
Multivariate time series models for asset prices

Christian M. Hafner¹ and Hans Manner²

¹ Institut de statistique and CORE, Université catholique de Louvain, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium. Tel +32 10 47 43 06, fax +32 10 47 30 32 christian.hafner@uclouvain.be

² Department of Quantitative Economics, Maastricht University, PoBOX 616, MD 6200, Maastricht, The Netherlands, h.manner@maastrichtuniversity.nl

1 Introduction

In this chapter we review recent developments in time series analysis of financial assets. We will focus on the multivariate aspect since in most applications the dynamics of a broad variety of assets is relevant. In many situations in finance, the high dimensional characteristics of the data can lead to numerical problems in estimation algorithms. As a motivating example, we show that an application of a standard multivariate GARCH type model in high dimensions to determine the minimum variance portfolio yields sub-optimal results due to biased parameter estimates. One possibility to avoid numerical problems is to impose more structure on the conditional covariance matrix of asset returns, for example a factor structure.

We first discuss recent advances in factor models, where factors can be observed as in the one-factor capital asset pricing model (CAPM) and the three-factor model of [57], or unobserved. The main idea of factor models is to capture common movements in asset prices while reducing the dimension substantially, allowing for flexible statistical modelling.

If factors exhibit specific dynamic features such as volatility clustering or fat tails, then these are typically inherited by the asset prices or returns. For example, fat tailed factor distributions may generate tail dependence and reduce the benefits of portfolio diversification. As for volatility clustering, the modelling of the volatility and the dependence between assets becomes essential for asset pricing models. We therefore review volatility models, again focusing on multivariate models. Since its introduction by [48] and [21], the generalized autoregressive conditional heteroscedastic (GARCH) model has dominated the empirical finance literature and several reviews appeared, e.g. [20] and [19]. We compare (multivariate) GARCH models to the alternative class of (multivariate) stochastic volatility (SV) models, where the volatility processes are driven by idiosyncratic noise terms. We consider properties and estimation of the alternative models.

With an increasing amount of intra-day data available, an alternative approach of volatility modelling using so-called realized volatility (RV) measures has become available. This approach goes back to an idea of [2]. Rather than modelling volatility as an unobserved variable, RV tries to make volatility observable by taking sums of squared intra-day returns, which converges to the daily integrated volatility if the time interval between observations goes to zero. A similar approach is available to obtain realized covariances, taking sums of intra-day cross-products of returns. While this approach delivers more precise measures and predictions of daily volatility and correlations, it also uses another information set and is hence difficult to compare with standard GARCH or SV type models.

Correlation-based models are models of linear dependence, which are sufficient if the underlying distributions have an elliptical shape. However, one often finds empirically that there is an asymmetry in multivariate return distributions and that correlations change over time. In particular, clusters of large negative returns are much more frequent than clusters of large positive returns. In other words, there is lower tail dependence but no upper tail dependence. Copulas are a natural tool to model this effect and have the additional advantage of decoupling the models for the marginal distributions from those for the dependence. We review recent research on dynamic copula models and compare them to correlation-based models.

Finally, we consider approaches how to evaluate the quality of fitted models from a statistical and economic perspective. Two important criteria are, for example, the Value-at-Risk of portfolios and the portfolio selection problem.

2 The investor problem and potential complications

Since the seminal work of Markowitz [90], portfolio selection has become one of the main areas of modern finance. Today, investment strategies based on mean-variance optimization are considered the benchmark. A first problem of the standard approach is that the obtained optimal portfolio weights depend on second moments (variances and covariances) of the underlying asset returns, which are notoriously time-varying. In other words, the optimal portfolio can only be considered optimal for a short period of time, after which a re-balancing becomes necessary. Another problem is that the formula for optimal portfolio weights depends on the inverse of the covariance matrix, and that in high dimensions the covariance matrix is typically ill-behaved. Hence, portfolio selection might lead to suboptimal results in high dimensions when the standard formulas are applied.

A somewhat related problem is the numerical complexity of standard multivariate volatility models, where the number of parameters may explode as the dimension increases, which leads to intractable estimation and inference of these models. Moreover, in those models where the number of parameters is constant (such as the DCC model of Engle [49] see Section 4.2), there is

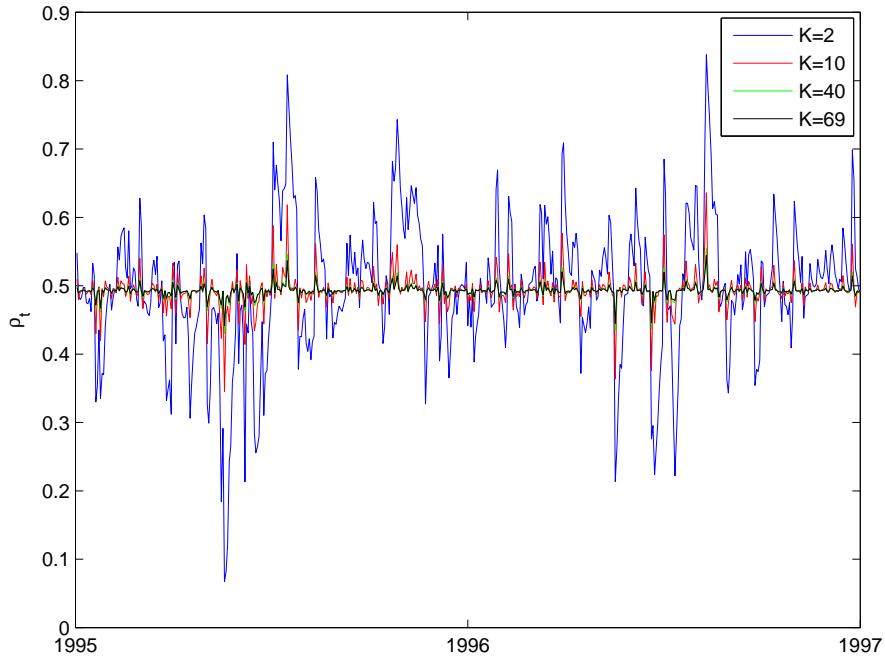


Fig. 1. Conditional correlations between two fixed assets for growing dimensions of the model

no problem in terms of model complexity, but another problem occurs: as the dimension increases, parameter estimates are downward biased and variation in correlations is underestimated, see e.g. [56]. In the following, we illustrate this effect using data of the London stock exchange.

We use the estimated (time varying) covariance matrix for the DCC model to construct the minimum variance portfolio (MVP). For the estimated covariance matrix \hat{H}_t , the MVP weights are

$$w_t = \frac{\hat{H}_t^{-1} \iota}{\iota^\top \hat{H}_t^{-1} \iota}, \quad (1)$$

where ι is an $(N \times 1)$ vector of ones.

The measure of interest is then the variance of the MVP, which should be minimal across different models, and the variance of the standardized portfolio returns given by $r_{p,t} = w_t^\top r_t / \sqrt{w_t^\top \hat{H}_t w_t}$, which should be close to one.

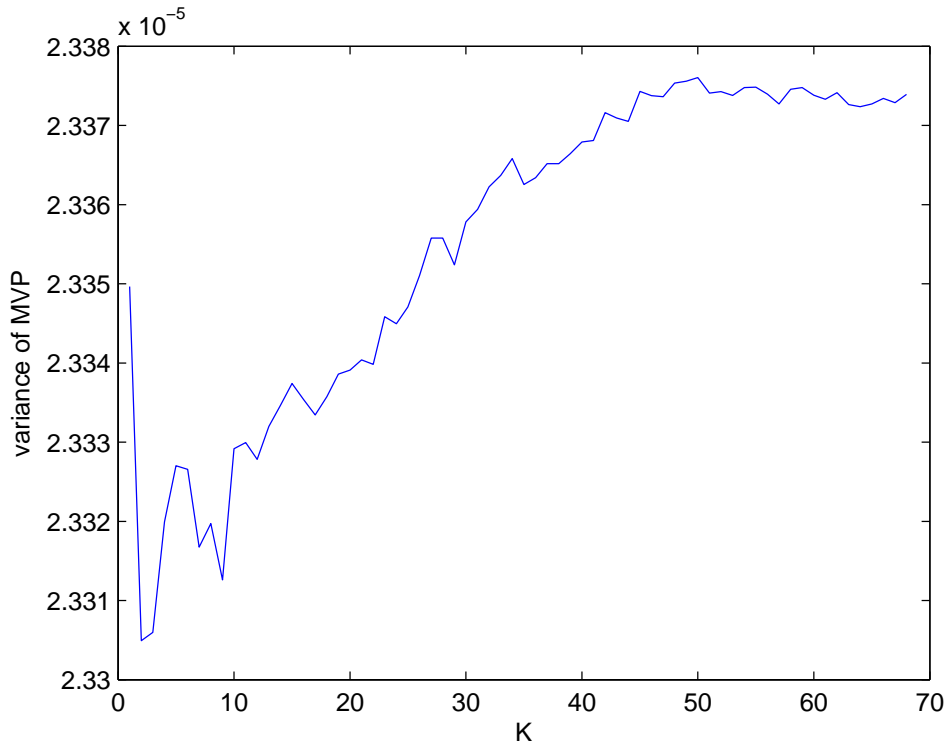


Fig. 2. Variance of the MVP of two fixed assets for growing dimensions of the model

To illustrate the potential problems that can occur when modelling large dimensional data sets we consider daily returns of 69 stocks that are part of the FTSE 100 index ranging from January 1995 until December 1996. We consider the problem of estimating conditional correlations and constructing the minimum variance portfolio (MVP) between the only the first two stocks in the data set. However, a model is fit to a larger data set and we look at the effect of including additional assets in the model.

Figure 2 shows the correlations between the first two assets of the sample estimated using the DCC Garch model by Engle in [49] as the number of assets in the sample K is increased. Surprisingly as the dimension of the data set increases the correlation dynamics are estimated with less precision and the conditional correlations become almost flat for K large as already noted in [56]. Using the covariance matrix estimated using the same sample we constructed the MVP for the first two assets using (1). The number of assets is increased from 2 to 69 and the variance of the resulting portfolio is plotted in Figure 2 as a function of K . The portfolio reaches the lowest variance for the model estimated using about 10 assets thus implying that the additional

information contained in the other series adds economic value. However, once K is increased further the variance grows again and the benefit of including more information in the data is outweighed by the numerical problems causing the flat estimates of the conditional correlations. As the dimension of the model grows further the problem is likely to become worse in addition to the computational complexity that makes estimating large dimensional models difficult.

3 Factor models for asset prices

Let $r_t = (r_{1t}, \dots, r_{Nt})^\top$ denote the vector of asset returns at time t , $t = 1, \dots, T$. Factor models assume that there is a small number K , $K < N$ of factors f_{kt} , $k = 1 \dots, K$, such that

$$r_t = a + Bf_t + \varepsilon_t, \quad (2)$$

where a is an $(N \times 1)$ vector, B an $(N \times K)$ loading matrix and ε_t a stochastic error term with mean zero and variance matrix Ω , uncorrelated with the factors. The idea of factor models is to separate common, non-diversifiable components from idiosyncratic, diversifiable ones. The idiosyncratic error terms are usually assumed to be uncorrelated so that Ω is diagonal, in which case one speaks of a *strict factor model*. If the factors are stationary with mean zero and variance matrix Σ , then returns are stationary with mean a and variance

$$H := \text{Var}(r_t) = B\Sigma B^\top + \Omega. \quad (3)$$

Dynamic properties of the factors typically carry over to returns. For example, if factors are nonstationary with time-varying variance Σ_t , then returns will also be nonstationary with variance $H_t = B\Sigma_t B^\top + \Omega$. Another example is that of conditional heteroscedasticity, where factors can be stationary but conditioned on the information of lagged factors, the variance Σ_t is time-varying. Models for Σ_t and H_t will be discussed in the next section.

Note that factor models are identified only up to an invertible rotation of the factors and the loading matrix. To see this, let G be an invertible $(K \times K)$ matrix and write (2) equivalently as $r_t = a + BGG^{-1}f_t + \varepsilon_t$, then we have the same model but with factors $\tilde{f}_t = G^{-1}f_t$ and loading matrix $\tilde{B} = BG$. Thus, only the K -dimensional factor space can be identified, not the factors themselves.

Two types of factor models are usually distinguished: those with observed and unobserved factors. When factors are observed, then simple estimation methods such as OLS can be used to estimate the parameters a and the loading matrix B . The most popular example of an observed one-factor model in finance is the capital asset pricing model (CAPM), developed by [100] and [86], where the single factor is the market portfolio, which is usually approximated

by an observable broad market index. Several empirical anomalies have been found which led to the three-factor model of [57], where additional to the market factor there is a second factor explaining differences in book to market values of the stocks, and a third factor controlling for differences in market capitalization or sizes of the companies. A general multifactor asset pricing model has been proposed by [98] in his arbitrage pricing theory (APT).

When factors are unobserved, estimation becomes more involved. Imposing structure on Ω and Σ it is possible to do maximum likelihood estimation, but in high dimensions this is often infeasible. On the other hand, [31] have shown that by allowing Ω to be non-diagonal and hence defining an *approximate factor model*, one can consistently estimate the factors (up to rotation) using principal components regression if both the time and cross-section dimension go to infinity. [13] provides inferential theory for this situation, whereas [37] and [14] propose tests for the number of factors in an approximate factor model.

In order to render the factor model *dynamic*, several approaches have been suggested recently. A stationary dynamic factor model specifies the loading matrix B as a lag polynomial $B(L)$ where L is the lag operator and factors follow a stationary process, for example a vector autoregression. [60] apply the dynamic principal components method by [26] to estimate the common component $B(L)f_t$ in the frequency domain. Forecasting using the dynamic factor model has been investigated e.g. by [101]. A recent review of dynamic factor models is given by [24].

Rather than considering stationary processes, [91] follow another approach where factors are stationary but the loading matrix B is a smooth function of time, and hence returns are non-stationary. Estimation is performed using localized principal components regression. To extend the idea of dynamic factor models to the nonstationary case, [45] let the lag polynomial $B(L)$ be a function of time and show asymptotic properties of the frequency domain estimator for the common components.

4 Volatility and dependence models

4.1 Univariate volatility models

In this section we review alternative univariate models for volatility: GARCH, stochastic volatility and realized volatility.

GARCH

The generalized autoregressive conditional heteroskedasticity (GARCH) model introduced by [48] and [21] suggests the following specification for asset returns r_t ,

$$\begin{aligned} r_t &= \mu_t + \varepsilon_t, & \varepsilon_t &= \sigma_t \xi_t \\ \sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \end{aligned} \quad (4)$$

where $\xi_t \sim N(0, 1)$ and μ_t is the mean, conditional on the information set at time $t - 1$. For example, the CAPM mentioned in Section 3 implies that for the return on the market portfolio, $\mu_t = r_f + \lambda \sigma_t^2$, where r_f is the risk free interest rate, λ the market price of risk and σ_t^2 market volatility that could be explained by the GARCH model in (4). This is the so-called GARCH-in-mean or GARCH-M model of [52].

For σ_t^2 in (4) to be a well defined variance, sufficient conditions for positivity are $\omega > 0$ and $\alpha \geq 0, \beta \geq 0$. Higher order models that include more lags of ε_t and σ_t^2 are possible but rarely used in practice. A more serious restriction of the standard GARCH model is that recent errors ε_t have a symmetric impact on volatility with respect to their sign. Empirically, one has often observed a *leverage effect*, meaning a higher impact of negative errors than positive ones. Many extensions of the standard GARCH model have been proposed, see e.g. [75] for a review of alternative specifications.

The GARCH(1,1) process in (4) is covariance stationary if and only if $\alpha + \beta < 1$, in which case the unconditional variance of ε_t is given by $\sigma^2 = \omega / (1 - \alpha - \beta)$. In the GARCH-M case with $\mu_t = r_f + \lambda \sigma_t^2$, the unconditional first two moments of r_t are given by $E[r_t] = r_f + \lambda \sigma^2$ and $\text{Var}(r_t) = \lambda^2 \text{Var}(\sigma_t^2) + \sigma^2$. Note that a positive autocorrelation of σ_t^2 induces a similar autocorrelation in returns in the GARCH-M model. This corresponds to empirical evidence of significant first order autocorrelations in daily or weekly stock returns, see e.g. Chapter 2 of [28]. Straightforward calculations show that the τ -order autocorrelation of r_t is given by

$$\rho(\tau) = (\alpha + \beta)^\tau \frac{\lambda^2 \text{Var}(\sigma_t^2)}{\lambda^2 \text{Var}(\sigma_t^2) + \sigma^2}, \quad \tau \geq 1.$$

Compared with an AR(1) model with $\mu_t = \phi r_{t-1}$ for which $\rho(\tau) = \phi^\tau$, these autocorrelations could be matched for $\tau = 1$, but at higher orders the GARCH-M model would imply higher autocorrelation than the AR(1) model. [70] compared the GARCH-M and AR(1) specifications and found that in most cases the AR(1) model, although without economic motivation, provides a better fit to the data. Obviously, if $\lambda = 0$, then r_t is white noise with $\rho(\tau) = 0$ for all $\tau \neq 0$. An effect of nonzero autocorrelation of returns does not violate the hypothesis of market efficiency, as the autocorrelation is explained by a time-varying risk premium, see e.g. [52].

The GARCH model implies that returns y_t have a fat tailed distribution, which corresponds to empirical observations already found by [58] and [89]. In particular, assuming $\xi_t \sim N(0, 1)$ and finite fourth moments of r_t by the condition $\beta^2 + 2\alpha\beta + 3\alpha^2 < 1$, the GARCH(1,1) process in (4) has an unconditional kurtosis given by

$$\kappa = 3 + \frac{6\alpha^2}{1 - \beta^2 - 2\alpha\beta - 3\alpha^2}$$

where the second term is positive such that $\kappa > 3$. Thus, while the conditional distribution of r_t is Gaussian, the unconditional one is fat-tailed. Furthermore, there is volatility clustering in the sense that there are periods of high volatility and other periods of low volatility. This reflected by a positive autocorrelation of squared error terms.

Estimation of GARCH models is rather straightforward. Suppose one can separate the parameter ϕ that describes the conditional mean μ_t from the volatility parameter $\theta = (\omega, \alpha, \beta)'$. Assuming normality of ξ_t , one can write the log likelihood function for a sample of T observations up to an additive constant as

$$L(\phi, \theta) = -\frac{1}{2} \sum_{t=1}^T \left[\log \sigma_t^2(\theta) + \frac{\{y_t - \mu_t(\phi)\}^2}{\sigma_t^2(\theta)} \right]$$

which is maximized numerically w.r.t. ϕ and θ . Under weak regularity conditions, [23] show that $\sqrt{T}(\hat{\theta} - \theta) \rightarrow N(0, J^{-1})$ where J is the Fisher information matrix.

Stochastic Volatility

Stochastic volatility (SV) models offer a good alternative to capture time-varying variances of asset returns. They originated in different branches of the literature such as financial economics, option pricing and the modelling of financial markets in order to relax the constant variances assumption. For example, [76] allow volatility to follow a general diffusion in their option pricing model. [36] introduced a model where the information flow to the market is specified as a log-normal stochastic process, which results in a mixture of normal distributions for asset prices. [102] accommodated the persistence in volatility and suggested the following autoregressive SV model, which is the most common formulation.

$$r_{it} = \mu_{it} + \exp(h_{it}/2)\xi_{it} \tag{5}$$

$$h_{it+1} = \delta_i + \gamma_i h_{it} + \sigma_{\eta_i} \eta_{it} \tag{6}$$

ξ_{it} and η_{it} are standard normal innovations and are potentially (negatively) correlated, which leads to a statistical leverage effect meaning that price drops lead to increases in future volatility. σ_{η_i} is assumed to be positive and for $|\gamma| < 1$ the returns r_{it} are strictly stationary. This basic specification is able to explain the fat-tailed return distributions and persistence in volatility well due to the flexibility introduced by the error term. In fact, the Gaussian SV model fit financial data considerably better than a Normal GARCH(1,1) model and it performs about as well as a GARCH model with Student-t innovations. [103], [66] and [8] are excellent reviews on SV models and some extensions. Estimation of SV models, which is reviewed in [27], is not trivial and probably the main reason why ARCH models are considered more often

in empirical studies. Estimation can be done by many different techniques such as the method of moments (see [102]), quasi maximum likelihood using the Kalman filter in [74], the simulated method of moments by [44], [68] and [62], Markov Chain Monte Carlo (MCMC) estimation by [77] and [81], and simulation based maximum likelihood estimations using importance sampling (IS) by [38], [40] and [84]. We recommend using either MCMC or IS methods for estimating the parameters and latent volatility process in a SV model, as these offer very efficient estimates and the considerable computational effort can be handled easily by modern computers.

Realized Volatility

With the availability of high-frequency data, by which we mean price data observed every 5 minutes or even more often, a new set of very powerful tools for volatility estimation and modelling has evolved, namely realized volatility and related concepts. The information contained in high-frequency data allows for improved estimation and forecasting of volatility compared to using only daily data. Furthermore, realized volatility measure relate closely to continuous time SV models and one only needs to assume that the return process is arbitrage free and has a finite instantaneous mean. This in turn implies that the price process is a semi-martingale that the returns can be decomposed into a predictable and integrable mean component and a local martingale. This includes the continuous time stochastic volatility diffusion

$$dp_t = \mu_t dt + \sigma_t dW_t, \quad (7)$$

where W_t denotes Brownian motion and the volatility process σ_t is assumed to be stationary. Denote the continuously compounded h period return by $r_{t+h,h} \equiv p_{t+h} - p_t$, where one usually chooses $h = 1$ to be one trading day. Consider a sample of $1/\Delta$ observations per day. In practice Δ is often chosen to be $1/288$ corresponding to 5-minute returns, although this clearly depends on the data set. Sampling too frequently can lead to a bias due to microstructure noise in the data. Then realized variance for day t is defined as

$$RV = \sum_{j=1}^{h/\Delta} r_{t+j\Delta,\Delta}^2. \quad (8)$$

This is a consistent estimator of the quadratic variation and, if the price process does not exhibit any jumps, also of the integrated variance $\int_0^h \sigma_{t+s}^2 ds$. However, in the presence of jumps quadratic variation decomposes into integrated variance and the quadratic variation of the jump component. [18] propose a measure that consistently estimates the integrated variance even in the presence of jumps. This estimator, called bipower variation, is defined as

$$BPV = \frac{\pi}{2} \sum_{j=2}^{h/\Delta} |r_{t+j\Delta,\Delta}| |r_{t+(j-1)\Delta,\Delta}|. \quad (9)$$

Thus it is possible to separate the continuous and the jump components of volatility by estimating both realized variance and bipower variation, and to identify the jumps by looking at the difference between the two.

Convergence in probability of RV was established by [5]. Empirical properties of RV are documented in [4] and [3], such as approximate log-normality, high correlation across different RV series, and long memory properties of volatilities. Forecasting of volatility and the gains that can be made by using high frequency data are discussed in [6]. [9] consider latent factor models for RV series and show that these can help forecasting volatilities. The asymptotic distribution of the RV measure and connections to SV models are provided in the notable contributions [15] and [16].

4.2 Multivariate volatility models

Multivariate GARCH models

GARCH models have been vastly applied to multivariate problems in empirical finance. The typically large number of assets, however, caused problems in early years where models were too complex with too many parameters to estimate. For example, the BEKK model of [51] specifies the conditional covariance matrix H_t as

$$H_t = C_0 C_0^\top + A \varepsilon_{t-1} \varepsilon_{t-1}^\top A^\top + B H_{t-1} B^\top \quad (10)$$

where C_0 , A and B are $N \times N$ parameter matrices and C_0 is upper triangular. The model (10) is the simplest version of a BEKK model, but higher order models are rarely used. An advantage of the classical BEKK model is its flexibility and generality while generating implicitly a positive definite H_t . However, the number of parameters to estimate is $O(N^2)$, which revealed to be infeasible in high dimensions.

In the following we will therefore concentrate on two model classes, factor GARCH and DCC models, that can be applied to hundreds or thousands of assets. Factor models can be shown to be restricted versions of the BEKK model in (10), while DCC type models form a separate, non-nested class of models. A broad overview of multivariate GARCH models has been given recently by [19].

Suppose there are N asset returns, r_{1t}, \dots, r_{Nt} , $t = 1, \dots, T$. A model with K factors can be written as

$$r_{it} = b_{i1} f_{1t} + \dots + b_{iK} f_{Kt} + \varepsilon_{it}, \quad i = 1, \dots, N$$

where ε_{it} is an idiosyncratic white noise sequence. In matrix notation this is just the model given in (2). If factors follow univariate GARCH processes with conditional variance σ_{it}^2 and are conditionally orthogonal, then the conditional variance of r_{it} can be written as

$$h_{it} = \sum_{k=1}^K b_{ik}^2 \sigma_{it}^2 + \omega_i$$

where $\omega_i = \text{Var}(\varepsilon_{it})$. Factors can be observed assets as in [54] or latent and estimated using statistical techniques. For example, the Orthogonal GARCH model of [1] uses principal components as factors and the eigenvalues of the sample covariance matrix to obtain the factor loadings, before estimating the univariate GARCH models of the factors. [105] generalizes the O-GARCH model to allow for multiplicities of eigenvalues while maintaining identifiability of the model.

A second class of models has attracted considerable interest recently, the class of dynamic conditional correlation (DCC) models introduced by [49] and [104]. In the standard DCC model of order (1,1), conditional variances h_{it} are estimated in a first step using e.g. univariate GARCH. Then, standardized residuals $e_{it} = (r_{it} - \mu_{it})/\sqrt{h_{it}}$ are obtained and the conditional correlation is given by

$$R_{ij,t} = \frac{Q_{ij,t}}{\sqrt{Q_{ii,t}Q_{jj,t}}}$$

where $Q_{ij,t}$ is the (i, j) -element of the matrix process Q_t ,

$$Q_t = S(1 - \alpha - \beta) + \alpha e_{t-1} e_{t-1}^\top + \beta Q_{t-1} \quad (11)$$

with S being the sample covariance matrix of e_{it} . In the special case of $\alpha = \beta = 0$, one obtains the constant conditional correlation (CCC) model of [22].

Splitting the joint likelihood into conditional mean, variance and correlation parameters, the part of the likelihood corresponding to the correlation parameters can be written as

$$\log L(\alpha, \beta) = -\frac{1}{2} \sum_{t=1}^T (\log |R_t| + e_t^\top R_t^{-1} e_t) \quad (12)$$

An interesting feature of estimators that maximize (12) is that for increasing dimension N the α estimates appear to go to zero, as noted already by [55]. [56] argue that this may be due to the first stage estimation of the conditional variance parameters and the sample covariance matrix S . The parameters of the first stage can be viewed as nuisance parameters for the estimation of the second stage. The covariance targeting idea used in the specification of (11) depends on one of these nuisance parameters, S . The effect, clearly demonstrated in simulations by [56] and [71], is a negative bias for the α estimate, thus delivering very smooth correlation processes in high dimensions and eventually estimates that converge to the degenerate case of a CCC model. [56] propose to use a so-called composed likelihood estimation, where the sum of quasi-likelihoods over subsets of assets is maximized. They show that this approach does not suffer from bias problems in high dimensions.

Another reason why maximization of (12) is not suitable in high dimensions is numerical instability due to almost singular matrices R_t and the problem of inverting this matrix at every t . The sample covariance matrix S is typically ill-conditioned, meaning that the ratio of its largest and smallest eigenvalue is huge. In this case, shrinkage methods as in [83] could possibly be applied to S to improve the properties of the DCC estimates.

A limitation of the classical DCC model in (11) is that only two parameters, α and β , drive the dynamic structure of a whole covariance matrix, possibly of high dimension. This seems implausible if N is large, say 50 or higher. [71] proposed to generalize the DCC model as

$$Q_t = S \odot (1 - \bar{\alpha}^2 - \bar{\beta}^2) + \alpha \alpha^\top \odot \varepsilon_{t-1} \varepsilon_{t-1}^\top + \beta \beta^\top \odot Q_{t-1}$$

where now α and β are $(N \times 1)$ vectors, \odot is the Hadamard product, i.e. elementwise multiplication, and $\bar{\alpha} = (1/N) \sum_i \alpha_i$ and $\bar{\beta} = (1/N) \sum_i \beta_i$. This generalized version of the DCC model has the advantage of still guaranteeing a positive definite Q_t and R_t while being much more flexible in allowing some correlations to be very smooth and others to be erratic.

Multivariate stochastic volatility models

The basic specification for a multivariate stochastic volatility model (MSV) introduced by [74] is given by

$$r_t = \mu_t + H_t^{1/2} \xi_t \tag{13}$$

$$H_t^{1/2} = \text{diag}\{\exp(h_{1t}), \dots, \exp(h_{Nt})\}$$

$$h_{it+1} = \delta_i + \gamma_i h_{it} + \eta_{it}, \text{ for } i = 1, \dots, N \tag{14}$$

$$\begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} P_\xi & 0 \\ 0 & \Sigma_\eta \end{pmatrix} \right] \tag{15}$$

where $\mu_t = (\mu_{1t}, \dots, \mu_{Nt})^\top$, $\xi_t = (\xi_{1t}, \dots, \xi_{Nt})^\top$ and $\eta_t = (\eta_{1t}, \dots, \eta_{Nt})^\top$. Σ_η is a positive-definite covariance matrix and P_ξ is a correlation matrix capturing the contemporaneous correlation between the return innovations. Of course, both correlations between the mean innovations and the volatility innovations can be restricted to be zero to reduce the number of parameters. If one only assumes that the off-diagonal elements of Σ_η are equal to zero this specification corresponds to the constant conditional correlation (CCC) GARCH model by [22], since no volatility spillovers are possible.

This basic model has relatively few parameters to estimate $(2N + N^2)$, but [39] shows that it outperforms standard Vector-GARCH models that have a higher number of parameters. Nevertheless, a number of extensions of this model are possible. First, one can consider heavy tailed distributions for the innovations in the mean equation ξ_t in order to allow for higher excess kurtosis compared to the Gaussian SV model, although in most cases this seems to be unnecessary. [74] suggest using a multivariate t-distribution for that purpose.

A second simple and natural extension of the basic model can be achieved by introducing asymmetries into the model. One possibility is to replace (15) by

$$\begin{aligned} \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} &\sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} P_\xi & L \\ L & \Sigma_\eta \end{pmatrix} \right] \\ L &= \text{diag}\{\lambda_1 \sigma_{\eta,11}, \dots, \lambda_N \sigma_{\eta,NN}\}, \end{aligned} \quad (16)$$

where $\sigma_{\eta,ii}$ denotes the i 'th diagonal element of Σ_η and λ_i is expected to be negative for $i = 1, \dots, N$. This specification allows for a statistical leverage effect. [12] distinguish between leverage, denoting negative correlation between current returns and future volatility, and general asymmetries meaning negative returns have a different effect on volatility than positive ones. These asymmetric effects may be modeled as a threshold effect or by including past returns and their absolute values, in order to incorporate the magnitude of the past returns, in equation (14). The latter extension was suggested by [38] and is given by

$$h_{it+1} = \delta_i + \phi_{i1} y_{it} + \phi_{i2} |y_{it}| + \gamma_i h_{it} + \sigma_{\eta_i} \eta_{it}. \quad (17)$$

A potential drawback of the basic models and its extensions is that the number of parameters grows with N and it may become difficult to estimate the model with a high dimensional return vector. Factor structures in MSV models are a possibility to achieve a dimension reduction and make the estimation of high dimensional systems feasible. Furthermore, factor structures can help identify common features in asset returns and volatilities and thus relate naturally to the factor models described in Section 3. [42] propose a multivariate ARCH model with latent factors that can be regarded as the first MSV model with a factor structure, although [74] are the first to propose the use of common factors in the SV literature. Two types of factor SV models exist: Additive factor models and multiplicative factor models. An additive K factor model is given by

$$\begin{aligned} r_t &= \mu_t + D f_t + e_t \\ f_{it} &= \exp(h_{it}/2) \xi_{it} \\ h_{it+1} &= \delta_i + \gamma_i h_{it} + \sigma_{\eta_i} \eta_{it}, \text{ for } i = 1, \dots, K, \end{aligned} \quad (18)$$

with $e_t \sim N(0, \text{diag}(\sigma_1^2, \dots, \sigma_N^2))$, $f_t = (f_{1t}, \dots, f_{Kt})^\top$, D is an $N \times K$ matrix of factor loadings and $K < N$. Identification is achieved by setting $D_{ii} = 1$ for all $i = 1, \dots, N$ and $D_{ij} = 0$ for all $j < i$. As mentioned in [12] a serious drawback of this specification is that homoscedastic portfolios can be constructed, which is unrealistic. Assuming a SV model for each element of e_t can solve this problem, although it does increase the number of parameters again. Furthermore, the covariance matrix of e_t is most likely not diagonal. A further advantage of the model is that it does not only accommodate time-varying volatility, but also time-varying correlations, which reflects the important stylized fact that

correlations are not constant over time. A multiplicative factor model with K factors is given by

$$\begin{aligned} r_t &= \mu_t + \exp\left(\frac{wh_t}{2}\right) \xi_t \\ h_{it+1} &= \delta_i + \gamma_i h_{it} + \sigma_{\eta_i} \eta_{it}, \text{ for } i = 1, \dots, K, \end{aligned} \quad (19)$$

where w is an $N \times K$ matrix of factor loadings that is of rank K and $h_t = (h_{1t}, \dots, h_{Kt})^\top$. This model is also called stochastic discount factor model.

Although factor MSV models allow for time-varying correlations these are driven by the dynamics in the volatility. Thus a further extension of the basic model is to let the correlation matrix P_ξ depend on time. For the bivariate case [106] suggest the following specification for the correlation coefficient ρ_t .

$$\begin{aligned} \rho_t &= \frac{\exp(2\lambda_t) - 1}{\exp(2\lambda_t) + 1} \\ \lambda_{t+1} &= \delta_\rho + \gamma_\rho \lambda_t + \sigma_\rho z_t, \end{aligned} \quad (20)$$

where $z_t \sim N(0, 1)$. A generalization to higher dimensions of this model is not straightforward. [106] propose the following specification following the DCC specification of [49].

$$\begin{aligned} P_{\xi t} &= \text{diag}(Q_t^{-1/2}) Q_t \text{diag}(Q_t^{-1/2}) \\ Q_{t+1} &= (\iota^\top - A - B) \odot S + B \odot Q_t + A \odot z_t z_t^\top, \end{aligned} \quad (21)$$

where $z_t \sim N(0, I)$, ι is a vector of ones. An alternative to this is the model by [11], which also uses the DCC specification, but the correlations are driven by a Wishart distribution.

Further specifications of MSV models along with a large number of references can be found in [12], whereas [106] compares the performance of a number of competing models. One main finding of this study is that models that allow for time-varying correlations clearly outperform constant correlation models.

Estimation can in principle be done using the same methods suggested for univariate models, although not each method may be applicable to every model. Still, simulated maximum likelihood estimation and MCMC estimation appear to be the most flexible and efficient estimation techniques available for MSV models.

Realized covariance

The definition of realized volatility extends to the multivariate case in a straightforward fashion and thus the additional information contained in high frequency data can also be exploited when looking at covariance, correlation and simple regressions. Some references are [3] and [4] providing definitions,

consistency results and empirical properties of the multivariate realized measures. [17] provide a distribution theory for realized covariation, correlation and regression, the authors discuss how to calculate confidence intervals in practice. A simulation study illustrates the good quality of their approximations in finite samples when Δ is small enough (about 1/288 works quite well). Let the h period return vector be $r_{t+h,h}$. Then realized covariance is defined as

$$RCOV = \sum_{j=1}^{h/\Delta} r_{t+j\Delta,\Delta} r_{t+j\Delta,\Delta}^\top. \quad (22)$$

The realized correlation between return on asset k , $r_{(k)t+h,h}$, and the return of asset l , $r_{(l)t+h,h}$, is calculated as

$$RCORR = \frac{\sum_{j=1}^{h/\Delta} r_{(k)t+j\Delta,\Delta} r_{(l)t+j\Delta,\Delta}}{\sqrt{\sum_{j=1}^{h/\Delta} r_{(k)t+j\Delta,\Delta}^2 \sum_{j=1}^{h/\Delta} r_{(l)t+j\Delta,\Delta}^2}}. \quad (23)$$

Finally, the regression slope when regressing variable l on variable k is given by

$$\hat{\beta}_{(lk),t} = \frac{\sum_{j=1}^{h/\Delta} r_{(k)t+j\Delta,\Delta} r_{(l)t+j\Delta,\Delta}}{\sum_{j=1}^{h/\Delta} r_{(k)t+j\Delta,\Delta}^2}. \quad (24)$$

All these quantities have been shown to follow a mixed normal limiting distribution. An application of the concept of realized regression is given in [7], where the authors compute the realized quarterly betas using daily data and discuss its properties.

Dynamic copula models

A very useful tool for specifying flexible multivariate versions of any class of distribution functions are copulas. A copula is, loosely speaking, that part of a multivariate distribution function that captures all the contemporaneous dependence. The most important results concerning copulas known as Sklar's theorem tells us that there always exists a copula such that any multivariate distribution function can be decomposed into the marginal distributions capturing the individual behavior of each series and a copula characterizing the dependence structure. This separation does not only allow for an easy and tractable specification of multivariate distributions, but also for a two-step estimation greatly reducing the computational effort. Thus any of the volatility models described above can be generalized to the multivariate case in a straightforward fashion by coupling the univariate models using copulas. Furthermore, dependence structures that go beyond linear correlation such as tail dependence and asymmetric dependencies, which is useful when markets or

stocks show stronger correlation for negative than for positive returns, can be allowed for. [92] provides a mathematical introduction to the topic, whereas [78] treats the topic from a statistical viewpoint. [35] and [61] look at copulas and their applications for financial problems.

Consider the N -dimensional return vector $r_t = (r_{1t}, \dots, r_{Nt})^\top$. Let F_i be the marginal distribution function of return i at let H be the joint distribution function of r_t . Then by Sklar's theorem there exists a copula function C such that

$$H(r_{1t}, \dots, r_{Nt}) = C\{F_1(r_{1t}), \dots, F_N(r_{Nt})\}. \quad (25)$$

Additionally, if the marginals are continuous the copula is unique. Recalling that by the probability integral transform the variable $u_{it} = F_i(r_{it})$ follows a standard uniform distribution it becomes clear that a copula is simply a multivariate distribution function with $U(0, 1)$ marginals.

A large number of examples of copula function and methods to simulate artificial data from them, which is extremely useful for the pricing of derivatives with multiple underlying assets, is discussed in the chapter "Copulae Modelling" in this handbook. However, here we focus our attention on the situation when the copula is allowed to vary over time, which accommodates the special case of time-varying correlations, a feature usually observed in financial data. Dynamic copulas can thus be used to construct extremely flexible multivariate volatility models that tend to fit the data better than models assuming a dependence structure that is fixed over time. In what follows we denote the time-varying parameter of a bivariate copula by θ_t .

Structural breaks in dependence: A formal test for the presence of a breakpoint in the dependence parameter of a copula was developed in [41]. Denote η_t 's the parameters of the marginal distributions, which are treated as nuisance parameters. Formally, the null hypothesis of no structural break in the copula parameter becomes

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_T \text{ and } \eta_1 = \eta_2 = \dots = \eta_T$$

whereas the alternative hypothesis of the presence of a single structural break is formulated as:

$$H_1 : \theta_1 = \dots = \theta_k \neq \theta_{k+1} = \dots = \theta_T \equiv \theta_k^* \text{ and } \eta_1 = \eta_2 = \dots = \eta_T.$$

In the case of a known break-point k , the test statistics can be derived as a generalized likelihood ratio test. Let $L_k(\theta, \eta)$, $L_k^*(\theta, \eta)$ and $L_T(\theta, \eta)$ be the log-likelihood functions corresponding to a copula based multivariate model using the first k observations, the observations from $k+1$ to T and all observations, respectively. Then the likelihood ratio statistic can be written as

$$LR_k = 2[L_k(\hat{\theta}_k, \hat{\eta}_T) + L_k^*(\hat{\theta}_k^*, \hat{\eta}_T) - L_T(\hat{\theta}_T, \hat{\eta}_T)],$$

where a hat denotes the maximizer of the corresponding likelihood function. Note that $\hat{\theta}_k$ and $\hat{\theta}_k^*$ denote the estimates of θ before and after the break, whereas $\hat{\theta}_T$ and $\hat{\eta}_T$ are the estimates of θ and η using the full sample. In the case of an unknown break date k , a recursive procedure similar to the one proposed in [10] can be applied. The test statistic is the supremum of the sequence of statistics for known k

$$Z_T = \max_{1 \leq k < T} LR_k \quad (26)$$

and the asymptotic critical values of [10] can be used. [29] extended the procedure to additionally allow for a breakpoint in the unconditional variance of the individual series at a (possibly) different point in time and they discuss how to estimate the breakpoints in volatility and in dependence sequentially.

The conditional copula model: [94] showed that Sklar's theorem still holds for conditional distributions and suggested the following time varying specification for copulas. For the Gaussian copula correlation evolves, similarly to the DCC model, as

$$\rho_t = \Lambda \left\{ \alpha + \beta_1 \cdot \rho_{t-1} + \beta_2 \cdot \frac{1}{p} \sum_{j=1}^p \Phi^{-1}(u_{1,t-j}) \cdot \Phi^{-1}(u_{2,t-j}) \right\}, \quad (27)$$

where, $\Lambda(x) = \frac{1-e^{-x}}{1+e^{-x}}$ is the inverse Fisher transformation. The number of lags p is chosen to be 10, although this is a rather arbitrary choice that may be varied. For copulas different from the Gaussian the sum in (27) is replaced by $\sum_{j=1}^p |u_{1,t-j} - u_{2,t-j}|$ and Λ has to be replaced by a transformation appropriate to ensure the dependence parameter is in the domain of the copula of interest.

Adaptive estimation of time-varying copulas: In order to save some space we refer to the chapter "Copulae Modelling" in this handbook for a description of these techniques to estimate dynamic copulas introduced by [67].

Stochastic dynamic copulas: While the model by Patton can be seen as the counterpart to a GARCH model, where correlations are a function of the past observation, in [72] we propose to let the dependence parameter of a copula follow a transformation of a Gaussian stochastic process. That has, similar to stochastic volatility models, the advantage of being a bit more flexible than a DCC model or the specification by Patton at the cost of being more difficult to estimate. Furthermore, it is a natural approach for a multivariate extension of stochastic volatility models.

We assume that θ_t is driven by an unobserved stochastic process λ_t such that $\theta_t = \Psi(\lambda_t)$, where $\Psi : \mathbb{R} \rightarrow \Theta$ is an appropriate transformation to ensure that the copula parameter remains in its domain and whose functional form depends on the choice of copula. The underlying dependence parameter λ_t ,

which is unobserved, is assumed to follow a Gaussian autoregressive process of order one,

$$\lambda_t = \alpha + \beta\lambda_{t-1} + \nu\varepsilon_t, \quad (28)$$

where ε_t is an i.i.d. $N(0, 1)$ innovation. Since λ_t is unobservable it must be integrated out of the likelihood function. Such a T dimensional integral cannot be solved analytically. However, λ_t can be integrated out by Monte Carlo integration using the efficient importance sampler of [84].

Local likelihood estimation of dynamic copulas: A model which allows θ_t to change over time in a non-parametric way is proposed in [73]. It is assumed that the copula parameter can be represented as a function $\theta(t/T)$ in rescaled time. If that function is sufficiently smooth then the bivariate return process is locally stationary. Estimation is done in two steps, where first GARCH models for the margins are estimated and in the second step the time-varying copula parameter is estimated by local maximum likelihood estimation. That means that the log-likelihood function is locally weighted by a kernel function. Additionally, a one step correction for the estimates of the GARCH parameters ensures semi-parametric efficiency of the estimator, which is shown to work well in simulations.

Assessing the quality of the models

For practical purposes it is important to have a way to distinguish among the many competing models. For testing a particular feature of a model such as the leverage effect one can often apply standard hypothesis tests such as t-tests or likelihood ratio tests. When competing models do not belong to the same model class and are non-nested this is usually not possible anymore. Here we do not only consider statistical criteria to assess how well a given model can describe that data, but we also look at some economic measures that compare the usefulness of competing models for certain investment decisions.

The simplest way to compare the in-sample fit of competing models is to look at the value of the log-likelihood function at the parameter estimates, which gives a good indication of how well the statistical model describes a given data set. Since not all models have the same number of parameters and since models with a larger number of parameters will most of the time fit the data better due to more flexibility, it is often recommendable to use some type of information criterion that penalizes a large number of parameters in a model. The two most commonly used information criteria are the Akaike information criterion given by

$$AIC = -2LL + 2p \quad (29)$$

and the Bayesian information criterion

$$BIC = -2LL + p \log(T), \quad (30)$$

where LL denotes the value of log-likelihood function, T is the sample size and p is the number of parameters in the model. The model with the smallest value for either AIC or BIC is then considered the best fitting one, where the BIC tends to favor more parsimonious models. However, even the best fitting model from a set of candidate models may not provide reasonable fit for the data, which is why distributional assumptions are often tested using specific goodness-of-fit tests such as the Jarque-Bera test for normality, the Kolmogorov-Smirnov test or the Anderson-Darling test. One may also want to test for i.i.d.'ness of the standardized residuals of the candidate model by testing for remaining autocorrelation and heteroscedasticity. Finally, one may be interested in comparing the out-of-sample performance of a number of models. We refer to [43] for possible procedures. When comparing the forecasting performance of volatility models realized volatility offers itself naturally as a measure for the (unobserved) variance of a series.

Although a good statistical fit of a model is a desirable feature of any model a practitioner may be more interested in the economic importance of using a certain model. A very simple, yet informative measure is the Value-at-Risk (VaR), which measures how much money a portfolio will lose at least with a given probability. For portfolio return y_t the VaR at quantile α is defined as $P[y_t < \text{VaR}_\alpha] = \alpha$. The VaR can be computed both in sample and out-of-sample and [53] suggest a test to assess the quality of a VaR estimate for both cases. A related measure is the expected shortfall (ES), which is the expected loss given that the portfolio return lies below a specific quantile, i.e. $\text{ES}_\alpha = E(y_t | y_t < \text{VaR}_\alpha)$. As portfolio managers are often interested to minimize the risk of their portfolio for a given target return models can be compared by their ability to construct the minimum variance portfolio as suggested by [32]. The minimum variance portfolio can be considered and the conditional mean can be ignored as it is agreed on that the mean of stock returns is notoriously difficult to forecast, especially for returns observed at a high frequency. A similar approach was taken in [59] to evaluate the economic values of using sophisticated volatility models for portfolio selection. Since portfolio manager often aim at reproducing a certain benchmark portfolio [32] also suggest to compare models by their ability to minimize the tracking error volatility, which is the standard deviation of the difference between the portfolio's return and the benchmark return.

5 Data illustration

In this section we want to illustrate some of the techniques mentioned above for modelling a multi-dimensional time series of asset prices. The data we consider are those 69 stocks from the FTSE 100 index that were included in that index over our whole sample period. We look at daily observations from the

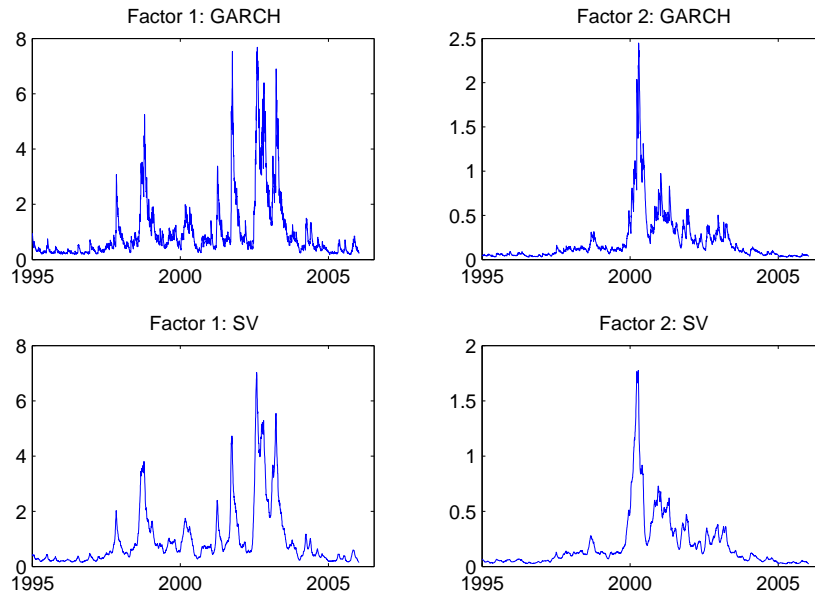


Fig. 3. Conditional volatilities of the first two factors

beginning of 1995 until the end of 2005 and calculate returns by taking the first difference of the natural logarithm. We multiply returns by 100 to ensure stability of the numerical procedures used for estimation. Modelling 69 assets is still less than the vast dimensions required for practical applicability, but it is already quite a large number for many multivariate time series models and much more than what is used in most studies. Fitting a 69 dimensional volatility model directly to the data is not possible for many of the models presented above, mainly because the number of parameters grows rapidly with the dimension of the problem and estimation becomes difficult or even impossible. We therefore impose a lower dimensional factor structure on the data in order to achieve a reduction of the dimension of the problem and fit different volatility models to the factors extracted by principal component analysis (PCA). The idiosyncratic components are assumed to be independent of each other and their time-varying volatilities are estimated by univariate GARCH and SV models. When estimating simple univariate GARCH or SV models to the factors this is very similar to the O-GARCH model of Alexander (2000), but we also consider multivariate GARCH and SV models to model the volatility of the factors jointly. Namely, we estimate DCC and BEKK GARCH models, and SV models with conditional correlations being described by the Patton and SCAR copula specification. For the last two cases conditional correlations can only be estimated for the case of two factors. Note that although

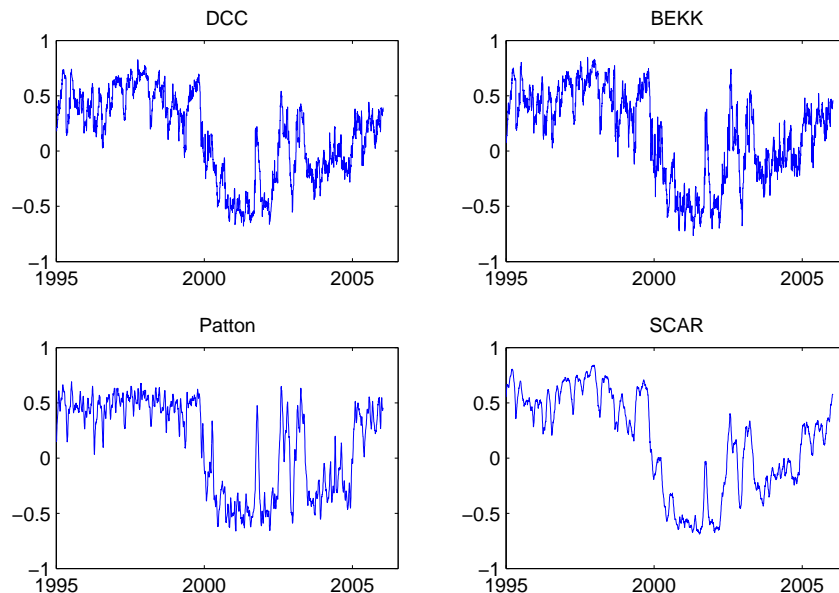


Fig. 4. Conditional correlations between the first two factors

the correlations between the factors extracted by PCA are unconditionally zero, conditional correlations may be different from zero and vary over time.

For the factor specification the covariance matrix for the full set of assets can be calculated using equation (3) in Section 3. For the number of factors we restrict our attention to a maximum of four factors. When estimating SV model the efficient importance sampler by [84] is used for estimation and for the time-varying volatility we consider the smoothed variance, i.e. an estimate of the volatility using the complete sample information. The volatilities of the first two factors estimated by GARCH and SV are shown in Figure 5, whereas the conditional correlation using the four competing bivariate models can be found in Figure 5. The correlation dynamics show that the factors are only unconditionally orthogonal, but show a strong variation over time and extremely high persistence ($\beta = 0.99$ for the SCAR model). It is also notable that the four models for produce estimates of the conditional correlation that very similar.

The results comparing the in-sample ability to compute the minimum variance portfolio (MVP) of the competing models can be found in Table 1. For comparison we also include the variance of the equally weighted portfolio to see how much can be gained by optimizing the portfolio. All models yield clear improvements over using the equally weighted portfolio. Furthermore, the ranking of the models is the same whether looking at the variance of the

Table 1. In-sample fit of competing volatility models

Model		σ_{MVP}	$\sigma_{MVP-std}$	Model		σ_{MVP}	$\sigma_{MVP-std}$
Equally weighted		0.821					
1 factor	GARCH	0.307	2.369	SV	0.143	1.341	
2 factors	O-GARCH	0.295	2.004	O-SV	0.135	1.100	
	DCC	0.292	1.965	SCAR	0.131	1.073	
	BEKK	0.292	1.959	Patton	0.133	1.081	
3 factors	O-GARCH	0.298	2.047	O-SV	0.137	1.129	
	DCC	0.297	2.006				
	BEKK	0.296	1.994				
4 factors	O-GARCH	0.302	2.041	O-SV	0.140	1.128	
	DCC	0.301	2.003				
	BEKK	0.299	1.990				

Note: Variance of the minimum variance portfolios (σ_{MVP}) and the standardized portfolio ($\sigma_{MVP-std}$) constructed using competing factor based multivariate volatility models for 69 stocks from the FTSE 100 during the period 1995 to 2005.

MVP, σ_{MVP} , or the variance of the standardized MVP, $\sigma_{MVP-std}$. The choice of the number of factors does not matter as much as one might expect. Still, two factor models give the best results and seem to be sufficient to estimate the volatility of the data set. Allowing for non-zero conditional correlations between the factors slightly improves the quality of the covariance matrix of the stock returns. Finally, the smoothed volatilities of the SV models seem to provide much better estimates of the covariance than volatilities estimated with GARCH models. This is not surprising, as the SV volatilities are estimated using the full information in the data, whereas the GARCH volatilities are based on 1-step ahead forecasts. Hence, the 2-Factor SV model with correlations estimated using a SCAR specification provides the best fit for our data set based on the the economic criteria we have chosen. Nevertheless one has to keep in mind that the analysis we have done is entirely an in-sample comparison. Out-of-sample the models may be ranked quite differently. In particular, when considering 1-step ahead forecasts the GARCH model is likely to perform quite well (in particular better than SV), due to the way it is designed. When considering multi-step ahead forecasts it is unclear which model will do better and this issue is worth investigating.

6 Outlook

In this chapter we have reviewed new developments in the dynamic modelling of financial asset returns. We have concentrated on the multivariate aspect, since the typical practical application concerns not only volatilities but also an adequate modelling of asset dependencies. We have paid attention to the use of factor structures in order to achieve some dimension reduction when modelling a large number of assets. Such factor structures combined with appropriate volatility models seem to provide a good fit to the data we examined and not too much information is lost when computing the minimum variance portfolio, compared to modelling the full set of assets directly.

In future research the choice of the number of factors, a problem that has been discussed extensively in a theoretical way and for macroeconomic applications, needs to be analyzed concerning the model performance using economic criteria such as the construction of the MVP. Also the use of the class of locally stationary factor models by [91] and [45] for financial applications needs to be considered. Furthermore, the modelling of vast dimensional data (i.e. over 100 assets) needs to be studied. Although some progress has been made for GARCH models, stochastic volatility models that are usable for such dimensions, and estimation techniques for them, need to be developed. Finally, time-varying copula models need to be extended to allow for dimensions larger than two in order to be relevant for realistic applications.

References

1. Alexander, C.: Orthogonal GARCH. In: *Mastering Risk*, Financial Times, vol 2. Prentice Hall, 21–38 (2001)
2. Andersen, T.G., Bollerslev, T.: Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, **39**, 885–905 (1998)
3. Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebends, H.: The Distribution of Realized Stock Volatility. *Journal of Financial Economics*, **61**, 43–76 (2001)
4. Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P.: The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association*, **96**, 42–55 (2001)
5. Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P.: Modeling and Forecasting Realized Volatility. *Econometrica*, **71**(2), 579–625 (2003)
6. Andersen, T.G., Bollerslev, T., Lange, S.: Forecasting Financial Market Volatility: Sample Frequency vis--vis Forecast Horizon. *Journal of Financial Economics*, **61**, 43–76 (1999)
7. Andersen, T.G., Bollerslev, T., Diebold, F.X., Wu, J.: Realized beta: persistence and predictability. In: Fomby, T., Terrell, D. (ed) *Advance in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger*, vol. B, 1–40 (2006)

8. Andersen, T.G., Shephard, N.: Stochastic volatility: origins and overview. In: *Handbook of Financial Time Series*, Springer, Berlin Heidelberg New York (forthcoming)
9. Anderson, H.M., Vahid, F.: Forecasting the Volatility of Australian Stock Returns: Do Common Factors Help? *Journal of Business and Economic Statistics*, **25**(1), 76–90 (2007)
10. Andrews, D.W.K.: Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica*, **61**, 821–856 (1993)
11. Asai, M., McAleer, M.: The structure of dynamic correlations in multivariate stochastic volatility models. *Journal of Econometrics*, forthcoming
12. Asai, M., McAleer, M., Yu, J.: Multivariate stochastic volatility: a review. *Econometric Reviews*, **25**(2-3), 145–175 (2006)
13. Bai, J.: Inferential theory for factor models of large dimensions. *Econometrica*, **71**, 135–171 (2003)
14. Bai, J., Ng, S.: Determining the number of factors in approximate factor models. *Econometrica*, **70**(1), 191–221 (2002)
15. Barndorff-Nielsen, O.E., Shephard, N.: Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society, Ser. B*, **64**, 253–280 (2002)
16. Barndorff-Nielsen, O.E., Shephard, N.: Estimating Quadratic Variation Using Realized Variance. *Journal of Applied Econometrics*, **17**, 457–477 (2002)
17. Barndorff-Nielsen, O.E., Shephard, N.: Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica*, **72**, 885–925 (2004)
18. Barndorff-Nielsen, O.E., Shephard, N.: Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, **2**, 1–37 (2004)
19. Bauwens, L., Laurent, S., Rombouts, J.: Multivariate garch models: a survey. *Journal of Applied Econometrics*, **7**, 79–109 (2006)
20. Bera, A., Higgins, M.: A survey of ARCH models: properties, estimation and testing. *Journal of Economic Surveys*, **7**, 305–366 (1993)
21. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327 (1986)
22. Bollerslev, T.: Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economics and Statistics*, **72**, 498–505 (1990)
23. Bollerslev, T., Wooldridge, J.M.: Quasi maximum likelihood estimation of dynamic models with time-varying covariances. *Econometric Reviews*, **11**, 143–172 (1992)
24. Breitung, J., Eickmeier, S.: Dynamic factor models. In: In O. Hübler and J. Frohn (ed) *Modern Econometric Analysis*. Springer, Berlin Heidelberg New York (2006)
25. Breymann, W., Dias, A., Embrechts, P.: Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, **3**(1), 1–14 (2003)
26. Brillinger, D.R.: *Time Series: Data Analysis and Theory*. Holt, Rinehart and Winston, New York (1981)
27. Broto, C., Ruiz, E.: Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys*, **18**(5), 613–649 (2004)

28. Campbell, J.Y., Lo, A.W., MacKinlay, A.C.: *The Econometrics of financial Markets*. Princeton University Press, Princeton (1997)
29. Candelon, B., Manner, H.: Testing for Asset Market Linkages: A new Approach based on Time-Varying Copulas. METEOR Research Memorandum, RM/07/052, Maastricht University (2007)
30. Carnero, M.A., Peña, D., Ruiz, E.: Persistence and Kurtosis in GARCH and Stochastic Volatility Models. *Journal of Financial Econometrics*, **2**, 319–342 (2004)
31. Chamberlain, G., Rothschild, M.: Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, **51**, 1281–1304 (1983)
32. Chan, L.K.C., Karceski, J., Lakonishok, J.: On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *The Review of Financial Studies*, **12(5)**, 937–974 (1999)
33. Chen, X., Fan, Y.: Estimation of Copula-Based Semiparametric Time Series Models. *Journal of Econometrics*, **135**, 307–335 (2006)
34. Chen, X., Fan, Y.: Estimation and Model Selection of Semiparametric Copula-based Multivariate Dynamic Models under Copula Misspecification. *Journal of Econometrics*, **135**, 125–154 (2006)
35. Cherubini G., Luciano, E., Vecchiato, W.: *Copula Methods in Finance*. John Wiley and Sons Ltd (2004)
36. Clark, P.K.: A subordinate stochastic process model with finite variance for speculative prices. *Econometrica*, **41**, 135–155 (1973)
37. Connor, G., Korajczyk, R.A.: A Test for the Number of Factors in an Approximate Factor Model. *Journal of Finance*, **48(4)**, 1263–1291 (1993)
38. Danielsson, J.: Stochastic volatility in asset prices estimation with simulated maximum likelihood. *Journal of Econometrics*, **64(1-2)**, 375–400 (1994)
39. Danielsson, J.: Multivariate stochastic volatility models: Estimation and comparison with VGARCH models. *Journal of Empirical Finance*, **5**, 155–173 (1998)
40. Danielsson, J., Richard, J.F.: Accelerated Gaussian importance sampler with application to dynamic latent variable models. *Journal of Applied Econometrics*, **8**, 153–174 (1993)
41. Dias, A., Embrechts, P.: Change-Point Analysis for Dependence Structures in Finance and Insurance. In: Szegoe, G. (ed) *Risk Measures for the 21st Century*, Wiley Finance Series, Chapter 16, 321–335 (2004)
42. Diebold, F.X., Nerlove, M.: The dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model. *Journal of Applied Econometrics*, **4**, 1–21 (1989)
43. Diebold, F.X., Gunther, T.A., Tay: Evaluating Density forecasts. *International Economic Review*, **39**, 863–883 (1998)
44. Duffie, D., Singleton, K.J.: Simulated moments estimation of markov models of asset prices. *Econometrica*, **61**, 929–952 (1993)
45. Eichler, M., Motta, G., von Sachs, R.: Fitting Dynamic Factor Models to Non-stationary Time Series. Discussion Paper 0829, Institute of Statistics, UCL (2008)
46. Embrechts, P., McNeil, A., Straumann, D.: Correlation and Dependence Properties in Risk Management: Properties and Pitfalls. In: Dempster, M.

- (ed) Risk Management: Value at Risk and Beyond, Cambridge University Press (2002)
47. Embrechts, P., Lindskog, F., McNeil, A.: Modelling Dependence with Copulas and Applications to Risk Management. In: Handbook of Heavy Tailed Distributions in Finance, Elsevier, Chapter 8, 329–384 (2003)
 48. Engle, R.F.: Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, **50**, 987–1008 (1982)
 49. Engle, R.F.: Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics*, **20(3)**, 339–350 (2002)
 50. Engle, R.F.: High Dimension Dynamic Correlations. Prepared for a Festschrift for David Hendry, Oxford University (2007)
 51. Engle, R.F., Kroner, K.F.: Multivariate simultaneous generalized ARCH. *Econometric Theory*, **11**: 122–150 (1995)
 52. Engle, R.F., Lilien, D.M., Robins, R.P.: Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica*, **55**, 391–407 (1987)
 53. Engle, R.F., Manganelli, S.: CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business and Economic Statistics*, **22(4)**, 367–381 (2004)
 54. Engle, R.F., Ng, V.K., Rothschild, M.: Asset Pricing With a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills. *Journal of Econometrics*, **45**, 213–238 (1990)
 55. Engle, R.F., Sheppard, K.: Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. NBER Working Paper 8554, National Bureau of Economic Research (2001)
 56. Engle, R.F., Shephard, N., Sheppard, K.: Fitting and testing vast dimensional time-varying covariance models. NYU Working Paper No. FIN-07-046 (2007)
 57. Fama, E.F., French, K.R.: Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, **33**, 3–56 (1993)
 58. Fama, E.F.: The behavior of stock market prices. *Journal of Business*, **38**, 34–105 (1965)
 59. Fleming, J., Kirby, C., Ostdiek, B.: The Economic Value of Volatility Timing. *The Journal of Finance*, **56(1)**, 329–352 (2001)
 60. Forni, M., Hallin, M., Lippi, M., Reichlin, L.: The generalized dynamic factor model: Identification and estimation.
 61. Franke, J., Härdle, W., Hafner, C.M.: *Statistics of Financial Markets An Introduction*. Springer (2008)
 62. Gallant, A.R., Tauchen, G.: Which moments to match. *Econometric Theory*, **12**, 657–681 (1996)
 63. Genest, C., Rivest, L.-P.: Statistical Inference Procedures for Bivariate Archimedean Copulas. *Journal of the American Statistical Association*, **88**, 1034–1043 (1993)
 64. Genest, C., Ghoudi, K., Rivest, L.-P.: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, **82**, 543–552 (1995)

65. Genest, C., Goorbergh, R.W.J., Werker, B.J.M.: Multivariate Options Pricing Using Dynamic Copula Models. *Insurance: Mathematics and Economics*, **37**, 101–114 (2005)
66. Ghysels, E., Harvey, A.C., Renault, E.: Stochastic volatility. In: Maddala, G., Rao, C.R. (ed) *Handbook of Statistics*, vol. 14. Elsevier Sciences, Amsterdam (1996)
67. Giacomini, E., Härdle, W., Spokoiny, V.: Inhomogeneous Dependency Modelling with Time Varying Copulae. *Journal of Business and Economic Statistics*, forthcoming (2009)
68. Gouriéroux, C., Monfort, A., Renault, E.: Indirect Inference. *Journal of Applied Econometrics*, **8**, 85–118 (1993)
69. Granger, C.W.J., Teräsvirta, T., Patton, A.: Common Factors in Conditional Distributions. *Journal of Econometrics*, **132**(1), 43–57 (2006)
70. Hafner, C.M., Herwartz, H.: Testing linear autoregressive dynamics under heteroskedasticity. *The Econometrics Journal*, **3**, 177–197 (2000)
71. Hafner, C.M., Franses, P.H.: A generalized dynamic conditional correlation model: Simulation and application to many assets. *Econometric Reviews*, forthcoming.
72. Hafner, C.M., Manner, H.: Dynamic stochastic copula models: Estimation, inference and applications. METEOR Research Memorandum, RM/08/043, Maastricht University (2008)
73. Hafner, C.M., Reznikova, O.: Efficient estimation of a semiparametric dynamic copula model. Manuscript UCL (2008)
74. Harvey, A.C., Ruiz, E., Shephard, N.: Multivariate stochastic variance models. *Review of Economic Studies*, **61**, 247–264 (1994)
75. Hentschel, L.: All in the family: Nesting symmetric and asymmetric garch models. *Journal of Financial Economics*, **39**, 71–104 (1995)
76. Hull, J., White, A.: The pricing of options with stochastic volatilities. *Journal of Finance*, **42**, 281–300 (1987)
77. Jacquier, E., Polson, N.G., Rossi, P.E.: Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics*, **12**, 371–389 (1994)
78. Joe, H.: *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC (1997)
79. Junker, M., May, A.: Measurement of Aggregate Risk with Copulas. *Econometrics Journal*, **8**, 428–454 (2005)
80. Junker, M., Szimayer, A., Wagner, N.: Nonlinear term structure dependence: Copula functions, empirics, and risk implications. *Journal of Banking & Finance*, **30**, 1171–1199 (2006)
81. Kim, S., Shephard, N., Chib, S.: Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies*, **65**, 361–393 (1998)
82. King, M., Sentana, E., Wadhvani, S.: Volatility and Links Between National Stock Markets. *Econometrica*, **62**, 901–934 (1994)
83. Ledoit, O, Santa-Clara, P, Wolf, M.: Flexible multivariate GARCH modeling with an application to international stock markets. *Review of Economics and Statistics*, **85**, 735–747 (2003)
84. Liesenfeld, R., Richard, J.F.: Univariate and multivariate volatility models: Estimation and diagnostics. *Journal of Empirical Finance*, **10**, 505–531 (2003)

85. Liesenfeld, R., Richard, J.F.: Classical and Bayesian Analysis of Univariate and Multivariate Stochastic Volatility Models. *Econometric Reviews*, **25(2-3)**, 335–360 (2006)
86. Lintner, J.: Security prices, risk and maximal gains from diversification. *Journal of Finance*, **20**, 587–615 (1965)
87. Longin, F., Solnik, B.: Is the Correlation in International Equity Returns Constant: 1960-1990. *Journal of International Money and Finance*, **14**, 3–26 (1995)
88. Longin, F., Solnik, B.: Extreme Correlation of International Equity Markets. *Journal of Finance*, **56**, 649–676 (2001)
89. Mandelbrot, B.: The variation of certain speculative prices. *Journal of Business*, **36**, 394–419 (1963)
90. Markowitz, H.: *Portfolio Selection: efficient diversification of investments*. John Wiley, New York (1959)
91. Motta, G., Hafner, C., von Sachs, R.: Locally stationary factor models: Identification and nonparametric estimation. *Econometric Theory* (forthcoming)
92. Nelsen, R.B.: *An Introduction to Copulas*. Springer (2006)
93. Patton, A.: On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation. *Journal of Financial Econometrics* **2**, 130–168 (2004)
94. Patton, A.: Modelling Asymmetric Exchange Rate Dependence. *International Economic Review*, **47(2)**, 527–556 (2006)
95. Patton, A.: Estimation of Multivariate Models for Time Series of Possibly Different Lengths. *Journal of Applied Econometrics*, **21**, 147–173 (2006b)
96. Ramchand L., Susmel, R.: Volatility and Cross Correlation across Major Stock Markets. *Journal of Empirical Finance* **5**, 397–416 (1998)
97. Richard, J.F., Zhang, W.: Efficient high-dimensional importance sampling. *Journal of Econometrics*, **141(2)**, 1385–1411 (2007)
98. Ross, S.A.: The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, **13**, 341–360 (1976)
99. Sandmann, G., Koopman, S.J.: Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics*, **87(2)**, 271–301 (1998)
100. Sharpe, W.F.: Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, **19**, 425–442 (1964)
101. Stock, J.H., Watson, M.W.: Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, **97**, 1167–1179 (2002)
102. Taylor, S.J.: *Modelling Financial Time Series*. JohnWiley and Sons, Chichester (1986)
103. Taylor, S.J.: Modelling stochastic volatility: a review and comparative study. *Mathematical finance*, **4**, 183–204, (1994)
104. Tse, Y.K., Tsui, A.K.C.: A multivariate GARCH model with time-varying correlations. *Journal of Business and Economic Statistics*, **20(3)**, 351–362 (2002)
105. van der Weide, R.: Go-garch: A multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics*, **17**, 549–564, (2002)

106. Yu, J., Meyer, R.: Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews*, **25(2-3)**, 361–384 (2006)