No. 9154
INFORMATION MATRIX TEST, PARAMETER HETEROGENEITY AND ARCH: A SYNTHESIS

R 46<br>by Anil K. Bera<br>and Sangkyu Lee<br>518.92

October 1991

Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis*<br>ANIL K. BERA<br>University of Illinois at Urbana-Champaign and CentER, Tilburg University<br>SANGKYU LEE<br>CNB Economic Research Institute at Seoul


#### Abstract

We apply the White information matrix (IM) test to the linear regression model with autocorrelated errors. A special case of one component of the test is found to be identical to the Engle Lagrange multiplier (LM) test for autoregressive conditional heteroskedasticity (ARCH). Given Chesher's interpretation of the IM test as a test for parameter heterogeneity, this establishes a connection among the IM test, ARCH and parameter variation. This also enables us to specify conditional heteroskedasticity in a more general and convenient way. Other interesting byproducts of our analysis are tests for the variation in conditional and static skewness which we call tests for "heterocliticity."


JEL Classification No. 210/Key Words: Autoregressive Conditional Heteroskedasticity; Information Matrix Test; Lagrange Multiplier Test; Parameter Heterogeneity.
*Correspondence should be addressed to Anil K. Bera, Department of Economics, University of Illinois at Urbana-Champaign, 1206 South Sixth Street, Champaign, IL 61820, U.S.A.

In a pioneering article, White (1982) suggested the information matrix (IM) test as a general test for model specification. In recent years, this test has received a lot of attention. In particular, Chesher (1984) demonstrated that this test can be viewed as a Lagrange multiplier (LM) test for specification error against the alternative of parameter heterogeneity. As a byproduct of this analysis, Chesher (1983) and Lancaster (1984) provided a " $n R^{2}$ ". version of the IM test. An application of the IM test to the linear regression model by Hall (1987) led to a very interesting result that the test decomposed asymptotically into three components, one testing heteroskedasticity and the other two testing some forms of normality. Engle (1982), in an apparently unrelated influential paper, introduced the autoregressive conditional heteroskedasticity (ARCH) model which characterizes explicitly the conditional variance of the regression disturbances. He also suggested an LM test for ARCH. The purpose of this paper is to establish a connection among the IM test, parameter heterogeneity and ARCH. And as far as the IM test is concerned, we examine only the algebraic structure of the test.

An important finding by Hall (1987) was that the components of the IM test are insensitive to serial correlation. Hall also commented "had our original specification included first order autoregressive errors, then the $I M$ test does not decompose asymptotically into the sum of our original three component test . . . 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." (p. 262). In the next section, we start with a linear regression model with autoregressive (AR) errors and apply the IM test to it. The indicator vector is found to have a block diagonal covariance matrix. And as the null model now has more parameters, naturally we get a few extra components in the IM test. From the additional components of the statistic, we can also obtain the Engle's LM test for $A R C H$ as a special case. The implication of this result is discussed in detail in section 3 . Given Chesher's interpretation of the IM test as a test for parameter heterogeneity or random coefficient, it is now easy to give a random coefficient interpretation to ARCH. This fact has been noted recently by several authors [see, e.g., Tsay (1987)]. This provides us with a convenient framework to extend $A R C H$ so that interaction factor between past residuals could also be considered and as a consequence we suggest an augmented ARCH (AARCH) model. The last section of the paper contains some concluding remarks.

## 2. THE IM TEST FOR THE LINEAR REGRESSION MODEL WITH AR ERRORS

We consider the 11 near regression model

$$
\begin{equation*}
y_{t}=x_{t}^{\prime} \beta+\varepsilon_{t} \tag{1}
\end{equation*}
$$

where $y_{t}$ is the $t-t h$ observation on the dependent variable, $x_{t}$ is a $k \times 1$ vector of fixed regressors and the $\varepsilon_{t}$ are assumed to follow a stationary $A R(p)$ process

$$
\begin{equation*}
\varepsilon_{t}=\sum_{j=1}^{p} \varphi_{j} \varepsilon_{t-j}+u_{t} \tag{2}
\end{equation*}
$$

with $u_{t} \sim \operatorname{NIID}\left(0, \sigma_{u}^{2}\right)$. We will write this $\operatorname{AR}(p)$ process as $\varepsilon_{t}=\underline{\varepsilon}_{t}^{\prime} \phi+u_{t}$ where $\varepsilon_{-}=\left(\varepsilon_{t-1}, \ldots, \varepsilon_{t-p}\right)^{\prime}$ and $\phi=\left(\varphi_{1}, \ldots, \phi_{p}\right)^{\prime}$. Assuming that $\varepsilon_{1}$ is given, the log-1ikelihood function for this model can be written as

$$
L(\theta)=\sum_{t=1}^{n} \ell_{t}(\theta)=-\frac{n}{2} \log 2 \pi-\frac{n}{2} \log \sigma_{u}^{2}-\frac{1}{2 \sigma_{u}^{2}} \sum_{t=1}^{n}\left(\varepsilon_{t}-\varepsilon_{t}^{\prime} \phi\right)^{2}
$$

where $\theta=\left(\beta^{\prime}, \phi^{\prime}, \sigma_{u}^{2}\right)^{\prime}$ is a $q^{\times}$vector of parameters with $q=k+p+1$. Note that $\left(\varepsilon_{t}-\underline{\varepsilon}_{t}^{\prime} \phi\right)$ involves $\beta$ since $\varepsilon_{t}-\underline{\varepsilon}_{t}^{\prime} \phi=\left(y_{t}-\underline{y}_{t}^{\prime} \phi\right)-\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right)^{\prime} \beta$, where $\underline{y}_{t}=\left(y_{t-1}, \ldots, y_{t-p}\right)^{\prime}$ and $\underline{x}_{t}=\left(x_{t-1}, \ldots, x_{t-p}\right)^{\prime}$.

Let $\hat{\theta}$ denote the maximum likelihood estimate (MLE) of $\theta$. Then White's IM test is constructed based on

$$
d(\hat{\theta})=\operatorname{vech} C(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n} d_{t}(\hat{\theta}) \quad \text { (say) }
$$

where

$$
C(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n}\left[\frac{\partial^{2} \ell_{t}(\hat{\theta})}{\partial \theta \partial \theta^{\prime}}+\left(\frac{\partial \ell_{t}(\hat{\theta})}{\partial \theta}\right)\left(\frac{\partial \ell_{t}(\hat{\theta})}{\partial \theta}\right)^{\prime}\right]=A(\hat{\theta})+B(\hat{\theta}) \quad \text { (say) }
$$

Note that $-A(\hat{\theta})^{-1}$ and $B(\hat{\theta})^{-1}$ are the two different estimators for the asymptotic variance of $\sqrt{n} \hat{\theta}$ using the Hessian matrix and the outer product form, respectively. Therefore, the $I M$ test principles can also be viewed as a test based on the difference of two estimators.

A consistent estimator of the variance matrix of $\sqrt{n} d(\hat{\theta})$ is [see White (1982, p. 11)]

$$
\begin{equation*}
\hat{v}(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n} a_{t}(\hat{\theta}) a_{t}^{\prime}(\hat{\theta}) \tag{3}
\end{equation*}
$$

where $a_{t}(\hat{\theta})=d_{t}(\hat{\theta})-\nabla d(\hat{\theta})_{A}(\hat{\theta})^{-1} \nabla \ell_{t}(\hat{\theta})$ with $\nabla d(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n} \frac{\partial d_{t}(\hat{\theta})}{\partial \theta}$ and $\nabla \ell_{t}(\hat{\theta})=\frac{\partial \ell_{t}(\hat{\theta})}{\partial \theta}$. Then the White IM test takes the form of

$$
\begin{equation*}
T_{W}=n d^{\prime}(\hat{\theta}) \hat{v}(\hat{\theta})^{-1} d(\hat{\theta}) \tag{4}
\end{equation*}
$$

When the model (1) is correct, $\mathrm{T}_{\mathrm{W}}$ follows an asymptotic $\mathrm{X}^{2}$ distribution with $\frac{q(q+1)}{2}$ degrees of freedom. If there is an intercept term in the regression model (1), the $x^{2}$ degrees of freedom should be reduced by one. Similar adjustments are necessary if the regressors contain some polynomial terms and a constant, or if some of the exogenous variables are binary [see White (1980, p. 825)]. It should also be noted that White (1982) derived the IM test for IID observations. However, as shown in White (1987), the IM equality holds under fairly general conditions. For our autoregressive case, mixing conditions stated in White (1987) are satisfied, and therefore the IM test remains valid.

After some algebra and rearranging the terms in $d(\hat{\theta})$, we can write (for algebraic derivations, see Appendices A and B), suppressing $\theta$ such that writing $\hat{d}$ for $d(\hat{\theta})$,

$$
\begin{equation*}
\hat{d}=\left(\hat{d}_{1}^{\prime}, \hat{d}_{2}^{\prime}, \hat{d}_{3}, \hat{d}_{4}^{\prime}, \hat{a}_{5}^{\prime}, \hat{d}_{6}^{\prime}\right) \cdot \tag{5}
\end{equation*}
$$

where

$$
\begin{aligned}
& \hat{d}_{1}:\left[\frac{1}{n \theta_{u}^{4}} \sum_{t=1}^{n}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t i}-\underline{x}_{t i}^{\prime} \hat{\phi}^{n}\right)\left(x_{t j}-\underline{x}_{t j}^{\prime} \hat{\phi}\right)\right] \quad, j=1,2, \ldots, k ; \quad 1 \leq j \\
& \hat{d}_{2}:\left[\frac{1}{n \hat{\sigma}_{u}^{4}} \sum_{t=1}^{n}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right) \hat{\varepsilon}_{t-1} \hat{\varepsilon}_{t-j}\right]_{1, j=1,2, \ldots, p ; \quad 1 \leq j} \\
& \hat{d}_{3}:\left[\frac{1}{4 n \theta_{u}^{8}} \sum_{t=1}^{n}\left(\hat{u}_{t}^{4}-3 \hat{\sigma}_{u}^{4}\right)\right] \\
& \hat{d}_{4}:\left[\frac{1}{n \hat{\theta}_{u}^{4}} \sum_{t=1}^{n}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t 1}-x_{t i}^{\prime} \hat{\phi}^{n}\right) \hat{\varepsilon}_{t-j}-\frac{1}{n \theta_{u}^{2}} \sum_{t=1}^{n} \hat{u}_{t} x_{t-j 1}\right] \\
& i=1,2, \ldots, k ; \quad j=1,2, \ldots, p \\
& \hat{\mathrm{~d}}_{5}:\left[\frac{1}{2 n \hat{\sigma}_{u}^{6}} \sum_{t=1}^{n} \hat{u}_{t}^{3}\left(x_{t i}-\underline{x}_{t i}^{\prime} \hat{\phi}\right)\right]_{1=1,2, \ldots, k} \\
& \hat{d}_{6}:\left[\frac{1}{2 n \hat{\theta}_{u}^{6}} \sum_{t=1}^{n} \hat{u}_{t}^{3 \hat{\varepsilon}}{ }_{t-1}\right]_{i=1,2, \ldots, p}
\end{aligned}
$$

Our expressions for $\hat{\mathrm{d}}_{1}, \hat{\mathrm{~d}}_{3}$ and $\hat{\mathrm{d}}_{5}$ are identical to those of $\Delta_{1}, \Delta_{3}$ and $\Delta_{2}$ of Hall (1987, pp. $259-260$ ) if we put $\hat{\phi}=0$. If it is desirable to test only in a certain direction, we can premultiply $\hat{d}$ by a selection matrix whose elements are either zero or unity [see White (1982, pp. 9-10) and Hall (1987, p. 258)].

Now to obtain the IM test statistic, all we need is to derive the variance matrix of $\hat{d}$. We find that the variance matrix is block diagonal (for detailed derivation, see Appendix B). Denote the estimator of the variance of $\sqrt{n} \hat{d}_{1}$ as $\hat{V}\left(\hat{d}_{1}\right) \equiv \hat{V}_{1}, 1=1,2, \ldots, 6$. To express $\hat{V}_{1}$ 's succinctly, we define the vectors whose typical elements are described as

$$
\begin{aligned}
& \underline{x}_{t}: \quad\left[\left(x_{t i}-\underline{x}_{t i}^{\prime} \hat{\phi}\right)\left(x_{t j}-\underline{x}_{t j}^{\prime} \hat{\phi}\right)\right. \\
&\left.-\frac{1}{n} \sum_{t=1}^{n}\left(x_{t i}-\underline{x}_{t i} \hat{\phi}\right)\left(x_{t j}-\underline{x}_{t j} \hat{\phi}\right)\right]_{i, j=1,2, \ldots, k ; \quad 1 \leq j} \\
& \underline{\xi}_{t}: \quad\left[\hat{\varepsilon}_{t-1} \hat{\varepsilon}_{t-j}-\frac{1}{n} \sum_{t=1}^{n} \hat{\varepsilon}_{t-1} \hat{\varepsilon}_{t-j}\right]_{i, j=1,2, \ldots, p ; \quad 1 \leq j} \\
& \underline{s}_{t}: \quad\left[\left(x_{t i}-\underline{x}_{t i}^{\prime} \hat{\phi}\right) \hat{\varepsilon}_{t-j}\right]_{i=1,2, \ldots, k ;} \quad j=1,2, \ldots, p \\
& \underline{z}_{t}: \quad\left[x_{t-j i}\right]_{i=1,2, \ldots, k ; \quad j=1,2, \ldots, p} \\
& \underline{r}_{t}: \quad\left[x_{t i}-\underline{x}_{t i}^{\prime} \hat{\phi}_{i=1,2}\right]_{i=1, \ldots, k}
\end{aligned}
$$

We also denote

$$
\hat{\mathrm{w}}=\nabla \hat{d}_{41} \hat{A}_{11}^{-1} \nabla \cdot \hat{d}_{41}+\frac{1}{n \theta_{u}^{2}} \sum_{t=1}^{n} \underline{z}_{t} \underline{z}_{t}^{\prime}
$$

where $\nabla d_{41}$ is the $(4,1)$ block of $\nabla d(\theta)$ and $A_{11}$ is the upper left-hand corner block of $A(\theta)$ [see Appendix B]. Then we have very concise forms of $\hat{\mathrm{V}}_{\mathrm{i}}$ 's as follows:

$$
\begin{aligned}
& \hat{v}_{1}=\frac{2}{n \theta_{u}^{4}} \sum_{t=1}^{n} \underline{x}_{t} \underline{x}_{t}^{\prime}, \quad \hat{v}_{2}=\frac{2}{n \theta_{u}^{4}} \sum_{t=1}^{n} \underline{\xi}_{t} \underline{\xi}_{t}^{\prime}, \quad \hat{v}_{3}=\frac{3}{2 \theta^{8}} \\
& \hat{v}_{4}=\frac{2}{n \theta_{u}^{4}} \sum_{t=1}^{n} \underline{s}_{t} \underline{s}_{t}^{\prime}+\hat{W}, \quad \hat{v}_{5}=\frac{3}{2 n \theta_{u}^{6}} \sum_{t=1}^{n} \underline{r}_{t} \underline{r}_{t}^{\prime}, \quad \hat{v}_{6}=\frac{3}{2 n \theta_{u}^{6}} \sum_{t=1}^{n} \hat{\varepsilon}_{t} \hat{\varepsilon}_{t}^{\prime}
\end{aligned}
$$

Given the block diagonality of the variance matrix of $\hat{d}$, we can write the $I M$ test as

$$
\begin{equation*}
T_{W}=\sum_{i=1}^{6} T_{i}=n \sum_{i=1}^{6} \hat{d}_{i}^{\prime} \hat{v}_{i}^{-1} \hat{d}_{i} \tag{6}
\end{equation*}
$$

that is, the derived IM test statistic is found to be decomposed as the sum of six quadratic forms. In the next section, we analyze these components of $T_{W}$ in detail.

## 3. INTERPRETATION OF THE COMPONENTS OF THE IM TEST

Using Chesher's analysis, we can say the statistic $T_{1}$ is a test for randomness of the regression parameters in the presence of autocorrelation. If we put $\hat{\phi}=0$, then this reduces to the White (1980) test for heteroskedasticity [and $\mathrm{T}_{1 \mathrm{n}}$ in Hall (1987, p. 261)]. Recently, there have been some robustness studies of various tests for heteroskedasticity in the presence of autocorrelation [see, e.g., Epps and Epps (1977), Bera and Jarque (1982), Godfrey and Wickens (1982), Bumb and Kelejian (1983), Bera and McKenzie (1986)] and their general conclusion is that various tests for heteroskedasticity are sensitive to the presence of autocorrelation. A byproduct of our analysis is that we have a simple test for heteroskedasticity in the presence of autocorrelation. All we need to do is to modify the White test slightly. Instead of regressing the squares of the least squares residuals on the squares and cross products of $x_{t}$ 's, we should regress $\hat{u}_{t}^{2}$ on the squares and cross products of $\left(x_{t}-\hat{x}_{t}^{\prime} \hat{\phi}\right)$ after estimating the model with an appropriate AR process. For example, if there is AR(1) error, then the regressors should be the squares and cross products of $\left(x_{t}-\hat{\phi}_{1} x_{t-1}\right)$. Similarly, the modification of $T_{2 n}$ in Hall (1987), which is our $T_{5}$, requires that we should replace $x_{t}$ by ( $x_{t}-\underline{x}_{t}^{\prime} \hat{\phi}$ ). Our $T_{3}$ is (kurtosis) test for normality, and it utilizes the conditional mean corrected residuals rather than the OLS residuals.
let us now concentrate on the new test statistics we obtain by including $\phi$ in our model. The statistic $T_{2}$ tests the randomness of $\phi=\left(\phi_{1}, \phi_{2}, \ldots, \phi_{\mathrm{p}}\right)^{\prime}$. Suppose that the parameters of autoregressive errors are varying around a mean value with finite variances. This
can be formulated as $\phi_{t} \sim(\phi, \Omega)$, where $\phi_{t}=\left(\phi_{1 t}, \phi_{2 t}, \ldots, \phi_{p t}\right)^{\prime}$. Then $T_{2}$ is the LM statistic for testing $H_{0}: S l=0$. Let us first consider a very special case in which $\phi=0$ and $\Omega$ is diagonal. Therefore, we have $\hat{\phi}_{1}=\hat{\phi}_{2}=\ldots=\hat{\phi}_{p}=0$, and $\hat{u}_{t-1}=\hat{\varepsilon}_{t-1}(1=1,2, \ldots, p)$, where the $\hat{\varepsilon}_{t}$ are the OLS residuals. Consequently, $T_{2}$ reduces to

$$
\begin{equation*}
T_{2}=\frac{1}{2}\left[\sum_{t=1}^{n} \hat{u}_{t}^{2}\left(\frac{\hat{u}_{t}^{2}}{\theta^{2}}-1\right)\right] \cdot\left[\sum_{t=1}^{n} \underline{\xi}_{t} \underline{\xi}_{t}^{\prime}\right]^{-1}\left[\sum_{t=1}^{n} \hat{u}_{t}^{2}\left(\frac{\hat{u}_{t}^{2}}{\hat{\theta}_{u}^{2}}-1\right)\right] \tag{7}
\end{equation*}
$$

where $\hat{u}_{t}^{2}=\left(\hat{u}_{t-1}^{2}, \hat{u}_{t-2}^{2}, \ldots, \hat{u}_{t-p}^{2}\right)$ ' and a typical elements of $\underline{\xi}_{t}$ is now $\left(\hat{u}_{t-1}^{2}-\frac{1}{n} \sum_{t=1}^{n} \hat{u}_{t-1}^{2}\right)$, for $i=1,2, \ldots, p$. This is identical to the Engle (1982) LM statistic for testing the pth-order linear ARCH disturbances, i.e., testing $H_{0}: \alpha_{1}=\alpha_{2}=\ldots=\alpha_{p}=0$ in the ARCH process specified as

$$
\operatorname{Var}\left(u_{t} \mid \underline{u}_{t}\right)=\sigma_{u}^{2}+\alpha_{1} u_{t-1}^{2}+\ldots+\alpha_{p} u_{t-p}^{2}
$$

where $\underline{u}_{t}=\left(u_{t-1}, u_{t-2}, \ldots, u_{t-p}\right)^{\prime}$. An asymptotically equivalent form of this statistic is $n R^{2}$ where $R^{2}$ is the coefficient of multiple determination from the regression of $\hat{u}_{t}^{2}$ on a unit term and $\left(\hat{u}_{t-1}^{2}, \hat{u}_{t-2}^{2}, \ldots, \hat{u}_{t-p}^{2}\right)$.

From our representation of test for $A R C H$ as a test for randomness of $\phi$ parameters and its equivalence to one component of the IM test, the consequence of the presence of ARCH is that the "usual" estimators for variance of $\hat{\phi}$ will be inconsistent if ARCH is ignored. This is similar to the case that the standard variance estimator for $\hat{\beta}$ is inconsistent in the presence of static heteroskedasticity. Therefore,
the standard tests for autocorrelation are not valld in the presence of ARCH [see, e.g., Diebold (1986) and Bera et al. (1990)]. This result is not entirely obvious since under $A R C H$, the disturbances are still unconditionally homoskedastic. Although the above point could be made without an IM test interpretation, IM test framework provides an easy guidance for checking whether the standard inference procedures fail.

We now relax the assumption of the diagnolaity of $\Omega$. The structure of the test statistic will remain the same except $R^{2}$ will be obtained by regressing $\hat{u}_{t}^{2}$ on a constant and the squares and cross products of the lagged residuals. $T_{2}$ will then be a LM statistic for testing $H_{0}: \quad \alpha_{i j}=0(i \geq j=1,2, \ldots, p)$ in

$$
\begin{equation*}
\operatorname{Var}\left(u_{t} \mid \underline{u}_{t}\right)=\sigma_{u}^{2}+\sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_{i j} u_{t-i} u_{t-j} \quad 1 \geq j \tag{8}
\end{equation*}
$$

The above specification of conditional variance generalizes the Engle ARCH model. This will be called the augmented ARCH (AARCH) process. Properties and testing of this model are discussed in Bera et al. (1990). Lastly, if we additionally relax the assumption of $\phi=0$, $\hat{u}_{t}$ will no longer be equal to $\hat{\varepsilon}_{t}$ and $T_{2}$ will have to be calculated from the regression of $\hat{u}_{t}^{2}$ on a constant and the squares and cross products of $\hat{\varepsilon}_{t-i}(i=1,2, \ldots, p)$. This will give us the LM statistics for testing $A R C H$ or $A A R C H$ in the presence of autocorrelation.

From the above discussion, it is clear that the Engle ARCH model can be viewed as a special case of random coefficient autoregressive (RCAR) model. To see this more clearly, let us write equation (2) as

$$
\varepsilon_{t}=\sum_{j=1}^{p} \Phi_{j t^{E} t-j}+u_{t} .
$$

If it is assumed that $\phi_{j t} \sim\left(0, \alpha_{j}\right)$ and $\operatorname{cov}\left(\phi_{j t}, \phi_{j} t^{\prime}\right)=0$, for $j \neq j^{\prime}$, then the conditional variance is given by

$$
\operatorname{Var}\left(\varepsilon_{t} \mid \varepsilon_{t}\right)=\sigma_{u}^{2}+\sum_{j=1}^{p} \alpha_{j} \varepsilon_{t-j}^{2}
$$

Here we observe that $A R C H$ and the above RCAR models have the same first two conditional moments as mentioned in Tsay (1987) where it is called as second-order equivalence. If we further assume that the $\Phi_{\mathrm{jt}}$ are normally distributed, then all the moments of ARCH and RCAR processes will be the same, e.g., for $p=1$, the first four moments are

$$
\mu_{1}=0, \mu_{2}=\frac{\sigma_{u}^{2}}{1-\alpha_{1}}, \mu_{3}=0 \text { and } \mu_{4}=\frac{3 \sigma_{u}^{4}\left(1+\alpha_{1}\right)}{\left(1-\alpha_{1}\right)\left(1-3 \alpha_{1}^{2}\right)}
$$

[see Engle (1982, p. 992)]. Here we should note that calculation of moments are much easier under the RCAR scheme.

By comparing $T_{1}$ and $T_{2}$, we note that they test for static and conditional heteroskedasticity, respectively. Given the block diagonality of covariance matrix of the IM test in our case, we can test for static and conditional heteroskedasticity simultaneously simply by adding up these two statistics. The statistic $T_{4}$ is also related to $T_{1}$ and $T_{2}$. From the expression of $\hat{d}_{4}$, we note that $\mathrm{T}_{4}$ has two components. The second component is based on $\left(n \theta_{u}^{2}\right)^{-1} \sum_{t} \hat{\mathrm{u}}_{\mathrm{t}} \underline{x}_{\mathrm{t}}^{\prime}$ and this can be viewed as some form of a test for exogeneity. In certain practical applications when the $x_{t}$ are known to be exogenous, this part can be ignored. Then $T_{4}$ will be based only on the first component of $\hat{\mathrm{d}}_{4}$ and the resulting test will be a test for conditional heteroskedasticity caused by the interaction between the
disturbance term and the regressors. We should, however, note that by excluding the second component of $\hat{d}_{4}$ we do not get exactly an IM test but something that is very close to an IM test. For the special case,

$$
\hat{v}_{4}=\frac{2}{n \theta_{u}^{2}} \sum_{t=1}^{n} \underline{s}_{t} \underline{s}_{t}^{\prime}
$$

and therefore, for obtaining $T_{4}$ we run the regression of $\hat{u}_{t}^{2}$ on a constant and cross products of lagged innovations $\hat{\varepsilon}_{t}$ and transformed , ^ exogenous variables $x_{t}-\underline{x}_{t} \phi$. As a natural consequence, a general test statistic for heteroskedasticity would be $T_{1}+T_{2}+T_{4}$ which under the null hypothesis will have an asymptotic $x^{2}$ distribution with $(k+p)(k+p+1) / 2$ degrees of freedom. To get reasonable power, we will have to make a judicious selection of the regressors from the set of squares and cross products of $x_{t}-\underline{x}_{t}^{\prime} \hat{\phi}$ and $\hat{\varepsilon}_{t}$, or make some adjustment to the test statistic [see Bera (1986)].

The last two statistics $\mathrm{T}_{5}$ and $\mathrm{T}_{6}$ can be viewed as the statistics for testing variation in the third moment of $u_{t}$. In $T_{5}$, the variation is assumed to depend on the exogenous variables $x_{t}$ and in $T_{6}$, on the lagged innovat ton process. In some sense, we could say that $\mathrm{T}_{5}$ and $T_{6}$ test for static and conditional heterocliticity, respectively. The term heterocliticity is used since when the skewness coefficient is plotted against $x_{t}$ or $\varepsilon_{t-i}$, we obtain the clitic curve [see Kendall and Stuart (1973, p. 362)]. As noted in Hall (1987), the test for normality (skewness part) proposed by Bowman and Shenton (1975) and Jarque and Bera (1987) is a special case of $T_{5}$ while $T_{3}$ which tests for the variation of $\sigma_{u}^{2}$ is a pure test for kurtosis. In this con-
nection, let us mention that if the IM test is applied to an ARCH model, that leads to a test for heterokurticity [for details see Bera and Zuo (1991)]. This provides a specification test for an estimated ARCH model.

## 4. CONCLUSION

Our application of the White IM test to the linear regression model with autoregressive errors provides many interesting results. The most important result is that a special case of one components of this test is identical to the Engle LM test for ARCH. Chesher's interpretation of the IM test as the test for parameter heterogenefty leads us naturally to specify the $A R C H$ processes as a random coefficient autoregressive (RCAR) model. From both theoretical and practical points of view, this representation of $A R C H$ is convenient and useful. As discussed in Bera et al. (1990), we can now easily verify the stationarity condition for $A R C H$ as a special case of RCAR model, study the robustness of test for AR process in the presence of ARCH and vice versa, and generalize the ARCH process to take account of interaction between the disturbance terms.

The difference between the static and conditional heteroskedasticity is now clear. The former could be related to the variation of the regression coefficients while the latter to the variation of the autoregressive parameters. A mixture of them is possible when the heteroskedasticity is caused by the interaction between exogenous variables and disturbances. We also discuss the possibilities of static and conditional variations in skewness, what we call heterocliticity.

## APPENDIX A

The Derivatives of the Log-1ikelihood Function

For our model, the vector of parameters is $\theta=\left(B^{\prime}, \phi^{\prime}, \sigma_{u}^{2}\right)^{\prime}$ and the log-likelihood function for the $t$-th observation conditional on the information set $\Psi_{t-1}$, in which $\varepsilon_{t}=\left(\varepsilon_{t-1}, \ldots, \varepsilon_{t-p}\right)$ ' is contained, is given by

$$
\ell_{t}(\theta)=-\frac{1}{2} \log 2 \pi-\frac{1}{2} \log \sigma_{u}^{2}-\frac{1}{2 \sigma_{u}^{2}}\left(\varepsilon_{t}-\varepsilon_{t}^{\prime} \phi\right)^{2} .
$$

Note that ${ }^{11} t=\varepsilon_{t}-\underline{\varepsilon}_{t}^{\prime} \phi=\left(y_{t}-\underline{y}_{t}^{\prime} \phi\right)-\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right)^{\prime} \beta$ where $\underline{y}_{t}=\left(y_{t-1}, \ldots, y_{t-p}\right)^{\prime}$ and $x_{t}=\left(x_{t-1}, \ldots, x_{t-p}\right)^{\prime}$. Then the first and second partial derivatives of $\ell_{t}(\theta)$ with respect to $\theta$ are easily obtained. The first derivatives are

$$
\begin{aligned}
& \frac{\partial \ell_{t}(\theta)}{\partial \beta}=\frac{1}{\sigma_{u}^{2}} u_{t}\left(x_{t}-\underline{x}_{t}^{\prime} \varphi\right), \frac{\partial \ell_{t}(\theta)}{\partial \phi}=\frac{1}{\sigma_{u}^{2}} u_{t} \underline{\varepsilon}_{t} \text { and } \\
& \frac{\partial \ell_{t}(\theta)}{\partial \sigma_{u}^{2}}=-\frac{1}{2 \sigma_{u}^{2}}+\frac{1}{2 \sigma_{u}^{4}} u_{t}^{2} .
\end{aligned}
$$

And the second derivatives are

$$
\begin{aligned}
& \frac{\partial^{2} \ell_{t}(\theta)}{\partial \beta \partial \beta^{\prime}}=-\frac{1}{\sigma_{u}^{2}}\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right)\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right)^{\prime}, \frac{\partial^{2} \ell_{t}(\theta)}{\partial \phi \partial \phi^{\prime}}=\frac{1}{\sigma_{u}^{2}} \underline{\varepsilon}_{t} \varepsilon_{t}^{\prime}, \\
& \frac{\partial^{2} \ell_{t}(\theta)}{\partial\left(\sigma_{u}^{2}\right)^{2}}=\frac{1}{2 \sigma_{u}^{4}}-\frac{1}{\sigma_{u}^{6}} u_{t}^{2}, \frac{\partial^{2} \ell_{t}(\theta)}{\partial \beta \partial \phi^{\prime}}=-\frac{1}{\sigma_{u}^{2}}\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right) \varepsilon_{t}^{\prime}-\frac{1}{\sigma_{u}^{2}} u_{t} \underline{x}_{t}^{\prime}, \\
& \frac{\partial^{2} \ell{ }_{t}(\theta)}{\partial \beta \partial \sigma_{u}^{2}}=-\frac{1}{\sigma_{u}^{4}} u_{t}\left(x_{t}-x_{t}^{\prime} \phi\right)^{\prime} \text { and }-\frac{\partial^{2} \ell_{t}(\theta)}{\partial \phi \partial \sigma_{u}^{2}}=-\frac{1}{\sigma_{u}^{4}} u_{t} \varepsilon_{t}^{\prime} .
\end{aligned}
$$

## APPENDIX B

## A Consistent Covariance Matrix Estimator for the Information Matrix Test

A consistent covariance estimator for the IM test proposed by White (1982) is stated as

$$
\begin{equation*}
\hat{v}(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n} a_{t}(\hat{\theta}) a_{t}^{\prime}(\hat{\theta}) \tag{B.1}
\end{equation*}
$$

where $a_{t}(\hat{\theta})=d_{t}(\hat{\theta})-\nabla \mathrm{d}(\hat{\theta}) \mathrm{A}(\hat{\theta})^{-1} \nabla \ell_{t}(\hat{\theta})$. Let us begin with the indicator vector $d(\hat{\theta})$ which is defined as

$$
d(\hat{\theta})=\operatorname{vech}[C(\hat{\theta})]=\operatorname{vech}[A(\hat{\theta})+B(\hat{\theta})]
$$

where

$$
\begin{aligned}
& A(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n}\left[\frac{\partial^{2} \ell_{t}(\theta)}{\partial \theta \partial \theta^{\prime}}\right] \theta=\hat{\theta} \\
& =\frac{1}{n} \sum_{t=1}^{n}
\end{aligned}
$$

and

$$
B(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n}\left[\left(\frac{\partial \ell_{t}(\theta)}{\partial \theta}\right)\left(\frac{\partial \ell_{t}(\theta)}{\partial \theta}\right)^{\prime}\right]_{\theta=\hat{\theta}}
$$


From $A(\hat{\theta})$ and $B(\hat{\theta}), C(\hat{\theta})$ is easily derived as

$$
\begin{aligned}
C(\hat{\theta}) & =A(\hat{\theta})+B(\hat{\theta}) \\
& =\frac{1}{n} \sum_{t=1}^{n}
\end{aligned}
$$

$$
\left\{\begin{array}{lll}
\frac{1}{\sigma_{u}^{4}}\left(u_{t}^{2}-\sigma_{u}^{2}\right)\left(x_{t}-x_{t}^{\prime} \phi\right)\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right)^{\prime} & \frac{1}{\sigma_{u}^{4}}\left(u_{t}^{2}-\sigma_{u}^{2}\right)\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right) \underline{\varepsilon}_{t}^{\prime}-\frac{1}{\sigma_{u}^{2}} u_{t} \underline{x}_{t}^{\prime} & \frac{1}{2 \sigma_{u}^{6}\left(x_{t}-x_{t}^{\prime} \phi\right)\left(u_{t}^{3}-3 \sigma_{u}^{2} u_{t}\right)} \\
\frac{1}{\sigma_{u}^{4}}\left(u_{t}^{2}-\sigma_{u}^{2}\right) \varepsilon_{t}\left(x_{t}-\underline{x}_{t}^{\prime} \phi\right)^{\prime}-\frac{1}{\sigma_{u}^{2}} u_{t} \underline{x}_{t} & \frac{1}{\sigma_{u}^{4}\left(u_{t}^{2}-\sigma_{u}^{2}\right) \varepsilon_{t} \underline{\varepsilon}_{t}^{\prime}} & \frac{1}{2 \sigma_{u}^{6} \underline{\varepsilon}_{t}\left(u_{t}^{3}-3 \sigma_{u}^{2} u_{t}\right)} \\
\frac{1}{2 \sigma_{u}^{6}\left(u_{t}^{3}-3 \sigma_{u}^{2} u_{t}\right)\left(x_{t}-x_{t}^{\prime} \phi\right)^{\prime}} & \frac{1}{2 \sigma_{u}^{6}\left(u_{t}^{3}-3 \sigma_{u}^{2} u_{t}\right) \varepsilon_{t}^{\prime}} & \frac{1}{4 \sigma_{u}^{8}\left(u_{t}^{4}-6 \sigma_{u}^{2} u_{t}^{2}+3 \sigma_{u}^{4}\right)}
\end{array}\right.
$$

Now it is stralghtforward to obtain $d(\hat{\theta})$. For analytical convenience, we rearrange $d(\hat{\theta})$ as described in the paper. Then the first part of $a_{t}(\hat{\theta})$ defined from $d(\hat{\theta})=\frac{1}{n} \sum_{t=1}^{n} d_{t}(\hat{\theta})$ can be written as

$$
\begin{equation*}
d_{t}(\hat{\theta})=\left(\hat{d}_{t 1}^{\prime}, \hat{d}_{t 2}^{\prime}, \hat{d}_{t 3}, \hat{d}_{t 4}^{\prime}, \hat{d}_{t 5}^{\prime}, \hat{d}_{t 6}^{\prime}\right)^{\prime} \tag{B.2}
\end{equation*}
$$

where

$$
\begin{aligned}
& \hat{d}_{t 1}=\left[\hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t 1}-x_{t 1}^{\prime} \hat{\phi}\right)^{2}, \ldots, \hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t k}-x_{t k}^{\prime} \hat{\phi}\right)^{2},\right. \\
& \hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t 1}-\underline{x}_{t 1}^{\prime} \hat{\phi}\right)\left(x_{t 2}-x_{t 2}^{\prime} \hat{\phi}\right), \ldots, \\
& \left.\hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t}(k-1)^{-x_{t}^{\prime}}(k-1)^{\prime}\right)\left(x_{t k}-x_{t k}^{\prime} \hat{\phi}\right)\right] '
\end{aligned}
$$

is a $\frac{k(k+1)}{2} \times 1$ vector,

$$
\begin{aligned}
& a_{t 2}=\left[\hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right) \hat{\varepsilon}_{t-1}^{2}, \ldots, \hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right) \hat{\varepsilon}_{t-p}^{2}, \hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right) \hat{\varepsilon}_{t-1} \hat{\varepsilon}_{t-2},\right. \\
& \left.\ldots, \hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)_{t-p+1} \hat{\varepsilon}_{t-p}\right]
\end{aligned}
$$

is a $\frac{p(p+1)}{2} \times 1$ vector,

$$
\hat{d}_{t 3}=\left(4 \hat{\sigma}_{u}^{8}\right)^{-1}\left(\hat{u}_{t}^{4}-6 \hat{\sigma}_{u}^{2} \hat{u}_{t}^{2}+3 \hat{\sigma}_{u}^{4}\right)
$$

is a scalar,

$$
\begin{aligned}
\hat{d}_{t 4}= & {\left[\hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t}-x_{t}^{\prime} \hat{\phi}\right)^{\prime} \hat{\varepsilon}_{t-1}, \ldots, \hat{\sigma}_{u}^{-4}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right)\left(x_{t}-x_{t}^{\prime} \hat{\phi}\right)^{\prime} \hat{\varepsilon}_{t-p}\right]^{\prime} } \\
& -\left[\theta_{u}^{-2} x_{t-1}^{\prime}, \ldots, \theta_{u}^{-2} x_{t-p}^{\prime}\right]^{\prime}
\end{aligned}
$$

is a $k p \times l$ vector,

$$
\hat{d}_{t 5}=\left(2 \hat{\sigma}_{u}^{6}\right)^{-1}\left(\hat{u}_{t}^{3}-3 \hat{\sigma}_{u}^{2} \hat{u}_{t}\right)\left(x_{t}-x_{t}^{\prime} \hat{\phi}\right)
$$

is a $k \times 1$ vector, and finally,

$$
\hat{d}_{t 6}=\left[\left(2 \hat{\sigma}_{u}^{6}\right)^{-1}\left(\hat{u}_{t}^{3}-3 \hat{\sigma}_{u}^{2} \hat{u}_{t}\right) \hat{\varepsilon}_{t-1}, \ldots,\left(2 \hat{\sigma}_{u}^{6}\right)^{-1}\left(\hat{u}_{t}^{3}-3 \hat{\sigma}_{u}^{2} \hat{u}_{t}\right) \hat{\varepsilon}_{t-p}\right]
$$

is a $p \times 1$ vector.
Next we consider

$$
\nabla d\left(\theta_{0}\right)=\lim _{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^{n} E\left[\frac{\partial d_{t}\left(\theta_{0}\right)}{\partial \theta}\right]
$$

Using the normality assumption of the $u_{t}$ and taking expectation conditional on the information set $\Psi_{t-1}$ iteratively, after some algebra we can get the following simple form of $\nabla \mathrm{d}\left(\theta_{0}\right)$

$$
\nabla d\left(\theta_{0}\right)=\left[\begin{array}{ccc}
0 & 0 & \nabla d_{13} \\
0 & 0 & \mathrm{~d}_{23} \\
0 & 0 & 0 \\
\nabla d_{41} & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{array}\right]
$$

where ${V d_{13}}=\left(m x_{11}, \ldots, m x_{k k}, m x_{12}, \ldots, m x_{(k-1){ }_{k}}\right)^{\prime}$ is a $\frac{k(k+1)}{2} \times 1$ vector with

$$
\begin{gathered}
m x_{i j}=-\frac{1}{\sigma_{u}^{4}} 1 i m_{n+\infty} \frac{1}{n} \sum_{t=1}^{n}\left(x_{t i}-x_{t i}^{\prime} \phi_{0}\right)\left(x_{t j}-x_{t j}^{\prime} \phi_{0}\right), \\
1, j=1,2, \ldots, k: \quad 1 \leq j, \\
\nabla d_{23}=\left(m \varepsilon_{11}, \ldots, m \varepsilon_{p p}, m \varepsilon_{12}, \ldots, m \varepsilon_{(p-1) p}\right)^{\prime} \text { is a } \frac{p(p+1)}{2} \times 1 \text { vector with } \\
m \varepsilon_{i j}=-\frac{1}{\sigma_{u}^{4}} 1 i m_{n+\infty} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t-1} \varepsilon_{t-j}, \\
1, j=1,2, \ldots, k: \quad i \leq j
\end{gathered}
$$

and $\nabla_{41}=\left(w_{11}, w_{12}, \ldots, w_{k p}\right)$ ' is a $k p \times k$ matrix with

$$
\begin{array}{r}
w_{i j}=\frac{1}{\sigma_{u}^{2}} 11 m_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^{n}\left(x_{t}-x_{t}^{\prime} \phi_{0}\right)^{\prime} x_{t-j, 1} \\
1=1,2, \ldots, k, j=1,2, \ldots, p .
\end{array}
$$

This implies that $\nabla \mathrm{d}\left(\theta_{0}\right)$ can be estimated consistently by the $\nabla \mathrm{d}(\hat{\theta})$ which is

$$
\operatorname{vd}(\hat{\theta})=\left[\begin{array}{ccc}
0 & 0 & \nabla \hat{\mathrm{~d}}_{13}  \tag{в.3}\\
0 & 0 & \nabla \hat{\mathrm{~d}}_{23} \\
0 & 0 & 0 \\
v \hat{\mathrm{~d}}_{41} & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{array}\right]
$$

where for example, $\nabla \hat{\mathrm{d}}_{13}=\left(\overline{m x}_{11}, \ldots, \overline{m x}_{\mathrm{kk}}, \overline{m x}_{12}, \ldots, \overline{m x}_{(k-1) k}\right)^{\prime}$ is a $\frac{k(k+1)}{2} \times 1$ vector with

$$
\begin{array}{r}
\overline{m x}_{i j}=-\frac{1}{n \theta_{u}^{4}} \sum_{t=1}^{n}\left(x_{t i}-\underline{x}_{t i}^{\prime} \hat{\phi}\right)\left(x_{t j}-\underline{x}_{t j}^{\prime} \hat{\phi}\right), \\
i, j=1,2, \ldots, k: \quad i \leq j .
\end{array}
$$

Similarly, we can simplify $A(\hat{\theta})$ as follows:

$$
A(\hat{\theta})=\left[\begin{array}{cccc}
-\frac{1}{n \theta_{u}^{2}} \sum_{t=1}^{n}\left(x_{t}-\underline{x}_{t}^{\prime} \hat{\phi}\right)\left(x_{t}-\underline{x}_{t}^{\prime} \hat{\phi}\right), & 0 & 0 \\
0 & -\frac{1}{n \theta_{u}^{2}} \sum_{t=1}^{n} \hat{\varepsilon}_{t} \hat{\varepsilon}_{t}^{\prime} & 0 \\
& 0 & 0 & -\frac{1}{2 \theta_{u}^{4}}
\end{array}\right] \text { (B.4) }
$$

For future use, let us denote the upper left-hand corner block of $A(\hat{\theta})$ as $A_{11}(\hat{\theta}) \equiv \hat{A}_{11}$. We can simplify the expression for $A(\hat{\theta})$ further by using analytic expectation of $\varepsilon_{t} \varepsilon_{t}^{\prime}$. For examplme, when $p=1$, $\mathrm{n}^{-1} \sum \varepsilon_{\mathrm{t}-1}^{2}$ can be replaced by $\delta_{\mathrm{u}}^{2} /\left(1-\hat{\phi}_{1}\right)$. This might provide better finite sample performance. However, then we will loose the $n R^{2}$ interpretation of our test statistics. Also, no general expression for the analytic expectation can be given for all values of $p$.

$$
\text { Finally, } \nabla \ell_{t}(\hat{\theta})=\frac{\partial \ell_{t}(\hat{\theta})}{\partial \theta} \text { is easily given from Appendix } A \text { by }
$$

$$
v v_{t}(\hat{\theta})=\left[\begin{array}{c}
\frac{1}{\theta_{u}^{2}} \hat{u}_{t}\left(x_{t}-\underline{x}_{t} \hat{\phi}\right)  \tag{B.5}\\
\frac{1}{\partial_{u}^{2}} \hat{u}_{t-\hat{\varepsilon}_{t}} \\
-\frac{1}{2 \hat{\theta}_{u}^{2}}+\frac{1}{2 \theta_{u}^{4}} \hat{u}_{t}^{2}
\end{array}\right]
$$

and we denote the first $(k \times 1)$ vector of $\nabla \ell_{t}(\hat{\theta})$ as $\nabla \ell_{t 1}(\hat{\theta}) \equiv \nabla \hat{\ell}_{t 1}$. For the following discussion, recall the definitions of $\underline{x}_{t}$, $\underline{\xi}_{t}$, $\underline{s}_{t}$ and $\underline{r}_{t}$, provided in the main text. From (B.2)-(B.5), $a_{t}(\hat{\theta})$ can be easily derived as

$$
\begin{equation*}
a_{t}(\hat{\theta})=d_{t}(\hat{\theta})-\nabla d(\hat{\theta}) A(\hat{\theta})^{-1} V_{t}(\hat{\theta})=\left(\hat{a}_{t 1}^{\prime}, \hat{a}_{t 2}^{\prime}, \hat{a}_{t 3}^{\prime}, \hat{a}_{t 4}^{\prime}, \hat{a}_{t 5}^{\prime}, \hat{a}_{t 6}^{\prime}\right) \prime \tag{B.6}
\end{equation*}
$$

where $\hat{a}_{t 1}=\frac{1}{\partial_{u}^{4}}\left(\hat{u}_{t}^{2}-\hat{a}_{u}^{2}\right)_{\underline{x}_{t}}, \hat{a}_{t 2}=\frac{1}{\partial_{u}^{4}}\left(\hat{u}_{t}^{2}-\hat{\sigma}_{u}^{2}\right) \underline{\xi}_{t}, \hat{a}_{t 3}=\hat{d}_{t 3}$, $\hat{a}_{t 4}=\hat{d}_{t 4}-v \hat{d}_{41} \hat{A}_{11}^{-1} \nabla \hat{\ell}_{t 1}, \hat{a}_{t 5}=\hat{d}_{t 5}$, and $\hat{a}_{t 5}=\hat{d}_{t 6}$.

Now we establish the block diagonality of the covariance matrix of the IM test, say $V\left(\theta_{0}\right)$. It is assumed that all conditions stated in White (1982) are satisfied. Given (B.2)-(B.6) with the normality assumption of the $u_{t}, V\left(\theta_{0}\right)$ takes the form of
$V\left(\theta_{0}\right)=\lim _{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^{n}$

where $W=V_{41} A_{11}^{-1} V d_{41}^{\prime}+n^{-1} \sigma_{u}^{-2} \Sigma \underline{z}_{t} \underline{z}_{t}^{\prime}$, and the diagonal elements are consistently estimated by $\hat{V}_{i}, i=1,2, \ldots, 6$, stated in the main text. To prove this result, let us consider

$$
\begin{equation*}
V\left(\theta_{0}\right)=\lim _{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^{n} E\left[a_{t}\left(\theta_{0}\right) a_{t}^{\prime}\left(\theta_{0}\right)\right] \tag{B.7}
\end{equation*}
$$

In the first stage, we evaluate $E\left[a_{t}\left(\theta_{0}\right) a_{t}^{\prime}\left(\theta_{0}\right)\right]$ conditional on the information set $\Psi_{t-1}$ using the normality assumption of the $u_{t}$ and taking expectation iteratively. In the next stage, we use the facts that at $\theta=\theta_{0}, E\left(\underline{\varepsilon}_{t}\right)=0, E\left(\underline{s}_{t}\right)=0$ and $E\left(\underline{\xi}_{t}\right)=0$ for all $t$. Then we have the result.

We are grateful to two referees for their very useful comments. Both of them pointed out some errors in our earlier derivation and made many helpful suggestions which improved the exposition of the paper. We also wish to express our appreciation to the participants of the 1988 North American Econometric Society Summer Meeting, Rob Engle, Bruce Hansen, Hal White, Jan Kmenta, Pravin Trivedi, Xiao-Lei Zuo, and in particular Alastair Hall for constructive comments on an earlier draft of the paper. All errors, of course, remain our own. Financial support from the Research Board and the Bureau of Economic and Business Research of the University of Illinois is gratefully acknowledged.

## REFERENCES

BERA, A. K. (1986), "Model Specification Test through Eigenvalues," Paper presented at the 1986 North-American Summer Meeting of the Econometric Society, Duke University, Durham.

BERA, A. K. and JARQUE, C. M. (1982), "Model Specification Tests: A Simultaneous Approach," Journal of Econometrics, 20, 59-82.

BERA, A. K. and MCKENZIE, C. R. (1986), "Alternative Forms and Properties of the Score Test," Journal of Applied Statistics, 13, 13-25. BERA, A. K. and ZUO, X-L. (1991), "Specification Tests for ARCH Models," mimeo, Department of Economics, University of Illinois at UrbanaChampaign.

BERA, A. K., LEE, S. and HIGGINS, M. L. (1990), "Interaction between Autocorrelation and Conditional Heteroskedasticity: A Random Coefficient Approach," Discussion Paper 90-25, Department of Economics, University of California at San Diego.

BOWMAN, K. O. and SHENTON, L. R. (1975), "Omnibus Contours for Departure from Normality Based on $b_{1}$ and $b_{2}$," Biometrika, 62, 243-250. BUMB, B. L. and KELEJIAN, H. H. (1983), "Autocorrelated and Heteroscedastic Disturbances in Linear Regression Analysis: A Monte Carlo Study," Sankhyā (Series B), 45, 257-270.

CHESHER, A. D. (1983), "The Information Matrix Test: Simplified Calculation via a Score Test," Economics Letters, 13, 45-48. CHESHER, A. D. (1984), "Testing for Neglected Heterogeneity," Econometrica, 52, 865-872.

DIEBOLD, F. X. (1986), "Testing for Serial Correlation in the Presence of ARCH," Proceedings of the American Statistical Association, Business and Economic Statistics Section, 323-328.

ENGLE, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," Econometrica, 50, 987-1007. EPPS, T. W. and EPPS, M. L. (1977), "The Robustness of Some Standard Tests for Autocorrelation and Heteroscedasticity When Both Problems are Present," Econometrica, 45, 745-753.

GODFREY, L. G. and WICKENS, M. R. (1982), "Tests of Misspecification Using Locally Equivalent Alternative Models," in Evaluating the Reliability of Macro-Economic Models, ed. by G. C. Chow and P. Corsi. New York: John Wiley \& Sons, pp. 71-103.

HALL, A. (1987), "The Information Matrix Test for the Linear Mode1," Review of Economic Studies, LIV, 257-263.

JARQUE, C. M. and BERA, A. K. (1987), "An Efficient Large-Sample Test for Normality of Observations and Regression Residuals," International Statistical Review, 55, 163-172.

KENDALI, M. G. and STUART, A. (1973), The Advanced Theory of
Statistics, Volume 2 (London: Charles Griffin).
LANCASTER, T. (1984), "The Covariance Matrix of the Information Matrix Test," Econometrica, 52, 1051-1053.

TSAY, R. S. (1987), "Conditional Heteroscedastic Time Series Models," Journal of the American Statistical Association, 82, 590-604.

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

WHITE, H. (1982), "Maximum Likelthood Estimation of Misspecified Models," Econometrica, 50, 1-25.

WHITE, H. (1987), "Specification Testing in Dynamic Mode1s," in
Advances in Econometrics: Fifth World Congress, Volume I, ed.
by T. F. Bewley, Cambridge: Cambridge University Press, pp. 1-58.

| (For previous papers please consult previous discussion papers.) |  |  |
| :---: | :---: | :---: |
| No. | Author(s) | Title |
| 9036 | J. Driffill | The Term Structure of Interest Rates: Structural Stability and Macroeconomic Policy Changes in the UK |
| 9037 | F. van der Ploeg | Budgetary Aspects of Economic and Monetary Integration in Europe |
| 9038 | A. Robson | Existence of Nash Equilibrium in Mixed Strategies for Games where Payoffs Need not Be Continuous in Pure Strategies |
| 9039 | A. Robson | An "Informationally Robust Equilibrium" for Two-Person Nonzero-Sum Games |
| 9040 | M.R. Baye, G. Tian and J. Zhou | The Existence of Pure-Strategy Nash Equilibrium in Games with Payoffs that are not Quasiconcave |
| 9041 | M. Burnovsky and <br> I. Zang | "Costless" Indirect Regulation of Monopolies with Substantial Entry Cost |
| 9042 | P.J. Deschamps | Joint Tests for Regularity and Autocorrelation in Allocation Systems |
| 9043 | S. Chib, J. Osiewalski and M. Steel | Posterior Inference on the Degrees of Freedom Parameter in Multivariate-t Regression Models |
| 9044 | H.A. Keuzenkamp | The Probability Approach in Economic Methodology: On the Relation between Haavelmo's Legacy and the Methodology of Economics |
| 9045 | I.M. Bomze and E.E.C. van Damme | A Dynamical Characterization of Evolutionarily Stable States |
| 9046 | E. van Damme | On Dominance Solvable Games and Equilibrium Selection Theories |
| 9047 | J. Driffill | Changes in Regime and the Term Structure: A Note |
| 9048 | A.J.J. Talman | General Equilibrium Programming |
| 9049 | H.A. Keuzenkamp and F. van der Ploeg | Saving, Investment, Government Finance and the Current Account: The Dutch Experience |
| 9050 | C. Dang and A.J.J. Talman | The $D_{1}$-Triangulation in Simplicial Variable Dimension Algorithms on the Unit Simplex for Computing Fixed Points |
| 9051 | M. Baye, D. Kovenock and C. de Vries | The All-Pay Auction with Complete Information |
| 9052 | H. Carlsson and E. van Damme | Global Games and Equilibrium Selection |

\(\left.$$
\begin{array}{lll}\text { No. } & \text { Author(s) } & \text { Title } \\
9053 & \begin{array}{l}\text { M. Baye and } \\
\text { D. Kovenock }\end{array} & \begin{array}{l}\text { How to Sell a Pickup Truck: "Beat-or-Pay" } \\
\text { Advertisements as Facilitating Devices }\end{array} \\
9054 & \text { Th. van de Klundert }\end{array}
$$ \quad \begin{array}{l}The Ultimate Consequences of the New Growth <br>

Theory; An Introduction to the Views of M.\end{array}\right\}\)| Fitzgerald Scott |
| :--- | :--- |


| No. | Author(s) | Title |
| :---: | :---: | :---: |
| 9072 | G. Alogoskoufis and <br> F. van der Ploeg | Endogenous Growth and Overlapping Generations |
| 9073 | K.C. Fung | Strategic Industrial Policy for Cournot and Bertrand Oligopoly: Management-Labor Cooperation as a Possible Solution to the Market Structure Dilemma |
| 9101 | A. van Soest | Minimum Wages, Earnings and Employment |
| 9102 | A. Barten and <br> M. McAleer | Comparing the Empirical Performance of Alternative Demand Systems |
| 9103 | A. Weber | EMS Credibility |
| 9104 | G. Alogoskoufis and <br> F. van der Ploeg | Debts, Deficits and Growth in Interdependent Economies |
| 9105 | R.M.W.J. Beetsma | Bands and Statistical Properties of EMS Exchange Rates |
| 9106 | C.N. Teulings | The Diverging Effects of the Business Cycle on the Expected Duration of Job Search |
| 9107 | E. van Damme | Refinements of Nash Equilibrium |
| 9108 | E. van Damme | Equilibrium Selection in $2 \times 2$ Games |
| 9109 | G. Alogoskoufis and <br> F. van der Ploeg | Money and Growth Revisited |
| 9110 | L. Samuelson | Dominated Strategies and Commom Knowledge |
| 9111 | F. van der Ploeg and <br> Th. van de Klundert | Political Trade-off between Growth and Government Consumption |
| 9112 | Th. Nijman, F. Palm and C. Wolff | Premia in Forward Foreign Exchange as Unobserved Components |
| 9113 | H. Bester | Bargaining vs. Price Competition in a Market with Quality Uncertainty |
| 9114 | R.P. Gilles, G. Owen and R. van den Brink | Games with Permission Structures: The Conjunctive Approach |
| 9115 | F. van der Ploeg | Unanticipated Inflation and Government Finance: The Case for an Independent Common Central Bank |
| 9116 | N. Rankin | Exchange Rate Risk and Imperfect Capital Mobility in an Optimising Model |
| 9117 | E. Bomhoff | Currency Convertibility: When and How? A Contribution to the Bulgarian Debate! |
| 9118 | E. Bomhoff | Stability of Velocity in the G-7 Countries: A Kalman Filter Approach |


| No. | Author(s) | Title |
| :---: | :---: | :---: |
| 9119 | J. Osiewalski and <br> M. Steel | Bayesian Marginal Equivalence of Elliptical Regression Models |
| 9120 | S. Bhattacharya, <br> J. Glazer and <br> D. Sappington | Licensing and the Sharing of Knowledge in Research Joint Ventures |
| 9121 | J.W. Friedman and <br> L. Samuelson | An Extension of the "Folk Theorem" with Continuous Reaction Functions |
| 9122 | S. Chib, J. Osiewalski and M. Steel | A Bayesian Note on Competing Correlation Structures in the Dynamic Linear Regression Model |
| 9123 | Th. van de Klundert and L. Meijdam | Endogenous Growth and Income Distribution |
| 9124 | S. Bhattacharya | Banking Theory: The Main Ideas |
| 9125 | J. Thomas | Non-Computable Rational Expectations Equilibria |
| 9126 | J. Thomas and T. Worrall | Foreign Direct Investment and the Risk of Expropriation |
| 9127 | T. Gao, A.J.J. Talman and $Z$. Wang | Modification of the Kojima-Nishino-Arima Algorithm and its Computational Complexity |
| 9128 | S. Altug and R.A. Miller | Human Capital, Aggregate Shocks and Panel Data Estimation |
| 9129 | H. Keuzenkamp and A.P. Barten | Rejection without Falsification - On the History of Testing the Homogeneity Condition in the Theory of Consumer Demand |
| 9130 | G. Mailath, L. Samuelson and J. Swinkels | Extensive Form Reasoning in Normal Form Games |
| 9131 | K. Binmore and <br> L. Samuelson | Evolutionary Stability in Repeated Games Played by Finite Automata |
| 9132 | L. Samuelson and <br> J. Zhang | Evolutionary Stability in Asymmetric Games |
| 9133 | J. Greenberg and <br> S. Weber | Stable Coalition Structures with Unidimensional Set of Alternatives |
| 9134 | F. de Jong and <br> $F$. van der Ploeg | Seigniorage, Taxes, Government Debt and the EMS |
| 9135 | E. Bomhoff | Between Price Reform and Privatization Eastern Europe in Transition |
| 9136 | H. Bester and <br> E. Petrakis | The Incentives for Cost Reduction in a Differentiated Industry |


| No. | Author(s) | Title |
| :---: | :---: | :---: |
| 9137 | L. Mirman, L. Samuelson and E. Schlee | Strategic Information Manipulation in Duopolies |
| 9138 | C. Dang | The $\mathrm{D}_{2}^{*}$-Triangulation for Continuous Deformation Algorithms to Compute Solutions of Nonlinear Equations |
| 9139 | A. de Zeeuw | Comment on "Nash and Stackelberg Solutions in a Differential Game Model of Capitalism" |
| 9140 | B. Lockwood | Border Controls and Tax Competition in a Customs Union |
| 9141 | C. Fershtman and <br> A. de Zeeuw | Capital Accumulation and Entry Deterrence: A Clarifying Note |
| 9142 | J.D. Angrist and <br> G.W. Imbens | Sources of Identifying Information in Evaluation Models |
| 9143 | A.K. Bera and A. Ullah | Rao's Score Test in Econometrics |
| 9144 | B. Melenberg and <br> A. van Soest | Parametric and Semi-Parametric Modelling of Vacation Expenditures |
| 9145 | G. Imbens and <br> T. Lancäster | Efficient Estimation and Stratified Sampling |
| 9146 | Th. van de Klundert and S. Smulders | Reconstructing Growth Theory: A Survey |
| 9147 | J. Greenberg | On the Sensitivity of Von Neuman and Morgenstern Abstract Stable Sets: The Stable and the Individual Stable Bargaining Set |
| 9148 | S. van Wijnbergen | Trade Reform, Policy Uncertainty and the Current Account: A Non-Expected Utility Approach |
| 9149 | S. van Wijnbergen | Intertemporal Speculation, Shortages and the Political Economy of Price Reform |
| 9150 | G. Koop and M.F.J. Steel | A Decision Theoretic Analysis of the Unit Root Hypothesis Using Mixtures of Elliptical Models |
| 9151 | A.P. Barten | Consumer Allocation Models: Choice of Functional Form |
| 9152 | R.T. Baillie, T. Bollerslev and M.R. Redfearn | Bear Squeezes, Volatility Spillovers and Speculative Attacks in the Hyperinflation 1920s Foreign Exchange |
| 9153 | M.F.J. Steel | Bayesian Inference in Time Series |
| 9154 | A.K. Bera and S. Lee | Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis |



17000011174829

