Spectral tests of the martingale hypothesis under conditional heteroscedasticity

Rohit S. Deo*

New York University, 8-57 KMEC, 44, West 4th Street, NY 10012, USA

Abstract

We study the asymptotic distribution of the sample standardized spectral distribution function when the observed series is a conditionally heteroscedastic martingale difference. We show that the asymptotic distribution is no longer a Brownian bridge but another Gaussian process. Furthermore, this limiting process depends on the covariance structure of the second moments of the series. We show that this causes test statistics based on the sample spectral distribution, such as the Cramér von-Mises statistic, to have heavily right skewed distributions, which will lead to over-rejection of the martingale hypothesis in favour of mean reversion. A non-parametric correction to the test statistics is proposed to account for the conditional heteroscedasticity. We demonstrate that the corrected version of the Cramér von-Mises statistic has the usual limiting distribution which would be obtained in the absence of conditional heteroscedasticity. We also present Monte Carlo results on the finite sample distributions of uncorrected and corrected versions of the Cramér von-Mises statistic. Our simulation results show that this statistic can provide significant gains in power over the Box–Ljung–Pierce statistic against long-memory alternatives. An empirical application to stock returns is also provided. © 2000 Elsevier Science S.A. All rights reserved.

JEL classification: C12

Keywords: Sample spectral distribution function; Martingale difference; Conditional heteroscedasticity; Cramér von-Mises statistic

* Tel.: + 1-212-998-0469; fax: + 1-212-995-4003.

E-mail address: rdeo@stern.nyu.edu (R.S. Deo).

1. Introduction

Economic theory suggests that many financial and economic time series like stock returns and exchange rate returns are uncorrelated. More specifically, the concept of market efficiency leads one to believe that future values of such series should be unpredictable given the past. This is the famous martingale difference hypothesis and it is a more serious restriction than mere absence of correlation. It implies that there is no non-trivial function of past data, linear or non-linear, which can be used to predict future values. Testing such a general hypothesis is however practically impossible, since it encompasses too many possibilities. A more realistic approach towards testing the martingale hypothesis has been through testing for the absence of correlation under various data-generating mechanisms.

Two popular tests for uncorrelatedness are the variance ratio test of Cochrane (1988) and the spectral based tests of Durlauf (1991). The spectral-based tests exploit the fact that under the null hypothesis of a martingale difference, the spectral distribution function is a straight line. Thus, deviations of the sample spectral distribution from the straight line may be used to test for the presence of correlation. Durlauf showed that under the null hypothesis of a martingale difference, a normalized version of the difference between the sample and theoretical standardized spectral distribution function converges to a Gaussian process. The asymptotic distributions of various functionals of this difference can then be obtained and tests for departure from the null of no correlation may be obtained. Durlauf (1991) obtained his limiting distribution under conditions which ruled out conditional heteroscedasticity. However, it is a well-accepted fact that most financial and economic series which are hypothesized to be martingale differences show conditional hetero-scedasticity. Thus, it is important to take this conditional heteroscedasticity into account when studying the behaviour of the spectral-based tests of the martin-gale difference hypothesis.

In this paper, we show that the spectral-based tests no longer have the usual limiting distribution when there is conditional heteroscedasticity. As a matter of fact, we show that, in general, the limiting distribution is heavily right skewed, with the amount of skewness depending on the degree of persistence in the second moments. This fact may explain why such tests tend to reject the martingale difference null in favour of mean reversion. We also suggest a way to correct these tests in a non-parametric way to account for the conditional heteroscedasticity. For one such corrected test, we prove that the asymptotic distribution is the same as what would be obtained in the absence of conditional heteroscedasticity. We provide Monte Carlo simulations for some of the uncor-rected and corrected tests and also provide an empirical application.

The layout of the paper is as follows. In Section 2, we state the assumptions we require for our results and discuss some popular models which satisfy these

assumptions. In Section 3, we derive the limiting distributions of spectral-based tests for the martingale difference hypothesis. Modified versions of the test statistics to account for the conditional heteroscedasticity are proposed and the limiting distribution for one of the modified test statistics is obtained. In Section 4, we present Monte Carlo simulation results for some of the uncorrected and corrected test statistics and in Section 5 apply these tests to real data. We finish the paper with a technical appendix containing the proofs of all our results.

2. Assumptions

We will derive the limiting behaviour of various spectral distribution tests under the hypothesis that the time series of interest, X_t, satisfies

$$X_t = \mu + \varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}$ is a martingale difference sequence and μ is some real number. Thus, X_t may be the first difference of a random walk with martingale difference innovations. Our specification allows the random walk to have a possible drift, implying a non-zero mean for the observed series X_t. The assumptions we make about the martingale difference series $\{\varepsilon_t\}$ in (1) are as follows:

Condition A. (i) $E(\varepsilon_t|F_{t-1}) = 0$, where $F_{t-1} = \sigma\{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$ is the sigma field generated by $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$.

(ii) $E(\varepsilon_t^2) = \sigma^2 < \infty$.

(iii) $\lim_{n\to\infty} n^{-1}\sum_{j=1}^{n} E(\varepsilon_j^2|F_{j-1}) = \sigma^2 > 0$ almost surely.

(iv) There exists a random variable U with $E(U^4) < \infty$ such that $P(|\varepsilon_t| > u) \leqslant cP(|U| > u)$ for some $0 < c < \infty$ and all t, all $u \geqslant 0$.

(v) $E(\varepsilon_t^2 \varepsilon_{t-r}\varepsilon_{t-s}) = \sigma^4 \tau_{rs}$ is finite and uniformly bounded for all t, $r \geqslant 1$, $s \geqslant 1$.

(vi) $\lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} \varepsilon_{t-r}\varepsilon_{t-s}E(\varepsilon_t^2|F_{t-1}) = \sigma^4 \tau_{rs}$ almost surely for any $r \geqslant 1$, $s \geqslant 1$.

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

(viii) $E(\varepsilon_t^8)$ is uniformly bounded for all t.

Assumptions (i)–(vi) are identical to the ones made in Hannan and Heyde (1972) and are essential for obtaining a central limit theorem for the sample autocorrelations. Assumptions (vii) and (viii) are required to obtain a functional limit theorem for the sample spectral distribution function in a random function space. This functional limit theorem then allows us to obtain the limiting distributions of a wide variety of popular goodness-of-fit tests of zero correlation in the frequency domain. Assumption (vii) essentially requires the product moments of the series $\{\varepsilon_t\}$ to behave similarly to those of an independent series. Note that Condition A does not require the series $\{\varepsilon_t\}$ to be strictly stationary (it does imply covariance stationarity). It also does not impose conditional homoscedasticity on the series $\{\varepsilon_t\}$.

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 of Condition A. 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

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

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

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. The validity of Condition A for a general GARCH(*p*, *q*) model can be demonstrated 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 $\omega > 0$, $\beta \geqslant 0$ and $\alpha > 0$. Furthermore, let α and β be such that $E\{\log_e(\beta + \alpha v_t^2)\} < 0$ and $E\{(\beta + \alpha v_t^2)^4\} < 1$. Then $\{\varepsilon_t\}$ satisfies the assumptions of Condition A.

The condition $E\{\log_e(\beta + \alpha v_t^2)\} < 0$ 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 $E\{(\beta + \alpha v_t^2)^4\} < 1$ will be satisfied by some non-empty subset of S. For example, values of α, β extremely close to the origin will certainly satisfy the second condition.

Assumption (viii) of Condition A, requiring the existence of at least eight moments for the martingale difference series $\{\varepsilon_t\}$ might seem strong considering that financial and economic series seem to exhibit thick tails. However, we feel that this assumption is essential to obtain a functional limit theorem for the sample spectral distribution function in a random function space. Furthermore, the existence of the eighth moment is not too restrictive if one can find a transformation $g(\cdot)$ such that $r_t = g(\varepsilon_t)$ satisfies Condition A when $\{\varepsilon_t\}$ itself has only a finite fourth moment. In such a situation, our results would then apply to the series $\{r_t\}$, which would be the series to be analysed. One such

transformation, suggested by one of the referees, is

$$r_t = |\varepsilon_t|^{1/2} \, \text{sign}(\varepsilon_t). \tag{4}$$

Supposing $\{\varepsilon_t\}$ were generated by the stochastic volatility model (2), where $\{v_t\}$ has a distribution which is symmetric around zero with only fourth moment finite. Then, by arguments similar to those used for Lemma 1, one can show that $\{r_t\}$ defined by (4) would satisfy Condition A.

To obtain our main results on the sample spectral distribution function of the process $\{X_t\}$, we need to know the limiting distribution of the sample autocorrelations of $\{X_t\}$. This is stated in the following theorem which follows directly from Theorem 2 of Hannan and Heyde (1972).

Theorem 2.1. Let assumptions (i)–(vii) *of Condition A hold. Define*

$$\bar{X} = n^{-1} \sum_{t=1}^{n} X_t,$$

$$\hat{\sigma}^2 = n^{-1} \sum_{t=1}^{n} (X_t - \bar{X})^2$$

and

$$\hat{\rho}_i = n^{-1} \hat{\sigma}^{-2} \sum_{t=1}^{n-i} (X_t - \bar{X})(X_{t+i} - \bar{X}), \qquad i \geqslant 1. \tag{5}$$

Then, for any finite fixed positive integer k, *we have*

$$n^{1/2} \hat{\boldsymbol{\rho}} \overset{D}{\to} \text{N}(\mathbf{0}, \, \boldsymbol{W}),$$

where $\hat{\boldsymbol{\rho}} = (\hat{\rho}_1, \, \hat{\rho}_2, \dots, \, \hat{\rho}_k)'$ *and* $\boldsymbol{W} = [w_{ij}]$ *is a* $k \times k$ *diagonal matrix with* $w_{ii} = \tau_{ii}$.

It should be noted that the normalized sample autocorrelations are not identically distributed under Condition A. Their asymptotic variance depends on the covariance in the second moments of the series $\{\varepsilon_t\}$ at the appropriate lag. For example, under the stochastic volatility model in (2), it can be easily shown that $\tau_{ii} = \exp\{4\, Cov(h_t, h_{t-i})\}$. Since there is no natural bound on the covariance of a stationary series, this implies that τ_{ii} (and hence the variance of $\hat{\rho}_i$) can be arbitrarily large under such a model. This anomalous behaviour of the sample autocorrelations arises due to the conditional heteroscedasticity that we are allowing in the series. The normalized sample autocorrelations will however have an asymptotic variance of 1 at all lags if the series $\{\varepsilon_t\}$ has a constant conditional variance. This can be seen from the fact that in such a case, $\tau_{ii} = \sigma^{-4} \text{E}(\varepsilon_t^2 \varepsilon_{t-i}^2) = \sigma^{-4} \text{E}\{\varepsilon_{t-i}^2 \text{E}(\varepsilon_t^2 | F_{t-i})\} = \sigma^{-4} \text{E}\{\varepsilon_{t-i}^2 \sigma^2\} = 1$.

Furthermore, it should also be noted that the asymptotic independence of $(\hat{\rho}_r, \hat{\rho}_s)$ for any $r > s > 0$ in Theorem 2.1 is entirely due to assumption (vii) of Condition A. This assumption implies that $\tau_{rs} = \sigma^{-4}E(\varepsilon_t^2 \varepsilon_{t-r} \varepsilon_{t-s}) = 0$ for any $r > s > 0$.

It is of interest to compare our Condition A and Theorem 2.1 with analogous assumptions and results in the seminal work of Durlauf (1991) regarding the sample spectral distribution function. The assumptions made in Durlauf (1991) are stated in his Definition 2.1. Durlauf's assumptions are identical to our assumptions (i)–(vi) and assumption (viii) of Condition A, and hence allow for conditional heteroscedasticity. Durlauf then states his Theorem 2.1, quoting Theorem 2 of Hannan and Heyde (1972), that the normalized sample autocorrelations of $\{X_t\}$ are both asymptotically independent and identically distributed with unit asymptotic variance at any lag. This application of Hannan and Heyde (1972) is incorrect. As demonstrated above, the sample correlations have variance depending on τ_{ii} in the presence of conditional heteroscedasticity and hence are not identically distributed. Furthermore, since Durlauf (1991) does not make any assumption similar to our assumption (vii) (which implies that $E(\varepsilon_t^2 \varepsilon_{t-r} \varepsilon_{t-s}) = 0$ for any $r > s > 0$), there is no guarantee that the sample correlations are asymptotically independent. Hence, the main results on the sample spectral distribution that Durlauf (1991) obtains in his Theorem 2.2 and subsequent Corollaries, which depend on his Definition 2.1 and Theorem 2.1, would not hold either in the presence of conditional heteroscedasticity or the absence of some restricted form of "independence" as defined through our assumption (vii).

We would also like to point out that our Theorem 2.1 requires assumption (vii) in Condition A to hold only for any $q \leqslant 4$. However, the stronger requirement that it hold for any $q \leqslant 8$ is essential in showing the tightness of the sample spectral distribution function to obtain our main result below.

In the next section, we study the asymptotic behaviour of the sample spectral distribution function in the presence of conditional heteroscedasticity. Our approach draws heavily on the work of Durlauf (1991).

3. Spectral-based tests of the martingale difference hypothesis

The correlation structure of a stationary time series is determined by its standardized spectral density defined by

$$f(\lambda) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \rho_h \cos \lambda h, \quad -\pi \leqslant \lambda \leqslant \pi,$$

where ρ_h is its correlation function at lag h. This theoretical standardized spectral density can be estimated based on the observed data $X_1, X_2, \ldots, X_n$ by

the sample standardized spectral density given by

$$I_{n,w}(\lambda) = (2\pi)^{-1} \sum_{j=-(n-1)}^{(n-1)} w_n(j)\hat{\rho}_j \cos \lambda j, \tag{6}$$

where $\hat{\rho}_j$ is the sample autocorrelation at lag j defined as in (5) above and $w_n(\cdot)$ is an appropriate sequence of weights symmetric about zero with $w_n(0) = 1$. If the form of the standardized spectral density function $f(\lambda)$ has been specified, then departures from it can be detected by studying the normalized cumulated deviations given by

$$U_{n,w}(t) = (2n)^{1/2} \int_0^{\pi t} (I_{n,w}(\omega) - f(\omega)) \, d\omega, \quad t \in [0, 1]. \tag{7}$$

Under some conditions on the weight sequence $w_n(\cdot)$ and assuming that $\{\varepsilon_t\}$ is a zero mean independent series, the result of Durlauf (1991) shows that $U_{n,w}(t)$ converges to a Brownian bridge. This result can then be used to obtain the limiting distributions of common goodness-of-fit test statistics like the Cramér von-Mises statistic, the Anderson Darling statistic, etc., which are all functionals of $U_{n,w}(t)$.

In this section, we show that when $\{\varepsilon_t\}$ is a conditionally heteroscedastic martingale difference, $U_{n,w}(t)$ no longer converges to a Brownian bridge but to another Gaussian process. To gain more insight into why this happens, we observe that under the null hypothesis of a martingale difference, all the correlations are zero and the standardized spectral density reduces to $(2\pi)^{-1}$. Thus, the normalized cumulated deviations reduce to

$$
\begin{aligned}
U_{n,w}(t) &= (2n)^{1/2} \int_0^{\pi t} \left(I_{n,w}(\omega) - \frac{1}{2\pi} \right) d\omega \\
&= \frac{\sqrt{2}}{\pi} \sum_{j=1}^{n-1} n^{1/2} \hat{\rho}_j w_n(j) \frac{\sin j\pi t}{j}, \quad t \in [0, 1].
\end{aligned} \tag{8}
$$

When $\{X_t\}$ is a conditionally homoscedastic martingale difference, we know that $n^{1/2}\hat{\rho}_j$ has asymptotically the same distribution as the sequence $\{\eta_j\}$, where $\{\eta_j\} \sim$ i.i.d. N(0, 1). Thus, assuming for now that $w_n(\cdot) \equiv 1$, we have heuristically for large n,

$$
\begin{aligned}
U_{n,w}(t) &= \frac{\sqrt{2}}{\pi} \sum_{j=1}^{n-1} n^{1/2} \hat{\rho}_j \frac{\sin j\pi t}{j}, \\
&\simeq \frac{\sqrt{2}}{\pi} \sum_{j=1}^{n-1} \eta_j \frac{\sin j\pi t}{j}, \\
&\simeq \frac{\sqrt{2}}{\pi} \sum_{j=1}^{\infty} \eta_j \frac{\sin j\pi t}{j}
\end{aligned}
$$

which is a Brownian bridge. This result breaks down when the sample correlations are asymptotically heteroscedastic, as happens in the case of a conditionally heteroscedastic series $\{\varepsilon_t\}$, yielding a different limiting process for $U_{n,w}(t)$. The new limiting process is given in the following Theorem. Henceforth, we will always assume that $U_{n,w}(t)$ is given by (8). i.e. under the null hypothesis of a martingale difference.

Theorem 3.1. *Assume Condition A holds. Furthermore, let the sequence of weights* $w_n(\cdot)$ *satisfy the following conditions*:
(i) $w_n(j)$ *is uniformly bounded in j and n for* $|j| \leqslant n$.
(ii) $\lim_{n \to \infty} w_n(j) = 1$ *for all fixed j.*
(iii) $w_n(0) = 1$ *and* $w_n(\cdot)$ *is symmetric around zero for all n.*
Then

$$U_{n,w}(t) \overset{\mathrm{D}}{\Rightarrow} U(t) \quad on\ t \in [0,\,1],$$

where

$$U(t) = \frac{\sqrt{2}}{\pi} \sum_{j=1}^{\infty} \eta_j \sqrt{\tau_{jj}}\, \frac{\sin j\pi t}{j}, \qquad t \in [0,\,1]$$

and $\{\eta_j\} \sim$ i.i.d. N(0, 1).

Note that the limiting distribution is invariant to the choice of the weight sequence $w_n(\cdot)$, since its effect washes out asymptotically. Thus, the choice of the weights only affects the small sample behaviour of $U_{n,w}$ and of any statistic depending on $U_{n,w}$.

On applying the continuous mapping theorem, we get the limiting distributions of various common spectral shape tests, which we state in the following Corollary.

Corollary 3.2. *Under the assumptions of Theorem 3.1, we have*
(i)   *Anderson darling statistic*

$$AD_n = \int_0^1 \frac{U_{n,w}^2(t)}{t(1-t)}\, \mathrm{d}t \overset{\mathrm{D}}{\to} \int_0^1 \frac{U^2(t)}{t(1-t)}\, \mathrm{d}t.$$

(ii) *Cramér von-Mises statistic*

$$CVM_n = \int_0^1 U_{n,w}^2(t)\, \mathrm{d}t \overset{\mathrm{D}}{\to} \int_0^1 U^2(t)\, \mathrm{d}t.$$

(iii) *Kolmogorov–Smirnov statistic*

$$KS_n = \sup_{t\in[0,1]} |U_{n,w}(t)| \overset{\mathrm{D}}{\to} \sup_{t\in[0,1]} |U(t)|.$$

(iv) *Kuiper statistic*

$$K_n = \sup_{[0\leqslant s,\ t\leqslant 1]} |U_{n,w}(t) - U_{n,w}(s)| \overset{\mathrm{D}}{\to} \sup_{[0\leqslant s,\ t\leqslant 1]} |U(t) - U(s)|.$$

From Corollary 3.2, it is clear that the distributions of the various test statistics depend crucially on the sequence $\{\tau_{ii}\}$ which is a measure of the dependence in the second moments of the series $\{X_t\}$. As mentioned earlier, for the stochastic volatility model in (2), it can be easily shown that $\tau_{ii} = \exp\{4\,Cov(h_t, h_{t-i})\}$. If the series $\{h_t\}$ is such that $Cov(h_t, h_{t-i})$ is positive, then the asymptotic variance of $\sqrt{n}\hat{\rho}_i$ will actually be bigger than 1, which is the asymptotic variance in the absence of conditional heteroscedasticity. To see the effect this has on the spectral-distribution-based test statistics, it is instructive to consider an alternative expression for one of them, the Cramér von-Mises statistic. When $w_n(\cdot) \equiv 1$, it is known (see for e.g., Anderson and You, 1996) that the Cramér von-Mises statistic can also be written as

$$CVM_n = \frac{1}{\pi^2} \sum_{j=1}^{n-1} \frac{1}{j^2} \left[\sqrt{n}\hat{\rho}_j\right]^2. \tag{9}$$

If the variance of $\sqrt{n}\hat{\rho}_i$ is bigger than 1, it is clear that the asymptotic distribution of CVM_n will have a thicker right tail compared to that of the usual distribution obtained in the absence of conditional heteroscedasticity. Furthermore, it is also clear that the rate of decay of $Cov(h_t, h_{t-i})$ to zero (and thus that of τ_{ii} to 1) will affect the thickness of the right tail, with a slower rate of decay leading to a thicker tail. Hence, using the usual cutoff points of the Cramér von-Mises statistic will lead to over-rejection of the martingale difference hypothesis in martingale difference series which show strong persistence in their second moments.

The other three test statistics in Corollary 3.2 also have distributions with thicker right tails under conditional heteroscedasticity due to the inflated variance of $U_{n,w}(t)$. The right tail of the limiting distribution of the Anderson Darling statistic is affected even more seriously than the Cramér von-Mises statistic by conditional heteroscedasticity. This is due to the fact that in the Anderson Darling statistic, the quantity $U_{n,w}^2(t)$, which has a larger variance when $\tau_{ii} > 1$, is weighted by $[t(1 - t)]^{-1}$ which gets large for values of t near 0 and 1.

It is clear from the above discussion that the limiting distributions of the spectral-based tests depend on the covariance of the second moments of the series. Thus, the critical values of these distribution will vary, depending on

the true parameters of the underlying process, making these tests infeasible from the practical point of view. To avoid this problem, we suggest the following non-parametric correction to the test statistics. To circumvent the dependence on τ_{ii}, we work with the following modified form of $U_{n,w}(t)$, given by

$$U_{n,w,C}(t) = \frac{\sqrt{2}}{\pi} \sum_{j=1}^{n-1} n^{1/2} \hat{a}_j w_n(j) \frac{\sin j\pi t}{j}, \qquad t \in [0, 1], \tag{10}$$

where

$$\hat{a}_j = \hat{\sigma}^2 \hat{\rho}_j \left[(n-j)^{-1} \sum_{t=1}^{n-j} (X_t - \bar{X})^2 (X_{t+j} - \bar{X})^2 \right]^{-1/2}. \tag{11}$$

In the proof of Theorem 3.3 below, we prove that for fixed j, $\hat{\sigma}^2 [(n-j)^{-1}\sum_{t=1}^{n-j} (X_t - \bar{X})^2 (X_{t+j} - \bar{X})^2]^{-1/2}$ is a consistent estimator of $\tau_{jj}^{1/2}$. It then follows from Theorem 2.1 that for any finite k, the collection of random variables $(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_k)'$ will be asymptotically independent and normal with zero mean and variance 1. Hence, we should expect the test statistics given in Corollary 3.2 but based on $U_{n,w,C}(t)$ to have the same limiting distributions as would be obtained if $\{X_t\}$ were an independent identically distributed white noise series. This will allow us to use tabulated critical values (see Anderson and You, 1996) of the standard limiting distributions when carrying out tests of the martingale hypothesis. In the following theorem, we state conditions under which such a result holds for the Cramér von-Mises test statistic based on $U_{n,w,C}(t)$.

Theorem 3.3. Let $U_{n,w,C}(t)$ be as in (10) and let the weights $w_n(\cdot)$ satisfy the conditions stated in Theorem 2.1. Let the assumptions of Condition A hold and assume that

$$\limsup_{n \to \infty} \sup_{1 \leqslant j \leqslant n} n\mathrm{E}(\hat{a}_j^2) < \infty. \tag{12}$$

Then

$$CVM_{n,C} \equiv \int_0^1 U_{n,w,C}^2(t)\,\mathrm{d}t \xrightarrow{\mathrm{D}} \int_0^1 B^2(t)\,\mathrm{d}t,$$

where $B(t)$ is a Brownian bridge of $[0, 1]$.

The next Corollary shows that under some conditions on the dependence in the second moments, Theorem 3.3 holds for the stochastic volatility model given in (2).

Corollary 3.4. Let the series $\{\varepsilon_t\}$ satisfy

$$\varepsilon_t = v_t \exp(h_t),$$

where $\{v_t\}$ is a sequence of independent $(0, \sigma_v^2)$ variables with $E(v_t^8) < \infty$. Let $h_t = \sum_{j=0}^{\infty} \alpha_j u_{t-j}$, where $\{u_t\}$ is a sequence of normally distributed $(0, \sigma_u^2)$ variables. Furthermore, let the coefficients α_j satisfy $|\alpha_j| \leqslant A\lambda^j$ for some positive constant A and some $0 < \lambda < 1$ and let the series $\{u_t\}$ and $\{v_t\}$ be independent. Then

$$CVM_{n,C} \equiv \int_0^1 U_{n,w,C}^2(t)\,dt \xrightarrow{D} \int_0^1 B^2(t)\,dt,$$

where $B(t)$ is a Brownian bridge on $[0, 1]$.

The condition that we have imposed on the coefficients $\{\alpha_j\}$ in Corollary 3.4 above will be satisfied if the series $\{h_t\}$ is a stationary autoregressive moving average (ARMA). However, the condition is not satisfied if $\{h_t\}$ is a stationary autoregressive fractionally integrated moving average (ARFIMA) or some analogous long-memory series. Furthermore, we have had to assume the finiteness of eight moments for the result to hold. Though this assumption allows for fairly thick-tailed distributions, it rules out infinite variance stochastic volatility models like those studied by de Vries (1991) and Deo (1997).

It is interesting to compare the Cramér von-Mises statistic given in Eq. (9) above with the Box–Ljung–Pierce statistic, which is another statistic commonly used to test for the presence of correlation. The Box–Ljung–Pierce statistic is given by

$$BL_n^K = \sum_{j=1}^{K} (\sqrt{n}\hat{\rho}_j)^2 \tag{13}$$

for some predetermined fixed positive integer K. If the underlying process is a conditionally homoscedastic martingale difference, then BL_n^K has asymptotically a χ^2 distribution with K degrees of freedom. However, the choice of the integer K naturally plays an important part in how well the finite sample distribution of the statistic approximates the limiting distribution and will generally result in a trade-off between the size and the power of the test. The Cramér von-Mises statistic has an advantage in this regard, since it takes into account all the $n-1$ sample autocorrelations which can be computed. From Eq. (9) above, it can be seen that the Cramér von-Mises statistic assigns the weight j^{-2} to the normalized sample autocorrelation at lag j. It is this declining set of weights which allow the Cramér von-Mises statistic to have a non-degenerate limiting distribution in spite of using all $n-1$ sample correlations. On the other hand, the Box–Ljung–Pierce statistic assigns the constant weight 1 to the normalized sample correlations at lag j, as seen from Eq. (13). As a result, the number of sample correlations used, K, must remain fixed as the sample size increases for the Box–Ljung–Pierce statistic to have a non-degenerate limiting distribution.

The fact that the Cramér von-Mises statistic uses all available $n - 1$ sample correlations might also be to its advantage in detecting long-memory time series. Long-memory series are those in which the correlations decay at a hyperbolic rate, as opposed to an exponential rate in short-memory time series. However, in a long-memory series, the individual correlations might all be small in magnitude, though decaying slowly to zero. Hence, one might expect the Cramér von-Mises statistic, which takes into account all sample correlations, to have superior power to the Box–Ljung–Pierce statistic in such cases. A similar comparison, naturally, can be made between the two test statistics when corrected for conditional heteroscedasticity. The corrected version of the Box–Ljung–Pierce statistic will be given by

$$BL_{n,C}^{K} = \sum_{j=1}^{K} (\sqrt{n}\hat{a}_j)^2,$$

where the $\hat{a}_j$ are given by Eq. (11) above. It is clear that under the assumptions of Condition A, for a fixed positive integer K, $BL_{n,C}^{K}$ will have an asymptotic χ^2 distribution with K degrees of freedom.

In the next section, we present Monte Carlo simulation results for both corrected and uncorrected versions of some of the test statistics considered above.

4. Simulation results

We conducted a simulation study to examine the size and power performance of some of the spectral based test statistics studied above. For both sample sizes $n = 100$ and 500, we generated 1000 realizations of the stochastic volatility model

$$X_t = v_t \exp(h_t),$$

$$h_t = \alpha h_{t-1} + 0.5u_t,$$

where $|\alpha| < 1$ and (u_t, v_t) are a sequence of independent bivariate normal random variables with zero mean and covariance matrix given by $diag(\sigma_u^2, 1)$. Note that this model is a martingale difference and satisfies the conditions of Corollary 3.4. The values of the pair (α, σ_u) that we used were (0.936, 0.424) and (0.951, 0.314). These are values which Shephard (1996, Table 1.6) obtained by fitting the above stochastic volatility model to real exchange rate data and thus reflect a practical situation. For each parameter configuration and sample size, we computed corrected and uncorrected versions of three test statistic. These were:

(i) The Cramér von-Mises statistic. The uncorrected version is denoted by CVM_n and the corrected version by $CVM_{n,C}$.

(ii) The Box–Ljung–Pierce statistic with $K = 5$. The uncorrected version is denoted by BL_n^5 and the corrected version by $BL_{n,C}^5$.

(iii) The Box–Ljung–Pierce statistic with $K = 15$. The uncorrected version is denoted by BL_n^{15} and the corrected version by $BL_{n,C}^{15}$.

In addition to the conditionally heteroscedastic data, we also studied the performance of these test statistics when the data was actually Gaussian white noise. This is necessary to see how the corrected statistics behave when the data is actually homoscedastic and the correction is unnecessary.

In Table 1, we compare the empirical sizes of the uncorrected versions of the statistics. The sizes were computed by comparing CVM_n, BL_n^5 and BL_n^{15} with the asymptotic 5% and 10% critical values of the Cramér von-Mises, the χ_5^2 and the χ_{15}^2 distributions, respectively. As is to be expected, the tests based on the uncorrected statistics are oversized when the data is a conditionally heteroscedastic martingale difference. This inflation in size can be quite severe and is greater for the larger sample size. It is interesting to note that among the three statistics, the Cramér von-Mises statistic suffers the least from the problem of size inflation. When the data are Gaussian white noise and therefore homoscedastic, all three statistics maintain approximately their nominal size, though BL_n^{15} is somewhat undersized when $n = 100$. This might explain why the size inflation in BL_n^{15} is lower than that in BL_n^5 when $n = 100$ but higher when $n = 500$.

In Table 2, we compare the sizes of the corrected versions of the statistics. As earlier, the sizes were computed by comparing $CVM_{n,C}$, $BL_{n,C}^5$, and $BL_{n,C}^{15}$ with the asymptotic 5% and 10% critical values of the Cramér von-Mises, the χ_5^2 and the χ_{15}^2 distributions, respectively. As can be seen from the table, the corrected statistics maintain their nominal size even under conditional heteroscedasticity though $BL_{n,C}^{15}$ is somewhat undersized when $n = 100$. Furthermore, it is reassuring to note that the corrected statistics retain the nominal size when the data is actually Gaussian white noise and the correction is unnecessary.

Table 1
Empirical sizes for uncorrected test statistics

n	Test	SV AR(1)		SV AR(1)		Gaussian white noise	
		$\alpha = 0.936$	$\sigma = 0.424$	$\alpha = 0.951$	$\sigma = 0.314$		
		5%	10%	5%	10%	5%	10%
	CVM_n	0.132	0.216	0.099	0.173	0.042	0.081
100	BL_n^5	0.179	0.265	0.137	0.208	0.039	0.080
	BL_n^{15}	0.148	0.216	0.123	0.186	0.036	0.067
	CVM_n	0.288	0.391	0.215	0.313	0.045	0.094
500	BL_n^5	0.405	0.500	0.325	0.414	0.053	0.100
	BL_n^{15}	0.472	0.586	0.398	0.496	0.047	0.092

Table 2
Empirical sizes for corrected test statistics

n	Test	SV AR(1)		SV AR(1)		Gaussian white noise	
		$\alpha = 0.936$ $\sigma = 0.424$		$\alpha = 0.951$ $\sigma = 0.314$			
		5%	10%	5%	10%	5%	10%
	$CVM_{n,C}$	0.036	0.083	0.038	0.087	0.041	0.087
100	$BL_{n,C}^5$	0.030	0.065	0.028	0.064	0.048	0.085
	$BL_{n,C}^{15}$	0.029	0.062	0.031	0.065	0.041	0.065
	$CVM_{n,C}$	0.048	0.090	0.057	0.092	0.038	0.094
500	$BL_{n,C}^5$	0.049	0.085	0.052	0.086	0.054	0.095
	$BL_{n,C}^{15}$	0.056	0.084	0.049	0.092	0.049	0.093

Table 3
Empirical power for uncorrected test statistics

n	Test	ARFIMA(0, d, 0)		ARFIMA(0, d, 0)		White noise + differenced AR(l)	
		$d = -0.10$		$d = 0.10$			
		5%	10%	5%	10%	5%	10%
	CVM_n	0.138	0.231	0.151	0.233	0.074	0.120
100	BL_n^5	0.067	0.146	0.103	0.168	0.055	0.103
	BL_n^{15}	0.055	0.101	0.076	0.115	0.014	0.077
	CVM_n	0.578	0.716	0.672	0.773	0.134	0.232
500	BL_n^5	0.416	0.571	0.581	0.679	0.132	0.236
	BL_n^{15}	0.233	0.367	0.453	0.566	0.105	0.192

In Tables 3 and 4, we compare the empirical power of the uncorrected and corrected versions of the three statistics, respectively. The power calculations were made when the data was generated by the following two alternative models:

(i) A fractionally integrated model (ARFIMA(0, d, 0)) given by $(1 - B)^d X_t = u_t$, where $\{u_t\}$ are i.i.d. N(0, 1). For the simulations, we chose two values of d, -0.1 and 0.1.

(ii) The sum of white noise and the first difference of a stationary autoregressive process of order one. i.e. $X_t = v_t + Y_t - Y_{t-1}$, where $Y_t = 0.85 Y_{t-1} + u_t$. The vector (u_t, v_t) was chosen to be a sequence of independent bivariate normal random variables with mean zero and variance covariance matrix given by $diag(\sigma_u^2, 1)$. The value of σ_u^2 was chosen such that the share of the variance of X_t due to the mean reverting component $Y_t - Y_{t-1}$, given by $2\sigma_u^2\{(1 + 0.85) + 2\sigma_u^2\}^{-1}$, was $\frac{1}{2}$.

Table 4
Empirical power for corrected test statistics

n	Test	ARFIMA(0, d, 0)		ARFIMA(0, d, 0)		White noise + differenced AR(l)	
		$d = -0.10$		$d = 0.10$			
		5%	10%	5%	10%	5%	10%
	$CVM_{n,C}$	0.117	0.204	0.158	0.247	0.056	0.117
100	$BL_{n,C}^{5}$	0.073	0.147	0.110	0.168	0.063	0.112
	$BL_{n,C}^{15}$	0.059	0.099	0.077	0.130	0.047	0.075
	$CVM_{n,C}$	0.550	0.698	0.691	0.787	0.140	0.245
500	$BL_{n,C}^{5}$	0.382	0.545	0.606	0.695	0.134	0.250
	$BL_{n,C}^{15}$	0.225	0.369	0.451	0.559	0.112	0.207

From Table 4 it is seen that among the three statistics, the corrected Cramér von-Mises statistic almost always has the highest power against all the three alternatives studied here. The gain in power for the Cramér von-Mises statistic is the most, however, against the fractionally integrated series, being as high as 17% when the sample size is 500. As discussed earlier, this was to be expected, since, for such models, the individual correlations are small in magnitude even though they decay very slowly to zero. Surprisingly, against the fractionally integrated alternative, the Box–Ljung–Pierce statistic at lag 15 has lower power than that at lag 5, even though it uses information from correlations upto a greater number of lags. Again, this might be due to the small magnitude of the individual correlations. When the alternative is the sum of white noise and the first difference of a stationary autoregressive process, both $CVM_{n,C}$ and $BL_{n,C}^{5}$ have comparable power, which is slightly higher than that of $BL_{n,C}^{15}$.

By comparing Tables 3 and 4, we also see that the power of the uncorrected and corrected versions of the statistics is virtually the same for the three alternatives considered. Thus, using the corrected statistics does not result in a loss of power when the data is conditionally homoscedastic.

From this small Monte Carlo study, it is clear that correcting the Cramér von-Mises statistic is essential when the data are conditionally hetero-scedastic martingale differences. Failure to account for the conditional heteroscedasticity can result in serious distortions in the size of the tests. Furthermore, there does not seem to be any distortion in either size or power, when using the corrected statistics for conditionally homoscedastic data. In addition, the corrected Cramér von-Mises statistic has much higher power than the Box–Ljung–Pierce statistic against fractionally integrated processes.

5. Application to stock prices

In this section, we apply our results to some time series of stock returns. The series that we analyse are weekly fluctuations for two CRSP NYSE-AMEX aggregate porfolios and also monthly returns on the CRSP-NYSE. Both of these data sets were analyzed in Durlauf (1991).

The first data sets consists of 1216 weekly returns on value-weighted and equal-weighted CRSP NYSE-AMEX portfolios from September 6, 1962 to December 26, 1985. Following Durlauf (1991), the weekly returns were computed using closing Wednesday prices. If the exchange was closed on a Wednesday, the Thursday price was used and if the exchange was also closed on Thursday, the previous Tuesday price was used. An examination of the sample autocorrelations of the squared returns (not presented here) showed that conditional heteroscedasticity is present in both the series. In Table 5 we report values of the CVM test statistic (CVM_n) as well as the corrected CVM statistic ($CVM_{n,C}$) for both these series. As can be seen from the table, we can reject the null hypothesis of zero correlation at the 1% level of significance for both the series, based upon the CVM statistic. The evidence is overwhelming in the case of the equal weighted returns. However, the corrected CVM statistics for both the series are much smaller. We are no longer able to reject the null hypothesis at the 1% level of significance for the value weighted returns. As a matter of fact, using the tables provided in Anderson and You (1996), one finds that the p-value for the corrected CVM statistic is between 2.5% and 5%. There is still strong evidence against the null for the equal-weighted returns however.

The second data set consisted of 780 monthly returns on the CRSP-NYSE value-weighted and equal-weighted portfolios from January 1926 to December 1990. As before, examination of the sample autocorrelations of the squared returns (not presented here) showed the presence of conditional heteroscedasticity in both the series. In Table 6, we report values of the CVM test statistic as well as the corrected CVM statistic for these series. The uncorrected CVM statistic leads to rejection of the null hypothesis at the 1% level of significance for both the series. However, as before, the corrected CVM statistics are much

Table 5
Tests for weekly CRSP returns (1962–1985)

	CVM_n	$CVM_{n,C}$
Value-weighted	0.857	0.533
Equal-weighted	11.213	5.988
Asymptotic critical values		
5%	0.46	
1%	0.74	

Table 6
Tests for monthly CRSP returns (1926–1990)

	CVM_n	$CVM_{n,C}$
Value-weighted	1.022	0.348
Equal-weighted	2.237	0.563
Asymptotic critical values		
5%	0.46	
1%	0.74	

smaller and we can no longer reject the null hypothesis at the 5% level for the value-weighted returns. There is some evidence against the null for the equal-weighted returns, though the associated *p*-value obtained from Anderson and You (1996) is greater than 2.5%.

6. Conclusions

We have shown that the distribution of the sample spectral distribution function for a white noise series is affected by the presence of conditional heteroscedasticity. The asymptotic distribution depends on the covariance structure of the second moments of the series. This causes test statistics based on the sample spectral distribution, such as the Cramér von-Mises statistic, to have heavily right skewed distributions, which will lead to over-rejection of the martingale hypothesis in favour of mean reversion. This phenomenon is confirmed by Monte Carlo simulations. A non-parametric correction to the test statistics is proposed to account for the conditional heteroscedasticity. The corrected version of the Cramér von-Mises statistic is shown to have an asymptotic distribution unaffected by conditional heteroscedasticity. A Monte Carlo study of the corrected version of the Cramér von-Mises statistic shows that the finite sample distribution behaviour is quite satisfactory for samples as small as a 100 observations. An empirical application to stock returns shows that evidence against the null hypothesis of the random walk can be considerably weakened after using the corrected test and accounting for conditional heteroscedasticity.

Acknowledgements

Appendix. (Technical)

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 $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ and $\{u_t\}$ is a sequence of independent standard normal variables. Furthermore, $\{u_t\}$ and $\{v_t\}$ will also be independent. It is trivial to check that $\{\varepsilon_t\}$ is a martingale difference and thus satisfies (i) of Condition A. Furthermore, by using the fact that $E\{\exp(aY)\} = \exp\{0.5a^2 \, Var(Y)\}$ for a zero mean Gaussian random variable Y, we get $\sigma_0^2 \equiv E(\varepsilon_t^2) = \sigma_v^2 \exp\{2\sum_{j=0}^{\infty} \alpha_j^2\} < \infty$, thus verifying condition (ii). Now, by Lemma 3.5.8 and Theorem 3.5.8 of Stout (1974), $z_t \equiv E(\varepsilon_t^2 | F_{t-1}) = \sigma_v^2 \exp(2\alpha_0^2) \exp(2\sum_{j=1}^{\infty} \alpha_j u_{t-j})$ is an ergodic sequence with mean $E(z_t^2) = \sigma_0^2 < \infty$. Hence, by Theorem 3.5.7 of Stout (1974), $\lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} z_t = \lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} E(\varepsilon_t^2 | F_{t-1}) = \sigma_0^2$, thereby satisfying (iii). Since $\{\varepsilon_t\}$ is a stationary sequence, (iv) is satisfied with $c = 1$ and $U = \varepsilon_1$. The stationarity of $\{\varepsilon_t\}$ coupled with the existence of its eighth moment also guarantees (v). By Lemma 3.5.8 and Theorem 3.5.8 of Stout (1974), $\varepsilon_{t-r}\varepsilon_{t-s}E(\varepsilon_t^2 | F_{t-1}) = \varepsilon_{t-r}\varepsilon_{t-s}z_t$ is a stationary ergodic series for any $r \geqslant 1$, $s \geqslant 1$ and hence by Theorem 3.5.7 of Stout (1974), assumption (vi) is also satisfied. The fact that $\{v_t\}$ is an independent zero mean series with finite eighth moment and also independent of $\{h_t\}$ guarantees that assumptions (vii) and (viii) are met. Hence, the stochastic volatility model (2) satisfies Condition A.

Proof of Lemma 2. Under the conditions of Lemma 2, it follows by Theorem 2 and the Corollary to Theorem 3 of Nelson (1990), that ε_t is a stationary ergodic martingale difference with finite eighth moments. Thus, $\{\varepsilon_t\}$ immediately satisfies assumptions (i), (ii), (iv), (v) and (viii) of Condition A. By Theorem 2 of Nelson (1990), σ_t^2 is stationary and ergodic and can be expressed as

$$\sigma_t^2 = \omega\left[1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} (\beta + \alpha v_{t-i}^2) \right]. \tag{A.1}$$

Hence, by Theorem 3.5.7 of Stout (1974), assumptions (iii) and (vi) of Condition A are also satisfied. It remains now to show that assumption (vii) of Condition A also holds for this model. For any q, $2 \leqslant q \leqslant 8$, consider $E(\prod_{i=1}^{q} \varepsilon_{t_i}^{s_i})$, where the s_i are non-negative integers such that at least one s_i is exactly one and $\sum_i s_i \leqslant 8$. We assume without loss of generality that $t_1 > t_2 > \cdots > t_q$. Then, there is some p, $1 \leqslant p \leqslant q$, such that $s_i = 2j_i$ for $1 \leqslant i < p$ and $s_p = 2j_p + 1$, where j_i are non-negative integers. Thus,

$$E\left(\prod_{i=1}^{q} \varepsilon_{t_i}^{s_i} \right) = E\left(\left\{ \prod_{i=p+1}^{q} \varepsilon_{t_i}^{s_i} \right\} E\left[\prod_{i=1}^{p} \varepsilon_{t_i}^{s_i} \mid F_{t_p - 1} \right] \right). \tag{A.2}$$

But

$$
\mathrm{E}\left[\prod_{i=1}^{p} \varepsilon_{t_i}^{s_i} \mid F_{t_p - 1}\right] = \mathrm{E}\left[\left(\prod_{i=1}^{p-1} \varepsilon_{t_i}^{2j_i}\right)\varepsilon_{t_p}^{2j_p + 1}\middle| F_{t_p - 1}\right]
$$

$$
= \sigma_{t_p}^{2j_p + 1}\mathrm{E}\left[\left(\prod_{i=1}^{p-1} \varepsilon_{t_i}^{2j_i}\right)v_{t_p}^{2j_p + 1} \mid F_{t_p - 1}\right]. \tag{A.3}
$$

From (A.1) and (A.3) it is clear that every term in $(\prod_{i=1}^{p-1}\varepsilon_{t_i}^{2j_i})v_{t_p}^{2j_p + 1}$ consists of an odd power of v_{t_p} multiplied by some function of all the other v_t. Each such term has an expectation of zero since the process $\{v_t\}$ is independent and symmetric around zero. Furthermore, the conditional expectation of all such terms is still zero by independence. Hence, $\mathrm{E}[\prod_{i=1}^{p}\varepsilon_{t_i}^{s_i} \mid F_{t_p - 1}] = 0$, which by (A.2) implies that (vii) of Condition A holds.

Proof of Theorem 3.1. Since $U_{n,w}(\cdot)$ is location and scale invariant, we will assume henceforth, without loss of generality, that $\mathrm{E}(X_t) = 0$ and $Var(X_t) = 1$. Note that $U_{n,w}$ can be written as

$$
U_{n,w}(t) = \frac{1}{\hat{\sigma}^2} \frac{\sqrt{2}}{\pi} \sum_{j=1}^{n-1} n^{1/2}\hat{\gamma}_j w_n(j) \frac{\sin j\pi t}{j}, \quad t\in[0, 1],
$$

where $\hat{\gamma}_j = \hat{\sigma}^2 \hat{\rho}_j$. We will first prove the weak convergence of $\hat{\sigma}^2 U_{n,w}(t)$ to $U(t)$. Since $p\lim_{n\to\infty} \hat{\sigma}^2 = 1$, the result will then also hold for $U_{n,w}(t)$.

Expanding $\hat{\gamma}_j$ for every j, we can express $\hat{\sigma}^2 U_{n,w}$ as

$$
\hat{\sigma}^2 U_{n,w}(t) = \sum_{p=1}^{4} T_{n,p}(t), \quad t\in[0, 1], \tag{A.4}
$$

where

$$
T_{n,p}(t) = \frac{\sqrt{2}}{\pi} \sum_{j=1}^{n-1} n^{1/2}\hat{c}_{j,p} w_n(j) \frac{\sin j\pi t}{j}
$$

and $\hat{c}_{j,1} = n^{-1}\sum_{t=j+1}^{n} X_t X_{t-j}$, $\hat{c}_{j,2} = n^{-1}\bar{X}\sum_{t=j+1}^{n} X_t$, $\hat{c}_{j,3} = n^{-1}\bar{X}\sum_{t=1}^{n-j} X_t$ and $\hat{c}_{j,4} = n^{-1}(n - j)\bar{X}^2$. Our first step is to prove the weak convergence of $T_{n,1}(t)$ to $U(t)$. In order to do this, it is sufficient (see Theorems 8.1 and 8.2 of Billingsley, 1968) to establish the following three conditions:

(i) For any $\delta > 0$, $\exists\ \xi > 0$ such that $\mathrm{P}(|T_{n,1}(0)| > \delta) < \xi\ \forall n \geqslant 1$.
(ii) $(T_{n,1}(t_1), T_{n,1}(t_2), \ldots, T_{n,1}(t_k)) \xrightarrow{\mathrm{D}} (U(t_1), U(t_2), \ldots, U(t_k))$ for any finite set $(t_1, t_2, \ldots, t_k)$.
(iii) For every $\gamma >$ and $\eta > 0$, $\exists\varepsilon\in(0, 1)$ and an integer N_0 such that $\mathrm{P}(\sup_{|p-t| < \varepsilon}|T_{n,1}(t) - T_{n,1}(p)| > \gamma) < \eta$ for all $n > N_0$.

Condition (i) holds trivially. We will demonstrate condition (ii) only for $k = 1$ since the argument for general k follows by applying the Cramér–Wold device. To prove condition (ii), we write $T_{n,1}(t)$ as

$$T_{n,1}(t) = \frac{\sqrt{2}}{\pi} \sum_{j=1}^{s} n^{1/2} \hat{c}_{j,1} w_n(j) \frac{\sin j\pi t}{j} + \frac{\sqrt{2}}{\pi} \sum_{j=s+1}^{n-1} n^{1/2} \hat{c}_{j,1} w_n(j) \frac{\sin j\pi t}{j}$$

$$\equiv T_{n,1}^s(t) + R_{n,1}^s(t) \tag{A.5}$$

for some integer s. From Theorem 2.1 and the fact that $w_n(j) \to 1$ for fixed j, it follows that for any fixed s,

$$T_{n,1}^s(t) \overset{D}{\to} T_1^s(t) \equiv \frac{\sqrt{2}}{\pi} \sum_{j=1}^{s} \eta_j \sqrt{\tau_{jj}} \frac{\sin j\pi t}{j}. \tag{A.6}$$

Furthermore, though the series $\{X_t\}$ is not independent, assumption (vii) of Condition A and the boundedness of $w_n(\cdot)$ allow us to exactly retrace the steps of Theorem 1 of Grenander and Rosenblatt (1957, p. 188) and conclude that for sufficiently large s, $R_{n,1}^s(t)$ is small in probability uniformly in n and t. More specifically, given $\delta > 0$ and $\eta > 0$, there exists an s and an N_0 such that for all $n > N_0$,

$$P\left(\sup_{t \in [0,1]} |R_{n,1}^s(t)| > \delta \right) < \eta. \tag{A.7}$$

Finally, we also have

$$T_1^s(t) \overset{D}{\to} U(t) \tag{A.8}$$

as $s \to \infty$. Eqs. (A.5)–(A.8) allow us to apply Proposition 6.3.9 of Brockwell and Davis (1991) and conclude that condition (ii) is satisfied. To prove condition (iii), we note that

$$P\left(\sup_{|p-t|<\varepsilon} |T_{n,1}(t) - T_{n,1}(p)| > \gamma \right)$$

$$\leqslant P\left(\sup_{|p-t|<\varepsilon} |T_{n,1}^s(t) - T_{n,1}^s(p)| > \gamma \right)$$

$$+ P\left(\sup_{|p-t|<\varepsilon} |R_{n,1}^s(t) - R_{n,1}^s(p)| > \gamma \right). \tag{A.9}$$

Given $\gamma > 0$ and $\eta > 0$, it follows from Eq. (A.7) that there exists an s sufficiently large and an N_1 such that the second term on the right-hand side of (A.9) is less than $\eta/2$ for all $n > N_1$. For this fixed s, the sequence of probability measures

associated with $T_{n,1}^s(\cdot)$ is tight and hence the first term on the right-hand side of (A.9) can be made smaller than $\eta/2$ for all $n > N_2$, for some N_2. Thus, condition (iii) holds for all $n > \max(N_1, N_2)$ and it follows that

$$T_{n,1}(t) \Rightarrow U(t), \quad t \in [0, 1]. \tag{A.10}$$

We now proceed to prove that for $2 \leqslant p \leqslant 4$, $T_{n,p}(t)$ converges weakly to zero. Since $\sqrt{n}\bar{X} = O_p(1)$, it suffices to prove the result for $n^{-1/2}\bar{X}^{-1}T_{n,p}(t)$. We will only demonstrate the proof for $p = 2$ since the same method applies in the other cases. The result is obtained by verifying the three conditions stated earlier. Condition (i) is again trivially satisfied. As before, we will prove condition (ii) only for $k = 1$ since the general result follows by applying the Cramér–Wold device. For some integer s, we can write $Q_n(t) = n^{-1/2}\bar{X}^{-1}T_{n,2}(t)$ as

$$Q_n(t) = \frac{\sqrt{2}}{\pi} \sum_{j=1}^{s} n^{-1/2}\hat{b}_j w_n(j) \frac{\sin j\pi t}{j} + \frac{\sqrt{2}}{\pi} \sum_{j=s+1}^{n-1} n^{-1/2}\hat{b}_j w_n(j) \frac{\sin j\pi t}{j}$$

$$\equiv Q_n^s(t) + R_n^s(t),$$

where $\hat{b}_j = n^{-1/2}\sum_{t=j+1}^{n} X_t$. Since $\hat{b}_j = O_p(1)$, it follows that for any fixed s,

$$p \lim_{n \to \infty} Q_n^s(t) = 0. \tag{A.11}$$

Also, assumption (vii) of Condition A implies that $E(\hat{b}_j^4) = O(1)$. This fact and the Cauchy–Schwarz inequality imply that

$$E(\hat{b}_i \hat{b}_j \hat{b}_k \hat{b}_l) = O(1) = O(n) \tag{A.12}$$

for any i, j, k, l. Once again we can retrace the steps in Theorem 1 of Grenander and Rosenblatt (1957, p. 188) and using (A.12) and the boundedness of $w_n(\cdot)$ conclude that (A.7) holds for $R_n^s(t)$. By using Proposition 6.3.9 of Brockwell and Davis (1991), we conclude that $p \lim_{n \to \infty} Q_n(t) = 0$. Finally, condition (iii) can be shown for $Q_n(t)$ in a manner similar to the one used for $T_{n,1}(t)$. Thus, we have the weak convergence of $Q_n(t)$ (and hence of $T_{n,2}(t)$) to zero.

Proof of Theorem 3.3. It is possible (see Anderson and You, 1996) to express $CVM_{n,C}$ as

$$CVM_{n,C} = \frac{2}{\pi^2} \sum_{j=1}^{n-1} n \frac{\hat{a}_j^2 w_n^2(j)}{j^2} = \frac{2}{\pi^2} \sum_{j=1}^{s} n \frac{\hat{a}_j^2 w_n^2(j)}{j^2} + \frac{2}{\pi^2} \sum_{j=s+1}^{n-1} n \frac{\hat{a}_j^2 w_n^2(j)}{j^2}$$

$$\equiv T_n^s + R_n^s,$$

where s is sone integer. By assumptions (iii) and (iv) of Condition A, we have $p \lim_{n \to \infty} \hat{\sigma}^2 = \sigma^2$. See Tanaka (1996, pp. 81, 82). A similar argument, used in

conjunction with assumptions (iv) and (vi) also gives, for fixed j, $p\lim_{n\to\infty}(n-j)^{-1}\sum_{t=1}^{n-j}(X_t-\bar{X})^2(X_{t+j}-\bar{X})^2 = E(\varepsilon_t^2\varepsilon_{t+j}^2)$. Using these results, it follows from Theorem 2.1 and the fact that $w_n(j)\to 1$ for fixed j, that for fixed s,

$$T_n^s \overset{\text{D}}{\to} \frac{2}{\pi^2}\sum_{j=1}^{s}\frac{\eta_j^2}{j^2},$$

where $\{\eta_j\}\sim$ i.i.d. N(0, 1). Also,

$$\frac{2}{\pi^2}\sum_{j=1}^{s}\frac{\eta_j^2}{j^2} \overset{\text{D}}{\to} \frac{2}{\pi^2}\sum_{j=1}^{\infty}\frac{\eta_j^2}{j^2}$$

as $s\to\infty$. Furthermore, it follows from (12) and the boundedness of $w_n(\cdot)$, that for any $\delta>0$

$$P(|R_n^s|>\delta) \leqslant \frac{2}{\pi^2}\sum_{j=s+1}^{n-1}n E(\hat{a}_j^2)\frac{w_n^2(j)}{j^2} \leqslant M\sum_{j=s+1}^{n-1}j^{-2}$$

for some constant M. Hence,

$$\lim_{s\to\infty}\lim_{n\to\infty}P(|R_n^s|>\delta)=0.$$

From Proposition 6.3.9 of Brockwell and Davis (1991), we thus get

$$CVM_{n,C} \overset{\text{D}}{\to} \frac{2}{\pi^2}\sum_{j=1}^{\infty}\frac{\eta_j^2}{j^2}.$$

However

$$\frac{2}{\pi^2}\sum_{j=1}^{\infty}\frac{\eta_j^2}{j^2} = \int_0^1 B^2(t)\,dt$$

with probability 1, where $B(\cdot)$ is a Brownian bridge on [0, 1]. See Eq. (3.8) of Anderson (1993).

Proof of Corollary 3.4. Since the test statistic is location and scale invariant, we will assume throughout this proof, without loss of generality, that $E(X_t)=0$ and $Var(X_t)=1$. By Lemma 1, the assumptions of Condition A are satisfied by this model. To prove the result, we thus have to only verify that (12) of Theorem 3.3 holds for the process. For $n-\sqrt{n}<j\leqslant n-1$, we get upon applying the Cauchy Schwarz inequality

$$n\hat{a}_j^2 = \frac{n-j}{n}\frac{[\sum_{t=1}^{n-j}(X_t-\bar{X})(X_{t+j}-\bar{X})]^2}{\sum_{t=1}^{n-j}(X_t-\bar{X})^2(X_{t+j}-\bar{X})^2} \leqslant \frac{n-j}{n}(n-j)\leqslant 1.$$

Thus, we have

$$\sup_{n-\sqrt{n}<j\leqslant n-1} n E(\hat{a}_j^2) < 1. \tag{A.13}$$

The theorem is proved if

$$\sup_{1 \leqslant j \leqslant n - \sqrt{n}} n\mathrm{E}(\hat{a}_j^2) < M < \infty \tag{A.14}$$

for some constant M, which we show next. To do this, we verify that the conditions of Theorem 5.4.3 of Fuller (1996) hold. We can write $n\hat{a}_j^2$ as

$$n\hat{a}_j^2 = (n - j)f(Z_{j,n}, Y_{j,n}), \tag{A.15}$$

where $f(x, y) = y^{-1}x^2$ for $x^2 \leqslant y$,

$$Z_{j,n} = \frac{1}{\sqrt{n(n - j)}} \sum_{t=1}^{n-j} (X_t - \bar{X})(X_{t+j} - \bar{X})$$

and

$$Y_{j,n} = \frac{1}{n - j} \sum_{t=1}^{n-j} (X_t - \bar{X})^2 (X_{t+j} - \bar{X})^2.$$

Note that by the Cauchy–Schwarz inequality, $f(Z_{j,n}, Y_{j,n})$ is a bounded function. By expanding the product in the numerator of $Z_{j,n}$, noting that $\mathrm{E}(\sum_{t=1}^{l} X_t)^2 = \mathrm{O}(l)$ and applying the Cauchy–Schwarz inequality, it follows that

$$\mathrm{E}(Z_{j,n}^2) = \mathrm{O}\left(\frac{1}{n - j}\right) \tag{A.16}$$

for $1 \leqslant j \leqslant n - \sqrt{n}$.

Expanding the squared term in the numerator of $Y_{j,n}$ and letting $\tau_j = \mathrm{E}(X_t^2 X_{t-j}^2)$, we can write $Y_{j,n} - \tau_j$ as

$$Y_{j,n} - \tau_j = \frac{1}{n - j} \sum_{t=1}^{n-j} (X_t^2 X_{t+j}^2 - \tau_j) + R_{j,n},$$

where

$$R_{j,n} = \frac{1}{n - j}$$

$$\times \left\{ \begin{array}{l} \bar{X}^2 \sum_{t=1}^{n-j} X_t^2 - 2\bar{X} \sum_{t=1}^{n-j} X_t^2 X_{t+j} + \bar{X}^2 \sum_{t=1}^{n-j} X_{t+j}^2 - 2\bar{X}^3 \sum_{t=1}^{n-j} X_{t+j} \\ + \bar{X}^4(n - j) - 2\bar{X} \sum_{t=1}^{n-j} X_t X_{t+j}^2 - 2\bar{X}^3 \sum_{t=1}^{n-j} X_t + 4\bar{X}^2 \sum_{t=1}^{n-j} X_t X_{t+j} \end{array} \right\}$$

is the remainder term. We first demonstrate that

$$\mathrm{E}(R_{j,n}^2) = \mathrm{O}\left(\frac{1}{n - j}\right). \tag{A.17}$$

By assumption (vii) of Condition A, it follows that both $E(\sum_{t=1}^{n-j} X_t^2 X_{t+j})^2$ and $E(\sum_{t=1}^{n-j} X_t X_{t+j}^2)^2$ are $O(n-j)$. By assumption (viii) of Condition A, it follows that both $E(\sum_{t=1}^{n-j} X_{t+j}^2)^2$ and $E(\sum_{t=1}^{n-j} X_t^2)^2$ are $O((n-j)^2)$ Also, $E(\bar{X}^{2s}) = O(n^{-s})$ for $1 \leqslant s \leqslant 4$. These facts together with the Cauchy–Schwarz inequality imply Eq. (A.17). We next show that

$$E\left(\frac{1}{n-j}\sum_{t=1}^{n-j}(X_t^2 X_{t+j}^2 - \tau_j)\right)^2 = O\left(\frac{1}{n-j}\right). \tag{A.18}$$

We have

$$E\left(\frac{1}{n-j}\sum_{t=1}^{n-j}(X_t^2 X_{t+j}^2 - \tau_j)\right)^2$$

$$= O\left(\frac{1}{n-j}\right)$$

$$+ \frac{2}{(n-j)^2}\sum_{t=1}^{n-j-1}\sum_{l=t+1}^{n-j} E(X_t^2 X_{t+j}^2 X_l^2 X_{l+j}^2 - \tau_j^2). \tag{A.19}$$

It can be easily shown that under the assumed model, $\tau_j = \sigma_v^4 \exp(4\sigma_h^2 + 4\gamma_j)$, where $\sigma_v^2 = Var(v_t)$, $\sigma_h^2 = Var(h_t)$ and $\gamma_j = Cov(h_t, h_{t-j})$. Furthermore, one can also show that

$$E(X_t^2 X_{t+j}^2 X_l^2 X_{l+j}^2) = \sigma_v^8 \exp(8\sigma_h^2 + 8\gamma_j + 8\gamma_{l-t} + 4\gamma_{|l-t-j|} + 4\gamma_{l-t+j}).$$

Thus,

$$|E(X_t^2 X_{t+j}^2 X_l^2 X_{l+j}^2 - \tau_j^2)|$$

$$= \sigma_v^8 \exp(8\sigma_h^2 + 8\gamma_j)|\exp(8\gamma_{l-t} + 4\gamma_{|l-t-j|} + 4\gamma_{l-t+j}) - 1|$$

$$\leqslant M_1|8\gamma_{l-t} + 4\gamma_{|l-t-j|} + 4\gamma_{l-t+j}|$$

$$= O(\lambda^{l-t}) + O(\lambda^{|l-t-j|}) + O(\lambda^{l-t+j}) \tag{A.20}$$

for some constant M_1, where the last inequality follows from the assumption that $\alpha_j = O(\lambda^j)$ for some $0 < \lambda < 1$. Using the bound (A.20) in Eq. (A.19) gives us

$$E\left(\frac{1}{n-j}\sum_{t=1}^{n-j}(X_t^2 X_{t+j}^2 - \tau_j)\right)^2 = O\left(\frac{1}{n-j}\right)$$

which in conjunction with (30) proves that

$$E(Y_{j,n} - \tau_j)^2 = O\left(\frac{1}{n-j}\right). \tag{A.21}$$

From Eqs. (A.16) and (A.21), it follows that condition (i) of Theorem 5.4.3 of Fuller (1996) is satisfied with $a_{n,j} = (n-j)^{-1/2}$, $\alpha = 1$ and $s = 2$. Since

$\tau_j = \sigma_v^4 \exp(4\sigma_h^2 + 4\gamma_j)$, we have $0 < C_1 < \inf_{j \geq 1} \tau_j \leq \sup_{j \geq 1} \tau_j < C_2 < \infty$ for some constants C_1 and C_2. Defining the set S_j by $S_j = \{x, y : |x| \leq 1, |y - \tau_j| \leq C_1\}$ and noting that the function $f(x, y)$ is a bounded function, we see that the remaining conditions of Theorem 5.4.3 of Fuller (1996) are satisfied and thus

$$\mathrm{E} f(Z_{j,n}, Y_{j,n}) = \mathrm{O}\left(\frac{1}{n-j}\right) \tag{A.22}$$

for $1 \leq j \leq n - \sqrt{n}$. Eq. (A.14) now follows from Eqs. (A.22) and (A.15).

References

Anderson, T.W., 1993. Goodness of fit tests for spectral distributions. Annals of Statistics 21, 830–847.

Anderson, T.W., You, L., 1996. Adequacy of asymptotic theory for goodness-of-fit criteria for spectral distributions. Journal of Time Series Analysis 17, 533–552.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Bougerol, P., Picard, N., 1992. Stationarity of GARCH processes and of some non-negative time series. Journal of Econometrics 52, 115–127.

Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods, 2nd Edition. Springer, New York.

Cochrane, J.H., 1988. How big is the random walk in GNP? Journal of Political Economy 96, 893–920.

Deo, R.S., 1997. Conditionally heteroscedastic stable processes. Working paper, New York University.

de Vries, C.G., 1991. On the relation between GARCH and stable processes. Journal of Econometrics 48, 313–324.

Durlauf, S.N., 1991. Spectral based testing of the martingale hypothesis. Journal of Econometrics 50, 355–376.

Fuller, W.A., 1996. Introduction to Statistical Time Series, 2nd Edition. Wiley, New York.

Grenander, U., Rosenblatt, M., 1957. Statistical Analysis of stationary Time Series. Chelsea, New York.

Hannan, E.J., Heyde, C.C., 1972. On limit theorems of quadratic functions of a discrete time series. Annals of Mathematical Statistics 43, 2058–2066.

Nelson, D., 1990. Stationarity and persistence in the GARCH(1, 1) model. Econometric Theory 6, 318–334.

Shephard, N., 1996. Statistical aspects of ARCH and stochastic volatility. In: Cox, D.R., et al. (Ed.), Time Series Models in Econometrics, Finance and Other Fields. Chapman & Hall, London.

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

Tanaka, K., 1996. Time Series Analysis. Wiley, New York.