Journal of the Korean Statistical Society (2002), **31**: *4, pp 459–468*

Asymmetric Modeling in Beta-ARCH Processes

S. Y. Hwang[1] and Myung-Wook Kahng[1]

ABSTRACT

A class of asymmetric beta-ARCH processes is proposed and connections to traditional ARCH models are explained. Geometric ergodicity of the model is discussed. Conditional least squares as well as maximum likelihood estimators of parameters and their limit results are also presented. A test for symmetry of the model is studied with limiting power of test statistic given.

Keywords. Asymmetric beta-ARCH, geometric ergodicity, test for symmetry.
AMS 2000 subject classifications. Primary 62M10.

1. Introduction

Since the seminal paper of Engle (1982), a lot of research has been directed to conditional heteroscedastic autoregressive (ARCH) models in which the conditional variance (or volatility) was specified as a linear combination of the squared residuals. Despite Engle's ARCH's usefulness in many applications, there has recently been growing interest in modeling the conditional variance via nonlinear ARCH using other functional forms rather than squared residuals, especially in the field of econometric applications. Guegan and Diebolt (1994) proposed the first order β-ARCH process and this was extended to m-order cases by An *et al.* (1997). The m-order β-ARCH process $\{X_t\}$ is defined as

$$\begin{aligned} X_t &= \sqrt{v_{t-1}} \cdot e_t, \\ v_{t-1} &= \alpha_0 + \alpha_1 |X_{t-1}|^\beta + \cdots + \alpha_m |X_{t-m}|^\beta \end{aligned} \tag{1.1}$$

where $0 < \beta \leq 2$, and $\{e_t\}$ is a sequence of *iid* random variables with mean zero and variance unity. It is obvious that $\beta = 2$ provides Engle's linear ARCH. Guegan and Diebolt (1994) and An *et al.* (1997) investigated probabilistic structures

Received May 2002; accepted August 2002.
[1]Department of Statistics, Sookmyung Women's University, Seoul 140-742, Korea

such as geometric ergodicity and existence of moments for β-ARCH processes. Hwang and Basawa (2001) studied nonlinear autoregressive processes generated by β-ARCH errors in (1.1) and derived asymptotic results concerning parameter estimation.

Li and Li (1996) introduced a class of double threshold ARCH where the "threshold" concept was emerged into conditional variance as well as conditional mean. They proposed detailed fitting strategies of the model including order identification, estimation and diagnostic checking with applications to Hong Kong Hang Seng Index. Recently Hwang and Woo (2001) investigated a first order threshold ARCH model applied to diverse Korean financial times series. The m-order threshold model is defined by the equation

$$\begin{aligned} X_t &= \sqrt{v_{t-1}} \cdot e_t, \\ v_{t-1} &= \alpha_0 + \alpha_{11}(X_{t-1}^+)^2 + \alpha_{21}(X_{t-1}^-)^2 + \cdots \\ &\quad + \alpha_{1m}(X_{t-m}^+)^2 + \alpha_{2m}(X_{t-m}^-)^2 \end{aligned} \tag{1.2}$$

where and throughout the paper the notation $a_t^+ = \max(a_t, 0)$, $a_t^- = \max(-a_t, 0)$ will be used. Li and Li (1996) argued that (1.2) was capable of capturing the asymmetric phenomena in the conditional variance when $a_{1i} \neq a_{2i}$. It seems though that (1.2) is intrinsically symmetric since only the slopes (a_{1i} and a_{2i}) are permitted differently while retaining the same square functional forms around zero.

In this article we propose a class of models possessing asymmetric conditional variances by combining (1.1) and (1.2). The developed model is given by

$$\begin{aligned} X_t &= \sqrt{v_{t-1}} \cdot e_t, \\ v_{t-1} &= \alpha_0 + \alpha_{11}(X_{t-1}^+)^{\beta_{11}} + \alpha_{21}(X_{t-1}^-)^{\beta_{21}} + \cdots \\ &\quad + \alpha_{1m}(X_{t-m}^+)^{\beta_{1m}} + \alpha_{2m}(X_{t-m}^-)^{\beta_{2m}}. \end{aligned} \tag{1.3}$$

Notice that different β_{1i} and β_{2i} entail asymmetry in the conditional variance, and the class of models in (1.3) is rich enough to include the following models as special cases.

Example 1 (Engle's ARCH). Setting $\beta_{1i} = \beta_{2i} = 2$ and $\alpha_{1i} = \alpha_{2i}$ ($i = 1, \ldots, m$) in (1.3) reduces to Engle's ARCH;

Example 2 (β-ARCH). (1.1) can be obtained by choosing $\beta_{1i} = \beta_{2i} = \beta$ and $\alpha_{1i} = \alpha_{2i} = \alpha$ ($i = 1, \ldots, m$);

Example 3 (Threshold ARCH). Substituting $\beta_{1i} = \beta_{2i} = 2$ in (1.3) yields Eq. (1.2).

The main objectives of the paper are not to present comprehensive statistical account of the model in (1.3) nor to give details on technicalities. Rather, our intention is to motivate the model to be an applicable class of asymmetric ARCH towards invoking future research topics. The rest of the paper is organized as follows. Section 2 briefly addresses the stationarity of the models and parameter estimation is discussed in Section 3 where the least squares and the maximum likelihood methods are investigated. Section 4 is concerned with testing symmetry of the model. It is to be noted that (1.3) exhibits a symmetric pattern, provided

$$H : \beta_{1i} = \beta_{2i} = 2 \ \text{ and } \ \alpha_{1i} = \alpha_{2i} \ (i = 1, \dots, m).$$

The Wald test statistic is proposed for testing H and the limiting power function is presented.

2. The Model and Stationarity

Consider the observation process $\{X_t\}$ generated by

$$\begin{aligned} X_t &= \sqrt{v_{t-1}} \cdot e_t, \\ v_{t-1} &= \alpha_0 + \alpha_{11}(X_{t-1}^{+})^{\beta_{11}} + \alpha_{21}(X_{t-1}^{-})^{\beta_{21}} + \cdots \\ &\quad + \alpha_{1m}(X_{t-m}^{+})^{\beta_{1m}} + \alpha_{2m}(X_{t-m}^{-})^{\beta_{2m}} \end{aligned} \tag{2.1}$$

where and in the sequel $\{e_t\}$ is possibly non-Gaussian *iid* sequence of random variables with zero mean and unit variance and the $(4m+1) \times 1$ vector of parameters is denoted by $\boldsymbol{\theta} = (\alpha_0, \alpha_{11}, \alpha_{21}, \dots, \alpha_{1m}, \alpha_{2m}, \beta_{11}, \beta_{21}, \dots, \beta_{1m}, \beta_{2m})^T$ with $\alpha_0 > 0$, $\alpha_{1i} \geq 0$, $\alpha_{2i} \geq 0$ and $0 < \beta_{1i} \leq 2$, $0 < \beta_{2i} \leq 2$.

(A.1) The common probability distribution of $\{e_t\}$ is absolutely continuous with respect to Lebesgue measure and is equipped with support on the whole real line $(-\infty, \infty)$.

It is obvious that (2.1) is a Markov process of order m, which facilitates the derivation of conditions for the geometric ergodicity of the model. Introduce $m \times 1$ vectors Y_t, $H(\cdot)$ and $V(\cdot)$ such that $Y_{t-1} = (X_t, X_{t-1}, \dots, X_{t-m+1})^T$, $H(Y_{t-1}) = (0, X_{t-1}, \dots, X_{t-m+1})^T$ and $V(Y_{t-1}) = (\sqrt{v_{t-1}}, 0, \dots, 0)^T$ which in turn yield the following first order $m \times 1$ vector Markov process:

$$Y_t = H(Y_{t-1}) + V(Y_{t-1}) \cdot e_t\,.$$

One may establish the geometric ergodicity of $\{Y_t\}$ (and hence for $\{X_t\}$) using various sets of conditions for Markovian time series. Refer to, for instance, Feigin and Tweedie (1985), and An *et al.* (1997). Let $\alpha(i) = \max(\alpha_{1i}, \alpha_{2i})$, $J \subset \{1, \ldots, m\}$ denote the collection of all subscripts i such that either $\beta_{1i} = 2$ or $\beta_{2i} = 2$. The lemma below presents a set of sufficient conditions for the geometric ergodicity of the model.

Lemma 1. *Under (A.1), $\{X_t\}$ in (2.1) is geometrically ergodic if* $\sum_{i \in J} \alpha(i) < 1$.

Remarks. For the special case of $m = 1$, the condition $\sum_{i \in J} \alpha(i) < 1$ is equivalent to

(i) $\max(\alpha_{1i}, \alpha_{2i}) < 1$ when either $\beta_1 = 2$ or $\beta_2 = 2$;

(ii) trivially satisfied when $0 < \beta_1 < 2$ and $0 < \beta_2 < 2$.

Condition (i) coincides with the result in Hwang and Woo (2001) and (ii) is analogous to those in An *et al.* (1997).

Proof. Arguing with Theorem 2 in Feigin and Tweedie (1985), lemma is essentially obtained by verifying that there exists a non-negative continuous real valued function $g : R^m \to R$ such that for all $||y_{t-1}||$ large,

$$\left| \frac{E(g(Y_t) \mid Y_{t-1} = y_{t-1})}{g(y_{t-1})} \right| < 1. \tag{2.2}$$

A straightforward manipulation of the proof of Theorem 3.2 of An *et al.* (1997) shows that under $\sum_{i \in J} \alpha(i) < 1$, (2.2) holds by taking $g(\cdot)$ as maximum norm, *viz.*, $g(y_{t-1}) = ||y_{t-1}||_\infty$. One can also deduce from (A.1) that $\{Y_t\}$ is ϕ-irreducible when ϕ is the Lebesgue measure on R^m and is a Feller chain (*cf.* An *et al.*, 1997), which completes the proof. □

Therefore, the following will be assumed throughout.

(C.1) The model (2.1) is geometrically ergodic.

3. Estimation of Parameters

We begin with the conditional least squares(CLS) estimation problems. Based on the data, $\{ X_{-m+1}, \ldots, X_1, \ldots, X_n\}$ the conditional least squares estimator $\widehat{\boldsymbol{\theta}}_n$ of parameters $\boldsymbol{\theta} = (\alpha_0, \alpha_{11}, \alpha_{21}, \ldots, \alpha_{1m}, \alpha_{2m}, \beta_{11}, \beta_{21}, \ldots, \beta_{1m}, \beta_{2m})^T$ is obtained by minimizing

$$\Psi_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} [\, X_t^2 - v_{t-1}(\boldsymbol{\theta})\,]^2. \tag{3.1}$$

In order to discuss the asymptotic properties for $\widehat{\boldsymbol{\theta}}_n$, we need to impose the following.

(C.2) The finiteness of eighth order moment of the stationary distribution: $EX_t^8 < \infty$.

Theorem 1. *Under (C.1) and (C.2) there exists, with probability tending to one, a sequence of estimators $\widehat{\boldsymbol{\theta}}_n$ minimizing $\Psi_n(\boldsymbol{\theta})$ and $\widehat{\boldsymbol{\theta}}_n$ is strongly consistent to $\boldsymbol{\theta}$. Furthermore,*

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0},\ \mathbf{A}^{-1}\mathbf{B}\,\mathbf{A}^{-1})$$

where $\mathbf{A}$ *and* $\mathbf{B}$ *are specified below.*

Proof. Conditions (C.1) and (C.2) imply the relevant three conditions in Theorem 3.1 of Tjøsthiem (1986) and hence, with probability converging to one, there exists a sequence of estimators $\widehat{\boldsymbol{\theta}}_n$ which is strongly consistent. Note that $\widehat{\boldsymbol{\theta}}_n$ solves $\partial\Psi_n(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$. Following the lines as in Klimko and Nelson (1978), and Tjøstheim (1986), it can be verified [under (C.1) and (C.2)] that

$$\sqrt{n}\,(\boldsymbol{\theta}_n - \boldsymbol{\theta}) = \left[n^{-1}\,\frac{\partial^2\Psi_n(\boldsymbol{\theta})}{\partial^2\boldsymbol{\theta}}\right]^{-1} \cdot \left[n^{-1/2}\,\frac{\partial\,\Psi_n(\boldsymbol{\theta})}{\partial\,\boldsymbol{\theta}}\right] + o_p(1). \tag{3.2}$$

The ergodic theorem tells us that $n^{-1}\,\partial^2\Psi_n(\boldsymbol{\theta})/\partial^2\boldsymbol{\theta}$ converges in probability to $\mathbf{A}$ where $\mathbf{A} = \text{plim}(n^{-1}\cdot\partial^2\Psi_n(\boldsymbol{\theta})/\partial^2\boldsymbol{\theta})$. It may also be noticed that $\partial\Psi_n(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ is a sum of zero mean martingale differences and thus it follows from the martingale central limit theorem that

$$n^{-1/2}\,\frac{\partial\,\Psi_n(\boldsymbol{\theta})}{\partial\,\boldsymbol{\theta}} \xrightarrow{d} N(\mathbf{0},\,\mathbf{B})$$

where

$$\mathbf{B} = 4\,E\,[\,(X_t^2 - v_{t-1}(\boldsymbol{\theta}))^2\cdot\partial v_{t-1}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\cdot(\partial v_{t-1}(\boldsymbol{\theta})/\partial\boldsymbol{\theta})^T\,]$$

of which the existence is secured by (C.2), which combined with (3.2) yields the theorem. □

We now turn to the maximum likelihood estimation problem. Denoting the density of e_t by $f(\cdot)$, the log-likelihood $l_n(\boldsymbol{\theta})$ at the parameter $\boldsymbol{\theta}$ is given by

$$l_n(\boldsymbol{\theta}) = \sum_{i=1}^{n}\,[\,\log\,f(e_t) - \frac{1}{2}\,\log\,v_{t-1}(\boldsymbol{\theta})\,]\,. \tag{3.3}$$

We introduce $S_n(\boldsymbol{\theta})$ and $F_n(\boldsymbol{\theta})$ for the normalized score function and the average sample Fisher information matrix, respectively, *viz.*,

$$S_n(\boldsymbol{\theta}) = n^{-1/2}\,\partial l_n(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\,, \quad F_n(\boldsymbol{\theta}) = -n^{-1}\,\partial S_n(\boldsymbol{\theta})/\partial\boldsymbol{\theta}. \tag{3.4}$$

Let $\boldsymbol{\theta}$ be arbitrarily chosen but a fixed parameter value which is assumed to be the interior point of the parameter space, and consider the moving parameters $\boldsymbol{\theta}_n = \boldsymbol{\theta} + \mathbf{h}/\sqrt{n}$ converging to $\boldsymbol{\theta}$ as n tends to infinity with $\mathbf{h}$ being a given vector of constants. Define the log-likelihood ratio $R_n(\boldsymbol{\theta},\,\mathbf{h})$ associated with $\boldsymbol{\theta}_n$, and $\boldsymbol{\theta}$ as $R_n(\boldsymbol{\theta},\,\mathbf{h}) = l_n(\boldsymbol{\theta}_n) - l_n(\boldsymbol{\theta})$. In order to obtain the asymptotic expansion of $R_n(\boldsymbol{\theta},\,\mathbf{h})$, assume that the condition below is satisfied.

(C.3) The density $f(\cdot)$ of e_t is such that $i_f = E\,[\,(d\,\log\,f(e_t)\,/\,d\,e_t)^2\,] < \infty$. Also, for the scale family of densities $\{k_b(e_t) = b^{-1}\cdot f(e_t/b),\ b>0\}$, $I(b)$ is continuous in b where $I(b) = E\,[\,(d\,\log\,k_b(e_t)\,/\,d\,b)^2\,] < \infty$.

All the probabilistic statements hereafter in this section will be made under the probability measure corresponding to $\boldsymbol{\theta}$ unless stated otherwise.

Theorem 2. *Assume that $EX_t^4 < \infty$ and, (C.1) and (C.3) hold. We then have,*

(i) $R_n(\boldsymbol{\theta},\,\mathbf{h}) = \mathbf{h}^T S_n(\boldsymbol{\theta}) - \mathbf{h}^T F(\boldsymbol{\theta})\mathbf{h}/2 + o_p(1)$;

(ii) $S(\boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0},\ F(\boldsymbol{\theta}))$,

where $F(\boldsymbol{\theta}) = \mathrm{plim}[F_n(\boldsymbol{\theta})]$.

Proof. The main arguments are straightforward adaptations of the works in Bickel *et al.* (1993, p. 14). We will thus provide outlines of the proof only. Consider the (square root) conditional density ratio of $\boldsymbol{\theta}^*$ to $\boldsymbol{\theta}$:

$$g_t(\boldsymbol{\theta}^*, \boldsymbol{\theta}) = [\,c_t(\boldsymbol{\theta}^*)\,/\,c_t(\boldsymbol{\theta})\,]^{1/2} \tag{3.5}$$

where $c_t(\boldsymbol{\theta})$ stands for the conditional density of X_t given $X_{t-1}, \ldots, X_{t-m}$, *i.e.*, $c_t(\boldsymbol{\theta}) = f(e_t(\boldsymbol{\theta})/\sqrt{v_{t-1}(\boldsymbol{\theta})})$. Provided one can show that $g_t(\boldsymbol{\theta}^*, \boldsymbol{\theta})$ is differentiable in quadratic mean with respect to $\boldsymbol{\theta}^*$ at $\boldsymbol{\theta}^* = \boldsymbol{\theta}$, *viz.*, as $\lambda \to 0$,

$$\lambda^{-1}\,[\,g_t(\boldsymbol{\theta} + \lambda\mathbf{h}, \boldsymbol{\theta}) - g_t(\boldsymbol{\theta}, \boldsymbol{\theta})\,] \longrightarrow \mathbf{h}^T \dot{g}_t(\boldsymbol{\theta}) \ \text{ in quadratic mean} \tag{3.6}$$

where

$$\dot{g}(\boldsymbol{\theta}) = [\,2\,c_t(\boldsymbol{\theta})\,]^{-1}\cdot\partial\,c_t(\boldsymbol{\theta})\,/\,\partial\,\boldsymbol{\theta}, \tag{3.7}$$

the assertion (i) and (ii) are then immediate consequences of quadratic mean differentiability of $g(\boldsymbol{\theta}^*, \boldsymbol{\theta})$. See Roussas (1972, Chapter 2). Hence, our task reduces to verifying (3.6) which is in turn implied by

$$\overline{\lim_{\lambda \to 0}}\, E[\,\lambda^{-2}\,\{\,g_t(\boldsymbol{\theta} + \lambda\,\mathbf{h},\,\boldsymbol{\theta}) - 1\,\}^2\,] \le E[\,\mathbf{h}^T\,\dot{g}_t(\boldsymbol{\theta})\,]^2. \tag{3.8}$$

Exploiting finite fourth order moment condition and continuity of $I(b)$ in (C.3), and employing together the lines in p. 14 of Bickel *et al.* (1993) provides (3.8) which completes the proof. □

The next theorem identifies the limiting distribution of the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}_{ML}$which solves the score equation $S_n(\boldsymbol{\theta}) = \mathbf{0}$.

Theorem 3. *Under the same conditions as in Theorem 2, we have the following as* $n \to \infty$

(i) *With probability tending to one, there exists a unique consistent solution of* $S_n(\boldsymbol{\theta}) = \mathbf{0}$;

(ii) *Let* $\widehat{\boldsymbol{\theta}}_{ML}$ *denote the consistent solution of* $S_n(\boldsymbol{\theta}) = \mathbf{0}$. *Then*

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}) \xrightarrow{d} N(\,\mathbf{0},\; F^{-1}(\boldsymbol{\theta})\,).$$

Proof. Denote the sequence of open δ-neighborhood around $\boldsymbol{\theta}$ by $N_n(\boldsymbol{\theta}, \delta) = \{\boldsymbol{\theta}^* | \sqrt{n}\,||\,\boldsymbol{\theta}^* - \boldsymbol{\theta}|| < \delta\}$. Note that the event $E_n(\boldsymbol{\theta},\, \delta) = \{l_n(\boldsymbol{\theta}^*) - l_n(\boldsymbol{\theta}) < 0,$ for all $\boldsymbol{\theta}^*$ in $N_n(\boldsymbol{\theta},\, \delta)\}$ ensures that the log-likelihood attains a unique local maximum at $\widehat{\boldsymbol{\theta}}_{ML}$, say, inside $N_n(\boldsymbol{\theta},\, \delta)$. From Theorem 2 it is obtained that

$$R_n(\boldsymbol{\theta},\, \mathbf{h}) = l_n(\boldsymbol{\theta}_n) - l_n(\boldsymbol{\theta}) \xrightarrow{d} N(\,-\tau/2,\; \tau)$$

where $\tau = \mathbf{h}^T F(\boldsymbol{\theta})\mathbf{h}$. One can then choose τ sufficiently large so that given $\epsilon > 0$, $P[E_n(\boldsymbol{\theta},\, \delta)] > 1 - \epsilon$, for all sufficiently large n, which implies the assertion (i). See Fahmeir and Kaufmann (1985) for similar arguments in the context of generalized linear models. Since $S_n(\widehat{\boldsymbol{\theta}}_{ML}) = o_p(1)$, together with the Taylor's expansion of $S_n(\boldsymbol{\theta})$ around $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_{ML}$, it can be verified via standard arguments that

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}) = F^{-1}(\boldsymbol{\theta}) \cdot S_n(\boldsymbol{\theta}) + o_p(1)$$

which concludes (ii). □

It follows from Theorem 2 that two probability measures associated with $\boldsymbol{\theta}_n = \boldsymbol{\theta} + \mathbf{h}/\sqrt{n}$ and $\boldsymbol{\theta}$ are mutually contiguous, and asymptotic normality of $\widehat{\boldsymbol{\theta}}_{ML}$ remains valid under $\boldsymbol{\theta}_n$:

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}_n) \xrightarrow{d} N(\mathbf{0},\ F^{-1}(\boldsymbol{\theta}))\ \text{ under }\ \boldsymbol{\theta}_n \tag{3.9}$$

which are related crucially to the non-null limiting distribution of various test statistics (*cf.* Hall and Mathiason, 1990).

4. Tests for Symmetry

This section is concerned with the problem of testing the symmetry of the model. Define $2m \times 1$ vector of constants,

$$\boldsymbol{\eta} = (\alpha_{11} - \alpha_{21}, \ldots, \alpha_{1m} - \alpha_{2m}, \beta_{11} - \beta_{21}, \ldots, \beta_{1m} - \beta_{2m})^T$$

and consider the problem of testing composite hypotheses H for symmetry against the sequence of alternatives K_n:

$$H : \boldsymbol{\eta} = \mathbf{0},\ K_n : \boldsymbol{\eta} = \boldsymbol{\ell}/\sqrt{n} \tag{4.1}$$

with $\boldsymbol{\ell}$ being a $2m \times 1$ vector of constants. The Wald statistic is proposed for testing (4.1)

$$T_n = n \cdot \widehat{\boldsymbol{\eta}}_{ML}^T \cdot [\,\mathbf{M} \cdot F^{-1}(\widehat{\boldsymbol{\theta}}_{ML}) \cdot \mathbf{M}^T\,]^{-1} \cdot \widehat{\boldsymbol{\eta}}_{ML} \tag{4.2}$$

with $\widehat{\boldsymbol{\eta}}_{ML} = \mathbf{M} \cdot \widehat{\boldsymbol{\theta}}_{ML}$, where $2m \times (4m+1)$ matrix $\mathbf{M}$ is defined as

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}$$

so that $\boldsymbol{\eta} = \mathbf{M} \cdot \boldsymbol{\theta}$. The following theorem derives both the null and non-null distributions of T_n.

Theorem 4. *Assume the same conditions as those for Theorem 2. The Wald statistic T_n, under K_n, has a limiting non-central chi-squared distribution with non-centrality parameter specified below. In particular, T_n converges (in distribution) to chi-squared distribution with $2m$ degrees of freedom under H.*

Proof. It is implied by (3.9) that under K_n

$$\sqrt{n}\,(\widehat{\boldsymbol{\eta}}_{ML} - \mathbf{M}\cdot\boldsymbol{\theta}_n) \xrightarrow{d} N(\mathbf{0},\ \mathbf{M}\cdot F^{-1}(\boldsymbol{\theta})\cdot\mathbf{M}^T). \tag{4.3}$$

Let $\boldsymbol{\theta}_H$ denote the parameter vector $\boldsymbol{\theta}$ with restrictions of $\alpha_{1i} = \alpha_{2i}$ and $\beta_{1i} = \beta_{2i}$ as imposed in the null hypothesis of symmetry. It then readily follows from (4.3) that T_n converges to, under K_n, non-central chi-squared distribution with $\boldsymbol{\ell}^T[\,M\cdot F^{-1}(\boldsymbol{\theta}_H)\cdot M^T\,]^{-1}\boldsymbol{\ell}$ as the non-centrality parameter. The null distribution is immediate by setting $\boldsymbol{\ell} = \mathbf{0}$ in K_n. □

Acknowledgements

We thank the two referees for careful reading of the paper. This work is supported by grants from KRF (2000-015-DP0052 and 99-041-D00083-D1200).

REFERENCES

An, H. Z., Chen, M. and Huang, F. C. (1997). "The geometric ergodicity and existence of moments for a class of non-linear time series models", *Statistics & Probability Letters,* **31**, 213–224.

Bickel, P. J., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models,* Johns Hopkins University Press, Baltimore.

Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation", *Econometrica,* **50** , 987–1008.

Guegan, D. and Diebolt, J. (1994). "Probabilistic properties of the β-ARCH model", *Statistica Sinica,* **4**, 71–87.

Fahmeir, L. and Kaufmann, H. (1985). "Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models", *The Annals of Statistics,* **13**, 342–368.

Feigin, P. D. and Tweedie, R. L. (1985). "Random coefficient autoregressive processes : a Markov chain analysis of stationarity and finiteness of moments", *Journal of Time Series Analysis,* **6**, 1–14.

Hall, W. J. and Mathiason, D. J. (1990). "On large sample estimation and testing in parametric models", *International Statistics Review*, **58**, 77–97.

Hwang, S. Y. and Basawa, I. V. (2001). "Estimation for nonlinear autoregressive models generated by beta-ARCH processes", To appear in *Sankhyā A*.

Hwang, S. Y. and Woo, Mi-Ja (2001). "Threshold ARCH(1) processes : asymptotic inference", *Statistics & Probability Letters,* **53**, 11–20.

Klimko, L. A. and Nelson, P. I. (1978). "On conditional least squares estimation for stochastic processes", *The Annals of Statistics,* **6**, 629–642.

Li, C. W. and Li, W. K. (1996). "On a double-threshold autoregressive heteroscedastic time series model", *Journal of Applied Econometrics,* **11**, 253–274.

Roussas, G. G. (1972). *Contiguity of Probability Measures : Some Applications in Statistics,* Cambridge University Press, London.

Tjøstheim, D. (1986). "Estimation in nonlinear time series models", *Stochastic Processes and Their Applications,* **21**, 251–273.