Faculty Working Papers

A WEAK MEAN SQUARE ERROR TEST
FOR STOCHASTIC RESTRICTIONS IN REGRESSION

T. A. Yancey, M. E. Bock and G. G. Judge

College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana - Champaign

March 19, 1971

A WEAK MEAN SQUARE ERROR TEST FOR STOCHASTIC RESTRICTIONS IN REGRESSION

T. A. Yancey, M. E. Bock and G. G. Judge

No. 8

A WEAK MEAN SQUARE ERROR TEST
FOR STOCHASTIC RESTRICTIONS IN REGRESSION

T.A. Yancey, M.E. Bock and G.G. Judge

The weak mean square error results of Wallace, relating to tests for linear restrictions in regression, are extended to include tests for stochastic linear restrictions.

1. Introduction

In a 1968 paper, Toro-Vizcarrondo and Wallace [6] showed that the non-central F distribution with a non-centrality parameter of one-half could be used to test the hypothesis that a linear equality restricted estimator, $\tilde{\beta}$, is better than an unrestricted least squares estimator, b, where better refers to the generalized mean square error (MSE) criterion. This means for the estimators $\tilde{\beta}$ and b that $\mathrm{MSE}(\ell\tilde{\beta}) \leq \mathrm{MSE}(\ell b)$ for every $\ell \neq o$. In a companion paper, the same authors [8] present tables of critical points, review the test procedure and give some examples of how the test may be used. In a 1970 paper, Yancey, Judge and Bock [9] present a test for determining the conditions under which stochastic linear prior information, which may be incorrect on the average, may improve the parameter estimates for the linear hypothesis model over conventional sample information estimates in the sense of having the same or smaller mean square errors for all estimates. In a 1971 paper in this Journal, Wallace [7] develops tests of betterness of the restricted estimator over the unrestricted least squares estimator where betterness is defined in terms of average squared Euclidean

distance (weak mean square error (WMSE)), such that in estimation for the vector $\underline{\theta}$,

$$(1.1)\qquad E(\hat{\underline{\theta}}-\underline{\theta})'(\hat{\underline{\theta}}-\underline{\theta}) = tr(\Sigma_{\hat{\underline{\theta}}\hat{\underline{\theta}}} + (\text{Bias } \hat{\underline{\theta}})(\text{Bias } \hat{\underline{\theta}})') = \sum_i MSE(\underline{\theta}_i).$$

In a companion paper, Goodnight and Wallace [1] develop tables to be used with the weak mean square error tests for linear restrictions in regression models.

In this note, we extend the weak mean square error results of Wallace [7] to include stochastic linear restrictions and show that the test statistic for the hypothesis

$$(1.2)\qquad E(\hat{\underline{\beta}}-\underline{\beta})'(\hat{\underline{\beta}}-\underline{\beta}) \leq E(\underline{b}-\underline{\beta})'(\underline{b}-\underline{\beta}),$$

where $\hat{\underline{\beta}}$ is the stochastic restricted estimator and $\underline{b}$ is the unrestricted least squares estimator, has a non-central F distribution with a non-centrality parameter of J/2, where J is the number of stochastic restrictions.

2. The Statistical Model

Consider the linear statistical model

$$(2.1)\qquad \underline{y} = X\underline{\beta} + \underline{u},$$

$$(2.2)^{1/}\qquad E(\underline{u}) = \underline{o} \quad \text{and} \quad E(\underline{u}\underline{u}') = \sigma^2 I,$$

where $\underline{y}$ is a (T × 1) vector of observations, X is a (T × K) matrix of non-

---

1/ Although we make use of the classical stochastic assumptions regarding the disturbances, the results of this paper are also valid for the more general specification $\sigma^2\theta$, where θ is some symmetric positive definite matrix.

stochastic variables of rank K, $\underline{u}$ is a $(T \times 1)$ vector of unobservable random normal variables, I is a unit matrix of order T, and $\underline{\beta}$ is a $(K \times 1)$ vector of unknown parameters.

Assume, in line with the work of Theil and Goldberger [5], Theil [4], Kakwani [3] and others, the following stochastic prior information about $\underline{\beta}$:

(2.3) $$\underline{r} = R\underline{\beta} + \underline{v},$$

(2.4) $$E(\underline{v}) = \underline{\delta}, \quad E(\underline{v}\underline{v}') = \sigma_o^2\Omega_o \quad \text{and} \quad E(\underline{u}\underline{v}') = 0,$$

where R is a $(J \times K)$ known matrix of rank J, $\underline{r}$ is a $(J \times 1)$ vector of prior estimates of $R\underline{\beta}$, $\underline{v}$ is a $(J \times 1)$ random vector of unobservable normal random variables with mean vector $\underline{\delta}$, representing the uncertainty about the possibly biased prior information $\underline{r}$, and $\sigma_o^2\Omega_o$ is a known non-singular matrix. The ratio $m = \sigma_o^2/\sigma^2$ is also assumed known. In the discussion to follow, we define $\Omega = m\Omega_o$.

Under these specifications, the stochastic restricted estimator (Theil and Goldberger [5] and Theil [4]), which makes use of the Aitken estimator to combine the sample and stochastic prior information, is

(2.5) $$\hat{\underline{\beta}} = [\sigma^{-2}X'X + \sigma_o^{-2}R'\Omega_o^{-1}R]^{-1}[\sigma^{-2}X'\underline{y} + \sigma_o^{-2}R'\Omega_o^{-1}\underline{r}]$$
$$= [X'X + R'\Omega^{-1}R]^{-1}[X'\underline{y} + R'\Omega^{-1}\underline{r}],$$

and the variance-covariance of $\hat{\underline{\beta}}$ is

(2.6) $$E(\hat{\underline{\beta}}-E\hat{\underline{\beta}})(\hat{\underline{\beta}}-E\hat{\underline{\beta}})' = \sigma^2[X'X + R'\Omega^{-1}R]^{-1} = \Sigma_{\hat{\underline{\beta}}\hat{\underline{\beta}}}.$$

The unrestricted least squares estimator, using only sample information, is

(2.7) $$\underline{b} = (X'X)^{-1}X'\underline{y},$$

where the variance-covariance of $\underset{\sim}{b}$ is

$$(2.8)\qquad E(\underset{\sim}{b}-E\underset{\sim}{b})(\underset{\sim}{b}-E\underset{\sim}{b})' = \sigma^2(X'X)^{-1} = \Sigma_{\underset{\sim}{b}\underset{\sim}{b}}.$$

If we make use of the mean square error criterion, and the corresponding expected mean square error matrix of the estimator $\hat{\underset{\sim}{\beta}}$, then we may write the risk matrix as

$$(2.9)\qquad E(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})' = MSE_{\hat{\underset{\sim}{\beta}}\hat{\underset{\sim}{\beta}}} = \sigma^2 W^{-1} + W^{-1}R\Omega^{-1}\underset{\sim}{\delta}\underset{\sim}{\delta}'\Omega^{-1}R'W^{-1},$$

where

$$W = (X'X + R'\Omega^{-1}R).$$

The difference between all linear combinations of the expected loss for the stochastic restricted estimator $\hat{\underset{\sim}{\beta}}$ and the unrestricted least squares estimator $\underset{\sim}{b}$ is

$$(2.10)\qquad \Delta_1 = \underset{\sim}{\ell}'E(\underset{\sim}{b}-\underset{\sim}{\beta})(\underset{\sim}{b}-\underset{\sim}{\beta})'\underset{\sim}{\ell} - \underset{\sim}{\ell}'E(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})'\underset{\sim}{\ell} = \underset{\sim}{\ell}'MSE_{\underset{\sim}{b}\underset{\sim}{b}}\underset{\sim}{\ell} - \underset{\sim}{\ell}'MSE_{\hat{\underset{\sim}{\beta}}\hat{\underset{\sim}{\beta}}}\underset{\sim}{\ell}$$

$$= \underset{\sim}{\ell}'(X'X + R'\Omega^{-1}R)^{-1}R'\Omega^{-1}[\sigma^2(R(X'X)^{-1}R' + \Omega) - \underset{\sim}{\delta}\underset{\sim}{\delta}'][R\Omega^{-1}(X'X + R'\Omega^{-1}R)^{-1}]\underset{\sim}{\ell},$$

which must be greater than or equal to zero if the stochastically restricted estimator is to be superior using the MSE criterion.

In order to test the compatibility of the stochastic prior and sample information using the mean square error criterion Yancey, Judge and Bock [9], following the reasoning of Toro-Vizcarrondo and Wallace [6], developed a test statistic which has a non-central $F(J,T-K,\lambda)$ distribution and showed that every linear combination of the elements of the mean square error

matrix of $\hat{\beta}$ is as small or smaller than the same linear combinations of the elements of the mean square error matrix of b if and only if the non-centrality parameter λ is $\leqq \frac{1}{2}$.

3. The Weak Mean Square Error Criterion and Test

As noted by Wallace [7], to have the mean square error of every combination of $\tilde{\beta}$ better than the corresponding linear combination of b is quite a strong requirement. Wallace [7] suggests what he terms a weak mean square error (WMSE) criterion as a reasonable but weaker requirement. For vector estimation, this criterion may be stated in general as

$$(3.1) \qquad \Delta_2 = E(\hat{\theta}-\theta)'(\hat{\theta}-\theta) = tr(\Sigma_{\hat{\theta}\hat{\theta}} + (\text{Bias } \hat{\theta})(\text{Bias } \hat{\theta})') = \sum_i MSE(\hat{\theta}_i),$$

where Δ_2 is the expected squared Euclidean distance from $\hat{\theta}$ to θ.

The WMSE criterion for the stochastic linear restriction estimator may be written as

$$(3.2) \qquad \begin{aligned} WMSE &= tr\, E(b-\beta)(b-\beta)' - tr\, E(\hat{\beta}-\beta)(\hat{\beta}-\beta)' \\ &= tr(X'X + R'\Omega^{-1}R)^{-1}R\Omega^{-1}[\sigma^2(R(X'X)^{-1}R' + \Omega) - \delta\delta'][\Omega^{-1}R(X'X \\ &\qquad + R'\Omega^{-1}R)^{-1}]. \end{aligned}$$

In order for the stochastic restricted estimator to be superior via this criterion, equation (3.2) must be equal to or greater than zero.

From (3.2), we can write

$$(3.3) \qquad \begin{aligned} \Delta_2 &= \ell'(X'X + R'\Omega^{-1}R)^{-1}R'\Omega^{-1}[\sigma^2(R(X'X)^{-1}R' + \Omega) - \delta\delta']\Omega^{-1}R(X'X \\ &\qquad + R'\Omega^{-1}R)^{-1}\ell \\ &= \ell'A'BA\ell, \end{aligned}$$

which is positive semi-definite for all non-zero $(J \times 1)$ vectors $\underset{\sim}{\ell}$ if and only if

(3.4) $$B = [\sigma^2(R(X'X)^{-1}R' + \Omega) - \underset{\sim}{\delta}\underset{\sim}{\delta}']$$

is positive semi-definite.[2/]

Hence,

(3.5) $$\text{tr } B = \text{tr}[\sigma^2(R(X'X)^{-1}R' + \Omega) - \underset{\sim}{\delta}\underset{\sim}{\delta}'] \geqq 0.$$

The

(3.6) $$\text{tr}[\sigma^2(R(X'X)^{-1}R' + \Omega) - \underset{\sim}{\delta}\underset{\sim}{\delta}']$$

can be written as

(3.7) $$\text{tr}[\sigma^2 I - \underset{\sim}{\delta}\underset{\sim}{\delta}'(R(X'X)^{-1}R' + \Omega)^{-1}](R(X'X)^{-1}R' + \Omega),$$

where $(R(X'X)^{-1}R' + \Omega)$ is positive definite. Since the characteristic roots of a positive definite matrix are all positive, from (3.5) the

(3.8) $$\text{tr}[\sigma^2 I - \underset{\sim}{\delta}\underset{\sim}{\delta}'(R(X'X)^{-1}R' + \Omega)^{-1}](R(X'X)^{-1}R' + \Omega)$$

is equal to

(3.9) $$\sum_{i=1}^{n} \psi_{j(i)}\phi_i \geqq 0$$

[2/] This can be seen since $A = (X'X + R'\Omega^{-1}R)^{-1}$ is positive definite by assumption, and we can let the $\underset{\sim}{s}$, a non-negative $(J \times 1)$ vector in the quadratic form, in $\underset{\sim}{s}'B\underset{\sim}{s} \geq 0$ equal $A\underset{\sim}{\ell}$. Given either $\underset{\sim}{s}$ or $\underset{\sim}{\ell}$, we can find $\underset{\sim}{\ell} = A\underset{\sim}{s}$ or $\underset{\sim}{s} = A^{-1}\underset{\sim}{\ell}$. Clearly, $\underset{\sim}{s} = \bar{0}$, a null vector, if and only if $\underset{\sim}{\ell} = \underset{\sim}{0}$. Thus, C is positive semi-definite if and only if $A'BA$ is positive semi-definite since $\underset{\sim}{\ell}'ABA\underset{\sim}{\ell} = \underset{\sim}{s}'B\underset{\sim}{s} \geq 0$ for all $\underset{\sim}{s}$ and $\underset{\sim}{\ell}$ not identically zero and $\underset{\sim}{\ell}'ABA\underset{\sim}{\ell} = \underset{\sim}{s}B\underset{\sim}{s} = \bar{0}$ for some $\underset{\sim}{s}$ and $\underset{\sim}{\ell} \neq 0$. Furthermore, if a matrix is positive semi-definite, its trace is non-negative (Graybill [2, p. 318]), and we have $\text{tr } A'BA \geqq 0$ if and only if $\text{tr } B \geqq 0$.

for some ordering of ψ_i, where ϕ_i are characteristic roots of $(R(X'X)^{-1}R' + \Omega)$ and ψ_i the characteristic roots of $[\sigma^2 I - \underline{\delta\delta}'(R(X'X)^{-1}R' + \Omega)^{-1}]$ (Graybill [2, p. 233]).

Since the ϕ_i are all positive and since $\Sigma\psi_{j(i)}\phi_i \geqq 0$ for all sets of positive definite and semi-definite matrices $(R(X'X)^{-1}R' + \Omega)$ and $[\sigma^2 I - \underline{\delta\delta}'(R(X'X)^{-1}R' + \Omega)^{-1}]$ respectively, the $\psi_{j(i)}$ are all non-negative. Furthermore, the characteristic roots of positive semi-definite matrices are non-negative, $[\sigma^2 I - \underline{\delta\delta}'(R(X'X)^{-1}R' + \Omega)^{-1}]$ is positive semi-definite, and $tr[\sigma^2 I - \underline{\delta\delta}'(R(X'X)^{-1}R' + \Omega)^{-1}] \geqq 0$ which implies

$$(3.10) \qquad J\sigma^2 - \underline{\delta}'(R(X'X)^{-1}R' + \Omega)^{-1}\underline{\delta} \geqq 0,$$

or

$$(3.11) \qquad \frac{J}{2} \geqq \frac{1}{2\sigma^2}\underline{\delta}'(R(X'X)^{-1}R' + \Omega)\underline{\delta} = \lambda.$$

Hence, if the non-centrality parameter λ of the distribution of $\hat{\gamma}$ is less than or equal to $\frac{J}{2}$, the estimator $\hat{\underline{\beta}}$ is better than $\underline{b}$ when judged by the weak MSE criterion. Similarly, $\lambda \geq \frac{J}{2}$ if and only if $E(\underline{b}-\underline{\beta})'(\underline{b}-\underline{\beta}) \leq E(\hat{\underline{\beta}}-\underline{\beta})'(\hat{\underline{\beta}}-\underline{\beta})$.

Given this result, one can test the hypothesis that the use of stochastic linear restrictions improves the estimator in a weak MSE sense over its unrestricted least squares counterpart, by testing the hypothesis $\lambda \leq \frac{J}{2}$ against the alternative $\lambda > \frac{J}{2}$ and by using the tables developed by Goodnight and Wallace [1], noting whether $\hat{\gamma} \geq T^{\alpha}(\frac{J}{2}, J, T-K)$ at the α test level.

4. Concluding Remarks

A test for stochastic restrictions in regression which recognizes the trade-off between bias and variance has been developed. It should be noted in closing that if we remove the assumption that m, the ratio of the variances of the prior restriction and sample errors, is known, statistics are created whose distributions are unknown and untabled. However, for the statistical model of this paper, preliminary sampling experiments indicate that the non-central F distribution is a reasonable approximation to the empirical distribution when m is replaced by an estimate, $\hat{m} = \tilde{\sigma}_0^2/\check{\sigma}^2$. Finally, let us note that as in the equality restricted estimator case, the properties of the preliminary test estimator,

$$\hat{\hat{\underline{\beta}}} = \hat{\underline{\beta}}\, I_{(0,c)}(\hat{\gamma}) + \underline{b}\, I_{(c,\infty)}(\hat{\gamma}),$$

which are based on the preliminary test about λ, where

$$I_{(0,c)}(\hat{\gamma}) = 1 \quad \text{and} \quad I_{(c,\infty)}(\hat{\gamma}) = 0, \quad \text{when } \hat{\gamma} < c,$$

and

$$I_{(0,c)}(\hat{\gamma}) = 0 \quad \text{and} \quad I_{(c,\infty)}(\hat{\gamma}) = 1, \quad \text{when } \hat{\gamma} \geq c,$$

for

$$\int_c^\infty f(\hat{\gamma})d\hat{\gamma} = \alpha,$$

are yet to be determined. The authors are currently working on the sampling properties of this and related preliminary test estimators.

University of Illinois

References

[1] Goodnight, J. and T.D. Wallace (1971). "Operational Techniques and Tables for Making Weak MSE Tests for Restrictions in Regression," Econometrica, _____________.

[2] Graybill, F.A. (1969). Introduction to Matrices with Applications in Statistics, Wadsworth Publishing Company, Belmont, California.

[3] Kakwani, N.C. (1968). "Note on the Unbiasedness of a Mixed Regression Estimator," Econometrica, Vol. 36, pp. 610-11.

[4] Theil, H. (1963). "On the Use of Incomplete Prior Information in Regression Analysis," Journal of the American Statistical Association, Vol. 58, pp. 401-14.

[5] Theil, H. and A.S. Goldberger (1961). "On Pure and Mixed Estimation in Economics," International Economic Review, Vol. 2, pp. 65-78.

[6] Toro-Vizcarrondo, C. and T.D. Wallace (1968). "A Test of the Mean Square Error Criterion for Restrictions in Linear Regression," Journal of the American Statistical Association, Vol. 63, pp. 558-572.

[7] Wallace, T.D. (1971). "Weaker Criteria and Tests for Linear Restrictions in Regression," Econometrica, ___________.

[8] Wallace, T.D. and C.E. Toro-Vizcarrondo (1969). "Tables for the Mean Square Error Test for Exact Linear Restrictions in Regression," Journal of the American Statistical Association, Vol. 64, pp. 1644-93.

[9] Yancey, T.A., G.G. Judge and M.E. Bock (1970). "A Mean Square Error Test When Stochastic Restrictions are Used in Regression," Workshop Paper No. 7001, Department of Economics, University of Illinois.

SELECTIVE CONTROL OF INDEPENDENT ACTIVITIES: LINEAR PROGRAMMING OF MARKOVIAN DECISIONS+

Hirohide Hinomoto
University of Illinois
Urbana, Illinois