A GMM Skewness and Kurtosis Ratio Test for Higher Moment Dependence

Woon K. Wong*

Cardiff Business School

Aberconway Building, Colum Drive, Cardiff CF10 3EU, UK.

10 November 2018

*Corresponding details: Tel: +44 29 20874900; Email: wongwk3@cardiff.ac.uk

A GMM Skewness and Kurtosis Ratio Test for Higher Moment Dependence

Abstract

This article extends the variance ratio test of Lo and MacKinlay (1988) to tests of skewness and kurtosis ratios using the generalized methods of moments. In particular, overlapping observations are used in which dependencies are explicitly modelled to make the tests more powerful and have better size properties. The proposed higher order ratio tests can be useful in risk management where risk models are estimated using daily data but multiperiod forecasts of tail risks are required for the determination of risk capital. Application of the tests finds significant higher moment dependence in the US stock market returns.

Keywords: Skewness, kurtosis, overlapping observations, moments, cumulants

JEL Classification: C10, G11

1 Introduction

This article extends the variance ratio test of Lo and MacKinlay (1988) to tests of skewness and kurtosis ratios. Specifically, under the independently and identically distributed (IID) assumption, the skewness and kurtosis of single-period returns are respectively $\sqrt{h}$ and h times the corresponding third- and fourth-moment statistics of h-period returns. One challenge to testing the validity of these relationships is the use of higher order statistics which are associated with large estimation errors. To circumvent the estimation problem, this paper employs the GMM approach used by Richardson and Smith (1991) in which the dependencies of overlapping observations are explicitly modelled in order to obtain the required weighting matrix for the test of higher order moments. Monte-Carlo simulations show that the analytically derived weighting matrix is significantly better than the popular Newey-West covariance matrix, for the former fully utilizes information from the data thereby giving rise to more powerful tests with relatively good size properties.

The use of higher-order statistics for testing nonlinearity can be traced back to Subba Rao and Gabr (1980) and Hinich (1982). The tests apply Fourier transform to third order covariances to obtain the bispectrum which varies with frequency if nonlinear dependence is present in the time series. Wong (1997) later extends the bispectral test from univariate to multivariate time series and shows that the component of a linear non-Gaussian multivariate process cannot be represented as a linear time series. Based on the bispectral test, Hinich and

Patterson (1985) find evidence of nonlinearity in the US daily stock returns. The bispectral test is also used together with other tests such as the BDS test of Brock *et al.* (1986), the neural network test of Lee *et al.* (1993), the Lyapunov exponent test of McCaffrey *et al.* (1992) and so on to find evidence of chaos due to deterministic nonlinearity in financial time series. Instead of chaos, Abhyankar *et al.* (1995, 1997) find heteroscedasticity as the main stylized feature in the S&P 500 and other major stock indices. Given the ubiquity of heteroscedasticity in financial processes, the squared residual autocorrelation tests of McLeao and Li (1983) and Li and Mak (1994) become the popular nonlinearity tests in the form of diagnostic checks for the residuals of GARCH-type processes; see for example Hsieh (1989) and Tse and Tsui (2002).

The proposed higher-order ratio tests can be a useful complement in relation to the existing nonlinear dependence tests. First, zero squared residual autocorrelations do not necessarily imply that the skewness-kurtosis ratio relations would hold. On the other hand, if higher-order ratios fail to hold, it will be of interest to investigate which higher-moment dependence is the cause as such information can be relevant to various financial applications. For example, correlation between squares of price innovations would render the kurtosis ratio invalid and such dynamics can be captured by GARCH models. Another example could take the form of a higher volatility followed by a higher price shock in the next period; such nonlinear relationship would give rise to a higher than expected multiperiod-return skewness and is related to the asset pricing literature specified by the GARCH-in-mean models. To this end, some relevant higher-moment dependence t-statistics in association with the proposed ratio tests are suggested.

Finally, since the proposed tests focus specifically on skewness and kurtosis, they are relevant to risk management which is concerned with tail events. In particular, the ratio tests can help identify an appropriate risk model for the determination of risk capital which is related to tail risk measured over a multiperiod horizon.[1] A study of the US stock market returns shows that the residuals of some popular GARCH models pass both Ljung and Box (1978) and Li and Mak (1994) tests but fail the skewness-kurtosis ratio tests. Furthermore, it is found that the multiperiod Value-at-Risk (VaR) obtained using the square root scaling law is over conservative, whereas the higher-order dependence remained in the GARCH residuals produces a multiperiod VaR that fails to provide sufficient coverage.

The rest of the paper is organized as follows. Section 2 introduces some preliminary properties that are useful for the derivation of the skewness and kurtosis ratio tests in Section 3. The next section investigates the size properties of the proposed tests by simulation analyses. The empirical results of applying the ratio tests to the US equity markets are

[1] Basel III stipulates the use of 10-day tail risk for the determination of risk capital.

reported in Section 5. Finally, a summary is provided in Section 6.

2 Some preliminaries

2.1 Cumulants

In this paper, the analyses and results are presented in terms of cumulants. Formally, the p-th order joint cumulant of p-variate random variable $(y_1, \ldots, y_p)$, denoted as $\mathrm{cum}(y_1, \ldots, y_p)$, is defined as the coefficient of $i^n t_1 \cdots t_p$ in the Taylor series expansion of the natural logarithm of $\mathrm{E}\left[\exp\left(i \sum_{j=1}^{p} y_j t_j\right)\right]$. For the special case $y_j = y$, $j = 1, \ldots, p$, $\mathrm{cum}(y_1, y_2, \ldots, y_p)$ is simply the p-th order cumulant of y. Note that $\mathrm{cum}(y) = \mathrm{E}(y)$ and $\mathrm{cum}(y, y) = \mathrm{var}(y)$.[2]

Listed below are some properties that motivate the use of cumulants in the subsequent analyses.

Lemma 1 Let z_1 and $y_1, \ldots, y_n$ be random variables whose joint cumulant exists. Then

1. $\mathrm{cum}(y_1, \ldots, y_n)$ is symmetric in its argument.
2. $\mathrm{cum}(y_1 + z_1, y_2, \ldots, y_n) = \mathrm{cum}(y_1, y_2, \ldots, y_n) + \mathrm{cum}(z_1, y_2, \ldots, y_n)$.
3. If any of $y_1, \ldots, y_n$ is independent of the remaining y's, $\mathrm{cum}(y_1, \ldots, y_n) = 0$.
4. If a is a constant, $\mathrm{cum}(a, y_1, \ldots, y_n) = 0$.
5. If $a_1, \ldots, a_n$ are constants, $\mathrm{cum}(a_1 y_1, \ldots, a_n y_n) = a_1 \cdots a_n \mathrm{cum}(y_1, \ldots, y_n)$.

2.2 Higher-order ratio relations

We shall now proceed to obtain the higher-order ratio relations based on which the proposed tests are formulated. Consider the log returns (r_t) of prices (P_t), with the former defined as $r_t = \ln(P_t / P_{t-1})$. Now define

$$\widetilde{r}_t = r_{t-h+1} + \cdots + r_t$$

as the h-period return at t. From now onwards, as in $\widetilde{r}_t$, we use '$\sim$' to indicate that the variable of interest is of h-period. For simplicity, h is suppressed in all multiperiod variables in this paper. Lo and MacKinlay (1988) made use of the fact that if r_t is IID, the stock price returns should pass the variance ratio test, i.e. the relationship

$$\mathrm{var}(\widetilde{r}_t) = h\mathrm{var}(r_t) \tag{1}$$

[2]The appendix at the end of the paper provides further relations between higher order central moments and cumulants.

holds. The variance ratio relation can now be easily extended to higher orders in terms of cumulants, as follows. Under the IID assumption of r_t, by virtue of properties 2 and 3 of Lemma 1,

$$\widetilde{\kappa}_p = h\kappa_p, \tag{2}$$

where $\widetilde{\kappa}_p$ and κ_p are the p-th order cumulant of $\widetilde{r}_t$ and r_t respectively. The result in (2) forms the basis for the higher-order ratio tests studied in this paper. If $p = 2$, (2) reduces to (1), as the second order cumulant is simply the variance.

Since skewness and kurtosis are now widely used, it is useful to relate the result of (2) to the two statistics. Let σ^2, ρ_3 and ρ_4 be the variance, skewness and kurtosis of r_t respectively. Then under the IID assumption,

$$\widetilde{\rho}_3 = \frac{\widetilde{\kappa}_3}{\widetilde{\sigma}^3} = \frac{h}{h^{3/2}}\frac{\kappa_3}{\sigma^3} = \frac{1}{\sqrt{h}}\rho_3, \tag{3}$$

$$\widetilde{\rho}_4 = \frac{\widetilde{\kappa}_4}{\widetilde{\sigma}^4} = \frac{h}{h^2}\frac{\kappa_4}{\sigma^4} = \frac{1}{h}\rho_4. \tag{4}$$

That is, as the holding interval h increases, $\widetilde{\rho}_3$ and $\widetilde{\rho}_4$ decline at a rate of $h^{-1/2}$ and h^{-1} respectively. This is the so-called intervalling effect on skewness and kurtosis that were studied by Hawawini (1980) and Lau and Wingender (1989).

Before we proceed to derive the required tests, it is worthwhile to provide an example to illustrate why the higher order relations may not hold. Consider the two-period overlapping returns $\widetilde{r}_t = r_{t-1} + r_t$. By virtue of Lemma 1, the third order cumulant of $\widetilde{r}_t$ is

$$\begin{aligned}
\operatorname{cum}(\widetilde{r}_t, \widetilde{r}_t, \widetilde{r}_t) &= \operatorname{cum}(r_{t-1}, r_{t-1}, r_{t-1}) + \operatorname{cum}(r_t, r_t, r_t) \\
&\quad + 3\operatorname{cum}(r_{t-1}, r_{t-1}, r_t) + 3\operatorname{cum}(r_{t-1}, r_t, r_t) \\
\widetilde{\kappa}_3 &= 2\kappa_3 + 3\operatorname{cum}(r_{t-1}, r_{t-1}, r_t) + 3\operatorname{cum}(r_{t-1}, r_t, r_t). \qquad (5)
\end{aligned}$$

So testing $\widetilde{\kappa}_3 = 2\kappa_3$ is equivalent to testing $\operatorname{cum}(r_{t-1}, r_{t-1}, r_t) + \operatorname{cum}(r_{t-1}, r_t, r_t) = 0$. That is, if higher order intertemporal dependency exists between r_{t-1} and r_t, the skewness ratio relation does not hold.

Now, suppose r_t follows an AR(1) process:

$$r_t = m + ar_{t-1} + e_t \tag{6}$$

where m and a are constants and the innovation e_t is an IID random variable which has a finite non-zero third order cumulant or moment. Then according to Lemma 1,

$$\operatorname{cum}(r_{t-1}, r_{t-1}, r_t) = a \cdot \operatorname{cum}(r_{t-1}, r_{t-1}, r_{t-1}) = a \cdot \kappa_3 \neq 0. \tag{7}$$

Thus, linear autocorrelation in r_t would also result in the rejection of the skewness ratio relation; similar arguments also apply to the kurtosis ratio test. In short, both linear and nonlinear dependence could render the higher-order relation in (2) invalid.

3 Higher-order ratio tests

Richardson and Smith (1991) proposed a GMM approach for the variance ratio test, using (1) as a restriction in the sample moment conditions. A major contribution by Richardson and Smith is the use of an analytically derived weighting matrix in the presence of overlapping returns for the GMM test. By explicitly modeling the dependencies of overlapping observations, the approach uses more information from the data and thus enjoys higher test powers and better size properties. This section extends Richardson and Smith's GMM approach to the skewness and kurtosis ratio tests.

3.1 GMM test

To apply the GMM test procedure, for each period t we construct an R-vector $f_t(r_t, \widetilde{r}_t, \theta)$ where θ is a P-vector of unknown parameters, namely μ, σ^2 and κ_j, to be determined. Each element of $f_t(\cdot)$ corresponds to a restriction, at least one of which is attributed to the higher order-ratio relation given in (2). Given the time series $\{r_t, \widetilde{r}_t\}_{t=1}^T$,

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} f_t(r_t, \widetilde{r}_t, \theta) \tag{8}$$

tends to zero as T tends to infinity if the higher order-ratio relation holds. The idea behind the GMM approach is to obtain the estimator $\widehat{\theta}$ such that it has a minimum variance-covariance matrix. Hansen (1982) showed that this can be achieved by solving the system of equations

$$D_0' S_0^{-1} g_T(\theta) = 0, \tag{9}$$

where

$$D_0 = E\left[\frac{\partial g_0(\theta)}{\partial\theta}\right], \tag{10}$$

$$S_0 = \sum_{l=-\infty}^{\infty} E\left[f_t(\cdot)\, f_{t-l}(\cdot)'\right]. \tag{11}$$

It can be shown that under the null hypothesis,

$$\sqrt{T}\left(\widehat{\theta}-\theta\right) \longrightarrow N\left(0, \left[D_0' S_0^{-1} D_0\right]^{-1}\right), \tag{12}$$

$$T g_T\left(\widehat{\theta}\right)' S_0^{-1} g_T\left(\widehat{\theta}\right) \longrightarrow \chi^2_{R-P}, \tag{13}$$

where $R > P$. One reason for the popularity of the GMM approach lies in its validity when D_0 and S_0 are replaced by their consistent estimators, denoted respectively as D_T and S_T. In particular, the S_T is often calculated by the two-step procedure of Hansen and Singleton (1982) or the Newey and West (1987) approach, which guarantees a positive definite weighting matrix based on the sample estimates of (11).

A contribution of this article is to derive analytically, under the IID assumption, the matrix S_0 when overlapping observations are used. As is shown in the following subsections, only certain cumulants are required to be estimated if S_0 is analytically derived.

3.2 Skewness ratio test

For the skewness ratio test, f_t and D_0 are

$$f_t = \begin{bmatrix} r_t-\mu \\ (r_t-\mu)^3-\kappa_3 \\ (\widetilde{r}_t-h\mu)^3-h\kappa_3 \end{bmatrix}, \quad D_0 = \begin{bmatrix} -1 & 0 \\ -3\sigma^2 & -1 \\ -3h^2\sigma^2 & -h \end{bmatrix}, \tag{14}$$

with $R=3$, $P=2$ and $\theta = (\mu \quad \kappa_3)'$. To derive the required covariance matrix S_0, consider for example the covariance between the second and last elements of f_t in (14), i.e. $\text{cov}((r_t-\mu)^3-\kappa_3, (\widetilde{r}_t-h\mu)^3-h\kappa_3)$. Since κ_3 is non-stochastic, by virtue of the properties in Lemma 1, the required covariance is simply $\text{cum}(x_t^3, \widetilde{x}_t^3)$ where

$$x_t = r_t-\mu, \tag{15}$$

$$\widetilde{x}_t = \widetilde{r}_t-h\mu. \tag{16}$$

So, the associated element of S_0 is $\sum_{l=-\infty}^{\infty}\text{cum}\left(x_t^3, \widetilde{x}_{t-l}^3\right)$, which can be denoted as $s_{1,h}^{3,3}$, where the superscripts refer to the powers of random variables and the subscripts to the periods over which the returns are measured. Using the same notation, the required covariance matrix can be written as

$$S_0 = \begin{bmatrix} s_{1,1}^{1,1} & s_{1,1}^{1,3} & s_{1,h}^{1,3} \\ s_{1,1}^{3,1} & s_{1,1}^{3,3} & s_{1,h}^{3,3} \\ s_{h,1}^{3,1} & s_{h,1}^{3,3} & s_{h,h}^{3,3} \end{bmatrix}.$$

Exploiting the overlapping dependencies and the IID assumption, the elements of S_0 are derived in Appendix as:

$$\begin{aligned} s_{1,1}^{1,1} &= \sigma^2, && (17) \\ s_{1,h}^{1,3} &= h\left[\kappa_4 + 3h\sigma^4\right], && (18) \\ s_{1,h}^{3,3} &= h\left[\kappa_6 + (3h+12)\,\kappa_4\sigma^2 + 9\kappa_3^2 + (9h+6)\,\sigma^6\right], && (19) \\ s_{h,h}^{3,3} &= h^2\kappa_6 + \left[6h^3 + 9A_h\right]\kappa_4\sigma^2 + 9A_h\kappa_3^2 + \left[9h^4 + 6B_h\right]\sigma^6, && (20) \end{aligned}$$

where $A_h = h\left(2h^2+1\right)/3$ and $B_h = h^2\left(h^2+1\right)/2$. Note that if $h = 1$, $A_h = B_h = 1$, (18) reduces to $s_{1,1}^{1,3}$ and both (19) and (20) simplify to $s_{1,1}^{3,3}$.

3.3 Kurtosis ratio test

For the kurtosis ratio test, the corresponding f_t and D_0 are

$$f_t = \begin{bmatrix} r_t - \mu \\ (r_t-\mu)^2 - \sigma^2 \\ (r_t-\mu)^4 - 3\sigma^4 - \kappa_4 \\ (\widetilde{r}_t - h\mu)^4 - 3h^2\sigma^4 - h\kappa_4 \end{bmatrix}, \quad D_0 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ -4\kappa_3 & -6\sigma^2 & -1 \\ -4h^2\kappa_3 & -6h^2\sigma^2 & -h \end{bmatrix}. \qquad (21)$$

Here, $R = 4$, $P = 3$ and $\theta = (\mu \quad \sigma^2 \quad \kappa_4)'$. Using the same notation as in the skewness ratio test, the associated weighting matrix is given by

$$S_0 = \begin{bmatrix} s_{1,1}^{1,1} & s_{1,1}^{1,2} & s_{1,1}^{1,4} & s_{1,h}^{1,4} \\ s_{1,1}^{2,1} & s_{1,1}^{2,2} & s_{1,1}^{2,4} & s_{1,h}^{2,4} \\ s_{1,1}^{4,1} & s_{1,1}^{4,2} & s_{1,1}^{4,4} & s_{1,h}^{4,4} \\ s_{h,1}^{4,1} & s_{h,1}^{4,2} & s_{h,1}^{4,4} & s_{h,h}^{4,4} \end{bmatrix},$$

where the required covariances are derived in Appendix as

$$
\begin{aligned}
s_{1,1}^{1,2} &= \kappa_3, && (22)\\
s_{1,1}^{2,2} &= \kappa_4 + 2\sigma^4, && (23)\\
s_{1,h}^{1,4} &= h\left[\kappa_5 + 10h\kappa_3\sigma^2\right], && (24)\\
s_{1,h}^{2,4} &= h\left[\kappa_6 + (6h+8)\,\kappa_4\sigma^2 + (4h+6)\,\kappa_3^2 + 12h\sigma^6\right], && (25)\\
s_{1,h}^{4,4} &= h\kappa_8 + (6h+22)\,\kappa_6\sigma^2 + (4h+52)\,\kappa_5\kappa_3 + 34\kappa_4^2 \\
&\quad + (84h+120)\,\kappa_4\sigma^4 + (100h+180)\,\kappa_3^2\sigma^2 + (72h+24)\,\sigma^8, && (26)\\
s_{h,h}^{4,4} &= h^2\kappa_8 + \left[12h^3 + 16A_h\right]\kappa_6\sigma^2 + \left[8h^3 + 48A_h\right]\kappa_5\kappa_3 + 34A_h\kappa_4^2 \\
&\quad + \left[36h^4 + 96hA_h + 72B_h\right]\kappa_4\sigma^4 + \left[64h^4 + 72hA_h + 144B_h\right]\kappa_3^2\sigma^2 \\
&\quad + \left[72h^2A_h + 24C_h\right]\sigma^8. && (27)
\end{aligned}
$$

In (27), $C_h = h\left(6h^4 + 10h^2 - 1\right)/15$. Similar to the case of the skewness ratio test, when $h = 1$, $C_h = 1$, (24) yields $s_{1,1}^{1,4}$, (25) yields $s_{1,1}^{2,4}$ and both (26) and (27) simplify to $s_{1,1}^{4,4}$.

3.4 Joint skewness and kurtosis ratio test

We also consider a joint test based on both skewness and kurtosis ratio relations, for the two statistics are often used together as in the case of normality test by Jarque and Bera (1980). For the joint skewness and kurtosis ratio test, we have

$$
f_t = \begin{bmatrix} r_t - \mu \\ (r_t-\mu)^2 - \sigma^2 \\ (r_t-\mu)^3 - \kappa_3 \\ (r_t-\mu)^4 - 3\sigma^4 - \kappa_4 \\ (\widetilde{r}_t - h\mu)^3 - h\kappa_3 \\ (\widetilde{r}_t - h\mu)^4 - 3h^2\sigma^4 - h\kappa_4 \end{bmatrix}, \quad D_0 = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ -3\sigma^2 & 0 & -1 & 0 \\ -4\kappa_3 & -6\sigma^2 & 0 & -1 \\ -3h^2\sigma^2 & 0 & -h & 0 \\ -4h^2\kappa_3 & -6h^2\sigma^2 & 0 & -h \end{bmatrix} \qquad (28)
$$

with $R = 6$, $P = 4$ and $\theta = (\mu \quad \sigma^2 \quad \kappa_3 \quad \kappa_4)'$. The covariance matrix is

$$S_0 = \begin{bmatrix} s_{1,1}^{1,1} & s_{1,1}^{1,2} & s_{1,1}^{1,3} & s_{1,1}^{1,4} & s_{1,h}^{1,3} & s_{1,h}^{1,4} \\ s_{1,1}^{2,1} & s_{1,1}^{2,2} & s_{1,1}^{2,3} & s_{1,1}^{2,4} & s_{1,h}^{2,3} & s_{1,h}^{2,4} \\ s_{1,1}^{3,1} & s_{1,1}^{3,2} & s_{1,1}^{3,3} & s_{1,1}^{3,4} & s_{1,h}^{3,3} & s_{1,h}^{3,4} \\ s_{1,1}^{4,1} & s_{1,1}^{4,2} & s_{1,1}^{4,3} & s_{1,1}^{4,4} & s_{1,h}^{4,3} & s_{1,h}^{4,4} \\ s_{h,1}^{3,1} & s_{h,1}^{3,2} & s_{h,1}^{3,3} & s_{h,1}^{3,4} & s_{h,h}^{3,3} & s_{h,h}^{3,4} \\ s_{h,1}^{4,1} & s_{h,1}^{4,2} & s_{h,1}^{4,3} & s_{h,1}^{4,4} & s_{h,h}^{4,3} & s_{h,h}^{4,4} \end{bmatrix}. \tag{29}$$

Most of the elements of S_0 in (29) have been provided in the preceding analyses. The remaining required covariance elements are (see Appendix for proofs)

$$\begin{aligned} s_{1,h}^{2,3} &= h\left[\kappa_5 + (3h+6)\,\kappa_3\sigma^2\right], & (30)\\ s_{1,h}^{4,3} &= h\left[\kappa_7 + (3h+18)\,\kappa_5\sigma^2 + 34\kappa_4\kappa_3 + (30h+72)\,\kappa_3\sigma^4\right], & (31)\\ s_{1,h}^{3,4} &= h\left[\kappa_7 + (6h+15)\,\kappa_5\sigma^2 + (4h+30)\,\kappa_4\kappa_3 + (66h+36)\,\kappa_3\sigma^4\right], & (32)\\ s_{h,h}^{3,4} &= h^2\kappa_7 + \left[9h^3 + 12A_h\right]\kappa_5\sigma^2 + \left[4h^3 + 30A_h\right]\kappa_4\kappa_3 \\ &\quad + \left[30h^4 + 36hA_h + 36B_h\right]\kappa_3\sigma^4. & (33) \end{aligned}$$

Again, setting $h = 1$ reduces (30) to $s_{1,1}^{2,3}$ whereas (31), (32) and (33) become $s_{1,1}^{3,4}$.

3.5 On heteroscedasticity

Lo and MacKinlay (1988) uses White's (1980) heteroscedastic-consistent covariance matrix estimator to calculate the standard errors of lagged serial correlation coefficients in order to make the variance ratio test heteroscedastic-consistent. However, it is worth noting that Richardson and Smith (1991) does not make its GMM procedure for overlapping observations heteroscedasticity-consistent. This is because, to make the GMM ratio test heteroscedasticity-consistent, additional estimation of numerous autocorrelation parameters are required, resulting in poor size properties in small samples. While it is theoretically possible to make the proposed skewness-kurtosis ratio tests heteroscedasticity-consistent, it will be either extremely complex (given the already complex covariance-matrix analytically derived in this section) or the higher-order ratio tests will have very poor size properties if Lo and MacKinlay's use of White's covariance matrix estimator is followed (as can be seen from the simulation study in the next section when Newey-West's (1987) method is employed to obtain the weighting matrix). The proposed higher-order ratio tests may be regarded as diagnostic tests similar to those of McLeod and Li (1983) and Li and Mak (1994) for IID processes or shocks. Indeed, it is the IID assumption that allows information to be retrieved fully in a simple manner from overlapping observations, thereby making the tests powerful with good size properties.

4 A simulation study of size properties

This section uses Monte Carlo simulations to investigate how well the asymptotic results derived in the last section would hold in practice. In particular, in order to demonstrate the advantage of the analytically derived S_0 over the widely used Newey-West covariance matrix, S_{nw}, we also consider ratio tests that use the latter covariance matrix in place of the former.[3] The empirical sizes are calculated as the proportion of rejections in 5,000 replications of the proposed tests on various supposedly IID processes of sample size N equals to 250 and 1,000. Table 1 provides the calculated test sizes at 10%, 5% and 1% levels with h equals 5 and 10 periods for the skewness (Skew), kurtosis (Kurt) and their joint (Joint) ratio tests.

< Table 1 Empirical sizes >

In Panel A, IID standard normal samples are generated and the entries in columns 3 to 5 are test sizes obtained using S_0. It can be seen that the tests are generally under-sized at 10% level but over-sized at 1% level. At 5% level, the empirical sizes are close to the theoretical value for both skewness and joint ratio tests but slightly under-sized for the kurtosis ratio test. The empirical sizes in the last three columns are obtained using the covariance matrix S_{nw}. It seems that the Newey-West approach estimates the required covariance matrix of higher-cumulants poorly, resulting in hugely under-sized skewness tests but over-sized kurtosis tests.

Now let $z_t \sim skst\,(\upsilon, \ln \xi)$ denote an IID zero-mean unit-variance skewed Student process where υ and $\ln \xi$ are the degree of freedom and skewness parameter respectively; see Hansen (1994) for further details. Since stock returns are well known to be skewed, leptokurtic and heteroscedastic, the process considered in Panel B is z_t whereas Panel C studies the estimated standardized residuals $\widehat{z}_t$ of a GARCH time series $r_t = \varepsilon_t = \sigma_t z_t$, $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$. In both cases, $z_t \sim skst\,(8, -0.1)$.[4] Possibly due to the non-normality of the simulated process, the empirical sizes in Panel B and C at 10% level tend to be slightly smaller than those of Panel A; the difference is smaller at 5% level and it disappears at 1% level. However, the entries in Panel C are qualitatively similar to those in Panel B, suggesting that GARCH does not produce the expected difference as explained by Li and Mak (1994) which shows that

[3] The Newey-West covariance matrix is estimated by $S_{nw} = \sum_{|l|<h} \sum_t (1 - l/h)\, f(t)\, f(t-l)'$.

[4] A negative $\ln \xi$ implies a left-skewed distribution. The sample of US stock market returns studied later in the next section is found to have similar υ and $\ln \xi$ values.

the sampling distribution of higher moments of $\widehat{z}_t$ is not necessarily the same as that of z_t. It is conjectured that the reason lies in the two different constructs of the traditional squared residual autocorrelation test and the proposed GMM ratio test. The former sums up the squares of the autocorrelations and hence its degree of freedom varies with the number of lags used. The GMM approach, on the other hand, estimates P parameters with R constraints giving rise to $R - P$ degree of freedom that is independent of h.

Finally, it is noted that noticeable improvements in empirical sizes are observed for all three processes as the sample size increases from 250 to 1,000. To confirm the validity of the asymptotic distribution, Table 2 below reports the empirical sizes when the sample size N increases from 2,500 to 10,000. Also reported in Panel D of the table are the averages of absolute differences or errors between the empirical sizes of the three ratio tests and their corresponding theoretical sizes. We can see that as N increases, the average absolute errors decline in almost all cases. The only exception is the case $h = 5$ at 10% level when N increases from 2,500 to 5,000, the average error becomes larger at 0.64. When the sample size increases to 10,000, the error reduces to 0.26.

$<$ Table 2 Asymptotic empirical sizes $>$

To sum up the above simulation study, the analytically derived covariance matrix S_0 gives rise to empirical test sizes that are reasonably close to their true value, especially when sample size is large. Moreover, the size properties of the proposed ratio tests remain good when the tests are applied to GARCH residuals.

5 Higher-moment dependence in stock markets

In this section, we apply the proposed GMM tests to the US stock markets and find significant presence of higher-order dependence even after fitting some of the most popular GARCH models. No attempt is made to identify the best GARCH model in terms of goodness of fit, forecast, or ability to pass diagnostic tests, for the aim here is to illustrate the complementary role of the skewness and kurtosis ratio tests. To explain the breakdown of higher-order ratio relations, we also provide simple t-tests of certain higher-moment statistics. Finally, the association of multiperiod tail risks with higher-order ratios is illustrated.

5.1 Data and descriptive statistics

Consider the S&P 500 stock index, a total of 2,516 log returns from 2 January 2006 to 31 December 2015.[5] Table 3 provides the descriptive statistics as well as the scaled standard

[5] For the empirical analysis, the log returns are calculated as $r_t = 100 \times \ln\left(P_t/P_{t-1}\right)$.

deviation (sd), skewness (sk) and kurtosis (ku) for various multi-period horizons $h = 5, 10$ and 20. Note that the scaling is achieved by setting $sd = h^{-1/2}\widetilde{\sigma}$, $sk = h^{1/2}\widetilde{\rho}_3$ and $ku = h\widetilde{\rho}_4$, so that their expected values would remain constant for different values of h if the ratio relations hold.

< Table 3 Basic statistics >

It can be seen from the sk and ku values that as h increases, the returns are increasingly more left-skewed and leptokurtic than would be the case if the returns were IID. To find out whether the changes in sk and ku are statistically significant, we apply the GMM ratio tests below.

5.2 Applying the skewness-kurtosis ratio tests

The skewness-kurtosis ratio tests, Li and Mak (1994) (LiMak) test as well as the Ljung and Box (1978) (LB) test are applied to the log returns, residuals of an AR(1) model, and standardized residuals of AR(1)-GARCH with Gaussian shocks (GARCH-g) and AR(1)-Asymmetric GARCH with skewed Student shocks (AGARCH-skst). In Table 4, Panel A reports the test results whereas Panel B provides the estimates of the models. Under the null hypothesis, the reported test statistics of Skew, Kurt and Joint are distributed as chi square with 1, 1, and 2 degree of freedom respectively; for both LB and LiMak, the degree of freedom is $h + 10$. In the last three columns of Panel A, sd, $k3$ and $k4$ are the scaled standard deviation and standardized third and fourth order cumulant statistics for $\widetilde{\sigma}/\sqrt{h}$, $\widetilde{\kappa}_3/(h\sigma^3)$ and $\widetilde{\kappa}_4/(h\sigma^4)$ respectively.[6] If the returns are IID, the expected values of these statistics will not vary with h. Hence any large changes in them, especially $k3$ and $k4$, would likely be accompanied by large, significant skewness and kurtosis test statistics.

< Table 4 GMM ratio tests >

First it can be seen that the squares of log returns of our sample are highly autocorrelated, as is evidenced from the large LiMak statistics (72.89 and 107.4). Similarly, the results of skewness and kurtosis ratio tests indicate the presence of third and fourth order dependence in the US stock markets.

After applying the AR(1) filter, the Ljung-Box test statistics have become lower but remain significant. As explained in Section 2, higher-order dependence could also be caused

[6] Note that $\widetilde{\sigma}^2$, $\widetilde{\kappa}_3$ and $\widetilde{\kappa}_4$ are estimated using the h-period returns $\widetilde{r}_t$ whereas σ is obtained from the daily returns r_t. Under the IID assumption, $k3/\sqrt{h}$ and $k4/h$ are respectively the skewness and kurtosis of h-period returns.

by linear autocorrelation. It is thus surprising to see that, instead of lower dependence, the AR residuals show signs of further deviation from the null hypothesis. Later in the next subsection, cumulant properties are used to explain the paradoxical evidence of higher skewness-kurtosis test statistics as well as larger magnitudes of $k3$ and $k4$ for the AR residuals.

Consistent with the literature, the standardized residuals of GARCH-g pass both Li-Mak and Ljung-Box tests. Also, the magnitudes of the $k3$ and $k4$ statistics as well as the skewness-kurtosis test statistics are now considerably smaller. Significant higher-order dependence, however, is still present in the residuals, for both skewness and joint ratio tests remain statistically significant. There is improvement when AGARCH-skst is fitted to the returns, as is evidenced from smaller ratio test statistics and lower variation in $k3$ and $k4$ values. Nevertheless, the skewness ratio relation breaks down for weekly residuals whereas the joint ratio test is significant for both weekly and fortnightly periods.

Also reported in Table 4 are the p-values of test statistics. The p-values and the empirical test sizes in Table 2 would ascertain the significance level of the higher-ratio tests. For example, the joint ratio test on weekly residuals of AGARCH-skst is more likely to be significant at 5% level since the p-value of 0.0097 is marginally below 0.01 and the corresponding empirical size in Table 2 is 2.1%.

5.2.1 Autocorrelation and intervalling effect

To explain why removing linear autocorrelation could result in larger variation in $k3$ and $k4$, consider the AR(1) process in (6) with parameters as shown in Panel A of Table 4. Suppose its innovations e_t have a finite nonzero k-th order cumulant denoted as $\kappa_{e,k}$. First note that $a < 0$ is consistent with declining scaled standard deviation (sd) with respect to h. Now, the k-th order cumulant of r_t can be written as

$$\kappa_k = \left(1 - a^k\right)^{-1} \cdot \kappa_{e,k}$$

For the weekly returns, the third order cumulant is

$$\operatorname{cum}\left(\widetilde{r}_t, \widetilde{r}_t, \widetilde{r}_t\right) = \sum_{i=1}^{5} \operatorname{cum}\left(r_i, r_i, r_i\right) + 3 \sum_{j \neq i} \operatorname{cum}\left(r_i, r_i, r_j\right) + \sum_{i \neq j \neq k} \operatorname{cum}\left(r_i, r_j, r_k\right). \qquad (34)$$

By virtue of Lemma 1, the summand in the second term on the right of (34) is either zero, $a\kappa_3$ or $a^2\kappa_3$. For small $a = -0.1$, a^2 is negligible and similar analyses show that $\operatorname{cum}(r_i, r_j, r_k)$ is of even smaller value, $a^3\kappa_3$ or less. Hence, the third order cumulant of log returns is approximately

$$\operatorname{cum}\left(\widetilde{r}_t, \widetilde{r}_t, \widetilde{r}_t\right) \approx 5\kappa_3 + 12a\kappa_3. \tag{35}$$

The corresponding third order cumulants of AR residuals are

$$\operatorname{cum}\left(\widetilde{e}_t, \widetilde{e}_t, \widetilde{e}_t\right) = 5\kappa_{e,3} = 5 \cdot \left(1 - a^3\right)\kappa_3 \approx 5\kappa_3, \tag{36}$$

since a is small. As $\kappa_3 < 0$, (36) is less than (35), hence the AR residuals have more negative $k3$ statistics than those of the weekly returns.

For the standardized kurtosis $k4$, we can use the same method of analysis and obtain for the weekly returns

$$\begin{aligned} \operatorname{cum}\left(\widetilde{r}_t, \widetilde{r}_t, \widetilde{r}_t, \widetilde{r}_t\right) &\approx \sum_{i=1}^{5} \operatorname{cum}\left(r_i, r_i, r_i, r_i\right) + 4\sum_{i=1}^{4} \operatorname{cum}\left(r_i, r_i, r_i, r_{i+1}\right) \\ &\approx 5\kappa_4 + 16a\kappa_4 \end{aligned} \tag{37}$$

whereas for the residuals, the cumulant is

$$\operatorname{cum}\left(\widetilde{e}_t, \widetilde{e}_t, \widetilde{e}_t, \widetilde{e}_t\right) = 5\kappa_{e,4} = 5 \cdot \left(1 - a^4\right)\kappa_4 \approx 5\kappa_4 \tag{38}$$

Since $\kappa_4 > 0$, a negative a implies that (38) is greater than (37), which is consistent with the reported $k4$ statistics in Table 4.

We end the discussion here by remarking that the proofs provided above for weekly returns can be easily extended to the fortnightly returns when $h = 10$.

5.2.2 Higher-moment t-tests

It would be interesting to find out which higher-moment dependence is responsible for the breakdown of the ratio relations as reported in Table 4. In particular, a significant GMM ratio test statistic could be due to one or more non-zero cumulants listed in the first column of Table 5 below. The second and third columns of the table provide the corresponding moment and variance, respectively, to be used in the test for zero cumulant under the null hypothesis that r_t is IID.[7] As an example, given a sample $\{r_{1,\ldots,}r_T\}$, first obtain the demeaned sample $\{x_{1,\ldots,}x_T\}$. The t-statistic for testing $H_0 : c011 = 0$ can then be calculated as $\sqrt{T-1}\sum x_t x_{t+1}^2/s$ where $s^2 = \sigma^2(\kappa_4 + 3\sigma^4)$.

[7] The cumulants for the calculation of the variance can be estimated using central moments as described in Appendix A.4.

Table 5: Higher moments for t-tests

Cumulant	Moment	Variance
$c011 = \text{cum}(r_t, r_{t+1}, r_{t+1})$	$E\left(x_t x_{t+1}^2\right)$	$\sigma^2\left(\kappa_4 + 3\sigma^4\right)$
$c001 = \text{cum}(r_t, r_t, r_{t+1})$	$E\left(x_t^2 x_{t+1}\right)$	$\sigma^2\left(\kappa_4 + 3\sigma^4\right)$
$c012 = \text{cum}(r_t, r_{t+1}, r_{t+2})$	$E\left(x_t x_{t+1} x_{t+2}\right)$	σ^6
$c0111 = \text{cum}(r_t, r_{t+1}, r_{t+1}, r_{t+1})$	$E\left(x_t x_{t+1}^3\right)$	$\sigma^2\left(\kappa_6 + 15\kappa_4\sigma^2 + 10\kappa_3^2 + 15\sigma^6\right)$
$c0001 = \text{cum}(r_t, r_t, r_t, r_{t+1})$	$E\left(x_t^3 x_{t+1}\right)$	$\sigma^2\left(\kappa_6 + 15\kappa_4\sigma^2 + 10\kappa_3^2 + 15\sigma^6\right)$
$c0011 = \text{cum}(r_t, r_t, r_{t+1}, r_{t+1})$	$E\left(x_t^2 x_{t+1}^2\right) - \left[E\left(x_t^2\right)\right]^2$	$\left(\kappa_4 + 2\sigma^4\right)^2$

To ensure a correct inference of the t-statistics, simulations similar to those conducted in Section 4 are carried out and the empirical sizes at 5% level are reported in Panel B of Table 6. With the exception of $c0011$ on residuals of GARCH-skst, all proposed t-tests have empirical sizes that are similar to the correct value. Now turning to Panel A which provides the test results, we can see that all six cumulants of S&P 500 returns are significantly different from zero. The t-statistics are reduced in size when the tests are applied to the AR(1) residuals. In particular, the $c0111$ is no longer significant. Consistent with the results reported in Table 4, much of the dependence has been removed by the GARCH-g model, except for $c011$ and $c0011$. Finally, the AGARCH-skst makes further but small improvement as the t-statistic of $c011$ is now marginally insignificant at 5% level.

$<$ Table 6 Tests on higher moment statistics $>$

To sum up, the significance of higher ratio test may be attributed to two forms of dependence ($c011$ and $c0011$) that remain even after taking into account stock market salient features such as heteroscedasticity, volatility asymmetry and non-normality of distribution. Negative t-statistics of $c011$ suggest that volatility tends to rise after negative shocks. As for $c0011$, it is interesting to note that after the GARCH filter, the squares of residuals are no longer persistent but become negatively autocorrelated. Finally, we remark that the under-sized issue of the $c0011$ test does not invalidate the above results as the magnitudes of the t-statistics concerned are quite large.

5.3 Implications for risk management

In practice, tail risks are often estimated using daily returns; see for example Hsieh (1993), Wong (2010), Dupuis *et al.* (2015) and Beckers *et al.* (2017). However, the Basel Committee stipulates that the risk capital for a bank's trading portfolio is to be determined by a tail risk that is measured over a 10-day horizon; see Basel Committee on Banking Supervision (2016) for details. It will be demonstrated here how the proposed ratio tests can help identify an appropriate risk model for the forecast of multiperiod tail risks.

There are two popular measures of tail risk, namely VaR and expected shortfall (ES). Although Wong (2008, 2010) have shown that backtests based on the latter are more powerful, for simplicity, the study below is carried out in terms of the former. In particular, a total of six multiperiod VaR models are constructed from the two estimated GARCH models (GARCH-g and AGARCH-skst), each with three ways of generating h-day ahead VaR forecasts at 99% coverage for downside risk. To serve as a benchmark for comparison, consider the *ex post* 1-day ahead VaR forecast on day t

$$\text{1-day VaR} = m\left(r_{t-1}\right) + \widehat{\sigma}_t \cdot q_{0.01} = m\left(r_{t-1}\right) + v\left(\overline{\sigma}_{t-1}, \overline{z}_{t-1}\right) \cdot q_{0.01}$$

where m and v^2 are respectively the estimated mean and variance function of the GARCH model, $\overline{\sigma}_{t-1}$ and $\overline{z}_{t-1}$ are respectively the estimated volatility and standardized shock, and $q_{0.01}$ is the first percentile of the shock distribution based on 5,000 bootstrap samples drawn from $\{\overline{z}_t\}_{t=1}^{T}$.[8] The first way of generating the h-day ahead VaR is to use the scaling law that assumes IID returns, i.e.,

$$h\text{-day VaR} = \sqrt{h} \times \text{1-day VaR}.$$

The second method is by block-bootstrapping from $\{\overline{z}_t\}_{t=1}^{T}$ with block length h to obtain the random vector $\left(\overline{z}_{(\tau)}, \ldots, \overline{z}_{(\tau+h-1)}\right)$. The h-period return $\widehat{r}_t^h = \widehat{r}_t + \cdots + \widehat{r}_{t+h-1}$ is then calculated from $\widehat{r}_t = m\left(r_{t-1}\right) + \widehat{\sigma}_t \overline{z}_{(\tau)}$ where $\widehat{\sigma}_t = v\left(\overline{\sigma}_{t-1}, \overline{z}_{t-1}\right)$, and $\widehat{r}_{t+j} = m\left(\widehat{r}_{t+j-1}\right) + \widehat{\sigma}_{t+j} \overline{z}_{(\tau+j)}$ where $\widehat{\sigma}_{t+j} = v\left(\widehat{\sigma}_{t+j-1}, \overline{z}_{(\tau+j-1)}\right)$, $j = 1, \ldots, h-1$. The required h-day VaR is obtained as the first percentile of the bootstrap distribution of $\widehat{r}_t^h$. The third approach differs from the second method only in its bootstrapping; instead of block-bootstraping, h random draws with replacement are carried out to form $\left(\overline{z}_{(\tau)}, \ldots, \overline{z}_{(\tau+h-1)}\right)$.

< Table 7: Backtesing of multiperiod VaR models >

The VaR forecasts are compared with the US stock market returns and the proportions of VaR exceptions are reported in Table 7 above. Since the 1-day ahead forecasts are *ex post*, the empirical proportions of 1-day VaR breaches for both GARCH models are close to 1%, the expected value if the risk model is true. For the h-day VaR forecasts, the number

[8] *Ex post* here refers to the fact that the forecasts are constructed from the parameters of GARCH models, and associated conditional volatilities and standardized residuals that are estimated using the full sample. In using *ex post* forecasts, the study focuses on how information from skewness-kurtosis ratio tests could help explain the performance of h-day VaR forecasts of GARCH models estimated using daily returns.

of observations reduces by a factor of h if non-overlapping forecasts are considered, in which case none of observed exception rates fail the Kupiec (1995) test of unconditional coverage.

Although a formal statistical inference cannot be made, the exception rates reported in Table 7 are consistent with the results of preceding analysis. First, when the 1-day VaR is multiplied by $\sqrt{h}$ to obtain the h-day VaR, the observed coverage of multiperiod tail risk is larger than 99%. One possible reason is due to a sample-specific property that is indicated by the negative autoregressive coefficient in the mean equation of the estimated models. Next, since the block bootstrap procedure retains the nonlinear dependence detected by the skewness-kurtosis ratio tests, the associated observed exception rates are generally higher than those using the IID bootstrap procedure. Finally, for both block and IID bootstrap h-day VaRs, the exception rates associated with AGARCH-skst are closer to 1% than those of GARCH-g. This observation is consistent with the skewness-kurtosis-ratio test results that the former GARCH model is relatively more successful in removing nonlinear dependence present in the stock returns.

In short, the empirical analysis in this subsection shows the importance of skewness-kurtosis ratio relations in constructing multiperiod tail risk forecasts from risk models based on daily returns. Therefore, the proposed higher-order ratio tests are useful in providing valuable information for the modelling and forecasting of multiperiod tail risks; see for example Mancini and Trojani (2011).

6 Conclusion

Skewness and kurtosis ratio tests are developed using a GMM technique in which overlapping observations are used so that more information can be utilized in the proposed tests. This is achieved by explicitly modelling the dependencies in the overlapping data under the IID assumption. Simulation experiments demonstrate that the proposed tests have relatively good size properties for residuals of GARCH processes as well as original time series.

Application of the higher ratio tests to the US stock market returns illustrates their complementary role to existing nonlinearity diagnostic tests. For example, the GARCH-filtered standardized residuals pass the Li and Mak (1994) test but fail the skewness-kurtosis ratio tests. The ability of the proposed tests to shed light on the nature of nonlinear dependence is particularly useful when multiperiod forecasts of tail events are required, for tail risks are closely associated with both the level of asymmetry and tail fatness of the distribution as measured by skewness and kurtosis respectively.

A Appendix

Analytical proofs for the covariance matrices S_0 used in the skewness-kurtosis ratio tests are provided here. The proofs are made simpler using x_t and $\widetilde{x}_t$ instead of r_t and $\widetilde{r}_t$, for the former have zero mean; see (15) and (16). The required covariances may be divided into three categories: covariance between products of single-period random returns (e.g. $s_{1,1}^{3,4}$), between products of single-period and h-period random returns (e.g. $s_{1,h}^{3,4}$), and between products of h-period random returns (e.g. $s_{h,h}^{3,4}$), with increasing level of complexity.

In all three cases, the required covariances can be obtained using the indecomposable partition method stated in Lemma 2. However, in order to facilitate an understanding (and cross verification) of the proofs, we first consider the results for the covariances between the products of single-period returns. These are provided in A.1 where relations between cumulants and moments are introduced. A.2 provides Lemma 2, which is required for the derivation of the covariances of the products of multiperiod random variables, and A.3 derives all the required covariances involving multiperiod returns. Finally, A.4 provides the formulae to estimate the cumulants from central moments in order to obtain the required covariance matrix S_0 for the proposed tests.

A.1 Proofs for $S_{1,1}^{p,q}$

First consider the following formulae provided by Kendall and Stuart (1969, p.70) for expressing higher-order central moments, μ_j, in terms of cumulants, κ_j:

$$\begin{aligned}
\mu_2 &= \kappa_2 = \sigma^2, && (39)\\
\mu_3 &= \kappa_3, && (40)\\
\mu_4 &= \kappa_4 + 3\sigma^4, && (41)\\
\mu_5 &= \kappa_5 + 10\kappa_3\sigma^2, && (42)\\
\mu_6 &= \kappa_6 + 15\kappa_4\sigma^2 + 10\kappa_3^2 + 15\sigma^6, && (43)\\
\mu_7 &= \kappa_7 + 21\kappa_5\sigma^2 + 35\kappa_4\kappa_3 + 105\kappa_3\sigma^4, && (44)\\
\mu_8 &= \kappa_8 + 28\kappa_6\sigma^2 + 56\kappa_5\kappa_3 + 35\kappa_4^2 + 210\kappa_4\sigma^4 + 280\kappa_3^2\sigma^2 + 105\sigma^8. && (45)
\end{aligned}$$

We shall now consider deriving an expression of $S_{1,1}^{p,q}$ $(1 \le p, q \le 4)$ in terms of cumulants using the above formulae. Under the IID assumption, x_t and x_{t-l} are independent for $l \neq 0$. Thus, by virtue of Property 3 in Lemma 1, $\text{cum}(x_t^p, x_{t-l}^q) = 0$ for $l \neq 0$. Using the above moment formulae, and exploiting the fact that $\text{E}(x_t) = 0$, $s_{1,1}^{p,q} = \sum \text{cum}(x_t^p, x_{t-l}^q) =$

$\mathrm{cov}(x_t^p, x_t^q) = \mu_{p+q} - \mu_p\mu_q$, it is straightforward that $s_{1,1}^{1,1} = \sigma^2$ and $s_{1,1}^{1,2} = \kappa_3$. For $s_{1,1}^{2,2}$,

$$s_{1,1}^{2,2} = \mathrm{cov}(x_t^2, x_t^2) = \mu_4 - \mu_2^2.$$

Substituting for μ_4 using (41) and replacing μ_2 with σ^2, we have

$$s_{1,1}^{2,2} = \kappa_4 + 3\sigma^4 - \sigma^4 = \kappa_4 + 2\sigma^4.$$

Using the same principle, the other more complex covariances are derived as follows.

$$\begin{aligned}
s_{1,1}^{1,3} &= \kappa_4 + 3\sigma^4, && (46)\\
s_{1,1}^{1,4} &= \kappa_5 + 10\kappa_3\sigma^2, && (47)\\
s_{1,1}^{2,3} &= \kappa_5 + 9\kappa_3\sigma^2, && (48)\\
s_{1,1}^{2,4} &= \kappa_6 + 14\kappa_4\sigma^2 + 10\kappa_3^2 + 12\sigma^6, && (49)\\
s_{1,1}^{3,3} &= \kappa_6 + 15\kappa_4\sigma^2 + 9\kappa_3^2 + 15\sigma^6, && (50)\\
s_{1,1}^{3,4} &= s_{1,1}^{4,3} = \kappa_7 + 21\kappa_5\sigma^2 + 34\kappa_4\kappa_3 + 102\kappa_3\sigma^4 && (51)\\
s_{1,1}^{4,4} &= \kappa_8 + 28\kappa_6\sigma^2 + 56\kappa_5\kappa_3 + 34\kappa_4^2 + 204\kappa_4\sigma^4 + 280\kappa_3^2\sigma^2 + 96\sigma^8. && (52)
\end{aligned}$$

Letting $h = 1$ in, for example, (19) and (20) will give rise to the same formula for $s_{1,1}^{3,3}$ in (50) above. One important observation to be made here is that $s_{1,1}^{p,q}$ contains the basic *structure* for $s_{1,h}^{p,q}$ and $s_{h,h}^{p,q}$. Take the case of $p = q = 4$ as an example; the right hand sides of (26) and (27) in the kurtosis ratio test share the same cumulant terms with $s_{1,1}^{4,4}$ in (52): κ_8, $\kappa_6\sigma^2$, $\kappa_5\kappa_3$, κ_4^2, $\kappa_4\sigma^4$, $\kappa_3^2\sigma^2$ and σ^8. Moreover, when $h = 1$, $A_h = B_h = C_h = 1$, yielding the same coefficients for all cumulant terms in $s_{1,1}^{p,q}$, $s_{1,h}^{p,q}$ and $s_{h,h}^{p,q}$, where $1 \le p, q \le 4$. Therefore, as can be seen in A.3 below, h^p $(1 \le p \le 4)$, A_h, B_h and C_h reflect the effects of having h-period returns in place of single-period returns under the null hypothesis of independent returns.

A.2 Cumulant of products of random variables

The above shows how $s_{1,1}^{p,q}$ can be obtained using the formulae provided by Kendall and Stuart (1969). However, things become complicated when multiperiod returns are involved. Since the required covariances are essentially the cumulants of products of random variables, we introduce here the concept of an indecomposable partition provided by Brillinger (1975, Section 2.3) in order to obtain the cumulants of products of x_t.

Definition *Consider a partition $P_1 \cup \cdots \cup P_M$ of the table of entries (not necessarily* *rectangular) given below*

$$\begin{array}{ccc} (1,1) & \cdots & (1,J_1) \\ \vdots & & \vdots \\ (I,1) & \cdots & (1,J_I) \end{array}$$

Sets $P_{m'}$ and $P_{m''}$ are said to hook if there exists $(i_1, j_1) \in P_{m'}$ and $(i_2, j_2) \in P_{m''}$ such that $i_1 = i_2$; that is (i_1, j_1) and (i_2, j_2) are from the same row. $P_{m'}$ and $P_{m''}$ are said to communicate if there exists a sequence of sets $P_{m_1} = P_{m'}, P_{m_2}, \ldots, P_{m_N} = P_{m''}$ such that P_{m_n} and $P_{m_{n+1}}$ hook for $n = 1, \ldots, N-1$. A partition is said to be indecomposable if all of its sets communicate.

Each row in the table above corresponds to a product of (random) returns in our paper. So, $I = 2$, as we need only covariances that are second order cumulants. Take the case of $\mathrm{cum}\big(x_t^3, \widetilde{x}_{t-l}^4\big)$ in $s_{1,h}^{3,4}$ for illustration, we can let the first row of entries in the above table correspond to x_t^3, whereas the second row correspond to $\widetilde{x}_{t-l}^4$, so that $J_1 = 3$ and $J_2 = 4$. An indecomposable partition as defined above is one that contains at least a set in which at least one element is from x_t^3 and the other from $\widetilde{x}_{t-l}^4$.

The result that can be used to obtain the joint cumulant of products of random variables may now be presented in Lemma 2 below.

Lemma 2 *Consider the (two way) I random variables*

$$Y_i = \prod_{j=1}^{J_i} X_{ij},$$

where $j = 1, \ldots, J_i$ and $i = 1, \ldots, I$. *The joint cumulant cum$(Y_1, \ldots, Y_I)$ is given by*

$$\sum_{\mathbf{P}} \mathrm{cum}\left(X_{ij}; ij \in P_1\right) \cdots \mathrm{cum}\left(X_{ij}; ij \in P_M\right)$$

where the summation is over all indecomposable partitions $P = P_1 \cup \cdots \cup P_M$.

Example 1 Consider the simple case of $\mathrm{cum}\big(x_t^2, x_{t-l}^2\big)$ in $s_{1,1}^{2,2}$. Then in the notation of Lemma 2, $Y_1 = X_{11}X_{12}$ and $Y_2 = X_{21}X_{22}$, which correspond to x_t^2 and x_{t-l}^2 respectively. Applying Lemma 1 and making use of the fact that $\mathrm{E}(x_t) = \mathrm{E}(\widetilde{x}_{t-l}) = 0$,

$$\begin{aligned} \mathrm{cum}\left(Y_1, Y_2\right) &= \mathrm{cum}\left(X_{11}, X_{12}, X_{21}, X_{22}\right) \\ &\quad + \mathrm{cum}\left(X_{11}, X_{21}\right)\mathrm{cum}\left(X_{12}, X_{22}\right) + \mathrm{cum}\left(X_{11}, X_{22}\right)\mathrm{cum}\left(X_{12}, X_{21}\right), \end{aligned}$$

which gives rise to

$$\operatorname{cum}\left(x_t^2, x_{t-l}^2\right) = \operatorname{cum}\left(x_t, x_t, x_{t-l}, x_{t-l}\right) + 2\operatorname{cum}\left(x_t, x_{t-l}\right)^2. \tag{53}$$

Note that $\operatorname{cum}(x_t, x_t)\operatorname{cum}(x_{t-l}, x_{t-l})$ is not an indecomposable partition because there is no cumulant term that links the x_t^2 and x_{t-l}^2 together.

A.3 Proofs for $S_{1,h}^{p,q}$ and $S_{h,h}^{p,q}$

Here, we first introduce some preliminary results, a notation to simplify the presentation of proofs, and then proceed to use Lemma 2 to derive the required covariances involving multiperiod returns.

A.3.1 Preliminary results

There are two properties of x_t which render the derivation of covariance matrices S_0 relatively straightforward. Firstly, $\mathrm{E}(x_t) = 0$. Secondly, x_t and x_{t-l} are independent except for $l = 0$. The first property enables us to ignore all indecomposable partitions that result in $\mathrm{E}(x_t)$ as a cumulant term. By virtue of Lemma 1, the second property implies that for j random variable x's at time t or $t-l$, we have

$$\operatorname{cum}\left(x_t, \ldots, x_{t-l}\right) = \begin{cases} \kappa_j & \text{if } l = 0, \\ 0 & \text{if } l \neq 0. \end{cases} \tag{54}$$

If the j random variables are a mixture of x_t's and h-period random returns $\widetilde{x}_{t-l}$'s,

$$\operatorname{cum}\left(x_t, \ldots, \widetilde{x}_{t-l}\right) = \begin{cases} \kappa_j & \text{for } 1-h \leq l \leq 0, \\ 0 & \text{for } l > 0. \end{cases} \tag{55}$$

Finally, for j h-period random returns $\widetilde{x}$'s at time t or $t-l$,

$$\operatorname{cum}\left(\widetilde{x}_t, \ldots, \widetilde{x}_{t-l}\right) = \begin{cases} (h - |l|)\,\kappa_j & \text{for } |l| < h, \\ 0 & \text{for } |l| \geq h. \end{cases} \tag{56}$$

A.3.2 Notation

To derive the required covariances of multiperiod returns, it is helpful to simplify the notation in the following way. We denote the j-th order joint cumulant of random variables $y_1, .., y_j$ by $\langle y_1 \cdots y_j \rangle$, that is

$$\operatorname{cum}\left(y_1, .., y_j\right) = \langle y_1 \cdots y_j \rangle .$$

Suppose for example $y_1 = y_2 = u$ and $y_3 = \cdots = y_j = v$. Then the cumulant can be simply written as

$$\mathrm{cum}\left(y_1, .., y_j\right) = \left\langle u^2 v^{j-2} \right\rangle .$$

Note that $\langle \cdot \rangle$ does not represent the cumulant of the products of random variables; for instance, $\langle x^3 \rangle = \mathrm{cum}(x, x, x) \neq \mathrm{cum}(x^3)$.

A.3.3 Covariances for skewness ratio test

The covariances between single-period returns are already provided in A.1. Next, we consider covariances that involve h-period returns. First, consider the simple case of $s_{1,h}^{1,3} = \sum \mathrm{cum}\left(x_t, \widetilde{x}_{t-l}^3\right)$. Applying Lemma 2,

$$\mathrm{cum}\left(x_t, \widetilde{x}_{t-l}^3\right) = \left\langle x_t \widetilde{x}_{t-l}^3 \right\rangle + 3 \left\langle x_t \widetilde{x}_{t-l} \right\rangle \left\langle \widetilde{x}_{t-l}^2 \right\rangle .$$

According to (55) and (56), $\mathrm{cum}\left(x_t, \widetilde{x}_{t-l}^3\right) = \kappa_4 + 3h\sigma^4$ for $1 - h \leq l \leq 0$, zero otherwise. So

$$s_{1,h}^{1,3} = h\left[\kappa_4 + 3h\sigma^4\right]$$

Similarly, for $1 - h \leq l \leq 0$,

$$\begin{aligned}
\mathrm{cum}\left(x_t^3, \widetilde{x}_{t-l}^3\right) &= \left\langle x_t^3 \widetilde{x}_{t-l}^3 \right\rangle + 3\left\langle x_t^3 \widetilde{x}_{t-l} \right\rangle \left\langle \widetilde{x}_{t-l}^2 \right\rangle + 3\left\langle x_t \widetilde{x}_{t-l}^3 \right\rangle \left\langle x_t^2 \right\rangle \\
&\quad +9\left\langle x_t^2 \widetilde{x}_{t-l}^2 \right\rangle \left\langle x_t \widetilde{x}_{t-l} \right\rangle + 9\left\langle x_t^2 \widetilde{x}_{t-l} \right\rangle \left\langle x_t \widetilde{x}_{t-l}^2 \right\rangle \\
&\quad +9\left\langle x_{t-l}^2 \right\rangle \left\langle x_t \widetilde{x}_{t-l} \right\rangle \left\langle \widetilde{x}_{t-l}^2 \right\rangle + 6\left\langle x_t \widetilde{x}_{t-l} \right\rangle \left\langle x_t \widetilde{x}_{t-l} \right\rangle \left\langle x_t \widetilde{x}_{t-l} \right\rangle \qquad (57) \\
&= \kappa_6 + (3h + 12)\kappa_4 \sigma^2 + 9\kappa_3^2 + (9h + 6)\sigma^6,
\end{aligned}$$

which if multiplied by h gives rise to $s_{1,h}^{3,3}$. To see how the number of each type of indecomposable partition is obtained in (57), take $\langle x_t^2 \widetilde{x}_{t-l} \rangle \left\langle x_t \widetilde{x}_{t-l}^2 \right\rangle$ as an example: there are three ways of choosing x_t^2 from x_t^3 and three ways of choosing $\widetilde{x}_{t-l}$ from $\widetilde{x}_{t-l}^3$ to yield $\langle x_t^2 \widetilde{x}_{t-l} \rangle$; there is only one left way for the remaining random variables to form $\left\langle x_t \widetilde{x}_{t-l}^2 \right\rangle$. So, the required number is $3 \times 3 \times 1 = 9$.

Now in the case of $\mathrm{cum}\left(\widetilde{x}_t^3, \widetilde{x}_{t-l}^3\right)$ in $s_{h,h}^{3,3}$, each term in the sum of products of cumulants will retain the *same form* as the right hand side of (57), and replacing x_t with $\widetilde{x}_t$ yields the expression for $\mathrm{cum}\left(\widetilde{x}_t^3, \widetilde{x}_{t-l}^3\right)$. So, making use of (56),

$$\begin{aligned}
s_{h,h}^{3,3} &= \sum (h - |l|)\kappa_6 + \left[6h \sum (h - |l|) + 9 \sum (h - |l|)^2\right] \kappa_4 \sigma^2 \\
&\quad +9 \sum (h - |l|)^2 \kappa_3^2 + \left[9h^2 \sum (h - |l|) + 6 \sum (h - |l|)^3\right] \sigma^6,
\end{aligned}$$

where the summation is from $l=-h+1,..,h-1$. Note that $\sum(h-|l|)=h^2$, $\sum(h-|l|)^2=A_h$ and $\sum(h-|l|)^3=B_h$, and this completes the proof for the expression of $s_{h,h}^{3,3}$ in (20).

A.3.4 Covariances for kurtosis ratio test

From the above derivations of $S_{1,h}^{1,3}$ and $S_{1,h}^{3,3}$, we can see that covariances between products of single- and h-period returns yield a simple multiple of h, and provide the basic form for more complex covariances between products of h-period returns. These steps of proof are similar for covariances in the kurtosis ratio test. So we have

$$\begin{aligned}
s_{1,h}^{1,4} &= \sum \operatorname{cum}\left(x_t,\widetilde{x}_{t-l}^4\right) \\
&= \sum\left[\left\langle x_t\widetilde{x}_{t-l}^4\right\rangle+4\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^3\right\rangle+6\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle\right] \\
&= h\left[\kappa_5+10h\kappa_3\sigma^2\right].
\end{aligned}$$

Also, multiplying by h the following cumulant

$$\begin{aligned}
\operatorname{cum}\left(x_t^2,\widetilde{x}_{t-l}^4\right) &= \left\langle x_t^2\widetilde{x}_{t-l}^4\right\rangle+6\left\langle x_t^2\widetilde{x}_{t-l}^2\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle+8\left\langle x_t\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle \\
&\quad +4\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^3\right\rangle+6\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle+12\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle \\
&= \kappa_6+(6h+8)\kappa_4\sigma^2+(4h+6)\kappa_3^2+12h\sigma^6
\end{aligned}$$

yields $s_{1,h}^{2,4}$. The case for $s_{1,h}^{4,4}$ is more complex; the indecomposable partitions of $\operatorname{cum}\left(x_t^4,\widetilde{x}_{t-l}^4\right)$ are

$$\begin{aligned}
&\left\langle x_t^4\widetilde{x}_{t-l}^4\right\rangle+6\left\langle x_t^4\widetilde{x}_{t-l}^2\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle+6\left\langle x_t^2\widetilde{x}_{t-l}^4\right\rangle\left\langle x_t^2\right\rangle+16\left\langle x_t^3\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle \\
&+4\left\langle x_t^4\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^3\right\rangle+4\left\langle x_t\widetilde{x}_{t-l}^4\right\rangle\left\langle x_t^3\right\rangle+24\left\langle x_t^3\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle+24\left\langle x_t^2\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle \\
&+18\left\langle x_t^2\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t^2\widetilde{x}_{t-l}^2\right\rangle+16\left\langle x_t^3\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}^3\right\rangle \\
&+36\left\langle x_t^2\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t^2\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle+48\left\langle x_t^3\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle \\
&+48\left\langle x_t\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t^2\right\rangle+72\left\langle x_t^2\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle \\
&+16\left\langle x_t^3\right\rangle\left\langle\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle+24\left\langle x_t^3\right\rangle\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle+24\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t^2\right\rangle \\
&+36\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle+36\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t^2\right\rangle+144\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle \\
&+72\left\langle x_t^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle\widetilde{x}_{t-l}^2\right\rangle+24\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle.
\end{aligned} \tag{58}$$

In the above, only $\left\langle \widetilde{x}_{t-l}^{2}\right\rangle$ and $\left\langle \widetilde{x}_{t-l}^{3}\right\rangle$ yield a factor h. Thus

$$\begin{aligned}\operatorname{cum}\left(x_t^4, \widetilde{x}_{t-l}^4\right) &= \kappa_8 + (6h+22)\,\kappa_6\sigma^2 + (4h+52)\,\kappa_5\kappa_3 + 34\kappa_4^2 \\ &\quad + (84h+120)\,\kappa_4\sigma^4 + (100h+180)\,\kappa_3^2\sigma^2 + (72h+24)\,\sigma^8,\end{aligned}$$

for $l = 1-h, .., 0$. Thus, multiplying the above by h yields $s_{1,h}^{4,4}$. Replacing x_t^4 with $\widetilde{x}_t^4$ in (58) gives us $\operatorname{cum}\left(\widetilde{x}_t^4, \widetilde{x}_{t-l}^4\right)$ which, after applying the result of (56), yields

$$\begin{aligned}&(h-|l|)\,\kappa_8 + \left[12h\,(h-|l|) + 16\,(h-|l|)^2\right]\kappa_6\sigma^2 \\ &+ \left[8h\,(h-|l|) + 48\,(h-|l|)^2\right]\kappa_5\kappa_3 + 34\,(h-|l|)^2\,\kappa_4^2 \\ &+ \left[36h^2\,(h-|l|) + 96\,(h-|l|)^2 + 72\,(h-|l|)^3\right]\kappa_4\sigma^4 \\ &+ \left[64h^2\,(h-|l|) + 72h\,(h-|l|)^2 + 144\,(h-|l|)^3\right]\kappa_3^2\sigma^2 \\ &+ \left[72h^2\,(h-|l|)^2 + 24\,(h-|l|)^4\right]\sigma^8.\end{aligned}$$

Summing the above from $l = -h+1$ to $h-1$ and noting $\sum_{l=-h+1}^{h-1}(h-|l|)^4 = C(h)$, we have the required covariance.

A.3.5 Covariances for the joint skewness and kurtosis ratio test

The remaining covariances to be derived for the joint skewness and kurtosis ratio test are $s_{1,h}^{2,3}$, $s_{1,h}^{4,3}$, $s_{1,h}^{3,4}$ and $s_{h,h}^{3,4}$. Using the same method as above,

$$\begin{aligned}s_{1,h}^{2,3} &= \sum \operatorname{cum}\left(x_t^2, \widetilde{x}_{t-l}^3\right) \\ &= \sum\left[\left\langle x_t^2\widetilde{x}_{t-l}^3\right\rangle + 6\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle + 3\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle \widetilde{x}_{t-l}^2\right\rangle\right] \\ &= h\left[\kappa_5 + (3h+6)\,\kappa_3\sigma^2\right]\end{aligned}$$

For $s_{1,h}^{4,3}$, applying the indecomposable partition method for $\operatorname{cum}\left(x_t^4, \widetilde{x}_{t-l}^3\right)$ yields

$$\begin{aligned}&\left\langle x_t^4\widetilde{x}_{t-l}^3\right\rangle + 3\left\langle x_t^4\widetilde{x}_{t-l}\right\rangle\left\langle \widetilde{x}_{t-l}^2\right\rangle + 6\left\langle x_t^2\widetilde{x}_{t-l}^3\right\rangle\left\langle x_{t-l}^2\right\rangle + 12\left\langle x_t^3\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle \\ &+4\left\langle x_t\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t^3\right\rangle + 12\left\langle x_t^3\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle + 18\left\langle x_t^2\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle \\ &+12\left\langle x_t^3\right\rangle\left\langle \widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle + 18\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle x_t^2\right\rangle\left\langle \widetilde{x}_{t-l}^2\right\rangle \\ &+36\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle + 36\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle \\ =\;& \kappa_7 + (3h+18)\,\kappa_5\sigma^2 + 34\kappa_4\kappa_3 + (30h+72)\,\kappa_3\sigma^4.\end{aligned}$$

Multiplying the above result by a factor of h gives rise to $s_{1,h}^{4,3}$. $s_{1,h}^{3,4}$ is a mirror image of $s_{1,h}^{4,3}$, so we have

$$\begin{aligned}
s_{1,h}^{3,4} &= \sum \Big[\left\langle x_t^3\widetilde{x}_{t-l}^4\right\rangle + 6\left\langle x_t^3\widetilde{x}_{t-l}^2\right\rangle\left\langle \widetilde{x}_{t-l}^2\right\rangle + 3\left\langle x_t\widetilde{x}_{t-l}^4\right\rangle\left\langle x_t^2\right\rangle + 12\left\langle x_t^2\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle \\
&\quad +4\left\langle x_t^3\widetilde{x}_{t-l}\right\rangle\left\langle \widetilde{x}_{t-l}^3\right\rangle + 12\left\langle x_t\widetilde{x}_{t-l}^3\right\rangle\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle + 18\left\langle x_t^2\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle \\
&\quad +12\left\langle \widetilde{x}_{t-l}^3\right\rangle\left\langle x_t^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle + 18\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t^2\right\rangle\left\langle \widetilde{x}_{t-l}^2\right\rangle \\
&\quad +36\left\langle x_t^2\widetilde{x}_{t-l}\right\rangle\left\langle \widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle + 36\left\langle x_t\widetilde{x}_{t-l}^2\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\left\langle x_t\widetilde{x}_{t-l}\right\rangle\Big] \\
&= h\left[\kappa_7 + (6h+15)\,\kappa_5\sigma^2 + (4h+30)\,\kappa_4\kappa_3 + (66h+36)\,\kappa_3\sigma^4\right]
\end{aligned}$$

Replacing the x_t in the above with $\widetilde{x}_t$ yields the required $s_{h,h}^{3,4}$:

$$\begin{aligned}
s_{h,h}^{3,4} &= \sum \Big[(h-|l|)\,\kappa_7 + \left(9h\,(h-|l|) + 12\,(h-|l|)^2\right)\kappa_5\sigma^2 \\
&\quad + \left(4h\,(h-|l|) + 30\,(h-|l|)^2\right)\kappa_4\kappa_3 \\
&\quad + \left(30h^2\,(h-|l|) + 36h\,(h-|l|)^2 + 36\,(h-|l|)^3\right)\kappa_3\sigma^4\Big] \\
&= h^2\kappa_7 + \left[9h^3 + 12A_h\right]\kappa_5\sigma^2 + \left[4h^3 + 30A_h\right]\kappa_4\kappa_3 \\
&\quad + \left[30h^4 + 36hA_h + 36B_h\right]\kappa_3\sigma^4,
\end{aligned}$$

and this completes the proofs.

A.4 Estimation of cumulants

The covariance matrix S_0 is expressed in terms of cumulants, which in practice can be estimated using central moments as shown below; see Kendall and Stuart (1977, p.71). Note that $\kappa_2 = \mu_2$ and $\kappa_3 = \mu_3$.

$$\kappa_4 = \mu_4 - 3\sigma^4, \tag{59}$$

$$\kappa_5 = \mu_5 - 10\mu_3\sigma^2, \tag{60}$$

$$\kappa_6 = \mu_6 - 15\mu_4\sigma^2 - 10\mu_3^2 + 30\sigma^6, \tag{61}$$

$$\kappa_7 = \mu_7 - 21\mu_5\sigma^2 - 35\mu_4\mu_3 + 210\mu_3\sigma^4, \tag{62}$$

$$\kappa_8 = \mu_8 - 28\mu_6\sigma^2 - 56\mu_5\mu_3 - 35\mu_4^2 + 420\mu_4\sigma^4 + 560\mu_3^2\sigma^2 - 630\sigma^8. \tag{63}$$

References

Abhyankar, A., L.S. Copeland, and W.K. Wong. 1995. Nonlinear dynamics in real-time equity market indices: Evidence from the United Kingdom. *The Economic Journal* 105: 864–880.

Abhyankar, A., L.S. Copeland, and W.K. Wong. 1997, Uncovering Nonlinear Structure In Real-Time Stock-Market Indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. *Journal of Business & Economic Statistics* 15: 1–14.

Basel Committee on Banking Supervision. 2016. Minimum capital requirements for market risk. Available at: *www.bis.org/bcbs/publ/d352.pdf* [Accessed 22 July 2018].

Beckers, B., H. Herwartz, and M. Seidel. 2017. Risk forecasting in (T)GARCH models with uncorrelated dependent innovations. *Quantitative Finance* 17: 121–137.

Brillinger, D.R. 1975. *Time Series: Data Analysis and Theory.* Holt, Rinehart and Winston, New York.

Brock, W.A., W. Dechert, J.A. Scheinkman, and B. LeBaron. 1996. A test for independence based on the correlation dimension. *Econometric Reviews* 15: 197–235.

Cont, R. 2001. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* 1: 223–236.

Dupuis, D.J., N. Papageorgiou, and B. Remillard. 2015. Robust conditional variance and Value-at-Risk estimation, *Journal Financial Econometrics* 13: 896 – 921.

Hansen, L.P. 1982. Large sample properties of generalized method of moment estimators. *Econometrica* 50: 1029–1054.

Hansen, B.E. 1994. Autoregressive Conditional Density Estimation. *International Economic Review* 35: 705–730.

Hansen, L., and K. Singleton. 1982. Generalized institutional variables estimation of non-linear rational expectations models. *Econometrica* 50: 1269–1286.

Hawawini, G.A. 1980. An analytical examination of the intervalling effect on skewness and other moments. *Journal of Financial and Quantitative Analysis* 15: 1121–1127.

Hinich, M.J. 1982. Testing for Gaussianity and linearity of a stationary time series. *Journal of Time Seies Analysis* 3: 169–176.

Hinich, M.J., and D.M. Patterson. 1985. Evidence of nonlinearity in daily stock returns. *Journal of Business & Economic Statistics* 3: 69–77.

Hsieh, D.A. 1989. Testing for nonlinear dependence in daily foreign exchange rates. *Journal of Business* 62: 339–368.

Hsieh, D.A. 1993. Implications of nonlinear dynamics for financial risk management. *Journal of Financial and Quantitative Analysis* 28: 41–64.

Jarque, C.M., A.K. Bera. 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* 6: 255–259.

Kendall, M.G., and A. Stuart. 1977. *The Advanced Theory of Statistics. Vol. 1: Distribution Theory.* London: Griffin.

Kupiec, P. 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 3: 73–84.

Lau, H., and J.R. Wingender. 1989. The Analytics of the intervalling effect on skewness and kurtosis of stock returns. *Journal of Financial and Quantitative Analysis* 24: 215–233.

Li, W.K., and T.K. Mak. 1994. On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. *Journal of Time Series Analysis* 15: 627–636.

Lo, A., and C. MacKinlay. 1988. Stock prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies* 1: 41–66.

Ljung, G.M., and G.E.P. Box. 1978. On a measure of a lack of fit in time series models. *Biometrika* 65: 297–303.

Mancini, L., and F. Trojani. 2011. Robust Value at Risk prediction. *Journal of Financial Econometrics* 9: 1–33.

McCaffrey, D.F., S. Ellner, A.R. Gallant, and D.W. Nychka. 1992. Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. J*ournal of the American Statistical Association* 87: 682–95.

McLeod, A.I., and W.K. Li. 1983. Diagnostic checking arma time series models using squared-residual autocorrelations. *Journal of Time Series Analysis* 4: 269–273.

Newey, W., and K. West. 1987. A simple, positive-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703 – 708.

Richardson, M., and T. Smith. 1991. Tests of financial models in the presence of overlapping observations. *Review of Financial Studies* 4: 227–254.

Subba Rao, T., and M.M. Gabr. 1980. A test for linearity of stationary time series. *Journal of Time Series Analysis* 1: 145–158.

Tse, Y.K., and A.K. Tsui. 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics* 20: 351–362.

Wong, W.K. 1997. Frequency domain tests of multivariate Gaussianity and linearity. *Journal of Time Series Analysis* 18: 181–194.

Wong, W.K. 2008. Backtesting trading risk of commercial banks using expected shortfall. *Journal of Banking and Finance* 32: 1404–1415.

Wong, W.K. 2010. Backtesting value-at-risk based on tail losses. *Journal of Empirical Finance* 17: 526–538