CONSISTENT OLS COVARIANCE ESTIMATOR AND
MISSPECIFICATION TEST FOR MODELS WITH
STATIONARY ERRORS OF UNSPECIFIED FORM

Howard E. Doran

No. 51 - May 1991
1. Introduction

When ordinary least squares is applied to a time series linear model, correct inference depends crucially on there being no autocorrelation in the disturbance. As standard textbook discussions show, autocorrelated disturbances result in inefficient but consistent parameter estimates. However, the estimated covariance matrix of the parameters is not consistent. When positive autocorrelation is present, standard errors are likely to be too small, leading to overly optimistic confidence intervals, and the 'discovery' of non-existent economic relationships [see Pierce (1977)]. If ordinary least squares is to be used when autocorrelated disturbances are a possibility, then correct inference depends on finding an estimator of the covariance matrix which is consistent when autocorrelation is present.

In considering the analogous problem in the case of heteroskedastic error models, White (1980) developed a consistent covariance estimator and a test of misspecification which do not depend on the particular form of heteroskedasticity.

In this paper White's approach is followed, and a consistent covariance estimator and misspecification test for the autocorrelated error model are provided. This is achieved by transforming to the frequency domain. This transformation converts a stationary time series error model to a heteroskedastic error model, and White's results are immediately available. Subsequent transformation back to the time domain yields a covariance estimator and misspecification test which do not require a parametric specification of the error process. The covariance estimator is thus more specialized than that obtained by White and Domowitz (1984) for a much more general error process.
The plan of the paper is as follows. In Section 2 the covariance estimator is developed, and in Section 3, a misspecification test is proposed. The work in these sections depends on asymptotic results. In order to assess the usefulness in small samples, a limited series of Monte Carlo experiments was conducted, with results reported in Sections 4 and 5. Section 6 is the conclusion.

For readers who are not familiar with frequency domain analysis, an Appendix contains the results which are used in Sections 2 and 3.

2. A Consistent Covariance Estimator

Let us consider the regression model

\[ y = X\beta + u, \]  

where \( y \) and \( u \) are \( T \)-vectors and \( X \) is \( T \times K \). We assume \( E(u) = 0, E(X'u) = 0 \) and \( E(uu') = \Omega. \) \hspace{1cm} (2.2)

We will further assume that the data are time series, and that \( u \) is generated by a covariance stationary ergodic process. It then follows that

\[ \Omega_{k\ell} = \gamma_{\tau}, \]  

where \( \tau = |k-\ell|, (\tau = 0,1,\ldots,T-1) \), and \( \gamma_{\tau} = E(u_t u_{t-\tau}). \) \hspace{1cm} (2.3)

As shown in the Appendix, the model may be expressed in the frequency domain by pre-multiplication by the matrix \( U \), defined in (A6). Thus,

\[ z = Z\beta + v, \]

where \( z = Uy, Z = UX \) and \( v = Uu. \)

The matrices \( z, Z \) and \( v \) are \( T \times 1, T \times K \) and \( T \times 1 \), respectively and are complex valued. Also, by (2.2), the covariance matrix of \( v \), denoted by \( \Phi \), is given
by

\[ \phi = E(vv^*) = U \Omega U^*, \]

where the symbol '*' denotes conjugate-transposition.

When T is large,

\[ \phi = 2\pi \text{diag}\left[f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_T)\right], \tag{2.4} \]

approximately, where \( f(\lambda) \) is the spectrum of the \( u \) process at frequency \( \lambda \).

The implication of (2.4) is that a covariance stationary error model in the time domain, becomes approximately a heteroskedastic error model in the frequency domain. The form of the heteroskedasticity is not completely general, however, because the symmetry of the spectrum about \( \lambda = 0 \) implies that \( f(\lambda_{2n-j}) = f(\lambda_j) \), where \( n = \lfloor T/2 \rfloor \), and hence, the error structure is completely described by only \( \lfloor T/2 \rfloor + 1 \) parameters.

The above property of approximate heteroskedasticity suggests that the ideas developed by White (1980) for consistent ordinary least squares (ols) covariance estimation and specification testing of heteroskedastic models can also be applied to stationary time series error models expressed in the frequency domain.

Following White, we define an estimator \( \hat{V} \) of the covariance of \( 4\hat{T}(\hat{\beta} - \beta) \), where \( \hat{\beta} \) is the ols estimator, by

\[ \hat{V} = \left( Z^*Z \right)^{-1} \left( Z^* \hat{\phi} Z \right) \left( Z^*Z \right)^{-1}, \tag{2.5} \]

where

\[ \hat{\phi} = \text{diag}(p_1, p_2, \ldots, p_T). \tag{2.6} \]

In the above, \( p_j \) is the \( j \)-th periodogram ordinate of \( \hat{u} \), given by
The $\lambda_j = -\pi + \pi j/n$.

The $p_j$ are the frequency domain analogues of the squared residuals $\hat{u}_j^2$ used in White's estimator. The estimator $\hat{V}$ can now be expressed in terms of time domain quantities. We consider first the matrix $T^{-1}(Z^* \hat{\Theta} Z)$.

$$T^{-1}(Z^* \hat{\Theta} Z) = T^{-1}(X'U^* \hat{\Theta} U X)$$
$$= T^{-1}(X'\hat{\Omega} X), \text{ say}$$

where

$$\hat{\Omega} = U^* \hat{\Theta} U.$$

The $(k, \ell)$ element of $\hat{\Omega}$ is given by

$$\hat{\Omega}_{k\ell} = \sum_{j=1}^{2n} \sum_{m=1}^{2n} U^*_j \hat{\Theta}_{jm} U_{ml}$$
$$= \sum_{j=1}^{2n} \sum_{m=1}^{2n} \hat{U}_{jk} \hat{\Theta}_{jm} U_{ml}$$
$$= \sum_{j=1}^{2n} \hat{U}_{jk} p_j U_{j\ell}$$
$$= \frac{1}{T} \sum_{j=1}^{2n} p_j e^{-i\lambda_j(k-\ell)}.$$

Thus, as far as $k$ and $\ell$ are concerned, $\hat{\Omega}_{k\ell}$ depends only on $(k-\ell)$. Also, from the symmetry of $p_j$ about $\lambda_n$ (that is, zero frequency), it is easily shown that $\hat{\Omega}_{\ell k} = \hat{\Omega}_{k\ell}$. It follows that $\hat{\Omega}_{k\ell}$ depends only on $\tau = |k-\ell|$, ($\tau = 0, 1, \ldots, T-1$) and we have
\[ \hat{\Omega}_{kl} = \frac{1}{T} \sum_{j=1}^{2n} p_j e^{-i\lambda_j T}. \] (2.9)

That is,

\[ \hat{\Omega}_{kl} = \frac{1}{T} \left\{ p(0) + 2 \sum_{j=1}^{n-1} p_j \cos \mu_j T + p_{2n} \cos \pi T \right\}, \] (2.10)

where

\[ p(0) = T u^2 \]

and

\[ \mu_j = \pi j / n. \]

If the model contains an intercept, then p(0) is identically zero. Returning to (2.5),

\[ Z^*Z = X'U^*U X = X'X, \]

and so, in terms of the original (time domain) observations

\[ \hat{\nu} = \left( \frac{X'X}{T} \right)^{-1} \left( \frac{X'\hat{\Omega} X}{T} \right) \left( \frac{X'X}{T} \right)^{-1}, \] (2.11)

with the elements of \( \hat{\Omega} \) being given by (2.10).

The expression (2.9), which defines the elements of \( \hat{\Omega} \), is capable of a very easy interpretation. We have from (2.3) that \( \Omega_{kl} = \gamma_{\tau} \). From spectral theory, the autocovariance \( \gamma_{\tau} \) can be written in the form

\[ \gamma_{\tau} = \int_{-\pi}^{\pi} f(\lambda) e^{-\lambda \tau} d\lambda, \] (2.12)

where \( f(\lambda) \) is the spectrum of the disturbance process at frequency \( \lambda \). If \( f(\cdot) \) were known at the T frequency 'points' \( \lambda_j \) (\( j = 1, 2, \ldots, T \)) then the integral (2.12) could be approximated by a sum to yield

\[ \gamma_{\tau} \approx \sum_{j=1}^{T} f(\lambda_j) e^{-i\lambda_j \tau} \delta \lambda_j \]
Thus,

\[ \gamma_\tau \approx \frac{1}{T} \sum_{j=1}^{T} \left[ 2\pi f(\lambda_j) \right] e^{-\lambda_j \tau}. \]  

By comparing (2.9) and (2.13) it is clear that \((X'\Omega X)\) is an estimator of \((X'\Omega X)\) obtained by approximating the integral of (2.12) by a sum, and replacing \(2\pi f(\lambda_j)\) by the corresponding periodogram ordinate \(p_j\). As \(p_j\) is an asymptotically unbiased, but inconsistent estimator of \(2\pi f(\lambda_j)\) [see Harvey (1981, 84-85)] the analogy with White's estimator which replaces \(\sigma_1^2\) by \(\hat{\sigma}_1^2\) is apparent.

3. A Misspecification Test

For the case of the heteroskedastic error model, White (1980) proposed a convenient misspecification test. Let us suppose the regression matrix of (2.1) is expressed in the form

\[ X = \begin{bmatrix} X_1, X_2, \ldots, X_k \end{bmatrix} \]

where each \(X_j\) is a T-vector. There are \(S = K(K+1)/2\) different product pairs of the type \(X_k X_\ell\). If the particular combination \((k, \ell)\) is indexed by the integer \(r (r = 1, 2, \ldots, S)\), we may define a T-vector \(\Psi_r\) whose \(j\)-th element is given by

\[ \Psi_{rj} = X_{kj} X_{\ell j}, \quad (j = 1, 2, \ldots, T). \]  

White showed that, assuming
Some general boundedness conditions on the regressors; $u_j$ is independent of $X_{kj}$ for all $j,k$; and $E(u_j^4)$ does not depend on $j$ under the null hypothesis of no heteroskedasticity, then if the artificial regression

$$\hat{u}_j^2 = \alpha_0 + \alpha_1 \psi_{1j} + \ldots + \alpha_S \psi_{Sj} \tag{3.2}$$

is performed,

$$T \hat{R}_2 \sim \chi^2_S \tag{3.3}$$

under the null hypothesis.

This test can be used as a general test for autocorrelation in the time series error model by using the frequency domain representation.

Here we define $\psi_r$ (complex) by

$$\psi_{rj} = Z_{kj} \tilde{Z}_{lj}, \quad \tag{3.4}$$

carry out the regression

$$p_j = \alpha_0 + \alpha_1 \psi_{1j} + \ldots + \alpha_S \psi_{Sj} \tag{3.5}$$

and use (3.3) to test departures from a white-noise error process. There is one minor modification which is needed. When the original model (2.1) has an intercept, because the FFT of a column of 1's has the form $(\sqrt{T}, 0, 0, \ldots, 0)'$ [see (A1)] all $\psi$'s of the form $Z_{1j} \tilde{Z}_{kj}$ ($k = 1, 2, \ldots, K$) will be perfectly collinear, and hence only one such $\psi$ is kept in the regression (3.5). This means the parameter $S$ is redefined as

$$S = K(K-1)/2 + 1, \tag{3.6}$$

and the $\psi$'s are given by

$$\psi_{1j} = |Z_{1j}|^2$$
Although the regressors $\psi_r$ are complex-valued, the regression (3.5) can be carried out using standard regression packages. Because $\lambda_{2n-j} = -\lambda_j$ it follows that

$$Z_{k,2n-j} = \tilde{Z}_{kj}.$$ 

Hence,

$$\psi_{r,2n-j} = Z_{k,2n-j} \tilde{Z}_{l,2n-j} = \tilde{Z}_{kj} Z_{lj} = \psi_{r,j} \quad (j = 1,2,\ldots,n-1).$$

Thus, apart from $j = n$ ($\lambda = 0$) and $j = 2n$ ($\lambda = \pi$) (at which frequencies the $\psi_r$ are all real), the observations on the regressors occur in conjugate pairs.

Setting

$$\psi_{r,j} = a_{r,j} + i b_{r,j},$$

it can easily be verified that if the regression (3.5) is set up as

$$\begin{bmatrix}
  p_n \\
  p_{2n} \\
  \vdots \\
  \sqrt{2} p_1 \\
  \vdots \\
  \sqrt{2} p_{n-1} \\
  0 \\
  \vdots \\
  0
\end{bmatrix} = \begin{bmatrix}
  1 & \psi_{rn} \\
  1 & \psi_{r,2n} \\
  \sqrt{2} & \sqrt{2} a_{r,1} \\
  \vdots & \vdots \\
  \sqrt{2} & \sqrt{2} a_{r,n-1} \\
  0 & \sqrt{2} b_{r,1} \\
  \vdots & \vdots \\
  0 & \sqrt{2} b_{r,n-1}
\end{bmatrix}$$

(3.7)

the error sum of squares (SSE) is unaltered. The total sum of squares (SST) is calculated directly from the $p_j$ ($j = 1,2,\ldots,2n$). Then,
\( T(1 - \text{SSE/SST}) \sim \chi^2 \) under \( H_0 \).

This procedure, which does not require a parameterization of the alternative hypothesis, can be seen as an alternative to the Portmanteau Q-test [Box and Pierce (1970) and Ljung and Box (1978)]. The Q-test has an element of arbitrariness in that the number of autocorrelations \( M \) which are used must be chosen in advance. The test procedure suggested here has no such disadvantage.

4. Monte Carlo Experiment - Covariance Estimator

The theory underlying the ols covariance estimator given in (2.11) and (2.10) is asymptotic. When applied to small samples, this estimator will almost certainly be biased, with biases being introduced from any (or all) of at least the four following sources:

(i) The matrix \( U \Omega U^* \) is only asymptotically diagonal;
(ii) The elements \( \hat{\Omega}_{k\ell} \) involve approximating an integral by a summation;
(iii) \( p_j \) is only an asymptotically unbiased estimator of \( 2\pi f(\lambda_j) \) when the disturbance process is not white noise; and
(iv) When the model contains an intercept, the periodogram ordinate corresponding to zero frequency (\( p_n \)) is identically zero. Thus, the summation which is used to approximate the integral has no contribution at \( \lambda=0 \). The resulting bias will be most severe when the disturbance process has most of its power in the neighbourhood of \( \lambda = 0 \). Such a process is an AR(1), with \( \rho \) close to +1.

In order to assess the usefulness of the covariance estimator, a very limited Monte Carlo experiment was undertaken.
Observations were generated from the model

\[ y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t, \]

where

\[ \beta_0 = \beta_1 = \beta_2 = 1; \]
\[ X_{1t} = \exp\left[1.56 + 0.05t + \epsilon_{1t}\right], \]
\[ \epsilon_{1t} \sim N(0, 0.015^2), \]
\[ X_{2t} = X_{2,t-1} + 0.1 + \epsilon_{2t}, \]
\[ X_{2,0} = 7, \]
\[ \epsilon_{2t} \sim N(0, 0.3^2). \]

Thus, \( X_{1t} \) represents a series with a growth rate of 5 per cent, and \( X_{2t} \) a random walk with a small upward drift. The regressors were held fixed in the replications. The disturbance \( u_t \) was generated from an AR(1) process, with values \( \rho = 0\)(0.2)0.8 and 0.99. The case \( \rho = 0.99 \) was included to observe behaviour at the boundary of stationarity. In every case \( T = 40 \) and 100 replications were generated.

The results for the standard errors of \( \hat{\beta}_0, \hat{\beta}_1 \) and \( \hat{\beta}_2 \) are given in Table 1 and those for the full covariance matrix in Table 2.

An examination of Table 1 in particular reveals the results to be (at least for this experiment) quite encouraging. As expected, the standard errors are biased and the direction of bias is consistently downwards. However, the magnitudes of the biases are not large, being of the order of 5 per cent when \( \rho = 0 \) and rising to about 10 per cent when \( \rho = 0.8 \). When \( \rho = 0.99 \) the magnitude rises to about 70 per cent for the intercept and 30 per cent for \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \).
Table 1

Estimates of standard deviations of ols estimators $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\hat{\beta}_0$</th>
<th>true standard deviation</th>
<th>average of estimates of standard deviation</th>
<th>standard error of average</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7.670</td>
<td>7.235</td>
<td>0.128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1560</td>
<td>0.1469</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1601</td>
<td>1.0961</td>
<td>0.0193</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>8.4751</td>
<td>7.8104</td>
<td>0.147</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1730</td>
<td>0.1592</td>
<td>0.0030</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.2814</td>
<td>1.1825</td>
<td>0.0223</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>9.7083</td>
<td>8.4192</td>
<td>0.191</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1989</td>
<td>0.1713</td>
<td>0.0040</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.4664</td>
<td>1.2723</td>
<td>0.0289</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>11.5566</td>
<td>10.3333</td>
<td>0.277</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2382</td>
<td>0.2123</td>
<td>0.0058</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.7422</td>
<td>1.5599</td>
<td>0.0419</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>15.007</td>
<td>13.520</td>
<td>0.401</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.3141</td>
<td>0.2805</td>
<td>0.0086</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.2302</td>
<td>2.0376</td>
<td>0.0605</td>
<td></td>
</tr>
<tr>
<td>0.99</td>
<td>22.992</td>
<td>16.816</td>
<td>0.589</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5182</td>
<td>0.3528</td>
<td>0.0128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.5286</td>
<td>2.5297</td>
<td>0.0884</td>
<td></td>
</tr>
</tbody>
</table>
Table 2
Estimates of Covariance Matrix of ols estimators $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$

<table>
<thead>
<tr>
<th></th>
<th>True Covariance Matrix</th>
<th>Average Estimated Covariance Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\beta}_0$</td>
<td>$\hat{\beta}_1$ $\hat{\beta}_2$</td>
</tr>
<tr>
<td></td>
<td>58.83</td>
<td>1.140 0.024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.95)</td>
</tr>
<tr>
<td>$\rho = 0$</td>
<td>$\hat{\beta}_2$</td>
<td>-8.866 -0.176 1.345</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.295) (0.006) (0.045)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_0$</td>
<td>71.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.48)</td>
</tr>
<tr>
<td>$\rho = 0.2$</td>
<td>$\hat{\beta}_2$</td>
<td>1.379 0.030</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.048) (0.001)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_0$</td>
<td>-10.81 -0.215 1.642</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.375) (0.007) (0.057)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_0$</td>
<td>94.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.42)</td>
</tr>
<tr>
<td>$\rho = 0.4$</td>
<td>$\hat{\beta}_2$</td>
<td>1.784 0.040</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.066) (0.002)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}_0$</td>
<td>-14.14 -0.280 2.150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.520) (0.010) (0.078)</td>
</tr>
</tbody>
</table>
The standard errors of the average covariances are given in parentheses.

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.0259)</td>
<td>(0.069)</td>
<td>(0.049)</td>
<td>(0.1262)</td>
<td>(0.045)</td>
<td>(0.1262)</td>
</tr>
<tr>
<td>7.173</td>
<td>0.931</td>
<td>12.451</td>
<td>1.224</td>
<td>-0.36</td>
<td>0.927</td>
</tr>
<tr>
<td>0.141</td>
<td>0.866</td>
<td>0.269</td>
<td>3.286</td>
<td>1.826</td>
<td>0.098</td>
</tr>
<tr>
<td>31.709</td>
<td>228.63</td>
<td>0.099</td>
<td>0.019</td>
<td>0.813</td>
<td>0.067</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.269)</td>
<td>(0.035)</td>
<td>(1.17)</td>
<td>0.588</td>
<td>2.974</td>
<td>4.974</td>
</tr>
<tr>
<td>4.514</td>
<td>2.232</td>
<td>0.278</td>
<td>0.617</td>
<td>-0.22</td>
<td>0.747</td>
</tr>
<tr>
<td>0.708</td>
<td>3.019</td>
<td>6.994</td>
<td>3.694</td>
<td>0.098</td>
<td>0.019</td>
</tr>
<tr>
<td>198.70</td>
<td>225.21</td>
<td>0.098</td>
<td>0.019</td>
<td>0.813</td>
<td>0.067</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.137)</td>
<td>(0.019)</td>
<td>(0.047)</td>
<td>(0.038)</td>
<td>(0.057)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>2.608</td>
<td>0.340</td>
<td>17.21</td>
<td>3.035</td>
<td>-0.389</td>
<td>0.900</td>
</tr>
<tr>
<td>0.08</td>
<td>0.124</td>
<td>0.185</td>
<td>0.241</td>
<td>0.451</td>
<td>0.057</td>
</tr>
<tr>
<td>0.418</td>
<td>114.38</td>
<td>133.56</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
</tbody>
</table>

The covariance matrix is estimated covariance matrix

Table 2 (cont.)
Thus, except at the boundary of stationarity, the biases are not large enough to seriously jeopardize inference.

In a discussion of the nature and extent of the bias for the heteroskedasticity consistent covariance estimator, Chesher and Jewitt (1987) conclude that where the design matrix $X$ has points of high leverage the estimator can be very severely biased downwards. In our Monte Carlo experiment, the high leverage 'point' is zero frequency, which is typical of economic data which often exhibit smoothness, due to trends or other long-run behavior. However, the zero frequency component only affects the intercept $\hat{\beta}_0$. Thus, at least as far as the slope coefficients are concerned, bias may not be as severe a problem with the time series error model.

One final matter should be mentioned. The regressors $X_{1t}$ and $X_{2t}$ have spectra which are unbounded at $\lambda = 0$. Thus, the model used does not fulfill White's conditions of boundedness. Despite this, the results are very reasonable. It appears that the covariance estimator may not be very sensitive to departures from these assumptions. Clearly, much more investigation could be done on this point.

5. Monte Carlo Experiments - Specification Test

In order to evaluate size and power characteristics of the test described in Section 3, two Monte Carlo experiments were performed. In the first experiment samples were generated from the same model as described in Section 4. Then, secondly, a lagged dependent variable was added to the model, with coefficient $\gamma$ taking two values $\gamma = 0.4$ and $\gamma = 0.8$. In both experiments sample size was fixed at $T = 40$ and 1000 replications were performed.
(a) **Fixed Regressors Experiment**

The frequency domain test, denoted by FQD, was compared with the following four standard tests:

1. A Lagrange Multiplier (LM) test assuming an AR(1) alternative, denoted by LMI;
2. A LM test, assuming an AR(2) alternative (LM2);
3. The Portmanteau test with M = 10 (Q10); and
4. The Portmanteau test with M = 20 (Q20).

We note that because the disturbance generating mechanism was in fact AR(1), the test LMI would be expected to perform best.

The empirical sizes, for a nominal size $\alpha = 0.05$ are given in Table 3. It appears that for samples as small as $T = 40$, FQD and LMI have the correct size, whereas the other tests have empirical size larger than the nominal size. This is particularly the case for the Portmanteau tests Q10 and Q20.

In order to compare power, the critical values were chosen to achieve an empirical size of 0.05 for each test. The results of this comparison are given in Table 4. The Lagrange Multiplier tests LMI and LM2 are clearly superior. FQD is slightly more powerful than Q10 and Q20 for $\rho \leq 0.4$ and slightly less powerful when $\rho > 0.4$.

When account is taken of the arbitrariness of the Q tests and the relatively poor size characteristics, the evidence from this experiment suggests that the FQD test may be superior to the Portmanteau tests.

In addition to the above experiment, samples with a first order moving average disturbance, with positive parameter, were generated. The power characteristics of the FQD test were extremely poor in this case - far worse than for the other tests. From this evidence it is inferred that the FQD test has no power against alternatives in which the disturbance spectral
### Table 3

Empirical sizes when $\alpha = 0.05$

Fixed regressor model, $T = 40$, 1000 replications$^a$

<table>
<thead>
<tr>
<th>Test</th>
<th>Empirical size</th>
</tr>
</thead>
<tbody>
<tr>
<td>FQD</td>
<td>0.044 (0.007)</td>
</tr>
<tr>
<td>LM1</td>
<td>0.065 (0.008)</td>
</tr>
<tr>
<td>LM2</td>
<td>0.097 (0.009)</td>
</tr>
<tr>
<td>Q10</td>
<td>0.101 (0.010)</td>
</tr>
<tr>
<td>Q20</td>
<td>0.116 (0.010)</td>
</tr>
</tbody>
</table>

$^a$ Standard errors are given in parentheses.

### Table 4

Power comparisons, empirical size 0.05

Fixed regressor model, $T = 40$, 1000 replications

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>FQD</th>
<th>LM1</th>
<th>LM2</th>
<th>Q10</th>
<th>Q20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.2</td>
<td>0.123</td>
<td>0.084</td>
<td>0.060</td>
<td>0.091</td>
<td>0.093</td>
</tr>
<tr>
<td>0.4</td>
<td>0.256</td>
<td>0.400</td>
<td>0.213</td>
<td>0.240</td>
<td>0.244</td>
</tr>
<tr>
<td>0.6</td>
<td>0.412</td>
<td>0.767</td>
<td>0.602</td>
<td>0.545</td>
<td>0.514</td>
</tr>
<tr>
<td>0.8</td>
<td>0.616</td>
<td>0.989</td>
<td>0.847</td>
<td>0.783</td>
<td>0.758</td>
</tr>
<tr>
<td>0.99</td>
<td>0.741</td>
<td>0.972</td>
<td>0.931</td>
<td>0.847</td>
<td>0.831</td>
</tr>
</tbody>
</table>
power is concentrated in the high frequencies. However, this situation is unlikely to occur with economic data, and even if it does, its effects on OLS estimators are not serious. We conclude that this defect of the FQD test is unlikely to be of practical significance.

(b) Lagged Dependent Variable Experiment

In this context Portmanteau tests are not valid [see Breusch and Pagan (1980) and Godfrey (1988, 121)], and they were replaced by one- and two-sided Durbin h-tests, denoted by h1 and h2, respectively. The size characteristics are shown in Table 5 and power comparisons in Table 6.

It appears that when a lagged dependent variable is introduced into the model, the sizes of all tests are incorrect, being too large for every test except h1. When $\gamma = 0.4$, the empirical sizes of LM2 and h2 are particularly large. LM1 and FQD are best, with sizes fairly close to the nominal size $\alpha = 0.05$. The considerable increase in size in comparing LM2 with LM1 suggests that over-parameterization of the alternative hypothesis assumptions may be quite serious when a lagged dependent variable is present.

The power comparisons (see Table 6) show h1 to be best, followed by LM1. This is to be expected as these two tests incorporate the most (correct) a priori information. The other three tests have similar power.

In view of the superior size characteristics, this experiment suggests that the FQD test may be preferable to h1, h2 and LM2.

As discussed earlier, the assumptions underlying the theory of FQD have been violated with this choice of model. Despite this violation, the size of the FQD test was still reasonably close to the nominal size. Again, this indicates considerable robustness to departures from the assumptions.
Table 5

Empirical sizes when $\alpha = 0.05$

Lagged dependent variable model, $T = 40$, 1000 replications$^a$

<table>
<thead>
<tr>
<th>Test</th>
<th>Empirical size</th>
<th>$\gamma = 0.4$</th>
<th>$\gamma = 0.8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FQD</td>
<td></td>
<td>0.078</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.008)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>LM1</td>
<td></td>
<td>0.070</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>LM2</td>
<td></td>
<td>0.113</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.010)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>h1</td>
<td></td>
<td>0.039</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.006)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>h2</td>
<td></td>
<td>0.102</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
</tr>
</tbody>
</table>

$^a$ Standard errors are given in parentheses.
Table 6
Power comparisons, empirical size 0.05
Lagged dependent variable model, T = 40, 1000 replications

(i) \( \gamma = 0.4 \)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>FQD</th>
<th>LM1</th>
<th>LM2</th>
<th>h1</th>
<th>h2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.2</td>
<td>0.087</td>
<td>0.082</td>
<td>0.078</td>
<td>0.194</td>
<td>0.043</td>
</tr>
<tr>
<td>0.4</td>
<td>0.207</td>
<td>0.298</td>
<td>0.241</td>
<td>0.511</td>
<td>0.172</td>
</tr>
<tr>
<td>0.6</td>
<td>0.349</td>
<td>0.530</td>
<td>0.448</td>
<td>0.729</td>
<td>0.354</td>
</tr>
<tr>
<td>0.8</td>
<td>0.505</td>
<td>0.648</td>
<td>0.592</td>
<td>0.811</td>
<td>0.475</td>
</tr>
<tr>
<td>0.99</td>
<td>0.505</td>
<td>0.557</td>
<td>0.502</td>
<td>0.712</td>
<td>0.390</td>
</tr>
</tbody>
</table>

(ii) \( \gamma = 0.8 \)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>FQD</th>
<th>LM1</th>
<th>LM2</th>
<th>h1</th>
<th>h2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.2</td>
<td>0.169</td>
<td>0.085</td>
<td>0.064</td>
<td>0.304</td>
<td>0.074</td>
</tr>
<tr>
<td>0.4</td>
<td>0.385</td>
<td>0.355</td>
<td>0.250</td>
<td>0.672</td>
<td>0.327</td>
</tr>
<tr>
<td>0.6</td>
<td>0.608</td>
<td>0.723</td>
<td>0.612</td>
<td>0.900</td>
<td>0.684</td>
</tr>
<tr>
<td>0.8</td>
<td>0.788</td>
<td>0.899</td>
<td>0.833</td>
<td>0.972</td>
<td>0.863</td>
</tr>
<tr>
<td>0.99</td>
<td>0.830</td>
<td>0.896</td>
<td>0.837</td>
<td>0.962</td>
<td>0.863</td>
</tr>
</tbody>
</table>
6. Summary and Conclusions

Using the device of transforming a time series regression model into the frequency domain, the results of White's (1980) analysis of heteroskedastic error models can be utilized for the case of stationary autocorrelated errors. In particular, a consistent estimator of the OLS covariance matrix is derived and a misspecification test which does not require knowledge of the error structure (other than stationarity) is provided. Because of constraints that stationarity implies on the nature of frequency domain heteroskedasticity and the typical nature of regressor leverage, the suggestion is made that White's methods may in fact perform better for the autocorrelated error model than for the original heteroskedastic case.

Monte Carlo experiments for both the covariance estimator and the specification test have provided encouraging results and shown that these procedures may be quite robust to departures from the underlying assumptions on the nature of the regressors.
References


Appendix

This discussion draws heavily on that of Fishman (1969).

**The Finite Fourier Transform**

Consider a $T \times 1$ vector $x = (x_1, x_2, \ldots, x_T)'$, which is a realization generated by a covariance stationary, ergodic time-series process. For ease of exposition we will assume $T$ is even, given by $T = 2n$. We define the Finite Fourier Transform (FFT) of $x$ to be another $T \times 1$ vector $w = (w_1, w_2, \ldots, w_T)'$, where

$$w_j = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_t e^{i(t-n)\lambda_j}, \quad (j = 1, 2, \ldots, T) \quad (A1)$$

and $\lambda_j$ is a 'frequency' in $(-\pi, \pi)$, given by

$$\lambda_j = -\pi + \frac{\pi j}{n}. \quad (A2)$$

The vector $w$ is sometimes called the 'frequency domain representation' of $x$.

**The Periodogram**

The quantity $|w_j|^2$ is called the periodogram ordinate of $x$ at frequency $\lambda_j$. The set $|w_j|^2, (j = 1, 2, \ldots, T)$, is the periodogram of $x$. From the definition (A2) of $\lambda_j$, it is clear that

$$\lambda_{2n-j} = -\lambda_j.$$ 

Hence $w_{2n-j}$ and $w_j$ are complex conjugates, implying that

$$|w_{2n-j}|^2 = |w_j|^2. \quad (A3)$$

If we set $s = j-n$, it follows that

$$|w_{n-s}|^2 = |w_{n+s}|^2$$

and the periodogram ordinates are seen to be symmetric about $\lambda_n$, that is,
zero frequency. Finally, we note that

\[ w_n = w(\lambda=0) = \frac{1}{T} \sum_{t=1}^{T} x_t^2 = TX^2, \]  

(A4)

showing that the periodogram ordinate at zero frequency is completely determined by the sample mean of \( x \).

**Matrix Notation**

Expressing (A1) in matrix notation,

\[ w = Ux, \]  

(A5)

where \( U \) is a \( T \times T \) symmetric complex-valued matrix, having \((k,\ell)\) element given by

\[ U_{k\ell} = \frac{1}{\sqrt{T}} e^{i\lambda_k(\ell-n)}. \]  

(A6)

The frequency domain representation is obtained by pre-multiplying the time domain vector by \( U \). Denoting the operation of transpose-conjugation, by \( ' * \), it is easily seen that

\[ U^*U = UU^* = I_T. \]  

(A7)

Thus, \( U \) is the complex analogue of an orthogonal matrix and

\[ U^{-1} = U^*. \]  

(A8)

For the purposes of this paper, the most important property of \( U \) is that it approximately diagonalizes all 'Toeplitz form' matrices. More specifically, if a \( T \times T \) matrix \( A \) has the property that \( A_{k\ell} \) depends only on \(|k-\ell|\), it is called a Toeplitz form matrix. It can be shown that \( UAU^* \) approaches a diagonal matrix as \( T \to \infty \). In particular, if \( A \) is the covariance matrix of a stationary process, then

\[ UAU^* = 2\pi \text{diag}(f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_T)). \]  

(A9)
as $T \to \infty$, where $f(\cdot)$ is the spectrum of the process and $\lambda_j$ is given in (A2).

**Regression in the Frequency Domain**

Suppose we have, in the usual notation, a time series linear model

$$y = X\beta + u.$$ 

Pre-multiplying by $U$, we obtain the frequency domain representation

$$z = Z\beta + v,$$

where $z = Uy$, $Z = UX$ and $v = Uu$. The ordinary least squares (ols) estimator of $\beta$, using frequency domain observations, is given by

$$\hat{\beta}_F = (Z^*Z)^{-1}Z^*z$$

$$= (X'U*UX)^{-1}X'U*Uy$$

$$= (X'X)^{-1}X'y = \hat{\beta},$$

where $\hat{\beta}$ is the usual (time domain) ols estimate. The frequency domain residual vector $\hat{v}$ is given by

$$\hat{v} = z - \hat{\beta}_F Z = Uy - \hat{\beta}UX$$

$$= U(y - \hat{\beta}X).$$

Thus

$$\hat{v} = U\hat{u}.$$  \hspace{1cm} (A10)

This implies that in order to obtain the frequency domain ols residual vector, it is only necessary to find the FFT of the time domain residual vector $\hat{u}$. For this purpose, regression in the frequency domain is unnecessary.


A Note on A Bayesian Estimator in an Autocorrelated Error Model. William Griffiths and Dan Dao, No. 3 - April 1979.


Bayesian Econometrics and How to Get Rid of Those Wrong Signs. William E. Griffiths, No. 31 - November 1987.


Using the Kalman Filter to Estimate Sub-Populations. Howard E. Doran, No. 44 - March 1990.


