Deo, Rohit S.; Chen, Willa W.

Working Paper

The Variance Ratio Statistic at Large Horizons

Papers / Humboldt-Universität Berlin, Center for Applied Statistics and Economics (CASE), No. 2004,04

Provided in cooperation with:
Humboldt-Universität Berlin

Suggested citation: Deo, Rohit S.; Chen, Willa W. (2003) : The Variance Ratio Statistic at Large Horizons, Papers / Humboldt-Universität Berlin, Center for Applied Statistics and Economics (CASE), No. 2004,04, http://hdl.handle.net/10419/22178

The Variance Ratio Statistic at Large Horizons

Willa W. Chen and Rohit S. Deo*
Texas A&M University and New York University

November 13, 2003

Abstract: We make three contributions to using the variance ratio statistic at large horizons. Allowing for general heteroscedasticity in the data, we obtain the asymptotic distribution of the statistic when the horizon k is increasing with the sample size n but at a slower rate so that $k/n \to 0$. The test is shown to be consistent against a variety of relevant mean reverting alternatives when $k/n \to 0$. This is in contrast to the case when $k/n \to \delta > 0$, where the statistic has been recently shown to be inconsistent against such alternatives. Secondly, we provide and justify a simple power transformation of the statistic which yields almost perfectly normally distributed statistics in finite samples, solving the well known right skewness problem. Thirdly, we provide a more powerful way of pooling information from different horizons to test for mean reverting alternatives. Monte Carlo simulations illustrate the theoretical improvements provided.

JEL classification: C12, C22

Key words: Mean reversion; Frequency domain; Power transformation

*Corresponding author. rdeo@stern.nyu.edu, Tel: (212) 998-0469, Stern School of Business, New York University, 8-57 KMEC, 44 W. 4'th Street, New York NY 10012, USA.

1 Introduction

The variance ratio (VR) statistic is one of the popular tests that has been employed in the literature to test the random walk hypothesis for financial and economic data. The statistic is obtained as the sample variance of k-period differences, $x_t - x_{t-k}$, of the time series x_t, divided by k times the sample variance of the first difference, $x_t - x_{t-1}$, for some integer k. The VR statistic has been found by several authors (see, for example, Faust (1992)) to be particularly powerful when testing against mean reverting alternatives to the random walk model, particularly when k is large. However, the practical use of the statistic has been impeded by the fact that the asymptotic theory provides a poor approximation to the small sample distribution of the VR statistic. More specifically, rather than being normally distributed as the theory states, the statistics are severely biased and right skewed for large k, (see Lo and MacKinlay, 1989) which makes application of the statistic problematic. To circumvent this problem, Richardson and Stock (1989) derived the asymptotic distribution of the VR statistic under the random walk null, assuming that both k and n increase to infinity but in such a way that k/n converges to a positive constant δ which is strictly less than 1. They showed that the VR statistic, without any normalization, converges to a functional of Brownian motion. Through Monte Carlo simulations, they demonstrated that this new distribution provides a far more robust approximation to the small sample distribution of the VR statistic. However, Deo and Richardson (2003) have recently shown that the VR statistic is inconsistent against an important class of mean reverting alternatives under this framework. Thus, though the VR statistic would have vastly improved size properties under the null hypothesis of a random walk if k were chosen to be a fraction of the sample size n, it would fail to detect such alternatives with probability approaching 1 as the sample size increased. Currently there is no proposal in the literature which provides a way of using the VR statistic without compromising either its finite sample size properties or its large sample power properties.

With this backdrop, we provide several contributions to the literature. First, it is intuitively appealing to maintain the assumption that the multiperiod horizon k is large, not least because longer horizons have a better chance of capturing mean reversion in the series. Thus, under general conditions which allow for conditional heteroscedasticity in the innovations, we study the limiting behaviour of the VR statistic for large k but now under the restriction that $k/n \to 0$. Specifically, we show that when $k \to \infty$, $n \to \infty$ but $k/n \to 0$, then under the null of a random walk, the VR statistic is asymptotically normal with a mean of 1. The requirement that k is large is important since, as stated above, previous authors have shown that large values of k are to be preferred when testing for mean reversion. Furthermore, we prove that under this alternative distribution theory, the test is consistent, in that the probability of it detecting a wide variety of mean reversion alternatives approaches one as the sample size n increases.

Unfortunately, this new distribution does not solve the well documented skewness problem of the VR statistic's sampling distribution. The second contribution of this paper is to propose a method which is shown to improve the asymptotic normal approximation to the distribution of the statistic by an order of magnitude in finite samples, via a simple power transformation of the VR statistic. Monte Carlo simulations confirm the theoretical assertion of the vast improvement of the normal approximation afforded by the power transformation. Our Monte Carlo simulations also show that this improvement in the normal approximation leads to significant gains in power against mean reverting alternatives.

The third contribution of this paper is to implement a new joint test which uses VR statistics computed at different differencing periods to test the random walk null hypothesis. The joint test statistic which has been studied so far in the literature is the Wald type chi-square test statistic which jointly tests whether a sequence of population variance ratios at several differencing periods all equal 1. However, this test is blind to the inherent one sided nature of a mean reverting alternative hypothesis, since under such an alternative all the population variance ratios should be less than 1. See Lo and MacKinlay(1989). In this paper, we adapt a test procedure proposed by Follmann (1996) for testing against one sided alternatives for the mean vector of a multivariate normal distribution. Our Monte Carlo simulations show that this adapted test in combination with the power transformation results in significant power gains over the usual chi-square test when testing for mean reverting alternatives, while retaining the appropriate size.

The paper is organized as follows. In section 2, we define the VR statistic and provide its asymptotic distribution under conditional heteroscedasticity for large k such that $k^{-1}+k/n \to 0$. We also demonstrate in that section that in this framework the VR statistic is consistent against a wide range of alternatives. In section 3, we provide an alternative equivalent representation of the VR statistic which motivates the power transformation that provides a better approximation to the normal distribution. A new joint test which combines information from several differencing periods and is useful against one sided alternatives is also introduced. Section 4 presents Monte Carlo results for the various statistics that we have proposed under two different null hypotheses and three alternative hypotheses. All technical proofs are relegated to the Appendix.

2 Asymptotic Theory for the Variance Ratio Statistic

Given $n+1$ observations $x_0, x_1, ..., x_n$ of a time series, the variance ratio statistic with a positive integer $k(< n)$ as differencing period is defined as

$$VR\left(k\right) = \hat{\sigma}_b^2\left(k\right)/\hat{\sigma}_a^2,$$

where

$$\hat{\sigma}_b^2(k) = \frac{n}{k(n-k+1)(n-k)} \sum_{t=k}^{n} (x_t - x_{t-k} - k\hat{\mu})^2 ,$$

$$\hat{\sigma}_a^2 = \frac{1}{n-1} \sum_{t=1}^{n} (x_t - x_{t-1} - \hat{\mu})^2$$

and

$$\hat{\mu} = n^{-1} \sum_{t=1}^{n} (x_t - x_{t-1}) .$$

In the usual fixed k asymptotic treatment, under the null hypothesis that the $\{x_t\}$ follow a random walk with possible drift, given by

$$x_t = \mu + x_{t-1} + \varepsilon_t \tag{1}$$

where μ is a real number and $\{\varepsilon_t\}$ is a sequence of zero mean independent random variables, it is possible to show (see, for example, Lo and Mackinlay (1988)) that

$$\sqrt{n}\,(VR\,(k) - 1) \xrightarrow{D} N\left(0,\ \sigma_k^2\right) ,$$

where σ_k^2 is some simple function of k. This result extends to the case where the $\{\varepsilon_t\}$ are a martingale difference series with conditional heteroscedasticity (see, for example, Campbell, Lo and MacKinlay 1997), though the variance σ_k^2 has to be adjusted to account for the conditional heteroscedasticity. However, the asymptotic behaviour of the variance ratio statistic for large values of k, such that $k^{-1} + k/n \to 0$, is not known when the innovations ε_t are conditionally heteroscedastic. In this section, we provide precisely this asymptotic distribution, in obtaining which the following assumptions on the series of innovations $\{\varepsilon_t\}$ are made:

(A1) $\{\varepsilon_t\}$ is ergodic and $E\left(\varepsilon_t | \mathfrak{F}_{t-1}\right) = 0$ for all t, where $\mathfrak{F}_t$ is a sigma field, ε_t is $\mathfrak{F}_t$ measurable and $\mathfrak{F}_{t-1} \subset \mathfrak{F}_t$ for all t.

(A2) $E\left(\varepsilon_t^2\right) = \sigma^2 < \infty$.

(A3) For any integer q, $2 \leq q \leq 8$, and for q non-negative integers s_i, $E\left(\Pi_{i=1}^{q} \varepsilon_{t_i}^{s_i}\right) = 0$ when at least one s_i is exactly one and $\sum_{i=1}^{q} s_i \leq 8$.

(A4) For any integer r, $2 \leq r \leq 4$, and for r non-negative integers s_i, $E\left(\Pi_{i=1}^{r} \varepsilon_{t_i}^{s_i} \, |\mathfrak{F}_t\right) = 0$ when at least one s_i is exactly one and $\sum_{i=1}^{r} s_i \leq 4$, for all $t < t_i$, $i = 1, 2, 3, 4$.

(A5) $\lim\limits_{n\to\infty} Var\left[E\left(\varepsilon_{t+n}^2 \varepsilon_{t+n+j}^2 |\mathfrak{F}_t\right)\right] = 0$ uniformly in j for every $j > 0$.

(A6) $\lim\limits_{n\to\infty} E\left(\varepsilon_t^2 \varepsilon_{t-n}^2\right) = \sigma^4$.

Conditions (A1) - (A6) allow the innovations ε_t to be a martingale difference sequence with conditional heteroscedasticity. As a matter of fact, lemmas 1 and 2 below show that the stochastic volatility model (see Shephard 1996) and the GARCH model (Bollerslev 1986), which are two of the most popular models in the literature for conditional heteroscedastic martingale differences, satisfy conditions (A1) - (A6). Conditions (A3) - (A4) state that the series $\{\varepsilon_t\}$ shows product moment behaviour similar to that of an independent white noise process. Conditions (A5) - (A6) state that ε_t and ε_{t-n} are roughly independent for large lags n.

The following two lemmas assert that two major models of conditionally heteroscedastic martingale differences, viz. the stochastic volatility model and the generalized autoregressive conditionally heteroscedastic (GARCH) model, satisfy the assumptions (A1)-(A6). The proofs of the lemmas are in the Technical Appendix at the end.

Lemma 1 *Let the series $\{\varepsilon_t\}$ be generated by the stochastic volatility model*

$$\varepsilon_t = v_t \exp\left(h_t\right), \tag{2}$$

where $\{v_t\}$ is an independent $\left(0, \sigma_v^2\right)$ stationary series, $\{h_t\}$ is a stationary zero mean Gaussian series and $\{v_t\}$ and $\{h_t\}$ are independent. Assume that $E\left(v_t^8\right) < \infty$. Then $\{\varepsilon_t\}$ satisfies the *assumptions (A1)-(A6).*

See Shephard (1996) for a discussion of the model (2) and its applications.

Our next lemma asserts that under some conditions the GARCH(1,1) family of models also satisfies Condition A. We have restricted attention to the GARCH(1,1) case for simplicity of exposition. We conjecture that conditions (A1) - (A6) will continue to hold for a general GARCH(p,q) model, the proof following along similar lines by referring to the work of Bougerol and Picard (1992).

Lemma 2 *Let the series* $\{\varepsilon_t\}$ *be a GARCH(1,1) process given by*

$$\varepsilon_t = \sigma_t v_t, \tag{3}$$

where $\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2$ *and* $\{v_t\}$ *is a sequence of independent standard normal variables. Let* $\omega > 0$, $\beta \geq 0$ *and* $\alpha > 0$. *Furthermore, let* α *and* β *be such that* $E\left\{\log_e\left(\beta + \alpha v_t^2\right)\right\} < 0$ *and* $E\left\{\left(\beta + \alpha v_t^2\right)^4\right\} < 1$. *Then* $\{\varepsilon_t\}$ *satisfies the assumptions (A1)-(A6).*

The condition $E\left\{\log_e\left(\beta + \alpha v_t^2\right)\right\} < 0$ in Lemma 2 is satisfied by any pair (α, β) in the set $S = \{(\alpha, \beta) : \alpha + \beta < 1\}$ (See Nelson, 1990) while the condition $E\left\{\left(\beta + \alpha v_t^2\right)^4\right\} < 1$ will be satisfied by a non-empty subset of S (see Bollerslev, 1986).

We now state our result on the limiting distribution of the VR statistic in the following theorem.

Theorem 3 *Let the series* $\{x_t\}$ *satisfy equation (1) and assume that conditions (A1)-(A6) hold. For a fixed positive integer* s, *let* $k_1 < k_2 < ... < k_s < n$ *be positive integers such that* $k_1 \to \infty$,

$k_s n^{-1} \to 0$ and $k_i k_j^{-1} \to a_{ij}$ for $1 \leq i \leq j \leq s$. Let $\mathbf{D}_n$ be an $s \times s$ diagonal matrix with diagonal elements $d_{ii} = \sqrt{n/k_i}$ for $i = 1, 2, ..., s$. Then

$$\mathbf{V}_n \overset{D}{\approx} N(\mathbf{1},\, \mathbf{D}_n^{-1}\mathbf{\Sigma}\mathbf{D}_n^{-1}),$$

where $\mathbf{V}_n = (VR\,(k_1)\,, VR\,(k_2)\,, ..., VR\,(k_s))'$, $\mathbf{1}$ is a $s \times 1$ vector of ones and $\mathbf{\Sigma} = (\sigma_{ij})$ is an $s \times s$ matrix such that $\sigma_{ij} = \sigma^4 4 a_{ij}^{1/2}\,(3 - a_{ij})\,/6$.

Note that the limiting distribution of the VR statistic is free of nuisance parameters and is identical to that obtained when the ε_t are assumed to be independent. See Theorem 9.4.1 of Anderson (1994). Furthermore, the VR statistics computed at different differencing periods k_i, are asymptotically independent when $k_i k_j^{-1} \to 0$ for $i < j$. Both of these results are in contrast to those obtained when the differencing periods are fixed and not allowed to increase to infinity with the sample size. See Lo and MacKinlay (1989). It is interesting to note that the limiting distribution of the VR statistic is free of nuisance parameters depending on higher moments which might arise due to conditional heteroscedasticity. This is quite different from the behaviour of other tests of the random walk hypothesis in the presence of conditional heteroscedasticity. See Deo (2000).

We have established the asymptotic distribution of the VR statistic under the null hypothesis of a random walk with conditional heteroscedasticity when $k \to \infty$, $n \to \infty$ and $k/n \to 0$. The next theorem states that under this framework, the VR statistic also provides a consistent test against a large class of mean reverting alternatives.

Theorem 4 *Let $\{e_t\}$ and $\{u_t\}$ be two series of zero mean independent processes with finite fourth moments and which are independent of each other. Define the processes $\{y_t\}$ and $\{z_t\}$ by $y_t = \sum_{j=0}^{\infty} a_j u_{t-j}$ and $z_t = \sum_{j=0}^{\infty} b_j e_{t-j}$, where $|a_j| \leq C\lambda^j$ and $|b_j| \leq C\lambda^j$ for some constant C*

and $0 < \lambda < 1$. *Let* $r_t = \mu + r_{t-1} + z_t$ *and* $x_t = r_t + y_t$. *If* $k \to \infty$, $n \to \infty$ *and* $k/n \to 0$, *then*

$$VR(k) \xrightarrow{P} \frac{\sigma_z^2 + 2\sum_{j=1}^{\infty} \gamma_z(j)}{\sigma_z^2 + 2\sigma_y^2 - 2\gamma_y(1)}.,$$

where σ_z^2 *and* σ_y^2 *are the variances of* z_t *and* y_t *respectively, while* $\gamma_z(j)$ *and* $\gamma_y(j)$ *are the respective autocovariances at lag* j.

Theorem 4 shows that the power properties of the VR statistic under the $k/n \to 0$ framework are markedly different from those when $k/n \to \delta > 0$, in which case Deo and Richardson (2003) have shown the VR statistic to be inconsistent against the alternatives considered in Theorem 4.

Though the VR statistic has an asymptotic normal distribution when $k/n \to 0$, it is obvious that in finite samples the normal distribution may not provide a good approximation since the statistic is a quadratic form and hence must be right skewed. A common method which has a long history in Statistics to reduce skewness and induce normality in such random variables is to consider power transformations. The obvious question, naturally, is which power one should use and we address this question for the VR statistic in the next section

3 Power Transformations of the Variance Ratio Statistic

In attempting to address the skewness of the finite sample distribution of the VR statistic, it helps to express the VR statistic in an alternative form, which lends more insight into how the normal distribution approximation can be improved. Inspection of the proof of Theorem 3 in the Appendix shows that

$$VR(k) = \hat{\sigma}^{-2} \sum_{|j| \leq k} (1 - |j|/k)\, \hat{\gamma}_j + o_p\left(\sqrt{k/n}\right), \tag{4}$$

where $\hat{\gamma}_j = \hat{\gamma}_{-j} = n^{-1} \sum_{t=j+1}^{n} \varepsilon_t \varepsilon_{t-j}$ for $j \geq 0$ and

$$\hat{\sigma}^2 = (n-1)^{-1} \sum_{t=1}^{n} (\varepsilon_t - \bar{\varepsilon})^2 = (n-1)^{-1} \sum_{t=1}^{n} (x_t - x_{t-1} - \hat{\mu})^2 .$$

Now, using the fact that

$$\hat{\gamma}_j = \int_0^{2\pi} I(\lambda) \exp(-ij\lambda) \, d\lambda,$$

where $I(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^{n} \varepsilon_t \exp(-it\lambda) \right|^2$ is the periodogram, we get from (4),

$$VR(k) = \hat{\sigma}^{-2} \int_0^{2\pi} W_k(\lambda) I(\lambda) \, d\lambda + o_p\left(\sqrt{k/n}\right), \tag{5}$$

where

$$W_k(\lambda) = \sum_{|j| \leq k} (1 - |j|/k) \exp(-ij\lambda) = k^{-1} \left\{ \frac{\sin(k\lambda/2)}{\sin(\lambda/2)} \right\}^2 .$$

As shown in part (i) of Lemma 7 in the Technical Appendix below, the integral in (5) can be approximated by a discrete sum over the Fourier frequencies $\lambda_j = 2\pi j/n$ with error $o_p\left(\sqrt{k/n}\right)$ and hence we get

$$VR(k) = \frac{4\pi}{n\hat{\sigma}^2} \sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j) I(\lambda_j) + o_p\left(\sqrt{k/n}\right). \tag{6}$$

The behaviour of $VR(k)$ is thus dictated by the behaviour of the periodogram values $I(\lambda_j)$ at the Fourier frequencies. If the ε_t series is Gaussian, then it is well known (Brockwell and Davis, 1996) that the variables $2\pi I(\lambda_j)/\sigma^2$ are exactly independent identically distributed standard exponential random variables for all sample sizes. This behaviour of the variables $2\pi I(\lambda_j)/\sigma^2$ can be shown to continue to hold asymptotically if the ε_t are a martingale difference sequence with finite fourth moment, by applying the Central Limit Theorem for martingale differences to $n^{-1/2} \sum_{t=1}^{n} \varepsilon_t \exp(-i\lambda_j t)$. These observations in conjunction with (6) and the fact that $\hat{\sigma}^2/\sigma^2 = 1 + O_p\left(n^{-1/2}\right)$ imply that, in general, we may think of the VR statistic as being of the form

$$VR(k) = \frac{2}{n} \sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j) V_j + o_p\left(\sqrt{k/n}\right), \tag{7}$$

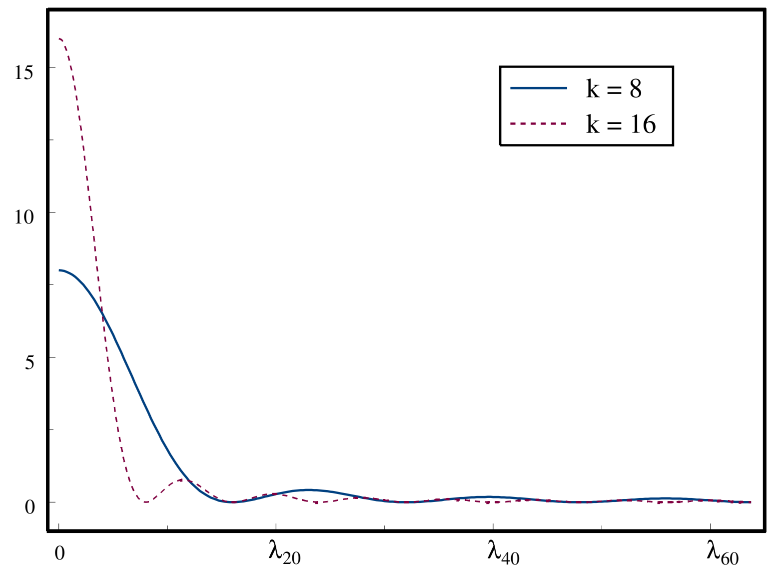

Figure 1: $W_k(\lambda)$ for $n = 128$ and $k = 8$ and 16.

where the V_j are independent standard exponential random variables. As we next show, this approximate expression for the VR statistic as a weighted linear combination of independent standard exponential random variables helps us both to understand why the normal distribution provides a bad approximation for large k as well as to obtain an appropriate power transformation which improves the normal approximation.

It is known (see, for example, page 509 of Anderson 1994) that $W_k(\lambda)$ has a peak at the origin and then damps down to zero for values of λ further from the origin. Furthermore, the larger k is, the more quickly $W_k(\lambda)$ damps down to zero, which can be seen in Figure 1, where we plot $W_k(\lambda)$ for $n = 128$ and $k = 8$ and 16. Thus, for large values of k, we see from (7) that $VR(k)$ will essentially be a sum of too few independent standard exponential random variables for the central limit theorem to properly take effect, resulting in right skewed distributions. However, Chen and Deo (2003) have recently shown that power transformations may be gainfully applied to random variables which have approximate linear representations of the form in (7), yielding much better normal approximations. Using their results (See equation 9 of Chen and Deo, 2003),

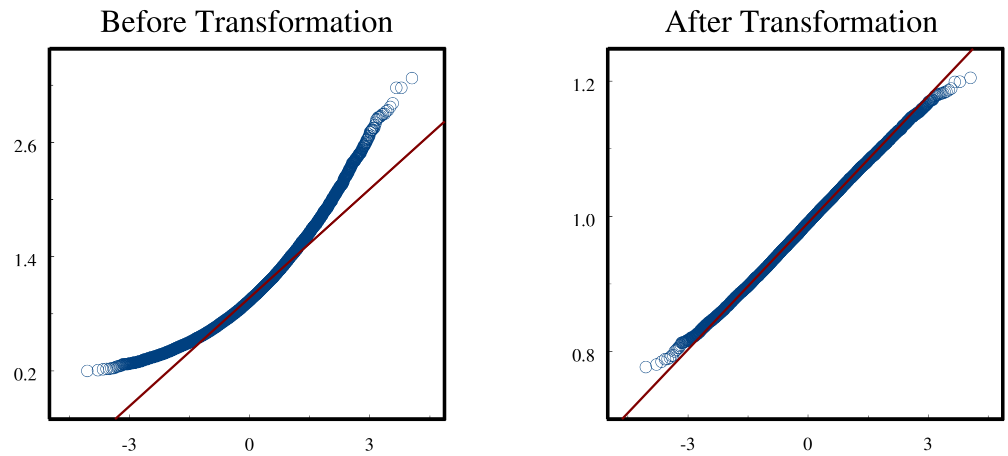

Figure 2: QQ plots of $VR(k)$ and $VR^{\beta}(k)$ on 20,000 replications with $n = 128$, $k = 16$ and $\varepsilon_t \sim N(0,1)$.

it follows that if one sets

$$\beta = 1 - \frac{2}{3}\frac{\left(\sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j)\right)\left(\sum_{j=1}^{[(n-1)/2]} W_k^3(\lambda_j)\right)}{\left(\sum_{j=1}^{[(n-1)/2]} W_k^2(\lambda_j)\right)^2}, \tag{8}$$

then the Gaussian distribution provides a better approximation to the distribution of $VR^{\beta}(k)$ than to that of $VR(k)$. A dramatic visual display of this improvement is shown in Figure 2 The plot on the left is a QQ plot of 20000 replications of the $VR(k)$ statistic, based on a sample size of $n = 128$ and k =16**,** where the ε_t are i.i.d. standard normal. The extreme curvature is indicative of the right skewness of the distribution of $VR(k)$. The plot on the right is a QQ plot of $VR^{\beta}(k)$, where β was computed using (8). The plot now shows a straight line as would be expected for observations from a normal distribution. The power transformation thus provides a very simple method of getting almost near perfect normality for the finite sample distribution of the VR statistic. A standard Taylor series argument applied to the result of Theorem 3 yields the asymptotic distribution of $VR^{\beta}(k)$ which can then be used for inference. However, we feel that since the power transformation is motivated by the representation (6), it might be preferable to re-define the VR statistic as well as its power transformation directly in terms of the leading term of that expression, thus avoiding any effects of the remainder term on its finite sample distribution. Towards that end, we now define the VR statistic based on the periodogram, for differencing period k, as

$$VR_p(k) = \frac{1}{(1-k/n)} \frac{4\pi}{n\hat{\sigma}^2} \sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j) I_{\triangle X}(\lambda_j), \tag{9}$$

where $I_{\triangle X}(\lambda_j) = (2\pi n)^{-1} \left|\sum_{t=1}^{n} (x_t - x_{t-1} - \hat{\mu}) \exp(-i\lambda_j t)\right|^2$. Since the periodogram is shift invariant at non-zero Fourier frequencies, we have $I_{\triangle X}(\lambda_j) = I(\lambda_j)$ and hence the $VR_p(k)$ statistic as defined in (9) based on the observed data $x_t - x_{t-1} - \hat{\mu}$ is identical to the first term in (6), which is based on the unobserved ε_t. It should be noted that this expression for the VR statistic, apart from the normalisation of $(1-k/n)^{-1}$ which is just a finite sample correction ensuring a unit mean, is precisely the normalised discrete periodogram average estimate of the spectral density of a stationary process at the origin and has a long tradition in time series analysis. See Brockwell and Davis, 1991. From (6) it follows that $VR_p(k)$ will have the same asymptotic distribution as that of $VR(k)$ given in Theorem (3) and hence, by the usual Taylor series argument, the asymptotic distribution of $VR_p^{\beta}(k)$ may be obtained. It is however preferable to have an expression for the variance of $VR_p(k)$, and thus for that of $VR_p^{\beta}(k)$, which is accurate in finite samples and accounts for the finite sample effects of conditional heteroscedasticity. Towards this end, we first define the quantities $C_{n,k} = n(n-k)^{-1}$ and

$$\hat{\tau}_j = \hat{\sigma}^{-4} (n-j-4)^{-1} \sum_{t=j+1}^{n} (x_t - x_{t-1} - \hat{\mu})^2 (x_{t-j} - x_{t-j-1} - \hat{\mu})^2,$$

where $\hat{\tau}_j$ is an estimator of $\sigma^{-4} E\left(\varepsilon_t^2 \varepsilon_{t-j}^2\right)$. In part (ii) of Lemma 7, we show that the finite sample variance covariance matrix of $\mathbf{V}_p = (VR_p(k_1), VR_p(k_2), ..., VR_p(k_s))'$ with remainder terms of order $o\left(k_s^2/n^2\right)$ is consistently estimated by

$$\hat{\mathbf{\Sigma}} = \mathbf{L}' \begin{bmatrix} \hat{\mathbf{A}} & \hat{\mathbf{b}} \\ \hat{\mathbf{b}}' & \hat{d} \end{bmatrix} \mathbf{L}, \tag{10}$$

where $\mathbf{L} = (l_{k_1}, ..., l_{k_s})$,

$$l'_{k_i} = \left(2C_{n,k_i}\left(1 - 1/k_i\right), ..., 2C_{n,k_i}\left(1 - \left(k_i - 1\right)/k_i\right),\ \overbrace{0, ..., 0,}^{k_s - k_i \text{ terms}}\ - \left(k_i C_{n,k_i} - n/\left(n-1\right)\right) \right), \tag{11}$$

$$\hat{\mathbf{A}} = diag\left(\frac{n-j}{n^2}\hat{\tau}_j + \frac{j}{n^2}\right) \qquad j = 1, ..., k_s,$$

$\hat{\mathbf{b}}$ is a $k_s \times 1$ vector such that its j^{th} element is given by $\left(2\left(n-j\right)n^{-3}\hat{\tau}_j + 2jn^{-3}\right)$ and $\hat{d}$=$2n^{-2}$. We are now in a position to state the following Theorem.

Theorem 5 *Let the series $\{x_t\}$ satisfy equation (1) and assume that conditions (A1)-(A6) hold. For a fixed positive integer s, let $k_1 < k_2 < ... < k_s < n$ be positive integers such that $k_1 \to \infty$, $k_s n^{-1} \to 0$ and $k_i k_j^{-1} \to a_{ij}$ for $1 \leq i \leq j \leq s$. For each k_i, let β_i be given by (8) and define $\mathbf{V}_{p,\beta} = \left(VR_p^{\beta_1}\left(k_1\right), VR_p^{\beta_2}\left(k_2\right), ..., VR_p^{\beta_s}\left(k_s\right)\right)'$. Then*

$$\mathbf{V}_{p,\beta} \overset{D}{\approx} N(\mu_\beta, \mathbf{\Sigma}_\beta),$$

where the $(i,j)^{th}$ element of $\mathbf{\Sigma}_\beta$ is

$$\beta_i \beta_j \hat{\sigma}_{i,j}$$

and the i^{th} element of μ_β is

$$1 + 0.5\beta_i\left(\beta_i - 1\right)\hat{\sigma}_{i,i},$$

where $\hat{\sigma}_{i,j}$ is the $(i,j)^{th}$ entry of $\hat{\mathbf{\Sigma}}$ given in (10).

It is trivially seen that both $VR_p \xrightarrow{P} 1$ and $VR_p^{\beta} \xrightarrow{P} 1$ under conditions (A1)-(A6). Our next Theorem shows that both VR_p as well as VR_p^{β} also retain the consistency of the VR statistic with regard to detecting the alternative hypotheses assumed in Theorem 4.

Theorem 6 *Let the assumptions of Theorem 4 hold. Then*

$$VR_p(k) \xrightarrow{P} \frac{\sigma_z^2 + 2\sum_{j=1}^{\infty}\gamma_z(j)}{\sigma_z^2 + 2\sigma_y^2 - 2\gamma_y(1)},$$

and

$$VR_p^{\beta}(k) \xrightarrow{P} \left(\frac{\sigma_z^2 + 2\sum_{j=1}^{\infty}\gamma_z(j)}{\sigma_z^2 + 2\sigma_y^2 - 2\gamma_y(1)}\right)^{\beta},$$

where σ_z^2 and σ_y^2 are the variances of z_t and y_t respectively, while $\gamma_z(j)$ and $\gamma_y(j)$ are the respective autocovariances at lag j

We have, so far, obtained the joint distribution of the VR_p statistic computed at various differencing periods. These VR statistics can be combined into one single statistic by computing the quadratic form

$$Q_n = (\mathbf{V}_p - \mathbf{E}(\mathbf{V}_p))' \, \mathbf{Var}(\mathbf{V}_p)^{-1} (\mathbf{V}_p - \mathbf{E}(\mathbf{V}_p)), \tag{12}$$

where $\mathbf{V}_p = (VR_p(k_1), ..., VR_p(k_s))'$. Due to the asymptotic normality of $\mathbf{V}_p$, this quadratic form will have an asymptotic chi-squared distribution with s degrees of freedom under the null hypothesis of a random walk. The test statistic Q_n can then be used to test whether the sequence of population variance ratios all equal one for $i = 1, 2, ..., s$. Since the quadratic form Q_n is always positive, rejection of the null hypothesis of a random walk occurs only in the upper tail of the distribution of Q_n. However, under the important alternative of mean reverting processes of the kind imposed in finance applications, the population variance ratios, given by $VRP(k) \equiv Var\left(\sum_{i=1}^{k}\varepsilon_i\right) / (kVar(\varepsilon_1))$ are generally expected to be less than 1 for large k. For example, it can be easily shown that for the alternative models which are the sum of permanent and transitory components (See Poterba and Summers, 1988, and Fama and French, 1988), $VRP(k)$ is less than 1 for all values of k. Hence, under such mean reverting processes, the alternative hypothesis actually has the one sided form $H_a : VRP(k) < 1$ for $i = 1, ..., s$. In such circumstances, ignoring the one sided nature of the alternative can lead to a loss of power of the test. However, Follmann (1996) has proposed a test for the null hypothesis that the mean vector of a multivariate normal random variable is zero, which has good power for alternatives where all the elements of the mean vector are negative. Thus, Follman's procedure would be directly applicable in the setting where the alternative of interest is a mean reverting process. We now adapt Follman's procedure to test for mean reverting alternatives using VR_p statistics as follows. In testing the null hypothesis of a random walk

$$H_0 : VRP(k_1) = ... = VRP(k_s) = 1 \qquad\qquad i = 1, 2, ..., s$$

versus the one sided alternative

$$H_a : VRP(k_1) < 1, ..., VRP(k_s) < 1 \qquad\qquad i = 1, 2, ..., s$$

at the α level of significance, reject the null hypothesis if

$$\sum_{i=1}^{s} [VR_p(k_i) - 1] < 0 \qquad \text{and} \qquad Q_n > \chi^2_{s,2\alpha}, \tag{13}$$

where $\chi^2_{s,2\alpha}$ is the upper 2α critical value of a chi-square distribution with s degrees of freedom. From the asymptotic normality of VR_p and Theorem 2.1 of Follmann (1996), it follows that the procedure given above has an asymptotic level of significance equal to α. An analogous procedure can be developed using the power transformation as follows: Reject the null hypothesis if

$$\sum_{i=1}^{s} \left[VR_p^{\beta_i}(k_i) - 1\right] < 0 \qquad \text{and} \qquad QP_n > \chi^2_{s,2\alpha}, \tag{14}$$

where

$$QP_n = (\mathbf{V}_{p,\beta} - \mu_\beta)' \, \mathbf{\Sigma}_\beta^{-1} (\mathbf{V}_{p,\beta} - \mu_\beta), \tag{15}$$

and μ_β, $\mathbf{\Sigma}_\beta$ are as in Theorem 5. The test procedure based on the power transformation would be expected to have better size and power properties compared to the one based on the original

VR_p statistics since the quadratic form QP_n should be expected to have a distribution closer to the expected chi-square distribution. In the next Section, we report the results from a Monte Carlo study, which evaluates the effectiveness of the new proposals we have made.

4 Simulation Results

We carried out Monte Carlo simulations to evaluate the finite sample performance of tests based on our modified variance ratio statistic. The size properties under the null hypothesis were evaluated using the following two models: (i) $x_t = x_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d\ N(0,1)$ (ii) $x_t = x_{t-1} + \varepsilon_t$, where $\varepsilon_t = \sigma_t v_t$, $v_t \sim i.i.d.\ N(0,1)$ and $\sigma_t^2 = 0.0001 + 0.8575\sigma_{t-1}^2 + 0.1171\varepsilon_{t-1}^2$. The parameter values for the GARCH(1,1) model in (ii) were chosen to reflect values obtained when fitting such models to real data. The sample sizes we considered were $n = 128$ and 512. For $n = 128$, we used $k_1 = 8$ and $k_2 = 16$, whereas for $n = 512$ we used $k_1 = 16$ and $k_2 = 32$. Table 1 reports the Monte Carlo sizes of the test statistics under the Gaussian white noise case whereas Table 2 is for the GARCH(1,1) model. The sizes are reported for the statistics VR_p and VR_p^β for each combination of sample size and k, where β was computed for each case using (8). For each nominal level of significance, the sizes are reported for both the left and right tail to demonstrate the skewness and the effect of the power transformation on it. We also report the sizes of the quadratic tests (12), denoted in the table by Q_n, based upon both the untransformed and transformed VR statistics. Sizes for the modified intersection tests given in (13) and (14), denoted in the table by IQ_n, are also shown.

It is immediately apparent from Table 1 that while the distribution of VR_p is very right skewed, as is well known, the power transformation is able to correct it and provide near perfect normality with sizes in each tail that are very close to nominal. Furthermore, it is also seen that the quadratic and the modified intersection tests based on the transformed VR statistics have much better size properties than those using their untransformed counterparts.

To evaluate the power properties of our tests, we generated data from the mean reverting process given by $x_t = r_t + y_t$, where $r_t = r_{t-1} + w_t$, $y_t = 0.96y_{t-1} + u_t$ and $u_t \sim i.i.d.\ N(0, 1)$ and also independent of $\{w_t\}$. The errors w_t were assumed to be $i.i.d.\ N\left(0, \sigma_w^2\right)$ where $\sigma_w^2 = 0.5, 1$ and 2. This model with the same parameter configuration was considered in Lo and MacKinlay (1989), while Richardson and Smith (1991) used the same model but with slightly different parameter values. Tables 3-5 report the Monte Carlo power values for this alternative model for the three different values of σ_w^2. As the value of σ_w^2 increases, the permanent component dominates the process and the power of all tests decreases, as is to be expected. However, similar behaviour of the tests is seen across all the three tables. It is clear that the individual tests based on the transformed VR statistics provide power which is significantly superior to that of the untransformed ones, in some cases increasing the power by as much as 10%. The quadratic test based on the transformed statistics also provides significant power gain over that based on the untransformed statistics. Furthermore, it is seen that the modified intersection test, which is specially geared to take into account the uni-directional nature of mean reverting alternatives, is able to provide a significant advantage over the quadratic test, when based on the transformed VR statistics. The overall conclusion from the Monte Carlo study is that the transformation of the VR statistic proposed in the paper is able to solve the problem of skewness, providing good size properties as well as significant power gains. The modified intersection test is also able to incorporate information from various differencing periods and yet maintain good power.

Appendix

Proof of Lemma 1:

Since $\{h_t\}$ is a Gaussian stationary series with zero mean, it can be expressed as $h_t = \sum_{j=0}^{\infty} \alpha_j u_{t-j}$, where $\sum \alpha_j^2 < \infty$ and $\{u_t\}$ is a sequence of independent standard normal variables. Furthermore, $\{u_t\}$ and $\{v_t\}$ will also be independent. Let $\mathfrak{F}_t = \sigma\left(u_t, u_{t-1,}u_{t-2}, ..., v_t, v_{t-1,}v_{t-2,...}\right)$. By Lemma 3.5.8 and Theorem 3.5.8 of Stout (1974), $\{\varepsilon_t\}$ is an ergodic sequence. Furthermore, Lemma 1 in Deo (2000) shows that ε_t satisfies (A1) - (A3). Since $\{v_t\}$ is an independent zero mean sequence, (A4) is trivially true. Also,

$$E\left(\varepsilon_{t+n}^2 \varepsilon_{t+n+j}^2 | \mathfrak{F}_t\right) =$$
$$E^2\left(v_t^2\right) \exp\left\{2\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right) u_{t-p}\right\} \exp\left(2\sum_{s=0}^{j-1}\alpha_s^2 + 2\sum_{p=0}^{n-1}\left(\alpha_p+\alpha_{p+j}\right)^2\right).$$

Since $\sum \alpha_j^2 < \infty$, to prove (A5) it suffices to show that

$$\lim_{n\to\infty} Var\left(\exp\left\{2\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right) u_{t-p}\right\}\right) = 0 \tag{16}$$

uniformly in j. But

$$\begin{aligned}
&Var\left(\exp\left\{2\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right) u_{t-p}\right\}\right) \\
&= \exp\left\{8\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right)^2\right\} - \exp\left\{4\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right)^2\right\} \\
&= \exp\left\{4\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right)^2\right\}\left(\exp\left\{4\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right)^2\right\} - 1\right).
\end{aligned}$$

Since $\sum_{p=0}^{\infty}\left(\alpha_{p+n}+\alpha_{p+n+j}\right)^2$ converges to 0 uniformly in j, (16) is established. The proof of (A6) follows along similar lines.

Proof of Lemma 2:

Lemma 2 in Deo (2000) proves (A1) - (A3). An argument similar to the one provided on page 309 in the proof of Lemma 2 of Deo (2000) also establishes (A4). We now turn to proving (A5). Iterating the expression for ε_t, we have

$$\begin{aligned}\varepsilon_{t+n}^2 &= v_{t+n}^2\omega + v_{t+n}^2\omega\sum_{k=1}^{n-2}\Pi_{i=1}^{k}\left(\alpha v_{t+n-i}^2+\beta\right) + v_{t+n}^2\sigma_{t+1}^2\Pi_{i=1}^{n-1}\left(\alpha v_{t+n-i}^2+\beta\right) \\ &\equiv T_{11}+T_{12}+T_{13}\end{aligned} \quad (17)$$

and

$$\begin{aligned}&\varepsilon_{t+n+j}^2 \\ &= v_{t+n+j}^2\omega + v_{t+n+j}^2\omega\sum_{k=1}^{n+j-2}\Pi_{i=1}^{k}\left(\alpha v_{t+n+j-i}^2+\beta\right) + v_{t+n+j}^2\sigma_{t+1}^2\Pi_{i=1}^{n-1}\left(\alpha v_{t+n+j-i}^2+\beta\right) \\ &\equiv T_{21}+T_{22}+T_{23}.\end{aligned}$$

Thus,

$$E\left(\varepsilon_{t+n}^2\varepsilon_{t+n+j}^2|\mathfrak{F}_t\right) = \sum_{p,q=1}^{3}E\left(T_{1p}T_{2q}|\mathfrak{F}_t\right). \quad (18)$$

Consider the term $T_{12}T_{23}$. Then we can easily see that we can express $T_{12}T_{23}$ as the product $T_{12}T_{23}=AB$, where

$$A = \omega\sigma_{t+1}^2v_{t+n+j}^2v_{t+n}^2\left(\alpha v_{t+n}^2+\beta\right)\Pi_{i=1}^{j-1}\left(\alpha v_{t+n+i}^2+\beta\right)$$

and

$$B = \Pi_{i=1}^{n-1}\left(\alpha v_{t+n-i}^2+\beta\right)\sum_{k=1}^{n-2}\Pi_{i=1}^{k}\left(\alpha v_{t+n-i}^2+\beta\right).$$

Letting $\theta_1 = E\left(\alpha v_t^2+\beta\right)$ and $\theta_2 = E\left(\alpha v_t^2+\beta\right)^2$ and noting that $Ev_{t+n}^4=3$, we get

$$E\left(T_{12}T_{23}|\mathfrak{F}_t\right) \le \omega\sigma_{t+1}^2\left(3\alpha+\beta\right)\theta_1^{j-1}\sum_{k=1}^{n-2}\theta_2^k\theta_1^{n-1-k}.$$

Since $\gamma=\max\left(\theta_1,\theta_2\right)<1$, it follows that for all $j\ge 1$ there exists some finite constant C such that

$$E\left(T_{12}T_{23}|\mathfrak{F}_t\right) \le C\sigma_{t+1}^2\left(n-2\right)\gamma^{n-1}$$

and hence

$$Var\left(E\left(T_{12}T_{23}|\mathfrak{F}_t\right)\right) \leq E\left(E\left(T_{12}T_{23}|\mathfrak{F}_t\right)\right)^2 \leq C^2 E\left(\sigma_{t+1}^4\right) n^2 \gamma^{2(n-1)}$$

uniformly in j. Thus,

$$\lim_{n\to\infty} Var\left(E\left(T_{12}T_{23}|\mathfrak{F}_t\right)\right) = 0$$

uniformly in j. Similar arguments yield

$$\lim_{n\to\infty} Var\left(E\left(T_{1p}T_{2q}|\mathfrak{F}_t\right)\right) = 0 \qquad 1 \leq p, q \leq 3 \tag{19}$$

uniformly in j. Thus, (A5) follows from (18), (19) and the Cauchy Schwarz inequality. To prove (A6), we first note that using (17),

$$E\left(\varepsilon_t^2|\mathfrak{F}_{t-n}\right) = \omega\left(\frac{1-\theta_1^n}{1-\theta_1}\right) + \theta_1^n \sigma_{t-n}^2.$$

Thus, $E\left(\varepsilon_t^2\varepsilon_{t-n}^2\right) = E\left(\varepsilon_{t-n}^2 E\left(\varepsilon_t^2|\mathfrak{F}_{t-n}\right)\right) = \omega\left(\frac{1-\theta_1^n}{1-\theta_1}\right)\sigma^2 + \theta_1^n E\left(\sigma_{t-n}^4\right)$ and so

$$\lim_{n\to\infty} E\left(\varepsilon_t^2\varepsilon_{t-n}^2\right) = \omega\left(\frac{1-\theta_1^n}{1-\theta_1}\right)\sigma^2 = \sigma^4.$$

Proof of Theorem 3:

By simple but tedious algebraic manipulation, it can be shown that

$$\begin{aligned}[VR\left(k_i\right) - 1] = {} & \frac{2n^2}{\hat{\sigma}^2\left(n-k_i+1\right)\left(n-k_i\right)} \sum_{j=1}^{k_i-1}\left(1-\frac{j}{k_i}\right)\hat{\gamma}_j - \frac{n\left(A_i+B_i\right)}{\hat{\sigma}^2 k_i\left(n-k_i+1\right)\left(n-k_i\right)} \\ & + o_p\left(\sqrt{\frac{k_i}{n}}\right)\end{aligned}$$

where

$$\hat{\gamma}_j = n^{-1}\sum_{t=j+1}^{n}\varepsilon_t\varepsilon_{t-j},$$

$$\begin{aligned}A_i &= -2\sum_{v=0}^{k_i-2}\sum_{p=1}^{k_i-1-v}\sum_{s=p+1}^{k_i-v-1}\varepsilon_s\varepsilon_{s-p} - 2\sum_{v=0}^{k_i-2}\sum_{p=1}^{k_i-1-v}\sum_{s=n-v+1}^{n}\varepsilon_s\varepsilon_{s-p} \\ &= A_{i1} + A_{i2}\end{aligned}$$

and

$$B_i = \sum_{v=0}^{k_i-1} \sum_{q=1}^{k_i-v-1} \varepsilon_q^2 + \sum_{v=0}^{k_i-1} \sum_{q=n-v+1}^{n} \varepsilon_q^2.$$

Since $E(B_i) = O\left(k_i^2\right)$ trivially, it follows that $\left[k_i\left(n-k_i+1\right)\left(n-k_i\right)\right]^{-1} nB_i = o_p\left(\left[n^{-1}k_i\right]^{\frac{1}{2}}\right)$. By condition (A1), we have $E(A_i) = 0$. Furthermore, by using condition (A3), it can be easily seen that $E\left(A_{i1}^2\right) = E\left(A_{i2}^2\right) = O\left(k_i^4\right)$. By the Cauchy Schwarz inequality, it follows that $Var(A_i) = O\left(k_i^4\right)$ and hence $\left[k_i\left(n-k_i+1\right)\left(n-k_i\right)\right]^{-1} nA_i = o_p\left(\left[n^{-1}k_i\right]^{\frac{1}{2}}\right)$. Since $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$, we have

$$\sqrt{\frac{n}{k_i}}\left[VR(k_i) - 1\right] = \frac{n^2}{\hat{\sigma}^2\left(n-k_i+1\right)\left(n-k_i\right)}\sqrt{\frac{n}{k_i}}2\sum_{j=1}^{k_i-1}\left(1-\frac{j}{k_i}\right)\hat{\gamma}_j + o_p(1). \tag{20}$$

Now consider

$$\sqrt{\frac{n}{k_i}}2\sum_{j=1}^{k_i-1}\left(1-\frac{j}{k_i}\right)\hat{\gamma}_j = \frac{2}{\sqrt{nk_i}}\sum_{j=1}^{k_i-1}\left(1-\frac{j}{k_i}\right)\sum_{q=1}^{n-j}\varepsilon_q\varepsilon_{q+j}$$

$$= \frac{2}{\sqrt{nk_i}}\sum_{q=1}^{n-k_i}\sum_{j=1}^{k_i}\left(1-\frac{j}{k_i}\right)\varepsilon_q\varepsilon_{q+j} + \frac{2}{\sqrt{nk_i}}\sum_{q=n-k_i+1}^{n}\sum_{j=1}^{n-q}\left(1-\frac{j}{k_i}\right)\varepsilon_q\varepsilon_{q+j}$$

$$= R_{i1} + R_{i2}.$$

By conditions (A1) and (A3) respectively, it follows that $E(R_{i2}) = 0$ and $E\left(R_{i2}^2\right) = o(1)$ and hence

$$\sqrt{\frac{n}{k_i}}2\sum_{j=1}^{k_i-1}\left(1-\frac{j}{k_i}\right)\hat{\gamma}_j = R_{i1} + o_p(1). \tag{21}$$

Now define $N = [\sqrt{nk_s}]$ and $M = \left[N^{-1}n\right]$. Then, $M \to \infty$, $N \to \infty$, $n^{-1}N \to 0$ and $N^{-1}k_i \to 0$ for $i = 1, 2, ..., s$. Also, define

$$W_{i,\,q} = \frac{1}{\sqrt{k_i}}\sum_{j=1}^{k_i}\left(1-\frac{j}{k_i}\right)\varepsilon_q\varepsilon_{q+j} \qquad q = 1, 2, ..., n-k_i,$$

$$Z_{i,\,p} = \frac{1}{\sqrt{N}}\left\{W_{i,\,(p-1)N+1} + ... + W_{i,\,pN-k_i}\right\} \qquad p = 1, 2, ..., M.$$

and

$$V_{i,\,l} = W_{i,\,lN-k_i+1} + ... + W_{i,\,lN} \qquad l = 1, 2, ..., M-1.$$

Then we can decompose R_{i1} as

$$R_{i1} = \frac{2}{\sqrt{M}} \sum_{p=1}^{M} Z_{i,\, p} + \frac{2}{\sqrt{n}} \sum_{l=1}^{M-1} V_{i,\, l} \tag{22}$$
$$\equiv U_{i1} + U_{i2}.$$

By condition (A3), it follows that $E\left(W_{i,\,a} W_{i,\,b}\right) = 0$ for $a < b$ and hence $E\left(V_{i,\,a} V_{i,\,b}\right) = 0$ for $a < b$. Thus,

$$E\left(U_{i2}^2\right) = \frac{4}{n} \sum_{l=1}^{M-1} E\left(V_{i,\,l}^2\right) = \frac{4}{n} \sum_{l=1}^{M-1} k_i E\left(W_{i,\,1}^2\right)$$
$$= O\left(\frac{k_i (M-1)}{n}\right) = o(1) \qquad i = 1, 2, ..., s. \tag{23}$$

From equations (21), (22) and (23) it follows that

$$\sqrt{\frac{n}{k_i}} 2 \sum_{j=1}^{k_i - 1} \left(1 - \frac{j}{k_i}\right) \hat{\gamma}_j = U_{i1} + o_p(1)$$

and hence, from equation (20),

$$\sqrt{\frac{n}{k_i}} \left[VR(k_i) - 1\right] = \frac{n^2}{\hat{\sigma}^2 (n - k_i + 1)(n - k_i)} U_{i1} + o_p(1) .$$

Since $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$ and $\left[(n - k_i + 1)(n - k_i)\right]^{-1} n^2 \to 1$, the Theorem will be proved if we show that the vector $(U_{11}, U_{21}, ..., U_{s1})'$ converges in distribution to a multivariate normal distribution with mean zero and variance covariance matrix $\sigma^4 \mathbf{\Sigma}$. To do this, it is sufficient to show that for any set of s real numbers c_i,

$$\sum_{i=1}^{s} c_i U_{i1} = 2M^{-\frac{1}{2}} \sum_{p=1}^{M} \sum_{i=1}^{s} c_i Z_{i,\, p} \xrightarrow{D} N\left(0, \sigma^4 \sum_{i,j} c_i c_j \sigma_{ij}\right), \tag{24}$$

which we now proceed to demonstrate.

Let $\mathfrak{G}_{p,\, n} = \sigma\left\{\varepsilon_{pN}, \varepsilon_{pN-1}, \varepsilon_{pN-2}, ...\right\}$ be the sigma algebra generated by $\left\{\varepsilon_{pN}, \varepsilon_{pN-1}, \varepsilon_{pN-2}, ...\right\}$. Then, for any set of s real numbers c_i, the sequence $\left\{\sum_{i=1}^{s} c_i Z_{i,\, p}\right\}$ forms a martingale difference with respect to $\mathfrak{G}_{p,\,n}$. To show (24), we first need to establish that

$$\left(\sum_{p=1}^{M} E\left(\sum_{i=1}^{s} c_i Z_{i,\,p}\right)^2\right)^{-1} \sum_{p=1}^{M} E\left[\left(\sum_{i=1}^{s} c_i Z_{i,\,p}\right)^2 |\mathfrak{G}_{p-1,\,n}\right] \xrightarrow{P} 1. \tag{25}$$

Now, by condition (A3)

$$E\left(\left[\sum_{i=1}^{s} c_i Z_{i,\,p}\right]^2\right) = \sum_{i=1}^{s} c_i^2 E\left(Z_{i,\,p}^2\right) + 2\sum_{i<u} c_i c_u E\left(Z_{i,\,p} Z_{u,\,p}\right)$$
$$= \sum_{i=1}^{s} c_i^2 \frac{N-k_i}{N} E\left(W_{i,\,1}^2\right) + 2\sum_{i<u} c_i c_u \frac{N-k_i}{N} E\left(W_{i,\,1} W_{u,\,1}\right).$$

By conditions (A3) and (A6),

$$E\left(W_{i,\,1}^2\right) = k_i^{-1} \sum_{j=1}^{k_i} \left(1 - \frac{j}{k_i}\right)^2 E\left(\varepsilon_1^2 \varepsilon_{1+j}^2\right) \to 4^{-1}\sigma^4 \sigma_{ii}$$

and

$$E\left(W_{i,\,1} W_{u,\,1}\right) = \left(k_u k_i\right)^{-\frac{1}{2}} \sum_{j=1}^{k_i} \left(1 - \frac{j}{k_i}\right)\left(1 - \frac{j}{k_u}\right) E\left(\varepsilon_1^2 \varepsilon_{1+j}^2\right) \to 4^{-1}\sigma^4 \sigma_{iu}$$

for $i < u$. Hence, we have

$$\lim_{M\to\infty} \frac{1}{M} \sum_{p=1}^{M} E\left(\left[\sum_{i=1}^{s} c_i Z_{i,\,p}\right]^2\right) = 4^{-1}\sigma^4 \sum_{i,j} c_i c_j \sigma_{ij}. \tag{26}$$

We now show that

$$M^{-1} \sum_{p=1}^{M} E\left[\left(\sum_{i=1}^{s} c_i Z_{i,\,p}\right)^2 |\mathfrak{G}_{p-1,\,n}\right] \xrightarrow{P} 4^{-1}\sigma^4 \sum_{i,j} c_i c_j \sigma_{ij}, \tag{27}$$

which along with (26) will prove (25). We have

$$E\left[\left(\sum_{i=1}^{s} c_i Z_{i,\,p}\right)^2 |\mathfrak{G}_{p-1,\,n}\right] = \sum_{i=1}^{s} c_i^2 E\left(Z_{i,\,p}^2 |\mathfrak{G}_{p-1,\,n}\right)$$
$$+ 2\sum_{i<u} c_i c_u E\left(Z_{i,\,p} Z_{u,\,p} |\mathfrak{G}_{p-1,\,n}\right).$$

Letting $f(x) = (1-x)$, $Y_{i,u,p} = E\left(Z_{i,\,p} Z_{u,\,p} |\mathfrak{G}_{p-1,\,n}\right)$ and using condition (A4), we get for $i \le u$,

$$Y_{i,u,p} = N^{-1}\left(k_i k_u\right)^{-\frac{1}{2}} \sum_{a=1}^{N-k_u} \sum_{b=1}^{k_i} f\left(\frac{b}{k_i}\right) f\left(\frac{b}{k_u}\right) E\left(\varepsilon_{pN-N+a}^2 \varepsilon_{pN-N+a+b}^2 |\mathfrak{G}_{p-1,\,n}\right).$$

By condition (A6), there exists $C < \infty$ such that

$$E\left|E\left(\varepsilon_{pN-N+a}^2\varepsilon_{pN-N+a+b}^2|\mathfrak{G}_{p-1,\,n}\right)-E\left(\varepsilon_t^2\varepsilon_{t+b}^2\right)\right| < C \tag{28}$$

for all p, a and b. Furthermore, given any $\delta > 0$, by condition (A5) and Jensen's inequality there exists an integer N_0 such that

$$\sup_{b>0} E\left|E\left(\varepsilon_{pN-N+a}^2\varepsilon_{pN-N+a+b}^2|\mathfrak{G}_{p-1,\,n}\right)-E\left(\varepsilon_t^2\varepsilon_{t+b}^2\right)\right| < \delta \tag{29}$$

for all $a > N_0$. Hence, letting $H_{p,a,b} = E\left(\varepsilon_{pN-N+a}^2\varepsilon_{pN-N+a+b}^2|\mathfrak{G}_{p-1,\,n}\right) - E\left(\varepsilon_t^2\varepsilon_{t+b}^2\right)$, we have for any $\varepsilon > 0$

$$P\left\{\left|M^{-1}\sum_{p=1}^{M}E\left(Z_{i,\,p}^2|\mathfrak{G}_{p-1,\,n}\right)-M^{-1}\sum_{p=1}^{M}N^{-1}k_i^{-1}\sum_{a=1}^{N-k_i}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)E\left(\varepsilon_t^2\varepsilon_{t+b}^2\right)\right|>\varepsilon\right\}$$

$$\leq P\left\{N^{-1}k_i^{-1}\sum_{a=1}^{N-k_i}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)M^{-1}\sum_{p=1}^{M}|H_{p,a,b}|>\varepsilon\right\}$$

$$\leq P\left\{N^{-1}k_i^{-1}\sum_{a=1}^{N_0}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)M^{-1}\sum_{p=1}^{M}|H_{p,a,b}|>2^{-1}\varepsilon\right\}$$
$$+P\left\{N^{-1}k_i^{-1}\sum_{a=N_0+1}^{N-k_i}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)M^{-1}\sum_{p=1}^{M}|H_{p,a,b}|>2^{-1}\varepsilon\right\}$$

$$\leq 2\varepsilon^{-1}N^{-1}k_i^{-1}\sum_{a=1}^{N_0}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)M^{-1}\sum_{p=1}^{M}E\left|H_{p,a,b}\right|$$
$$+2\varepsilon^{-1}N^{-1}k_i^{-1}\sum_{a=N_0+1}^{N-k_i}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)M^{-1}\sum_{p=1}^{M}E\left|H_{p,a,b}\right|$$

$$\leq 2\varepsilon^{-1}N^{-1}N_0k_i^{-1}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)C+2\varepsilon^{-1}N^{-1}\left(N-k_i\right)k_i^{-1}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)\delta, \tag{30}$$

where the last inequality follows from equations (28) and (29). Since δ can be chosen to be arbitrarily small and N large enough that $N^{-1}N_0 \to 0$, it follows from equation (30) that

$$M^{-1}\sum_{p=1}^{M}E\left(Z_{i,\,p}^2|\mathfrak{G}_{p-1,\,n}\right)-M^{-1}\sum_{p=1}^{M}N^{-1}k_i^{-1}\sum_{a=1}^{N-k_i}\sum_{b=1}^{k_i}f^2\left(\frac{b}{k_i}\right)E\left(\varepsilon_t^2\varepsilon_{t+b}^2\right)\xrightarrow{P}0.$$

Since, by condition (A6) we also have

$$M^{-1}\sum_{p=1}^{M} N^{-1}k_i^{-1}\sum_{a=1}^{N-k_i}\sum_{b=1}^{k_i} f^2\left(\frac{b}{k_i}\right) E\left(\varepsilon_t^2\varepsilon_{t+b}^2\right) \to \sigma^4 4^{-1}\sigma_{ii},$$

we obtain

$$M^{-1}\sum_{p=1}^{M} E\left(Z_{i,\,p}^2|\mathfrak{G}_{p-1,\,n}\right) \xrightarrow{P} \sigma^4 4^{-1}\sigma_{ii}.$$

A similar argument as above in conjunction with the fact that $k_u^{-1}k_i \to a_{iu}$ for $i < u$ yields

$$M^{-1}\sum_{p=1}^{M} E\left(Z_{i,\,p}Z_{u,\,p}|\mathfrak{G}_{p-1,\,n}\right) \xrightarrow{P} \sigma^4 4^{-1}\sigma_{iu}.$$

Thus, (27) is established giving equation (25).

By using condition (A3), one can employ the same argument given on page 539 of Anderson (1994) to show that $E\left(Z_{i,\,p}^4\right)$ is uniformly bounded in n for $i = 1, 2, ..., s$. This implies that $E\left(\sum_{i=1}^s c_i Z_{i,\,p}\right)^4$ is also uniformly bounded in n from whence we get

$$M^{-1}\sum_{p=1}^{M} E\left[\left(\sum_{i=1}^{s} c_i Z_{i,\,p}\right)^2 I\left(|\sum_{i=1}^{s} c_i Z_{i,\,p}| > \varepsilon\sqrt{M}\right)\right] \to 0 \qquad (31)$$

for every $\varepsilon > 0$. By Chebyshev's inequality, equation (31) implies that

$$M^{-1}\sum_{p=1}^{M} E\left[\left(\sum_{i=1}^{s} c_i Z_{i,\,p}\right)^2 I\left(|\sum_{i=1}^{s} c_i Z_{i,\,p}| > \varepsilon\sqrt{M}\right)|\mathfrak{G}_{p-1,\,n}\right] \xrightarrow{P} 0. \qquad (32)$$

Hence, equation (24) follows from equations (25) and (32) and Theorem 5.3.4 of Fuller (1996).

Proof of Theorem 4:

We first note that by the weak law of large numbers, $\hat{\sigma}_a^2 \xrightarrow{P} Var\left(z_t\right) + Var\left(y_t - y_{t-1}\right)$. Now, letting $V_{n,k} \equiv n\left(\hat{\sigma}_a^2 k\left(n-k+1\right)\left(n-k\right)\right)^{-1}$, we get

$$\begin{aligned} VR\left(k\right) &= V_{n,k}\sum_{t=k}^{n}\left(x_t - x_{t-k} - k\hat{\mu}\right)^2 \\ &= V_{n,k}\sum_{t=k}^{n}\left[\left\{\sum_{j=t-k+1}^{t} z_j - k\bar{z}\right\} + \sum_{t=k}^{n}\left(y_t - y_{t-k} - \frac{k}{n}\left\{y_n - y_0\right\}\right)\right]^2. \end{aligned} \qquad (33)$$

It is trivial to show that

$$\sum_{t=k}^{n}\left(y_t - y_{t-k} - \frac{k}{n}\{y_n - y_0\}\right)^2 = o_p(nk). \tag{34}$$

Now

$$\begin{aligned}\sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j - k\bar{z}\right\}^2 &= \sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j\right\}^2 + (n-k)k^2\bar{z}^2 - 2k\bar{z}\sum_{t=k}^{n}\sum_{j=t-k+1}^{t} z_j \\ &= \sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j\right\}^2 + (n-k)k^2 O_p\left(n^{-1}\right) \\ &\quad - 2k\bar{z}\left(\sum_{j=0}^{k} j z_j + \sum_{j=k}^{n-k} k z_j + \sum_{j=n-k}^{n}(n-j) z_j\right) \\ &= \sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j\right\}^2 + O_p\left(k^2\right) + O_p\left(k^2\right). \end{aligned} \tag{35}$$

From (35), we get

$$V_{n,k}\sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j - k\bar{z}\right\}^2 = V_{n,k}\sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j\right\}^2 + o_p(1). \tag{36}$$

Letting $\hat{\gamma}_j = \hat{\gamma}_{-j} = n^{-1}\sum_{t=j+1}^{n} z_t z_{t-j}$, some tedious algebra yields

$$\sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j\right\}^2 = nk\sum_{j=-(k-1)}^{k-1}\left(1 - |j|/k\right)\hat{\gamma}_j - A - B, \tag{37}$$

where

$$\begin{aligned} A &= -2\sum_{v=0}^{k-2}\sum_{p=1}^{k-1-v}\sum_{s=p+1}^{k-v-1} z_s z_{s-p} - 2\sum_{v=0}^{k-2}\sum_{p=1}^{k-1-v}\sum_{s=n-v+1}^{n} z_s z_{s-p} \\ &= A_1 + A_2 \end{aligned}$$

and

$$B = \sum_{v=0}^{k-1}\sum_{q=1}^{k-v-1} z_q^2 + \sum_{v=0}^{k-1}\sum_{q=n-v+1}^{n} z_q^2.$$

Now

$$E\left(A_1^2\right) = 4\sum_{p=1}^{k-1}\sum_{v=1}^{k-1-p}\sum_{s=1}^{k-1}\sum_{j=1}^{k-1-s}(k-v-p)(k-j-s)E\left(z_v z_{v+p} z_j z_{j+s}\right).$$

From equation (6.2.5) of page 315 of Fuller (1996), we have $\left|E\left(z_v z_{v+p} z_j z_{j+s}\right)\right| = O\left(\lambda^{|v|+|p|+|j|+|s|}\right)$ and hence

$$E\left(A_1^2\right) = O\left(k^4\right).$$

A similar argument shows that $E\left(A_2^2\right) = O\left(k^4\right)$ and hence, by the Cauchy Schwarz and Chebyshev inequalities, we get

$$A = O_p\left(k^2\right). \tag{38}$$

Since $E(B) = O\left(k^2\right)$ trivially, it follows from (38), (37) and (36) that

$$V_{n,k}\sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j - k\bar{z}\right\}^2 = V_{n,k}nk\sum_{j=-(k-1)}^{k-1}\left(1-|j|/k\right)\hat{\gamma}_j + o_p(1).$$

From Theorem 9.3.3 and Theorem 9.4.1 of Anderson (1994), it follows that

$$\sum_{j=-(k-1)}^{k-1}\left(1-|j|/k\right)\hat{\gamma}_j \xrightarrow{P} \sum_{|j|<\infty}\gamma_z(j)$$

and hence

$$V_{n,k}\sum_{t=k}^{n}\left\{\sum_{j=t-k+1}^{t} z_j - k\bar{z}\right\}^2 \xrightarrow{P} \left(Var(z_t) + Var(y_t - y_{t-1})\right)^{-1}\sum_{|j|<\infty}\gamma_z(j). \tag{39}$$

From (33), (34), (39) and the Cauchy-Schwarz inequality, we get

$$VR(k) \xrightarrow{P} \left(Var(z_t) + Var(y_t - y_{t-1})\right)^{-1}\sum_{|j|<\infty}\gamma_z(j).$$

Lemma 7 **(i)** $\int_0^{2\pi} W_k(\lambda) I(\lambda) d\lambda = \frac{4\pi}{n}\sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j) I(\lambda_j) + o_p\left(\sqrt{k/n}\right).$

(ii) *The finite sample variance covariance matrix of* $\mathbf{V}_p = \left(VR_p(k_1), VR_p(k_2), ..., VR_p(k_s)\right)'$ *with remainder terms of order* $o\left(k_s^2/n^2\right)$ *is estimated consistently by the matrix* $\hat{\mathbf{\Sigma}}$ *in (10).*

Proof of (i): Using the fact that $I(\lambda) = (2\pi)^{-1} \sum_{|s|<n} \hat{\gamma}_s \exp(-is\lambda)$ and that

$$\sum_{j=0}^{n-1} \exp(-i(s-p)\lambda_j) = n \qquad \text{if } s-p=0,\ \pm n$$
$$= 0 \text{ otherwise,}$$

we get

$$\begin{aligned}
\frac{2\pi}{n}\sum_{j=1}^{n-1} W_k(\lambda_j) I(\lambda_j) &= \frac{2\pi}{n}\sum_{j=0}^{n-1} W_k(\lambda_j) I(\lambda_j) - \frac{2\pi k}{n} I(0) \\
&= \frac{2\pi}{n}\sum_{j=0}^{n-1}\sum_{|p|<k} (1-|p|/k)\exp(ip\lambda_j)(2\pi)^{-1}\sum_{|s|<n}\hat{\gamma}_s \exp(-is\lambda_j) - \frac{2\pi k}{n} I(0) \\
&= \frac{1}{n}\sum_{|p|<k}\sum_{|s|<n}(1-|p|/k)\hat{\gamma}_s \sum_{j=0}^{n-1}\exp(-i(s-p)\lambda_j) - \frac{2\pi k}{n} I(0) \\
&= \sum_{|p|<k}(1-|p|/k)\hat{\gamma}_p + 2\sum_{p=1}^{k}(1-p/k)\hat{\gamma}_{n-p} - \frac{2\pi k}{n} I(0) \qquad (40) \\
&= \int_0^{2\pi} W_k(\lambda) I(\lambda)\, d\lambda + 2\sum_{p=1}^{k}(1-p/k)\hat{\gamma}_{n-p} - \frac{2\pi k}{n} I(0)
\end{aligned}$$

where the last step follows from the identity $\hat{\gamma}_j = \int_0^{2\pi} I(\lambda)\exp(-ij\lambda)\, d\lambda$. We now note that since $I(\lambda) = (2\pi n)^{-1} \left|\sum_{t=1}^{n} \varepsilon_t \exp(i\lambda t)\right|^2$, it follows that $(2\pi k/n) I(0) = k\bar{\varepsilon}^2 = O_p(kn^{-1}) = O_p\left(\sqrt{k/n}\right)$. Furthermore, $Var(\hat{\gamma}_{n-p}) = O(pn^{-2})$ while $Cov(\hat{\gamma}_{n-p},\ \hat{\gamma}_{n-s}) = 0$ which implies that $2\sum_{p=1}^{k}(1-p/k)\hat{\gamma}_{n-p} = O_p\left(\sqrt{k/n}\right)$. Part (i) of the lemma now follows by noting that

$$(2\pi/n)\sum_{j=1}^{n-1} W_k(\lambda_j) I(\lambda_j) = (4\pi/n)\sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j) I(\lambda_j) + \frac{2\pi}{nk} I\left(\lambda_{n/2}\right)\delta_{\{n \text{ even}, k \text{ odd}\}},$$

where δ is the indicator function due to the periodicity of the sine and cosine functions on $[0, 2\pi]$.

Proof of (ii): Using a Taylor series expansion and the equation (40) in the proof of part (i) above, we get

$$VR_p(k) = 1 + C_{n,k}(4\pi/n) \sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j) I(\lambda_j) - \hat{\sigma}^2 + O_p\left(k^{1/2}/n\right)$$
$$= 1 + 2C_{n,k} \sum_{j=1}^{k-1} (1 - j/k)(\hat{\gamma}_j + \hat{\gamma}_{n-j}) - (kC_{n,k} - n/(n-1))\bar{\varepsilon}^2 + O(k/n)\hat{\gamma}_0 + O_p\left(k^{1/2}/n\right).$$

Now define the random vector $\mathbf{U} = \left(\hat{\gamma}_1 + \hat{\gamma}_{n-1},\ \hat{\gamma}_2 + \hat{\gamma}_{n-2}, ..., \hat{\gamma}_s + \hat{\gamma}_{n-s},\ \bar{\varepsilon}^2\right)$. Since $Var(\hat{\gamma}_0) = O\left(n^{-1}\right)$, it is seen that

$$Var(VR_p(k)) = l_k' Var(\mathbf{U})\, l_k + o\left(k^2/n^2\right), \tag{41}$$

where l_k is as defined in (11). Letting $\tau_j = \sigma^{-4} E\left(\varepsilon_t^2 \varepsilon_{t-j}^2\right)$, tedious but elementary calculation shows that

$$Var(\mathbf{U}) = \begin{bmatrix} \mathbf{A}_0 & \mathbf{b}_0 \\ \mathbf{b}_0' & d_0 \end{bmatrix}, \tag{42}$$

where $\mathbf{A}_0 = diag\left(\frac{n-j}{n^2}\tau_j + \frac{j}{n^2}\tau_{n-j}\right)$ for $j = 1, ..., k_s$, $\mathbf{b}_0$ is a $k_s \times 1$ vector such that its j^{th} element is given by $\left(2(n-j)n^{-3}\tau_j + 2jn^{-3}\tau_{n-j}\right)$ and $d_0 = n^{-3}\tau_0 + 6n^{-4}\sum_{u=1}^{n-1}(n-u)\tau_u - n^{-2}$. Using the fact that by Assumption (A6) $\tau_j \to 1$ as $j \to \infty$, it is easily seen that $6n^{-2}\sum_{u=1}^{n-1}(n-u)\tau_u = 3 + o(1)$ and these facts in conjunction with substituting (42) in (41), we get

$$Var(VR_p(k)) = l_k' \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}' & d \end{bmatrix} l_k + o(k^2/n^2),$$

where $\mathbf{A} = diag\left(\frac{n-j}{n^2}\tau_j + \frac{j}{n^2}\right)$ for $j = 1, ..., k_s$, $\mathbf{b}$ is a $k_s \times 1$ vector such that its j^{th} element is given by $\left(2(n-j)n^{-3}\tau_j + 2jn^{-3}\right)$ and $d = 2n^{-2}$. The estimated variance covariance matrix is now obtained by replacing τ_j in the entries of $\mathbf{A}$ and $\mathbf{b}$ by $\hat{\tau}_j$ and standard arguments from smoothing theory establish consistency of the resulting estimated covariance matrix.

Proof of Theorem 6: In the proof of Lemma 7, we noted that

$$(4\pi/n) \sum_{j=1}^{[(n-1)/2]} W_k(\lambda_j) I(\lambda_j) = \sum_{|p|<k} (1 - |p|/k)\hat{\gamma}_p + 2\sum_{p=1}^{k} (1 - p/k)\hat{\gamma}_{n-p} - k\bar{\varepsilon}^2.$$

It is trivially true that under the assumptions of Theorem **6**, $\bar{\varepsilon}^2 = O_p\left(n^{-1}\right)$. The result for $VR_p\left(k\right)$ now follows by noting that $\sum_{p=1}^{k}\left(1-p/k\right)\hat{\gamma}_{n-p} = o_p\left(1\right)$, that $\hat{\sigma}^2 \xrightarrow{p} \left(Var\left(z_t\right)+Var\left(y_t-y_{t-1}\right)\right)$ and that by Theorem 9.3.3 and Theorem 9.4.1 of Anderson (1994),

$$\sum_{j=-(k-1)}^{k-1}\left(1-\left|j\right|/k\right)\hat{\gamma}_j \xrightarrow{P} \sum_{|j|<\infty}\gamma_z\left(j\right).$$

The result for $VR_p^{\beta}\left(k\right)$ follows by continuity.

REFERENCES

Anderson, T. W. (1994). The Statistical Analysis of Time Series. Wiley Classics edition.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics,* 31, 307-327.

Bougerol, P. and Picard, N. (1992). "Stationarity of GARCH processes and of Some Non-Negative Time Series", *Journal of Econometrics*, 52, 115-127.

Campbell, J., Lo, A. and MacKinlay, A. (1997). The Econometrics of Financial Markets. Princeton University Press, Princeton.

Campbell, J. and Mankiw, N. G. (1987). "Are Output Fluctuations Transitory?", *Quarterly Journal of Economics*, 102, 857-880.

Cochrane, J. (1988). "How Big is the Random Walk in GNP?", *Journal of Political Economy,* 96, 893-920.

Chen, W. and Deo, R. (2003). "Power transformations to induce normality and their applications," *Journal of the Royal Statistical Society, Series B*, To Appear.

Deo, R. (2000). “Spectral Tests of the Martingale Hypothesis under Conditional Heteroscedasticity”, *Journal of Econometrics*, 99, 291-315.

Deo, R. and Richardson, M. (2003). ”On the Asymptotic power of the variance ratio test”, *Econometric Theory,* 19, 231-239.

Fama, E. and French, K. (1988). “Permanent and Temporary Components of Stock Prices”, *Journal of Political Economy”,* 96, 246-273.

Faust, J. (1992). “When Are Variance Ratio Tests For Serial Dependence Optimal?”, *Econometrica,* 60, 1215-1226.

Follmann, D. (1996). “A Simple Multivariate Test For One-Sided Alternatives”, *Journal of the American Statistical Association*, 91, 854-861.

Fuller, W. (1996). Introduction to Time Series Analysis. Second Edition. Wiley, New York.

Lo, A. and MacKinlay, A. C. (1988). “Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test”, *Review of Financial Studies,* 1,41-66.

Lo, A. and MacKinlay, A. C. (1989). The Size and Power of the Variance Ratio Test in Finite Samples: a Monte Carlo Investigation. *Journal of Econometrics,* 40, 203-238.

Nelson, D. (1990). “Stationarity and Persistence in the GARCH(1,1) model”. *Econometric Theory*, 6, 318-334.

Poterba, J. and Summers, L. (1988). “Mean Reversion in Stock returns: Evidence and Implications”, *Journal of Financial Economics*, 22, 27-60.

Richardson, M and Smith, T. (1991). “Tests of Financial Models with the Presence of Overlapping Observations”, *Review of Financial Studies*, 4, 227-254.

Richardson, M. and Stock, J. (1989). "Drawing Inferences from Statistics based on Multi-Year Asset Returns", *Journal of Financial Economics,* 25, 323-348.

Shephard, N. (1996). "Statistical Aspects of ARCH and Stochastic Volatility". Time Series Models in Econometrics, Finance and Other Fields. Ed. Cox, D. R. et al. Chapman & Hall, London.

Stout, W. (1974). Almost Sure Convergence. Academic Press, New York.

Table I. Sizes in Percentage under the Null of Random Walk with Gaussian White Noise

			5%						10%					
			Before transformation			After transformation			Before transformation			After transformation		
n		k	Lower	Upper	Size	Lower	Upper	Size	Lower	Upper	Size	Lower	Upper	Size
128	VR_p	8	0.58	3.82	4.39	2.44	2.29	4.73	2.54	6.28	8.81	4.94	4.73	9.67
		16	0.05	4.51	4.56	2.36	2.36	4.71	0.92	7.02	7.94	4.92	4.94	9.86
	Q_n				5.80			5.13			9.43			9.85
	IQ_n				1.52			5.71			4.95			10.86
512	VR_p	16	1.04	3.59	4.62	2.41	2.38	4.79	3.18	6.26	9.44	4.86	4.89	9.75
		32	0.05	4.19	4.69	2.41	2.52	4.92	2.24	6.62	8.86	4.83	5.01	9.84
	Q_n				5.21			4.79			9.36			9.66
	IQ_n				2.17			5.45			6.49			10.74

Data are generated from $x_t = \mu + x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$

Table II. Sizes in Percentage under the Null of Random Walk with GARCH(1,1) White Noise

			5%						10%					
			Before transformation			After transformation			Before transformation			After transformation		
n		k	Lower	Upper	Size	Lower	Upper	Size	Lower	Upper	Size	Lower	Upper	Size
128	VR_p	8	0.31	4.08	4.38	1.96	2.44	4.40	1.74	6.45	8.19	4.46	4.86	9.32
		16	0.02	4.64	4.66	1.96	2.56	4.51	0.55	6.90	7.45	4.50	4.95	9.45
	Q_n				6.21			4.64			9.52			9.21
	IQ_n				1.34			4.86			3.86			10.09
512	VR_p	16	0.63	3.86	4.49	2.03	2.45	4.48	2.36	6.10	8.45	4.45	4.87	9.32
		32	0.19	4.40	4.58	1.73	2.45	4.17	1.17	6.53	7.70	4.06	5.04	9.10
	Q_n				5.68			4.32			9.29			8.90
	IQ_n				1.50			4.45			4.94			9.40

Data are generated from $x_t = \mu + x_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2)$, $\sigma_t^2 = .0001 + .8575\sigma_{t-1}^2 + .1171\varepsilon_{t-1}^2$

Table III. Power in Percentage against the Alternative of Random Walk + AR(1)

			5%						10%					
			Before transformation			After transformation			Before transformation			After transformation		
n		k	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power
128	VR_p	8	0.98	2.12	3.10	3.67	1.03	4.70	3.78	3.86	7.64	7.17	2.86	10.02
		16	0.09	1.97	2.06	3.63	0.91	4.54	1.66	3.17	4.83	7.37	2.20	9.57
	Q_n				3.94			5.52			6.96			10.59
	IQ_n				2.12			7.88			6.79			14.58
512	VR_p	16	5.70	0.21	5.91	10.95	0.13	11.08	13.53	4.95	14.02	18.32	0.35	18.67
		32	3.45	0.09	3.54	13.47	0.05	13.52	12.83	0.21	13.03	22.69	0.12	22.81
	Q_n				3.24			12.10			8.28			20.18
	IQ_n				7.13			19.24			19.97			31.91

Data are generated from $x_t = r_t + y_t, r_t = r_{t-1} + w_t$, $w_t \sim N(0, 0.5)$, $y_t = 0.96y_{t-1} + u_t$, $u_t \sim N(0, 1)$

Table IV. Power in Percentage against the Alternative of Random Walk + AR(1)

			5%						10%					
			Before transformation			After transformation			Before transformation			After transformation		
n		k	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power
128	VR_p	8	0.89	2.62	3.51	3.21	1.45	4.65	3.34	4.70	8.04	6.43	3.51	9.94
		16	0.09	2.53	2.61	3.28	1.19	4.46	1.43	4.28	5.71	6.57	2.80	9.37
	Q_n				4.22			5.22			7.69			10.16
	IQ_n				2.05			7.10			6.22			13.45
512	VR_p	16	3.77	0.51	4.28	7.58	0.31	7.89	9.56	1.07	10.62	13.32	0.79	14.11
		32	2.12	0.30	2.41	8.63	0.13	8.76	8.09	0.57	8.65	15.44	0.41	15.85
	Q_n				2.75			8.55			6.93			15.05
	IQ_n				5.12			13.79			14.48			23.45

Data are generated from $x_t = r_t + y_t, r_t = r_{t-1} + w_t$, $w_t \sim N(0,1)$, $y_t = 0.96y_{t-1} + u_t$, $u_t \sim N(0,1)$

Table V. Power in Percentage against the Alternative of Random Walk + AR(1)

			5%						10%					
			Before transformation			After transformation			Before transformation			After transformation		
n		k	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power
128	VR_p	8	0.81	3.20	4.01	2.92	1.82	4.73	3.04	5.43	8.47	5.91	4.05	9.95
		16	0.08	3.22	3.30	2.94	1.60	4.53	1.29	5.25	6.54	6.08	3.59	9.67
	Q_n				4.85			5.22			8.48			10.15
	IQ_n				1.86			6.64			5.70			12.60
512	VR_p	16	2.54	1.11	3.65	5.29	0.63	5.92	6.71	2.01	8.72	9.64	1.59	11.23
		32	1.20	0.83	2.02	5.68	0.47	6.15	5.42	1.67	7.09	10.21	1.09	11.30
	Q_n				2.97			6.29			6.45			11.76
	IQ_n				3.69			9.90			10.96			17.62

Data are generated from $x_t = r_t + y_t, r_t = r_{t-1} + w_t$, $w_t \sim N(0,2)$, $y_t = 0.96y_{t-1} + u_t$, $u_t \sim N(0,1)$

Table VI. Power in Percentage against the Alternative of AR(1)

			5%						10%					
			Before transformation			After transformation			Before transformation			After transformation		
n		k	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power	Lower	Upper	Power
128	VR_p	8	1.20	1.36	2.55	4.38	0.70	5.08	4.47	2.56	7.03	8.42	1.80	10.22
		16	0.10	1.05	1.15	4.67	0.42	5.08	2.18	1.93	4.10	9.52	1.24	10.76
	Q_n				3.09			6.10			5.87			11.32
	IQ_n				2.26			9.31			7.55			17.33
512	VR_p	16	13.07	0.04	13.10	22.44	0.01	22.45	26.39	0.11	26.50	33.95	0.07	34.02
		32	12.93	0.00	12.93	34.24	0.00	34.24	33.08	0.01	33.08	48.96	0.00	48.96
	Q_n				5.99			28.10			17.15			40.84
	IQ_n				16.68			40.39			39.07			56.85

Data are generated from $x_t = 0.96x_{t-1} + u_t$, $u_t \sim N(0,1)$