The Korean Communications
in Statistics Vol. 12 No. 3, 2005
pp. 557-587

Two Sample Test Procedures for Linear Rank Statistics for Garch Processes

S. Ajay Chandra, Jito Vanualailai, and Sushil D. Raj[1)]

Abstract

This paper elucidates the limiting Gaussian distribution of a class of rank order statistics $\{T_N\}$ for two sample problem pertaining to empirical processes of the squared residuals from two independent samples of GARCH processes. A distinctive feature is that, unlike the residuals of ARMA processes, the asymptotics of $\{T_N\}$ depend on those of GARCH volatility estimators. Based on the asymptotics of $\{T_N\}$, we empirically assess the relative asymptotic efficiency and effect of the GARCH specification for some GARCH residual distributions. In contrast with the independent, identically distributed or ARMA settings, these studies illuminate some interesting features of GARCH residuals.

Keywords : GARCH process; squared residuals; empirical processes; rank order statistic; asymptotic relative efficiency; GARCH volatility effect.

1. Introduction

Analysis of financial data has received a considerable amount of attention in the literature during the past two decades. Several models have been suggested to capture special features of financial data and most of these models have the property that the conditional variance depends on the past values. One of the well known and most heavily used examples is the class of ARCH(p) processes, introduced by Engle (1982). This process was generalized by Bollerslev (1986) in a manner analogous to the extension from AR to ARMA models in traditional time series by allowing past conditional variances equation. The resulting process is called GARCH(p,q). Since then, ARCH and GARCH related processes have become perhaps the most popular and extensively studied financial econometric models (Comte and Lieberman (2003), Francq and Zakoian (2004), Lee and Taniguchi (2005)). For a class of GARCH(p,q) processes, Bougerol and Picard (1992a, b) established necessary and sufficient conditions for

1) Department of Mathematics and Computing Science University of the South Pacific Suva, Fiji
E-mail : chandra_sa@usp.ac.fj

the existence of a unique stationary solution and gave its explicit representation.

For an ARCH(p) process, Horváth *et al.* (2001) derived the limiting distribution of the empirical process based on the squared residuals. Then they showed that, unlike the residuals of ARMA models, these residuals do not behave in this context like asymptotically independent random variables, and the asymptotic distribution involves a term depending on estimators of the volatility parameters of the process. Later, Berkes *et al.* (2003a) extended the same result to the class of GARCH(p,q) processes. Also Lee and Taniguchi (2005) proved local asymptotic normality for ARCH(∞) processes, and discussed the residual empirical process for an ARCH(1) model with stochastic mean.

Two sample problem in the i.i.d. settings is one of the important statistical problems. For such a problem, the study of the asymptotic properties based on rank order statistics is fundamental and an essential part of nonparametric statistics. Many researchers have contributed to their development, and numerous theorems have been formulated in various testing problems. The classical limit theorem of a properly normalized two sample rank order statistic which generated much interest in this area is the celebrated Chernoff Savage (1958) theorem. They have also proved that the asymptotic normality property is valid in the non null case, subject to certain regularity conditions relating mainly to the smoothness and size of the weights. The Chernoff savage (1958) theorem provides a useful guide in investigating consistency and efficiency properties of most two sample linear rank statistics. Moreover, under less stringent conditions on the score generating functions, Puri and Sen (1993) formulated the same theorem for the one sample, two sample and k−sample problems. Thus, this study motivates us to consider two independent samples from GARCH(p,q) processes $\{X_1\}$, a target process and $\{Y_1\}$. The corresponding squared innovation processes are, say, $\{\xi^2_{x,t}\}$ and $\{\xi^2_{y,t}\}$ with possibly non Gaussian distributions F and G. In order to highlight the problem of testing the equality between these distributions, a nonparametric technique is employed based on the two sample rank order statistics. Such statistics serves as a basis for the comparison in terms of tests of goodness of fit.

The object of this paper is to elucidate the asymptotic theory of the two sample rank order statistics $\{T_N\}$ for GARCH residual empirical processes based on the techniques of Chernoff and Savage (1958) and Berkes and Horváth (2003a). Since the asymptotics of the residual empirical processes are different from those for the usual ARMA case, the limiting distribution of $\{T_N\}$ is greatly different from that of ARMA case (of course the i.i.d. case). More concretely, the paper is organized as follows. Section 2 gives the setting of $\{T_N\}$ pertaining to empirical processes of the squared residuals from two independent samples of GARCH(p,q) processes and derives its asymptotic distribution. This result, in Section 3, facilitates the study of asymptotic performance of $\{T_N\}$, for which, we assess the asymptotic relative efficiency and effect of the GARCH specification for some GARCH residual distributions. These studies help to highlight some important features of GARCH residuals in comparison with the independent, identically distributed or ARMA settings.

2. Two sample rank order statistics and results

In this section we consider an important class of nonparametric tests based on rank order statistics (see Chernoff and Savage (1958)) for two sample problem pertaining to empirical processes based on the squared residuals from two independent samples of GARCH processes.

A class of GARCH(p,q) processes is defined by the equations

$$X_t = \begin{cases} \sigma_t(\theta_x)\varepsilon_t,\ \sigma_t^2(\theta_x) = \alpha_{x,0} + \sum_{i=1}^{p_x} \alpha_{x,i} X_{t-i}^2 + \sum_{j=1}^{q_x} \beta_{x,j} \sigma_{t-j}^2(\theta_x),\ t = 1,\ldots,m, \\ 0, \qquad t = -l_x + 1,\ldots,0, \end{cases} \tag{1}$$

where $l_x = \max(p_x, q_x)$, $\{\varepsilon_t\}$ is a sequence of independent, identically distributed random variables,

$$\theta_x = \left(\theta_{x,0},\ldots,\theta_{x,p_x+q_x}\right)^T \equiv \left(\alpha_{x,0},\alpha_{x,1},\ldots,\alpha_{x,p_x},\beta_{x,1},\ldots,\beta_{x,q_x}\right)^T \qquad \in \Theta \subset \mathbf{R}^{p_x+q_x+1}$$

is an unknown parameter vector satisfying $\alpha_{x,0} > 0$, $\alpha_{x,i} \geq 0$, $i = 1,\ldots p_x$, $\beta_{x,i} \geq 0$, $j = 1,\ldots q_x$ and ε_t is independent of $X_s, s < t$. When $\beta_{x,j} = 0$, $j = 1,\ldots q_x$, the process $\{X_t\}$ reduces to ARCH(p_x). Nelson (1990) showed that in the case of GARCH(1,1), the process $\{X_t\}$ has a unique stationary solution if and only if $E(\beta_{x,1} + \alpha_{x,1}\varepsilon_t^2) < 0$. Bougerol and Picard (1992a, b) established strict stationarity and ergodicity of a general GARCH(p,q) process in terms of the top Lyapunov exponent

$$\lambda_{L,x} = \inf_{0 \le t < \infty} (1+t)^{-1} E\left\{\log\left\|A_x(\varepsilon_0)A_x(\varepsilon_{-1})\cdots A_x(\varepsilon_{-t})\right\|\right\} < 0,$$

where $A_x(\varepsilon_t)$ is a matrix composed of the coefficients of $\{X_t\}$ and ε_t, and $\|\cdot\|$ is the Euclidean norm. Henceforth, denote by $F(x)$ the distribution function of ε_t^2 and we assume the existence of a density $f = F'$, which is continuous on $(0,\infty)$.

Another class of GARCH(p,q) processes, independent of $\{X_t\}$, is defined similarly by the equations

$$Y_t = \begin{cases} \sigma_t(\theta_y)\xi_t,\ \sigma_t^2(\theta_y) = \alpha_{y,0} + \sum_{i=1}^{p_y} \alpha_{y,i} Y_{t-i}^2 + \sum_{j=1}^{q_y} \beta_{y,j} \sigma_{t-j}^2(\theta_y),\ t = 1,\ldots,n, \\ 0, \qquad t = -l_y + 1,\ldots,0, \end{cases} \tag{2}$$

where $l_y = \max(p_y, q_y)$, $\{\xi_t\}$ is a sequence of independent, identically distributed random variables,

$$\theta_y = \left(\theta_{y,0},\ldots,\theta_{y,p_y+q_y}\right)^T \equiv \left(\alpha_{y,0},\alpha_{y,1},\ldots,\alpha_{y,p_y},\beta_{y,0},\beta_{y,1},\ldots,\beta_{y,q_y}\right)^T \quad \in \Theta \subset \mathbf{R}^{p_y+q_y+1},$$

$\alpha_{y,0} > 0$, $\alpha_{y,i} \geq 0$, $i = 1,\ldots p_y$, $\beta_{y,j} \geq 0$, $j = 1,\ldots q_y$, are unknown parameters, and ξ_t is independent of $Y_s, s < t$. It is assumed that

$$\lambda_{L,y} = \inf_{0 \le t < \infty} (1+t)^{-1} E\left\{\log\left\|A_y(\xi_0)A_y(\xi_{-1})\cdots A_y(\xi_{-t})\right\|\right\} < 0$$

for the stationarity of $\{Y_t\}$ where $A_y(\xi_t)$ is a matrix composed of the coefficients of $\{Y_t\}$ and ξ_t. In the case of GARCH(1,1), this condition reduces to $E\left(\beta_{y,1} + \alpha_{y,1}\xi_t^2\right) < 0$. The distribution function of ξ_t^2 will be denoted by $G(x)$ and we assume that $g = G'$ exists, which is continuous on $(0,\infty)$.

In the following, we are concerned with the two sample problem of testing

$$H_0 : F(x) = G(x) \text{ for all } x \text{ against } H_A : F(x) \neq G(x) \text{ for some } x. \tag{3}$$

First we consider the estimation of θ_x and θ_y. Suppose that observed stretches $(X_1,\ldots,X_m)$ and $(Y_1,\ldots,Y_n)$ are available. Then we can rewrite $\sigma_t^2(\theta_x)$ and $\sigma_t^2(\theta_y)$ as a linear function of $(X_{t-1}^2, X_{t-2}^2,\ldots)$ and $(Y_{t-1}^2, Y_{t-2}^2,\ldots)$, respectively. Following Berkes *et al.* (2003a), we define a sequence of functions by recursion. Write

$$u_x = \left(a_{x,0},\, a_{x,1},\ldots,\, a_{x,p_x},\, b_{x,1},\ldots,\, b_{x,q_x}\right)^T \in \mathbf{R}^{p_x+q_x+1}$$

and

$$u_y = \left(a_{y,0},\, a_{y,1},\, \ldots,\, a_{y,p_y},\, b_{y,1},\, \ldots,\, b_{y,q_y}\right)^T \in \mathbf{R}^{p_y+q_y+1}$$

If $q_x \geq p_x$, then

$$
\begin{aligned}
& c_0(u_x) = a_{x,0} \Big/ \left(1 - \left(b_{x,1} + \cdots + b_{x,q_x}\right)\right) \\
& c_1(u_x) = a_{x,1} \\
& c_2(u_x) = a_{x,2} + b_{x,1} c_1(u_x) \\
& \vdots \\
& c_{p_x}(u_x) = a_{x,p_x} + b_{x,1} c_{p_x-1}(u_x) + \cdots + b_{x,p_x-1} c_1(u_x) \\
& c_{p_x+1}(u_x) = b_{x,1} c_{p_x}(u_x) + \cdots + b_{x,p_x} c_1(u_x) \\
& \vdots \\
& c_{q_x}(u_x) = b_{x,1} c_{q_x-1}(u_x) + \cdots + b_{x,q_x-1} c_1(u_x)
\end{aligned}
$$

and if $q_x < p_x$, the above equations are replaced with

$$
\begin{aligned}
& c_0(u_x) = a_{x,0} / (1 - (b_{x,1} + \cdots + b_{x,q_x})) \\
& c_1(u_x) = a_{x,1} \\
& c_2(u_x) = a_{x,2} + b_{x,1} c_1(u_x) \\
& \vdots \\
& c_{q_x+1}(u_x) = a_{x,q_x+1} + b_{x,1} c_{q_x}(u_x) + \cdots + b_{x,q_x} c_1(u_x) \\
& \vdots \\
& c_{p_x}(u_x) = a_{x,p_x} + b_{x,1} c_{p_x-1}(u_x) + \cdots + b_{x,q_x} c_{p_x-q_x}(u_x).
\end{aligned}
$$

In general, if $i > l_x$, then

$$c_i(u_x) = b_{x,1} c_{i-1}(u_x) + b_{x,2} c_{i-2}(u_x) + \cdots + b_{x,q_x} c_{i-q_x}(u_x).$$

They showed for the true value θ_x that

$$\sigma_t^2(\theta_x) = c_0(\theta_x) + \sum_{i=1}^{\infty} c_i(\theta_x) X_{t-i}^2,$$

and analogously for θ_y,

$$\sigma_t^2(\theta_y) = c_0(\theta_y) + \sum_{i=1}^{\infty} c_i(\theta_y) Y_{t-i}^2,$$

The corresponding logarithm of the quasi maximum likelihood functions of $(X_1, \cdots, X_m)$ and $(Y_1, \cdots, Y_n)$ is given by

$$\hat{L}_m(u_x) = -\frac{1}{2}\sum_{t=2}^{m}\left\{\log \hat{h}_t(u_x) + \frac{X_t^2}{\hat{h}_t(u_x)}\right\}$$

and

$$\hat{L}_n(u_y) = -\frac{1}{2}\sum_{t=2}^{n}\left\{\log \hat{h}_t(u_y) + \frac{Y_t^2}{\hat{h}_t(u_y)}\right\}$$

where

$$\hat{h}_t(u_x) = c_0(u_x) + \sum_{i=1}^{t-1} c_i(u_x) X_{t-i}^2, \qquad 2 \le t \le m$$

and

$$\hat{h}_t(u_y) = c_0(u_y) + \sum_{i=1}^{t-1} c_i(u_y) Y_{t-i}^2, \qquad 2 \le t \le n$$

The quasi maximum likelihood estimators, respectively, are defined as

$$\hat{\theta}_{x,m} = \arg \max \left\{\hat{L}_m(u_x),\, u_x \in U_x\right\} \quad \text{and} \quad \hat{\theta}_{y,n} = \arg \max \left\{\hat{L}_n(u_y),\, u_y \in U_y\right\}$$

where

$$U_z = \Big\{ u_z : \beta_{z,1} + \cdots + \beta_{z,q_z} \le \rho_{z,0} \text{ and } r_z < \min\left(a_{z,0}, a_{z,1}, \ldots, a_{z,p_z}, b_{z,1}, \ldots, b_{z,q_z}\right) \le \max(a_{z,0}, a_{z,1}, \ldots, a_{z,p_z}, b_{z,1}, \ldots, b_{z,q_z}) \le r_z^* \Big\}$$

with $0 < r_z < r_z^*$, $0 < \rho_{z,0} < 1$ and $q_z r_z < \rho_{z,0}$. Here, it is assumed that $\theta_x \in U_x$ and $\theta_y \in U_y$. We also assume that

$$m^{1/2}\left\|\hat{\theta}_{x,m}-\theta_x\right\|=O_p(1) \quad \text{and} \quad n^{1/2}\left\|\hat{\theta}_{y,n}-\theta_y\right\|=O_p(1). \tag{4}$$

For the validity of (4), Berkes *et al.* (2003a) gave a set of sufficient conditions without assuming that ε_t and ξ_t are standard normal (see also Comte and Lieberman (2003)). The corresponding empirical squared residuals are given by

$$\hat{\varepsilon}_t^2=X_t^2/\hat{h}_t\left(\hat{\theta}_{x,m}\right), 2\le t\le m \quad \text{and} \quad \hat{\xi}_t^2=Y_t^2/\hat{h}_t\left(\hat{\theta}_{y,n}\right), 2\le t\le n$$

For the testing problem (3), we begin by describing our approach in line with Chernoff and Savage (1958). Let $N=m+n$, $\lambda_N=m/N$, and we assume that the inequalities $0<\lambda_0\le\lambda_N\le 1-\lambda_0<1$ hold for all N and some fixed $\lambda_0\le\frac{1}{2}$. Then the combined distribution is defined by

$$H_N(x)=\lambda_N F(x)+(1-\lambda_N)G(x)$$

In the same way, if $\hat{F}_m(x)$ and $\hat{G}_n(x)$ denote the respective empirical distribution functions $\{\hat{\varepsilon}_t\}$ and $\{\hat{\xi}_t\}$, the corresponding empirical distribution is

$$\hat{H}_N(x)=\lambda_N\hat{F}_m(x)+(1-\lambda_N)\hat{G}_n(x) \tag{5}$$

Note that

$$\hat{U}_m(x)=m^{1/2}\left(\hat{F}_m(x)-F(x)\right)=m^{-1/2}\sum_{t=1}^{m}\left[\mathbf{I}\left(\hat{\varepsilon}_t^2\le x\right)-F(x)\right] \tag{6}$$

and

$$\hat{V}_n(x)=n^{1/2}\left(\hat{G}_n(x)-G(x)\right)=n^{-1/2}\sum_{t=1}^{n}\left[\mathbf{I}\left(\hat{\xi}_t^2\le x\right)-G(x)\right] \tag{7}$$

where $\mathbf{I}(A)$ is the indicator function of the event A. Berkes et al. (2003a) showed that (6) has the following representation

$$\hat{U}_m(x)=U_m(x)+A_x xf(x)+\eta_m(x) \tag{8}$$

where

$$U_m(x) = m^{-1/2} \sum_{t=1}^{m} \left[\mathbf{I}\left(\varepsilon_t^2 \le x\right) - F(x) \right], \quad A_x = \sum_{0 \le i \le p_x + q_x} m^{1/2} \left(\hat{\theta}_{x,i} - \theta_{x,i}\right) \tau_{x,i} \tag{9}$$

and $\sup_x |\eta_m(x)| = o_p(1)$ with $\tau_{x,0} = E\left[1/\sigma_t^2(\theta_x)\right]$ and $\tau_{x,i} = E\left[X_{t-i}^2/\sigma_t^2(\theta_x)\right]$, $1 \le i \le p_x$, $\tau_{x,p_x+i} = E\left[\sigma_{t-i}^2(\theta_x)/\sigma_t^2(\theta_x)\right]$, $1 \le i \le q_x$. In the same way, the corresponding representation of (7) is given by

$$\hat{V}_n(x) = V_n(x) + A_y x g(x) + \eta_n^*(x) \tag{10}$$

where

$$V_n(x) = n^{-1/2} \sum_{t=1}^{n} \left[\mathbf{I}\left(\xi_t^2 \le x\right) - G(x) \right], \quad A_y = \sum_{0 \le i \le p_y + q_y} n^{1/2} \left(\hat{\theta}_{y,i} - \theta_{y,i}\right) \tau_{y,i} \tag{11}$$

and $\sup_x |\eta_n^*(x)| = o_p(1)$ with $\tau_{y,0} = E\left[1/\sigma_t^2(\theta_y)\right]$ and $\tau_{y,i} = E\left[X_{t-i}^2/\sigma_t^2(\theta_y)\right]$, $1 \le i \le p_y$, $\tau_{y,p_y+i} = E\left[\sigma_{t-i}^2(\theta_y)/\sigma_t^2(\theta_y)\right]$, $1 \le i \le q_y$. Write $F_m(x) = m^{-1} \sum_{t=1}^{m} \mathbf{I}(\varepsilon_t^2 \le x)$ and $G_n(x) = n^{-1} \sum_{t=1}^{n} \mathbf{I}(\xi_t^2 \le x)$. From (8) and (10), the expression (5) then becomes

$$\hat{H}_N(x) = \mathcal{H}_N(x) + m^{-1/2} \lambda_N A_x x f(x) + n^{-1/2} (1 - \lambda_N) A_y x g(x) + \eta_N^{**}(x) \tag{12}$$

where $\mathcal{H}_N(x) = \lambda_N F_m(x) + (1 - \lambda_N) G_n(x)$ and $\eta_N^{**}(x) = m^{-1/2} \lambda_N \eta_m(x) + n^{-1/2} (1 - \lambda_N) \eta_n^*(x)$. The decomposition (12) is important and will be used repeatedly in the sequel.

Define $S_{N,i} = 1$, if the i^{th} smallest one in the combined residuals $\hat{\varepsilon}_2^2, \ldots, \hat{\varepsilon}_m^2$, $\hat{\xi}_2^2, \ldots, \hat{\xi}_n^2$ is from $\hat{\varepsilon}_2^2, \ldots, \hat{\varepsilon}_m^2$ and otherwise define $S_{N,i} = 0$. Now, let us consider the two sample rank order statistics of the form

$$T_N = \frac{1}{m} \sum_{i=1}^{N} R_{N,i} S_{N,i}$$

where the $R_{N,i}$ are given constants called scores. A key feature in the Chernoff Savage theory is that a linear rank statistic can be represented in the form of a Stieltjes integral. Thus, if the weights for a linear rank order are functions of the ranks, an equivalent representation of T_N is

$$T_N = \int J\left[\frac{N}{N+1}\hat{H}_N(x)\right]d\hat{F}_m(x), \tag{13}$$

where $R_{N,i} = J(i/(N+1))$, and $J(u), 0 < u < 1$, is a continuous function. Several typical examples of J which have been reported in Puri and Sen (1993) are stated below:

(i) Wilcoxon's two sample test with $J(u) = u, 0 < u < 1$,

(ii) Van der Waerden's two sample test with $J(u) = \Phi(u), 0 < u < 1$, where

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{0} \exp(-t^2/2)dt,$$

(iii) Mood's two sample test with $J(u) = (u - \frac{1}{2})^2, 0 < u < 1,$

(iv) Klotz's two sample normal scores test with $J(u) = (\Phi^{-1}(u))^2, 0 < u < 1.$

To elucidate the asymptotics of (13), we impose the following regularity conditions.

Assumption 1.

(A.1) $J(u)$ is not constant and has a continuous derivative $J'(u)$ on (0,1).

(A.2) $|J(u)| \le K[u(1-u)]^{-1/2+\delta}$ and $|J'(u)| \le K[u(1-u)]^{-3/2+\delta}$ for some $\delta > 0$ and K is any constant which does not depend on m, n, N, F or G.

(A.3) $xf(x)$, $xg(x)$, $xf'(x)$ and $xg'(x)$ are uniformly bounded continuous, and integrable functions on $(0,\infty)$.

(A.4) There exists $c > 0$ such that $F(x) \ge c\{xf(x)\}$ and $G(x) \ge c\{xg(x)\}$ for all $x > 0$.

(A.5) $|xf(x)| < KH_N(x)(1-H_N(x))$ and $|xg(x)| < KH_N(x)(1-H_N(x))$ for all $x > 0$ and $K > 0$.

Note that the above examples of J satisfy the conditions (A.1) and (A.2). We also require the following assumption.

Assumption 2.

(B.1) ε_t^2 and ξ_t^2 are nondegenerate random variables.

(B.2) $E(\varepsilon_t^2)=1$, $E(\xi_t^2)=1$, $E|\varepsilon_t^2|^{2+\delta}<\infty$ and $E|\xi_t^2|^{2+\delta}<\infty$ for some $\delta>0$.

The conditions (B.1) and (B.2) of Assumption 2 uniquely identify θ_x and θ_y (see Berkes *et al.* (2003a)). In order to state the result, we introduce the following:

$$U_x = E\left\{\frac{1}{\sigma_t^4(\theta_x)}\left(\frac{\partial}{\partial\theta_x}\sigma_t^2(\theta_x)\right)\left(\frac{\partial}{\partial\theta_x}\sigma_t^2(\theta_x)\right)^T\right\}$$

$$U_y = E\left\{\frac{1}{\sigma_t^4(\theta_y)}\left(\frac{\partial}{\partial\theta_y}\sigma_t^2(\theta_y)\right)\left(\frac{\partial}{\partial\theta_y}\sigma_t^2(\theta_y)\right)^T\right\}$$

$$R_x = \left(E\varepsilon_t^4 - 1\right)U_x,\quad R_y = \left(E\xi_t^4 - 1\right)U_y$$

$$\gamma_{x,t} = \frac{1}{\sigma_t^2(\theta_x)}\frac{\partial}{\partial\theta_x}\sigma_t^2(\theta_x),\quad \gamma_{y,t} = \frac{1}{\sigma_t^2(\theta_y)}\frac{\partial}{\partial\theta_y}\sigma_t^2(\theta_y)$$

Under certain regularity conditions, Berkes et al. (2003b) showed that U_x, U_y and R_x, R_y are non singular. Now, using standard arguments, it is seen that the i^{th} element of $\hat{\theta}_{x,m}$ and $\hat{\theta}_{y,n}$ admits the representations

$$\hat{\theta}^i_{x,m} - \theta^i_x = \frac{1}{m}\sum_{t=2}^{m} Z^i_{x,t}(\varepsilon_t^2 - 1) + o_p(m^{-1/2}), \qquad 0 \le i \le p_x + q_x$$

and

$$\hat{\theta}^i_{y,n} - \theta^i_y = \frac{1}{n}\sum_{t=2}^{n} Z^i_{t,t}(\xi_t^2 - 1) + o_p(n^{-1/2}), \qquad 0 \le i \le p_y + q_y$$

where $Z^i_{x,t}$ and $Z^i_{y,t}$ are the i^{th} element of $U_x^{-1}\gamma_{x,t}$ and $U_y^{-1}\gamma_{y,t}$, respectively. Write $\delta_{x,i} = E\left(Z^i_{x,t}\right)$, $0 \le i \le p_x + q_x$, $\delta_{y,i} = E\left(Z^i_{y,t}\right)$, $0 \le i \le p_y + q_y$, and $\tau_x = \left(\tau_{x,0}, \dots, \tau_{x,p_x+q_x}\right)^T$ and $\tau_y = \left(\tau_{y,0}, \dots, \tau_{y,p_y+q_y}\right)^T$ (recall (9) and (11)). Then we have the following result, whose proof is given in Section 4.

Theorem 1. *Suppose that Assumptions 1 and 2 hold and that, in addition,* $\hat{\theta}_{x,m}$ *and* $\hat{\theta}_{y,n}$ *are the quasi maximum likelihood estimators of* θ_x *and* θ_y *satisfying (4). Then*

$$N^{1/2}(T_N - \mu_N)/\sigma_N \xrightarrow{d} \mathcal{N}(0,1) \quad as \quad N \to \infty,$$

where

$$\mu_N = \int J[H_N(x)]\, dF(x) \quad and \quad \sigma_N^2 = \sigma_{1N}^2 + \sigma_{2N}^2 + \sigma_{3N}^2 + \gamma_N \neq 0$$

with

$$\sigma_{1N}^2 = 2(1-\lambda_N)\left\{ \iint_{x<y} A_N(x,y)dF(x)dF(y) + \frac{1-\lambda_N}{\lambda_N} \iint_{x<y} B_N(x,y)dG(x)dG(y) \right\},$$

$$\sigma_{2N}^2 = \varphi_{x,N}^T U_x^{-1} R_x U_x^{-1} \varphi_{x,N}, \qquad \sigma_{3N}^2 = \varphi_{y,N}^T U_y^{-1} R_y U_y^{-1} \varphi_{y,N}$$

and

$$\gamma_N = 2(1-\lambda_N)\left\{ \frac{1-\lambda_N}{\lambda_N} L_{x,N} \sum_{0 \le i \le p_x+q_x} \tau_{x,i}\delta_{x,i} + L_{y,N} \sum_{0 \le i \le p_y+q_y} \tau_{y,i}\delta_{y,i} \right\},$$

where

$$A_N(x,y) = G(x)(1-G(y))J'[H_N(x)]J'[H_N(y)],$$

$$B_N(x,y) = F(x)(1-F(y))J'[H_N(x)]J'[H_N(y)],$$

$$\varphi_{x,N} = -(\lambda_N)^{-1/2}(1-\lambda_N)\int xf(x)J'[H_N(x)]dG(x) \times \tau_x,$$

$$\varphi_{y,N} = (1-\lambda_N)^{1/2} \int zg(z)J'[H_N(z)]\,dF(z) \times \tau_y,$$

$$L_{x,N} = \iint \psi_x(x) yf(y) J'[H_N(x)]J'[H_N(y)]dG(x)dG(y),$$

$$L_{y,N} = \iint \psi_y(x) zg(z) J'[H_N(x)]J'[H_N(z)]\,dF(x)dF(z).$$

with

$$\psi_x(x) = \int_0^x (u-1) f(u) du \quad \text{and} \quad \psi_y(x) = \int_0^x (u-1) g(u) du$$

Remark 2.1. If $\{\varepsilon_t\}$ and $\{\xi_t\}$ are Gaussian, then $R_x = 2U_x$ and $R_y = 2U_y$, respectively.

Remark 2.2. The terms σ_{2N}^2, σ_{3N}^2 and γ_N depend on the GARCH volatility estimators $\hat{\theta}_{x,m}$ and $\hat{\theta}_{y,n}$. Hence, the asymptotics of $\{T_N\}$ are greatly different in comparison with the independent, identically distributed or ARMA settings.

Remark 2.3 For $\{T_N\}$ to be practically feasible, it is necessary to replace σ_N^2 which depends on several unknown parameters and functions by a consistent estimator $\hat{\sigma}_N^2$. Observe that $\delta_{x,i}, \tau_{x,i}, \delta_{y,j}, \tau_{y,j};\ 0 \le i \le p_x + q_x,\ 0 \le j \le p_y + q_y,$ and $\psi_x(x)$ and $\psi_y(x)$, are expected values and can be consistently estimated by the corresponding averages. Note also that $U_x^{-1} R_x U_x^{-1}$ and $U_y^{-1} R_y U_y^{-1}$ are the asymptotic covariance matrices of $\sqrt{m}(\hat{\theta}_{x,m} - \theta_x)$ and $\sqrt{n}(\hat{\theta}_{y,n} - \theta_y)$, respectively, and their estimation is discussed in Gouriéroux (1997).

3. Asymptotic performance of $\{T_N\}$

The limiting distribution of $\{T_N\}$ given in the preceding section facilitates the study of relative asymptotic efficiency and GARCH volatility effect. Thus we may proceed to demonstrate empirically these aspects of $\{T_N\}$ for some GARCH residual distributions.

3.1 Asymptotic relative efficiency

This subsection considers the assessment of asymptotic relative efficiency among different tests based on $\{T_N\}$. To begin with, let us state a set of Pitman regularity conditions which makes the calculation of efficiency for two test sequences quite easy in the case of finite sample sizes. Suppose that T_N is a test statistic based on the first N observations for testing $H_0 : \theta = \theta_0$ against $H_A : \theta > \theta_0$ with critical region $T_N \ge \lambda_{N,\alpha}$. Further, suppose

(i) $\lim_{N \to \infty} P_{\theta_0}(T_N \ge \lambda_{N,\alpha}) = \alpha$, where $0 < \alpha < 1$ is a given level;

(ii) there exist function $\mu_N(\theta)$ and $\sigma_N(\theta)$ such that $N^{1/2}(T_N - \mu_N(\theta))/\sigma_N(\theta) \xrightarrow{d} \mathcal{N}(0,1)$ uniformly in $\theta \in [\theta_0, \theta_0 + \varepsilon], \varepsilon > 0$;

(iii) $\mu_N'(\theta) > 0$;

(iv) for a sequence $\{\theta_N = \theta_0 + N^{-1/2}\delta, \delta > 0\}$, $\lim_{N\to\infty}[\mu_N'(\theta_N)/\mu_N'(\theta_0)] = 1$, $\lim_{N\to\infty}[\sigma_N(\theta_N)/\sigma_N(\theta_0)] = 1$;

(v) $\lim_{N\to\infty}[\mu_N'(\theta_0)/\sigma_N(\theta_0)] = c > 0$

Then the asymptotic power is given by $1-\Phi(\lambda_\alpha - \delta c)$, where $\lambda_\alpha = \Phi^{-1}(1-\alpha)$. The quantity c defined by (v) is called the efficacy of T_N. It is known that the asymptotic power, in addition to providing a measure of performance, also serves as a basis for the comparison of different tests.

Let $T^{(1)} = \{T_N^{(1)}\}$ and $T^{(2)} = \{T_N^{(2)}\}$ be test sequences with efficacies c_1 and c_2, respectively. Then the asymptotic relative efficiency (ARE) of $T^{(1)}$ relative to $T^{(2)}$ is given by

$$e\left(T^{(1)}, T^{(2)}\right) = c_1^2 / c_2^2$$

Now consider the GARCH(1,1) process given by

$$X_t = \begin{cases} \sigma_t(\theta_x)\varepsilon_t, \ \sigma_t^2(\theta_x) = \alpha_{x,0} + \alpha_{x,1}X_{t-1}^2 + \beta_{x,1}\sigma_{t-1}^2(\theta_x), & t = 1,\ldots,m, \\ 0, & t \le 0, \end{cases}$$

where $\{\varepsilon_t\}$ is a sequence of independent, identically distributed random variables, $\alpha_{x,0} > 0$, $\alpha_{x,1} \ge 0$, $\beta_{x,1} \ge 0$, $E(\beta_{x,1} + \alpha_{x,1}\varepsilon_t^2) < 0$, and ε_t is independent of $X_s, s < t$.

Another GARCH(1,1) process, independent of $\{X_t\}$, is given by

$$Y_t = \begin{cases} \sigma_t(\theta_y)\xi_t, \ \sigma_t^2(\theta_y) = \alpha_{y,0} + \alpha_{y,1}Y_{t-1}^2 + \beta_{y,1}\sigma_{t-1}^2(\theta_y), & t = 1,\ldots,n, \\ 0, & t \le 0, \end{cases}$$

where $\{\xi_t\}$ is a sequence of independent, identically distributed random variables, $\alpha_{y,0} > 0$, $\alpha_{y,1} \ge 0$, $\beta_{y,1} \ge 0$, $E(\beta_{y,1} + \alpha_{y,1}\xi_t^2) < 0$, and ξ_t is independent of $Y_s, s < t$.

Based on the results of Section 2, we consider the scale problem in the case of $G(x) = F(\theta x)$, $\theta > 0$, when F is arbitrary and has finite variance σ_F^2. This motivates the two sample testing problem for scale as follows;

$$H_0 : \theta = 1 \quad \text{against} \quad H_A : \theta > 1$$

For this we first evaluate the efficacy of the Wilcoxon, Van der Waerden, Mood and Klotz tests by applying Theorem 1 to the corresponding J function. We begin with Wilcoxon's W test, whose score function is given by $J(u) = u, \quad 0 < u < 1$. Henceforth, for the sake of simplicity , assume that $m = n = N/2$. Then the mean is

$$\mu_W(\theta) = \frac{1}{2}\int \left(F(x) + F(\theta x)\right) dF(x)$$

and that $\mu_W(\theta)$ is continuously differentiable with respect to θ at $\theta = 1$ under the integral sign, we have $\mu'_W(1) = \frac{1}{2}\int x f^2(x) dx$, and the variance under $H_0 : \theta = 1$ is

$$\sigma^2_W(1) = \sigma^2_{W,1}(x) + \sigma^2_{W,2}(x) + \sigma^2_{W,3} + \gamma_W(1),$$

where

$$\sigma^2_{W,1}(1) = \int_0^1 u^2 du - \left(\int_0^1 u du\right)^2 = \frac{1}{12},$$

$$\sigma^2_{W,2}(1) = \frac{C_x}{2}\left(\int x f^3(x) dx\right)^2, \qquad \sigma^2_{W,3}(1) = \frac{C_y}{2}\left(\int z f^3(z) dz\right)^2$$

and

$$\gamma_W(1) = \tilde{C}_x \iint \psi_x(x) y f^2(x) f^3(y) dx dy + \tilde{C}_y \iint \psi_y(x) z f^2(x) f^3(z) dx dz$$

with

$$C_x = \left[\tau_{x,0}, \tau_{x,1}, \tau_{x,2}\right] U_x^{-1} R_x U_x^{-1} \left[\tau_{x,0}, \tau_{x,1}, \tau_{x,2}\right]^T,$$

$$C_y = \left[\tau_{y,0}, \tau_{y,1}, \tau_{y,2}\right] U_y^{-1} R_y U_y^{-1} \left[\tau_{y,0}, \tau_{y,1}, \tau_{y,2}\right]^T,$$

$$\tilde{C}_x = \tau_{x,0}\delta_{x,0} + \tau_{x,1}\delta_{x,1} + \tau_{x,2}\delta_{x,2} \quad \text{and} \quad \tilde{C}_y = \tau_{y,0}\delta_{y,0} + \tau_{y,1}\delta_{y,1} + \tau_{y,2}\delta_{y,2}.$$

Thus the efficacy of the W test is

$$c_W = \mu'_W(1)/\sigma_W(1).$$

In the same way for Van der Waerden's VW test with $J(u) = \Phi^{-1}(u), \quad 0 < u < 1$, the efficacy is

$$c_{VW} = \mu'_{VW}(1)/\sigma_{VW}(1),$$

where

$$\mu'_{VW}(1) = \tfrac{1}{2}\int xh(x)dx \quad \text{and} \quad \sigma^2_{VW}(1) = \sigma^2_{VW,1}(1) + \sigma^2_{VW,2}(1) + \sigma^2_{VW,3}(1) + \gamma_{VW}(1)$$

with

$$\sigma^2_{VW,1}(1) = \int_0^1 \left(\Phi^{-1}(u)\right)^2 du - \left(\int_0^1 \Phi^{-1}(u)du\right)^2 = 1,$$

$$\sigma^2_{VW,2}(1) = \frac{C_x}{2}\left(\int xh(x)f(x)dx\right)^2, \qquad \sigma^2_{VW,3}(1) = \frac{C_y}{2}\left(\int zh(z)f(z)dz\right)^2$$

and

$$\gamma_{VW}(1) = \tilde{C}_x \iint \psi_x(x)yh(x)h(y)f(y)dxdy + \tilde{C}_y \iint \psi_y(x)zh(x)h(z)f(z)dxdz$$

and

$$h(s) = f^2(s)/\varphi\{\Phi^{-1}[F(s)]\},$$

where

$$\varphi(x) = \Phi'(x).$$

For Mood's M test with $J(u) = \left(u - \frac{1}{2}\right)^2, \quad 0 < u < 1$, the efficacy is

$$c_M = \mu'_M(1)/\sigma_M(1).$$

where,

$$\mu'_M(1) = \int x\tilde{h}(x)dx \text{ and } \sigma_M^2(1) = \sigma_{M,1}^2(1) + \sigma_{M,2}^2(1) + \sigma_{M,3}^2(1) + \gamma_M(1)$$

with

$$\sigma_{M,1}^2(1) = \int_0^1 \left(u - \tfrac{1}{2}\right)^4 du - \left(\int_0^1 \left(u - \tfrac{1}{2}\right)^2 du\right)^2 = \frac{1}{180},$$

$$\sigma_{M,2}^2(1) = C_x\left(\int x\tilde{h}(x)f(x)dx\right)^2, \quad \sigma_{M,3}^2(1) = C_y\left(\int z\tilde{h}(z)f(z)dz\right)^2 \text{ and}$$

$$\gamma_M(1) = 4\left(\tilde{C}_x \iint \psi_x(x)y\tilde{h}(x)\tilde{h}(y)f(y)dxdy + \tilde{C}_y \iint \psi_y(x)z\tilde{h}(x)\tilde{h}(z)f(z)dxdz\right)$$

and $\tilde{h}(s) = \left(F(s) - \tfrac{1}{2}\right)f^2(s)$.

Finally, for Klotz's K test with $J(u) = \left(\Phi^{-1}(u)\right)^2, \quad 0 < u < 1$, the efficacy is

$$c_K = \mu'_K(1)/\sigma_K(1),$$

where,

$$\mu'_K(1) = \int xh^*(x)dx \text{ and } \sigma_K^2(1) = \sigma_{K,1}^2(1) + \sigma_{K,2}^2(1) + \sigma_{K,3}^2(1) + \gamma_K(1)$$

with

$$\sigma_{K,1}^2(1) = \int_0^1 \left(\Phi^{-1}(u)\right)^4 du - \left(\int_0^1 \left(\Phi^{-1}(u)\right)^2 du\right)^2 = 2$$

$$\sigma_{K,2}^2(1) = C_x\left(\int xh^*(x)f(x)dx\right)^2, \quad \sigma_{K,3}^2(1) = C_y\left(\int zh^*(z)f(z)dz\right)^2 \text{ and}$$

$$\gamma_K(1) = 4\left(\tilde{C}_x \iint \psi_x(x)yh^*(x)h^*(y)f(y)dxdy + \tilde{C}_y \iint \psi_x(x)zh^*(x)h^*(z)f(z)dxdz\right)$$

and $h^*(s) = h(s)\Phi^{-1}\left(F(s)\right)$.

We are now ready to list the ARE formulas for the above tests. Hence, from the definition we have the following:

$$e(W,VW)=c_W^2/c_{VW}^2,\ \ e(M,VW)=c_M^2/c_{VW}^2,\ \ e(K,VW)=c_K^2/c_{VW}^2,$$
$$e(W,K)=c_W^2/c_K^2,\ e(M,K)=c_M^2/c_K^2,\ \ e(M,W)=c_M^2/c_W^2,$$

These formulas provide a basis for comparing the six tests for different distribution F. In order to approximate values of $e(\cdot,\ \cdot)$, we need to specify F. For this, let us suppose that $\{\varepsilon_t\}$ is a sequence of i.i.d.(0,1) random variables with continuous symmetric distribution F^* and density f^*. Then

$$F(x)=\begin{cases}2F^*\left(\sqrt{x}\right)-1, & x>0,\\ 0 & x\le 0.\end{cases}$$

We now compute the preceding distribution function in the following particular choices of F^*.

(i) F^*(Normal) : $F_{\mathbf{N}}^*(y)=\int_{-\infty}^{y}(2\pi)^{-1/2}e^{-t^2/2}dt$, $f_{\mathbf{N}}^*(y)=(2\pi)^{-1/2}e^{-y^2/2}, y\in\mathbf{R}$.

In this case, $F_{\mathbf{N}}(x)=2F_{\mathbf{N}}^*\left(\sqrt{x}\right)-1$, $f_{\mathbf{N}}(x)=(2\pi x)^{-1/2}e^{-x/2}, x>0$

(ii) F^*(Double exponential) : $F_{\mathbf{DE}}^*(y)=\int_{-\infty}^{y}2^{-1}e^{-|t|}dt=1-2^{-1}e^{-y}$, $f_{DE}^*(y)=2^{-1}e^{-|y|}, y\in\mathbf{R}$.

In this case, $F_{\mathbf{DE}}(x)=1-e^{-\sqrt{x}}$, $f_{\mathbf{DE}}(x)=\left(2\sqrt{x}\right)^{-1}e^{-\sqrt{x}}, x>0$

(iii) F^*(Logistic): $F_{\mathbf{L}}^*(y)=1/\left(1+e^{-y}\right)$, $f_{\mathbf{L}}^*(y)=e^{-y}/\left(1+e^{-y}\right)^2, y\in\mathbf{R}$.

In this case, $F_{\mathbf{L}}(x)=\left(1-e^{-\sqrt{x}}\right)/\left(1+e^{-\sqrt{x}}\right)$, $f_{\mathbf{L}}(x)=e^{-\sqrt{x}}/\sqrt{x}\left(1+e^{-\sqrt{x}}\right)^2\ x>0$.

In our case, $T_N=T_N(F)$ was constructed from the empirical residuals $\{\hat{\varepsilon}_t^{\,2}\}$ and $\{\hat{\xi}_t^{\,2}\}$. In the same way, if we construct it replacing $\{\hat{\varepsilon}_t^{\,2}\}$ and $\{\hat{\xi}_t^{\,2}\}$ by $\{\varepsilon_t^{\,2}\}$ and $\{\xi_t^{\,2}\}$, respectively, then it becomes the usual rank order statistic given in Puri and Sen (1993) which is denoted by T_N^{PS}. To facilitate comparisons among the tests for $m = n = N/2$, $F_{\mathbf{N}}$, $F_{\mathbf{DE}}$ and $F_{\mathbf{L}}$, and parameters, set $\alpha_0=\alpha_{x,0}=\alpha_{y,0}, \alpha_1=\alpha_{x,1}=\alpha_{y,1}$, and $\beta_1=\beta_{x,1}=\beta_{y,1}$. Then the protocol of the study is the following. For α_0 = 0.1, α_1 = 0.2, β_1 = 0.2 and $m = n$ = 100, 300, we generated realizations of X_t and Y_t. All the estimation results below are based on 100 replications. Table 1 reports various values of $e(\cdot,\ \cdot)$ in the i.i.d. setting, that is, T_N^{PS} for $F = F_{\mathbf{N}}$, $F_{\mathbf{DE}}$ and $F_{\mathbf{L}}$, and $\sigma_1^2(1)=\sigma_{W,1}^2(1), \sigma_{VW,1}^2(1), \sigma_{M,1}^2(1)$ and $\sigma_{K,1}^2(1)$. Based on our setting, that is, $T_N(F)$, Tables 2, 3 and 4 provide these values in the cases of $m = n$ = 100, 300, $(\alpha_0,\alpha_1,\beta_1)$ = (0.1, 0.2.0.2), and

$\sigma^2(1) = \sigma_W^2(1), \sigma_{VW}^2(1), \sigma_M^2(1)$ and $\sigma_K^2(1)$.

Table 1 Various values of $e(\cdot,\cdot)$ based on T_N^{PS}

ARE	Distribution		
	F_N	F_{DE}	F_L
$e(W,VW)$	0.4689	1.7273	2.2320
$e(M,W)$	0.3662	0.2492	0.2170
$e(W,K)$	0.9978	1.0200	1.5631
$e(M,VW)$	0.1717	0.4305	0.6145
$e(K,VW)$	0.4699	1.0934	1.2118
$e(M,K)$	0.2542	0.3392	0.3654

Table 2 Various values of $e(\cdot,\cdot)$ based on $T_N(F_N)$

ARE	ARCH(1,1)		GARCH(1,1)	
	$m = n = 100$	$m = n = 300$	$m = n = 100$	$m = n = 300$
	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.2$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.2$
$e(W,VW)$	0.4982	0.4784	0.4812	0.4751
$e(M,W)$	0.4288	0.4032	0.3982	0.3863
$e(W,K)$	0.9399	0.9581	0.9421	0.9706
$e(M,VW)$	0.2068	0.1864	0.1932	0.1819
$e(K,VW)$	0.5033	0.4815	0.4932	0.4786
$e(M,K)$	0.3108	0.2854	0.2998	0.2771

Table 3 Various values of $e(\cdot,\cdot)$ based on $T_N(F_{DE})$

ARE	ARCH(1,1)		GARCH(1,1)	
	$m = n = 100$	$m = n = 300$	$m = n = 100$	$m = n = 300$
	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.2$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.2$
$e(W,VW)$	1.7525	1.7348	1.7481	1.7311
$e(M,W)$	0.3057	0.2833	0.2963	0.2712
$e(W,K)$	1.0042	1.0116	1.0098	1.0166
$e(M,VW)$	0.5304	0.4936	0.5082	0.4642
$e(K,VW)$	1.0612	1.0723	1.0685	1.0794
$e(M,K)$	0.3831	0.3614	0.3756	0.3522

Table 4 Various values of $e(\cdot,\cdot)$ based on $T_N(F_L)$

ARE	ARCH(1,1)		GARCH(1,1)	
	$m = n = 100$	$m = n = 300$	$m = n = 100$	$m = n = 300$
	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.2$	$\alpha_0 = 0.1, \alpha_1 = 0.2, \beta_1 = 0.2$
$e(W,VW)$	2.7142	2.7480	2.7331	2.7671
$e(M,W)$	0.2594	0.2349	0.2402	0.2302
$e(W,K)$	1.5917	1.5774	1.5812	1.5716
$e(M,VW)$	0.6527	0.6302	0.6441	0.6262
$e(K,VW)$	1.1423	1.1583	1.1513	1.1878
$e(M,K)$	0.4228	0.4047	0.4191	0.3938

An examination of the ARE values in Tables 1-4 reveals some distinctive features. Evidently, the values in Tables 2-4 are intrinsically stable with respect to the choice of $m = n$, the parameters and distributions, but differ from those in Table 1 because of the asymptotics of the volatility estimators $\hat{\theta}_{x,m}$ and $\hat{\theta}_{y,n}$. Interestingly, however, when $m = n$ is increased, the values for the tests in the GARCH(1,1) case are closer to those in Table 1 than for the same tests in the ARCH(1) case. Roughly speaking, the contributing factor for this closeness is the weights $1/\sigma_t^2(\theta_x), \beta_{x,1} > 0$ and $1/\sigma_t^2(\theta_y), \beta_{y,1} > 0$, in the quantities $C_x, \tilde{C}_x$ and $C_y, \tilde{C}_y$, which reduce the effect of $\hat{\theta}_{x,m}$ and $\hat{\theta}_{y,n}$ respectively, while causing only little change in $F = F_*$. On the contrary, the tests in the latter are strong competitors to the tests in the former for all $m = n$, the parameters and distributions. In addition, for distributions $F = F_{DE}$ and $F = F_L$ with heavier tails than $F = F_N$, the W test is superior in that the values are greater than or equal to one. Likewise, the VW and K tests outperform their counterparts in the case of $F = F_N$. We also observe that the VW test is much less effcient than the K test for $F = F_{DE}$ and $F = F_L$. The M test performance is the poorest for all $m = n$, the parameters and distributions. We therefore summarize by saying that the VW and K tests are preferable if the distribution is normal while the W test is preferable if the distribution is logistic.

3.2 GARCH volatility effect

In this subsection we study a distinction of T_N and T_N^{PS} in terms of their levels of tests. For $\alpha \in (0,1)$, write $\lambda_\alpha = \Phi^{-1}(1-\alpha)$. Denote by μ_1 the mean of each test and write $\sigma_1^2 = \sigma_1^2(1)$. Suppose that $N^{1/2}(T_N^{PS} - \mu_1)/\sigma_1 \xrightarrow{d} \mathcal{N}(0,1)$ holds. Then the test $N^{1/2}(T_N^{PS} - \mu_1)/\sigma_1 \geq \lambda_\alpha$ has nominal asymptotic level α as $N \to \infty$. We assume α to be less than 0.5 so that $\lambda_\alpha > 0$.

For this λ_α, let $\tilde{\alpha}_N = P\{N^{1/2}(T_N - \mu_1)/\sigma \geq \lambda_\alpha\}$, where $\sigma^2 = \sigma^2(1)$. Then $\tilde{\alpha} = \lim_{N\to\infty} \tilde{\alpha}_N$ exists and is given by $\tilde{\alpha} = 1 - \Phi(\lambda_\alpha \delta)$, where $\delta = \sigma_1/\sigma$. Henceforth, we write $\sigma = \sigma_F$ if the concerned distribution is F. Since $\sigma_F \geq \sigma_1$, $\tilde{\alpha} \geq \alpha$.

To distinguish how much the actual $\tilde{\alpha}$ varies from the nominal α, we use the level α = 0.05 for which $\lambda_{0.05}$ = 1.645. Based on the preceding realizations of X_t and Y_t, and $F = F_N$, F_{DE} and F_L, we provide the results in Tables 5 - 7.

Table 5 Actual $\tilde{\alpha} = 1 - \Phi(\lambda_\alpha \delta), \delta = \sigma_1/\sigma$, when nominal level α = 0.05 based on $T_N(F_N)$

$\tilde{\alpha}$	ARCH(1,1)		GARCH(1,1)	
	$m = n = 100$	$m = n = 300$	$m = n = 100$	$m = n = 300$
	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.2$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.2$
δ_W	0.8712	0.8917	0.8858	0.9106
$\tilde{\alpha}_W$	0.0759	0.0712	0.0725	0.0671
δ_{VW}	0.9150	0.9246	0.9151	0.9316
$\tilde{\alpha}_{VW}$	0.0661	0.0641	0.0661	0.0627
δ_M	0.7841	0.8014	0.7913	0.8193
$\tilde{\alpha}_M$	0.0986	0.0937	0.0965	0.0889
δ_K	0.9123	0.9221	0.9124	0.9309
$\tilde{\alpha}_K$	0.0667	0.0647	0.0667	0.0628

Table 6 Actual $\tilde{\alpha} = 1 - \Phi(\lambda_\alpha \delta), \delta = \sigma_1/\sigma$, when nominal level α = 0.05 based on $T_N(F_{DE})$

$\tilde{\alpha}$	ARCH(1,1)		GARCH(1,1)	
	$m = n = 100$	$m = n = 300$	$m = n = 100$	$m = n = 300$
	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.2$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.2$
δ_W	0.9042	0.9186	0.9072	0.9248
$\tilde{\alpha}_W$	0.0685	0.0654	0.0678	0.0641
δ_{VW}	0.8681	0.8752	0.8702	0.8843
$\tilde{\alpha}_{VW}$	0.0766	0.0750	0.0762	0.0729
δ_M	0.8163	0.8272	0.8194	0.8391
$\tilde{\alpha}_M$	0.0897	0.0868	0.0888	0.0837
δ_K	0.8726	0.8833	0.8881	0.8963
$\tilde{\alpha}_K$	0.0756	0.0731	0.0720	0.0702

Table 7 Actual $\tilde{\alpha} = 1 - \Phi(\lambda_\alpha \delta), \delta = \sigma_1 / \sigma$, when nominal level α = 0.05 based on $T_N(F_L)$

$\tilde{\alpha}$	ARCH(1,1)		GARCH(1,1)	
	$m = n = 100$	$m = n = 300$	$m = n = 100$	$m = n = 300$
	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.0$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.2$	$\alpha_0 = 0.1, \alpha_1 = 0.2,$ $\beta_1 = 0.2$
δ_W $\tilde{\alpha}_W$	0.9263 0.0638	0.9341 0.0622	0.9275 0.0635	0.9417 0.0607
δ_{VW} $\tilde{\alpha}_{VW}$	0.8418 0.0831	0.8576 0.0792	0.8492 0.0812	0.8636 0.0777
δ_M $\tilde{\alpha}_M$	0.8378 0.0841	0.8412 0.0832	0.8380 0.0840	0.8462 0.0820
δ_K $\tilde{\alpha}_K$	0.8514 0.0807	0.8682 0.0766	0.8594 0.0787	0.8777 0.0744

Tables 5 - 7 show that the values of $\tilde{\alpha}_*$ differ from the nominal level α = 0.05 with respect to the choice of $m = n$, the parameters and distributions. It is also seen that these values decrease moderately as $m = n$ increases. This decrease is due to the asymptotics of $\hat{\theta}_{x,m}$ and $\hat{\theta}_{y,n}$. In the case of GARCH(1,1), the values for the *VW*, *K* and *W* tests in Tables 5, 6 and 7, respectively, are closer to α = 0.05 than those in the ARCH(1) case. This closeness is due to the effect of $\hat{\theta}_{x,m}$ and $\hat{\theta}_{y,n}$ in the GARCH(1,1) case, which is less severe than in the ARCH(1) case when $m = n$ increases. For the *M* test, the difference between $\tilde{\alpha}_*$ and α = 0.05 is quite substantial for all $m = n$, the parameters and distributions. Moreover, it shows the effect of skewness on the level. As is typically the case when $F = F_*$ is skewed to the right, $\tilde{\alpha}_* >$ α for the lower tail rejection region. It must be stressed that, in general, the closeness of $\tilde{\alpha}_*$ to α depends not only on the parameters but also on other aspects of $F = F_*$. The overall conclusion is that the asymptotic level of T_N is different from that of T_N^{PS} because of the effect of the GARCH specification.

4. Concluding Remarks

In this paper, the limiting Gaussian distribution of the two sample rank order statistics $\{T_N\}$ pertaining to empirical processes of the squared residuals from two independent samples of GARCH processes has been elucidated. A striking feature is that, unlike the residuals of ARMA processes, the asymptotics of $\{T_N\}$ depend on those of GARCH volatility estimators. It is well known that these results are widely used to study the asymptotic power and power efficiency of a class of two sample tests. These aspects of $\{T_N\}$, such as the asymptotic relative efficiency and GARCH volatility effect for some GARCH residual distributions have been illustrated numerically, which highlight some interesting features in the i.i.d. and in the GARCH residual settings.

It is possible in this framework to cover a broader class of distributions with nuisance prameters. Then the result would be applied to some time series processes. This subject merits further research.

5. Proof

In this section we provide the proof of Theorem 1. Write

$$J\left[\frac{N}{N+1}\hat{H}_N\right]=J[H_N]+\left(\hat{H}_N-H_N\right)J'[H_N]-\frac{\hat{H}_N}{N+1}J[H_N]$$
$$+\left\{J\left[\frac{N}{N+1}\hat{H}_N\right]-J[H_N]-\left(\frac{N}{N+1}\hat{H}_N-H_N\right)J'[H_N]\right\}$$

and $d\hat{F}_m=d\left(\hat{F}_m-F+F\right)$. Then the statistics (13) becomes

$$T_N=\mu_N+B_{1N}+B_{2N}+C_{1N}+C_{2N}+C_{3N}\,,$$

where

$$B_{1N}=\int J[H_N]d\left(\hat{F}_m-F\right)(x),$$
$$B_{2N}=\int\left(\hat{H}_N-H_N\right)J'[H_N]dF(x),$$
$$C_{1N}=\frac{-1}{N+1}\int\hat{H}_NJ'[H_N]d\hat{F}_m(x),$$
$$C_{2N}=\int\left(\hat{H}_N-H_N\right)J'[H_N]d\left(\hat{F}_m-F\right)(x),$$
$$C_{3N}=\int\left\{J\left[\frac{N}{N+1}\hat{H}_N\right]-J[H_N]-\left(\frac{N}{N+1}\hat{H}_N-H_N\right)J'[H_N]\right\}d\hat{F}_m(x),$$

To proof this theorem, we are required to show that

(i) $B_{1N}+B_{2N}$ has a limiting Gaussian distribution, and

(ii) the C_* terms are of higher order.

To begin with let us shoe statement (i). From (8) we notice that

$$B_{1N} = \int J\left[H_N\right] d\left(F_m - F\right)(x) + m^{-1/2} A_x \int J\left[H_N\right] d[xf(x)] + \text{higher order terms} \tag{15}$$

Then by partial integration of B_{2N}, and adding it to (15), we obtain

$$N^{1/2}\left(B_{1N} + B_{2N}\right) = N^{1/2}\left(1-\lambda_N\right)\left\{ \int s(x) d\left(F_m - F\right)(x) - \int s^*(x) d\left(G_n - G\right)(x) \right\}$$
$$- m^{-1/2} A_x \int x f(x) J'\left[H_N\right] dG(x) + n^{-1/2} A_y \int z g(z) J'\left[H_N\right] dF(z)$$
$$+ \text{ higher order terms.}$$
$$= a_N + b_N + c_N + d_N + \text{ higher order terms,} \quad (\text{say}), \tag{16}$$

where

$$s(x) = \int_{x_0}^{x} J'\left[H_N(y)\right] dG(y), \quad s^*(x) = \int_{x_0}^{x} J'\left[H_N(y)\right] dF(y),$$

and $\lambda_N s^*(x) + \left(1-\lambda_N\right) s(x) = J\left[H_N(x)\right] - J\left[H_N(x_0)\right]$ with x_0 determined somewhat arbitrarily, say by $H_N(x_0) = \frac{1}{2}$.

Let us compute the variance of (16). From the result by Puri and Sen (1993, p.97 99), we obtain

$$\sigma_{1N}^2 = Var\left(a_N + b_N\right) \tag{17}$$

Similarly, we can compute the same for c_N and d_N by first noting the result of TjΦstheim (1986) that

$$Var\left(m^{1/2}\left(\hat{\theta}_{x,m} - \theta_x\right)\right) = U_x^{-1} R_x U_x^{-1} \text{ and } Var\left(n^{1/2}\left(\hat{\theta}_{y,n} - \theta_y\right)\right) = U_y^{-1} R_y U_y^{-1}$$

Hence from (9), (11) and (16), we obtain

$$\sigma_{2N}^2 = Var\left(c_N\right) \text{ and } \sigma_{3N}^2 = Var\left(d_N\right) \tag{18}$$

As part of the main variance terms, we have only to evaluate

$$K_{1N} = 2E[a_N c_N] \text{ and } K_{2N} = 2E[b_N d_N],$$

since $\{X_t\}$ and $\{Y_t\}$ are independent. From (16) we obtain

$$K_{1N} = 2Nm^{-1}(1-\lambda_N)^2 \iint E[(m^{1/2}(F_m - F)(x))A_x] yf(y)J'[H_N(y)]dG(x)dG(y).$$

But by result of Horvath *et al.* (2001), it follows from (9) and (14) that

$$E\left[\left(m^{1/2}(F_m - F)(x)\right)A_x\right] = \psi_x(x)\sum_{i=0}^{p_x} \tau_{x,i}\delta_{x,i}$$

Thus,

$$K_{1N} = 2\frac{(1-\lambda_N)^2}{\lambda_N} L_{x,N}\sum_{i=0}^{p_x} \tau_{x,i}\delta_{x,i}$$

and analogously,

$$K_{2N} = 2(1-\lambda_N)L_{y,N}\sum_{i=0}^{p_y} \tau_{y,i}\delta_{y,i}$$

Adding K_{1N} and K_{2N} yields γ_N defined in Theorem 1.

Hence, using (17), (18), the term γ_N, and the central theorems given by Horvath *et al.* (2001), and Tjφstheim (1986), we may conclude that

$$N^{1/2}(B_{1N} + B_{2N})/\sigma_N \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as} \quad N \to \infty$$

Next we show statement (ii). For this purpose we need the following elementary results (see Puri and Sen (1993, p. 400)).

(E.1) $dH_N \geq \lambda_N dF \geq \lambda_0 dF$

(E.2) $F(1-F) \leq H_N(1-H_N)/\lambda_N^2 \leq H_N(1-H_N)/\lambda_0^2$

Let (α_N, β_N) be the interval S_{N_ϵ}, where

$$S_{N_\epsilon} = \{x : H_N(1-H_N) > \eta_\epsilon \lambda_0 / N\} \tag{19}$$

Then η_ϵ can be chosen independently of *F, G,* and λ_N so that

$$P\left[\varepsilon_t^2 \in S_{N_\epsilon}, t=1,\ldots,m, \quad \xi_t^2 \in S_{N_\epsilon}, t=1,\ldots,n\right] \geq 1-\epsilon$$

Let us first evaluate C_{1N}. Substituting (12) and $d\hat{F}_m = d\left(\hat{F}_m - F + F\right)$ into C_{1N} produces

$$\begin{aligned}
C_{1N} &= \frac{-1}{N+1}\int \tilde{H}_N(x)J'\left[H_N(x)\right]dF_m(x) \\
&- \frac{m^{1/2}}{N(N+1)}A_x \int xf(x)J'\left[H_N(x)\right]dF_m(x) \\
&- \frac{n^{1/2}}{N(N+1)}A_y \int xg(x)J'\left[H_N(x)\right]dF_m(x) \\
&- \frac{m^{-1/2}}{N+1}A_x \int \mathcal{H}_N(x)J'\left[H_N(x)\right]d[xf(x)] \\
&- \frac{m^{-1/2}n^{1/2}}{N(N+1)}A_xA_y \int xg(x)J'\left[H_N(x)\right]d[xf(x)] \\
&- \frac{1}{N(N+1)}A_x^2 \int xf(x)J'\left[H_N(x)\right]d[xf(x)] + \text{ higher order terms.} \\
&= \sum_{i=1}^{6} C_{1iN} + \text{ higher order terms, } \quad \text{(say).}
\end{aligned}$$

The proof of $C_{11N} = o_p\left(N^{-1/2}\right)$ is identical to that of Puri and Sen (1993, p.401). Next consider C_{12N}. From (A.3) it is seen that

$$\left|C_{12N}\right| \leq K\frac{m^{1/2}}{N}\left|A_x\right|\frac{1}{N}\int J'\left[H_N(x)\right]dF_m(x)$$

In the same way as the proof for C_{11N}, we have

$$\frac{1}{N}\int\left|J'\left[H_N(x)\right]\right|dF_m(x) = o_p\left(N^{-1/2}\right),$$

which, together with the fact $m^{1/2}N^{-1}\left|A_x\right| = O_p\left(m^{1/2}N^{-1}\right)$, implies $C_{12N} = o_p\left(N^{-1}\right)$. Similarly, we can prove that $C_{13N} = o_p\left(N^{-1}\right)$. Next we turn to C_{14N}. In what it follows, we mean that all mathematical relations, e.g., $\leq, =$ etc hold with probability $1-\epsilon$. In view of (A.2) - (A.4) and (19), it follows that

$$|C_{14N}| \le K\frac{m^{-1/2}}{cN}|A_x|\int_{S_{N_\epsilon}}\left[H_N(x)\left(1-H_N(x)\right)\right]^{-3/2+\delta} dF(x)$$
$$\le K\frac{m^{-1/2}}{cN}|A_x| O\left\{\left[H_N(\beta_N)\left(1-H_N(\beta_N)\right)\right]^{-1/2+\delta}\right\}$$
$$= O_p\left\{\frac{m^{-1/2}}{N}N^{1/2-\delta}\right\} = o_p\left(N^{-1}\right) \tag{20}$$

hence, $C_{14N} = o_p\left(N^{-1/2}\right)$. The proof for $C_{15N} = C_{16N} = o_p\left(N^{-1}\right)$ is similar to (20). Consequently, we have

$$C_{1N} = o_p\left(N^{-1/2}\right).$$

Next we consider C_{2N} By analogy with the first C term, we have

$$C_{2N} = \int[\mathcal{H}_N(x) - H_N(x)]J'[H_N(x)]d(F_m - F)(x)$$
$$+\frac{m^{1/2}}{N}A_x\int xf(x)J'[H_N(x)]d(F_m - F)(x)$$
$$+\frac{n^{1/2}}{N}A_y\int xg(x)J'[H_N(x)]d(F_m - F)(x)$$
$$+\frac{m^{-1/2}n^{1/2}}{N}A_xA_y\int xg(x)J'[H_N(x)]d[xf(x)]$$
$$+\frac{1}{N}A_x^2\int xf(x)J'[H_N(x)]d[xf(x)]$$
$$+m^{-1/2}A_x\int[\mathcal{H}_N(x) - H_N(x)]J'[H_N(x)]d[xf(x)] + \text{ higher order terms.}$$
$$= \sum_{i=1}^{6} C_{2iN} + \text{ higher order terms,} \quad \text{(say).}$$

The proof of $C_{21N} = o_p\left(N^{-1/2}\right)$ follows precisely on the same line as in Puri and Sen (1993, p.401). Next consider C_{22N}, for which, it suffices to show

$$\int_{S_{N_\epsilon}} xf(x)J'[H_N(x)]d(F_m - F)(x) = o_p(1) \tag{21}$$

From (A.3) - (A.5), (19), and Theorem 2.11.6 of Puri and Sen (1993), we see that (21) is dominated by

$$\int_{S_{N_\epsilon}} O\left\{\left[H_N(1-H_N)\right]^{\delta-1/2}\right\}\left|d(F_m-F)(x)\right|$$
$$= m^{-1/2}\int_{S_{N_\epsilon}} O\left\{N^{1/2-\delta}\right\}\left|d\left[m^{1/2}(F_m-F)(x)\right]\right| = O_p(1)$$

Therefore $C_{22N} = o_p\left(N^{-1/2}\right)$. Similarly, we can prove that $C_{23N} = o_p\left(N^{-1/2}\right)$. The proof for C_{24N} and C_{25N} is analogous to (20). To complete the assertion for C_{2N}, we evaluate C_{26N}. Let

$$I_N(\delta') = \sup_x N^{1/2}\left|\mathcal{H}_N(x) - H_N(x)\right| \le C^*\left[H_N(x)\left(1-H_N(x)\right)\right]^{1/2-\delta'} \tag{22}$$

where $\delta' > 0$, $C^* > 0$, so that $P\left(I_N(\delta')\right) \ge 1-\epsilon$ (see Puri and Sen (1993, p.401). Then from (A.2) - (A.4) and (20), it follows that

$$\left|C_{26N}\right| \le \frac{m^{-1/2}}{cN^{1/2}}\,|A_x|\int_{S_{N_\epsilon}}\left|N^{1/2}\left(\tilde{H}_N - H_N\right)J'\left[H_N\right]\right|dF(x)$$
$$\le \frac{C^* m^{-1/2}}{cN^{1/2}}\,|A_x|\,O\left\{\left[H_N(\beta_N)\left(1-H_N(\beta_N)\right)\right]^{\delta^*}\right\}$$
$$= O_p\left\{\frac{m^{-1/2}}{N^{1/2+\delta^*}}\right\} = o_p\left(N^{-1}\right),\ \delta^* > 0.$$

Hence, $C_{26N} = o_p\left(N^{-1/2}\right)$. Consequently we have

$$C_{2N} = o_p\left(N^{-1/2}\right)$$

Finally, we consider C_{3N}. Following the preceding C_{2N}, and using

$$J\left[\frac{N}{N+1}\hat{H}_N\right] = J[H_N] + \left(\frac{N}{N+1}\hat{H}_N - H_N\right)J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right],\quad 0<\varphi<1,$$

we obtain

$$C_{3N} = \int\left(\frac{N}{N+1}\mathcal{H}_N - H_N\right)\left\{J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'[H_N]\right\}dF_m(x)$$
$$+\frac{m^{1/2}}{N+1}A_x\int xf(x)\left\{J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'[H_N]\right\}dF_m(x)$$
$$+\frac{n^{1/2}}{N+1}A_y\int xg(x)\left\{J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'[H_N]\right\}dF_m(x)$$
$$+m^{-1/2}A_x\int\left(\frac{N}{N+1}\mathcal{H}_N - H_N\right)\left\{J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'[H_N]\right\}d[xf(x)]$$
$$+\frac{m^{-1/2}n^{1/2}}{N+1}A_xA_y\int xg(x)\left\{J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'[H_N]\right\}d[xf(x)]$$
$$+\frac{1}{N+1}A_x^2\int xf(x)\left\{J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'[H_N]\right\}d[xf(x)]$$
$$+ \text{ higher order terms.}$$
$$= \sum_{i=1}^{6} C_{3iN} + \text{ higher order terms,} \quad \text{(say).}$$

Let us first consider C_{31N}. From (12) we first note that

$$\left[\varphi H_N(x) + (1-\varphi)\frac{N}{N+1}\hat{H}_N(x)\right]\Big/H_N(x) = \varphi + (1-\varphi)\frac{N}{N+1}\frac{\mathcal{H}_N(x)}{H_N(x)}$$
$$+(1-\varphi)(N+1)^{-1}\left[m^{1/2}A_x xf(x) + n^{1/2}A_y xg(x)\right]\Big/H_N(x)$$
$$+\text{higher order terms} \quad (23)$$

From (23) we can write

$$\mathcal{H}_N(x) = H_N(x) + N^{-1/2}O(1) \quad (24)$$

with probability $\geq 1-\in$ where $O(1)$ is uniform in x. Hence by (A.3), (A.4) and (24), we obtain

$$(23) = 1 + O\left(N^{-1/2}\right)$$

with probability $\geq 1-\epsilon$. Similarly we can prove

$$\left\{1-\left[\varphi H_N(x)+(1-\varphi)\frac{N}{N+1}\hat{H}_N(x)\right]\right\}\times\left[1-H_N(x)\right]^{-1} = 1+O\left(N^{-1/2}\right)$$

with probability $\geq 1-\epsilon$. Thus, for sufficiently large $N>0$, we can find $\beta>0$ such that

$$\inf_x\left[\varphi H_N(x)+(1-\varphi)\frac{N}{N+1}\hat{H}_N(x)\right]\left[1-\left\{\varphi H_N(x)+(1-\varphi)\frac{N}{N+1}\hat{H}_N(x)\right\}\right]$$
$$\times\left[H_N(x)\left(1-H_N(x)\right)\right]^{-1}>\beta$$

with probability $\geq 1-\epsilon$. From the preceding arguments, observe that

$$\begin{aligned}|C_{31N}| &\leq \frac{1}{N^{1/2}}\int N^{1/2}\left|\frac{N}{N+1}\mathcal{H}_N-H_N\right|\left|J'\left[\varphi H_N+(1-\varphi)\frac{N}{N+1}\hat{H}_N\right]-J'\left[H_N(x)\right]\right|dF_m(x)\\ &=\frac{1}{N^{1/2}}\int I_N(x)dF_m(x), \qquad \text{(say).}\end{aligned} \tag{25}$$

It is easy to show

$$\int_{S_{N\epsilon}} I_N(x)dF_m(x)\leq C^*K\left[1+\beta^{\delta-3/2}\right]\int_{S_{N\epsilon}}\left[H_N(1-H_N)\right]^{\delta^*-1}dF_m(x) \tag{26}$$

and

$$E\int_{S_{N\epsilon}} I_N(x)dF_m(x)\leq C^*K\left[1+\beta^{\delta-3/2}\right]\int_{S_{N\epsilon}}\left[H_N\left(1-H_N\right)\right]^{\delta^*-1}dF(x) \tag{27}$$

Hence, $I_N(x)$ is integrable. Recalling (22), it is seen that $I_N(x)\to 0$ in probability. By the dominated convergence theorem, (25) - (27), we get $C_{31N}=o_p\left(N^{-1/2}\right)$. Next, consider C_{32N} Using the arguments of C_{31N}, (A.3) and (A.5), we obtain

$$\left|C_{32N}\right| \leq \frac{m^{1/2}}{N}\left|A_x\right| \int_{S_{N\epsilon}} |xf(x)| \left|J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'\left[H_N(x)\right]\right| dF_m(x)$$

$$\leq K\frac{m^{1/2}}{N}\left|A_x\right|\left[1+\beta^{\delta-3/2}\right]\int_{S_{N\epsilon}}\left[H_N\left(1-H_N\right)\right]^{\delta-1/2} dF_m(x) \tag{28}$$

with

$$E\int_{S_{N\epsilon}}\left[H_N\left(1-H_N\right)\right]^{\delta-1/2} dF_m(x) = \int_{S_{N\epsilon}}\left[H_N\left(1-H_N\right)\right]^{\delta-1/2} dF(x) \tag{29}$$

Then, by the dominated convergence theorem, $\left(m^{1/2}N^{-1} \mid A_x \mid\right) = O_p\left(m^{1/2}N^{-1}\right)$, (28) and (29), we have $C_{32N} = o_p\left(N^{-1/2}\right)$. Similarly, we can show $C_{33N} = o_p\left(N^{-1/2}\right)$. Next, we turn to C_{34N}. Following the arguments of C_{31N}, and using (A.2) - (A.4), we obtain

$$\left|C_{34N}\right| \leq \frac{m^{-1/2}}{cN^{1/2}}\left|A_x\right| \int N^{1/2}\left|\frac{N}{N+1}\tilde{H}_N - H_N\right|\left|J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'\left[H_N(x)\right]\right| dF(x)$$

$$\leq K\frac{m^{-1/2}}{cN^{1/2}}\left|A_x\right|\left[1+\beta^{\delta-3/2}\right]\int_{S_{N\epsilon}}\left[H_N\left(1-H_N\right)\right]^{\delta^*-1} dF(x).$$

Then, by the dominated convergence theorem, similarly as in C_{31N}, we can prove $C_{34N} = o_p\left(N^{-1/2}\right)$. Next, consider C_{35N}. From (A.2) - (A.4) and (28), we observe that

$$\left|C_{35N}\right| \leq K\frac{m^{-1/2}n^{-1/2}}{cN}\left|A_x\right|\left|A_y\right| \int H_N\left(1-H_N\right)\left|J'\left[\varphi H_N + (1-\varphi)\frac{N}{N+1}\hat{H}_N\right] - J'\left[H_N(x)\right]\right| dF(x)$$

$$\leq O_P\left\{m^{-1/2}n^{-1/2}N^{-1}\right\}\int_{S_{N\epsilon}}\left[H_N\left(1-H_N\right)\right]^{\delta-1/2} dF(x).$$

Hence, we have $C_{35N} = o_p\left(N^{-1/2}\right)$. Similarly we can show $C_{36N} = o_p\left(N^{-1/2}\right)$. Consequently, we have

$$C_{3N} = o_p\left(N^{-1/2}\right)$$

This completes the proof of the theorem.

Acknowledgement

The authors express their gratitude to the referee for his/her comments, which improved the original version of this paper. We also wish to thank Professor Sangyeol Lee for his comments.

References

[1] Berkes, I. and Horváth, L. (2003a). Limit results for the empirical process of squared residuals in GARCH models. *Stoch. Processes Appl.* 105, 271 298.

[2] Berkes, I., Horváth, L. and Kokoszka, P. (2003b). GARCH processes: structure and estimation. *Bernoulli* 9, 201 227.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307 328.

[4] Bougerol, P. and Picard, N. (1992a). Strict stationarity of generalized autoregressive processes. *Annals of Probability* 20, 1714 1730.

[5] Bougerol, P. and Picard, N. (1992b). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115 127.

[6] Chen, M. and An, H. Z. (1998). A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica* 8, 505 510.

[7] Chernoff, H. and Savage, I. R. (1958). Asymptotic normality and effciency of certain nonparametric test statistics. *Ann. Math. Statist.* 29, 972 994.

[8] Comte, F. and Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *J. Multi. Anal.* 84, 61 84.

[9] Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50, 987 1007.

[10] Francq, C. and Zakoian, J.M. (2004). Maximum likelihood estimation of pure GARCH and ARMA GARCH processes. *Bernoulli* 10, No.4, 605 637.

[11] Gouriéroux, C. (1997). *ARCH Models and Financial Applications*. New York: Springer.

[12] Lee, S. and Taniguchi, M. (2005). Asymptotic theory for ARCH models: LAN and residual empirical processes. *Statist. Sinica* 22, 215 234.

[13] Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, 318 334.

[14] Horváth, L., Kokoszka, P. and Teyssiere, G. (2001). Empirical process of the squared residulas of an ARCH sequence. *Ann. Statist.* 29, 445 469.

[15] Puri, M. L. and Sen, P. K. (1993). *Nonparametric Methods in Multivariate Analysis.Wiley*: New York.

[Received March 2005, Accepted July 2005]