The Information Matrix Test for the Linear Model

Author(s): Alastair Hall

Source: *The Review of Economic Studies*, Apr., 1987, Vol. 54, No. 2 (Apr., 1987), pp. 257-263

Published by: Oxford University Press

Stable URL: https://www.jstor.org/stable/2297515

Review of Economic Studies (1987) LIV, 257-263 0034-6527/87/00170257$02.00

The Information Matrix Test for the Linear Model

ALASTAIR HALL
North Carolina State University

First version received August 1983; *final version accepted August* 1986 (*Eds.*)

We derive the information matrix test, suggested by White, for the normal fixed regressor linear model, and show that the statistic decomposes asymptotically into the sum of three independent quadratic forms. One of these is White's general test for heteroscedasticity and the remaining two components are quadratic forms in the third and fourth powers of the residuals respectively. Our results show that the test will fail to detect serial correlation and never be asymptotically optimal against heteroscedasticity, skewness and non-normal kurtosis.

1. INTRODUCTION

In a recent study of the properties of maximum likelihood estimators when the statistical model is misspecified, White (1982) proposes the information matrix (IM) test for misspecification. This test is based on the fact that if the model is correct then the information matrix is equal to the covariance of the score vector. If it is assumed that the sequences of exogenous and error variables in our model satisfy the Cesaro summability conditions of Gallant and Holly (1980), then the identity can be written as follows,

$$\lim_{n\to\infty} \textstyle\sum_{t=1}^{n} E_t\left[\dfrac{\partial^2 L_t(\theta_0)}{\partial\theta\partial\theta'}\right] = \lim_{n\to\infty} \textstyle\sum_{t=1}^{n} E_t\left[\dfrac{\partial L_t(\theta_0)}{\partial\theta}\cdot\dfrac{\partial L_t(\theta_0)}{\partial\theta'}\right], \tag{1}$$

where L_t is the conditional log likelihood of the t-th observation on the endogenous variable y_t indexed by the unknown parameter vector θ whose true value is θ_0, n is the sample size and $E_t(\cdot)$ denotes expectations taken with respect to the conditional distribution of y_t. Evidence of violation of this information matrix identity is consequently indicative of model misspecification. The familiar forms of the Wald (W) and Lagrange Multiplier (LM) tests are constructed on the assumption that this identity holds, and so as White (1982) observes, the IM test can be used as a preliminary test before conducting inference with the usual procedures.

In this paper we consider the possible implications for our specification if a significant IM test statistic is recorded in the normal linear fixed regressor model. The approach taken is to condition on all but one of the assumptions of the model being correct and to examine in turn whether violation of the remaining assumption would induce a significant statistic. In this way we can compare the performance of the IM test with the appropriate LM (or W) test that is designed to be asymptotically most powerful against this particular misspecification given the validity of the rest of the model.

Our results show that the test statistic derived under the null hypothesis that the model is correctly specified, is asymptotically the sum of three independent components. The first of these is White's direct test for heteroscedasticity (White (1980)). The other two are quadratic forms in the third and fourth powers of the residuals respectively. Our analysis shows that the IM test is sensitive to heteroscedasticity and nonnormality, but does not detect serial correlation. The latter would cause tests based on the assumption of serial independence to be misleading, and so White's conjecture that "it is reasonable to expect the test will be consistent against any alternative which renders the usual maximum likelihood inference techniques invalid" (White (1982), p. 12) would appear to be inappropriate for this model.

Section 2 briefly restates the information matrix test with slight adaptation to allow for independently but not identically distributed random variables. In Section 3 the test statistic for our model is derived, and in Section 4 the potential causes of misspecification to which the test is sensitive are discussed. Finally Section 5 contains some conclusions.

2. THE INFORMATION MATRIX TEST

The IM test is based on a comparison of a subset of the elements of the matrices in equation (1). The model we consider is

$$y_t \sim \mathrm{IN}\,(x_t'\beta, \sigma^2), \qquad t=1,\ldots,n, \tag{2}$$

and so L_t represents the log likelihood function of y_t conditional on the vector of p exogenous variables x_t. Let $\hat{\theta}_n$ be the maximum likelihood estimator of $\theta'=(\beta', \sigma^2)$ from a sample of n observations. The matrices in equation (1) are unobservable, but can be consistently estimated by their sample analogues evaluated at $\hat{\theta}_n$, namely,

$$A_n(\hat{\theta}_n)=n^{-1}\textstyle\sum_{t=1}^{n}\dfrac{\partial^2 L_t}{\partial\theta\partial\theta'}(\hat{\theta}_n),$$

and

$$B_n(\hat{\theta}_n)=n^{-1}\textstyle\sum_{t=1}^{n}\dfrac{\partial L_t}{\partial\theta}(\hat{\theta}_n)\cdot\dfrac{\partial L_t}{\partial\theta'}(\hat{\theta}_n).$$

Due to the symmetry of $A_n(\hat{\theta}_n)$ and $B_n(\hat{\theta}_n)$ a test of the complete IM identity can be based on the lower triangular elements of $\{A_n(\hat{\theta}_n)+B_n(\hat{\theta}_n)\}$. Accordingly define d_t as the vector whose typical element is of the form

$$d_{ts}(\hat{\theta}_n)=\frac{\partial^2 L_t}{\partial\theta_i\partial\theta_j}(\hat{\theta}_n)+\frac{\partial L_t}{\partial\theta_i}(\hat{\theta}_n)\frac{\partial L_t}{\partial\theta_j}(\hat{\theta}_n), \tag{3}$$

where $j\leqq i$, $i=1,2,\ldots,p$ and $s=1,2,\ldots,(p+2)(p+1)/2$. However it may only be desired to test part of the information matrix identity, and so we base the test on the indicator vector,

$$D_n(\hat{\theta}_n)=n^{-1}\textstyle\sum_{t=1}^{n} Sd_t(\hat{\theta}_n)=S\Delta \tag{4}$$

where S is a $q\times(p+2)(p+1)/2$ selection matrix (whose elements are either zero or unity) and Δ is a $(p+2)(p+1)/2$ vector.

Subject to certain regularity conditions, a central limit theorem, (for instance, see White (1984 page 113)), can be applied to show that as n tends to infinity, $\sqrt{n}D_n(\hat{\theta}_n)$ converges to a multivariate normal distribution with mean zero if the model is correctly specified. The original formula for the asymptotic covariance of $\sqrt{n}D_n$ is analytically complicated, requiring third derivatives of the log likelihood function. Chesher (1983) and Lancaster (1984) have shown that these calculations can be simplified as the asymptotic covariance of $\sqrt{n}D_n$ can be consistently estimated by

$$V_n(\hat{\theta}_n) = n^{-1}\sum_{t=1}^{n} Sd_t(\hat{\theta}_n)d_t'(\hat{\theta}_n)S' - [n^{-1}\sum_{t=1}^{n} Sd_t(\hat{\theta}_n)F_{1t}'(\hat{\theta}_n)] \\ \times [n^{-1}\sum_{t=1}^{n} F_{1t}(\hat{\theta}_n)F_{1t}'(\hat{\theta}_n)]^{-1}[n^{-1}\sum_{t=1}^{n} F_{1t}(\hat{\theta}_n)d_t'(\hat{\theta}_n)S'] \tag{5}$$

where

$$F_{1t}(\hat{\theta}_n) = \partial L_t(\hat{\theta}_n)/\partial\theta,$$

if the model is correctly specified. Therefore under the null hypothesis that the model specification is correct, the information matrix test statistic,

$$T_n = nD_n'(\hat{\theta}_n)[V_n(\hat{\theta}_n)]^{-1}D_n(\hat{\theta}_n), \tag{6}$$

is asymptotically distributed χ_q^2. Chesher (1983) has shown that T_n is asymptotically equivalent to n times the uncentered R^2 from the regression of a constant on $Sd_t(\hat{\theta}_n)$ and $F_{1t}(\hat{\theta}_n)$.

3. DERIVATION OF THE TEST STATISTIC FOR THE FIXED REGRESSOR NORMAL LINEAR MODEL

The assumed model specification is as follows

$$y_t = x_t'\beta + u_t,$$

where

$c(i)\, u_t \sim \mathrm{IN}\,(0, \sigma^2), \qquad t = 1, 2, \ldots, n$

$c(ii)\, \lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} x_t x_t' = M,$ a $(p \times p)$ positive definite matrix,

and we further require for the existence of the variance of $\sqrt{n}D_n(\hat{\theta}_n)$ that

$c(iii)\, \lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} x_{it}x_{jt}x_{rt}x_{st}$ is finite for all $i, j, r, s = 1, 2, \ldots, p$.

The log likelihood function of y_t conditional on x_t is

$$L_t = \text{constant} - \tfrac{1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(y_t - x_t'\beta)^2, \tag{7}$$

and performing the appropriate partial differentiation yields

$$\hat{\theta}_n' = (\hat{\beta}_n', \hat{\sigma}_n^2) = (\sum y_t x_t'[\sum x_t x_t']^{-1}, \sum \hat{u}_t^2/n), \tag{8}$$

where $\hat{u}_t = (y_t - x_t'\hat{\beta}_n)$ and all summations are from $t = 1$ to n.

From equation (4) the indicator vector, D_n, can be written as

$$D_n(\hat{\theta}_n) = \begin{bmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_3 \end{bmatrix}\begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{bmatrix}$$

where S is partitioned to conform with $\Delta' = [\Delta_1' : \Delta_2' : \Delta_3']$. By taking first and second derivatives of L_t, in equation (7), with respect to θ it can be seen that for this model the vector Δ is of the following form:

(i) Δ_1 is a $p(p+1)/2 \times 1$ vector comparing the two estimators of the covariance of the $\hat{\beta}_n$ coefficients. Its typical element is of the form

$$n^{-1}\sum_{t=1}^{n}(\hat{u}_t^2 - \hat{\sigma}_n^2)x_{it}x_{jt}/\hat{\sigma}_n^4$$

(ii) Δ_2 is a $(p\times 1)$ vector comparing the estimators of the covariance of $\hat{\beta}_n$ and $\hat{\sigma}_n^2$, and has typical element

$$n^{-1}\sum_{t=1}^n \hat{u}_t^3 x_{it}/2\hat{\sigma}_n^6$$

(iii) Δ_3 is the scalar which compares two estimators of the covariance of $\hat{\sigma}_n^2$, and can be written as

$$n^{-1}\sum_{t=1}^n (\hat{u}_t^4-3\hat{\sigma}_n^4)/4\hat{\sigma}_n^8.$$

To demonstrate that T_n decomposes asymptotically into the sum of three independent quadratic forms under the null hypothesis, H_0, of no misspecification we examine the probability limit of $V_n(\hat{\theta}_n)$ when the model specification is correct. From equation (5) we have

$$\text{plim}_{n\to\infty,H_0} V_n(\hat{\theta}_n)=V(\theta_0)=S\{E_t^*(d_t d_t')-E_t^*(d_t F_{1t}')[E_t^*(F_{1t}F_{1t}')]^{-1}E_t^*(F_{1t}d_t')\}S'$$

where $E_t^*(\cdot)$ denotes the operator $\lim_{n\to\infty} n^{-1}\sum_{t=1}^n E_t(\cdot)$, and all expressions are evaluated at $\theta=\theta_0$.

For this model

$$E_t^*(d_t F_{1t}')=\begin{bmatrix} 0_{r\times p} & m \\ 0_{p\times p} & 0_{p\times 1} \\ 0_{1\times p} & 0_{1\times 1}\end{bmatrix},\qquad E_t^*(F_{1t}F_{1t}')=\begin{bmatrix}\sigma^{-2}M & 0_{p\times 1}\\ 0_{1\times p} & 1/2\sigma^4\end{bmatrix}$$

where the k-th element of m is of the form $\lim_{n\to\infty} n^{-1}\sum_{t=1}^n x_{it}x_{jt}/\hat{\sigma}_n^4$, $0_{a\times b}$ denotes an $(a\times b)$ matrix of zeros and $r=p(p+1)/2$. It therefore follows that

$$V(\theta_0)=S\{E_t^*(d_t d_t')-2\sigma^4 zz'/2\}S'$$

where z is a $(p+1)(p+2)/2$ vector whose first $p(p+1)/2$ elements are the vector m and the remaining elements are zeros.

We can now demonstrate the block diagonality of $V(\theta_0)$. Let d_{it} be the vector given by $\Delta_i=n^{-1}\sum_{t=1}^n d_{it}$ and d_{itj} be the j-th element of d_{it}. It is important to note first that

$$d_{2t}=(\hat{u}_t^2-3\hat{\sigma}_n^2)\hat{u}_t x_t/2\hat{\sigma}_n^6,$$
$$d_{3t}=(3\hat{\sigma}_n^4-6\hat{u}_t^2\hat{\sigma}_n^2+\hat{u}_t^4)/4\hat{\sigma}_n^8,$$

which are consistent with the quoted expressions for Δ_2 and Δ_3 as $\sum_{t=1}^n \hat{u}_t x_t=0$ and $\hat{\sigma}_n^2=n^{-1}\sum_{t=1}^n \hat{u}_t^2$. From the properties of a mean zero normal distribution, namely that: (a) $E(u_t^a)=0$, for $a=3$, 5, 7; (b) $E(u_t^6)=15\sigma^6$; and using the independence of x_t and u_t it can be shown that, for $j=1,2,\ldots,p(p+1)/2$ and $k=1,2,\ldots,p$:

(i) $$\text{plim}_{n\to\infty,H^0} n^{-1}\sum_{t=1}^n d_{1tj}d_{2tk}=\text{plim}_{n\to\infty} n^{-1}\sum_{t=1}^n(\hat{u}_t^2-\hat{\sigma}_n^2)\times(\hat{u}_t^3-3\hat{\sigma}_n^2\hat{u}_t)x_{it}x_{jt}x_{kt}/2\hat{\sigma}_n^{10}=0,$$

(ii) $$\text{plim}_{n\to\infty,H_0} n^{-1}\sum_{t=1}^n d_{1tj}d_{3t}=\text{plim}_{n\to\infty} n^{-1}\sum_{t=1}^n(\hat{u}_t^2-\hat{\sigma}_n^2)\times(\hat{u}_t^4-6\hat{u}_t^2\hat{\sigma}_n^2+3\hat{\sigma}_n^4)x_{it}x_{jt}/4\hat{\sigma}_n^{12}=0,$$

(iii) $$\text{plim}_{n\to\infty,H_0} n^{-1}\sum_{t=1}^n d_{2tk}d_{3t}=\text{plim}_{n\to\infty} n^{-1}\sum_{t=1}^n(\hat{u}_t^4-6\hat{u}_t^2\hat{\sigma}_n^2+3\hat{\sigma}_n^4)\times(\hat{u}_t^3-3\hat{\sigma}_n^2\hat{u}_t)x_{kt}/8\hat{\sigma}_n^{14}$$
$$0.$$

This implies that if the model specification is correct then the limiting covariance matrix of $\sqrt{n}D_n$ can be consistently estimated by a block diagonal matrix. Therefore the IM test is asymptotically equivalent to the sum of the following three statistics,

(i) $T_{1n}=[\sum_{t=1}^{n}(\hat{u}_t^2-\hat{\sigma}_n^2)\xi_t']S_1'[S_1\sum_{t=1}^{n}\xi_t\xi_t'S_1']^{-1}S_1[\sum_{t=1}^{n}(\hat{u}_t^2-\hat{\sigma}_n^2)\xi_t]/2\hat{\sigma}_n^4$,

where ξ_t is a $p(p+1)/2$ vector consisting of the lower triangular elements of $(x_tx_t'-n^{-1}\sum_{t=1}^{n}x_tx_t')$,

(ii) $T_{2n}=[\sum_{t=1}^{n}\hat{u}_t^3x_t']S_2'[S_2\sum_{t=1}^{n}x_tx_t'S_2']^{-1}S_2[\sum_{t=1}^{n}\hat{u}_t^3x_t]/6\hat{\sigma}_n^6$.

(iii) $T_{3n}=n^{-1}[\sum_{t=1}^{n}(\hat{u}_t^4-3\hat{\sigma}_n^4)]^2/24\hat{\sigma}^8$.

These three statistics contain tests familiar in the literature. The first, T_{1n}, is identical to the LM test against heteroscedasticity of the form

$$\sigma_t^2=h(\alpha_0+\textstyle\sum_{j\leq i}^{p}\alpha_r s_{1r}x_{it}x_{jt}), \tag{9}$$

where s_{1r} takes the value one or zero depending on whether the appropriate element of d_{1t} was included in the test, (see Breusch and Pagan (1979)). Under the assumption that u_t is a sequence of serially uncorrelated normal random variables, T_{1n} is also asymptotically equivalent to White's (1980) direct test for heteroscedasticity. In the special case in which $x_{1t}=1$, for all t, and $S_2=(1,0,0,\ldots,0)$ and $S_3=1$, then T_{2n} and T_{3n} are asymptotically equivalent to the tests for skewness and kurtosis suggested by Bowman and Shenton (1975) and the LM tests for normality based on (i) the Edgeworth expansion (Keifer and Salmon (1983)) and (ii) the Pearson family of distributions (Bera and Jarque (1980)).

4. ANALYSIS OF THE INFORMATION MATRIX TEST AS A TEST OF MISSPECIFICATION

The asymptotic decomposition of our test statistic into the sum of three components facilitates an analysis of the model misspecifications to which it is sensitive. The first component, T_{1n}, is the locally optimal test against heteroscedasticity of the form in equation (9). Koenker (1981) has shown that the Breusch Pagan LM statistic is distributed asymptotically as a noncentral χ^2 under local alternatives of the form,

$$\sigma_t^2=\sigma^2(1+g(z_t'\gamma_0/\sqrt{n})). \tag{10}$$

These arguments are extended in Hall (1982) to show that the LM test is asymptotically distributed as a noncentral χ^2, when the assumed heteroscedasticity specification of the test, $[\sigma_t^2]_{H_1}$, and the true σ_t^2 specification, $[\sigma_t^2]_T$, do not coincide. As one would expect, the greater the discrepancy between $[\sigma_t^2]_{H_1}$ and $[\sigma_t^2]_T$, the greater the loss of power of the test. This effect is accentuated by the inclusion of the remaining components. Using Amemiya's (1978) residual decomposition and similar arguments to Koenker, Hall (1982) has shown that T_{2n} and T_{3n} are distributed asymptotically as central χ^2 variates with p and 1 degrees of freedom respectively under the alternative in (10). Further the three quadratic forms are asymptotically independent under this alternative. Therefore the effect of including T_{2n} and T_{3n} in the test is to increase the degrees of freedom of the test but leave the noncentrality parameter unaltered. This causes a loss of power that increases with the number of regressors.[1]

None of the components is sensitive to serial correlation. The statistic T_{1n} is based on a comparison between two estimators of the variance of $\hat{\beta}_n$, but both are inconsistent in the presence of serial correlation and both converge to the same probability limit in this case. The other two components are similarly insensitive implying the information matrix test has power equal to its size against serial correlation. However all three components are sensitive to non-normality. The quadratic forms T_{2n} and T_{3n} are sensitive to skewness and non-normal kurtosis of the errors respectively, although the appropriate form of the critical region requires investigation. Koenker notes that the Breusch Pagan LM statistic is sensitive to non-normal kurtosis, and suggests "studentizing" the test by dividing the quadratic form by $n^{-1}\sum_{t=1}^{n}(\hat{u}_t^2-\hat{\sigma}_n^2)^2$ rather than $2\hat{\sigma}_n^4$ to ensure robustness against non-normal errors. If this alteration is applied to T_{1n} then the statistic becomes White's (1980) direct test for heteroscedasticity, which in turn implies the latter is the "studentized" LM test against the alternative in (10).

It is interesting to compare our results with two recent papers in the specification testing literature by Chesher (1984) and Bera and Jarque (1982). Chesher (1984) has shown that the IM test can be derived as an LM test against parameter variation. This interpretation is clearly reflected by our derived indicator vector. In Chesher's framework the nature of the parameter variation under the alternative is essentially unrestricted apart from some regularity conditions. Our analysis of the test's properties against particular heteroscedasticity specifications suggests the complete IM test may not be the most efficient method of detecting parameter variation in our model. It also restates the importance of seeking any additional information on the nature of the parameter variation, which may allow the development of more powerful tests: either through a reduction of the indicator vector to include only those elements deemed appropriate or through the construction of different tests. Bera and Jarque (1982) propose a general test of model adequacy which is a linear combination of the LM tests against heteroscedasticity, serial correlation, incorrect functional form, skewness and non normal kurtosis of the errors. Our analysis in section 3 shows that an appropriate choice of the selection matrix, S, allows three of these components to be derived from the IM test principle. It may be thought that the exclusion of the remaining two LM tests is due to the nature of our assumed model. However it is worth noting that had our original specification included first order autoregressive errors, then the IM test does not decompose asymptotically into the sum of our original three component test (derived in Section 3) plus the LM test against first-order serial correlation. In this more general framework the indicator vector no longer has a block diagonal covariance matrix due to the inclusion of the autoregressive coefficient in the parameter vector.

5. CONCLUSIONS

Any test of model adequacy can broadly be interpreted in two ways: as a pure significance test or a constructive test. With a pure significance test rejection of the null hypothesis implies only that some aspect of the model specification is incorrect. A constructive test can be interpreted as a test of one aspect of that specification because it is known to be asymptotically most powerful against a particular class of alternatives given the validity of the rest of the model. The form of the IM test used in this paper would be considered a pure significance test, as violation of the IM identity is indicative of model misspecification, but is not immediately interpretable as the consequence of one erroneous assumption in the manner of a constructive test. We have shown that the IM test asymptotically decomposes into the sum of three independent statistics and described the particular misspecifications against which each will be powerful. The insensitivity of each component to serial correlation suggests that the IM test alone is unlikely to be a satisfactory check of model adequacy.

Acknowledgement. This research was undertaken during the tenure of a Social Science Council studentship and is a revision of Warwick Economic Research Paper No. 250, University of Warwick.

I am greatly indebted to Grayham Mizon for his very valuable comments which, along with those of two anonymous referees, have greatly improved the presentation of the paper. I have also benefitted from discussions with John Aldrich, Trevor Breusch, Les Godfrey, Ray O'Brien, Mark Salmon and Ken Wallis on various aspects of the work.

An earlier version of this paper was presented at the 1983 European Meeting of the Econometric Society held at Pisa.

NOTE

1. This follows because if $\chi_{p_i}(m_i)$ are a sequence of independent χ^2 random variables with degrees of freedom, p_i and noncentrality parameter m_i, then $\sum \chi_{p_i}(m_i) = \chi_{p^*}(m^*)$ where $m^* = \sum m_i$, $p^* = \sum p_i$.

REFERENCES

AMEMIYA, T. (1977), "A Note on the Heteroscedastic Model", *Journal of Econometrics*, **6**, 365-370.

BERA, A. K. and JARQUE, C. M. (1980) "An Efficient Large-Sample Test for Normality of Observations and Regression Residuals" (mimeo, Australian National University).

BERA, A. K. and JARQUE, C. M. (1982), "Model Specification Tests: A Simultaneous Approach" (Paper presented at Third Latin American Meeting of the Econometric Society, Mexico City).

BOWMAN, K. O. and SHENTON, L. R. (1975) "Omnibus Contours for Departures from Normality Based on b_1 and b_2", *Biometrika*, **62**, 243-250.

BREUSCH, T. S. and PAGAN, A. R. (1979), "A Simple Test for Heteroscedasticity and Random Coefficient Variation", *Econometrica*, **47**, 1287-1294.

CHESHER, A. (1983), "The Information Matrix Test: Simplified Calculation via a Score Test Interpretation", *Economic Letters*, **13**, 45-48.

CHESHER, A. (1984), "Testing for Neglected Heterogeneity", *Econometrica*, **52**, 865-872.

GALLANT, A. R. and HOLLY, A. (1980), "Statistical Inference in an Implicit, Nonlinear Simultaneous Equation Model in the Context of Maximum Likelihood Estimation", *Econometrica*, **48**, 697-720.

HALL, A. R. (1982), "The Information Matrix Test for the General Linear Model" (unpublished MSc dissertation, University of Southampton).

KEIFER, N. M. and SALMON, M. (1983) "Testing Normality in Econometric Models", *Economics Letters*, **11**, 123-127.

KOENKER, R. (1981), "A Note on Studentising a Test of Heteroscedasticity", *Journal of Econometrics*, **17**, 107-112.

LANCASTER, T. (1984), "The Covariance Matrix of the Information Matrix Test", *Econometrica*, **52**, 1051-1054.

WHITE, H. (1980), "A Heteroscedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity", *Econometrica*, **48**, 817-838.

WHITE, H. (1982), "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, **50**. 1-26.

WHITE, H. (1984) *Asymptotic Theory for Econometricians* (New York: Academic Press).