if and only if

$$M(\hat{\beta}_1,\beta) - M(\hat{\beta}_2,\beta) \geq 0 \tag{1.4}$$

in the Löwner ordering.

2. ESTIMATION UNDER LINEAR STOCHASTIC RESTRICTIONS

In addition to the sample model (1.1), suppose there is additional information available for the parameter vector β in form of $m < p$ independent linear stochastic restrictions

$$\mathbf{r} = \mathbf{R}\beta + \nu \, ; \; \nu \sim N(\mathbf{0}, \sigma^2\Omega), \; E(\varepsilon\nu') = 0 \tag{2.1}$$

with rank $(\mathbf{R}) = m$, where $\mathbf{r}$ and $\mathbf{R}$ are a known vector and matrix, respectively, and Ω is assumed to be positive definite, and is also a known matrix. But the use of stochastic restrictions allows to describe situations which are more flexible than those underlying exact restrictions. Note that if $\nu = 0$, then the restrictions become exact.

By combining the sample model (1.1) with the above specific information (2.1) Theil and Goldberger (1961) developed the mixed estimator

$$\hat{\beta}_M = (\mathbf{S} + \mathbf{R}'\Omega^{-1}\mathbf{R})^{-1}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} + \mathbf{R}'\Omega^{-1}\mathbf{r}) \, , \tag{2.2}$$

which is unbiased for β, where $\mathbf{S} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$.

According to Toutenburg (1975) the mixed estimator $\hat{\beta}_M$ can also be written as

$$\hat{\beta}_M = \hat{\beta} + \mathbf{S}^{-1}\mathbf{R}'\left(\Omega + \mathbf{R}\mathbf{S}^{-1}\mathbf{R}'\right)^{-1}\left(\mathbf{r} - \mathbf{R}\hat{\beta}\right) \tag{2.3 a}$$

$$= \hat{\beta} + \mathrm{H}\hat{\delta} \, , \tag{2.3 b}$$

where $\mathbf{H} = \mathbf{S}^{-1}\mathbf{R}'\left(\Omega + \mathbf{R}\mathbf{S}^{-1}\mathbf{R}'\right)^{-1}$ and $\hat{\delta} = \mathbf{r} - \mathbf{R}\hat{\beta}$,

and in this case the matrix Ω is not necessarily regular. This may happen when exact and stochastic restrictions are used simultaneously. Henceforth we assume that Ω is nonsingular.

In statistical practise, the assumption that the prior information is unbiased; ie. $E(\mathbf{r}) = \mathbf{R}\beta$ is often violated.Then the true but unknown prior information can be formulated as

$$\mathbf{r} = \mathbf{R}\beta + \delta + \phi \, ; \; \phi \sim N(\mathbf{0}, \sigma^2\Omega) \, , \tag{2.4}$$

where δ is an unknown fixed $(m \times 1)$ vector.

Teräsvirta (1980) and Hill and Ziemer (1983) have given some interesting examples for the above type of prior information.

Note that when $\delta \neq \mathbf{0}$ in (2.4) then the mixed estimator $\hat{\beta}_M$ becomes biased with

$$\mathbf{B}\left(\hat{\beta}_M\right) = \mathbf{H}\delta \, , \tag{2.5}$$

and its dispersion matrix is

$$\mathbf{D}\left(\hat{\beta}_M\right) = \sigma^2\left(\mathbf{S} + \mathbf{R}'\Omega^{-1}\mathbf{R}\right)^{-1} \tag{2.6}$$

3. STATISTICAL IMPLICATIONS

Let us now turn to the question of the statistical evaluation of the compatibility of sample and stochastic prior information. The classical procedure is to test the hypothesis

$$H_0 : \delta = 0 \quad \text{against} \quad H_1 : \delta \neq 0 \tag{3.1}$$

under linear model (1.1) and stochastic prior information (2.1).
(cf: Judge and Bock, 1978, pp 121-124) .

The decision rule of this testing problem is

$$\text{Accept } H_0; \text{ if } u \leq c, \tag{3.2}$$

where
$$u = \frac{(\mathbf{r}-\mathbf{R}\hat{\beta})'(\Omega+\mathbf{RS}^{-1}\mathbf{R}')^{-1}(\mathbf{r}-\mathbf{R}\hat{\beta})}{m\,\hat{\sigma}^2} \tag{3.3}$$

has a noncentral **F** (m, n-p, λ) distribution under $H_1 : \delta \neq 0$, with noncentrality parameter

$$\lambda = \frac{\delta'(\Omega+\mathbf{RS}^{-1}\mathbf{R}')^{-1}\delta}{2\sigma^2} \tag{3.4}$$

and
$$\hat{\sigma}^2 = \frac{(\mathbf{Y}-\mathbf{X}\hat{\beta})'(\mathbf{Y}-\mathbf{X}\hat{\beta})}{n-p}, \tag{3.5}$$

and c is the $(1\text{-}\alpha)$- quantile of the $F(m, n\text{-}p)$ distribution.

In view of the MSE- matrix criterion we can compare the two estimators $\hat{\beta}$ and $\hat{\beta}_M$
Note that

$$\mathbf{M}(\hat{\beta},\beta)-\mathbf{M}(\hat{\beta}_M,\beta) = \sigma^2\,\mathbf{G} - \mathrm{H}\delta\delta'\mathrm{H}', \tag{3.6}$$

where $\mathbf{G} = \mathbf{S}^{-1}\mathbf{R}'\left(\Omega + \mathbf{RS}^{-1}\mathbf{R}'\right)^{-1}\mathbf{RS}^{-1}$.
Since the matrix **G** is n.n.d., and

$$B(\hat{\beta}_M) = \mathbf{H}\delta = \mathbf{GSR}^{+}\delta \tag{3.7}$$

we conclude

$$B(\hat{\beta}_M) \in \Re(\sigma^2\mathbf{G}),$$

where $\mathbf{R}^{+}$ is the Moore-Penrose inverse of **R** and $\Re(.)$ denotes the column space of a matrix.

Using lemma 1 in the appendix we can prove that $\hat{\beta}_M$ is superior to $\hat{\beta}$ with respect to the MSE-matrix criterion if and only if

$$\lambda = \frac{\delta'(\Omega+\mathbf{RS}^{-1}\mathbf{R}')^{-1}\delta}{2\sigma^2} \leq \frac{1}{2}. \tag{3.8}$$

Although the classical hypothesis can be used to test the above condition, it does not answer the question how far the biased restriction combined with the sample information improves the estimation.Therefore we may use the following most powerful test for the MSE-matrix dominance of $\hat{\beta}_M$ over $\hat{\beta}$.

$$\text{Accept } H_0: \lambda \leq 1/2 \text{ if and only if } u < c\,(1/2) \tag{3.9}$$

where $c\,(1/2)$ is the $(1\text{-}\alpha)$- quantile of the F (m, n-p, 1/2) distribution.

Note that under the null hypothesis $H_0: \lambda \leq 1/2$, the test statistic u follows a noncentral F (m, n-p, λ) distribution with λ as the noncentrality parameter.

4. PRE-TEST ESTIMATOR AND ITS STOCHASTIC PROPERTIES

The result of the testing problem (3.9) can be used to construct the so-called ordinary stochastic pre-test estimator (OSPE)

$$\beta^* = I_{[0,c)}(u)\hat{\beta}_M + I_{[c,\infty)}(u)\hat{\beta} , \tag{4.1}$$

where $c = c\,(1/2)$, and $I_{[0,c)}(.)$ and $I_{[c,\infty)}(.)$ are indicator functions with values one or zero depending on whether u falls in the subscribed interval or not . The estimator β^* will choose the mixed estimator $\hat{\beta}_M$ if H_0 is accepted and $\hat{\beta}$ otherwise.

The stochastic properties of the OSPE have been first investigated by Judge, Yancey and Bock (1973). Their method depended on partitioned matrices, and therefore problems may arise when comparing the MSE-matrices. Hence we develop a new method which can be used generally for all pre-test estimators depending on the mixed estimators. Note that using (2.3b), OSPE β^* (4.1) can be written as

$$\beta^* = \hat{\beta} + I_{[0,c)}(u)\mathbf{H}\,\hat{\delta} \ . \tag{4.2}$$

This form of β^* also implies that using the test statistic u one can test the bias-term of the mixed estimator $\hat{\beta}_M$ .Now to find the stochastic properties of β^* we use theorem 1 in Appendix which can be used generally for most pre-test estimators.

First rewrite the numerator of the test statistic u as

$$\frac{\hat{\delta}'\mathbf{A}^{-1}\hat{\delta}}{\sigma^2} = \frac{1}{\sigma^2}(\varepsilon',\phi')\,\mathbf{D}\begin{pmatrix}\varepsilon\\ \phi\end{pmatrix} \tag{4.3}$$

where $\mathbf{A} = \mathbf{RS}^{-1}\mathbf{R}' + \Omega$, $\phi = \Omega^{-1/2}(\delta+v)$ and

$$\mathbf{D} = \begin{pmatrix}-\mathbf{XS}^{-1}\mathbf{R}'\\ \Omega^{1/2}\end{pmatrix}\mathbf{A}^{-1}\left(-\mathbf{RS}^{-1}\mathbf{X}' , \Omega^{1/2}\right). \tag{4.4}$$

Note that $\mathbf{D}$ is a symmetric and idempotent matrix of rank m. Then there exists an orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{QDQ}' = \begin{pmatrix}\mathbf{I}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}. \tag{4.5}$$

Now define a random variable $\mathbf{W}$ as

$$\mathbf{W} = \frac{\mathbf{Q}}{\sigma}\begin{pmatrix}\varepsilon\\ \phi\end{pmatrix} \sim N\left(\mathbf{Q}\begin{pmatrix}0\\ a\end{pmatrix}, \mathbf{I}\right), \tag{4.6}$$

where $a = \Omega^{-1/2}\delta/\sigma$. Using (4.6) we can obtain

$$\mathbf{H}\,\hat{\delta} = \sigma\left(-\mathbf{RS}^{-1}\mathbf{X}' , \Omega^{1/2}\right)\mathbf{Q}'\mathbf{W} \ . \tag{4.7}$$

Also $u = \dfrac{\mathbf{W}'\mathbf{QDQ}'\mathbf{W}}{\gamma} = \mathbf{W}'\begin{pmatrix}\mathbf{I}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{W}/\gamma$

$$u = \frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma} \sim F(m, n-p, \lambda) ,$$

where $\gamma = m\hat{\sigma}^2/\sigma^2$

Now applying theorem 1 in Appendix we will derive the following results.

Theorem

Let the random variables $\varepsilon \sim N(0\,,\,\sigma^2 \mathbf{I}_n)$ and $\nu \sim N(0\,,\,\sigma^2\Omega)$. Then

(i) $E\left[\mathbf{I}_{[0,c)}(u)\mathbf{H}\hat{\delta}\right] = p_2\mathbf{H}\delta$

(ii) $E\left[\mathbf{I}_{[0,c)}(u)\mathbf{H}\,\hat{\delta}\,\hat{\delta}'\,\mathbf{H}'\right] = \sigma^2 p_2\mathbf{G} + p_4\mathbf{H}\delta\delta'\mathbf{H}'$

(iii) $E\left[\mathbf{I}_{[0,c)}(u)(\hat{\beta}-\beta)\hat{\delta}'\mathbf{H}'\right] = -\sigma^2 p_2\mathbf{G} - (p_4-p_2)\,\mathbf{H}\,\delta\,\delta'\,\mathbf{H}'$.

Therefore $E\left[\beta^*\right] = E\left[\hat{\beta} + \mathbf{I}_{[0,c)}(u)\mathbf{H}\hat{\delta}\right]$

$$E\left[\hat{\beta}\right] = \beta + p_2\,\mathbf{H}\,\delta\,. \tag{4.9}$$

Also $\mathbf{M}(\beta^*,\beta) = E\left[(\beta^*-\beta)(\beta^*-\beta)'\right]$

$$\mathbf{M}(\beta^*,\beta) = E\left[\left(\hat{\beta}-\beta+\mathbf{I}_{[0,c)}(u)\mathbf{H}\hat{\delta}\right)\left(\hat{\beta}-\beta+\mathbf{I}_{[0,c)}(u)\mathbf{H}\hat{\delta}\right)'\right]$$

$$= E(\hat{\beta}-\beta)(\hat{\beta}-\beta)' + 2E\left[\mathbf{I}_{[0,c)}(u)(\hat{\beta}-\beta)\,\hat{\delta}'\mathbf{H}'\right] + E\left[\mathbf{I}_{[0,c)}(u)\mathbf{H}\delta\hat{\delta}'\mathbf{H}'\right]$$

Now applying the above theorem, we obtain

$$\mathbf{M}(\beta^*,\beta) = \sigma^2\mathbf{S}^{-1} - \sigma^2 p_2\,\mathbf{G} + (2p_2-p_4)\mathbf{H}\delta\delta'\mathbf{H}', \tag{4.10}$$

which is similar to the expression derived by Judge and Bock (1978, Chapter 6).

5. CONCLUDING REMARKS

This method is applicable for all pre-test estimators in which the disturbance vector Ω of the prior information is nonsingular. To apply this method one has to rewrite the pre-test estimator in the form

$$\phi^* = \hat{\phi} + \mathrm{I}_{[0,c)}(T)\;\mathbf{M}\,\hat{\alpha}$$

where ϕ^* , $\hat{\phi}, c, T, \mathbf{M}$ and $\hat{\alpha}$ stand for the following terms in the pre-test estimator.

- ϕ^* - a biased estimator
- $\hat{\phi}$ - an unbiased estimator
- c - $(1 - \alpha)$ - quantile of the corresponding F- distribution
- T - test statistic
- $\mathbf{M}$ - a nonnegative definite matrix which depends on the pre-test estimator
- $\hat{\alpha}$ - the term which represents the incorrectness of the prior information

After deriving the stochastic properties of pre-test estimators it is an easy task to compare their mean square error matrices. This can be done by using lemma 1 in Appendix and by the tools derived in Trenkler (1985).

6. APPENDIX

Lemma 1 (Baksalary and Kala, 1983)

Let **A** be an $n \times n$ symmetric matrix, $\boldsymbol{a}$ an $n \times 1$ vector and $\gamma > 0$ a real number. Then the following conditions are equivalent:

(i) $\gamma A - \boldsymbol{aa}' \ddagger \mathbf{0}$

(ii) $\mathbf{A} \geq \mathbf{0}$, $\boldsymbol{a} \in \Re(\mathbf{A})$, $\boldsymbol{a}'\mathbf{A}^{-}\boldsymbol{a} \leq \gamma$,

where $\mathbf{A}^{-}$ is any g-inverse of **A**.

Lemma 2 (Trenkler and Toutenburg, 1989)

Let ω and ξ be independent random variables such that $\omega \sim N(\theta, I_m)$ and $\xi \sim \chi^2_q / d$ for some positive scalar d. Then

(i) $E\left[\mathbf{I}_{[0,c)}\left(\frac{\omega'\omega}{\xi}\right)\omega\right] = \theta h_\lambda(2)$

(ii) $E\left[\mathbf{I}_{[0,c)}\left(\frac{\omega'\omega}{\xi}\right)\omega\omega'\right] = h_\lambda(2)\mathbf{I}_m + \theta\theta' h_\lambda(4)$,

where $\lambda = \frac{\theta'\theta}{2}$ and $h_\lambda(l) = P\left[\frac{\chi^2(m+l,\lambda)}{\chi^2(q)} \leq \frac{c}{d}\right]$ for $l \in N$.

Theorem 1

Let $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2]$ be a random variable with distribution

$$\mathbf{W} \sim N\left(\mathbf{Q}\begin{pmatrix} 0 \\ a \end{pmatrix}, \mathbf{I}\right), \qquad \text{(A-1)}$$

where $\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix}$ is an orthogonal matrix such that $\mathbf{QDQ}' = \begin{pmatrix} \mathbf{I}_h & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, and **D** is a symmetric and idempotent matrix of rank h.

Moreover the test statistic

$$u = \frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma} \sim F(h, n-p, \lambda), \qquad \text{(A-2)}$$

where $\gamma = \hat{\sigma}^2 h / \sigma^2$ Then

$$\text{(i)} \quad \mathbf{Q}' \mathrm{E}\left[\mathrm{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{W}\right] = p_2\begin{pmatrix} 0 \\ a \end{pmatrix} \qquad \text{(A-3)}$$

(ii) $$\mathbf{Q'E}\left[\mathbf{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{WW'}\right]\mathbf{Q} = p_2\mathbf{D} + (p_4 - p_2)\,\mathbf{D}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & aa'\end{pmatrix}\mathbf{D} + p_2\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & aa'\end{pmatrix},$$ (A-4)

where $$p_l = P\left[\chi^{2}(h+l,\lambda) \Big/ \chi^{2}(n-p) \le \frac{ch}{n-p}\right]$$ for $l \in N$.

(A-5)

$I_{[0,c)}(.)$ is an indicator function, and the critical value c is the $(1-\alpha)$ quantile of the noncentral F-distribution with parameters h, n-p and λ.

Proof

(i) Following lemma 2 of the appendix we obtain

$$\mathbf{E}\left[\mathbf{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{W}_1\right] = p_2\,\mathbf{E}(\mathbf{W}_1)\,. \tag{A-6}$$

Generally the indicator function $\mathbf{I}_{[0,c)}(\omega^2)$ has a $\chi^2_{(3,\,\sigma^2/2)}$ distribution, where ω^2 is a scalar (see Judge and Bock 1978, p. 321)

Therefore

$$\mathrm{E}\left[\mathrm{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\right] = p_2\,. \tag{A–7}$$

Since $\mathbf{W}_1 \sim N\left(\mathrm{E}(\mathbf{W}_1),\ \mathbf{I}_h\right)$ and $\mathbf{W}_2 \sim N\left(\mathrm{E}(\mathbf{W}_2),\ \mathbf{I}_{p+m-h}\right)$

are independent, we have

$$\mathrm{E}\left[\mathrm{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{W}_2\right] = p_2\,\mathrm{E}(\mathbf{W}_2) \tag{A-8}$$

and that implies

$$\begin{aligned}\mathbf{Q}'\,\mathrm{E}\left[\mathrm{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{W}\right] &= \mathbf{Q}'\,p_2\,\mathrm{E}(\mathbf{W})\\ &= p_2\,\mathbf{Q'Q}\begin{pmatrix}0\\ a\end{pmatrix} = p_2\begin{pmatrix}0\\ a\end{pmatrix}.\end{aligned} \tag{A-9}$$

(ii) Considering the independence of $\mathbf{W}_1$ and $\mathbf{W}_2$, and consulting lemma 2 of the appendix we derive

$$\mathrm{E}\left[\mathrm{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{WW'}\right] = \begin{pmatrix}p_2\mathbf{I} + p_4\mathbf{Q}_{12}aa'\mathbf{Q}_{12}' & p_2\mathbf{Q}_{12}aa'\mathbf{Q}_{22}'\\ p_2\mathbf{Q}_{22}aa'\mathbf{Q}_{12}' & p_2\mathbf{I} + p_2\mathbf{Q}_{22}aa'\mathbf{Q}_{22}'\end{pmatrix}. \tag{A-10}$$

Since $\mathbf{Q}$ is an orthogonal matrix, we obtain

$$\mathbf{Q'E}\left[\mathbf{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{WW'}\right]\mathbf{Q} = \begin{pmatrix}\mathbf{S}_1 & \mathbf{S}_2\\ \mathbf{S}_3 & \mathbf{S}_4\end{pmatrix} = \mathbf{Q}'\begin{pmatrix}\mathbf{K} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{Q} + p_2\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & aa'\end{pmatrix}, \tag{A-11}$$

where

$$\mathbf{S}_1 = p_2 \mathbf{I} + (p_4 - p_2)\mathbf{Q}'_{11}\mathbf{Q}_{12}\, aa'\, \mathbf{Q}'_{12}\mathbf{Q}_{11} \tag{A-12a}$$
$$\mathbf{S}_2 = (p_4 - p_2)\mathbf{Q}'_{11}\mathbf{Q}_{12}\, aa'\, \mathbf{Q}'_{12}\mathbf{Q}_{12} \tag{A-12b}$$
$$\mathbf{S}_3 = (p_4 - p_2)\mathbf{Q}'_{21}\mathbf{Q}_{12}\, aa'\, \mathbf{Q}'_{12}\mathbf{Q}_{11} \tag{A-12c}$$
$$\mathbf{S}_4 = p_2 \mathbf{I} + (p_4 - p_2)\mathbf{Q}'_{21}\mathbf{Q}_{12}\, aa'\, \mathbf{Q}'_{12}\mathbf{Q}_{12} + p_2 aa', \tag{A-12d}$$

and

$$\mathbf{K} = p_2 \mathbf{I} + (p_4 - p_2)\mathbf{Q}_{12}\, aa'\, \mathbf{Q}'_{12}\ . \tag{A-12e}$$

Rewriting $E(\mathbf{W}) = \mu = (\mu'_1, \mu'_2)'$, we get

$$\begin{aligned}
\mathbf{Q}'\begin{pmatrix}\mathbf{K} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{Q} &= p_2\mathbf{Q}'\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{Q} + (p_4 - p_2)\mathbf{Q}'\begin{pmatrix}\mu_1\mu_1' & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{Q}\\
&= p_2\mathbf{D} + (p_4 - p_2)\mathbf{Q}'\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mu\mu'\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{Q}\\
&= p_2\mathbf{D} + (p_4 - p_2)\mathbf{D}\mathbf{Q}'\,\mu\mu'\,\mathbf{Q}\mathbf{D}\\
&= p_2\mathbf{D} + (p_4 - p_2)\mathbf{D}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & aa'\end{pmatrix}\mathbf{D}\ .
\end{aligned} \tag{A-13}$$

Consequently applying (A-13) to (A-11) we obtain

$$\mathbf{Q}'E\left[\mathbf{I}_{[0,c)}\left(\frac{\mathbf{W}_1'\mathbf{W}_1}{\gamma}\right)\mathbf{W}\mathbf{W}'\right]\mathbf{Q} = p_2\mathbf{D} + (p_4 - p_2)\mathbf{D}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & aa'\end{pmatrix}\mathbf{D} + p_2\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & aa'\end{pmatrix}.$$

7. REFERENCES

1. Baksalary,J.K. and Kala,R. Partial orderings between matrices one of which is of rank one. *Bull. Polish Acad. Sci. Math.* **31**, 5-7 (1983)
2. Judge,G.G. and Bock,M.E. The statistical implications of Pre-test and Stein-rule Estimators in Econometrics, *North-Holland, Amsterdam*. (1978)
3. Theil,H. and Goldberger,A.S. On pure and mixed estimation in economics. *Internat.Econom.Rev.* **2**, 65-78 (1961)
4. Toutenburg,H. Mixed prior information in regression analysis, *Biometrical Journel*, **17**, 363-372 (1975)
5. Hill,C. and Ziemer,R.F. Missing regressor values under conditions of multicollinearity, *Communications in Statistics*, **A12**,2557-2573. (1983)
6. Judge,G.G. , Yancey,T.A.and Bock,M.E. Properties of estimators after preliminary tests of significance when stochastic restrictions are used in regression, *Journal of Econometrics*, **1**, 29-48 (1973)
7. Teräsvirta,T. Linear restrictions in misspecified linear models and polynomial distributed lag estimation, *Research report no.16, Department of Statistics, University of Helsinki*, 1-39, (1980)
8. Trenkler,G. and Toutenburg,H. Pretesting procedures and forcasting in the regression model under restrictions, *Discussion paper, Department of Statistics, University of Dortmund.* (1989)
9. Trenkler,G. Mean square error matrix comparisons of estimators in linear regression, *Comm. Statist. A-Theory Methods*, **10**, 2495-2509 (1985)