

# Evaluating Portfolio Value-at-Risk using Semi-Parametric GARCH Models

Jeroen V.K. Rombouts<sup>1</sup>      Marno Verbeek<sup>2</sup>

December 8, 2004

## Abstract

In this paper we examine the usefulness of multivariate semi-parametric GARCH models for portfolio selection under a Value-at-Risk (VaR) constraint. First, we specify and estimate several alternative multivariate GARCH models for daily returns on the S&P 500 and Nasdaq indexes. Examining the within sample VaRs of a set of given portfolios shows that the semi-parametric model performs uniformly well, while parametric models in several cases have unacceptable failure rates. Interestingly, distributional assumptions appear to have a much larger impact on the performance of the VaR estimates than the particular parametric specification chosen for the GARCH equations. Finally, we examine the economic value of the multivariate GARCH models by determining optimal portfolios based on maximizing expected returns subject to a VaR constraint, over a period of 500 consecutive days. Again, the superiority and robustness of the semi-parametric model is confirmed.

Keywords: multivariate GARCH, semi-parametric estimation, Value-at-Risk, asset allocation

---

<sup>1</sup>HEC Montreal, 3000 chemin de la Cte-Sainte-Catherine, H3T 2A7, Montreal, Canada,  
jeroen.rombouts@hec.ca

<sup>2</sup>Rotterdam School of Management and Econometric Institute, Erasmus University Rotterdam, P.O.B.  
1738, 3000 DR Rotterdam, Netherlands, M.Verbeek@fbk.eur.nl.

# 1 Introduction

Models that forecast returns and volatility play an important role in financial decision-making. Recently, several studies investigate the significance of forecasting volatility for economic agents. For example, Fleming, Kirby, and Ostdiek (2001) examine the economic value of volatility timing, using a class of rolling estimators for the covariance matrix of returns. They forecast the covariance matrix one-day ahead and construct optimal portfolio weights, based on a mean-variance approach. Marquering and Verbeek (2004), using simple recursive linear models, analyze the economic value of volatility timing, jointly with return timing, at a monthly frequency. De Goeij and Marquering (2004) examine the economic value of forecasts from a multivariate asymmetric (parametric) GARCH model. These studies limit themselves to the mean-variance framework, where optimal portfolios are determined based on a trade off between the expected return on a portfolio and its variance.

The focus on variance as the relevant risk measure is appropriate if returns are (conditionally) normally distributed or if it can be assumed that investors only care about the first two moments of the distribution of their portfolio return. This assumption is restrictive, because it states that investors do not attach particular weight to skewness and kurtosis or to specific quantiles of the distribution. This is stressed by several recent papers. Harvey and Siddique (2000), for example, argue that conditional skewness is an important factor explaining the cross-section of expected returns, while Barone Adesi, Gagliardini, and Urga (2004) investigate the role of co-skewness for testing asset pricing models. In this paper we analyze optimal portfolio choice focussing explicitly on downside risk. In particular, we investigate the economic value of multivariate volatility models when optimal portfolios are constructed under a Value-at-Risk (VaR) constraint. Value-at-Risk defines the maximum expected loss on an investment over a specified horizon at a given confidence level, and is used by many banks and financial institutions as a key measure for market risk (see Jorion (2000) for an extensive introduction to VaR methodology).

Several recent papers analyze the risk-return trade off from a VaR perspective, for ex-

ample Duffie and Pan (1997), Lucas and Klaassen (1998), Gouriéroux, Laurent, and Scaillet (2000), Campbell, Huisman, and Koedijk (2001) and Alexander and Baptista (2002). These studies require an appropriate model for the tail behavior of the return distributions and their interdependence. Typically, the simultaneous distribution of the innovations in these models is assumed to be multivariate normal or Student  $t$ . This is restrictive, for example in the presence of nonzero third moments or in cases where the tail behavior is different across the portfolio components.

In this paper we investigate the implications on conditional Value-at-Risk (VaR) calculations for a portfolio when returns are described by a multivariate GARCH model, with an unrestricted distribution for the innovations. We specify and estimate several semi-parametric multivariate GARCH models for the returns on the S&P 500 and Nasdaq indexes, and compare their implied Value-at-Risk of several portfolios with those obtained from some parametric counterparts. Further, over a period of 500 days we determine the optimal dynamic trading rule that maximizes the expected return over the next trading day, subject to the constraint that the expected loss, with probability 1% or 5%, does not exceed a given VaR level. This allows us to compare dynamic portfolios based on different model specifications, for example by determining how frequently the VaR boundary is violated within the 500 day period. Because Value-at-Risk depends upon the joint tail behavior of the conditional distribution of asset returns, we expect that the parametric specifications only perform well in particular cases and settings, while the semi-parametric approach is expected to be robust against distributional misspecifications. Given the empirical evidence of asymmetries and – most importantly – excess kurtosis in the (conditional) distribution of stock returns, this is a potentially important advantage.

In the empirical section, we estimate a semi-parametric multivariate GARCH model for the returns on two broad stock market indexes (S&P 500 and Nasdaq). The GARCH parameters are estimated without making restrictive assumptions about the distributions of the innovations, while the latter are estimated non-parametrically using a technique proposed by Hafner and Rombouts (2004). This way we obtain a model where we specify the first two conditional moments of the returns jointly in a parametric way while the rest of

the return distribution is determined non-parametrically. The advantage of the multivariate approach is that the Value-at-Risk of any portfolio of assets can be determined from the GARCH estimates and the corresponding non-parametric estimate of the multivariate distribution of the innovations. Most importantly, this allows us to employ the framework of Campbell, Huisman, and Koedijk (2001) to determine optimal portfolio weights taking into account a Value-at-Risk constraint. In contrast, many existing approaches, including the regime-switching model of Billio and Pelizzon (2000) and the semi-parametric approach of Fan and Gu (2003), only allow one to determine the Value-at-Risk of a given asset or portfolio of assets.

The rest of this paper is organized as follows. Section 2 describes a number of alternative multivariate GARCH (MGARCH) specifications, and explains how the VaR of a portfolio can be calculated on the basis of these models. Section 3 describes the data and reports the estimation results for the MGARCH models for the S&P 500 and Nasdaq indexes over the period January 1988 – August 2001. Section 4 focuses on the VaR calculations and summarizes the results, by means of failure rates, for the different MGARCH models. Section 5 describes how optimal portfolio weights can be determined and analyzes the performance of the different models for VaR calculations. Finally, Section 6 concludes.

## 2 Multivariate GARCH models

In this section, we describe several alternative semi-parametric multivariate GARCH models and link them to the conditional Value-at-Risk of a portfolio constructed from the different asset categories. The GARCH model describes the conditional distribution of a vector of returns, from which quantiles of the distribution of portfolio returns can be derived. Let  $r_t$  denote the  $N$ -dimensional vector of stationary returns. The model can be written as follows

$$r_t = \mu_t(\theta) + H_t^{1/2}(\theta)\xi_t \quad t = 1, \dots, T, \quad (1)$$

where  $\mu_t(\theta)$  is an  $N$ -dimensional vector of conditional mean returns,  $\xi_t$  is an *i.i.d.* vector white noise process with identity covariance matrix and density  $g(\cdot)$ , and the symmetric  $N \times N$  matrix  $H_t(\theta)$  denotes the conditional covariance matrix of  $r_t$ . Unknown parameters are collected in the vector  $\theta$ . Both the mean and covariance matrix are conditional upon the information set  $I_{t-1}$ , containing at least the entire history of  $r_t$  until  $t - 1$ . Returns are thus assumed to be generated by a parameterized time varying location scale model. The conditional density of  $r_t$  is given by

$$f_{r_t|I_{t-1}}(r) = |H_t(\theta)|^{-1/2} g\left(H_t^{-1/2}(\theta)(r - \mu_t(\theta))\right). \quad (2)$$

The expression in (2) shows how the conditional distribution of the returns varies over time and allows one to estimate the time-varying quantiles of the return vector by replacing the parameter vector  $\theta$  and the unknown density  $g(\cdot)$  by their estimates,  $\hat{\theta}$  and  $\hat{g}(\cdot)$ , respectively.

We consider three multivariate GARCH models, that specify different functional forms for the conditional covariance matrix  $H_t(\theta)$  and how it depends upon the information set  $I_{t-1}$ . Further, we combine these specifications with alternative assumptions about the density  $g(\cdot)$  of  $\xi_t$ . The MGARCH models we consider are the diagonal VEC (DVEC) model and the dynamic conditional correlation (DCC) models of Tse and Tsui (2002) and Engle (2002). The specific assumptions of these three models are given in Definitions 1 to 3 below. We consider these alternative MGARCH models to make sure that our results are not specific to one particular, perhaps inappropriate, specification. Moreover, it allows us to analyze how sensitive the VaR estimates are with respect to the choice of the multivariate GARCH model.

**Definition 1** *The DVEC(1,1) model is defined as:*

$$h_t = c + A \eta_{t-1} + G h_{t-1}, \quad (3)$$

where

$$h_t = \text{vech}(H_t) \quad (4)$$

$$\eta_t = \text{vech}(\epsilon_t \epsilon_t'), \quad (5)$$

and  $\text{vech}(\cdot)$  denotes the operator that stacks the lower triangular portion of a  $N \times N$  matrix as a  $N(N+1)/2 \times 1$  vector.  $A$  and  $G$  are diagonal parameter matrices of order  $(N+1)N/2$  and  $c$  is a  $(N+1)N/2 \times 1$  parameter vector.

**Definition 2** The DCC model of Tse and Tsui (2002) or  $DCC_T(M)$  is defined as:

$$H_t = D_t R_t D_t, \quad (6)$$

where  $D_t = \text{diag}(h_{11t}^{1/2} \dots h_{NNt}^{1/2})$ ,  $h_{iit}$  can be defined as any univariate GARCH model, and

$$R_t = (1 - \theta_1 - \theta_2)R + \theta_1 \Psi_{t-1} + \theta_2 R_{t-1}. \quad (7)$$

In (7),  $\theta_1$  and  $\theta_2$  are non-negative parameters satisfying  $\theta_1 + \theta_2 < 1$ ,  $R$  is a symmetric  $N \times N$  positive definite correlation matrix with diagonal elements  $\rho_{ii} = 1$ , and  $\Psi_{t-1}$  is the  $N \times N$  sample correlation matrix of  $\epsilon_\tau$  for  $\tau = t - M, t - M + 1, \dots, t - 1$ . Its  $i, j$ -th element is given by:

$$\psi_{ij,t-1} = \frac{\sum_{m=1}^M u_{i,t-m} u_{j,t-m}}{\sqrt{(\sum_{m=1}^M u_{i,t-m}^2)(\sum_{h=1}^M u_{j,t-h}^2)}}, \quad (8)$$

where  $u_{it} = \epsilon_{it} / \sqrt{h_{iit}}$ . The matrix  $\Psi_{t-1}$  can be expressed as:

$$\Psi_{t-1} = B_{t-1}^{-1} L_{t-1} L_{t-1}' B_{t-1}^{-1}, \quad (9)$$

with  $B_{t-1}$  a  $N \times N$  diagonal matrix with  $i$ -th diagonal element being  $(\sum_{h=1}^M u_{i,t-h}^2)^{1/2}$  and  $L_{t-1} = (u_{t-1}, \dots, u_{t-M})$ , an  $N \times M$  matrix.

**Definition 3** The DCC model of Engle (2002) or  $DCC_E(S, L)$  is defined as:

$$H_t = D_t R_t D_t \quad (10)$$

where  $D_t = \text{diag}(h_{11t}^{1/2} \dots h_{NNt}^{1/2})$ ,  $h_{iit}$  can be defined as any univariate GARCH model, and

$$R_t = (\text{diag } Q_t)^{-1/2} Q_t (\text{diag } Q_t)^{-1/2}. \quad (11)$$

where the  $N \times N$  symmetric positive definite matrix  $Q_t$  is given by:

$$Q_t = (1 - \alpha - \beta) \bar{Q} + \alpha u_{t-1} u'_{t-1} + \beta Q_{t-1}, \quad (12)$$

where  $u_{it} = \epsilon_{it} / \sqrt{h_{iit}}$ ,  $\bar{Q}$  is the  $N \times N$  unconditional variance matrix of  $u_t$ , and  $\alpha (\geq 0)$  and  $\beta (\geq 0)$  are scalar parameters satisfying  $\alpha + \beta < 1$ .

From the joint return distribution we can calculate the quantiles of the marginal distributions  $r_{it} | I_{t-1}$ ,  $i = 1, \dots, N$ . These marginal densities are given by

$$f_{r_{it}|I_{t-1}}(r_i) = \int_{R^{N-1}} f_{r_t|I_{t-1}}(r_i, \bar{r}_{-i}) d\bar{r}_{-i}, \quad (13)$$

where  $\bar{r}_{-i}$  indicates everything in  $r$  except  $r_i$ . The main interest, however, lies in the distribution of a linear combination of the vector of returns,  $w'_t r_t$  or a portfolio, which depends upon the salient dependencies between the different returns. Because  $g(\cdot)$  is left unspecified the distribution of a linear combination of  $r_t$  can be calculated by the following well known result. Consider a random vector  $(X, Y) \sim f_{X,Y}$  and  $(U, V) = (R(X, Y), S(X, Y))$  a new random vector as a function of the previous vector. Suppose that  $R$  and  $S$  are functions such that we can calculate  $(X, Y) = (L(U, V), T(U, V))$ . Then we have

$$f_{U,V}(u, v) = |\det J| \cdot f_{X,Y}(L(u, v), T(u, v)) \quad (14)$$

where

$$J = \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{pmatrix}. \quad (15)$$

We are interested in a single linear combination, corresponding to an asset portfolio, so we can take  $Y = V = T(U, V)$  and integrate this part out of the multivariate density. In the

bivariate case, for example, the density at time  $t$  of the return on a portfolio with weights  $w_{1t} \neq 0$  and  $w_{2t} = 1 - w_{1t}$  is

$$f_{w_{1t}r_{1t}+w_{2t}r_{2t}|I_{t-1}}(r_p) = \frac{1}{w_{1t}} |H_t(\theta)|^{-1/2} \int g \left( H_t^{-1/2}(\theta) \left[ \begin{pmatrix} \frac{r_p - w_{2t}v}{w_{1t}} \\ v \end{pmatrix} - \mu_t(\theta) \right] \right) dv. \quad (16)$$

Because in our framework  $g(\cdot)$  is unknown, numerical integration techniques will be used to obtain the distribution of the portfolio return. See for example Bauwens, Lubrano, and Richard (1999) for details on numerical integration. At a given confidence level  $1 - \alpha$ , the Value-at-Risk (VaR) of a portfolio with weights  $w_t$  is defined as follows.

**Definition 4** *The VaR at level  $\alpha$  is the solution to*

$$P(w'_t r_t < VaR_\alpha) = \alpha \quad (17)$$

or

$$\alpha = \int_{-\infty}^{VaR_\alpha} f_{w'_t r_t | I_{t-1}}(r_p) dr_p. \quad (18)$$

The VaR is a measure of the market risk of the portfolio and measures the loss that it could generate (over a given time horizon) with a given degree of confidence. Above, we have expressed the VaR in relative terms as the quantile at level  $\alpha$  of the distribution of portfolio returns. With probability  $1 - \alpha$ , the losses on the portfolio will be smaller than  $VaR_\alpha$ . The VaR is widely adopted by banks and financial institutions to measure and manage market risk, as it reflects downside risk of a given portfolio or investment. In general, the VaR is a function of the confidence level  $\alpha$ , the density  $g(\cdot)$ , the portfolio weights  $w_t$ , the functional form of the mean vector  $\mu_t$  and of the covariance matrix  $H_t$ , where the latter three are time dependent. In the case where  $g(\cdot)$  is the multivariate normal density the definition of the VaR reduces to the well known formula  $VaR_\alpha = w'_t \mu_t + (w'_t H_t w_t)^{1/2} z_\alpha$  where  $z_\alpha$  is the  $\alpha$ -th quantile of the univariate standard normal distribution.



The parameter vector  $\theta$  is estimated by quasi maximum likelihood (QML) which implies that during estimation we suppose that  $g(x) \propto \exp(-\frac{x'x}{2})$ . The relevant part of the loglikelihood function for a sample  $t = 1, \dots, T$  then becomes

$$-\sum_{t=1}^T (\ln |H_t(\theta)| + (y_t - \mu_t(\theta))' H_t^{-1}(\theta) (y_t - \mu_t(\theta))), \quad (19)$$

conditional on some starting value for  $\mu_0$  and  $H_0$ . Equation (19) can be maximized with respect to  $\theta$  using a numerical algorithm, which results in a consistent and asymptotically normally distributed estimator, provided that  $\mu_t(\cdot)$  and  $H_t(\cdot)$  are correctly specified (Bollerslev and Wooldridge (1992)). Alternatively, it is possible to make other parametric assumptions about the distribution  $g(\cdot)$ , for instance the multivariate  $t$ -distribution with arbitrary degrees of freedom  $\nu$ . For more information on the estimation of MGARCH models we refer to Bauwens, Laurent, and Rombouts (2004).

The density  $g(\cdot)$  in (2) is estimated by a kernel density estimator. A general multivariate kernel density estimator with bandwidth matrix  $H$  and multivariate kernel  $\mathcal{K}$  can be written as

$$\hat{g}_H(x) = \frac{1}{T|H|} \sum_{t=1}^T \mathcal{K}(H^{-1}(\xi_t - x)).$$

Since the variance of the innovations should be the same in all directions, it is reasonable to use a scalar bandwidth,  $H = hI_N$ , with  $h > 0$ . It is well known that by requiring  $Th^N \rightarrow \infty$  and  $h \rightarrow 0$  as  $T \rightarrow \infty$ , the multivariate kernel density estimates are consistent and asymptotically normally distributed. The MSE-optimal rate for the bandwidth is  $T^{-1/(4+N)}$  which is a rule of thumb bandwidth proposed by Silverman (1986). Furthermore, we use a product kernel  $\mathcal{K}(x) = \prod_{i=1}^N K(x_i)$  and some univariate kernel function  $K$  such as Gaussian, quartic or Epanechnikov. Our density estimate becomes

$$\hat{g}_h(x) = \frac{1}{Th^N} \sum_{t=1}^T \prod_{i=1}^N K\left(\frac{\xi_{i,t} - x_i}{h}\right).$$

More details on multivariate kernel density estimation can be found in Scott (1992). In our application we will use a Gaussian kernel.

### 3 Data and estimation of the MGARCH models

We consider daily returns on two stock market indexes, namely the Standard & Poor’s 500 (S&P 500) index and the Nasdaq index. We have data from 04/01/1988 to 31/08/2001, which results in 3567 daily observations. Both daily log-prices and returns are plotted in Figure 1 and descriptive statistics are given in Table 1. There is a clear presence of fat tails in the return distributions. The kurtosis of the S&P 500 index and the Nasdaq index are 8.45 and 11.7, respectively. Even after estimation of a multivariate GARCH model to these data, we may expect that the nonparametrically estimated innovation density still features quite some pronounced departures from normality. The estimated unconditional correlation coefficient is 0.779.

**Table 1: Summary statistics**

	04/01/1988–31/08/2001	
	$T = 3567$	
	Nasdaq	S&P 500
Mean (%)	0.0462	0.0416
Standard Deviation (%)	1.4194	0.9560
Maximum	13.255	4.9887
Minimum	−10.168	−7.1127
Skewness	−0.1021	−0.4006
Kurtosis	11.698	8.4475

Daily Nasdaq and S&P 500 index returns descriptive statistics. The estimated correlation coefficient is 0.779.

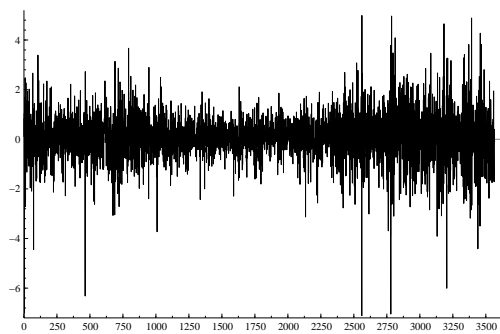
We estimate the DVEC, DCC Tse and DCC Engle models, described in Section 2, over the whole sample period by QML. The parameter estimates and corresponding robust standard errors and  $t$ -statistics for the DVEC model are given in Table 2. Note that we are close to the unit root case because the maximum eigenvalue is equal to 0.997. Generally



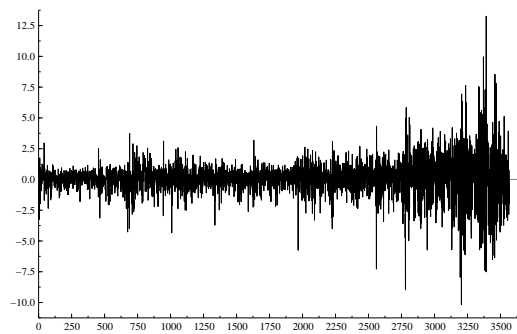
(a) daily Standard & Poor's 500 index log-prices



(b) daily Nasdaq index log-prices



(c) daily Standard & Poor's 500 index returns



(d) daily Nasdaq index returns

Figure 1: Sample period: 04/01/1988–31/08/2001 or 3567 observations. The returns are measured by their log-differences.

speaking, the QML standard errors are small. There is also quite some persistence in the conditional correlation process.

The DCC models are estimated in one step. We first estimate the DCC model of Tse, where we choose GARCH(1,1) models for the conditional variances of both series and we set  $M = 2$  in the correlation specification (8). The estimation results are given in Table 3. As already mentioned for the DVEC model, there is a high persistence in the conditional variance series. The same still holds for the conditional correlation series. We remark that the parameter estimates related to the conditional variances change only marginally between the DVEC and the DCC model. This is not surprising given the identical functional forms for the variances in both models.

The third model is the DCC model of Engle. The elements of  $\bar{Q}$  in (12) are set to their empirical counterparts to render the estimation simpler because there are less parameters to estimate. The estimation results are given in Table 4. Comparing the estimates with the corresponding estimates of the DVEC and the DCC Tse model, we only observe small differences. The correlation model parameter estimates of the DCC models are slightly different. The next two sections investigate whether these small differences have consequences for VaR computations. The estimated conditional correlation series is plotted in Figure 3. While the estimated unconditional correlation coefficient equals 0.779, there is substantial time variation in the conditional correlations. A test for constant conditional correlations, that is  $\theta_1 = \theta_2 = 0$ , would easily be rejected. Because the three MGARCH specifications are non-nested, a direct comparison based on statistical tests is complicated and we shall evaluate them on the basis of their performance regarding the implied Value-at-Risk.

Finally, we estimate the bivariate density  $g(\cdot)$  of the innovations  $\xi_t$ , for each of the three MGARCH specifications, on the basis of the estimated standardized residuals. These may be computed from (1) as  $\hat{H}_t^{-1/2}(r_t - \hat{\mu}_t)$  where the hats indicate that the unknown parameters in  $\theta$  are replaced by their estimates. As mentioned in Section 2 we use a Gaussian product kernel for the nonparametric estimation of the innovation density. The bandwidth is obtained by the conventional rule of thumb and in our application equals 0.26332. The estimated innovation density for the DCC model of Engle is displayed in Figure 2. We do

not show the estimated densities for the DVEC and the DCC Tse model because there are no marked differences. Figure 2 exhibits clear departures from normality. The skewness for the S&P 500 and Nasdaq residuals are  $-0.55279$  and  $-0.42581$ , respectively, while the kurtosis are  $5.4258$  and  $8.5048$ , respectively. These deviations from normality may have an important impact upon the Value-at-Risk that is implied by the distribution of portfolio returns, and suggest that the normal distribution may provide inaccurate VaR estimates. Below we shall determine and evaluate the VaRs on the basis of the semi-parametric distribution, as well as the bivariate normal and  $t$ -distributions.

**Table 2: Parameter estimates for the DVEC model**

	Coefficient	Std error	$t$ -statistic
$\omega_{11}$	0.005134	(0.00185)	2.779
$\omega_{21}$	0.004144	(0.00141)	2.934
$\omega_{22}$	0.005112	(0.00172)	2.979
$\alpha_{11}$	0.049925	(0.00846)	5.904
$\alpha_{22}$	0.044300	(0.00754)	5.873
$\alpha_{33}$	0.042482	(0.00752)	5.652
$\beta_{11}$	0.946960	(0.00869)	108.9
$\beta_{22}$	0.951263	(0.00812)	117.2
$\beta_{33}$	0.952473	(0.00822)	115.8

QML estimates for the DVEC model.  
QML standard errors in the Std error column. Sample of 3567 observations (04/01/1988–31/08/2001).

**Table 3: Parameter estimates for the DCC Tse model**

	Coefficient	Std error	<i>t</i> -statistic
$\omega_{11}$	0.004801	(0.00192)	2.502
$\omega_{22}$	0.003553	(0.00154)	2.309
$\alpha_{11}$	0.049029	(0.00979)	5.009
$\alpha_{22}$	0.034355	(0.00825)	4.163
$\beta_{11}$	0.944110	(0.01082)	87.24
$\beta_{22}$	0.959410	(0.00985)	97.41
$\rho$	0.156868	(0.07895)	1.987
$\theta_1$	0.023196	(0.00414)	5.601
$\theta_2$	0.979040	(0.00400)	244.6

QML estimates for the DCC Tse model.  
QML standard errors in the Std error  
column. Sample of 3567 observations  
(04/01/1988–31/08/2001).

**Table 4: Parameter estimates for the DCC Engle model**

	Coefficient	Std error	<i>t</i> -statistic
$\omega_{11}$	0.005651	(0.00204)	2.775
$\omega_{21}$	0.005045	(0.00185)	2.724
$\alpha_{11}$	0.052459	(0.00906)	5.790
$\alpha_{22}$	0.041059	(0.00807)	5.085
$\beta_{11}$	0.943694	(0.00946)	99.75
$\beta_{22}$	0.953667	(0.00908)	105.1
$\theta_1$	0.034216	(0.00700)	4.888
$\theta_2$	0.953263	(0.01034)	92.17

QML estimates for the DCC Engle model.  
QML standard errors in the Std error  
column. Sample of 3567 observations  
(04/01/1988–31/08/2001).

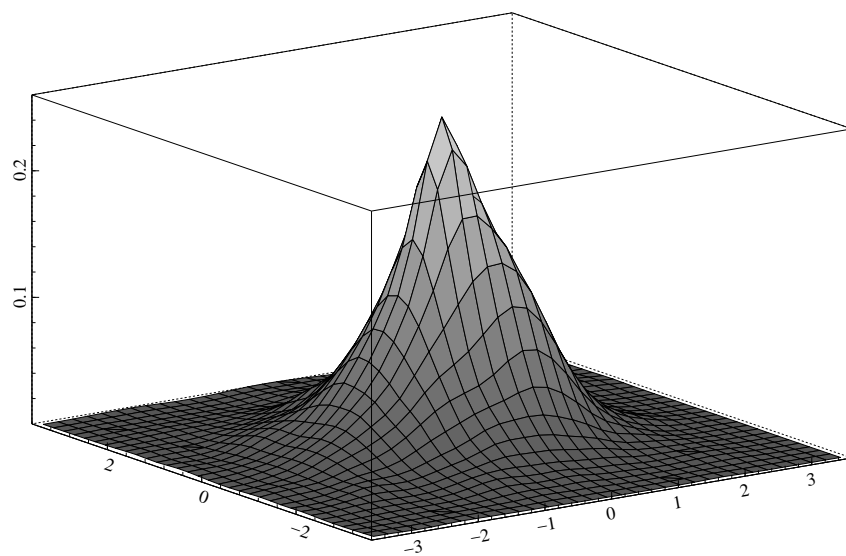


Figure 2: Plot of  $\hat{g}(\cdot)$ , the estimated innovation density of  $\xi_t$  implied by the DCC model of Engle. The skewness for the first and second component are  $-0.55$  and  $-0.43$  respectively and the kurtosis is  $5.43$  and  $8.51$  respectively.



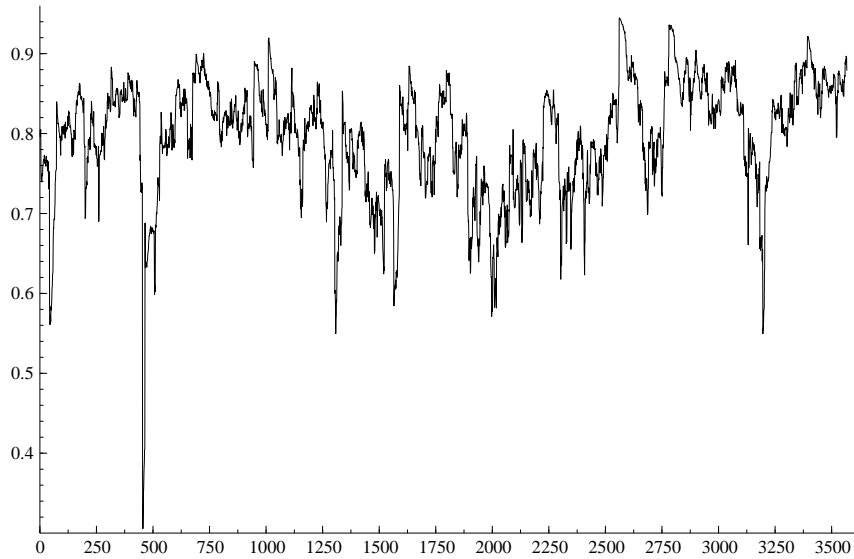


Figure 3: Plot of conditional correlations implied by the DCC model of Engle.

## 4 Value-at-Risk with fixed portfolio weights

This section explores the Value-at-Risk measures corresponding to the models estimated above. We investigate the Value-at-Risk at the 1% and 5% levels, denoted  $\text{VaR}_{0.01}$  and  $\text{VaR}_{0.05}$ , respectively, for the last 2000 trading days of the sample. We consider the three different MGARCH models defined in Section 2 and estimated by QML, and three time invariant portfolios with weights  $w^1 = (0.25, 0.75)'$ ,  $w^2 = (0.5, 0.5)'$  and  $w^3 = (0.75, 0.25)'$ . Furthermore, we distinguish between different innovation densities: the Gaussian, the student  $t$  and the nonparametric density. The degrees of freedom of the bivariate student  $t$  distribution are estimated by maximum likelihood as 6.61.

To compare the VaR levels we calculate failure rates for the different specifications. The failure rate (FR) is defined as the proportion of  $r_t$ 's smaller than the VaR. For a correctly specified model, the empirical failure rate should be close to the specified VaR level  $\alpha$ . We compare the empirical failure rate to its theoretical value by means of the Kupiec likelihood ratio test, see Kupiec (1995). The failures rates and the  $p$ -values for the Kupiec test are

displayed in Tables 5 and 6 for the five and one percent VaR level respectively.

From the failure rates and the  $p$ -values in Table 5, we observe that the normal distribution performs reasonably well for the 5 percent VaR level. The failure rates that are calculated on the basis of the student distribution are too low, while the semi-parametric procedure works reasonably well. Notice that the results for the DVEC and the DCC Engle model are very similar, while the DCC Tse model produces slightly different results. Table 6 displays the results for the 1 percent VaR level. In this case, the normal distribution has a difficult job in providing failure rates close to the VaR level. Overall, the empirical failure rates are too high, which means that we overestimate the first percentile of the distribution of the portfolio return. The Kupiec likelihood ratio test consistently rejects the normal distribution. The student distribution generates failure rates that are consistently too low, although reasonably close to the theoretical values. Contrary to the five percent VaR level case, this suggests that the degrees of freedom are correctly estimated. Again, the semi-parametric procedure works well.

The above results show that the semi-parametric procedure proposed in this paper is a promising tool for risk management analysis. Firstly, the procedure is based on a natural idea. We do not impose a specific functional form on the innovation distribution when we calculate the VaR. Secondly, we do not have to worry which innovation distribution to use for which specific VaR level. The fact that semi-parametric estimation of the VaR dominates parametric approaches is also demonstrated by Fan and Gu (2003) in the univariate case. Obviously, one can always impose other parametric distributional assumptions for the innovations that are more flexible than the normal and student  $t$  distributions. For example, Mittnik and Paoletta (2000) and Giot and Laurent (2003) work with a skewed  $t$  distribution, but there are always chances of severe misspecifications.

Table 5: VaR<sub>0.05</sub> results

		$w^1$		$w^2$		$w^3$	
		FR	$p$ -value	FR	$p$ -value	FR	$p$ -value
DVEC	Normal	0.0585	(0.089)	0.0550	(0.312)	0.0595	(0.058)
	$t_{\hat{\nu}}$	0.0375	(0.007)	0.0365	(0.004)	0.0385	(0.014)
	semi-parametric	0.0555	(0.267)	0.0515	(0.760)	0.0565	(0.191)
DCC <sub>T</sub>	Normal	0.0640	(0.006)	0.0600	(0.046)	0.0680	(0.000)
	$t_{\hat{\nu}}$	0.0455	(0.349)	0.0400	(0.034)	0.0445	(0.250)
	semi-parametric	0.0595	(0.058)	0.0585	(0.089)	0.0610	(0.029)
DCC <sub>E</sub>	Normal	0.0585	(0.089)	0.0545	(0.362)	0.0595	(0.058)
	$t_{\hat{\nu}}$	0.0380	(0.010)	0.0365	(0.004)	0.0390	(0.019)
	semi-parametric	0.0555	(0.267)	0.0515	(0.760)	0.0565	(0.191)

This table presents failure rates (FR) and  $p$ -values for the Kupiec LR test. We report this for the DVEC, the DCC Tse (DCC<sub>T</sub>) and the DCC Engle (DCC<sub>E</sub>) model. We distinguish between the normal, the student and the nonparametric innovation density, for three portfolio weights.

Table 6: VaR<sub>0.01</sub> results

		$w^1$		$w^2$		$w^3$	
		FR	$p$ -value	FR	$p$ -value	FR	$p$ -value
DVEC	Normal	0.0175	(0.002)	0.0195	(0.000)	0.0170	(0.004)
	$t_{\hat{\nu}}$	0.0080	(0.352)	0.0065	(0.093)	0.0075	(0.240)
	semi-parametric	0.0115	(0.510)	0.0095	(0.821)	0.0125	(0.279)
DCC <sub>T</sub>	Normal	0.0210	(0.000)	0.0240	(0.000)	0.0215	(0.000)
	$t_{\hat{\nu}}$	0.0095	(0.821)	0.0085	(0.489)	0.0090	(0.648)
	semi-parametric	0.0125	(0.279)	0.0125	(0.279)	0.0130	(0.197)
DCC <sub>E</sub>	Normal	0.0175	(0.002)	0.0195	(0.000)	0.0175	(0.002)
	$t_{\hat{\nu}}$	0.0075	(0.240)	0.0070	(0.154)	0.0075	(0.240)
	semi-parametric	0.0115	(0.510)	0.0095	(0.821)	0.0125	(0.280)

This table presents failure rates (FR) and  $p$ -values for the Kupiec LR test. We report this for the DVEC, the DCC Tse (DCC<sub>T</sub>) and the DCC Engle (DCC<sub>E</sub>) model. We distinguish between the normal, the student and the nonparametric innovation density, for three portfolio weights.

## 5 Optimal portfolio selection

To determine the optimal portfolio that takes into account downside risk by means of a VaR constraint, we combine the approach of Campbell, Huisman, and Koedijk (2001) with the multivariate GARCH models. The portfolio model allocates financial wealth by maximizing the expected return subject to a risk constraint, measured by the Value-at-Risk (VaR). The optimal portfolio is such that the maximum expected loss would not exceed the VaR for a chosen investment horizon at a given confidence level  $1 - \alpha$ . We consider the possibility of borrowing and lending according to the investor's preferences given her utility function and given the (riskless) interest rate prevailing in the market.

Let  $W_t$  denote the investor's wealth at time  $t$ , and  $b_t$  the amount of money that is borrowed ( $b_t > 0$ ) or lent ( $b_t < 0$ ) at the risk free rate  $r_f$ . In general, we consider  $N$  financial assets with prices at time  $t$  given by  $p_{i,t}$ ,  $i = 1, \dots, N$ . Define  $X_t \equiv [x_t \in R^N : \sum_{i=1}^N x_{i,t} = 1]$  as the set of portfolio weights at time  $t$ , with well-defined expected rates of return, where  $w_{i,t} = x_{i,t}(W_t + b_t)/p_{i,t}$  is the number of shares of asset  $i$  at time  $t$ . The budget constraint of the investor is given by

$$W_t + b_t = \sum_{i=1}^N w_{i,t} p_{i,t} = w'_t p_t. \quad (20)$$

The value of her portfolio at  $t + 1$  is

$$W_{t+1}(w_t) = (W_t + b_t)(1 + r_{t+1}(w_t)) - b_t(1 + r_f), \quad (21)$$

where  $r_{t+1}(w_t)$  is the portfolio return between time  $t$  and  $t+1$  (period  $t+1$ ). The VaR of the portfolio is defined as the maximum expected loss over a given investment horizon and for a given confidence level  $1 - \alpha$ . Denote the desired Value-at-Risk of the investor at a given confidence level by  $VaR^*$ . This specifies the (negative) dollar amount corresponding to the  $\alpha$ -th percentile of the distribution of future wealth  $W_{t+1}$ . For a given  $\alpha$ , the Value-at-Risk constraint specifies that the portfolio weights should be chosen such that

$$P_t[W_{t+1}(w_t) \leq W_t - VaR^*] \leq 1 - \alpha, \quad (22)$$

where  $P_t$  is the probability conditional on the available information at time  $t$ . Equation (22) represents the second constraint that the investor has to take into account. The portfolio optimization problem then corresponds to maximizing the expected value of (21), subject to (22). This optimization problem may be rewritten in an unconstrained way. To do so, we substitute (20) in (21) and take expectations, which yields

$$E_t W_{t+1}(w_t) = w'_t p_t (E_t r_{t+1}(w_t) - r_f) + W_t(1 + r_f). \quad (23)$$

Equation (23) shows that a risk-averse investor wants to invest a fraction of her wealth in risky assets if the expected return of the portfolio exceeds the risk free rate. Substituting (23) in (22) gives:

$$P_t [w'_t p_t (r_{t+1}(w_t) - r_f) + W_t(1 + r_f) \leq W_t - VaR^*] \leq 1 - \alpha, \quad (24)$$

so that,

$$P_t \left[ r_{t+1}(w_t) \leq r_f - \frac{VaR^* + W_t r_f}{w'_t p_t} \right] \leq 1 - \alpha. \quad (25)$$

This defines the quantiles  $q(w_t, \alpha)$  of the distribution of the return of the portfolio at a given confidence level  $\alpha$ . The portfolio can then be expressed as:

$$w'_t p_t = \frac{VaR^* + W_t r_f}{r_f - q(w_t, \alpha)}. \quad (26)$$

Finally, substituting (26) in (23) we obtain:

$$E_t(W_{t+1}(w_t)) = \frac{VaR^* + W_t r_f}{r_f - q(w_t, \alpha)} (E_t r_{t+1}(w_t) - r_f) + W_t(1 + r_f), \quad (27)$$

and therefore

$$w_t^* \equiv \arg \max_{w_t} \frac{E_t r_{t+1}(w_t) - r_f}{r_f - q(w_t, \alpha)}. \quad (28)$$

The two fund separation theorem applies, i.e. the investor's initial wealth and her desired VaR ( $VaR^*$ ) do not affect the maximization problem. Thus, the optimal asset allocation

of the risky part of the investor's portfolio does not depend upon her initial wealth or upon the Value-at-Risk that is imposed. This is similar to the results for the traditional case of mean-variance optimization. However, in the VaR case, different levels of confidence lead to different risky asset allocations. At a given  $\alpha$ , the desired VaR determines the amount of borrowing or lending of the investor, which is chosen so as to (theoretically) equal the portfolio VaR with the desired VaR. The amount of money that the investor wants to borrow or lend is found by substituting (20) in (26) and is given by

$$b_t = \frac{VaR^* + W_t q(w_t^*, \alpha)}{r_f - q(w_t^*, \alpha)}. \quad (29)$$

In order to solve the optimization problem (28) we need the expected returns for each of the  $N$  assets and the predicted covariance matrix from the multivariate GARCH model. From this we can compute the predicted portfolio returns and the quantiles  $q(w_t, \alpha)$  for each vector of weights  $w_t$  given a specified  $\alpha$ . Note that the use of a multivariate GARCH model in a portfolio optimization framework like this has an important benefit over univariate approaches where for every portfolio weight the asset return series are combined into a new univariate series that is used to fit a GARCH type model from which the portfolio VaR may be computed, see Rengifo and Rombouts (2004) for such an example. This becomes computationally cumbersome when the portfolio optimization is repeated several times.

For our application we obtain optimal portfolio weights by applying (28) to each of the last 500 trading days of the sample. Since we consider only two assets, the optimization problem in (28) becomes one dimensional. The expected return is estimated by the empirical mean of the sample and the wealth at day 0 is normalized at 1000. To compute the amount of borrowing  $b_t$ , the desired Value-at-Risk level ( $VaR^*$ ) over a one-day horizon is fixed at 1% of current wealth. The annual risk-free interest rate is set to 0.04. The portfolio VaR is calculated under the assumption of the normal, the  $t$  and the nonparametric innovation density. The underlying multivariate GARCH model is the DCC model of Engle, noting that in the previous section the alternative parameterizations produced very similar results. The estimated innovation density we used to compute the VaRs is

the same as in Section 4. Note that this density estimate is kept fixed for the 500 trading days. The failure rates and the final wealth are reported in Table 7 for both the 1% and 5% VaR level.

Table 7: Optimal portfolio results

	0.05 VaR level			0.01 VaR level		
	FR	$p$ -value	Final wealth	FR	$p$ -value	Final wealth
Normal	0.044	(0.530)	1281	0.026	(0.003)	1220
$t_{\hat{\nu}}$	0.032	(0.048)	1251	0.006	(0.331)	1186
semi-parametric	0.044	(0.530)	1254	0.012	(0.663)	1186

This table presents failure rates (FR), the  $p$ -values for the Kupiec LR test and the final wealth over 500 trading days for application 1 using the DCC model of Engle. We distinguish between the normal, the student and the nonparametric innovation density. The portfolio weights are obtained by maximizing (28) for each trading day.

The failure rates for this application indicate that the multivariate GARCH model based on the normal distribution produces too risky asset allocations at the 99% confidence level, with the 1% target VaR being exceeded during 13 out of 500 trading days. On the other hand, the model based on the  $t$  distribution produces too conservative allocations at both levels of confidence. The semi-parametric specification leads to acceptable failure rates at both the 1% and the 5% level. Apparently, the conclusions of Section 4, where the VaR is calculated using fixed portfolio weights, carry over to the case where the portfolios are constructed using time varying weights obtained by maximizing the expected returns subject to a VaR constraint. That is, the semi-parametric approach works well for all VaR levels.

Figure 4 displays the evolution of the weight of the Nasdaq index in the risky portfolio, obtained for the 0.05 VaR level and using the semi-parametric procedure. We observe substantial time-variation in the composition of the risky assets portfolio during the 500 trading days. Note that the weights obtained using different distributional assumptions



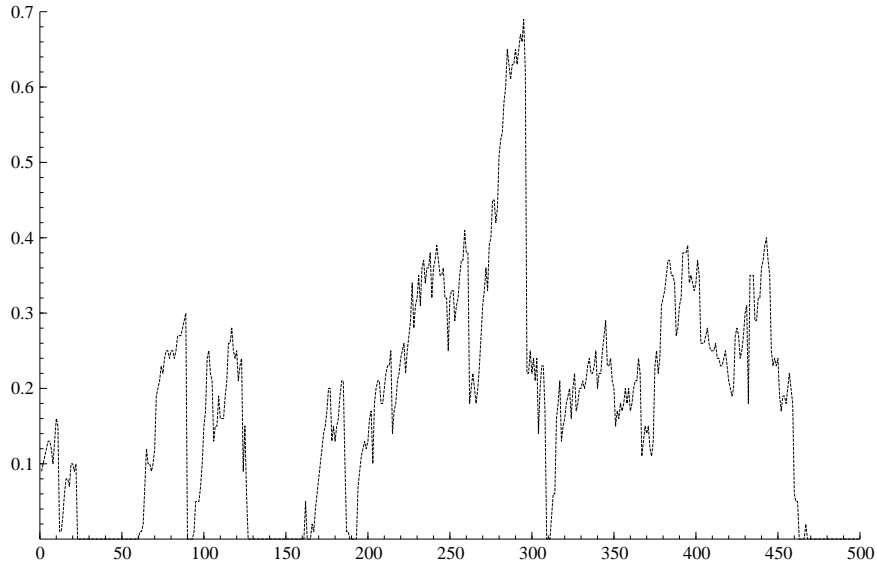


Figure 4: Plot of Nasdaq weights for the 0.05 VaR level (semi-parametric procedure).

are very similar, which explains why the resulting final wealth in Table 7 is close similar over the different rows.

Figure 5 presents the implied Value-at-Risk (at the 95% confidence level) of the risky portfolio (excluding riskless borrowing or lending). Because the composition of the risky portfolio varies from one day to the next, and because the conditional distribution of returns is time-varying, the VaR varies over time. At each trading day, the amount of borrowing or lending,  $b_t$ , should be such that the VaR of the entire investor's portfolio corresponds to her desired level  $VaR^*$ .

We also did the same portfolio optimization exercise using another data set that consisted of a stock and a bond index. The estimated innovation density showed less departures from normality but the results were basically the same as those illustrated above for the S&P 500 and the Nasdaq data. That is, based on inspection of the failure rates, the semi-parametric procedure dominates the parametric procedures.

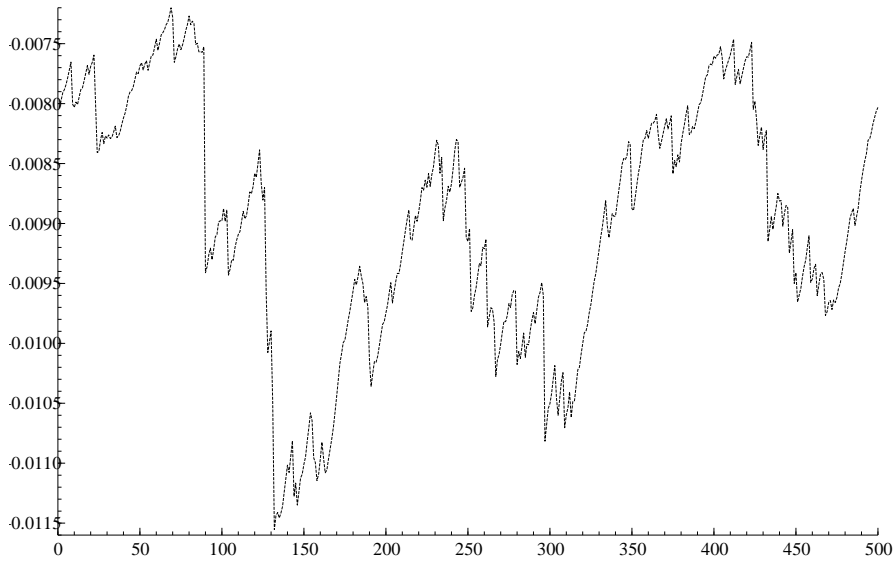


Figure 5: Plot of portfolio VaR's at the 0.05 level (semi-parametric procedure, risky portfolio only).

## 6 Conclusion

Analyzing the Value-at-Risk of a portfolio of assets with arbitrary holdings requires information about the (conditional) joint distribution of returns. In this paper we explored the usefulness of semi-parametric multivariate GARCH models for asset returns for evaluating the Value-at-Risk of a portfolio. We also illustrated how such models can be used to determine an optimal portfolio that is based on maximizing expected returns subject to a downside risk constraint, measured by Value-at-Risk. While parametric multivariate GARCH models impose strong distributional assumptions about the joint distribution of the innovations, the semi-parametric approach allows us to estimate the joint distribution without making restrictive assumptions. While this is theoretically superior, its performance in finite samples, and taking into account realistic conditions, is not necessarily optimal.

We examined the usefulness by considering the joint distribution of the returns on the S&P 500 and Nasdaq indexes. Our analyses of the 1% and 5% Value-at-Risk for a set

of three different portfolio holdings show that the semi-parametric multivariate GARCH models perform well and consistently over the different models and significance levels. This is a promising result. Interestingly, the sensitivity of the failure rates with respect to the distributional assumptions is larger than that with respect to the parametric specification that was chosen for the conditional covariance matrix (diagonal VEC model and two variants of dynamic conditional correlations models). When we determine an optimal portfolio where expected returns are maximized subject to a Value-at-Risk constraint, the results also show that the semi-parametric procedure works well, irrespective of the chosen confidence levels. While the normal distribution and the  $t$  distribution are rejected in specified cases, the semi-parametric approach passes the Kupiec likelihood ratio test in all situations.

## References

- ALEXANDER, G., AND A. BAPTISTA (2002): “Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis,” *Journal of Economic Dynamics and Control*, 26, 1159–1193.
- BARONE ADESI, G., P. GAGLIARDINI, AND G. URGÀ (2004): “Testing Asset Pricing Models with Coskewness,” *Journal of Business and Economic Statistics*, 22, 474–485.
- BAUWENS, L., S. LAURENT, AND J. ROMBOUTS (2004): “Multivariate GARCH Models: A Survey,” Forthcoming in *Journal of Applied Econometrics*.
- BAUWENS, L., M. LUBRANO, AND J. RICHARD (1999): *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, Oxford.
- BILLIO, M., AND L. PELIZZON (2000): “Value-at-Risk: a multivariate switching regime approach,” *Journal of Empirical Finance*, 7, 531–554.
- BOLLERSLEV, T., AND J. WOOLDRIDGE (1992): “Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances,” *Econometric Reviews*, 11, 143–172.
- CAMPBELL, R., R. HUISMAN, AND K. KOEDIJK (2001): “Optimal portfolio selection in a Value-at-Risk framework,” *Journal of Banking and Finance*, 25, 1789–1804.
- DE GOEIJ, P., AND W. MARQUERING (2004): “Modeling the Conditional Covariance Between Stock and Bond returns: A Multivariate GARCH Approach,” *Journal of Financial Econometrics*, 2, 531–564.
- DUFFIE, D., AND J. PAN (1997): “An Overview of Value at Risk,” *Journal of Derivatives*, 4, 7–49.
- ENGLE, R. (2002): “Dynamic Conditional Correlation : a Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models,” *Journal of Business and Economic Statistics*, 20, 339–350.

- FAN, J., AND J. GU (2003): “Semiparametric estimation of Value at Risk,” *Econometrics Journal*, 6, 261–290.
- FLEMING, J., C. KIRBY, AND B. OSTDIEK (2001): “The Economic Value of Volatility Timing,” *The Journal of Finance*, 56, 329–352.
- GIOT, P., AND S. LAURENT (2003): “Value-at-risk for long and short trading positions,” *Journal of Applied Econometrics*, 18, 641–663.
- GOURIEROUX, C., J. LAURENT, AND O. SCAILLET (2000): “Sensitivity analysis of Values at Risk,” *Journal of Empirical Finance*, 7, 225–245.
- HAFNER, C., AND J. ROMBOUTS (2004): “Semiparametric Multivariate Volatility Models,” Econometric Institute Report 21, Erasmus University Rotterdam.
- HARVEY, C., AND A. SIDDIQUE (2000): “Conditional Skewness in Asset Pricing Tests,” *Journal of Finance*, 55, 1263–1295.
- JORION, P. (2000): *Value-at-Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill.
- KUPIEC, P. (1995): “Techniques for Verifying the Accuracy of Risk Measurement Models,” *Journal of Derivatives*, 2, 173–84.
- LUCAS, A., AND P. KLAASSEN (1998): “Extreme Returns, Downside Risk, and Optimal Asset Allocation,” *Journal of Portfolio Management*, 25, 71–79.
- MARQUERING, W., AND M. VERBEEK (2004): “The Economic Value of Predicting Stock Index Returns and Volatility,” *Journal of Financial and Quantitative Analysis*, 39, 407–429.
- MITNIK, S., AND M. PAOLELLA (2000): “Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates,” *Journal of Forecasting*, 19, 313–333.

RENGIFO, E., AND J. ROMBOUITS (2004): “Dynamic Optimal Portfolio Selection In a VaR Framework,” Core Discussion Paper 2004/57.

SCOTT, D. (1992): *Multivariate Density Estimation: Theory, Practice and Visualisation*. John Wiley and Sons.

SILVERMAN, B. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall:London.

TSE, Y., AND A. TSUI (2002): “A multivariate Generalized Auto-regressive Conditional Heteroskedasticity model with time-varying correlations,” *Journal of Business and Economic Statistics*, 20, 351–362.