FACULTY WORKING PAPER 93-0112

Specification Test for a Linear Regression Model with Arch Process

Anil K. Bera
Department of Economics
University of Illinois

Xiao-Lei Zuo
Department of Economics and Statistics
National University of Singapore

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

BEBR

FACULTY WORKING PAPER NO. 93-0112

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

February 1993

Specification Test for a Linear Regression Model with Arch Process

Anil K. Bera
Department of Economics

Xiao-Lei Zuo
Department of Economics and Statistics

This version: January 1993

SPECIFICATION TEST FOR A LINEAR REGRESSION MODEL WITH ARCH PROCESS*

ANIL K. BERA
DEPARTMENT OF ECONOMICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

AND

XIAO-LEI ZUO
DEPARTMENT OF ECONOMICS AND STATISTICS
NATIONAL UNIVERSITY OF SINGAPORE

ABSTRACT

ARCH models are used widely in analyzing economic and financial tme series data. Many tests are available to detect the presence of ARCH; however, there is no acceptable procedure available for testing an estimated ARCH model. In this paper we develop a test for a linear regression model with ARCH disturbances using the framework of the information matrix (IM) test. For the ARCH specification, the covariance matrix of the indicator vector is not block diagonal, and the IM test is turned out to be a test for variation in the fourth moment, i.e, a test for heterokurtosis. An illustrative example is provided to demonstrate the usefulness of the proposed test.

Key Words: Autoregressive Conditional Heteroskedasticity; Information Matrix Test; Double Length Regression.

AMS Classification Number: 62J02

* An earlier version of this paper was presented at the 1992 Midwest Econometrics Group Meeting, Federal Reserve Bank of Minneapolis, U.S.A. We would like to thank the participants, and especially the discussant Danny Quah for helpful comments. Thanks are also due to Sungsup Ra for research assistance. Any remaining errors are our responsibility. Financial support from the Research Board of the University of Illinois is gratefully acknowledged. .

1.INTRODUCTION

In a seminal paper, Engle (1982) introduced the autoregressive conditional heteroscedastic (ARCH) models. These models are now very popular in analyzing financial and economic time series data [for a recent review, see Bera and Higgins (1993)]. There are many procedures available to detect the presence of ARCH. However, estimated ARCH models are not, in general, tested thoroughly, possibly because there is no acceptable procedure for doing that. In this paper we derive a simple specification test for an estimeted ARCH model in the linear regression framework using White's (1982) information matrix (IM) test principle.

The plan of the paper is as follows. In section 2, we specify the model and derive an algebraic structure of the IM test. Comparing with Hall (1987) and Bera and Lee (1992) who applied the IM test to linear regression model without and with serial correlation respectively, in the ARCH framework, the covariance matrix of the indicator vector no longer has a block diagonal structure due to the inclusion of the ARCH coefficients in the parameter vector. The algebraic structure of the test is much more complicated. First we derive a joint test and then concentrate on the components corresponding to the ARCH coefficients. The test turns out to be a test for time varying fourth moment, i.e a test for heterokurtosis. The test statistic can be computed by running a simple regression, and it can be given Chesher (1984)'s interpretation of Lagrange multiplier (LM) test for parameter heterogeneity. Local power of the test is also considered here. In section 3, the results are interpreted. An alternative form of the IM test is also computed by using the double-length regression proposed by Davidson and MacKinnon (1992) and is presented in section 4. Section 5 discusses an empirical example to illustrate the usefulness of our test. A concluding summary is given in section 5.

2. MODEL AND TEST

We consider the linear regression model

$$y_t = x_t'\beta + \epsilon_t \tag{1}$$

where y_t is the t-th observation on endogenous variable, x_t is a $k \times 1$ vector of exogenous variables, and ϵ_t's are assumed to follow an ARCH process. As specified in Engle (1982), an ARCH(p) process conditional on the information set Φ_{t-1} is described as

$$\epsilon_t \mid \Phi_{t-1} \sim N(0, h_t)$$

where

$$h_t = V(\epsilon_t \mid \Phi_{t-1}) = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \alpha_2\epsilon_{t-2}^2 + \ldots + \alpha_p\epsilon_{t-p}^2, \tag{2}$$

and $\alpha_0 > 0$, $\alpha_i \geq 0$, $\sum_{i=1}^{p} \alpha_i < 1$. Let $\underline{\epsilon_t} = (\epsilon_{t-1}^2, \epsilon_{t-2}^2, \ldots, \epsilon_{t-p}^2)'$, and $\alpha = (\alpha_1, \ldots, \alpha_p)'$. Assuming that $\underline{\epsilon_t}$ is given, the loglikelihood function is the sum of the conditional normal loglikelihoods function corresponding to (1) and (2). For our ARCH case, all assumptions mentioned in White (1987) are satisfied and the IM test can be applied to this model.

Let $L(\theta)$ be the average loglikelihood function and l_t be the log density function of the t-th observation and T be the sample size. Then

$$L(\theta) = \frac{1}{T}\sum_{t=1}^{T} l_t(\theta)$$

$$l_t(\theta) = -\frac{1}{2}\log\pi - \frac{1}{2}\log h_t - \frac{1}{2}\frac{\epsilon_t^2}{h_t}$$

where $\theta = (\beta', \alpha_0, \alpha')'$ is a $(k+p+1) \times 1$ vector of parameters.

Suppose $\hat{\theta}$ be the maximum likelihood estimator (MLE) of θ. Then, the IM test is based on the distinct elements of the matrix

$$C(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} [\frac{\partial^2 l_t(\hat{\theta})}{\partial\theta\partial\theta'} + (\frac{\partial l_t(\hat{\theta})}{\partial\theta})(\frac{\partial l_t(\hat{\theta})}{\partial\theta})']$$
$$= A(\hat{\theta}) + B(\hat{\theta}).$$

where

$$A(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} (\frac{\partial^2 l_t(\hat{\theta})}{\partial\theta\partial\theta'})$$

$$B(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} [(\frac{\partial l_t(\hat{\theta})}{\partial\theta})(\frac{\partial l_t(\hat{\theta})}{\partial\theta})']$$

Since $C(\hat{\theta})$ is symmetric, IM test just depends on

$$d(\hat{\theta}) = vechC(\hat{\theta}) = vech(A(\hat{\theta}) + B(\hat{\theta})),$$

which is a $m \times 1$ vector where $m = \frac{(p+k+1)(p+k+2)}{2}$. Subject to certain regularity conditions, it can be shown that the asymptotic covariance of $d(\hat{\theta})$ can be consistently estimated by [see White (1982, p. 11)]

$$Cov(\hat{d}) = V(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} a_t(\hat{\theta}) a_t(\hat{\theta})'$$

where

$$a_t(\hat{\theta}) = d_t(\hat{\theta}) - \nabla d(\hat{\theta}) A(\hat{\theta})^{-1} \nabla l_t(\hat{\theta})$$
$$\nabla d(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} \frac{\partial d_t(\hat{\theta})}{\partial\theta}$$
$$\nabla l_t(\hat{\theta}) = \frac{\partial l_t(\hat{\theta})}{\partial\theta}$$

The IM test can be written as

$$T_{IM} = Td'(\hat{\theta}) V(\hat{\theta})^{-1} d(\hat{\theta})$$

Therefore, under null hypothesis that the model (1)-(2) is a correct specification, the IM test statistic is asymptotically distributed as χ^2 with m degrees of freedom. We define $d(\hat{\theta})$ explicitly as follows:

$$d(\hat{\theta}) = (\hat{d_1}', \hat{d_2}, \hat{d_3}', \hat{d_4}', \hat{d_5}', \hat{d_6}')'$$

where $\hat{d}_1$ is a $\frac{k(k+1)}{2} \times 1$ vector; $\hat{d}_2$ is a scalar; $\hat{d}_3$ is a $\frac{p(p+1)}{2} \times 1$ vector; $\hat{d}_4$ is a $k \times 1$ vector; $\hat{d}_5$ is a $pk \times 1$ vector; $\hat{d}_6$ is a $p \times 1$ vector. The typical elements of $\hat{d}_{i,i=1,\ldots,6}$ are given below [for derivation, see Appendices A and B]:

$$\hat{d}_1 : \frac{1}{T}\sum_{t=1}^{T}[2(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\sum_{h=1}^{p}\hat{\alpha}_h\hat{\epsilon}_{t-h}^{*}x_{t-hi}^{*}\sum_{h=1}^{p}\hat{\alpha}_h\hat{\epsilon}_{t-h}^{*}x_{t-hj}^{*} + (\hat{\epsilon}_t^{*2} - 1)(x_{ti}^{*}x_{tj}^{*} - \sum_{h=1}^{p}\hat{\alpha}_h x_{t-hi}^{*}x_{t-hj}^{*})$$
$$- (\hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1) - 2\hat{\epsilon}_t^{*})\sum_{h=1}^{p}\hat{\alpha}_h\hat{\epsilon}_{t-i}^{*}x_{ti}^{*}x_{t-hj}^{*} - (\hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1) - 2\hat{\epsilon}_t^{*})\sum_{h=1}^{p}\hat{\alpha}_h\hat{\epsilon}_{t-h}^{*}x_{t-hi}^{*}x_{tj}^{*}]_{i\le j,i,j=1,2,\ldots,k}$$

$$\hat{d}_2 : \frac{1}{T}\sum_{t=1}^{T}[\frac{1}{2\hat{h}_t^2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})]$$

$$\hat{d}_3 : \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2}[(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\hat{\epsilon}_{t-i}^{*2}\hat{\epsilon}_{t-j}^{*2}]_{i\le j,i,j=1,2,\ldots,p}$$

$$\hat{d}_4 : -\frac{1}{T}\sum_{t=1}^{T}[\frac{1}{2\hat{h}_t}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\sum_{h=1}^{p}\hat{\alpha}_h\hat{\epsilon}_{t-h}^{*}x_{t-hi}^{*}$$
$$+ \frac{1}{2\hat{h}_t}(\hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1) - 2\hat{\epsilon}_t^{*})x_{ti}^{*}]_{i=1,2,\ldots,k}$$

$$\hat{d}_5 : -\frac{1}{T}\sum_{t=1}^{T}[\frac{1}{2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\sum_{h=1}^{p}\hat{\alpha}_h\hat{\epsilon}_{t-h}^{*}x_{t-hi}^{*}\hat{\epsilon}_{t-j}^{*2}$$
$$- \frac{1}{2}(\hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1) - 2\hat{\epsilon}_t^{*})x_{ti}^{*}\hat{\epsilon}_{t-j}^{*2}$$
$$+ \frac{1}{2}(\hat{\epsilon}_t^{*2} - 1)\hat{\epsilon}_{t-j}^{*}x_{t-ji}^{*}]_{i=1,2,\ldots,k,j=1,2,\ldots,p}$$

$$\hat{d}_6 : \frac{1}{T}\sum_{t=1}^{T}[\frac{1}{2\hat{h}_t}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}*2_t + \frac{3}{2})\hat{\epsilon}_{t-i}^{*2}]_{i=1,2,\ldots,p}$$

where

$$\hat{\epsilon}_t^* = \frac{\hat{\epsilon}_t}{\sqrt{\hat{h}_t}}, \quad \hat{\epsilon}_{t-i}^* = \frac{\hat{\epsilon}_{t-i}}{\sqrt{\hat{h}_t}}, \quad x_{ij}^* = \frac{x_{ij}}{\sqrt{\hat{h}_t}}, \quad x_{t-ij}^* = \frac{x_{t-ij}}{\sqrt{\hat{h}_t}}.$$

The variance matrix of $d(\hat{\theta})$ is not block diagonal. [For detailed derivation of $Var(d(\hat{\theta}))$, see Appendix B]. Therefore, the derived IM test statistics can not be written as the sum of quadratic forms as in Hall (1987) and Bera and Lee (1992).

As is well known, the IM formulaton tests for the full specification of the model (1)-(2). If we are interested only in the specification of ARCH part, we may concentrate in those particular components corresponding to the ARCH parameters. These components are based on the indicator vectors $d_3(\hat{\theta}), d_5(\hat{\theta})$ and $d_6(\hat{\theta})$. One of these three indicator vectors, $d_3(\hat{\theta})$ is related solely to the ARCH parameter vector α, and now we formulate a test based on this component. The other two components will be discussed in the next section.

To find the asymptotic variance of $\hat{d}_3$, we use a result from Pierce (1981). Let $y_1, y_2, ..., y_n$ be a sequence of random variables whose joint distribution depends on a parameter θ . Let $\hat{\theta}_n = \hat{\theta}_n\ (y_1, y_2, ..., y_n)$ be an asymptotically normal and efficient sequence of estimators. It is desired to find the limiting distribution of a statistic $\hat{T}_n = T_n(y_1, y_2, ..., y_n; \hat{\theta}_n)$, where at the true θ the corresponding sequence $T_n = T_n(y_1, y_2, ..., y_n; \theta)$ has a known limiting normal distribution. We assume that for every θ there is joint convergence in law to normality:

$$\begin{pmatrix} \sqrt{n}T_n \\ \sqrt{n}(\hat{\theta} - \theta) \end{pmatrix} \sim N\left[0, \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}\right].$$

and that there is a matrix B, possibly depending continuously on θ, such that

$$\sqrt{n}\hat{T}_n = \sqrt{n}T_n + B\sqrt{n}(\hat{\theta} - \theta) + o_p(1)$$

where $B = \lim E(\frac{\partial T_n}{\partial \theta})$. Under the above assumptions,

$$\sqrt{n}\hat{T}_n \sim N(0, V_{11} - BV_{22}B').$$

Using this result, an estimate of the variance of $\hat{d}_3$ has the following form:

$$\hat{V}_3 = \frac{3}{2T}\sum_{t=1}^{T}(\hat{\epsilon}^{*2}_{t-i}\hat{\epsilon}^{*2}_{t-j})(\hat{\epsilon}^{*2}_{t-i}\hat{\epsilon}^{*2}_{t-j})_{i\le j, i,j=1,2,\ldots,p}$$

[The derivation is in Appendix C: Part I]. Therefore, the test statistic can be written as

$$T_3 = \hat{d}_3'\hat{V}_3^{-1}\hat{d}_3$$

which asymptotically follows a chi-square distribution with $\frac{p(p+1)}{2}$ degrees of freedom.

The test T_3 can be perfomed as TR^2 of the ordinary least squares regression of $\hat{v}$ on $\hat{S}$, i.e.,

$$T_3 = \frac{1}{6}\hat{v}'\hat{S}(\hat{S}'\hat{S})^{-1}\hat{S}'\hat{v}$$

where $\hat{v} = (\hat{v}_1, \hat{v}_2, ..., \hat{v}_T)'$ is a $T \times 1$ vector with $\hat{v}_t = (\frac{1}{2}\hat{\epsilon}^{*4}_t - 3\hat{\epsilon}^{*2}_t + \frac{3}{2})$ and

$$\hat{S} = \begin{pmatrix} \hat{\epsilon}^{*2}_{1-1}\hat{\epsilon}^{*2}_{1-1} & \cdots & \hat{\epsilon}^{*2}_{1-1}\hat{\epsilon}^{*2}_{1-p} & \cdots & \hat{\epsilon}^{*2}_{1-p+1}\hat{\epsilon}^{*2}_{1-p} \\ . & \cdots & . & \cdots & . \\ . & \cdots & . & \cdots & . \\ . & \cdots & . & \cdots & . \\ \hat{\epsilon}^{*2}_{T-1}\hat{\epsilon}^{*2}_{T-1} & \cdots & \hat{\epsilon}^{*2}_{T-1}\hat{\epsilon}^{*2}_{T-p} & \cdots & \hat{\epsilon}^{*2}_{T-p+1}\hat{\epsilon}^{*2}_{T-p} \end{pmatrix}$$

is a $T \times \frac{p(p+1)}{2}$ matrix. [The detailed derivation is given in Appendix C: Part II].

3. INTERPRETATION OF THE INFORMATION MATRIX TEST

From the last section, each of the six components of the indicator vector contains the common element $v_t = \frac{1}{2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})$. Under normality and correct specification of ARCH model, $Ev_t = 0$. . So the sample moment $\frac{1}{T}\sum_{t=1}^{T} v_t$ would be expected to be close to zero. Hence a test for model being correct can be based on $\frac{1}{T}\sum_{t=1}^{T} v_t$, a measure of sample kurtosis.

Here we are interested in the special components related to the ARCH parameter α. As we mentioned in the last section, there are three such components, $\hat{d}_3, \hat{d}_5, \hat{d}_6$ which are related to the ARCH parameter α and have a special form $\sum_{t=1}^{T}(v_t g_t)$ where g_t is some function. Therefore, these can be considered different tests for heterokurtosis and each emphasize the effects of different aspects of heterokurtosis.

The component $\hat{d}_3$ is related to the parameter vector of α. It is clear from the expression $C(\hat{\theta})$ that $\hat{d}_3$ is based on the two estimates of variance of $\hat{\alpha}$. Taking $E(v_t)$ as a measure of kurtosis. $\hat{d}_3$ measures the change in the kurtosis. more precisely it tests whether the kurtosis depends on the cross product of the lagged residual squares. Following Chesher (1984), we can also give a test for heterogeneity interpretation to $\hat{d}_3$. Suppose the ARCH parameters α are varying around a mean with finite variance. This can be formulated as $\alpha_t \sim (\alpha, \Omega)$. Then T_3 is the LM test for testing $H_0 : \Omega = 0$, i.e. it tests the randomness of the ARCH parameters.

Next, $\hat{d}_5$ is based on the relationship between $\hat{\beta}$ and $\hat{\alpha}$. The assumption of a symmetric distribution implies $E\hat{\epsilon}_t^{*3} = 0$, $E\hat{\epsilon}_t^{*} = 0$ and that allows us to omit the second part of $\hat{d}_5$. The third part can also be omitted because it deals with heteroskedasticity arising from the cross products between $\hat{\epsilon}^*_{t-j}$ and x^*_{t-ji} and is not very important from a practical point of view. Then, $\hat{d}_5$ reduces to $\tilde{d}_5 = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2}(\frac{1}{2}(\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} - \frac{3}{2})\sum_{l=1}^{p}\hat{\alpha}_l\hat{\epsilon}^*_{t-l}x^*_{t-li}\hat{\epsilon}^{*2}_{t-j})_{i=1,2,\ldots,k,j=1,2,\ldots,p}$ and this describes a relationship between v_t and $\sum_{l=1}^{p}\hat{\alpha}_l\hat{\epsilon}^*_{t-l}x^*_{t-li}\hat{\epsilon}^{*2}_{t-j}$ and allows us to test heterokurtosis caused by cross products between lagged error terms and lagged exogenous variables. $\hat{d}_6$ is an expression arising from the two estimators of the covariance between $\hat{\sigma}^2$ and $\hat{\alpha}$, and can be used to test the heterokurtosis due to the square lagged errors.

Since the covariance matrix $\hat{V}$ is not block diagonal, these three test statistics are definitely correlated with each other and also with the other three components. To get overall test of the model, it is necessary to have a joint specification test. This can be obtained by using the results in Appendix B.

4.DOUBLE LENGTH REGRESSION FORM OF THE TEST

Davidson and MacKinnon (1992) proposed a double-length regression (DLR) to perform the IM tests on models which can be expressed as

$$f_t(y_t, \theta) = \epsilon_t, \quad t=1,\ldots,T, \quad \epsilon_t \sim N(0,1).$$

For this class of models, the contribution to the loglikelihood function from observation t is

$$l_t = -\frac{1}{2}log(2\pi) - \frac{1}{2}f_t^2 + k_t,$$

where $k_t = log \mid \frac{\partial f_t}{\partial y_t} \mid = log \mid f_t' \mid$ is the Jacobian contribution to l_t. The DLR uses 2T "observations." The regressand is $\hat{f}_t$ for "observation" t and one for "observation" $T+n$, $n = 1, 2, \ldots, T$, and the corresponding regressors are respectively $-\hat{F}_{ti}$ and $\hat{K}_{ti}$, where $F_{ti} = \frac{\partial f_t}{\partial \theta_i}$ and $K_{ti} = \frac{\partial k_t}{\partial \theta_i}$. The test statistic is then the explained sum of squares from this artificial regression. In order to obtain the DLR form of the IM test, an explicit alternative hypothesis of the model $f_t(y_t, \theta) = \epsilon_t$ is needed. Chesher's (1984) result suggests the use of the following model:

$$f_t(y_t, \theta + \zeta_t) = \epsilon_t, \epsilon_t \sim N(0, 1), \tag{3}$$

where $\theta = (\beta', \alpha_0, \alpha')'$ is a $k + 1 + p$ dimensional vector , and $\zeta_t = (0', 0, \eta_t')'$ and η_t are $k + 1 + p$ and p dimensional random vectors respectively, with η_t being distributed independently of ϵ_t and of η_s, $s \neq t$. $\eta_t \sim N(0, 2\Omega)$. By taking a second order Taylor series expansion of (3) in ζ_t, we have

$$f_t(y_t, \theta) \equiv \epsilon_t - \overline{F}_t\zeta_t - \frac{1}{2}\zeta_t'\overline{F}_t^*\zeta_t = \epsilon_t - F_t\eta_t - \frac{1}{2}\eta_t'F_t^*\eta_t = g_t \ (say),$$

where $\overline{F}_t$ is a $1 \times (k + 1 + p)$ row vector with typical element $\frac{\partial f_t}{\partial \theta_i}$ and $\overline{F}_t^*$ is a $(k + 1 + p) \times (k + 1 + p)$ matrix with typical element $\frac{\partial^2 f_t}{\partial \theta_i \partial \theta_j}$. F_t is a $1 \times p$ row vector with typical element $\frac{\partial f_t}{\partial \alpha_i}$ and F_t^* is a $p \times p$ matrix with typical element $\frac{\partial^2 f_t}{\partial \alpha_i \partial \alpha_j}$. Note that

$$E(g_t \mid y_t) = -tr(\Omega F_t^*).$$

and

$$Var(g_t \mid y_t) = 1 + 2tr(\Omega F_t{}^T F_t).$$

where "T" denotes transpose of a matrix. Thus locally in the neighborhood of $\Omega = 0$, the model is equivalent to

$$\frac{f_t(y_t, \theta) - E(g_t \mid y_t)}{\sqrt{V(g_t \mid y_t)}} = \vartheta_t, \ \vartheta_t \sim N(0, 1). \tag{4}$$

i.e,

$$q_t(y_t, \theta, \Omega) \equiv \frac{f_t(y_t, \theta) + tr(\Omega F_t^*)}{(1 + 2tr(\Omega F_t{}^T F_t))^{\frac{1}{2}}} = \vartheta_t.$$

The loglikelihood function of this model can be written as

$$l_t = -\frac{1}{2}log(2\pi) - \frac{1}{2}q_t^2 + r_t,$$

where $r_t = log \mid \frac{\partial q_t}{\partial y_t} \mid$. Now a LM test for $\Omega = 0$ in the above model can be computed from the following regression

$$\begin{pmatrix} f_1 \\ . \\ . \\ . \\ f_T \\ 1 \\ . \\ . \\ . \\ 1 \end{pmatrix} = \begin{pmatrix} Q_{1ij} \\ . \\ . \\ . \\ Q_{Tij} \\ R_{1ij} \\ . \\ . \\ . \\ R_{Tij} \end{pmatrix} (\Omega_{ij}) + \omega_t$$

$$= \begin{pmatrix} -\frac{1}{4h_t^2}(3\frac{\epsilon_t}{\sqrt{h_t}} - \frac{\epsilon_t^3}{h_t^{\frac{3}{2}}})\tilde{\epsilon}_t \\ \frac{1}{4h_t^2}(3\frac{1}{\sqrt{h_t}} + \frac{1}{h_t} - (1 + \frac{2}{\sqrt{h_t}})\frac{\epsilon_t^2}{h_t})\tilde{\epsilon}_t \end{pmatrix} (\Omega_{ij}) + \omega_t$$

where $i = 1, ..., p, j \leq i$, and $\tilde{\epsilon}_t = (\epsilon_{t-1}^4, ..., \epsilon_{t-p+1}^2 \epsilon_{t-p}^2)$ is a $1 \times \frac{p(p+1)}{2}$ vector. [Detailed derivation is given in Appendix D: Part I].

Since the loglikelihood functions of the Davidson and MacKinnon's and Chesher's models are the same under $H_0 : \Omega = 0$ although they differ under $H_1 : \Omega \neq 0$, the constrained MLE obtained from those loglikelhood functions are the same. And the score vectors of those loglikelihood functions at the constrained MLE are also equal. Using the terminology of Godfrey (1988, p.72) we can then say that the models (3) and (4) are locally equivalent alternative (LEA) to the ARCH model given in (1)-(2). Hence the LM test statistic for H_0 should be the same for both the loglikelihood functions.

Given the above results, the theoretical local power of T_3 test is equivalent to that of the LM test with the Davidson and MacKinnon's model. To derive the noncentrality parameter of the LM test we consider the simple case of $p = 1$. Note that now we have $\theta = (\beta', \alpha_0, \alpha_1)'$ is a $(k + 1 + 1) \times 1$ vector, $\zeta = \alpha_1 + \eta$ and $\eta \sim N(0, 2\sigma^2)$. For the LM test, the null and alternative hypotheses are

$$H_0 : V(\eta) = 2\sigma^2 = 0$$

and

$$H_a : V(\eta) = 2\sigma^2 = \frac{\delta}{\sqrt{T}}.$$

and similarly for the IM test

$$H_0 : h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$$

and

$$H_a : h_t = \alpha_0 + (\alpha_1 + \eta)\epsilon_{t-1}^2, \tag{5}$$

where $\eta \sim N(0, 2\sigma^2)$. Thus, locally in the neighborhood of $\sigma^2 = 0$, the model is eqivalent to

$$q_t(y_t, \theta, \sigma^2) \equiv \frac{f_t(y, \theta) + \sigma^2 F_{t33}}{(1 + 2\sigma^2 F_{t3} F_{t3})^{\frac{1}{2}}} = \mu_t.$$

with log-density function

$$l_t = -\frac{1}{2} log(2\pi) - \frac{1}{2} q_t^2 + r_t.$$

The variance of the test can be obtained using $E[-\frac{\partial^2 l_t}{\partial(\sigma^2)^2}]$. Therefore, the test statistic under the local alternative follows a noncentral χ^2 with 1 degree of freedom and noncentral parameter $\delta^2 \frac{1}{T} E \sum_{t=1}^{T} -\frac{\partial^2 l_t}{\partial(\sigma^2)^2}$. [Detail derivation is in the Appendix D: Part II.] The variance we got here is the same as we obtained for T_3 in Appendix C.

Following the procedure of Engle (1982), it is easy to find the second and fourth moments of a first-order random coefficient ARCH process i.e. of model (5). These moments will give some idea about the nature of the implicit alternatives for the IM test. Letting $\lambda_t = (\epsilon_t^4, \epsilon_t^2)'$,

$$E(\lambda_t \mid \Phi_{t-1}, \zeta) = \begin{pmatrix} 3\alpha_0^2 \\ \alpha_0 \end{pmatrix} + \begin{pmatrix} 3\zeta^2 & 6\alpha_0\zeta \\ 0 & \zeta \end{pmatrix} \lambda_{t-1}.$$

The condition for finite unconditional variance is same as in the non-random coefficient case, that is, $\alpha_1 < 1$, while to have a finite fourth moment it is now required that $3(\alpha_1^2 + 2\sigma^2) < 1$. For the standard ARCH(1) model, this condition is $3\alpha_1^2 < 1$. If these conditions are met, the moments can be computed as [For derivation, see Appendix D: Part III].

$$E(\epsilon_t^4) = [\frac{3\alpha_0^2}{(1-\alpha_1)^2}][\frac{1-\alpha_1^2}{1-3(\alpha_1^2+2\sigma^2)}]$$

$$E(\epsilon_t^2) = \frac{\alpha_0}{1-\alpha_1}.$$

Therefore, the kurtosis of the random coefficient ARCH(1) model is $3(1-\alpha_1^2)/[1-3(\alpha_1^2 + 2\sigma^2)]$ which is higher than $3(1-\alpha_1^2)/[1-3\alpha_1^2]$, the kurtosis for the standard ARCH(1) model, for $\sigma^2 \neq 0$. Since the unconditional variance remains the same, our alternative hypothesis can take acount of higher degree of nonnormality.

5. AN EMPIRICAL ILLUSTRATION

Engle(1983) estimated the following reduced form equation for inflation using quarterly data from 1947-IV-1979-IV

$$\dot{P}_t = \beta_1 + \beta_2(\dot{P}_{t-1}) + \beta_3(\dot{P}_{t-2} + \beta_4(\dot{P}M_{t-1})$$
$$+ \beta_5(\dot{W}_{t-1}) + \beta_6(\dot{M}_{t-1}) + \beta_7(t) + \epsilon_t$$

$$\epsilon_t \mid \Phi_{t-1}, \sim N(0, h_t)$$
$$h_t = \sigma^2 + \gamma \sum_{j=1}^{8} (\frac{9-j}{36})\epsilon_{t-j}^2$$

where $\dot{P}, \dot{P}M, \dot{W}, \dot{M}$ and t are respectively the rates of change of the GNP deflator, the rate of change of the import deflator, the rate of change of wages, the rate of change of money supply and a time trend. This inflation equation includes two lagged dependent variables and the conditional variance h_t is assumed to exhibit a two-parameter ARCH process of eight-order with linearly declining weights.

As reported in Engle (1983), the above model satisfies the standard diagnostic checks. Also Bera, Higgins and Lee (1990) found that above model passes the Pagan and Sabau (1987) test for correctly specified conditional variance. The T_3 test examines this model at $p = 1, 2, 3, 4, 5$. For $p = 1$ and 2, the values of T_3 are respectively 0.445 and 6.569.

These are not significant at both 1% and 5% significant levels for which the asymptotic critical values are 6.635 and 3.841 with 1 degree of freedom, and 11.345 and 7.815 with 3 degrees of freedom respectively. However, for $p = 3, 4, 5$, the test statistics, 61.98, 139.60, and 199.34 respectively, are highly significant at both 1% and 5% levels. This may imply that T_3 can reveal model misspecification, while the standard diagnostic checks may fail to do so.

5. CONCLUSION

In this paper, we have provided an application of IM test to the linear model with ARCH process. We give the computation and interpretation of the resulting test. Because ARCH error are involved, the framework of the information matrix test is much more complicated than those derived by Hall (1987) and Bera and Lee (1992) for simpler models. In our case the variance of the indicator vector is no longer a block diagonal matrix, and therefore, the components of the indicate vector are not asymptotically independent. We applied one component of IM test to the Engle (1983) model. The test we use above may have higher power than the tests which Engle (1983) had used for his model in the sense that the proposed test rejects the model specification while other diagnostic checks can not.

REFERENCES

Bera, A. K. and Higgins, M. L., (1993), A Survey of ARCH Models: Properties, Estimation and Testing, *Journal of Economics Surveys*, forthcoming.

Bera, A. K. & Lee, S., (1992), Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis, *Review of Economic Studies*, forthcoming.

Bera, A. K., Lee, S. & Higgins, M. L., (1990), On the Formulation of a General Structure for Conditional Heteroskedasticity, Discussion paper $90 - 41$, Department of Economics, University of California at San Diego.

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

Davidson, R. and Mackinnon, J.G., (1992), A New Form of the Information Matrix Test, *Econometrica*, 60, 1, 145-157.

Engle, R. F., (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, $50, 987 - 1007$.

Engle, R. F., (1983), Estimates of the Variance of U. S. Inflation Based upon the ARCH Model, *Journal of Money, Credit and Banking*, $15, 286 - 301$.

Godfrey, L. G.,(1988), *Misspecification Test in Econometrics: The Lagrange multiplier Principle and other Approaches*, Cambridge University Press, Cambridge, UK.

Hall, A., (1987), Information Matrix Test for the Linear Model, *Review of Econmic Studies*, $54, 257 - 263$.

Lancaster, T., (1984), The Covariance Matrix of the Information Matrix Test, *Econometrica*, $52, 1051 - 1053$.

Pagan, A. R., and Sabau, H. C. L., (1987), On the Inconsistency Tests for Heteroskedastic and Risk models, Mimeo, University of Rochester.

Pierce, D. A. (1982), The Asymptotic Effect of Substituting Estimators for Parameters in Certain Types of Statistics, *The Annals of Statistics*, $10, 475 - 478$.

White, H. (1982), Maximum Likelihood Estimation of Misspecified Models, *Econometrica*, $50, 1 - 25$.

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

APPENDIX A

The Derivatives of the Log-likelihood Function

For our model, the parameter vector is $\theta = (\beta', \alpha_0, \alpha')$' and the log-likelihood function for the t-th observation conditional on the information set Φ_{t-1} is given by

$$l(\theta) = \frac{1}{2}log2\pi - \frac{1}{2}logh_t - \frac{1}{2h_t}\epsilon_t^2.$$

The first derivatives are

$$\frac{\partial l_t}{\partial\beta} = (\frac{\epsilon_t^2}{h_t})\frac{1}{2h_t}\frac{\partial h_t}{\partial\beta} - \frac{1}{2h_t}\frac{\partial h_t}{\partial\beta} + \frac{\epsilon_t}{h_t}x_t$$

$$\frac{\partial l_t}{\partial\alpha_0} = \frac{1}{2h_t}(\frac{\epsilon_t^2}{h_t} - 1)$$

$$\frac{\partial l_t}{\partial\alpha} = \frac{1}{2h_t}(\frac{\epsilon_t^2}{h_t} - 1)\underline{\epsilon_t}$$

and second order derivatives are

$$\frac{\partial^2 l_t}{\partial\beta\partial\beta'} = -\frac{\epsilon_t}{h_t^2}x_t\frac{\partial h_t}{\partial\beta'} - \frac{\epsilon_t}{h_t^2}\frac{\partial h_t}{\partial\beta}x_t' - \frac{1}{h_t}x_tx_t' - \frac{1}{h_t^2}\frac{\epsilon_t^2}{2h_t}\frac{\partial h_t}{\partial\beta}\frac{\partial h_t}{\partial\beta'}$$
$$- \frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\frac{\partial h_t}{\partial\beta}\frac{\partial h_t}{\partial\beta'} + (\frac{\epsilon_t^2}{h_t} - 1)\frac{1}{2h_t}\frac{\partial^2 h_t}{\partial\beta\partial\beta'}$$

$$\frac{\partial^2 l_t}{\partial\alpha_0\partial\alpha_0} = -\frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1) - \frac{1}{2h_t^2}\frac{\epsilon_t^2}{h_t}$$

$$\frac{\partial^2 l_t}{\partial\alpha\partial\alpha'} = -\frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\underline{\epsilon_t}\underline{\epsilon_t}' - \frac{1}{2h_t^2}\frac{\epsilon_t^2}{h_t}\underline{\epsilon_t}\underline{\epsilon_t}'$$

$$\frac{\partial^2 l_t}{\partial\beta\partial\alpha_0} = -\frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\frac{\partial h_t}{\partial\beta} - \frac{1}{2h_t^2}\frac{\epsilon_t^2}{h_t}\frac{\partial h_t}{\partial\beta} - \frac{1}{h_t^2}\epsilon_tx_t$$

$$\frac{\partial^2 l_t}{\partial\beta\partial\alpha'} = -\frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\frac{\partial h_t}{\partial\beta}\underline{\epsilon_t}' - \frac{1}{2h_t^2}\frac{\epsilon_t^2}{h_t}\frac{\partial h_t}{\partial\beta}\underline{\epsilon_t}'$$
$$- \frac{1}{h_t^2}\epsilon_tx_t\underline{\epsilon_t}' + \frac{1}{2h_t}(\frac{\epsilon_t^2}{h_t} - 1)\frac{\partial^2 h_t}{\partial\beta\partial\alpha'}$$

$$\frac{\partial^2 l_t}{\partial\alpha_0\partial\alpha'} = -\frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\underline{\epsilon_t}' - \frac{1}{2h_t^2}\frac{\epsilon_t^2}{h_t}\underline{\epsilon_t}'$$

where

$$\frac{\partial h_t}{\partial \beta} = -\sum_{i=1}^{p} 2\alpha_i \epsilon_{t-i} x_{t-i}$$

and

$$\frac{\partial^2 h_t}{\partial \beta \partial \beta'} = \sum_{i=1}^{p} 2\alpha_i x_{t-i} x'_{t-i}.$$

The outer products are

$$\frac{\partial l_t}{\partial \beta}\frac{\partial l_t}{\partial \beta}' = (\frac{\epsilon_t^2}{h_t} - 1)^2 \frac{1}{4h_t^2}\frac{\partial h_t}{\partial \beta}\frac{\partial h_t}{\partial \beta}' + \frac{\epsilon_t}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\frac{\partial h_t}{\partial \beta} x'_t$$

$$+ \frac{\epsilon_t}{2h_t^2} x_t \frac{\partial h_t}{\partial \beta}' (\frac{\epsilon_t^2}{h_t} - 1) + \frac{1}{h_t}\frac{\epsilon_t^2}{h_t} x_t x'_t$$

$$\frac{\partial l_t}{\partial \beta}\frac{\partial l_t}{\partial \alpha_0}' = \frac{1}{4h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)^2 \frac{\partial h_t}{\partial \beta} + \frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\epsilon_t x_t$$

$$\frac{\partial l_t}{\partial \beta}\frac{\partial l_t}{\partial \alpha}' = \frac{1}{4h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)^2 \frac{\partial h_t}{\partial \beta}\underline{\epsilon_t}' + \frac{1}{2h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)\epsilon_t x_t \underline{\epsilon_t}'$$

$$\frac{\partial l_t}{\partial \alpha_0}\frac{\partial l_t}{\partial \alpha_0}' = \frac{1}{4h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)^2$$

$$\frac{\partial l_t}{\partial \alpha_0}\frac{\partial l_t}{\partial \alpha}' = \frac{1}{4h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)^2 \underline{\epsilon_t}'$$

$$\frac{\partial l_t}{\partial \alpha}\frac{\partial l_t}{\partial \alpha}' = \frac{1}{4h_t^2}(\frac{\epsilon_t^2}{h_t} - 1)^2 \underline{\epsilon_t \epsilon_t}'.$$

APPENDIX B

Covariance Matrix for the Information Matrix Test

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

$$V(\hat{\theta}) = \frac{1}{T}\sum_{t=1}^{T} a_t(\hat{\theta})a_t(\hat{\theta})' \tag{B.1}$$

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

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

where

$$A(\hat{\theta}) = \frac{1}{T}\sum_{t=1}^{T}[\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta'}]_{\theta=\theta} = \frac{1}{T}\sum_{t=1}^{T}$$

$$\begin{pmatrix} \frac{2\epsilon_t}{h_t^2}x_t\xi_t' + \frac{2\epsilon_t}{h_t^2}\xi_t x_t' - \frac{x_t x_t'}{h_t} - \frac{2}{h_t^2}\kappa_t\chi_t + \frac{1}{h_t}\delta_t\zeta_t & \frac{1}{h_t^2}\kappa_t\xi_t - \frac{\epsilon_t x_t}{h_t^2} & \frac{1}{h_t^2}\kappa_t\xi_t\underline{\epsilon_t}' - \frac{\epsilon_t x_t}{h_t^2}\underline{\epsilon_t}' - \frac{1}{h_t}\delta_t\eta_t \\ \frac{1}{h_t^2}\kappa_t\xi_t' - \frac{\epsilon_t}{h_t^2}x_t' & -\frac{1}{2h_t^2}\kappa_t & -\frac{1}{2h_t^2}\kappa_t\underline{\epsilon_t}' \\ \frac{1}{h_t^2}\kappa_t\underline{\epsilon_t}\xi_t' - \frac{\epsilon_t}{h_t^2}\underline{\epsilon_t}x_t' - \frac{1}{h_t}\delta_t\eta_t' & -\frac{1}{h_t^2}\kappa_t\underline{\epsilon_t} & -\frac{1}{h_t^2}\kappa_t\underline{\epsilon_t}\underline{\epsilon_t}' \end{pmatrix}_{\theta=\hat{\theta}}$$

and

$$B(\hat{\theta}) = \frac{1}{T}\sum_{t=1}^{T}[(\frac{\partial l_t}{\partial\theta})(\frac{\partial l_t}{\partial\theta})']_{\theta=\theta} = \frac{1}{T}\sum_{t=1}^{T}$$

$$\begin{pmatrix} \frac{1}{h_t^2}\delta_t^2\chi_t - \frac{\epsilon_t}{h_t^2}\delta_t\xi_t x_t' - \frac{\epsilon_t}{h_t^2}x_t\xi_t'\delta_t + \frac{\epsilon_t^2}{h_t^2}x_t x_t' & -\frac{1}{2h_t^2}\delta_t^2\xi_t + \frac{\epsilon_t x_t}{2h_t^2}\delta_t & -\frac{1}{2h_t^2}\delta_t^2\xi_t\underline{\epsilon_t}' + \frac{\epsilon_t}{2h_t^2}x_t\delta_t\underline{\epsilon_t}' \\ -\frac{1}{2h_t^2}\delta_t^2\xi_t' + \frac{\epsilon_t}{2h_t^2}x_t'\delta_t & \frac{1}{4h_t^2}\delta_t^2 & \frac{1}{4h_t^2}\delta_t^2\underline{\epsilon_t}' \\ -\frac{1}{2h_t^2}\delta_t^2\underline{\epsilon_t}\xi_t' + \frac{\epsilon_t}{2h_t^2}\delta_t\underline{\epsilon_t}x_t' & \frac{1}{4h_t^2}\delta_t^2\underline{\epsilon_t} & \frac{1}{4h_t^2}\delta_t^2\underline{\epsilon_t}\underline{\epsilon_t}' \end{pmatrix}_{\theta=\hat{\theta}}$$

where

$$\frac{\epsilon_t^2}{h_t} + (\frac{\epsilon_t^2}{h_t} - 1) = \kappa_t$$

$$(\frac{\epsilon_t^2}{h_t} - 1) = \delta_t$$

$$\sum_{i=1}^{p} \alpha_i \epsilon_{t-i} x_{t-i} = \xi_t$$

$$\sum_{i=1}^{p} \alpha_i x_{t-i} x'_{t-i} = \zeta_t$$

$$\sum_{i=1}^{p} \alpha_i^2 \epsilon_{t-i}^2 x_{t-i} x'_{t-i} + \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_i \alpha_j \epsilon_{t-i} \epsilon_{t-j} x_{t-i} x'_{t-j} = \chi_t$$

and η_t is a $k \times p$ matrix with (i.j)th element $\epsilon_{t-i} x_{t-ij}{}_{,i=1,2,\dots,p}$ and ${}_{j=1,2,\dots,k}$. From $A(\hat{\theta})$ and $B(\hat{\theta})$, $C(\hat{\theta})$ is derived as

$$C(\hat{\theta}) = A(\hat{\theta}) + B(\hat{\theta}) = \frac{1}{T} \sum_{t-1}^{T}$$

$$\begin{pmatrix} \frac{1}{h_t}\delta_t\mu_t + \frac{2}{h_t^2}\lambda_t\chi_t + \frac{1}{h_t}x_t v_t \xi'_t + \frac{1}{h_t} v_t \xi_t x'_t & -\frac{1}{h_t^2}\lambda_t\xi_t - \frac{1}{2h_t}x_t v_t & -\frac{1}{h_t^2}\lambda_t \xi_t \underline{\epsilon_t}' - \frac{1}{2h_t} v_t x_t \underline{\epsilon_t}' - \frac{1}{h_t}\delta_t\eta_t \\ -\frac{1}{h_t^2}\lambda_t \xi'_t - \frac{1}{2h_t} v_t x'_t & \frac{1}{2h_t^2}\lambda_t & \frac{1}{2h_t^2}\lambda_t \underline{\epsilon_t}' \\ -\frac{1}{h_t^2}\lambda_t \underline{\epsilon_t} \xi'_t - \frac{1}{2h_t} v_t \underline{\epsilon_t} x'_t - \frac{1}{h_t}\delta_t \eta'_t & \frac{1}{2h_t^2}\lambda_t \underline{\epsilon_t} & \frac{1}{2h_t^2}\lambda_t \underline{\epsilon_t}\underline{\epsilon_t}' \end{pmatrix}$$

where

$$(\frac{1}{2}\frac{\epsilon_t^4}{h_t^2} - 3\frac{\epsilon_t^2}{h_t} + \frac{3}{2}) = \lambda_t$$

$$\frac{2\epsilon_t}{h_t} - \frac{\epsilon_t}{h_t}(\frac{\epsilon_t^2}{h_t} - 1) = v_t$$

$$[x_t x'_t + \sum_{i=1}^{p} \alpha_i x_{t-i} x'_{t-i}] = \mu_t$$

Therefore, $d(\hat{\theta})$ is given by

$$d(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} d_t(\hat{\theta})$$

where

$$d_t(\hat{\theta}) = (\hat{d}'_{t1}, \hat{d}'_{t2}, \hat{d}'_{t3}, \hat{d}'_{t4}, \hat{d}'_{t5}, \hat{d}'_{t6})'$$

and

$$\hat{d}_{t1} = [(\hat{\epsilon}_t^{*2} - 1)[x_{t1}^{*2} + \sum_{i=1}^{p} \hat{\alpha}_i x_{t-i1}^{*2}] + 2(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})$$

$$(\sum_{i=1}^{p} \hat{\alpha}_i^2 \hat{\epsilon}_{t-i}^{*2} x_{t-i1}^{*2} + 2\sum_{i\leq j}\sum_{j=1}^{p} \hat{\alpha}_i\hat{\alpha}_j\hat{\epsilon}_{t-i}^{*}\hat{\epsilon}_{t-j}^{*} x_{t-i1}^{*} x_{t-j1}^{*})$$

$$+ 2(2\hat{\epsilon}_t^{*} - \hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1))(\sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i}^{*} x_{t1}^{*} x_{t-i1}^{*} + \sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i}^{*} x_{t-i1}^{*} x_{t1}^{*}), \ldots,$$

$$(\hat{\epsilon}_t^{*2} - 1)[x_{tk}^{*2} + \sum_{i=1}^{p} \hat{\alpha}_i x_{t-ik}^{*2}] + 2(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})$$

$$(\sum_{i=1}^{p} \hat{\alpha}_i^2 \hat{\epsilon}_{t-i}^{*2} x_{t-ik}^{*2} + 2\sum_{i\leq j}\sum_{j=1}^{p} \hat{\alpha}_i\hat{\alpha}_j\hat{\epsilon}_{t-i}^{*}\hat{\epsilon}_{t-j}^{*} x_{t-ik}^{*} x_{t-jk}^{*})$$

$$+ 2(2\hat{\epsilon}_t^{*} - \hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1))(\sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i}^{*} x_{tk}^{*} x_{t-ik}^{*} + \sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i}^{*} x_{t-ik}^{*} x_{tk}^{*}),$$

$$(\hat{\epsilon}^{*2} - 1)[x_{t1}^{*} x_{t2}^{*} + \sum_{i=1}^{p} \hat{\alpha}_i x_{t-i1}^{*} x_{t-i2}^{*}] + 2(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})$$

$$(\sum_{i=1}^{p} \hat{\alpha}_i^2 \hat{\epsilon}_{t-i}^{*2} x_{t-i1}^{*} x_{t-i2}^{*} + 2\sum_{i\leq j}\sum_{j=1}^{p} \hat{\alpha}_i\hat{\alpha}_j\hat{\epsilon}_{t-i}^{*}\hat{\epsilon}_{t-j}^{*} x_{t-i1}^{*} x_{t-j2}^{*})$$

$$+ 2(2\hat{\epsilon}_t^{*} - \hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1))(\sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i}^{*} x_{t1}^{*} x_{t-i2}^{*} + \sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i}^{*} x_{t-i1}^{*} x_{t2}^{*}), \ldots,$$

$$(\hat{\epsilon}_t^{*2} - 1)[x_{t(k-1)}^{*} x_{tk}^{*} + \sum_{i=1}^{p} \hat{\alpha}_i x_{t-i(k-1)}^{*} x_{t-ik}^{*}] + 2(\frac{1}{2}\hat{\epsilon}^{*4} - 3\hat{\epsilon}^{*2} + \frac{3}{2})$$

$$(\sum_{i=1}^{p} \hat{\alpha}_i^2 \hat{\epsilon}_{t-i}^{*2} x_{t-i(k-1)}^{*} x_{t-ik}^{*} + 2\sum_{i\leq j}\sum_{j=1}^{p} \hat{\alpha}_i\hat{\alpha}_j\hat{\epsilon}_{t-i}^{*}\hat{\epsilon}_{t-j}^{*} x_{t-i(k-1)}^{*} x_{t-jk}^{*})$$

$$+ 2(2\hat{\epsilon}_t^{*} - \hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1))(\sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i} x_{t(k-1)}^{*} x_{t-ik}^{*} + \sum_{i=1}^{p} \hat{\alpha}_i\hat{\epsilon}_{t-i}^{*} x_{t-i(k-1)}^{*} x_{tk}^{*})]'$$

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

$$\hat{d}_{t2} = \frac{1}{2\hat{h}_t^2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})$$

is a scalar,

$$\hat{d}_{t3} = [\frac{1}{2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\hat{\epsilon}_{t-1}^{*4}, \ldots,$$
$$\frac{1}{2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\hat{\epsilon}_{t-p}^{*4},$$
$$\frac{1}{2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\hat{\epsilon}_{t-1}^{*2}\hat{\epsilon}_{t-2}^{*2}, \ldots,$$
$$\frac{1}{2}(\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\hat{\epsilon}_{t-p+1}^{*2}\hat{\epsilon}_{t-p}^{*2}]'$$

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

$$\hat{d}_{t4} = [-\frac{1}{2\hat{h}_t}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\sum_{i=1}^{p}\hat{\alpha}_i\hat{\epsilon}_{t-i}^{*}x_{t-i}^{*} - \frac{1}{2\hat{h}_t}(2\hat{\epsilon}_t^{*} - \hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1))x_t^{*}]$$

is a $k \times 1$ vector,

$$\hat{d}_{t5} = [-\frac{1}{2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\sum_{i=1}^{p}\hat{\alpha}_i\hat{\epsilon}_{t-i}^{*}x_{t-i}^{*}\hat{\epsilon}_{t-1}^{*2} - \frac{1}{2}(2\hat{\epsilon}_t^{*} - \hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1))x_t^{*}\hat{\epsilon}_{t-1}^{*2}$$
$$- \frac{1}{2}(\hat{\epsilon}_t^{*2} - 1)x_{t-1}^{*}\hat{\epsilon}_{t-1}^{*}, \ldots,$$
$$- \frac{1}{2}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\sum_{i=1}^{p}\hat{\alpha}_i\hat{\epsilon}_{t-i}^{*}x_{t-p}^{*}\hat{\epsilon}_{t-p}^{*2} - \frac{1}{2}(2\hat{\epsilon}_t^{*} - \hat{\epsilon}_t^{*}(\hat{\epsilon}_t^{*2} - 1))x_t^{*}\hat{\epsilon}_{t-p}^{*2}$$
$$- \frac{1}{2}(\hat{\epsilon}_t^{*2} - 1)x_{t-p}^{*}\hat{\epsilon}_{t-p}^{*}]'$$

is a $kp \times 1$ vector,

$$\hat{d}_{t6} = [\frac{1}{2\hat{h}_t}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}^{*2} + \frac{3}{2})\hat{\epsilon}_{t-1}^{*2}, \ldots, \frac{1}{2\hat{h}_t}(\frac{1}{2}\hat{\epsilon}_t^{*4} - 3\hat{\epsilon}_t^{*2} + \frac{3}{2})\hat{\epsilon}_{t-p}^{*2}]'$$

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

$$\nabla d(\theta_0) = \lim_{n \to \infty}\frac{1}{T}\sum_{t=1}^{T}E[\frac{\partial d_t(\theta_0)}{\partial\theta}].$$

Using the normality assumption of the ϵ_t and taking expectation conditional on the information set Φ_{t-1} iteratively, after some algebra we can get the following simple form of $\nabla d(\theta_0)$

$$\nabla d(\theta_0) = \begin{pmatrix} 0 & \nabla d_{12} & \nabla d_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \nabla d_{51} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

where $\nabla d_{12} = (mx_{11}, ..., mx_{kk}, mx_{12}, ..., mx_{(k-1)k})'$ is a $\frac{k(k+1)}{2} \times 1$ vector with

$$mx_{ij} = -\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{1}{\hat{h}_t^2}(x_{ti}x_{tj} + \sum_{l=1}^{p} \alpha_l x_{(t-l)i}x_{(t-l)j})$$

$i \leq j, i, j = 1, 2, ..., k$.

$\nabla d_{13} = (w_{11}, ..., w_{kk}, w_{12}, ..., w_{(k-1)k})'$ is a $\frac{k(k+1)}{2} \times p$ matrix with

$$w_{ij} = -\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{1}{\hat{h}_t^2}(x_{ti}x_{tj} + \sum_{l=1}^{p} \alpha_l x_{(t-l)i}x_{(t-l)j})\underline{\epsilon}_t$$

$i \leq j, i, j = 1, 2, ..., k$

and $\nabla d_{51} = (z_{11}, z_{22}, ..., z_{pp})'$ is a $pk \times k$ matrix with

$$z_{ii} = -\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{2}{\hat{h}_t^2}[\epsilon_{t-i}x_{t-i}(\sum_{j=1}^{p} \alpha_j \epsilon_{t-j}x'_{t-j})]$$

$i = 1, 2, ..., p$

. This implies that $\nabla d(\theta_0)$ can be estimated consistently by the $\nabla d(\hat{\theta})$ which is

$$\nabla d(\hat{\theta}) = \begin{pmatrix} 0 & \nabla \hat{d}_{12} & \nabla \hat{d}_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \nabla \hat{d}_{51} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where for example, $\nabla \hat{d}_{12} = (\overline{mx}_{11}, ..., \overline{mx}_{kk}, \overline{mx}_{12}, ..., \overline{mx}_{(k-1)k})'$ is a $\frac{k(k+1)}{2} \times 1$ vector with

$$\overline{mx}_{ij} = -\frac{1}{T\hat{h}_t} \sum_{t=1}^{T} (x^*_{ti}x^*_{tj} + \sum_{l=1}^{p} \hat{\alpha}_l x^*_{(t-l)i}x^*_{(t-l)j})$$

$$i \leq j, i, j = 1, 2, \ldots, k.$$

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

$$A(\hat{\theta}) = \begin{pmatrix} -\frac{1}{T}\sum_{t=1}^{T}(x_t^* x_t^{*\prime} + 2\sum_{i=1}^{p}\hat{\alpha}_i^2 \epsilon_{t-i}^{*2} x_{t-i}^* x_{t-i}^{*\prime}) & 0 & 0 \\ 0 & -\frac{1}{2h_t^2} & -\frac{1}{2h_t^2}\underline{\epsilon}_t^{\prime} \\ 0 & -\frac{1}{2h_t^2}\underline{\epsilon}_t & -\frac{1}{2h_t^2}\underline{\epsilon}_t\underline{\epsilon}_t^{\prime} \end{pmatrix}$$

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

$$\nabla l_t(\hat{\theta}) = \begin{pmatrix} -(\hat{\epsilon}^{*2} - 1)\sum_{i=1}^{p}\hat{\alpha}_i \hat{\epsilon}_{t-i}^* x_{t-i}^* + \epsilon_t^* x_t^* \\ \frac{1}{2h_t}(\hat{\epsilon}_t^{*2} - 1) \\ \frac{1}{2h_t}(\hat{\epsilon}_t^{*2} - 1)\underline{\epsilon}_t \end{pmatrix}$$

Given the above expressions for $d(\hat{\theta})$, $\nabla d(\hat{\theta})$, $A(\hat{\theta})$, and $\nabla l_t(\hat{\theta})$, using the formula (B.1) we can obtain an estimate of the covariance matrix for the IM test. Unlike the cases of Hall (1987) and Bera and Lee (1992), here $A(\hat{\theta})$ is not block-diagonal, and this results in $V(\hat{\theta})$ to be non-block-diagonal. And the final expression of $V(\hat{\theta})$ is very complicated and is omitted.

APPENDIX C

Part I.

Here

$$d_3 = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2}\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2}(\frac{1}{2}\epsilon_t^{*4} - 3\epsilon_t^{*2} + \frac{3}{2})$$

$$i \leq j, i, j = 1, \ldots p$$

$$\epsilon_t^* = \frac{\epsilon_t}{\sqrt{h_t}} \sim N(0,1)$$

Using

$$V(x) = E[V(x \mid y)] + V[E(x \mid y)],$$

we obtain

$$V(d_3) = V_{11} = E[\frac{1}{T}\sum_{t=1}^{T}\frac{1}{4}(\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2})'(\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2})6].$$

Defining

$$\hat{V}_{11} = \frac{1}{T}\sum_{t=1}^{T}(\hat{\epsilon}_{t-i}^{*2}\hat{\epsilon}_{t-j}^{*2})'(\hat{\epsilon}_{t-i}^{*2}\hat{\epsilon}_{t-j}^{*2})\frac{3}{2}_{i \leq j, i, j = 1, \ldots, p}$$

We have

$$plim_{t \to \infty}\hat{V}_{11} = V_{11}$$

$$\hat{d}_3'\hat{V}_{11}^{-1}\hat{d}_3 \sim \chi^2_{\frac{p(p+1)}{2}}$$

Also note that

$$\begin{aligned}
\frac{\partial d_3}{\partial \beta} &= (2\epsilon_t^{*3}\frac{\partial \epsilon_t^*}{\partial \beta} - 6\epsilon_t^*\frac{\partial \epsilon_t^*}{\partial \beta})(\frac{\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2}}{2}) + (\frac{1}{2}\epsilon_t^{*4} - 3\epsilon_t^{*2} + \frac{3}{2})\epsilon_{t-j}^{*2}2\epsilon_{t-i}^*\frac{\partial \epsilon_{t-i}^*}{\partial \beta} \\
&\quad + (\frac{1}{2}\epsilon_t^{*4} - 3\epsilon_t^{*2} + \frac{3}{2})\epsilon_{t-i}^{*2}2\epsilon_{t-j}^*\frac{\partial \epsilon_{t-j}^*}{\partial \beta} - \frac{1}{h_t}(-\epsilon_t^{*4} + 3\epsilon^{*2})\frac{\partial h_t}{\partial \beta}\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2} \\
&\quad - \frac{2}{h_t}(\frac{1}{2}\epsilon_t^{*4} - 3\epsilon_t^{*2} + \frac{3}{2})\frac{\partial h_t}{\partial \beta}\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2} \\
\frac{\partial d_3}{\partial \alpha_0} &= \frac{1}{h_t}(-\epsilon_t^{*4} + 3\epsilon_t^{*2})\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2} - \frac{2}{h_t}(\frac{1}{2}\epsilon_t^{*4} - 3\epsilon_t^{*2} + \frac{3}{2})\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2} \\
\frac{\partial d_3}{\partial \alpha} &= \frac{1}{h_t}(-\epsilon_t^{*4} + 3\epsilon_t^{*2})\underline{\epsilon_t}\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2} - \frac{2}{h_t}(\frac{1}{2}\epsilon_t^{*4} - 3\epsilon_t^{*2} + \frac{3}{2})\underline{\epsilon_t}\epsilon_{t-i}^{*2}\epsilon_{t-j}^{*2}
\end{aligned}$$

$$E(\frac{\partial d_3}{\partial \beta}) = 0, E(\frac{\partial d_3}{\partial \alpha_0}) = 0, E(\frac{\partial d_3}{\partial \alpha}) = 0$$

Part II.

The test T_3 which we got in section 2 can be written as nR^2 form by running a regression $\hat{v}$ on $\hat{S}$.

$$\hat{d}_3 = \frac{1}{2T}\hat{S}'\hat{v},$$

where

$$\hat{v} = \begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \\ . \\ . \\ . \\ \hat{v}_T \end{pmatrix},$$

is a $T \times 1$ vector, $v_t = (\frac{1}{2}\epsilon_t^{*4} - 3\epsilon_t^{*2} + \frac{3}{2})$ and

$$\hat{S} = \begin{pmatrix} \hat{\epsilon}^{*2}_{1-1}\hat{\epsilon}^{*2}_{1-1} & \cdots & \hat{\epsilon}^{*2}_{1-1}\hat{\epsilon}^{*2}_{1-p} & \cdots & \hat{\epsilon}^{*2}_{1-p+1}\hat{\epsilon}^{*2}_{1-p} \\ \hat{\epsilon}^{*2}_{2-1}\hat{\epsilon}^{*2}_{2-1} & \cdots & \hat{\epsilon}^{*2}_{2-1}\hat{\epsilon}^{*2}_{2-p} & \cdots & \hat{\epsilon}^{*2}_{2-p+1}\hat{\epsilon}^{*2}_{2-p} \\ . & \cdots & . & \cdots & . \\ . & \cdots & . & \cdots & . \\ . & \cdots & . & \cdots & . \\ \hat{\epsilon}^{*2}_{T-1}\hat{\epsilon}^{*2}_{t-1} & \cdots & \hat{\epsilon}^{*2}_{T-1}\hat{\epsilon}^{*2}_{t-p} & \cdots & \hat{\epsilon}^{*2}_{T-p+1}\hat{\epsilon}^{*2}_{T-p} \end{pmatrix}$$

is a $T \times \frac{p(p+1)}{2}$ matrix. From the result above, we know that

$$\begin{aligned} V(\hat{d}_3) &= E(V(\hat{d}_3 \mid \Phi_{t-1})) \\ &= \frac{1}{4T^2}\hat{S}'E(\hat{v}\hat{v}')\hat{S} \\ &= \frac{6}{4T^2}\hat{S}'\hat{S}. \end{aligned}$$

$$\begin{aligned} T_3 &= \hat{d}_3'V(\hat{d}_3)^{-1}\hat{d}_3 \\ &= \frac{1}{4T^2}\hat{v}'\hat{S}(\hat{S}'\hat{S})^{-1}\hat{S}'\hat{v}\frac{4T^2}{6} \\ &= \frac{1}{6}\hat{v}'\hat{S}(\hat{S}'\hat{S})^{-1}\hat{S}'\hat{v}. \end{aligned}$$

$$\begin{aligned} TR^2 &= T\frac{\hat{v}'\hat{S}(\hat{S}'\hat{S})^{-1}\hat{S}'\hat{v}}{\hat{v}'\hat{v}} \\ &= \frac{1}{6}\hat{v}'\hat{S}(\hat{S}'\hat{S})^{-1}\hat{S}'\hat{v} \end{aligned}$$

$$plim_{T \to \infty} \frac{\hat{v}'\hat{v}}{T} = 6$$

$$T_3 = TR^2.$$

where R^2 is uncentered coefficient of determination of regression $\hat{v}$ on $\hat{S}$.

APPENDIX D

Part I.

To derive the double length regression (DLR) form of the IM test, we first note that

$$y_t - x_t\beta = \epsilon_t, \quad \epsilon_t \mid \Phi_{t-1} \sim N(0, h_t).$$

and

$$f_t(y_t, \theta) = \frac{y_t - x_t\beta}{\sqrt{h_t}} = \mu_t, \quad \mu_t \mid \Phi_{t-1} \sim N(0, 1).$$

Following Davidson and MacKinnon (1992), a locally equivalent model, under the alternative hypothesis that $Var(\alpha) = 2\Omega$, can be written as

$$\hat{q}_t(y_t, \theta, \Omega) \equiv \frac{\hat{f}_t(y_t, \theta) + tr(\Omega \hat{F}_t^*)}{(1 + 2tr(\Omega \hat{F}_t^T \hat{F}_t))^{\frac{1}{2}}} = \vartheta_t.$$

Using the notation as defined in section 3, we can write

$$\hat{F}_{ti} = \frac{\partial f_t}{\partial \alpha_i} \mid_{\theta=\theta} = -\frac{\hat{\epsilon}_t}{2\hat{h}_t^{\frac{3}{2}}} \hat{\epsilon}_{t-i}^2.$$

$$\hat{F}_{tij} = \frac{\partial^2 f_t}{\partial \alpha_i \partial \alpha_j} \mid_{\theta=\theta} = \frac{3\hat{\epsilon}_t}{4\hat{h}_t^{\frac{5}{2}}} \hat{\epsilon}_{t-i}^2 \hat{\epsilon}_{t-j}^2.$$

$$\hat{f}_t' = \frac{\partial \hat{f}_t}{\partial y_t} = \frac{1}{\sqrt{\hat{h}_t}}.$$

$$\hat{K}_{ti} = \frac{\partial log f_t'}{\partial \alpha_i} \mid_{\theta=\theta} = -\frac{1}{2\hat{h}_t^{\frac{3}{2}}} \hat{\epsilon}_{t-i}^2.$$

$$\hat{K}_{tij} = \frac{\partial^2 log f_t'}{\partial \alpha_i \partial \alpha_j} \mid_{\theta=\theta} = \frac{3}{4\hat{h}_t^{\frac{5}{2}}} \hat{\epsilon}_{t-i}^2 \hat{\epsilon}_{t-j}^2.$$

$$\begin{aligned}\hat{Q}_{tij} &= Q_{tij} \mid_{\Omega=0,\theta=\theta} \\ &= \frac{\partial q_t}{\partial \Omega_{ij}} \mid_{\Omega=0,\theta=\theta} \\ &= \hat{F}_{tij} - \hat{f}_t \hat{F}_{ti} \hat{F}_{tj} \\ &= \frac{3\hat{\epsilon}_t}{4\hat{h}_t^{\frac{5}{2}}} \hat{\epsilon}_{t-i}^2 \hat{\epsilon}_{t-j}^2 - \frac{\hat{\epsilon}_t^3}{4\hat{h}_t^{\frac{7}{2}}} \hat{\epsilon}_{t-i}^2 \hat{\epsilon}_{t-j}^2 \\ &= \frac{1}{4\hat{h}_t^2} \left(\frac{3\hat{\epsilon}_t}{\sqrt{\hat{h}_t}} - \frac{\hat{\epsilon}_t^3}{\hat{h}_t^{\frac{3}{2}}} \right) \hat{\epsilon}_{t-i}^2 \epsilon_{t-j}^2 .\end{aligned}$$

$$\begin{aligned}\hat{Q}_t^* &= \frac{1}{4\hat{h}_t^2} \left(\frac{3\hat{\epsilon}_t}{\sqrt{\hat{h}_t}} - \frac{\hat{\epsilon}_t^3}{\hat{h}_t^{\frac{3}{2}}} \right) (\hat{\epsilon}_{t-1}^4, \ldots, \hat{\epsilon}_{t-1}^2 \hat{\epsilon}_{t-2}^2, \ldots, \hat{\epsilon}_{t-p+1}^2 \hat{\epsilon}_{t-p}^2) \\ &= \frac{1}{4\hat{h}^2} \left(\frac{3\hat{\epsilon}_t}{\sqrt{\hat{h}_t}} - \frac{\hat{\epsilon}_t^3}{\hat{h}_t^{\frac{3}{2}}} \right) \tilde{\hat{\epsilon}}_t .\end{aligned}$$

where $\tilde{\hat{\epsilon}}_t = (\hat{\epsilon}_{t-1}^4, \ldots, \hat{\epsilon}_{t-1}^2 \hat{\epsilon}_{t-2}^2, \ldots, \hat{\epsilon}_{t-p+1}^2 \hat{\epsilon}_{t-p}^2)$ is a $\frac{p(p+1)}{2} \times 1$ vector.

$$\begin{aligned}\hat{R}_{tij} &= \hat{R}_{tij} \mid_{\Omega=0,\theta=\theta} \\ &= \frac{\partial log q_t'}{\partial \Omega_{ij}} \mid_{\Omega=0,\theta=\theta} \\ &= \hat{K}_{tij} + \hat{K}_{ti} \hat{K}_{tj} - \hat{F}_{ti} \hat{F}_{tj} - \hat{f}_t (\hat{K}_{ti} \hat{F}_{tj} + \hat{F}_{ti} \hat{K}_{tj}) \\ &= \frac{1}{4\hat{h}^2} \left(\frac{3}{\sqrt{\hat{h}_t}} + \frac{1}{\hat{h}_t} - (1 + \frac{2}{\sqrt{\hat{h}}}) \frac{\epsilon_t^2}{\hat{h}_t} \right) \hat{\epsilon}_{t-i}^2 \hat{\epsilon}_{t-j}^2 .\end{aligned}$$

$$\hat{R}_t^* = \frac{1}{4\hat{h}_t^2} \left(\frac{3}{\sqrt{\hat{h}_t}} + \frac{1}{\hat{h}_t} - (1 + \frac{2}{\sqrt{\hat{h}_t}}) \frac{\hat{\epsilon}_t^2}{\hat{h}_t} \right) \tilde{\hat{\epsilon}}_t .$$

Given the above results, the DLR form of the test can be expressed in terms of the following regression

$$\begin{aligned}\begin{pmatrix} \frac{\epsilon_t}{\sqrt{h_t}} \\ 1 \end{pmatrix} &= \begin{pmatrix} -\hat{Q}_t^* \\ \hat{R}_t^* \end{pmatrix} \Omega_{ij} + \omega_t \\ &= \begin{pmatrix} \frac{1}{4h_t^2} \left(\frac{3\epsilon_t}{\sqrt{h_t}} - \frac{\epsilon_t^3}{h_t^{\frac{3}{2}}} \right) \tilde{\hat{\epsilon}}_t \\ \frac{1}{4h_t^2} \left(\frac{3}{\sqrt{h_t}} + \frac{1}{h_t} - (1 + \frac{2}{\sqrt{h_t}}) \frac{\epsilon_t^2}{h_t} \right) \tilde{\hat{\epsilon}}_t \end{pmatrix} \Omega_{ij} + \omega_t\end{aligned}$$

Part II.

To calculate the non-centrality parameter for local alternative, we consider the simple case with $p = 1$. The model under H_0: $\sigma^2 = 0$, is

$$f_t(y_t, \theta) = \frac{y_t - x_t\beta}{\sqrt{h_t}} = \frac{\epsilon_t}{\sqrt{h_t}} = \mu_t$$

and $\mu_t \sim N(0,1)$. The loglikelihood of this model is

$$l_t = \frac{1}{2}log(2\pi) - \frac{1}{2}f_t^2 + k_t.$$

and therefore,

$$F_{t1} = \frac{\partial f_t}{\partial \alpha_1} = -\frac{\epsilon_t}{2h^{\frac{3}{2}}}\epsilon_{t-1}^2.$$

$$F_{t11} = \frac{\partial^2 f_t}{\partial \alpha_1^2} = \frac{3\epsilon_t}{4h_t^{\frac{5}{2}}}\epsilon_{t-1}^4.$$

$$f_t' = \frac{\partial f_t}{\partial y_t} = \frac{1}{\sqrt{h_t}}.$$

$$F_{t1}' = \frac{\partial F_{t1}}{\partial y_t} = -\frac{1}{2h_t^{\frac{3}{2}}}\epsilon_{t-1}^2.$$

$$F_{t11}' = \frac{\partial F_{t11}}{\partial y_t} = \frac{3}{4h_t^{\frac{5}{2}}}\epsilon_{t-1}^4.$$

$$\frac{\partial f_t}{\partial \beta} = -\frac{\epsilon_t}{2h_t^{\frac{3}{2}}}\frac{\partial h_t}{\partial \beta} - \frac{x_t}{\sqrt{h_t}}.$$

$$\frac{\partial f_t'}{\partial \beta} = -\frac{1}{2h_t^{\frac{3}{2}}}\frac{\partial h_t}{\partial \beta}.$$

$$\frac{\partial F_{t1}}{\partial \beta} = \frac{3\epsilon_t}{4h_t^{\frac{5}{2}}}\epsilon_{t-1}^2\frac{\partial h_t}{\partial \beta} + \frac{x_t}{2h_t^{\frac{3}{2}}}\epsilon_{t-1}^2 + \frac{\epsilon_t}{h_t^{\frac{3}{2}}}\epsilon_{t-1}x_{t-1}.$$

$$\frac{\partial F_{t1}'}{\partial \beta} = \frac{3}{4h_t^{\frac{5}{2}}}\epsilon_{t-1}^2\frac{\partial h_t}{\partial \beta} + \frac{1}{h_t^{\frac{3}{2}}}\epsilon_{t-1}x_{t-1}.$$

$$\frac{\partial F_{t11}}{\partial \beta} = -\frac{15\epsilon_t}{8h_t^{\frac{7}{2}}}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta} - \frac{3x_t}{4h_t^{\frac{5}{2}}}\epsilon_{t-1}^4 - \frac{3\epsilon_t}{h_t^{\frac{5}{2}}}\epsilon_{t-1}^3x_{t-1}.$$

$$\frac{\partial F'_{t11}}{\partial \beta} = -\frac{15}{8h_t^{\frac{7}{2}}}\epsilon_{t-1}^4 \frac{\partial h_t}{\partial \beta} - \frac{3}{h_t^{\frac{5}{2}}}\epsilon_{t-1}^3 x_{t-1}.$$

$$\frac{\partial f_t}{\partial \underline{\alpha}} = -\frac{\epsilon_t}{2h_t^{\frac{3}{2}}}\frac{\partial h_t}{\partial \underline{\alpha}}.$$

$$\frac{\partial f'_t}{\partial \underline{\alpha}} = -\frac{1}{2h_t^{\frac{3}{2}}}\frac{\partial h_t}{\partial \underline{\alpha}}.$$

$$\frac{\partial F_{t1}}{\partial \underline{\alpha}} = \frac{3\epsilon_t}{4h_t^{\frac{5}{2}}}\epsilon_{t-1}^2 \frac{\partial h_t}{\partial \underline{\alpha}}.$$

$$\frac{\partial F'_{t1}}{\partial \underline{\alpha}} = \frac{3}{4h_t^{\frac{5}{2}}}\epsilon_{t-1}^2 \frac{\partial h_t}{\partial \underline{\alpha}}.$$

$$\frac{\partial F_{t11}}{\partial \underline{\alpha}} = -\frac{15\epsilon_t}{8h_t^{\frac{7}{2}}}\epsilon_{t-1}^4 \frac{\partial h_t}{\partial \underline{\alpha}}.$$

$$\frac{\partial F'_{t11}}{\partial \underline{\alpha}} = -\frac{15}{8h_t^{\frac{7}{2}}}\epsilon_{t-1}^4 \frac{\partial h_t}{\partial \underline{\alpha}}.$$

Here $\underline{\alpha} = (\alpha_0, \alpha_1)'$. The model under H_a: $\sigma^2 = \frac{\delta}{\sqrt{T}}$, is

$$q_t(y_t, \theta, \sigma^2) = \frac{f_t(y_t, \theta) + \sigma^2 F_{t11}}{(1 + 2\sigma^2 F_{t1} F_{t1})^{\frac{1}{2}}} = \mu_t,$$

and $\mu_t \sim N(0, 1)$. The loglikelihood function of this model is

$$l_t = -\frac{1}{2}log(2\pi) - \frac{1}{2}q_t^2 + r_t.$$

We denote the information matrix of the model as

$$I = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

where

$$V_{11} = -E\begin{pmatrix} \frac{\partial^2 l_t}{\partial\beta\partial\beta'} & \frac{\partial^2 l_t}{\partial\beta\partial\underline{\alpha}} \\ \frac{\partial^2 l_t}{\partial\underline{\alpha}\partial\beta} & \frac{\partial^2 l_t}{\partial\underline{\alpha}\partial\underline{\alpha}'} \end{pmatrix},$$

is a $(k+2) \times (k+2)$ matrix,

$$V_{12} = V'_{21} = -E\begin{pmatrix} \frac{\partial^2 l_t}{\partial\beta\partial\sigma^{2'}} \\ \frac{\partial^2 l_t}{\partial\underline{\alpha}\partial\sigma^{2'}} \end{pmatrix},$$

is a $(k+2)\times 1$ vector, and

$$V_{22} = \frac{\partial^2 l_t}{\partial(\sigma^2)^2},$$

is a scalar. Now

$$\frac{\partial q_t}{\partial\beta} = \frac{(1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}(\frac{\partial f_t}{\partial\beta}+\sigma^2\frac{\partial F_{t11}}{\partial\beta}) - (f_t+\sigma^2\frac{\partial F_{t11}}{\partial\beta})(1+2\sigma^2 F_{t1}F_{t11})^{-\frac{1}{2}}2\sigma^2 F_{t1}\frac{\partial F_{t1}}{\partial\beta}}{(1+2\sigma^2 F_{t1}F_{t1})}.$$

$$\frac{\partial q_t}{\partial\sigma^2} = \frac{(1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}F_{t11} - (f_t+\sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}F_{t1}F_{t1}}{(1+2\sigma^2 F_{t1}F_{t1})}.$$

$$\frac{\partial q_t}{\partial\beta}\,|_{\sigma^2=0} = \frac{\partial f_t}{\partial\beta},$$

$$\frac{\partial q_t}{\partial\sigma^2}\,|_{\sigma^2=0} = F_{t11} - f_t F_{t1}F_{t1}.$$

$$\begin{aligned}-\frac{\partial q_t}{\partial\beta}\frac{\partial q_t}{\partial\sigma^2}\,|_{\sigma^2=0} &= -\frac{\partial f_t}{\partial\beta}F_{t11} + \frac{\partial f_t}{\partial\beta}f_t F_{t1}F_{t1} \\ &= \frac{3\epsilon_t^2}{8h_t^4}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta} + \frac{3x_t\epsilon_t}{4h_t^3}\epsilon_{t-1}^4 \\ &\quad - \frac{\epsilon_t^4}{8h_t^5}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta} - \frac{\epsilon_t^3 x_t}{h_t^4}\epsilon_{t-1}^4,\end{aligned}$$

$$E(-\frac{\partial q_t}{\partial\beta}\frac{\partial q_t}{\partial\sigma^2}\,|_{\sigma^2=0}|\,\Phi_{t-1}) = \frac{3}{8h_t^5}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta} - \frac{3}{8h_t^5}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta} = 0$$

$$\frac{\partial^2 q_t}{\partial\beta\partial\sigma^2}\,|_{\sigma^2=0} = \frac{\partial F_{t11}}{\partial\beta} - 2f_t F_{t1}\frac{\partial F_{t1}}{\partial\beta} - F_{t1}F_{t1}\frac{\partial f_t}{\partial\beta}.$$

$$\begin{aligned}-q_t\frac{\partial^2 q_t}{\partial\beta\partial\sigma^{2\prime}}\,|_{\sigma^2=0} &= -f_t\frac{\partial F_{t11}}{\partial\beta} + 2f_t^2 F_{t1}\frac{\partial F_{t1}}{\partial\beta} + f_t F_{t1}F_{t1}\frac{\partial f_t}{\partial\beta} \\ &= \frac{15\epsilon_t^2}{8h_t^4}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta} + \frac{3\epsilon_t x_t}{4h_t^3}\epsilon_{t-1}^4 + \frac{3\epsilon_t^2}{h_t^3}\epsilon_{t-1}^3 x_{t-1} \\ &\quad - \frac{3\epsilon_t^4}{4h_t^5}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta} - \frac{\epsilon_t^3 x_t}{2h_t^4}\epsilon_{t-1}^4 - \frac{\epsilon_t^4}{h_t^4}\epsilon_{t-1}^3 x_{t-1} \\ &\quad - \frac{\epsilon_t^4}{8h_t^5}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta} - \frac{\epsilon_t^3 x_t}{4h_t^4}\epsilon_{t-1}^4,\end{aligned}$$

$$E(-q_t\frac{\partial^2 q_t}{\partial\beta\partial\sigma^{2\prime}}\,|_{\sigma^2=0}|\,\Phi_{t-1}) = -\frac{3}{4h_t^3}\epsilon_{t-1}^4\frac{\partial h_t}{\partial\beta}.$$

$$\frac{\partial r_t}{\partial \beta} = [(1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}(\frac{\partial f_t'}{\partial \beta} + \sigma^2 \frac{\partial F_{t11}'}{\partial \beta})$$

$$+ (f_t' + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}} 2\sigma^2 F_{t1}\frac{\partial F_{t1}}{\partial \beta}$$

$$- (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}} 2\sigma^2 \frac{\partial F_{t1}'}{\partial \beta} F_{t1}$$

$$- (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}} 2\sigma^2 F_{t1}'\frac{\partial F_{t1}}{\partial \beta}$$

$$- (\frac{\partial f_t}{\partial \beta} + \sigma^2 \frac{\partial F_{t11}}{\partial \beta})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}} 2\sigma^2 F_{t1}' F_{t1}$$

$$+ (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{3}{2}} 2\sigma^2 F_{t1}' F_{t1} 2\sigma^2 F_{t1}\frac{\partial F_{t1}}{\partial \beta}]$$

$$- \frac{4\sigma^2 F_{t1}\frac{\partial F_{t1}}{\partial \beta}}{(1+2\sigma^2 F_{t1}F_{t1})}.$$

$$\frac{\partial^2 r_t}{\partial \beta \partial \sigma^2}\,|_{\sigma^2=0} = \frac{1}{f_t'}\frac{\partial F_{t11}'}{\partial \beta} - 2\frac{f_t{}'}{f_t} F_{t1}\frac{\partial F_{t1}'}{\partial \beta} - 2\frac{f_t{}'}{f_t} F_{t1}'\frac{\partial F_{t1}}{\partial \beta}$$

$$- 2\frac{1}{f_t{}'} f_{t1}' F_{t1}\frac{\partial f_t}{\partial \beta} - \frac{1}{f_t^{2\prime}} F_{t11}'\frac{\partial f_t}{\partial \beta}$$

$$+ 2\frac{f_t}{f_t^{2\prime}} F_{t1}' F_{t1}\frac{\partial f_t}{\partial \beta} - 2F_{t1}\frac{\partial F_{t1}}{\partial \beta}$$

$$= -\frac{15}{8h_t^3}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta} - \frac{3}{h_t^2}\epsilon_{t-1}^3 x_{t-1} + \frac{3\epsilon_t^2}{4h_t^4}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta}$$

$$+ \frac{\epsilon_t^2}{h_t^3}\epsilon_{t-1}^3 x_{t-1} + \frac{3\epsilon_t^2}{4h_t^4}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta} + \frac{\epsilon_t x_t}{2h_t^3}\epsilon_{t-1}^4$$

$$+ \frac{\epsilon_t^2}{h_t^3}\epsilon_{t-1}^3 x_{t-1} + \frac{\epsilon_t^2}{4h_t^4}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta} + \frac{\epsilon_t x_t}{4h_t^3}\epsilon_{t-1}^4$$

$$+ \frac{3}{8h_t^3}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta} - \frac{\epsilon_t^2}{4h_t^4}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta} + \frac{3\epsilon_t^2}{4h_t^4}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta}$$

$$+ \frac{\epsilon_t x_t}{2h_t^3}\epsilon_{t-1}^4 + \frac{\epsilon^2}{h_t^3}\epsilon_{t-1}^3 x_{t-1},$$

$$E(\frac{\partial^2 r_t}{\partial \beta \partial \sigma^2}\,|_{\sigma^2=0}|\,\Phi_{t-1}) = \frac{3}{4h_t^3}\epsilon_{t-1}^4\frac{\partial h_t}{\partial \beta}.$$

Then

$$E[(-q_t\frac{\partial^2 q_t}{\partial \beta \partial \sigma^2} + \frac{\partial^2 r_t}{\partial \beta \partial \sigma^2})\,|_{\sigma^2=0}|\,\Phi_{t-1}] = 0.$$

Since

$$\frac{\partial^2 l_t}{\partial\beta\partial\sigma^2} = -q_t\frac{\partial^2 q_t}{\partial\beta\partial\sigma^2} - \frac{\partial q_t}{\partial\beta}\frac{\partial q_t}{\partial\sigma^2} + \frac{\partial^2 r_t}{\partial\beta\partial\sigma^2},$$

$$E(\frac{\partial^2 l_t}{\partial\beta\partial\sigma^2}\mid_{\sigma^2=0}\mid \Phi_{t-1}) = 0,$$

Using the same procedure, we can obtain the $E(\frac{\partial^2 l_t}{\partial\underline{\alpha}\partial\sigma^2}\mid_{\sigma^2=0}\mid \Phi_{t-1}) = 0$. Then, we have $V_{12} = V_{21}' = 0$. Because the inverse variance of the non-central parameter is equal to $V^{22} = (V_{22} - V_{21}V_{11}^{-1}V_{12})^{-1}$, depending on the above results, $V^{22} = (V_{22})^{-1}$. Now we calculate V_{22}, using the following derivatives

$$\frac{\partial q_t}{\partial\sigma^2} = \frac{(1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}F_{t11} - (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}F_{t1}F_{t1}}{(1+2\sigma^2 F_{t1}F_{t1})}.$$

$$\frac{\partial^2 q_t}{\partial(\sigma^2)^2}\mid_{\sigma^2=0} = -2F_{t11}F_{t1}F_{t1} - 3f_t(F_{t1}F_{t1})^2.$$

$$\begin{aligned}\frac{\partial q_t}{\partial\sigma^2}\frac{\partial q_t}{\partial\sigma^2}\mid_{\sigma^2=0} &= (F_{t11} - f_tF_{t1}F_{t1})^2\\ &= F_{t11}^2 - 2f_tF_{t11}F_{t1}F_{t1} + f_t^2(F_{t1}F_{t1})^2.\end{aligned}$$

$$\begin{aligned}q_t' &= \frac{\partial q_t}{\partial y_t}\\ &= \frac{(1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}(f_t' + \sigma^2 F_{t11}') - (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}2\sigma^2 F_{t1}'F_{t1}}{(1+2\sigma^2 F_{t1}F_{t1})}.\end{aligned}$$

$$\begin{aligned}r_t &= log q_t'\\ &= log[(1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}(f_t' + \sigma^2 F_{t11}')\\ &- (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}2\sigma^2 F_{t1}'F_{t1}]\\ &- log(1+2\sigma^2 F_{t1}F_{t1})\end{aligned}$$

$$\frac{\partial r_t}{\partial \sigma^2} = [(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}(f_t' + \sigma^2 F_{t11}')F_{t1}F_{t1}$$

$$+ (1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}F_{t11}' - F_{t11}(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}2\sigma^2 F_{t1}' F_{t1}$$

$$+ (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{3}{2}}F_{t1}F_{t1}2\sigma^2 F_{t1}'F_{t1}$$

$$- (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}2F_{t1}'F_{t1}]$$

$$/[(1+2\sigma^2 F_{t1}F_{t1})^{\frac{1}{2}}(f_t' + \sigma^2 F_{t11}')$$

$$- (f_t + \sigma^2 F_{t11})(1+2\sigma^2 F_{t1}F_{t1})^{-\frac{1}{2}}2\sigma^2 F_{t1}'F_{t1}]$$

$$- 2F_{t1}F_{t1}/(1+2\sigma^2 F_{t1}F_{t1}).$$

$$\frac{\partial^2 r_t}{\partial(\sigma^2)^2}\,|_{\sigma^2=0} = \frac{1}{(f_t')^2}[f_t'(-f_t'(F_{t1}F_{t1})^2 + F_{t11}'F_{t1}F_{t1}$$

$$+ F_{t1}F_{t1}F_{t11}' - 2F_{t11}F_{t1}'F_{t1} + 2f_t F_{t1}F_{t1}F_{t1}'F_{t1}$$

$$- 2F_{t11}F_{t1}'F_{t1} + 2f_t F_{t1}F_{t1}F_{t1}'F_{t1})$$

$$- (f_t'F_{t1}F_{t1} + F_{t11}' - 2f_t F_{t1}'F_{t1})^2] + 4(F_{t1}F_{t1})^2$$

$$= \frac{1}{f_t^{2\prime}}[-f_t'^2(F_{t1}F_{t1})^2 + 8f_t'f_t(F_{t1})^3 F_{t1}'$$

$$- 4f_t'F_{t11}F_{t1}'F_{t1} - F_{t11}'F_{t11}' - 4f_t^2(F_{t1}'Ft1)^2$$

$$+ 4f_t F_{t11}'F_{t1}'F_{t1}] + 4(F_{t1}F_{t1})^2.$$

$$-\frac{\partial^2 q_t}{\partial(\sigma^2)^2}\,|_{\sigma^2=0} = -2f_t F_{t11}F_{t1}F_{t1} + 3f_t^2(F_{t1}F_{t1})^2.$$

$$\frac{\partial^2 l_t}{\partial(\sigma^2)^2} = -q_t\frac{\partial^2 q_t}{\partial(\sigma^2)^2} - (\frac{\partial q_t}{\partial\sigma^2})^2 + \frac{\partial^2 r_t}{\partial(\sigma^2)^2}.$$

and

$$\frac{\partial^2 l_t}{\partial(\sigma^2)^2}\mid_{\sigma^2=0} = 4f_t F_{t11}F_{t1}F_{t1} - 4f_t^2(F_{t1}F_{t1})^2 - F_{t11}^2$$
$$+ 8\frac{f_t}{f_t'}(F_{t1})^3 F_{t1}' - 4\frac{1}{f_t'}F_{t11}F_{t1}'F_{t1} - \frac{1}{f_t^{2'}}F_{t11}'F_{t11}$$
$$- 4\frac{f_t^2}{f_t^{2'}}(F_{t1}'F_{t1})^2 + 4\frac{f_t}{f_t^{2'}}F_{t11}'F_{t1}'F_{t1} + 2(F_{t1}F_{t1})^2$$
$$= \frac{12\epsilon_t^4}{16h_t^6}\epsilon_{t-1}^8 - \frac{4\epsilon_t^6}{16h_t^7}\epsilon_{t-1}^8 - \frac{9\epsilon_t^2}{16h_t^5}\epsilon_{t-1}^8$$
$$+ \frac{8\epsilon_t^4}{16h_t^6}\epsilon_{t-1}^8 - \frac{12\epsilon_t^2}{16h_t^5}\epsilon_{t-1}^8 - \frac{9\epsilon_t^2}{16h_t^5}\epsilon_{t-1}^8$$
$$- \frac{4\epsilon_t^4}{16h_t^6}\epsilon_{t-1}^8 + \frac{12\epsilon_t^2}{16h_t^5}\epsilon_{t-1}^8 + \frac{2\epsilon_t^4}{16h_t^6}\epsilon_{t-1}^8$$

Finally,

$$E\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\partial^2 l_t}{\partial(\sigma^2)^2}\right)\mid_{\sigma^2=0,\theta=\theta} = -\frac{1}{T}\sum_{t=1}^{T}\frac{3}{2\hat{h}^4}\hat{\epsilon}_{t-1}^8 = -\frac{1}{T}\sum_{t=1}^{T}\frac{3}{2}\hat{\epsilon}_{t-1}^{*8}.$$

Part III.

To compute the moments of ϵ_t, let us define

$$\lambda_t = (\epsilon_t^{2m}, \epsilon_t^{2(m-1)}, \ldots, \epsilon_t^2).$$

Following Engle(1982), we have

$$E(\epsilon_T^{2m} \mid \Phi_{t-1}) = h^m \prod_{j=1}^{m}(2j-1) = (\alpha_0 + \zeta\epsilon_{t-1}^2)^m \prod_{j-1}^{m}(2j-1).$$

Expanding this expression establishes that the moment is a linear combination of λ_{t-1}.

$$E(\lambda_t \mid \Phi_{t-2}) = b + A(b + A\lambda_{t-2}).$$

Only powers of ϵ less than or equal to 2m are required, therefore A is a upper triangular matrix and b is a $m \times 1$ vector. In general,

$$E(\lambda_t \mid \Phi_{t-k}) = (I + A + A^2 + \ldots + A^{k-1})b + A^k\lambda_{t-k}.$$

Because the series starts indefinitely far in the past with 2r finite moments, the limit as k goes to infinity exists if, and only if, all the eigenvalues of A lie within the unit circle. Now,

$$\lim_{k \to \infty} E(\lambda_t \mid \Phi_{t-k}) = E(\lambda_t),$$

is an expression for the stationary moments of the unconditional distribution of ϵ. We have

$$E(\lambda_t) = EE(\lambda_t \mid \Phi_{t-1}) = (I - E(A))^{-1} b.$$

Since

$$\begin{aligned}(I - E(A))^{-1} &= \begin{pmatrix} 1 - 3E(\zeta)^2 & -6\alpha_0 E(\zeta^2) \\ 0 & 1 - E(\zeta^2) \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \frac{1-\alpha_1}{(1-\alpha_1)(1-3(\alpha_1^2+2\sigma^2))} & \frac{6\alpha_0\alpha_1}{(1-\alpha_1)(1-3(\alpha_1^2+2\sigma^2))} \\ 0 & \frac{1-\alpha_1}{(1-\alpha_1)(1-3(\alpha_1^2+2\sigma^2))} \end{pmatrix},\end{aligned}$$

and

$$b = (3\alpha_0^2, \alpha_0)',$$

the expression for the fourth and seconds moments are as given in section 3.