Chapter 2
Multivariate Volatility Models

M.R. Fengler, H. Herwartz and F.H.C. Raters

Abstract Multivariate volatility models are widely used in finance to capture both volatility clustering and contemporaneous correlation of asset return vectors. Here, we focus on multivariate GARCH models. In this common model class, it is assumed that the covariance of the error distribution follows a time dependent process conditional on information which is generated by the history of the process. To provide a particular example, we consider a system of exchange rates of two currencies measured against the US Dollar (USD), namely the Deutsche Mark (DEM) and the British Pound Sterling (GBP). For this process, we compare the dynamic properties of the bivariate model with univariate GARCH specifications where cross sectional dependencies are ignored. Moreover, we illustrate the scope of the bivariate model by ex-ante forecasts of bivariate exchange rate densities.

2.1 Introduction

Volatility clustering, i.e. positive correlation of price variations observed on speculative markets, motivated the introduction of autoregressive conditionally heteroskedastic (ARCH) processes by Engle (1982) and its popular generalizations

M.R. Fengler
Faculty of Mathematics and Statistics, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland
e-mail: matthias.fengler@unisg.ch

H. Herwartz · F.H.C. Raters (✉)
Department of Economics, University of Göttingen, Humboldtallee 3, 37073 Göttingen, Germany
e-mail: hherwartz@uni-goettingen.de

F.H.C. Raters
e-mail: raters@uni-goettingen.de

by Bollerslev (1986) (Generalized ARCH, GARCH) and Nelson (1991) (exponential GARCH, EGARCH). Being univariate in nature, however, such models neglect a further stylized fact of empirical price variations, namely contemporaneous cross correlation e.g. over a set of assets, stock market indices, or exchange rates.

Cross section relationships are often implied by economic theory. Interest rate parities, for instance, provide a close relation between domestic and foreign bond rates. Assuming absence of arbitrage, the so-called triangular equation formalizes the equality of an exchange rate between two currencies on the one hand and an implied rate constructed via exchange rates measured towards a third currency. Furthermore, stock prices of firms acting on the same market often show similar patterns in the sequel of news that are important for the entire market (Hafner and Herwartz 1998). Similarly, analyzing global volatility transmission Engle et al. (1990) and Hamao et al.(1990) found evidence in favor of volatility spillovers between the world's major trading areas occurring in the sequel of floor trading hours. From this point of view, when modeling time varying volatilities, a multivariate model appears to be a natural framework to take cross sectional information into account. Moreover, the covariance between financial assets is of essential importance in finance. Effectively, many problems in financial practice like portfolio optimization, hedging strategies, or Value-at-Risk evaluation require multivariate volatility measures (Bollerslev et al. 1988; Cecchetti et al. 1988).

2.1.1 Model Specifications

Let $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Nt})^\top$ denote an N-dimensional error process, which is either directly observed or estimated from a multivariate regression model. The process ε_t follows a multivariate GARCH process if it has the representation

$$\varepsilon_t = \Sigma_t^{1/2} \xi_t, \tag{2.1}$$

where Σ_t is measurable with respect to information generated up to time $t - 1$, denoted by the filtration $\mathcal{F}_{t-1}$. By assumption, the N components of ξ_t follow a multivariate Gaussian distribution with mean zero and a covariance matrix equal to the identity matrix.

The conditional covariance matrix, $\Sigma_t = \mathsf{E}[\varepsilon_t \varepsilon_t^\top | \mathcal{F}_{t-1}]$, has typical elements σ_{ij} with σ_{ii}, $i = 1, \ldots, N$, denoting conditional variances and off-diagonal elements σ_{ij}, $i, j = 1, \ldots, N$, $i \neq j$, denoting conditional covariances. To make the specification in (2.1) feasible, a parametric description relating Σ_t to $\mathcal{F}_{t-1}$ is necessary. In a multivariate setting, however, dependencies of the second order moments in Σ_t on $\mathcal{F}_{t-1}$ become easily computationally intractable for practical purposes.

Let vech(A) denote the half-vectorization operator stacking the elements of a quadratic ($N \times N$)-matrix A from the main diagonal downwards in a $\frac{1}{2}N(N + 1)$ dimensional column vector. Within the so-called half-vec representation of the GARCH(p, q) model Σ_t is specified as follows:

$$\text{vech}(\Sigma_t) = c + \sum_{i=1}^{q} \widetilde{A}_i \text{vech}(\varepsilon_{t-i}\varepsilon_{t-i}^{\top}) + \sum_{i=1}^{p} \widetilde{G}_i \text{vech}(\Sigma_{t-i}). \qquad (2.2)$$

In (2.2), the matrices $\widetilde{A}_i$ and $\widetilde{G}_i$ each contain $\{N(N+1)/2\}^2$ elements. Deterministic covariance components are collected in c, a column vector of dimension $N(N+1)/2$. We consider in the following the case $p = q = 1$ since in applied work the GARCH(1,1) model has turned out to be particularly useful to describe a wide variety of financial market data (Bollerslev et al., 1994).

On the one hand, the half-vec model in (2.2) allows for a very general dynamic structure of the multivariate volatility process. On the other hand, this specification suffers from high dimensionality of the relevant parameter space, which makes it almost intractable for empirical work. In addition, it might be cumbersome in applied work to restrict the admissible parameter space such that the implied matrices Σ_t, $t = 1, \ldots, T$, are positive definite. These issues motivated a considerable variety of competing multivariate GARCH specifications.

Prominent proposals reducing the dimensionality of (2.2) are the constant correlation model (Bollerslev et al. 1988) and the diagonal model (Bollerslev et al. 1988). Specifying diagonal elements of Σ_t both of these approaches assume the absence of cross equation dynamics, i.e. the only dynamics are

$$\sigma_{ii,t} = c_{ii} + a_i \varepsilon_{i,t-1}^2 + g_i \sigma_{ii,t-1}, \ i = 1, \ldots, N. \qquad (2.3)$$

To determine off-diagonal elements of Σ_t, Bollerslev (1990) proposes a constant contemporaneous correlation,

$$\sigma_{ij,t} = \rho_{ij}\sqrt{\sigma_{ii}\sigma_{jj}}, \ i, j = 1, \ldots, N, \qquad (2.4)$$

whereas Bollerslev et al. (1988) introduce an ARMA-type dynamic structure as in (2.3) for $\sigma_{ij,t}$ as well, i.e.

$$\sigma_{ij,t} = c_{ij} + a_{ij}\varepsilon_{i,t-1}\varepsilon_{j,t-1} + g_{ij}\sigma_{ij,t-1}, \ i, j = 1, \ldots, N. \qquad (2.5)$$

For the bivariate case ($N = 2$) with $p = q = 1$, the constant correlation model contains only 7 parameters compared to 21 parameters encountered in the full model (2.2). The diagonal model is specified with 9 parameters. The price that both models pay for parsimony is in ruling out cross equation dynamics as allowed in the general half-vec model. Positive definiteness of Σ_t is easily guaranteed for the constant correlation model ($|\rho_{ij}| < 1$), whereas the diagonal model requires more complicated restrictions to provide positive definite covariance matrices.

The so-called BEKK model (Baba et al. 1990) provides a richer dynamic structure compared to both restricted processes mentioned before. Defining $N \times N$ matrices A_{ik} and G_{ik} and an upper triangular matrix C_0, the BEKK model reads in a general version as follows (see Engle and Kroner 1995):

$$\Sigma_t = C_0^\top C_0 + \sum_{k=1}^{K} \sum_{i=1}^{q} A_{ik}^\top \varepsilon_{t-i} \varepsilon_{t-i}^\top A_{ik} + \sum_{k=1}^{K} \sum_{i=1}^{p} G_{ik}^\top \Sigma_{t-i} G_{ik}. \qquad (2.6)$$

If $K = q = p = 1$ and $N = 2$, the model in (2.6) contains 11 parameters and implies the following dynamic model for typical elements of Σ_t:

$$\begin{aligned}
\sigma_{11,t} &= c_{11} + a_{11}^2 \varepsilon_{1,t-1}^2 + 2a_{11}a_{21}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + a_{21}^2 \varepsilon_{2,t-1}^2 \\
&\quad + g_{11}^2 \sigma_{11,t-1} + 2g_{11}g_{21}\sigma_{21,t-1} + g_{21}^2 \sigma_{22,t-1}, \\
\sigma_{21,t} &= c_{21} + a_{11}a_{22}\varepsilon_{1,t-1}^2 + (a_{21}a_{12} + a_{11}a_{22})\varepsilon_{1,t-1}\varepsilon_{2,t-1} + a_{21}a_{22}\varepsilon_{2,t-1}^2 \\
&\quad + g_{11}g_{22}\sigma_{11,t-1} + (g_{21}g_{12} + g_{11}g_{22})\sigma_{12,t-1} + g_{21}g_{22}\sigma_{22,t-1}, \\
\sigma_{22,t} &= c_{22} + a_{12}^2 \varepsilon_{1,t-1}^2 + 2a_{12}a_{22}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + a_{22}^2 \varepsilon_{2,t-1}^2 \\
&\quad + g_{12}^2 \sigma_{11,t-1} + 2g_{12}g_{22}\sigma_{21,t-1} + g_{22}^2 \sigma_{22,t-1}.
\end{aligned}$$

Compared to the diagonal model, the BEKK–specification economizes on the number of parameters by restricting the half-vec model within and across equations. Since A_{ik} and G_{ik} are not required to be diagonal, the BEKK model is convenient to allow for cross dynamics of conditional covariances. The parameter K governs to which extent the general representation in (2.2) can be approximated by a BEKK-type model. In the following we assume $K = 1$. Note that in the bivariate case with $K = p = q = 1$ the BEKK model contains 11 parameters. If $K = 1$, the matrices A_{11} and $-A_{11}$ imply the same conditional covariances. Thus, for uniqueness of the BEKK-representation $a_{11} > 0$ and $g_{11} > 0$ is assumed. Note that the right hand side of (2.6) involves only quadratic terms and, hence, given convenient initial conditions, Σ_t is positive definite under the weak (sufficient) condition that at least one of the matrices C_0 or G_{ik} has full rank (Engle and Kroner 1995). It is worthwhile to mention that in a similar way the univariate GARCH volatility model can be augmented by threshold specifications (Glosten et al. 1993), a generalization for asymmetric effects in a BEKK-type model is discussed in Kroner and Ng (1998).

2.1.2 Estimation of the BEKK Model

As in the univariate case, the parameters of a multivariate GARCH model are estimated by maximum likelihood (ML) optimizing numerically the Gaussian log-likelihood function.

With f denoting the multivariate normal density, the contribution of a single observation, l_t, to the log-likelihood of a sample is given as:

$$\begin{aligned}
l_t &= \ln\{f(\varepsilon_t|\mathcal{F}_{t-1})\} \\
&= -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(|\Sigma_t|) - \frac{1}{2}\varepsilon_t^\top \Sigma_t^{-1} \varepsilon_t.
\end{aligned}$$

Maximizing the log-likelihood, $l = \sum_{t=1}^{T} l_t$, requires nonlinear maximization methods. Involving only first order derivatives, the BHHH algorithm introduced by Berndt et al. (1974) is easily implemented and particularly useful for the estimation of multivariate GARCH processes.

If the actual error distribution differs from the multivariate normal, maximizing the Gaussian log-likelihood has become popular as Quasi ML (QML) estimation. In the multivariate framework, results for the asymptotic properties of the (Q)ML-estimator have been derived by Jeantheau (1998) who proves the QML-estimator to be consistent under the main assumption that the considered multivariate process is strictly stationary and ergodic. Further assuming finiteness of moments of ε_t up to order eight, Comte and Lieberman (2003) derive asymptotic normality of the QML-estimator. The asymptotic distribution of the rescaled QML-estimator is analogous to the univariate case and discussed in Bollerslev and Wooldridge (1992).

2.2 An Empirical Illustration

2.2.1 Data Description

We analyze daily quotes of two European currencies measured against the USD, namely the DEM and the GBP. The sample period is December 31, 1979 to April 1, 1994, covering $T = 3720$ observations. Note that a subperiod of our sample has already been investigated by Bollerslev and Engle (1993) discussing common features of volatility processes.

Let the bivariate vector R_t denote the exchange rates (DEM/USD and GBP/USD) at time t. Before inspecting the sample statistics (⊘XFGmvol101.R), we take the first differences of the log exchange rates, $\varepsilon_t = \ln(R_t) - \ln(R_{t-1})$. These log-differences are shown in Fig. 2.1. Evidently, the empirical means of both processes are very close to zero ($-4.72\text{e-}06$ and $1.10\text{e-}04$, respectively). Also minimum, maximum and standard errors are of similar size. As is apparent from Fig. 2.1, variations of exchange rate log-differences exhibit an autoregressive pattern: Large log-differences of foreign exchange rates are followed by large log-differences of either sign. This is most obvious in periods of excessive log-differences. Note that these volatility clusters tend to coincide in both series. It is precisely this observation that justifies a multivariate GARCH specification.

2.2.2 Estimating Bivariate GARCH

A fast algorithm is used to estimate the BEKK representation of a bivariate GARCH (1,1) model: QML-estimation is implemented by means of the BHHH-algorithm which minimizes the negative Gaussian log-likelihood function. The algorithm
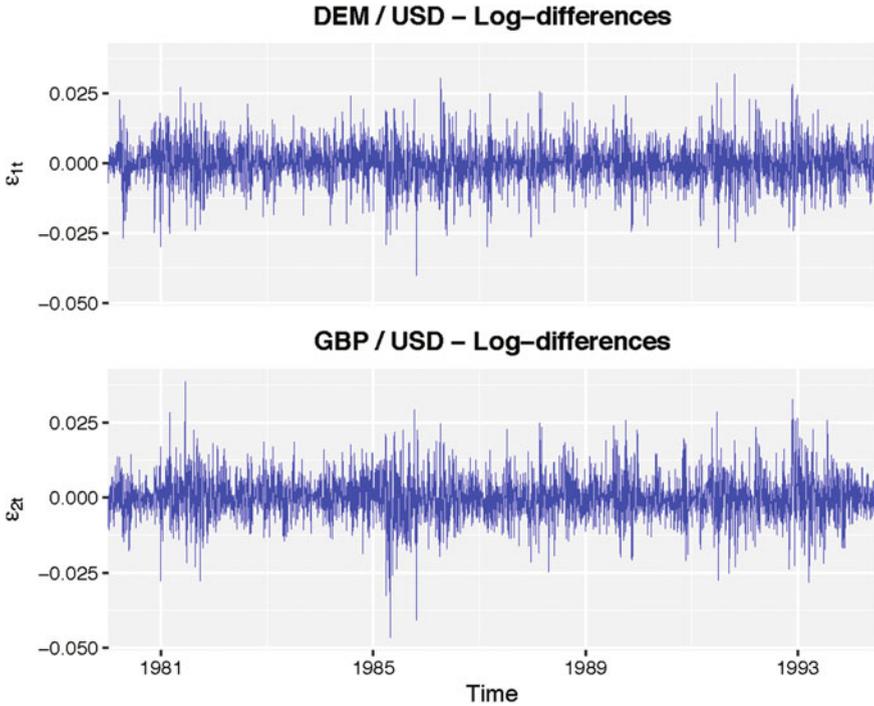
Fig. 2.1 Foreign exchange rate data: log-differences. 🔍 `XFGmvol01`

employs analytical first order derivatives of the log-likelihood function Lütkepohl (1996) with respect to the 11-dimensional vector of parameters containing the elements of C_0, A_{11} and G_{11} as given in (2.6). Alternatively, the R package `mgarchBEKK` Schmidbauer et al. (2016) might be considered when estimating this model in *R*. Section 2.3 contains further references for implementations of the BEKK model in widely used numerical programming environments.

The estimation output contains the stacked elements of the parameter matrices C_0, A_{11} and G_{11} in (2.6) after numerical optimization of the Gaussian log-likelihood function. Being an iterative procedure, the algorithm requires to determine suitable initial parameters. For the diagonal elements of the matrices A_{11} and G_{11} values around 0.3 and 0.9 appear reasonable, since in univariate GARCH(1,1) models parameter estimates for a_1 and g_1 in (2.3) often take values around $0.3^2 = 0.09$ and $0.81 = 0.9^2$. There is no clear guidance how to determine initial values for off diagonal elements of A_{11} or G_{11}. Therefore, it might be reasonable to try alternative initializations of these parameters. Given an initialization of A_{11} and G_{11}, the starting values for the elements in C_0 are determined by the algorithm assuming the unconditional covariance of ε_t to exist (Engle and Kroner 1995).

Given our example under investigation, the bivariate GARCH estimation yields
a vector of coefficient estimates,

$$\hat{\theta} = (.00115, .00031, .00076, .2819, -.0572, -.0504, .2934, .9389, .0251, .0275, .9391),$$

and a corresponding log-likelihood value $\hat{l} = 28599$ at the optimum. The first three
estimates are the parameters of the upper triangular matrix C_0, the following four
belong to the ARCH (A_{11}) and the last four to the GARCH parameters (G_{11}), i.e. for
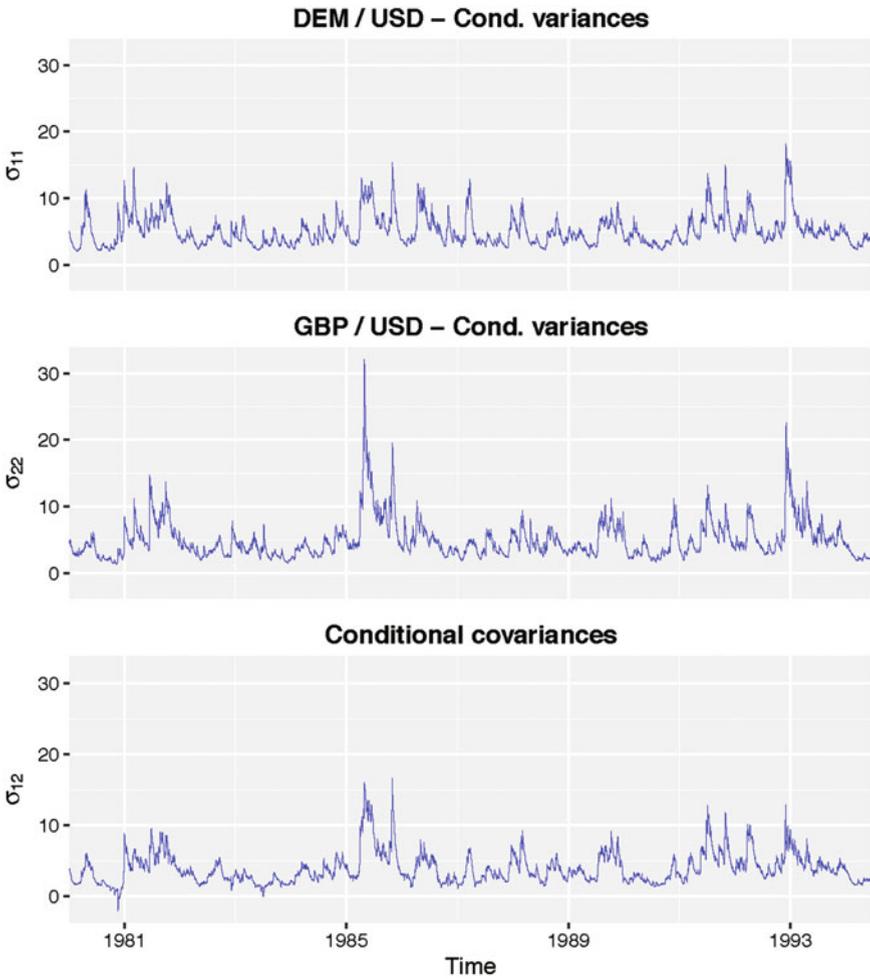our model,

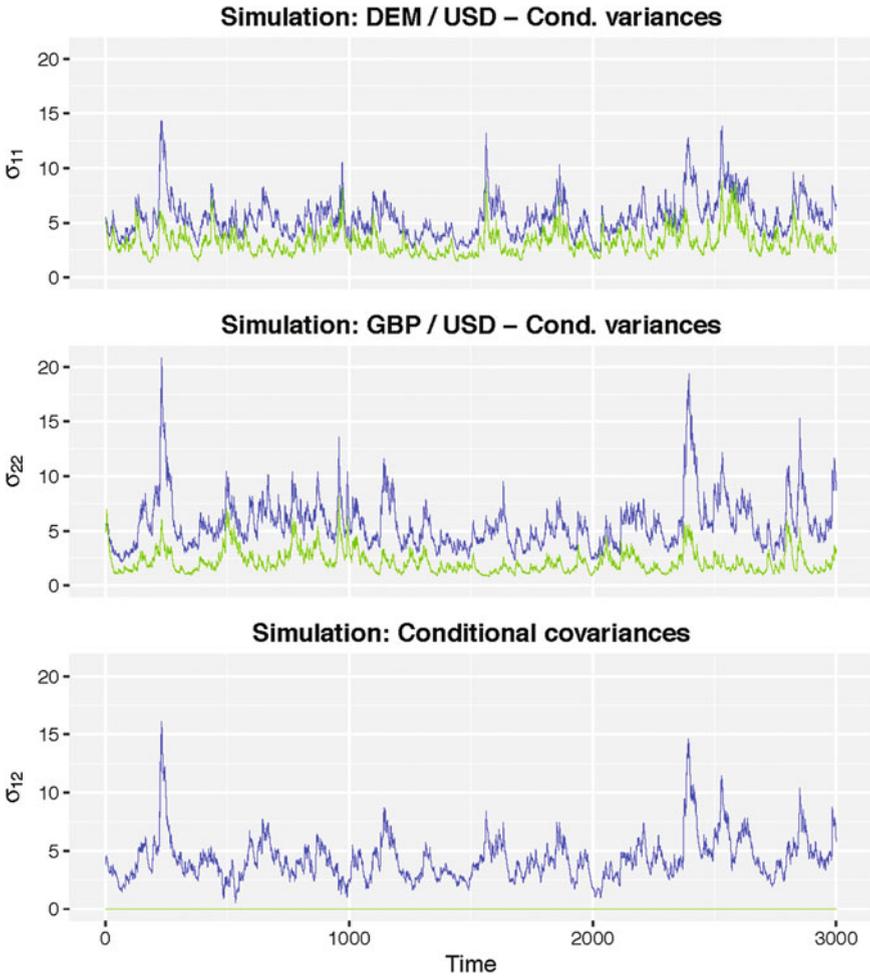**Fig. 2.2** Estimated variance and covariance processes, $10^5 \hat{\Sigma}_t$. XFGmvol02

Fig. 2.3 Simulated variance and covariance processes, both bivariate (*blue*) and univariate case (*green*), $10^5 \hat{\Sigma}_t$. 🔲 XFGmvol03

$$\Sigma_t = C_0^\top C_0 + A_{11}^\top \varepsilon_{t-1} \varepsilon_{t-1}^\top A_{11} + G_{11}^\top \Sigma_{t-1} G_{11}, \tag{2.7}$$

stated again for convenience, we find the matrices C_0, A_{11}, G_{11} to be:

$$C_0 = 10^{-3} \begin{pmatrix} 1.15 & .31 \\ 0 & .76 \end{pmatrix}, \ A_{11} = \begin{pmatrix} .282 & -.050 \\ -.057 & .293 \end{pmatrix}, \ G_{11} = \begin{pmatrix} .939 & .028 \\ .025 & .939 \end{pmatrix}. \tag{2.8}$$

2.2.3 Estimating the (co)variance Processes

The (co)variance is obtained by sequentially calculating the difference equation (2.7) where we use the estimator for the unconditional covariance matrix as initial value ($\Sigma_0 = \frac{E^\top E}{T}$). Here, the $T \times 2$ matrix E contains log-differences of our foreign exchange rate data.

We display the estimated variance and covariance processes in Fig. 2.2. The quantlet ⓠ`XFGmvol02.R` ss contains the code. The two upper panels of Fig. 2.2 show the variances of the DEM/USD and GBP/USD log-differences respectively, whereas in the lower panel we see the covariance process. Except for a very short period in the beginning of our sample, the covariance is positive and of non-negligible size throughout. This is evidence for cross sectional dependencies in currency markets which we mentioned earlier to motivate multivariate GARCH models.

Instead of estimating the realized path of variances as shown above, we could also use the estimated parameters to *simulate* volatility paths (ⓠ`XFGmvol03.R`). For this, at each point in time an observation ε_t is drawn from a multivariate normal distribution with variance Σ_t. Given these observations, Σ_t is updated according to (2.7). Then, a new residual is drawn with covariance Σ_{t+1}. We apply this procedure for $T = 3000$. The results, displayed in the three panels of Fig. 2.3, show a similar pattern as the original process given in Fig. 2.2. For the upper two panels, we generate two variance processes from the *same* set of simulated residuals ξ_t. In this case, however, we set off-diagonal parameters in $C_0^\top C_0$, A_{11} and G_{11} to zero to illustrate how the unrestricted BEKK model incorporates cross equation dynamics. As can be seen, both approaches are convenient to capture volatility clustering. Depending on the particular state of the system, spillover effects operating through conditional covariances, however, have a considerable impact on the magnitude of conditional volatility.

2.3 Forecasting Exchange Rate Densities

The preceding section illustrated how the GARCH model may be employed effectively to describe empirical price variations of foreign exchange rates. For practical purposes, as for instance scenario analysis, Value-at-Risk estimation (Chap. 1), option pricing (see the corresponding chapter), one is often interested in the future joint density of a set of asset prices. Continuing the comparison of the univariate and bivariate approach to model volatility dynamics of exchange rates, it is thus natural to investigate the properties of these specifications in terms of forecasting performance.

We implement an iterative forecasting scheme along the following lines: Given the estimated univariate and bivariate volatility models and the corresponding information sets $\mathcal{F}_{t-1}, t = 1, \ldots, T - 5$ (Fig. 2.2), we employ the identified data generating processes to simulate one-week-ahead forecasts of both exchange rates. To get a reliable estimate of the future density, we set the number of simulations to 5000 for each

initial scenario. This procedure yields two bivariate samples of future exchange rates, one simulated under bivariate, the other one simulated under univariate GARCH assumptions.

A review of evaluating competing density forecasts is offered by Tay and Wallis (2000). Adopting a Bayesian perspective the common approach is to compare the expected loss of actions evaluated under alternative density forecasts. In our pure time series framework, however, a particular action is hardly available for forecast density comparisons. Alternatively, one could concentrate on statistics directly derived from the simulated densities, such as first and second order moments or even quantiles. Due to the multivariate nature of the time series under consideration, it is a nontrivial issue to rank alternative density forecasts in terms of these statistics. Therefore, we regard a particular volatility model to be superior to another if it provides a higher simulated density estimate of the actual bivariate future exchange rate. This is accomplished by evaluating both densities at the actually realized exchange rate obtained from a bivariate kernel estimation. Since the latter comparison might suffer from different unconditional variances under univariate and multivariate volatility, the two simulated densities were rescaled to have identical variance. Performing the latter forecasting exercises iteratively over 3714 time points, we can test if the bivariate volatility model outperforms the univariate one.

To formalize the latter ideas, we define a success ratio SR_J as

$$SR_J = \frac{1}{|J|} \sum_{t \in J} \mathbf{1}\{\hat{f}_{biv}(R_{t+5}) > \hat{f}_{uni}(R_{t+5})\}, \tag{2.9}$$

where J denotes a time window containing $|J|$ observations and $\mathbf{1}$ an indicator function. $\hat{f}_{biv}(R_{t+5})$ and $\hat{f}_{uni}(R_{t+5})$ are the estimated densities of future exchange rates which are simulated by the bivariate and univariate GARCH processes, respectively, and which are evaluated at the actual exchange rate levels R_{t+5}. The simulations are performed in ◘XFGmvol104.

Our results show that the bivariate model indeed outperforms the univariate one when both likelihoods are compared under the actual realizations of the exchange rate process. In 82.3% of all cases across the sample period, $SR_J = 0.823$, $J = \{t : t = 1, ..., T - 5\}$, the bivariate model provides a better forecast. This is highly significant. In Table 2.1, we show that the overall superiority of the bivariate volatility

Table 2.1 Time varying frequencies of the bivariate GARCH model outperforming the univariate one in terms of one-week-ahead forecasts (success ratio)

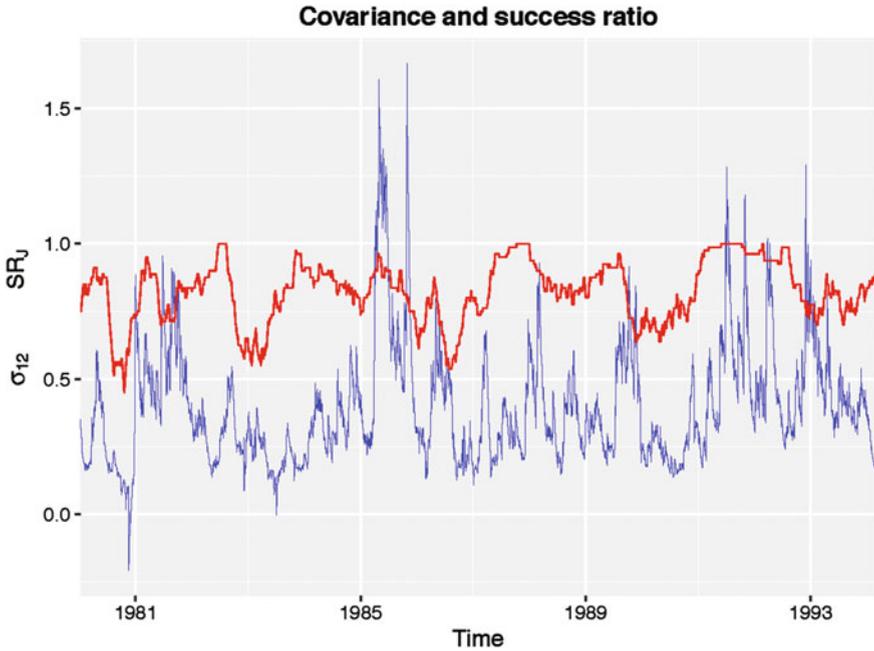| Time window J		Success ratio SR_J
1980	1981	0.762
1982	1983	0.786
1984	1985	0.868
1986	1987	0.780
1988	1989	0.872
1990	1991	0.835
1992	04/1994	0.854

Fig. 2.4 Estimated covariance process from the bivariate GARCH model ($10^4\hat{\sigma}_{12}$, *blue*) and success ratio over overlapping time intervals with window length 80 days (*red*). ▣ XFGmvol04

approach is confirmed when considering subsamples of two-years length. A-priori, one may expect the bivariate model to outperform the univariate one the larger (in absolute value) the covariance between both log-difference processes is. To verify this argument, we display in Fig. 2.4 the empirical covariance estimates from Fig. 2.2 jointly with the success ratio evaluated over overlapping time intervals of length $|J| = 80$.

As is apparent from Fig. 2.4, there is a close co-movement between the success ratio and the general trend of the covariance process, which confirms our expectations: the forecasting power of the bivariate GARCH model is particularly strong in periods where the DEM/USD and GBP/USD exchange rate log-differences exhibit a high covariance. For completeness, it is worthwhile to mention that similar results are obtained if the window width is varied over reasonable choices of $|J|$ ranging from 40 to 150.

With respect to financial practice and research we take our results as strong support for a multivariate approach towards asset price modeling. Whenever contemporaneous correlation across markets matters, the system approach offers essential advantages. To name a few areas of interest, multivariate volatility models are supposed to yield useful insights for risk management, scenario analysis and option pricing.

Appendix: Software Packages

This section gives a brief overview of BEKK model implementations for the numerical programming languages and environments R, MATLAB and Stata. Built-in functions and external packages for estimating univariate and further multivariate volatility models are briefly reviewed in Chap. 1 Appendix.

There exist two publicly available R packages which attempt to implement the BEKK approach. Both implementations are in early stages and, therefore, computed results need to be critically reviewed by the user. The package mgarchBEKK Schmidbauer et al. (2016) might be used for simulating, estimating and predicting BEKK models. The estimation of simulated data returns plausible results. In contrast, the package MTS by Tsay (2015) contains a single function BEKK11 for estimating two- or three-dimensional BEKK(1,1) models only.

MATLAB offers methods to assess univariate GARCH-type models by means of its Econometrics Toolbox. However, there is no official MATLAB Toolbox that implements the BEKK model. As described in Chap. 1 Appendix, the MFE Toolbox tries to fill the gap of assessing of multivariate volatility models in MATLAB. It is the direct successor to the UCSD Toolbox by Kevin Sheppard which is not being further developed. The codebase might help getting insights into the technical details of the BEKK approach. Because the toolbox is still under development, an optimized, error-free use can not be guaranteed.

Currently, Stata supports only the analysis of univariate volatility models, diagonal half-vec models, which are restricted versions of the half-vec model in (2.2), and conditional correlation models. It seems that there exists no publicly available extension to estimate a BEKK model. As an alternative, users might employ the tools of the independent software package JMulTi,[1] which is closely related to Lütkepohl and Krätzig (2004), for BEKK model estimation and investigation in combination with Stata.

References

Baba, Y., Engle, R. F., Kraft, D. F., & Kroner, K. F. (1990). *Multivariate simultaneous generalized ARCH, mimeo*. Department of Economics: University of California, San Diego.

Berndt, E. K., Hall, B. H., Hall, R. E., & Hausman, J. A. (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement*, *3*, 653–665.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics*, *72*, 498–505.

Bollerslev, T., & Engle, R. F. (1993). Common persistence in conditional variances. *Econometrica*, *61*, 167–186.

Bollerslev, T., Engle, R.F. and Nelson, D.B. (1994). ARCH Models. In R.F. Engle, D.L. McFadden (Eds.), Handbook of Econometrics, Vol. 4, (pp. 2959–3038). Elsevier.

---

[1] Project website: http://www.jmulti.de.

Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, *96*, 116–131.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, *11*, 143–172.

Cecchetti, S. G., Cumby, R. E., & Figlewski, S. (1988). Estimation of the optimal futures hedge. *The Review of Economics and Statistics*, *70*, 623–630.

Comte, F., & Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis*, *84*, 61–84.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*, 987–1007.

Engle, R. F., Ito, T., & Lin, W. L. (1990). Meteor showers or heat waves? *Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market, Econometrica*, *58*, 525–542.

Engle, R. F., & Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, *11*, 122–150.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, *48*, 1779–1801.

Hafner, C. M., & Herwartz, H. (1998). Structural analysis of portfolio risk using beta impulse response functions. *Statistica Neerlandica*, *52*, 336–355.

Hamao, Y., Masulis, R. W., & Ng, V. K. (1990). Correlations in price changes and volatility across international stock markets. *Review of Financial Studies*, *3*, 281–307.

Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric Theory*, *14*, 70–86.

Kroner, K. F., & Ng, V. K. (1998). Modeling asymmetric comovements of asset returns. *Review of Financial Studies*, *11*, 817–844.

Lütkepohl, H. (1996). *Handbook of matrices*. Chichester: Wiley.

Lütkepohl, H., & Krätzig, M. (2004). *Applied time series econometrics*. Cambridge: Cambridge University Press.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, *59*, 347–370.

Schmidbauer, H., Roesch, A., & Tunalioglu, V. S. (2016). mgarchBEKK: Simulating, Estimating and Diagnosing MGARCH (BEKK and mGJR) Processes. *R package version*, 2.

Tay, A. S., & Wallis, K. F. (2000). Density forecasting: a survey. *Journal of Forecasting*, *19*, 235–254.

Tsay, R. S. (2015). MTS: All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models. *R package version*, 33.