J. R. Statist. Soc. B (1989)
51, *No.* 1, *pp.* 71–79

On the Optimality of some Tests of the Error Covariance Matrix in the Linear Regression Model

By YUZO HONDA†

Kobe University, Japan

[Received August 1987. Final revision May 1988]

SUMMARY

This paper adopts the approach of conditional distributions to investigate the optimality of some tests of the error covariance matrix in the linear regression model. Specifically, we show that point optimal tests, advocated by King and Evans and King, are the most powerful similar test. We also derive the locally best similar and the locally best unbiased similar tests. Finding the latter cumbersome to apply, we then propose the asymptotically best similar and the asymptotically best unbiased similar (ABUS) tests as alternative criteria. We show that the two-sided Durbin–Watson test is ABUS against serial correlation and that the two-sided Lagrange multiplier test is ABUS against heteroscedastic disturbances.

Keywords: ASYMPTOTICALLY BEST SIMILAR TESTS; ASYMPTOTICALLY BEST UNBIASED SIMILAR TESTS; DURBIN-WATSON TEST; LAGRANGE MULTIPLIER TEST; LOCALLY BEST SIMILAR TESTS; LOCALLY BEST UNBIASED SIMILAR TESTS

1. INTRODUCTION

There are two important approaches to 'optimal tests' in composite hypotheses. The first is the use of conditional distributions, and the second appeals to the argument of invariance. Although quite a few researchers write about the optimality of certain tests of the error covariance matrix in the linear regression model, most, except for Anderson (1948), discuss the optimality of tests from the viewpoint of invariance alone. This tendency has been conspicuous especially after Durbin and Watson (1971) gave their justifications for their test from the standpoint of invariance.

This paper takes the first approach, i.e. the use of conditional distributions, and investigates some of these issues. Specifically, we give an alternative justification for the so-called point optimal tests, advocated by Evans and King (1985) and King (1985), by showing that they are the most powerful similar (MPS) test. We also derive the locally best similar (LBS) test for the one-sided alternative problem and the locally best unbiased similar (LBUS) test for the two-sided alternative.

Finding the LBUS test cumbersome to apply, we then propose the asymptotically best similar (ABS) test and the asymptotically best unbiased similar (ABUS) test as alternative criteria of optimality. These new criteria enable us to justify certain tests from an optimal viewpoint. For example, a two-sided version of the Durbin–Watson test is known to be approximately locally best unbiased invariant (LBUI) against autoregressive (AR(1)) disturbances only when certain restrictions on a design matrix are satisfied. However, these restrictions do not hold in standard regressions. We demonstrate that the two-sided Durbin–Watson test is ABUS and asymptotically best

†*Address for correspondence*: School of Business Administration, Kobe University, Nada-ku, Kobe 657, Japan.

 0035–9246/89/51071 $2.00

unbiased invariant (ABUI) even when these restrictions are violated. The second example is testing for heteroscedasticity. Although the two-sided Lagrange multiplier (LM) test is neither LBUS nor LBUI for testing additive or multiplicative heteroscedasticity, it is shown to be ABUS and ABUI.

The outline of the paper is as follows. Section 2 obtains the conditional density, applies the Neyman–Pearson lemma to this density and provides the MPS test. Section 3 derives the LBS and the LBUS tests and motivates the ABS and the ABUS tests. Section 4 applies asymptotically optimal tests to detecting serial correlation. The final section employs these ideas for testing for heteroscedasticity.

2. POINT OPTIMAL TESTS

Let us consider the normal linear model

$$y = X\beta + u; \; u \sim N(0, \sigma^2\Omega(\theta)), \tag{1}$$

where y is an $n \times 1$ random vector, X is an $n \times k$ fixed matrix of rank $k < n$, u is an $n \times 1$ disturbance vector, $\Omega(\theta)$ is a symmetric positive definite matrix for all θ with the condition $\Omega(0) = I_n$, and the column vector β and scalar σ^2 are unknown parameters. Our problem is to test the null hypothesis $H{:}\theta = 0$ against the one-sided alternative hypotheses $K_1^+{:}\theta > 0$, $K_1^-{:}\theta < 0$, or the two-sided alternative hypothesis $K_2{:}\theta \neq 0$.

The joint density of y in equation (1) is given by

$$f_1(y; \beta, \sigma^2, \theta) = (2\pi)^{-n/2}\sigma^{-n}|\Omega(\theta)|^{-1/2}\exp[-(y - X\beta)'\Omega^{-1}(\theta)(y - X\beta)/2\sigma^2]. \tag{2}$$

Since we are mainly interested in the local optimality of tests in this paper, we consider testing $H{:}\theta = 0$ against a specific alternative $K_s{:}\theta = \theta_1$ (where θ_1 is a specific value of θ, not equal to zero) and then let θ_1 approach zero. For this we assume for the moment that θ is known. With the value of θ known (to be either zero or θ_1), the statistics $(\hat{\beta}, \hat{\sigma}^2)$ are jointly sufficient and complete for (β, σ^2), where $\hat{\beta}$ and $\hat{\sigma}^2$ are defined by

$$\left.\begin{aligned} \hat{\beta} &= [X'\Omega^{-1}(\theta)X]^{-1}X'\Omega^{-1}(\theta)y, \\ \hat{u} &= y - X\hat{\beta}, \\ \hat{\sigma}^2 &= \hat{u}'\Omega^{-1}(\theta)\hat{u}/(n - k). \end{aligned}\right\} \tag{3}$$

Hence our approach is to use the conditional density of y given the values of $\hat{\beta}$ and $\hat{\sigma}^2$, $g(y|\hat{\beta}, \hat{\sigma}^2; \theta)$, to find the best test of $H{:}\theta = 0$. This approach is based on Cox and Hinkley (1974), Kendall and Stuart (1979) and Ferguson (1967). In particular, the reader is referred to pp. 134–136 in Cox and Hinkley (1974) for details. To obtain the conditional density $g(y|\hat{\beta}, \hat{\sigma}^2; \theta) = f_1(y; \beta, \sigma^2, \theta)/f_2(\hat{\beta}, \hat{\sigma}^2; \beta, \sigma^2, \theta)$, we need to find the joint density of $(\hat{\beta}, \hat{\sigma}^2)$, $f_2(\hat{\beta}, \hat{\sigma}^2; \beta, \sigma^2, \theta)$, first.

Note the following three facts. First, $\hat{\beta}$ has a k-variate normal distribution with mean β and covariance matrix $\sigma^2(X'\Omega^{-1}X)^{-1}$. Secondly, $(n - k)\hat{\sigma}^2/\sigma^2$ has a chi-squared distribution with $n - k$ degrees of freedom. Thirdly, $\hat{\beta}$ and $\hat{\sigma}^2$ are statistically independent. Combining these three facts, we find the joint density of $(\hat\beta, \hat\sigma^2)$ to be

$$f_2(\hat\beta, \hat\sigma^2; \beta, \sigma^2, \theta) = \frac{|X'\Omega^{-1}X|^{1/2} m^{m/2} (\hat\sigma^2)^{(m/2)-1}}{(2\pi)^{k/2}\sigma^{k+m}\Gamma(m/2)2^{m/2}} \times \exp[-(y - X\beta)'\Omega^{-1}(y - X\beta)/2\sigma^2], \tag{4}$$

where $m = n - k$. In obtaining equation (4), we have made use of the fact that

$$(y - X\beta)'\Omega^{-1}(y - X\beta) = m\hat\sigma^2 + (\hat\beta - \beta)'X'\Omega^{-1}X(\hat\beta - \beta).$$

It follows that the conditional density of y is

$$g(y|\hat\beta, \hat\sigma^2; \theta) = \frac{\Gamma(m/2)}{m\pi^{m/2}|X'\Omega^{-1}X|^{1/2}|\Omega|^{1/2}(\hat u'\Omega^{-1}\hat u)^{(m/2)-1}}. \tag{5}$$

Now apply the Neyman–Pearson lemma to conditional density (5). Then the best critical region of the null hypothesis $H{:}\theta = 0$ against a specific alternative hypothesis $K_s{:}\theta = \theta_1$ of a particular size α is given by

$$\frac{g(y|\hat\beta, \hat\sigma^2; \theta_1)}{g(y|b, s^2; 0)} > w_\alpha \tag{6}$$

where

$$\left.\begin{aligned} b &= (X'X)^{-1}X'y, \\ e &= y - Xb, \\ s^2 &= e'e/(n - k), \end{aligned}\right\} \tag{7}$$

and w_α is an appropriate constant. (Because the statistics $(\hat\beta, \hat\sigma^2)$ are jointly sufficient for (β, σ^2) under both the null and the alternative hypotheses, the arguments in the numerator in equation (6) do not contain particular values of nuisance parameters β and σ^2, unlike the more general case considered in equation (4) in Cox and Hinkley (1974), p. 135.)

Since equation (6) is equivalent to the inequality

$$\hat u'\Omega^{-1}(\theta_1)\hat u/e'e < w_1, \tag{8}$$

where w_1 is another constant, we have proven the following theorem.

Theorem 1. The test whose critical regions take the form of equation (8) is the MPS test of the null hypothesis $H{:}\theta = 0$ against a specific alternative hypothesis $K_s{:}\theta = \theta_1$ (known).

The test based on equation (8) is called the point optimal test, since Durbin and Watson (1971), pp. 9–10, Kadiyala (1970), Berenblut and Webb (1973), pp. 49–50, and King (1980), theorem 3, show that the test is most powerful invariant (MPI) for known θ_1. King (1985) and Evans and King (1985) advocate these point optimal tests for detecting first-order AR disturbances and heteroscedastic disturbances respectively.

3. LOCALLY BEST (UNBIASED) SIMILAR TESTS AND ASYMPTOTICALLY BEST (UNBIASED) SIMILAR TESTS

In the previous section we treated parameter θ_1 as known. In this section we shall make θ_1 sufficiently small and derive the LBS test for the one-sided alternative and the LBUS test for the two-sided case.

Define $A(\theta) = \partial\Omega^{-1}(\theta)/\partial\theta$ and $A_0 = A(0)$. Then we see that

$$\hat{u}'\Omega^{-1}(\theta_1)\hat{u} = e'e + \theta_1 e'A_0 e + o_p(\theta_1) \tag{9}$$

for small θ_1. Substituting equation (9) into equation (8), we have an inequality $\theta_1 q < w_2$ for small θ_1, where q is defined by

$$q = e'A_0 e/e'e, \tag{10}$$

and w_2 is another constant. We now have the following theorem.

Theorem 2. The test which rejects $H:\theta = 0$ for small values of q in equation (10) is the LBS test of H against the alternative hypothesis $K_1^+:\theta > 0$. Similarly, the test which rejects $H:\theta = 0$ for large values of q is the LBS test of H against $K_1^-:\theta < 0$.

Alternatively, we can derive the first part of theorem 2 directly by taking the logarithm of the conditional density $g(y|\hat{\beta}, \hat{\sigma}^2; \theta)$ in equation (5) and applying condition (5.78) in Ferguson (1967), p. 236, to it. Also note that theorem 2 implies that the LBS test coincides with the locally best invariant (LBI) test in King and Hillier (1985) for the one-sided alternative hypothesis.

For two-sided tests, consider a local quadratic approximation to the power function in the neighbourhood of $\theta = 0$. Then the first term is the size of the tests, α, which is a constant. The second term is zero, because we require the tests to be unbiased. Thus the maximization of local power subject to the conditions of size α and unbiasedness is equivalent to the maximization of the second derivative of the power function subject to these conditions. The maximization of the second derivative of the power function problem in turn can be solved with the aid of a generalized version of the Neyman–Pearson lemma stated in p. 213 of Neyman and Pearson (1966) (or in Neyman and Pearson (1936)), yielding a locally best unbiased test as in Neyman and Pearson (1966), p. 212, or in Ferguson (1967), p. 238.

We derive this locally best unbiased test for our problem by taking the first two partial derivatives of the logarithm of the conditional density $g(y|\hat{\beta}, \hat{\sigma}^2; \theta)$ in equation (5). The resulting functions are evaluated at $\theta = 0$. The quantities found are substituted into equation (5.88) in Ferguson (1967), p. 238. This gives the following theorem.

Theorem 3. The LBUS test of $H:\theta = 0$ against $K_2:\theta \neq 0$ takes critical regions of the form

$$\frac{2e'A_0GA_0e - e'D_0e}{e'e} + \frac{m}{2}q^2 + (2E + w_4)q > \frac{2(F - E^2 + w_3 - Ew_4)}{m - 2}, \tag{11}$$

where $D(\theta) = \partial A(\theta)/\partial\theta$, $D_0 = D(0)$, $G = X(X'X)^{-1}X'$, $R(\theta) = \partial(\log|X'\Omega^{-1}X|)/\partial\theta$, $S(\theta) = \partial(\log|\Omega|)/\partial\theta$, $E = [R(0) + S(0)]/2$, $2F = \partial(R + S)/\partial\theta|_{\theta=0}$, and w_3 and w_4 are constants chosen to satisfy the conditions of size and local unbiasedness.

Theorem 3 shows that the first term in inequality (11) is an obstacle hindering us from writing critical regions of the LBUS test as

$$(q + w_5)^2 > w_6, \tag{12}$$

where w_5 and w_6 are appropriate constants. The LBUI test in King and Hillier (1985) involves exactly the same term. Thus King and Hillier (1985) conclude that a two-sided version of the Durbin–Watson test can be justifiable from an optimal viewpoint only when the condition $(I_n - G)A_0G = 0$ (which implies that $e'A_0GA_0e = 0$) holds. However, this last condition does not hold in the usual regressions.

In the following sections we introduce the ABS test and the ABUS test as alternative criteria of optimality. The definition of these optimal tests is as follows. For a local test to be useful we are really assuming that the sample size is sufficiently large for only alternative hypotheses in the neighbourhood of the null to be of interest (I owe this idea to one of the referees). Specifically, we consider a sequence of local alternatives converging to the null at the rate $1/\sqrt{n}$ and find the locally best (unbiased) tests for each point of this sequence. The proposed ABS and ABUS tests are the limit of this sequence of locally optimal tests.

4. TESTING FOR SERIAL CORRELATION

Thus far we have developed our optimal tests in the general setting (1). Now we apply these general results to particular problems. In this section, we consider testing for serial correlation in the disturbance. In the next section, we investigate the optimality of the LM test for heteroscedasticity.

When the components of an $n \times 1$ disturbance vector u follow a first-order AR process in the alternative, i.e.

$$u_t = \rho u_{t-1} + \varepsilon_t, \ |\rho| < 1, \ (t = 2, 3, \ldots, n), \tag{13}$$

with $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \sim N(0, \sigma^2 I_n)$ and $u_1 = \varepsilon_1/(1 - \rho^2)^{1/2}$, we have

$$\Omega(\rho) = [(1 - \rho)^2 I_n + \rho B_1 + \rho(1 - \rho)C_1]^{-1}, \tag{14}$$

where B_1 and C_1 are $n \times n$ matrices such that

$$B_1 = \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 & 0 \\ -1 & 2 & -1 & \ldots & 0 & 0 \\ 0 & -1 & 2 & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \ddots & \cdot \\ \cdot & \cdot & \cdot & & & \cdot \\ 0 & 0 & 0 & \ldots & 2 & -1 \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{pmatrix} \text{ and } C_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & & \ddots & \cdot \\ \cdot & \cdot & & \ddots & \cdot \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix}, \tag{15}$$

respectively (see King (1980, 1981)).

Anderson (1948) finds that there is no uniformly most powerful similar (UMPS) test of $H{:}\rho = 0$ against either $K_1^+{:}\rho > 0$ or $K_1^-{:}\rho < 0$. He also shows the following result. If the covariance matrix takes the form

$$\Omega(\rho) = [(1 + \rho^2)I_n - 2\rho\Sigma]^{-1}, \tag{16}$$

where Σ is an $n \times n$ matrix, and if the column space of a design matrix X is spanned by some k latent vectors of a matrix Σ, then the UMPS test of the hypothesis $H{:}\rho = 0$ against $K_1^+{:}\rho > 0$ is given by $r > w_7$, where $r = e'\Sigma e/e'e$ and w_7 is a constant such that it gives a critical region of an appropriate size.

Our discussion in the previous sections indicates that the relevant quantity involving the covariance matrix (14) is the term $\hat{u}'\Omega^{-1}(\rho)\hat{u}$, which is the sum of the three quadratic forms. The last term of these three, $\rho(1-\rho)\hat{u}'C_1\hat{u}$, is of smaller order than the other two for large n. Thus we approximate the covariance matrix (14) by

$$[(1-\rho)^2 I_n + \rho B_1]^{-1}. \tag{17}$$

Then Anderson's result with $\Sigma = I_n - B_1/2$ implies that rejecting $H{:}\rho = 0$ for small values of

$$d = \frac{e'B_1 e}{e'e} \tag{18}$$

is approximately UMPS against $K_1^+{:}\rho > 0$, when the column space of a design matrix X is spanned by k latent vectors of B_1. The statistic (18) is known as the Durbin–Watson (1950, 1951) test statistic. Durbin and Watson (1971), pp. 9–10, furthermore justify their statistic from a standpoint of invariance. They show that their test is approximately uniformly most powerful invariant (UMPI) with this condition on X. Further, the test is shown to be approximately LBI in the neighbourhood of $\rho = 0$ for the one-sided alternative hypothesis.

Now let us assume that $(1/n)X'X$ converges to a positive definite matrix. The convergence of $(1/n)X'X$ together with the Schwarz inequality imply that

$$\frac{1}{n}\sum_{t=1}^{n-1} x_{ti}x_{(t+1)j} \text{ is } O(1) \quad (i, j = 1, 2, \ldots, k). \tag{19}$$

(I owe this to one of the referees.) Making use of these conditions, we have the following theorem.

Theorem 4. The Durbin–Watson test, which rejects $H{:}\rho = 0$ for small values of d in equation (18), is ABS against $K_1^+{:}\rho > 0$. Similarly, the test, which rejects $H{:}\rho = 0$ for large values of d, is ABS against $K_1^-{:}\rho < 0$.

Proof. Partially differentiate the inverse of equation (14) and evaluate the resulting function at $\rho = 0$. Then we have $A_0 = -2I_n + B_1 + C_1$. Substituting this A_0 into equation (10), we see that the LBS test takes the form

$$\frac{e'B_1 e + e'C_1 e}{e'e} < w_8 + 2 \tag{20}$$

against K_1^+, and also, with the inequality reversed, against K_1^-. w_8 in inequality (20) represents a sequence of constants chosen to satisfy the size condition. However, inequality (20) is asymptotically equivalent to the corresponding inequality on equation (18), since the term $e'C_1 e/e'e$ vanishes asymptotically. This completes the theorem.

Define the two-sided Durbin–Watson test as the test whose critical regions take the form

$$(d + w_9)^2 > w_{10},$$

where w_9 and w_{10} are constants. King and Hillier (1985) show that the two-sided Durbin–Watson test is approximately LBUI if the condition $(I_n - G)A_0 G = 0$ holds. However, this condition usually does not hold in an ordinary regression. Thus, we propose the following theorem.

Theorem 5. The two-sided Durbin–Watson test is ABUS.

Proof. For brevity we just give an outline of the proof. Consider a sequence of optimal tests in inequality (11). Corresponding to each test we have a sequence of constants chosen to satisfy the size and the local unbiasedness conditions of each test. We write a sequence of these numbers as (w_3, w_4) without an index, for simplicity. w_3 and w_4 are of the order of n^2 and n respectively.

Now the term $e'A_0GA_0e/e'e$ in inequality (11) can be shown to be $o_p(1)$ by Chebyshev's inequality. The term $m^{1/2}q$ in inequality (11) is $O_p(1)$. Both E and F turn out to be $O(1)$ by corollaries 30 and 38 and propositions 100 and 105 in Dhrymes (1978), p. 534, p. 540, p. 532 and p. 538 respectively.

Making use of the relationship $q = -2 + d + o_p(1)$, we thus have

$$\left[d - \left(2 - \frac{w_4}{n}\right)\right]^2 > \frac{1}{n^2}(4w_3 + w_4^2). \tag{21}$$

Putting $w_9 = -2 + n^{-1}w_4$ and $w_{10} = n^{-2}(4w_3 + w_4^2)$ proves the theorem. □

A few remarks are needed here. First, we have derived an ABS test in inequality (20) and an ABUS test in inequality (21). In a similar fashion we can extend an LBI test and an LBUI test by King and Hillier (1985) to an asymptotically best invariant (ABI) test and an ABUI test respectively with appropriate convergence conditions on X. It turns out that our ABS test coincides with an ABI test and that our ABUS test agrees with an ABUI test.

Secondly, some tests have the same logical foundation as that of the Durbin–Watson test. For example, Bhargava *et al.* (1982) generalize the Durbin–Watson statistic to panel data. Wallis (1972) considers testing for fourth-order autocorrelation. From the discussion in this section these tests are also ABS and ABI for the one-sided alternatives and ABUS and ABUI for the two-sided alternative.

Finally King (1981) investigates the power properties of a slightly modified test statistic

$$d' = \frac{e'B_0e}{e'e}, \tag{22}$$

where $B_0 = B_1 + C_1$, and B_1 and C_1 are defined in equation (15). Since d' and the Durbin–Watson d are asymptotically equivalent, the tests using d and d' share the same asymptotic optimal properties.

5. TESTING FOR HETEROSCEDASTICITY

King and Hillier (1985) prove that the one-sided LM test is LBI. Our theorem 2 shows that it is LBS as well. These results can be used, for instance, to show that the one-sided LM test for the error components model by Honda (1985) is both LBS and LBI.

Another example is testing for heteroscedasticity. Consider the model (1) with u_ts normally and independently distributed with mean zero and variance

$$\sigma_t^2 = \alpha_1 + \alpha_2 z_t = \sigma^2(1 + \theta z_t) \quad (t = 1, 2, \ldots, n), \tag{23}$$

where the z_ts are the exogenous variables and α_1 and α_2 are scalar unknown parameters. We are interested in testing $H{:}\alpha_2 = 0$. Farebrother (1987) intensively investigates alternative tests for this sort of model, including the LM test by Breusch and Pagan (1979). Here again, the one-sided LM test is both LBS and LBI.

Let us define the two-sided LM test as the test whose critical regions take the form

$$(q + w_{11})^2 > w_{12},$$

where w_{11} and w_{12} are constants. On the basis of the results of this paper as well as those by King and Hillier (1985), the two-sided LM test is neither LBUS nor LBUI in testing $H{:}\alpha_2 = 0$ against $K_2{:}\alpha_2 \neq 0$ in equation (23).

Now assume that $(1/n)X'X$ converges to a positive definite matrix, that

$$\sum_{t=1}^{n} z_t^{2r} x_{ti}^2 / z_s^{2r} x_{si}^2 \to \infty \quad (i = 1, 2, \ldots, k;\ s = 1, 2, \ldots n;\ r = 1, 2), \tag{24}$$

and that

$$\frac{1}{n} \sum_{t=1}^{n} x_{ti} x_{tj} z_t^r \text{ is } O(1) \quad (i, j = 1, 2, \ldots k;\ r = 1, 2). \tag{25}$$

Using similar arguments to the serial correlation case in Section 4, we obtain the following theorem.

Theorem 6. The two-sided LM test is ABUS.

When the alternative hypothesis takes a multiplicative form of heteroscedasticity

$$\sigma_t^2 = \exp(\alpha_1 + \alpha_2 z_t) = \sigma^2 \exp(\theta z_t) \quad (t = 1, 2, \ldots, n), \tag{26}$$

the ABUS test remains unchanged. Hence the two-sided LM test is ABUS against the alternative, equation (26), as well.

Finally we can extend the results in King and Hillier (1985) and derive the ABUI test. It turns out that the ABUI test coincides with the ABUS test for both additive and multiplicative heteroscedastic models.

ACKNOWLEDGEMENTS

In preparing this paper, I learned a great amount on the topic through correspondence with Professor Geoffrey S. Watson. One of the referees provided me with many perceptive comments on earlier versions of the paper. The referee, Professor D. M. Titterington and Professor Peter C. Mayer helped me in improving the presentation. I sincerely thank all of them. However, any remaining errors are solely my own. Financial support from the Japan Economic Research Foundation is also gratefully acknowledged.

REFERENCES

Anderson, T. W. (1948) On the theory of testing serial correlation. *Skand. Akt.*, **31**, 88–116.

Berenblut, I. I. and Webb, G. I. (1973) A new test for autocorrelated errors in the linear regression model. *J. R. Statist. Soc.* B, **35**, 33–50.

Bhargava, A., Franzini, L. and Narendranathan, W. (1982) Serial correlation and the fixed effects model. *Rev. Econ. Stud.*, **49**, 533–549.

Breusch, T. S. and Pagan, A. R. (1979) A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, **47**, 1287–1294.

Cox, D. R. and Hinkley, D. V. (1974) *Theoretical Statistics*. London: Chapman and Hall.
Dhrymes, P. J. (1978) *Introductory Econometrics*. New York: Springer.
Durbin, J. and Watson, G. S. (1950) Testing for serial correlation in least squares regression I. *Biometrika*, **37**, 409–428.
——(1951) Testing for serial correlation in least squares regression II. *Biometrika*, **38**, 159–178.
——(1971) Testing for serial correlation in least squares regression III. *Biometrika*, **58**, 1–19.
Evans, M. A. and King, M. L. (1985) A point optimal test for heteroscedastic disturbances. *J. Economet.*, **27**, 163–178.
Farebrother, R. W. (1987) The statistical foundations of a class of parametric tests for heteroscedasticity. *J. Economet.*, **36**, 359–368.
Ferguson, T. S. (1967) *Mathematical Statistics: a Decision Theoretic Approach*. New York: Academic Press.
Honda, Y. (1985) Testing the error components model with non-normal disturbances. *Rev. Econ. Stud.*, **52**, 681–690.
Kadiyala, K. R. (1970) Testing for the independence of regression disturbances. *Econometrica*, **38**, 97–117.
Kendall, M. and Stuart, A. (1979) *The Advanced Theory of Statistics*, vol. 2, *Inference and Relationships*, 4th edn. London: Griffin.
King, M. L. (1980) Robust tests for spherical symmetry and their application to least squares regression. *Ann. Statist.*, **8**, 1265–1271.
——(1981) The alternative Durbin–Watson test: an assessment of Durbin and Watson's choice of test statistic. *J. Economet.*, **17**, 51–66.
——(1985) A point optimal test for autoregressive disturbances. *J. Economet.*, **27**, 21–37.
King, M. L. and Hillier, G. H. (1985) Locally best invariant tests of the error covariance matrix of the linear regression model. *J. R. Statist. Soc.* B, **47**, 98–102.
Neyman, J. and Pearson, E. S. (1936) Contributions to the theory of testing statistical hypotheses. *Statist. Res. Mem.*, **1**, 1–37.
——(1966) *Joint Statistical Papers*. Berkeley: University of California Press.
Wallis, K. F. (1972) Testing for fourth order autocorrelation in quarterly regression equations. *Econometrica*, **40**, 617–636.