

Fourth moments of multivariate GARCH processes

Christian M. Hafner*

September 2000

Abstract

This paper derives conditions for the existence of fourth moments of multivariate GARCH processes in the general vector specification and gives explicit results for the fourth moments and autocovariances of the squares and cross-products. Results are provided for the kurtosis and co-kurtosis between components. An impulse response function for kurtosis and co-kurtosis is defined that allows to analyse the expectation of the (co-)kurtosis conditional on an initial shock. For a bivariate exchange rate series, these functions indicate that there is a trade-off between conditional variance and conditional kurtosis: the conditional variance increases with the size of the shocks, but the conditional kurtosis decreases.

Keywords: *multivariate GARCH, fourth moments, kurtosis, co-kurtosis, impulse response function*

JEL classification: C22

*Electrabel, R&D Energy Markets, Traverse d'Esoppe 6, B-1348 Louvain-la-Neuve, Belgium.

1 Introduction

Multivariate GARCH models have been frequently used in modelling volatility of multivariate time series. To give one example of the numerous applications in finance, Bollerslev, Engle and Wooldridge (1988) used a multivariate GARCH-M model to estimate time-varying betas within the capital asset pricing model. GARCH-type models capture conditional heteroskedasticity and leptokurtosis, two common features of financial data. Moreover, the statistical theory is developing. For example, Jeanteau (1998) has proved strong consistency of quasi maximum likelihood (QML) estimators of multivariate GARCH models, extending previous results of Lee and Hansen (1994) and Lumsdaine (1996) on univariate GARCH models.

An often used measure for the thickness of the tails is the kurtosis, or the standardized fourth moment. When fitting a model, one may therefore want to reproduce not only the first two moments, but also the fourth moment. For example, having estimated the model based on QML, the question arises which innovation distribution one should use in simulation studies such that the empirical kurtosis is approximated. In finance, this is particularly important for option pricing, where the degree of excess kurtosis determines the shape of the so-called smile. A second example is the estimation by generalized method of moments (GMM), where the efficiency may often be improved using correctly specified skewness and kurtosis, see e.g. Meddahi and Renault (1997) and Broze, Francq and Zakoian (2000). A last example is to use the kurtosis in hedging strategies, as shown by Mahieu (1995, ch.6) for foreign exchange risks.

Fourth moments of univariate GARCH models were investigated by He and Teräsvirta (1999) and Karanasos (1999). This paper provides results for the general multivariate GARCH(p, q) process in vector specification. This specification nests all popular linear specifications such as the BEKK specification of Engle and Kroner (1995), the constant conditional correlation model of Bollerslev (1990), and the factor GARCH models of Diebold and Nerlove (1989) and Engle, Ng and Rothschild (1990). In Section 2, we provide explicit results for the case of Gaussian innovations. In Section 3, we discuss the relationship between the (co-)kurtosis and measures for conditional heteroskedasticity. Furthermore, impulse response functions are defined for the kurtosis and the co-kurtosis. In Section 4, an empirical example of a bivariate exchange rate series shows that multi-step predictions of kurtosis have the inverse shape of corresponding volatility predictions: they are small for large initial shocks and large for small shocks. Proofs of the theorems are provided in the appendix.

2 Main results

Consider the general multivariate GARCH(p, q) model for the vector of a random error term ε_t with K components,

$$\varepsilon_t = H_t^{1/2} \xi_t \tag{1}$$

$$h_t = \text{vech}(H_t) = \omega + \sum_{i=1}^q A_i \eta_{t-i} + \sum_{j=1}^p B_j \text{vech}(H_{t-j}) \quad (2)$$

where $\omega = \text{vech}(\Omega)$, $\eta_t = \text{vech}(\varepsilon_t \varepsilon_t')$ and $N \times N$, $N = K(K+1)/2$ parameter matrices Ω , A_i , B_j . Throughout the paper, vec denotes the operator that stacks all columns of a matrix into a vector, and vech denotes the operator that stacks only the lower triangular part including the diagonal of a symmetric matrix into a vector. The innovation vector ξ_t is assumed to be i.i.d. with mean zero and identity covariance matrix. Denoting the information set available at time t by \mathcal{F}_t , we have the conditional moments $E[\varepsilon_t | \mathcal{F}_{t-1}] = 0$ and $\text{Var}(\varepsilon_t | \mathcal{F}_{t-1}) = H_t$. A sufficient condition for the conditional covariance matrix H_t to be positive definite is that each of the parameter matrices is positive definite. Note that (1) is a strong GARCH model by the definition of Drost and Nijman (1993).

By rearranging terms, the multivariate GARCH(p, q) model can be represented as a VARMA ($\max(p, q), p$) model,

$$\eta_t = \omega + \sum_{i=1}^{\max(p, q)} (A_i + B_i) \eta_{t-i} - \sum_{j=1}^p B_j u_{t-j} + u_t, \quad (3)$$

where $u_t = \eta_t - h_t$ is a white noise vector, i.e., $E[u_t] = 0$, $E[u_t u_t'] = \Sigma_u$ and $E[u_t u_s'] = 0$ for $s \neq t$. In (3), we set $A_{q+1} = \dots = A_p = 0$ if $p > q$ and $B_{p+1} = \dots = B_q = 0$ if $q > p$.

Assumption 1 *All eigenvalues of the matrix $\sum_{i=1}^{\max(p, q)} (A_i + B_i)$ have modulus smaller than one.*

The multivariate GARCH(p, q) model is covariance stationary if and only if Assumption 1 holds. In that case, the unconditional covariance matrix $\Sigma = \text{Var}(\varepsilon_t)$ is given by

$$\sigma = \text{vech}(\Sigma) = \left(I_N - \sum_{i=1}^{\max(p, q)} (A_i + B_i) \right)^{-1} \omega.$$

From the VARMA representation (3) it is possible to obtain the pure VMA(∞) representation

$$\eta_t = \sigma + \sum_{i=0}^{\infty} \Phi_i u_{t-i},$$

where the $N \times N$ matrices Φ_i can be determined recursively by $\Phi_0 = I_N$, $\Phi_i = -B_i + \sum_{j=1}^i (A_j + B_j) \Phi_{i-j}$, $i = 1, 2, \dots$, see Lütkepohl (1993, pp. 220). Now, the autocovariance of η_t is given by

$$\begin{aligned} \Gamma(\tau) &= E[(\eta_t - \sigma)(\eta_{t-\tau} - \sigma)'] \\ &= \sum_{i=0}^{\infty} \Phi_{\tau+i} \Sigma_u \Phi_i'. \end{aligned} \quad (4)$$

Using the notation $\Sigma_\eta = E[\eta_t \eta_t']$ and $\Sigma_h = E[h_t h_t']$, we can write $\Sigma_u = \Sigma_\eta - \Sigma_h$, and $\Gamma(0) = \Sigma_\eta - \sigma \sigma'$.

Autocorrelations are obtained by

$$R(\tau) = W^{-1}\Gamma(\tau)W^{-1}, \quad (5)$$

where W is a diagonal matrix with the square roots of the diagonal elements of $\Gamma(0)$ on its diagonal.

In order to give explicit results, one has to assume a specific distribution for the innovations ξ_t . In the following, we will assume the most popular one, a multivariate normal distribution. Extensions to other distributions such as the multivariate t -distribution are left to future research.

Assumption 2 *The innovations ξ_t are normally distributed, $\xi_t \sim N(0, I_K)$.*

Theorem 1 *Under Assumption 2,*

$$\text{vec}(\Sigma_\eta) = G_K \text{vec}(\Sigma_h), \quad (6)$$

where

$$G_K = 2(L_K \otimes L_K)(I_{K^2} \otimes N_K)(I_K \otimes C_{KK} \otimes I_K)(D_K \otimes D_K) + I_{N^2},$$

with L_m, D_m, C_{mn} denoting the elimination, duplication and commutation matrices, respectively, and $N_n = (I_{n^2} + C_{mn})/2$.

See Magnus (1988) for definition and properties of L_m, D_m, C_{mn} , and N_n . The matrix G_K is square of order N^2 . For example, in the bivariate case it is given by

$$G_2 = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

As a corollary, we obtain $\text{vec}(\Sigma_u) = (G_K - I_{N^2})\text{vec}(\Sigma_h)$. This can be used to calculate the autocovariances according to (4).

Theorem 2 *Under Assumption 2, fourth moments of the multivariate GARCH(p, q) model are finite, if and only if all eigenvalues of the matrix $\sum_{i=1}^{\infty} (\Phi_i \otimes \Phi_i)(G_K - I_{N^2})$ have modulus smaller than one. In that case, the fourth moment is given by*

$$\text{vec}(\Sigma_\eta) = G_K \left(I_{N^2} - \sum_{i=1}^{\infty} (\Phi_i \otimes \Phi_i)(G_K - I_{N^2}) \right)^{-1} \text{vec}(\sigma\sigma'). \quad (7)$$

Note that in the special case where all A_i and B_j are zero, i.e. the case of a Gaussian white noise vector process, all Φ_i 's are zero and, hence, $\text{vec}(\Sigma_\eta) = G_K \text{vec}(\sigma\sigma')$.

If models of small order are considered, there are ways to find simpler expressions than (7). For the popular multivariate GARCH(1,1) model, we can derive the following results.

Theorem 3 *Under Assumption 2, fourth moments of the multivariate GARCH(1,1) model are finite, if and only if all eigenvalues of the matrix*

$$Z = (A \otimes A)G_K + A \otimes B + B \otimes A + B \otimes B \quad (8)$$

have modulus smaller than one. In that case, the fourth moment is given by

$$\text{vec}(\Sigma_\eta) = G_K (I_{N^2} - Z)^{-1} \text{vec}(\omega\omega' + \omega\sigma'(A+B)' + (A+B)\sigma\omega'). \quad (9)$$

Note that in the univariate case, $K = 1$, we have $G_1 = 3$ and obtain as a special case of equation (9) with obvious notation

$$\mathbb{E}[\varepsilon_t^4] = 3 \frac{\omega^2 + 2\omega\sigma(\alpha + \beta)}{1 - 3\alpha^2 - 2\alpha\beta - \beta^2},$$

where $\sigma = \omega/(1 - \alpha - \beta)$, and finally the well-known formula for the kurtosis of a univariate GARCH(1,1) process

$$\frac{\mathbb{E}[\varepsilon_t^4]}{\mathbb{E}[\varepsilon_t^2]^2} = 3 + \frac{6\alpha^2}{1 - 3\alpha^2 - 2\alpha\beta - \beta^2}.$$

3 Impulse response functions for kurtosis and co-kurtosis

In practice, one often considers fourth moments that are normalized with the square of the variance. In the multivariate case, Mardia (1970) has defined a measure of multivariate kurtosis for i.i.d. random variables. Suppose that X is a p -dimensional random vector with mean μ and covariance matrix Σ . Mardia's measure is given by

$$\beta_{2,p} = \mathbb{E}[\{(X - \mu)' \Sigma^{-1} (X - \mu)\}^2],$$

which is the mean of the squared Mahalanobis distance of X from its mean μ . The measure $\beta_{2,p}$ is invariant with respect to affine transformations and reduces to the standard kurtosis in the univariate case.

In our setting with $\mu = 0$, conditional and unconditional covariance matrix Σ_t and Σ , respectively, an analogous measure would be

$$\beta_{2,p} = \mathbb{E}[\{\varepsilon_t' \Sigma^{-1} \varepsilon_t\}^2]. \quad (10)$$

The reason for using Σ rather than Σ_t is that in the latter case the multivariate kurtosis would be trivial, i.e. only depending on the distribution of ξ_t , but not on the dynamics of Σ_t . For the notion (10), however, it is difficult to derive explicit results in the multivariate GARCH case.

In the following, we will therefore consider alternative measures based on the fourth moments of single components of the process. These measures are not invariant with respect to general affine transformations, but only with respect to scalar transformations.

Definition 1 *Let the kurtosis of the i – th component of ε_t be defined as*

$$k_{ii} = \frac{E[\varepsilon_{it}^4]}{E[\varepsilon_{it}^2]^2}.$$

The extension to a co-kurtosis between two components is given in the following definition.

Definition 2 *Let the co-kurtosis of the i -th and j -th component of ε_t be defined as*

$$k_{ij} = \frac{E[\varepsilon_{it}^2 \varepsilon_{jt}^2]}{E[\varepsilon_{it}^2]E[\varepsilon_{jt}^2]}.$$

Let us elaborate a little further on the notion of co-kurtosis. For a Gaussian i.i.d. process, it is well-known that the co-kurtosis takes the value $k_{ij} = 1 + 2\rho_{ij}^2$, where ρ_{ij} is the correlation between the i – th and j – th component. In the case of conditional heteroskedasticity, the following theorem gives useful expressions for the kurtosis and the co-kurtosis.

Theorem 4 *Under assumption 2, the kurtosis and co-kurtosis can be expressed as*

$$k_{ii} = 3 + \frac{3 \text{Var}(h_{ii,t})}{\sigma_{ii}^2},$$

$$k_{ij} = 1 + 2\rho_{ij}^2 + \frac{2 \text{Var}(h_{ij,t}) + \text{Cov}(h_{ii,t}, h_{jj,t})}{\sigma_{ii}\sigma_{jj}},$$

where $h_{ij,t}$ is the ij -th element of H_t and σ_{ii} the ii -th element of Σ .

The formula for k_{ii} can already be found in Gouriéroux (1997, pp.38). The kurtosis is linked to a natural measure of conditional heteroskedasticity, $\text{Var}(h_{ii,t})/\sigma_{ii}^2$. Whenever the conditional variance is stochastic, the kurtosis is larger than three.

The co-kurtosis is linked to the covariance between the conditional variances and to the variance of the conditional covariance. The stronger the conditional variances are correlated and the higher the variance of the conditional covariance, the higher also the co-kurtosis.

In the following, we investigate the behavior of the expected kurtosis conditional on an initial shock. In a multivariate framework, a shock can occur in either one of the components, provided that it is independent of the other components. This is essentially the definition of news given by Hafner and Herwartz (1998a), as opposed to the definition of Gallant, Rossi and Tauchen (1993), who let shocks occur in the dependent error term ε_t . In our case, a shock appears in the i.i.d. innovation vector ξ_t , where one can arbitrarily consider alternative scenarios without getting into conflict with dependence.

Definition 3 We define the impulse response function of the kurtosis as

$$\kappa_{s,ii}(\xi_t) = \frac{E[\varepsilon_{i,t+s}^4 \mid \xi_t, H_t]}{E[\varepsilon_{i,t+s}^2 \mid \xi_t, H_t]^2} \quad (11)$$

and of the co-kurtosis as

$$\kappa_{s,ij}(\xi_t) = \frac{E[\varepsilon_{i,t+s}^2 \varepsilon_{j,t+s}^2 \mid \xi_t, H_t]}{E[\varepsilon_{i,t+s}^2 \mid \xi_t, H_t] E[\varepsilon_{j,t+s}^2 \mid \xi_t, H_t]}.$$

The state of H_t will usually be fixed at the steady state, i.e., $H_t = \Sigma$. Then, a variety of independent shocks ξ_t may be considered.

Note that $\kappa_{1,ii}(\xi_t) = 3$ and $\kappa_{1,ij}(\xi_t) = 1 + 2\text{Corr}(\varepsilon_{i,t+1}, \varepsilon_{j,t+1} \mid \mathcal{F}_t)^2$. Furthermore, it is interesting to investigate the behavior of $\kappa_{s,ii}(\xi_t)$ and $\kappa_{s,ij}(\xi_t)$ for $s \rightarrow \infty$. In analogy to the definition of persistence in variance by Bollerslev and Engle (1993), we give in the following a definition of persistence in kurtosis.

Definition 4 The stochastic process ε_t is defined to be persistent in kurtosis if

$$\limsup_{s \rightarrow \infty} |\kappa_{s,ij}(\xi_t) - k_{ij}| \neq 0 \quad a.s.$$

for some $i, j = 1, \dots, K$.

The multivariate GARCH(p, q) model is persistent in kurtosis if at least one of the eigenvalues of the matrix $\sum_{i=1}^{\infty} (\Phi_i \otimes \Phi_i)(G_K - I_{N^2})$ has a modulus equal or greater than one. In the case of the multivariate GARCH(1, 1) model, the relevant matrix simplifies to Z , given in (8). If ε_t is not persistent in kurtosis, the closeness of the eigenvalue with maximum norm to unity may be considered as a measure for the ‘degree of persistence’.

The nominator of (11) is the square of the volatility impulse response function as defined by Hafner and Herwartz (1998a). It is given by the ii -th element of $V_s(\xi_t)$,

$$V_s(\xi_t) = E[\varepsilon_{t+s} \varepsilon'_{t+s} \mid \xi_t, H_t = \Sigma].$$

The impulse response functions are easily calculated recursively. For example, for the multivariate GARCH(1,1) model we have

$$\text{vech}(V_1(\xi_t)) = \omega + A \text{vech}(\Sigma^{1/2} \xi_t \xi_t' \Sigma^{1/2}) + B \text{vech}(\Sigma),$$

and, for $s > 1$,

$$\text{vech}(V_s(\xi_t)) = \omega + (A + B) \text{vech}(V_{s-1}(\xi_t)).$$

For the fourth moment, we denote

$$F_s(\xi_t) = E[\eta_{t+s} \eta'_{t+s} \mid \xi_t, H_t = \Sigma]$$

and obtain $\text{vec}(F_s(\xi_t)) = G_K \text{vec}(F_s^*(\xi_t))$ with

$$\text{vec}(F_1^*(\xi_t)) = \text{vec}(\text{vech}(V_1(\xi_t))\text{vech}(V_1(\xi_t))')$$

and, for $s > 1$,

$$\text{vec}(F_s^*(\xi_t)) = \gamma + Z \text{vec}(F_{s-1}^*(\xi_t)),$$

where Z is given in (8), and $\gamma = \text{vec}(\omega\omega' + (A+B)\sigma\omega' + \omega\sigma'(A+B)')$. The impulse response functions are then given by

$$\kappa_{s,ii}(\xi_t) = \frac{F_{s,l(i)l(i)}(\xi_t)}{V_{s,ii}^2(\xi_t)},$$

where $l(i) = 1 + (i-1)(K+1)$, and

$$\kappa_{s,ij}(\xi_t) = \frac{F_{s,l(i)l(j)}(\xi_t)}{V_{s,ii}(\xi_t)V_{s,jj}(\xi_t)}.$$

4 Empirical Example

For an empirical illustration, let us consider the bivariate exchange rate series DEM/USD and GBP/USD from december 31, 1979 to april 1, 1994, as discussed by Hafner and Herwartz (1998a). We fitted a BEKK-GARCH(1,1) model to the residuals of a VAR(5) model. The eigenvalues of the matrix Z all lie between 0.914 and 0.976, implying finite fourth moments but a high persistence in kurtosis. The kurtosis as implied by the parameter estimates and assuming that the innovations are Gaussian are 3.50 for DEM/USD, 3.80 for GBP/USD and 2.52 for the co-kurtosis. This is considerably smaller than the empirical kurtosis of 4.94 for the residuals of DEM/USD, 5.47 for the residuals of GBP/USD, and 3.78 for the co-kurtosis of the residuals. It is a well-known empirical effect: conditional heteroskedasticity is often not sufficient to explain the excess kurtosis of financial returns. If one wants to capture distributional properties such as the kurtosis, one may prefer to use fat-tailed distributions for the innovations, such as a multivariate t-distribution.

Table 1 reports the autocorrelations of η_t , calculated using equation (5). For this, the infinite sum was truncated at 1000, with $\Phi_i, i \geq 1000$ being very close to zero. Clearly, the diagonal elements of $R(\tau)$ are higher than the off-diagonal elements. For example, the first order autocorrelation of ε_{1t}^2 (DEM/USD) is 0.1370, of ε_{2t}^2 (GBP/USD) it is 0.1573, and of the cross-product it is 0.1281. For all autocorrelations, the decay of the autocorrelation function is slow, corresponding to the large eigenvalues of the matrix Z .

Impulse response functions for the kurtosis and co-kurtosis are calculated and visualized in the following way: First, one component of ξ_t is fixed to zero, the other varied on an interval $[-5,5]$. Then, $\kappa_{s,ii}$ and $\kappa_{s,ij}$ are calculated as a function of this initial shock vector over a horizon of 100 days, $s = 1, 2, \dots, 100$. The corresponding functions are shown in Figure 1 to Figure 6. Starting from their initial values, i.e., 3 for $\kappa_{s,ii}$ and $1 + 2\text{Corr}(\varepsilon_{i,t+1}, \varepsilon_{j,t+1} \mid \mathcal{F}_t)^2$ for $\kappa_{s,ij}$,

τ	1	2	3	4	5	10	50	100
$\text{Corr}[\varepsilon_{1t}^2, \varepsilon_{1,t-\tau}^2]$	0.137	0.132	0.128	0.123	0.119	0.100	0.026	0.007
$\text{Corr}[\varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{1,t-\tau}^2]$	0.050	0.048	0.047	0.046	0.045	0.040	0.016	0.006
$\text{Corr}[\varepsilon_{2t}^2, \varepsilon_{1,t-\tau}^2]$	0.043	0.042	0.041	0.040	0.039	0.034	0.013	0.005
$\text{Corr}[\varepsilon_{1t}^2, \varepsilon_{1,t-\tau}\varepsilon_{2,t-\tau}]$	0.026	0.027	0.028	0.028	0.029	0.031	0.027	0.014
$\text{Corr}[\varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{1,t-\tau}\varepsilon_{2,t-\tau}]$	0.128	0.125	0.123	0.120	0.117	0.106	0.047	0.019
$\text{Corr}[\varepsilon_{2t}^2, \varepsilon_{1,t-\tau}\varepsilon_{2,t-\tau}]$	0.060	0.060	0.059	0.059	0.058	0.055	0.033	0.016
$\text{Corr}[\varepsilon_{1t}^2, \varepsilon_{2,t-\tau}^2]$	0.037	0.038	0.039	0.039	0.040	0.042	0.040	0.023
$\text{Corr}[\varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{2,t-\tau}^2]$	0.045	0.046	0.047	0.048	0.049	0.052	0.049	0.028
$\text{Corr}[\varepsilon_{2t}^2, \varepsilon_{2,t-\tau}^2]$	0.157	0.155	0.152	0.150	0.148	0.136	0.072	0.033

Table 1: *Autocorrelations of η_t with lag τ , calculated using equation (5) and parameter estimates for the bivariate exchange rate series.*

the functions exhibit a typical shape: although increasing for most of the shocks, the predicted kurtosis increases much stronger when the initial shock is close to zero than when it is large. This is the inverse of the typical shape of volatility impulse response functions (Hafner and Herwartz, 1998a). They were shown to increase for large positive or negative shocks. Thus, we can state that for the exchange rates considered, there is a trade-off between predicted volatilities and predicted kurtosis at a horizon larger than one time period.

The multivariate strong GARCH model restricts the conditional one-step kurtosis to a constant, 3 in the case of normal innovations. On the other hand, the above analysis shows that multi-step predictions of the kurtosis are not constant but depend on the initial shock. One may ask if the restriction implied by a strong GARCH model is justified. Indeed, there is empirical evidence for high-frequency exchange rates that also the one-step prediction of the kurtosis is non-trivial, see Bossaerts, Härdle and Hafner (1996) and Hafner (1998, pp.144). In a nonparametric framework, they show that the one-step predicted kurtosis has an inverted U-shape, much like the ones we find here for longer horizons.

A second possibility to visualize impulse response functions is to fix a time horizon and consider alternative scenarios for the two components of ξ_t . This is done in Figure 7 to 9, where we fix s at 10, which corresponds to two weeks, and let $\xi_{1,t}$ and $\xi_{2,t}$ vary between -5 and 5. Both kurtosis functions decrease with large shocks of the same sign in both components. The maximum appears to be close to zero shocks and on a ridge combining shocks of opposite sign. This is different for the co-kurtosis (Figure 9), where the impulse response function increases slightly with shocks of the same sign and decreases strongly with shocks of opposite sign.

5 Concluding remarks

This paper provides explicit results for the fourth moment structure of multivariate GARCH(p, q) processes in their general vector specification. Unlike second moments, fourth moments crucially depend on the innovation distribution, for which we assume a Gaussian distribution. Future research will have to investigate other distributions, in particular distributions with fat tails that allow for a better fit with financial time series. Furthermore, our empirical example of a bivariate exchange rate series showed that there seems to be a trade-off between conditional variance and conditional kurtosis. However, the one-step conditional kurtosis in the strong GARCH specification is restricted to be a constant. It might be interesting to investigate weak GARCH models that explicitly model the conditional kurtosis. The issue of temporal aggregation of multivariate GARCH models has not been addressed yet. Results as in Drost and Nijman (1993) and Drost and Werker (1996) for univariate GARCH models would certainly be useful to investigate the change of kurtosis when the sampling frequency is changed. If the analogy to univariate models holds, then the kurtosis and co-kurtosis should converge to the values for a Gaussian distribution under temporal aggregation. Last, but not least, the results of this paper will have to be extended to multivariate GARCH processes allowing for asymmetric impacts of positive and negative shocks to volatility, as in Hafner and Herwartz (1998b) and Kroner and Ng (1998), who both use threshold variables.

Appendix

Proof of Theorem 1: Applying Theorem 10.2 of Magnus (1988), we have

$$\text{Var}(\text{vec}(\varepsilon_t \varepsilon_t') \mid \mathcal{F}_{t-1}) = 2N_K(H_t \otimes H_t).$$

By definition, $\eta_t = L_K \text{vec}(\varepsilon_t \varepsilon_t')$, so that

$$\text{Var}(\eta_t \mid \mathcal{F}_{t-1}) = 2L_K N_K(H_t \otimes H_t) L_K'.$$

Applying the vec-operator, we get

$$\text{vec}(\text{Var}[\eta_t \mid \mathcal{F}_{t-1}]) = 2(L_K \otimes L_K)(I_{K^2} \otimes N_K) \text{vec}(H_t \otimes H_t).$$

Now, $\text{vec}(H_t \otimes H_t)$ can be expressed as the product of the vecs using the commutation matrix C to obtain

$$\text{vec}(\text{Var}[\eta_t \mid \mathcal{F}_{t-1}]) = 2(L_K \otimes L_K)(I_{K^2} \otimes N_K)(I_K \otimes C_{KK} \otimes I_K)(\text{vec}(H_t) \otimes \text{vec}(H_t)).$$

By definition, $\text{vec}(H_t) = D_K h_t$, so that $\text{vec}(H_t) \otimes \text{vec}(H_t) = (D_K \otimes D_K) \text{vec}(h_t h_t')$. Recalling the notation

$$\Sigma_\eta = \text{E}[\text{Var}(\eta_t \mid \mathcal{F}_{t-1})] + \Sigma_h$$

with $\Sigma_h = E[h_t h_t']$, and taking vecs, the result is obtained. \square

Proof of Theorem 2: Since $\Gamma(0) = \Sigma_\eta - \sigma\sigma'$ and $\Sigma_u = \Sigma_\eta - \Sigma_h$, we can write

$$\Sigma_\eta = \sum_{i=0}^{\infty} \Phi_i (\Sigma_\eta - \Sigma_h) \Phi_i' + \sigma\sigma'.$$

Taking vecs and applying Theorem 1, we obtain

$$G_K \text{vec}(\Sigma_h) = \sum_{i=0}^{\infty} (\Phi_i \otimes \Phi_i) (G_K - I_{N^2}) \text{vec}(\Sigma_h) + \text{vec}(\sigma\sigma'),$$

and, since $\Phi_0 = I_N$,

$$\left(I_{N^2} - \sum_{i=1}^{\infty} (\Phi_i \otimes \Phi_i) (G_K - I_{N^2}) \right) \text{vec}(\Sigma_h) = \text{vec}(\sigma\sigma').$$

The matrix $I_{N^2} - \sum_{i=1}^{\infty} (\Phi_i \otimes \Phi_i) (G_K - I_{N^2})$ is invertible if and only if all its eigenvalues have modulus smaller than one. \square

Proof of Theorem 3: Taking expectations, we have

$$\begin{aligned} E[\text{vec}(h_t h_t')] &= \gamma + (A \otimes A) E[\text{vec}(\eta_{t-1} \eta_{t-1}')] \\ &+ \{(A \otimes B) + (B \otimes A) + (B \otimes B)\} E[\text{vec}(h_{t-1} h_{t-1}')], \end{aligned}$$

where $\gamma = \text{vec}(\omega\omega' + (A+B)\sigma\omega' + \omega\sigma'(A+B)')$. Since the unconditional expectations are independent of time and $\text{vec}(\Sigma_\eta) = G_K \text{vec}(\Sigma_h)$ by Theorem 1, one obtains

$$(I_{N^2} - Z) \text{vec}(\Sigma_h) = \gamma$$

with $Z = (A \otimes A)G_K + (A \otimes B) + (B \otimes A) + (B \otimes B)$. The matrix $I_{N^2} - Z$ is invertible if and only if all eigenvalues of Z have modulus smaller than one. \square

Proof of Theorem 4: Only the equation for k_{ij} will be proved. We have

$$\begin{aligned} E[\varepsilon_{it}^2 \varepsilon_{jt}^2] &= E[E[\varepsilon_{it}^2 \varepsilon_{jt}^2 \mid \mathcal{F}_{t-1}]] \\ &= E[h_{it} h_{jt} + 2h_{ij,t}^2] \\ &= \sigma_{ii} \sigma_{jj} + \text{Cov}(h_{it}, h_{jt}) + 2 \left(\text{Var}(h_{ij,t}) + \sigma_{ij}^2 \right). \end{aligned}$$

Dividing by $\sigma_{ii} \sigma_{jj}$, the result is obtained. \square

References

- Bollerslev, T. (1990) Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economics and Statistics* 72, 498–505.
- Bollerslev, T. & R.F. Engle (1993) Common persistence in conditional variances. *Econometrica* 61, 167–186.
- Bollerslev, T., R.F. Engle & J.M. Wooldridge (1988) A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96, 116–131.
- Bossaerts, P., W. Härdle & C.M. Hafner (1996) Foreign exchange rates have surprising volatility. In P. Robinson (ed.), *Athens Conference on Applied Probability and Time Series*, vol.2, number 115 of *Lecture Notes in Statistics*, pp. 55–72. New York: Springer.
- Broze, L., C. Francq & J.-M. Zakoian (2000) Non redundancy of high order moment conditions for efficient GMM estimation of weak AR processes. CORE discussion paper 2000/33, Université catholique de Louvain, Louvain-la-Neuve, Belgium.
- Diebold, F.X. & M. Nerlove (1989) The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. *Journal of Applied Econometrics* 4, 1–21.
- Drost, F. C. & T. E. Nijman (1993) Temporal aggregation of GARCH processes. *Econometrica* 61, 909–927.
- Drost, F. C. & B. J. M. Werker (1996) Closing the GARCH gap: continuous time GARCH modeling. *Journal of Econometrics* 74, 31–57.
- Engle, R.F. & K.F. Kroner (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122–150.
- Engle, R.F., V.M. Ng & Rothschild (1990) Asset pricing with a factor-ARCH structure: empirical estimates for Treasury Bills. *Journal of Econometrics* 45, 213–237.
- Gallant, A.R., P.E. Rossi & G. Tauchen (1993) Nonlinear dynamic structures. *Econometrica* 61, 871–907.
- Gouriéroux, C. (1997) *ARCH Models and Financial Applications*. New York: Springer.
- Hafner, C.M. (1998) *Nonlinear time series analysis with applications to foreign exchange rate volatility*. Heidelberg: Physica.
- Hafner, C.M. & H. Herwartz (1998a) Volatility impulse response functions for multivariate GARCH models. CORE discussion paper 9847, Université catholique de Louvain, Louvain-la-Neuve, Belgium.

- Hafner, C.M. & H. Herwartz (1998b) Structural analysis of portfolio risk using beta impulse response functions. *Statistica Neerlandica* 52, 336–355.
- He, C. & T. Teräsvirta (1999) Fourth moment structure of the GARCH(p, q) process. *Econometric Theory* 15, 824–846.
- Jeantheau, T. (1998) Strong consistency of estimators for multivariate ARCH models, *Econometric Theory* 14, 70–86.
- Karanasos, M. (1999) The second moment and the autocovariance function of the squared errors of the GARCH model. *Journal of Econometrics* 90, 63–76.
- Kroner, K.F. & V.K. Ng (1998) Modeling asymmetric comovements of asset returns. *The Review of Financial Studies* 11, 817–844.
- Lee, S.W. & B.E. Hansen (1994) Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* 10, 29–52.
- Lumsdaine, R.L. (1996) Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. *Econometrica* 64, 575–596.
- Lütkepohl, H. (1993) *Introduction to Multiple Time Series Analysis*, 2nd ed., Berlin, New York: Springer.
- Magnus, J.R. (1988) *Linear Structures*. London: Griffin.
- Mahieu, R. (1995) *Financial Market Volatility – Statistical Models and Empirical Results*, PhD thesis, Universitaire Pers Maastricht, The Netherlands.
- Mardia, K.V. (1970) Measures of multivariate skewness and kurtosis with applications. *Biometrika* 57, 519–530.
- Meddahi, N. & E. Renault (1997) Linear statistical inference for ARCH-type processes, GREMAQ, Université de Toulouse, France.

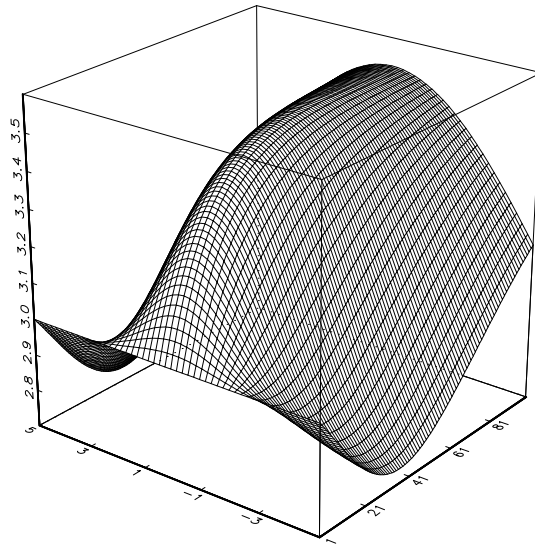


Figure 1: *Impulse response function of the kurtosis of DEM/USD to a shock in DEM/USD. The right axis displays the time horizon $s = 1 \dots, 100$, the left axis the shock $\xi_{1,t}$.*

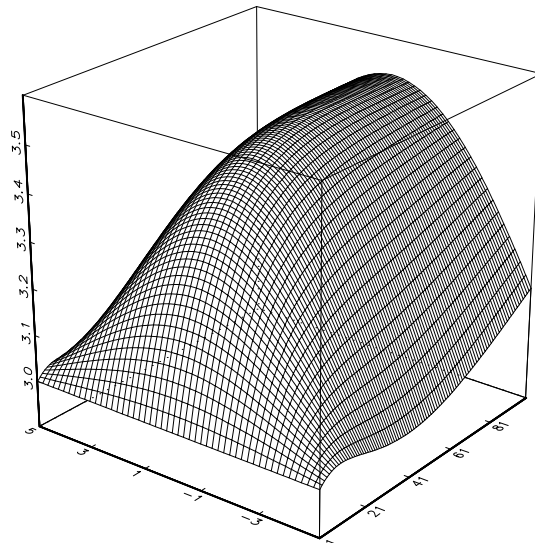


Figure 2: *Impulse response function of the kurtosis of DEM/USD to a shock in GBP/USD. The right axis displays the time horizon $s = 1 \dots, 100$, the left axis the shock $\xi_{2,t}$.*

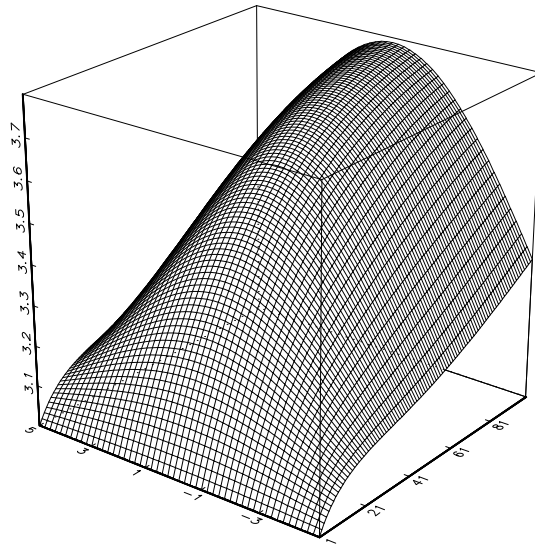


Figure 3: *Impulse response function of the kurtosis of GBP/USD to a shock in DEM/USD. The right axis displays the time horizon $s = 1 \dots, 100$, the left axis the shock $\xi_{1,t}$.*

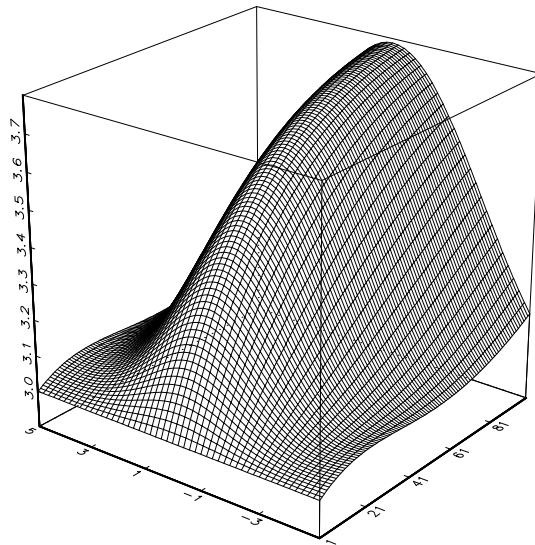


Figure 4: *Impulse response function of the kurtosis of GBP/USD to a shock in GBP/USD. The right axis displays the time horizon $s = 1 \dots, 100$, the left axis the shock $\xi_{2,t}$.*

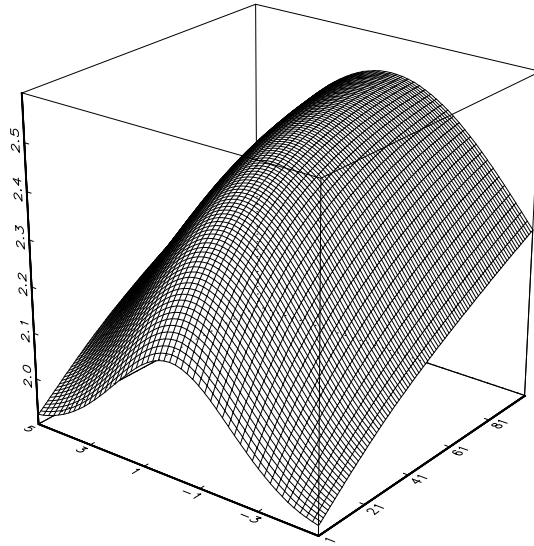


Figure 5: *Impulse response function of the co-kurtosis to a shock in DEM/USD. The right axis displays the time horizon $s = 1 \dots, 100$, the left axis the shock $\xi_{1,t}$.*

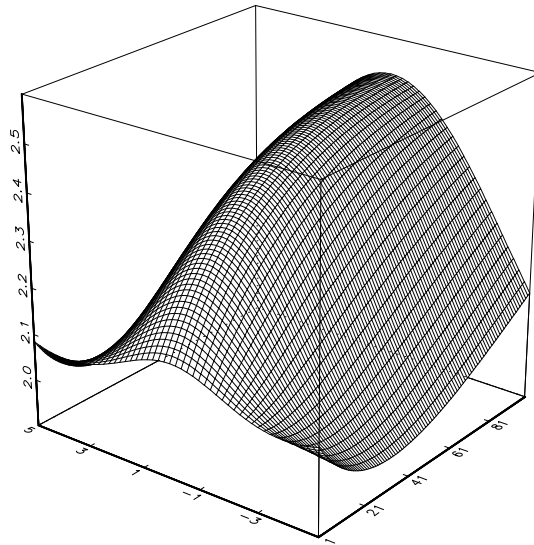


Figure 6: *Impulse response function of the co-kurtosis to a shock in GBP/USD. The right axis displays the time horizon $s = 1 \dots, 100$, the left axis the shock $\xi_{2,t}$.*

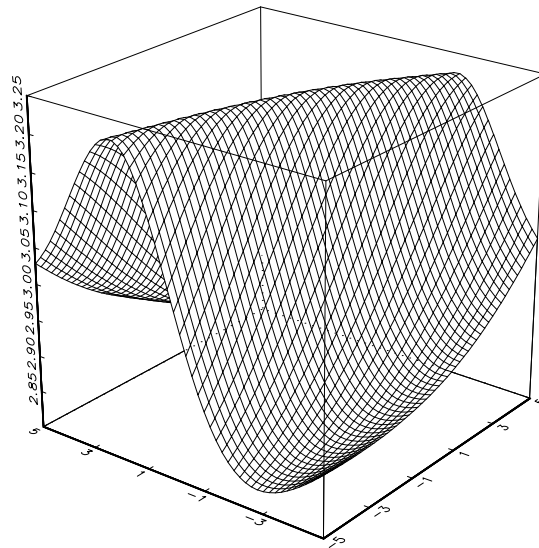


Figure 7: *Impulse response function at $s = 10$ of the kurtosis of DEM/USD to a simultaneous shock. The right axis displays a shock in GBP/USD, the left axis a shock in DEM/USD.*

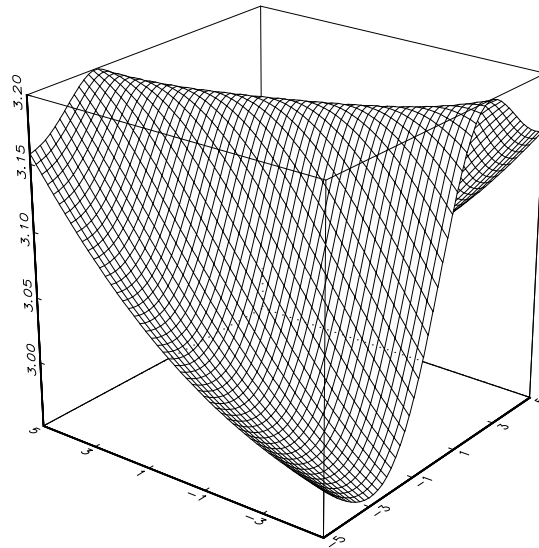


Figure 8: *Impulse response function at $s = 10$ of the kurtosis of GBP/USD to a simultaneous shock. The right axis displays a shock in GBP/USD, the left axis a shock in DEM/USD.*

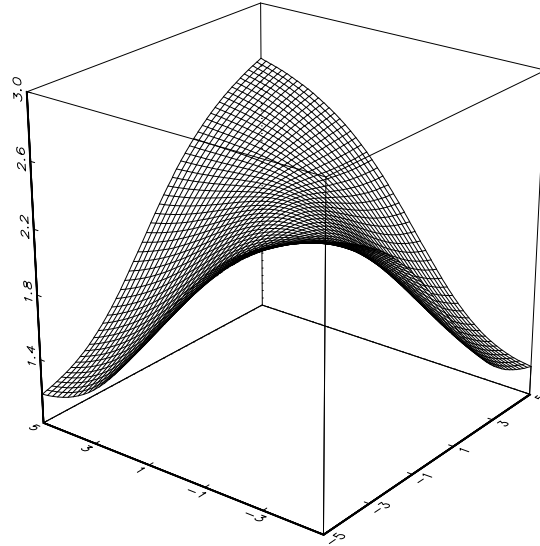


Figure 9: *Impulse response function at $s = 10$ of the co-kurtosis of DEM/USD to a simultaneous shock. The right axis displays a shock in GBP/USD, the left axis a shock in DEM/USD.*