Portfolio Value at Risk Forecasting with GARCH-Type Models

Bachelor's Thesis

Submitted in partial fulfillment of the requirements for the degree of Bachelor of Arts in Business and Economics

Author

Jan Heinrich Schlegel

Dorfstrasse 3, 9479 Malans SG

19-747-096

janheinrich.schlegel@uzh.ch

Supervisor

Prof. Dr. Marc Paolella

Professor of Empirical Finance

Department of Banking and Finance

University of Zurich

Co-Supervisor

Soros Chitsiripanich

Date of Submission: 16.02.2023

Abstract

This thesis examines the value at risk (VaR) forecasting ability of various univariate and multivariate models for a long equity portfolio. All of the considered models involve a generalized autoregressive conditional heteroskedasticity (GARCH)-type structure. The resulting forecasts are checked for desirable properties using violation-based backtests and compared in terms of predictive ability. We find that the VaR forecasts of almost all univariate models are inadequate, while the multivariate models have few problems passing these backtests. However, we do not find evidence that the multivariate models systematically outperform their univariate counterparts with regards to predictive accuracy, or vice versa.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The last four decades were shaped by many extreme events on the financial markets. Especially noteworthy incidents were the sudden market crash on the Black Monday, the dot-com bubble and the global financial crisis of 2008. All of these supposedly rare events pointed out that reducing systematic risk is paramount to ensure the stability of the financial system. Thus, the demand for more stringent regulation strongly increased which led to the introduction and consequent tightening of the Basel Accords. This framework makes use of so-called risk measures to determine the appropriate amount of risk capital that a institution has to hold. Even though a shift towards the severity-based expected shortfall (ES) can be observed, the value at risk (VaR) to this day remains the most popular measure for downside market risk (see Embrechts et al. (2014) for the differences between the VaR and the ES). When considering a long equity portfolio, the $p\%$ VaR for period t forecasted at time $t-1$ is defined as the negative p-quantile of the conditional portfolio return distribution, i.e.

$$\mathrm{VaR}_t^p = -Q_p(r_{PF,t}|\mathcal{F}_{t-1}) = -\inf_x\{x \in \mathbb{R} : \mathbb{P}(r_{PF,t} \leq x|\mathcal{F}_{t-1}) \geq p\}, \quad p \in (0,1). \tag{1.1}$$

Hereby $Q_p(\cdot)$ denotes the quantile function and $\mathcal{F}_{t-1}$ is a filtration that represents all the information available at time $t-1$. The parameter p is the level of the VaR and indicates that with target probability p the losses of the portfolio will exceed the VaR (Paolella et al., 2006). We will follow Santos et al. (2013) and consider the 1% and the 5% VaR level in our empirical analysis.

Due to the practical relevance of this risk measure it is essential to determine how the VaR should be estimated such that it neither severely underestimates nor overestimates future losses. To this end, a plethora of models have been proposed to generate VaR forecasts, many of which involve the generalized autoregressive conditional heteroskedasticity (GARCH) model by Bollerslev (1986) or an extension thereof. These GARCH structures should account for volatility clustering and the so-called "leverage effect" which both are typically inherent in financial time series (see e.g. McNeil et al. (2015, chap. 3)). Moreover, one can either choose a univariate approach, where only the conditional portfolio variance is modelled, or a multivariate approach, where the joint dynamics of the portfolio constituents are modelled.

The question whether the more complex multivariate models are to be preferred over the simpler univariate alternatives has been extensively discussed in literature. Santos et al. (2013), who considered three large portfolios, found that the multivariate models significantly outperform their univariate counterparts. Similarly, Kole et al. (2017) found that multivariate models have greater predictive ability but these differences are in most cases not high enough to be deemed significant. Further, they find that the choice of return frequency is more important than the choice of the model itself. In addition to the data frequency and the volatility or correlation model, the assumed distribution of the error terms often majorly impacts VaR forecasts. The importance of the assumed distribution of the innovations is highlighted in Paolella et al. (2006), Slim et al. (2017), Diks and Fang (2020) or Paolella and Polak (2022) among others.

A middle ground to fitting a GARCH-type model to all the constituents and a univariate model is to only provide equity factors, which should capture the main risks of the portfolio, with a GARCH structure. Fortin et al. (2022) introduced such a framework but failed to find significant advantages over more parsimonious univariate models. Moreover, none of the models they considered adequately forecasted the one-week-ahead VaR of a portfolio in terms of conditional coverage. We replicate their models and consider an equally weighted portfolio consisting of the same ten large cap single stocks they used. However, instead of the weekly return data Fortin et al. (2022) used, we based our analysis on daily returns as this is, as identified in Kole et al. (2017), the data frequency which yields the most adequate VaR forecasts (out of the time intervals they considered).

We extend the existing literature by comparing these factor copula-DCC-NGARCH models introduced in Fortin et al. (2022) to more established models such as the diagonal MixN(k)-GARCH presented in Haas et al. (2004) or the COMFORT model class by Paolella and Polak (2015). We show that, contrary to most univariate models, the forecasts generated by the multivariate models display desirable VaR properties in the form of correct unconditional coverage and independence of the violations. Further, we do not find enough evidence to justify stating that the multivariate approaches outperform the univariate procedures in terms of forecast ability, or vice versa. Additionally, we present a slightly modified version of the model class presented in Fortin et al. (2022) and compare it to the original models. Finally, we come up with a plausible explanation for the differences in forecasting ability between this original and the modified model class.

The remainder of this thesis is structured as follows. Chapter 2 is dedicated to the univariate and multivariate models. In chapter 3 we present the data, shine a light on how the VaR estimates were created for each model and explain how the forecasts are backtested and compared. In chapter 4 we discuss our empirical results. In chapter 5 we conclude.

Chapter 2

Theoretical Framework

This chapter will introduce the models that were used to forecast the value at risk (VaR) of a portfolio in our empirical application. First, some of the most popular univariate GARCH-type structures will be presented. Next, the multivariate models in form of the factor copula model class from Fortin et al. (2022) and the COMFORT model by Paolella and Polak (2015) are discussed. Note that there are far more GARCH-type models than the few that will be mentioned in this chapter (see e.g. Bollerslev (2008) for an extensive compendium). This chapter often makes use of the notion of a information set, denoted $\mathcal{F}_{t-1}$, which in simple terms is just the information available at the end of time $t-1$ on which the forecast will be based i.e. in our case an observed sequence of past returns.

2.1 Univariate Models

For our univariate models we assume the following portfolio return dynamics:

$$r_{PF,t} = \mu + \epsilon_t, \tag{2.1}$$

where for all but the MixN(k)-GARCH we let

$$\epsilon_t = \sigma_t z_t, \quad z_t \overset{iid}{\sim} F(0,1), \tag{2.2}$$

where $F(0,1)$ is some standardized distribution (i.e. zero-location and unit-scale), μ is the unconditional location and σ_t is the scale parameter. In this framework the conditional variance

$\sigma_t^2 = \mathbb{V}[r_{PF,t}|\mathcal{F}_{t-1}]$, where $\mathcal{F}_{t-1} = \{r_{PF,1}, ..., r_{PF,t-1}\}$ denotes the information available at time $t-1$, is assumed to be non-constant.

A simple way to model this conditional variance is by means of a generalized autoregressive conditional heteroskedasticity (GARCH) process. This GARCH model introduced in Bollerslev (1986) is a generalisation of the ARCH model by Engle (1982). The most prominent version, the GARCH(1,1), can be formulated via

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{2.3}$$

where $\omega > 0, \alpha \geq 0$, and $\beta \geq 0$. In other words, in a GARCH(1,1) model the conditional volatility behaves similarly to how an ARMA(1,1) process would for the conditional mean. For covariance stationarity of the process the parameters have to fulfill $\alpha + \beta < 1$. In our empirical application, we will only consider the GARCH(1,1) model with normal innovation terms i.e. $z_t \overset{iid}{\sim} \mathcal{N}(0,1)$ which will serve as our univariate benchmark. However, it is possible to define higher order GARCH(p,q) processes, but we will only regard the case where $p = q = 1$. Thus, henceforth whenever a GARCH-type model is mentioned without explicitly stating p and q, we will refer to the $p = q = 1$ case.

A relevant special case of the GARCH model is the exponentially weighted moving average (EWMA). It sets $\omega = 0$ and the weights α are exponentially decaying and sum up to one:

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)\epsilon_t^2, \quad \lambda \in (0,1). \tag{2.4}$$

This model formulation puts more weight on the most recent observations which can be beneficial in some cases. But note that this process is not covariance stationary and the variance is thus not mean-reverting since $\lambda + (1-\lambda) = 1$. This can pose a problem when large returns are observed as they will influence conditional volatility estimates for a long time (Lucas and Zhang, 2016). For our model comparison, we will be using daily data and therefore we will set the decaying factor $\lambda = 0.94$ in accordance with J.P. Morgan (1996).

A well known stylized fact of stock returns is that negative news tends to increase volatility more than positive news of equal magnitude. To account for this so-called "leverage effect" Glosten et al. (1993) created an asymmetric extension of the standard GARCH model which they called GJR-GARCH(1,1):

$$\sigma_t^2 = \omega + (\alpha + \gamma I_{t-1})\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{2.5}$$

where $\omega > 0$, $\alpha + \gamma \geq 0$, $\beta \geq 0$ and

$$I_{t-1} = \mathbb{1}_{\{\epsilon_{t-1}<0\}} = \begin{cases} 0 & \text{if } \epsilon_{t-1} \geq 0 \\ 1 & \text{if } \epsilon_{t-1} < 0 \end{cases}$$

The process is thereby covariance stationary if $\alpha+\beta+\frac{1}{2}\gamma < 1$. Obviously, if $\gamma = 0$ this model reduces to the GARCH(1,1) from equation 2.3 which treats news symmetrically, i.e. bad news ($\epsilon_{t-1} < 0$) have the same impact as good news ($\epsilon_{t-1} \geq 0$). In practice however, γ is usually found to be positive (Bollerslev, 2008). This leads to an asymmetric relationship since negative news is weighted with $\alpha + \gamma$ and positive news only with α. We decided to include this model in our comparison due to the great performance of the GJR with Student t innovations in Santos et al. (2013). Additionally, we tried out a GJR with the skewed-t distribution by Hansen (1994) for the error terms.

Another way to capture the leverage effect is presented in Engle and Ng (1993) which they termed the NGARCH(1,1) model:

$$\sigma_t^2 = \omega + \alpha\sigma_{t-1}^2(\epsilon_{t-1} - \theta)^2 + \beta\sigma_{t-1}^2, \tag{2.6}$$

where $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$. For covariance stationarity $\alpha(1 + \theta^2) + \beta < 1$ is required. It is apparent that for $\theta > 0$ negative innovations ($\epsilon_{t-1} < 0$) have a larger impact on the conditional variance than positive errors of the same magnitude which causes this model to be asymmetric too (Christoffersen and Langlois, 2013). We will incorporate this model with a skewed-t distribution for the innovations as this is the same specification that Fortin et al. (2022) and Christoffersen and Langlois (2013) used in their factor copula models.

The last class of univariate models that we included in our empirical comparison is the k-component mixed normal GARCH(1,1) (MixN(k)-GARCH) as introduced in Haas et al. (2004). We will only consider the case which was referred to as diagonal by Haas et al. (2004) as they identified it to be the superior choice in terms of performance and interpretability. In this framework, the conditional distribution of the error term ϵ_t is assumed to be mixed normal (see appendix C.1) with zero mean,

$$\epsilon_t|\mathcal{F}_{t-1} \sim \text{Mix}_k\text{N}(p_1, ..., p_k, \mu_1, ..., \mu_k, \sigma^2_{1,t}, ..., \sigma^2_{k,t}), \quad \sum_{i=1}^{k} p_i\mu_i = 0, \tag{2.7}$$

where $p_i \in (0,1)\,\forall i$, $\sum_{i=1}^{k} p_i = 1$ and $\mathcal{F}_{t-1} = \{r_{PF,1}, ..., r_{PF,t-1}\}$ is the information set at time $t-1$. The associated conditional variances are given by GARCH processes:

$$\sigma_{i,t} = \omega_i + \alpha_i\epsilon^2_{i,t-1} + \beta_i\sigma^2_{i,t-1}, \quad i = 1, ..., k. \tag{2.8}$$

The $k \geq 2$ components that are used for the conditional variance each represent a different market condition (Slim et al., 2017). Hence, often a low number of components such as 2 or 3 is sufficent as discovered in Haas et al. (2004). In case of different means for the different components, a normal mixture density incorporates skewness and fat tails (or thin tails) (Alexander and Lazar, 2006). This property enables the MixN(k)-GARCH model to produce sound VaR estimates as showcased in Haas et al. (2004) and Paolella et al. (2006).

2.2 A Multivariate Factor Copula-DCC-GARCH Model

The models described in the previous section all forecasted future returns of a portfolio based on past portfolio returns. Alternatively, we can take a multivariate approach where we first model the constituents of the portfolio and only in a second step draw conclusions about the portfolio. One such model was proposed in Fortin et al. (2022) following the discoveries made about asymmetric tail dependency of weekly factor returns in Christoffersen and Langlois (2013). Consequently, Fortin et al. (2022) make use of equity factors that should capture the main risks of stock returns to forecast the VaR of a portfolio of single stocks. More precisely, they utilize the Carhart four-factor model suggested in Carhart (1997) which adds a momentum factor to the linear three-factor model of Fama and French (1993). Within this model, the return of a single stock k in excess of the weekly risk free rate $r_{f,t}$ is given by

$$\begin{aligned} r_{k,t} - r_{f,t} &= \alpha_{k,t} + \beta_{k,\text{RMRF}}\text{RMRF}_t + \beta_{k,\text{SMB}}\text{SMB}_t + \beta_{k,\text{HML}}\text{HML}_t + \beta_{k,\text{MOM}}\text{MOM}_t + \varepsilon_{k,t} \quad (2.9) \\ &= \alpha_{k,t} + \boldsymbol{\beta}_k' \boldsymbol{r}_{F,t} + \varepsilon_{k,t}, \qquad t = 1, 2, ..., T, \quad (2.10) \end{aligned}$$

where both the intercept $\alpha_{k,t}$ and the vector of factor loadings $\boldsymbol{\beta}_k$ are assumed to be constant over time. The vector of equity factors at time t, $\boldsymbol{r}_{F,t} = (\text{RMRF}_t, \text{SMB}_t, \text{HML}_t, \text{MOM}_t)'$, consists of the market factor (RMRF), the size factor (SMB), the value factor (HML) and the momentum factor (MOM) (cf. Carhart (1997)). This leads to a reduction in dimensionality of the problem since only four factors have to be modelled instead of all constituents of the stock portfolio.

Fortin et al. (2022) decided to not only model the conditional variances but also the dependence between the factors. One possible way to do this is by using the Dynamic Conditional Correlation (DCC) structure[1] introduced in Engle (2002) which is a dynamic extension of the Constant Conditional Correlation (CCC) model by Bollerslev (1990). Let $\boldsymbol{Y_t} = (y_{1,t}, y_{2,t}, ..., y_{n,t})'$ be a vector consisting of the returns of n assets or of n factors at time $t = 1, ..., T$ and $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{Y_t}|\mathcal{F}_{t-1}]$ the corresponding mean vector which is assumed to be constant. The information set at time $t-1$ once again is denoted by $\mathcal{F}_{t-1}$ and consists of the past returns $\{\boldsymbol{Y_1}, ..., \boldsymbol{Y_{t-1}}\}$. We can then define the conditional covariance matrix of $\boldsymbol{r_t}$ as $\boldsymbol{\Sigma_t} := \mathbb{E}[(\boldsymbol{Y_t} - \boldsymbol{\mu})(\boldsymbol{Y_t} - \boldsymbol{\mu})'|\mathcal{F}_{t-1}]$. The DCC model decomposes the dynamics of the conditional covariance matrix into standard deviations and correlation and can be written as:

$$\boldsymbol{Y}_t|\mathcal{F}_{t-1} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma_t}), \quad \boldsymbol{\Sigma_t} = \boldsymbol{D_t}\boldsymbol{\Gamma_t}\boldsymbol{D_t} \tag{2.11}$$

where $\boldsymbol{D_t} = diag\,(\sigma_{1,t}, \sigma_{2,t}, ..., \sigma_{n,t})$ i.e. a $n \times n$ diagonal matrix with the diagonal consisting of the square roots of the conditional variances of the assets or the factor returns. These conditional variances can be modelled using one of the univariate models from section 2.1. Fortin et al. (2022) chose

[1] In contrast to Fortin et al. (2022), we will be using the DCC model of Engle (2002) instead of the cDCC model of Aielli (2013) due to its more established R implementation. The differences between these two models are relatively small according to Christoffersen and Langlois (2013).

an AR(3)-NGARCH(1,1) model with the skewed-t distribution of Hansen (1994) for the innovations of all factor returns. We however tried out two different univariate models in our empirical application: a NGARCH(1,1) with skewed-t and a GARCH(1,1) with normal innovations, both without an ARMA term. Henceforward, these multivariate models will be named by first specifying the conditional correlation structure followed by the GARCH-type process used for the conditional variances e.g. DCC-GARCH or CCC-GJR.

The $n \times n$ conditional correlation matrix $\boldsymbol{\Gamma_t}$ from equation 2.11 is symmetric positive-definite and given by

$$\boldsymbol{\Gamma_t} := \mathbb{E}[\boldsymbol{z_t z_t'} | \mathcal{F}_{t-1}] = diag\left(\boldsymbol{Q_t}^{-1/2}\right) \boldsymbol{Q_t}\, diag\left(\boldsymbol{Q_t}^{-1/2}\right) \tag{2.12}$$

where $\boldsymbol{z_t} := \boldsymbol{D_t}^{-1}(\boldsymbol{Y_t} - \boldsymbol{\mu})$ are the standardized residuals. $\boldsymbol{Q_t}$ also is a $n \times n$ symmetric positive-definite matrix that fulfills

$$\boldsymbol{Q_t} = (1 - \alpha - \beta)\bar{\boldsymbol{Q}} + \alpha \boldsymbol{z_{t-1} z_{t-1}'} + \beta \boldsymbol{Q_{t-1}}, \quad \alpha, \beta \geq 0, \tag{2.13}$$

where $\bar{\boldsymbol{Q}} := \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{z_t z_t'}$ is the $n \times n$ unconditional covariance matrix. Setting $\alpha = \beta = 0$ in equation 2.13 yields the CCC model of Bollerslev (1990) where $\boldsymbol{\Gamma_t} = \boldsymbol{\Gamma} = \bar{\boldsymbol{Q}}$ for all t.

Possibly the biggest drawback of the DCC model as specified above is the underlying assumption of multivariate normality. In particular, Christoffersen and Langlois (2013) discovered that ignoring the multivariate non-normality inherent in weekly factor returns leads to severe underestimation of the Expected Shortfall of an equally weighted portfolio of factors. To encompass this non-normality and the aforementioned asymmetric tail dependence of weekly equity factors, Fortin et al. (2022) stick to Christoffersen and Langlois (2013) and make use of so-called copulas to fit the joint conditional distribution of the factor returns. Copulas are functions that allow us to model the marginals (in our case the standardized residuals of the DCC-(N)GARCH model) individually and independently of the multivariate distribution making them incredibly flexible. The notion of Copulas is based on Sklar's theorem which when applied to the standardized DCC-(N)GARCH residuals of our n factor returns states that

$$\boldsymbol{F_t}(\boldsymbol{z_t}) = \boldsymbol{C_t}(F_{1,t}(z_{1,t}), ..., F_{n,t}(z_{n,t})), \tag{2.14}$$

where $\boldsymbol{F_t}(\boldsymbol{z_t})$ is the joint conditional distribution of the standardized residuals at time t, $F_{i,t}(\cdot)$ is the conditional marginal distribution of the standardized residual of factor i at time t and $\boldsymbol{C_t} : [0,1]^n \rightarrow [0,1]$ is the conditional copula that links these marginal distributions. As highlighted in Heinen and Valdesogo (2012), we can only apply Sklar's theorem to conditional distributions if we condition all marginals and the copula on the same information. This can be done by assuming that each marginal only depends on its own past and that the copula depends on the history of all four factors (Heinen and Valdesogo, 2012). As explained in e.g. Heinen and Valdesogo (2012) a copula can also be expressed as a multivariate distribution with $\mathcal{U}(0,1)$ margins

$$\boldsymbol{C_t}(u_1, ..., u_n) = \boldsymbol{F_t}(F_{1,t}^{-1}(u_{1,t}), ..., F_{n,t}^{-1}(u_{n,t})), \tag{2.15}$$

where $u_{i,t} = F_{i,t}(\epsilon_{i,t})$ is the probability integral transform (PIT) (see appendix A.1) of the standardized residuals of factor i at time t. This formulation shows that if we can sample from the copula, we can sample from the corresponding multivariate distribution $\boldsymbol{F_t}(\cdot)$. Thus, for Monte Carlo simulations one can first generate a vector of probabilities $\boldsymbol{u_t} := (u_{1,t}, ..., u_{n,t})'$ with conditional distribution $\boldsymbol{C_t}(\cdot)$ and then apply the quantile transform (see appendix A.2) to these $u_{i,t}$ to return a sample vector $(F_{1,t}^{-1}(u_{1,t}), ..., F_{n,t}^{-1}(u_{n,t}))'$ from $\boldsymbol{F_t}(\cdot)$.

In summary, we follow Christoffersen (2011, chap. 9) and apply the following scheme to simulate future factor returns:

1. Fit a DCC model with a GARCH or NGARCH model for the conditional variances and extract the standardized residuals $\boldsymbol{z_t} = (z_{1,t}, ..., z_{n,t})'$.

2. Calculate probabilities $u_{i,t} = F_{i,t}(z_{i,t})$, where $F_{i,t}(\cdot)$ is the conditional distribution that was estimated in the (N)GARCH model for the i'th factor.

3. Fit a copula to these probabilities. In our empirical application we consider the normal copula (see appendix B.1), the Student t copula (see appendix B.2) and the skewed-t copula. For

more information about the Student t copula and the skewed-t copula consult Demarta and McNeil (2005) and Yoshiba (2014).

4. Simulate a vector of probabilities $(\tilde{u}_{1,t}, ..., \tilde{u}_{n,t})'$ from the conditional copula.

5. Create simulated standardized residuals from the simulated copula probabilities using quantile transforms: $\widetilde{\boldsymbol{z}_t} := (F_{1,t}^{-1}(\tilde{u}_{1,t}), ..., F_{n,t}^{-1}(\tilde{u}_{n,t}))'$.

6. To create factor returns from the simulated standardized residuals we use the forecasted dynamics obtained from the DCC model:

$$\widetilde{\boldsymbol{r_{F,t}}} = \boldsymbol{\mu} + \boldsymbol{\Sigma_t}^{1/2} \widetilde{\boldsymbol{z}_t}, \tag{2.16}$$

where $\boldsymbol{\Sigma_t}^{1/2}$ denotes the matrix square root obtained through a Cholesky decomposition of the one-period-ahead forecast of the (DCC) covariance matrix of the factor returns and $\boldsymbol{\mu}$ the unconditional mean vector of the factor returns.

Applying equation 2.10 to these simulated one-period-ahead factor returns yields forecasts of the returns of the single stocks (see section 3.2 for more information). Finally, it is straightforward to calculate the simulated one-day-ahead portfolio returns and the desired risk measures since this approach is simulation-based.

2.3 The COMFORT Model

In the previous section we presented the approach of using copulas to introduce non-normality (and heterogeneous tails when using e.g. the skewed-t copula) into a CCC or DCC structure. But one could also directly choose a non-Gaussian distribution for the CCC or DCC model. One possible way to implement this is by first fitting a normal CCC or DCC model to the stock returns and then in a second step fitting a multivariate, non-normal distribution to the filtered residuals obtained in the first step (Paolella, 2018, chap. 11). But this so-called quasi maximum likelihood approach is inferior to joint maximum likelihood estimation of all parameters (Paolella, 2018, chap. 11). Consequently, Paolella and Polak (2015) developed an efficient expectation-maximization (EM) type algorithm that allows for full maximum likelihood estimation of all parameters of the multivariate generalized hyperbolic (MGHyp) distribution. This uni-modal distribution is commonly used in finance literature and offers great flexibility (Paolella and Polak, 2022). In particular, many popular distribution choices such as the multivariate normal, the multivariate Student t, the multivariate variance gamma or the normal inverse Gaussian (and many more) can be expressed as a limiting or special case of the MGHyp (Paolella and Polak, 2022). For a more detailed breakdown of this distribution we refer to McNeil et al. (2015, chap. 6), Paolella and Polak (2015) and Paolella and Polak (2022). Applying the aforementioned multi-stage EM algorithm to a MGHyp distribution, for whose covariance matrix a CCC or DCC structure was used, results in the so-called Common Market Factor Non-Gaussian Returns (COMFORT) model as introduced in Paolella and Polak (2015).

Let $\boldsymbol{Y_t} = (y_{1,t}, y_{2,t}, ..., y_{n,t})'$ be an n-dimensional vector of returns of the constituents of a portfolio at time $t = 1, ..., T$. The COMFORT model then has the form $\boldsymbol{Y_t} \sim \text{MGHyp}(\boldsymbol{\mu}, \boldsymbol{\gamma}, \boldsymbol{\Sigma_t}, \lambda_t, \chi_t, \psi_t)$ and can be expressed as a continuous normal mixture in the following way:

$$\boldsymbol{Y_t} = \boldsymbol{\mu} + \boldsymbol{\gamma} G_t + \boldsymbol{\varepsilon_t}, \qquad \boldsymbol{\varepsilon_t} = \boldsymbol{\Sigma_t}^{1/2} \sqrt{G_t} \boldsymbol{Z_t}, \tag{2.17}$$

where the mean vector $\boldsymbol{\mu}$ and the asymmetry vector $\boldsymbol{\gamma}$ are in $\mathbb{R}^n$ and $\boldsymbol{Z_t} \overset{iid}{\sim} \mathcal{N}_n(\boldsymbol{0}, \boldsymbol{I_n})$ (Paolella and Polak, 2022). Further, the symmetric positive-definite covariance matrix $\boldsymbol{\Sigma_t} = \boldsymbol{D_t \Gamma_t D_t}$ is modelled using a CCC or DCC structure (see section 2.2) where the matrix square root is taken by means of a Cholesky decomposition. Lastly, the mixing random variables $G_t | \mathcal{F}_{t-1} \sim \text{GIG}(\lambda_t, \chi_t, \psi_t)$, where $\mathcal{F}_{t-1} = \{\boldsymbol{Y_1}, ..., \boldsymbol{Y_{t-1}}\}$ is the information set at time $t-1$, follow a generalized inverse Gaussian (GIG) random variable and are independent of $\boldsymbol{Z_t}$ (Paolella and Polak, 2022). Consult Barndorff-Nielsen et al. (1992) for a detailed account on the GIG distribution or Paolella and Polak (2022) for more information about the GIG in the COMFORT model. This sequence $\{G_t\}$ can be interpreted as a common market factor because conditional on this common market factor, which incorporates jumps and news arrival, the returns are multivariate normal distributed (Paolella, 2018, chap. 11).

In our empirical application in chapter 4 we will only consider the CCC augmentation in which $\boldsymbol{\Gamma_t} = \boldsymbol{\Gamma}$ is time invariant. For the dynamics of the diagonal elements of $\boldsymbol{D_t}$ adjusted GARCH-type processes, which incorporate the common market factor, are used. More precisely, we will utilize the GARCH(1,1) and the GJR-GARCH(1,1) from section 2.1 but with error term $\epsilon_{i,t} = y_{i,t} - \mu_i - \gamma_i G_t$ instead of $\epsilon_{i,t} = y_{i,t} - \mu_i$ for the i'th portfolio constituent. Furthermore, we will only study the so-called multivariate variance-gamma (MVG) distribution which can be expressed as a MGHyp distribution with $\lambda_t > 0$, $\chi_t = 0$ and ψ_t=2 for all t (see Paolella and Polak (2022)). This restriction of the shape parameters χ_t and ψ_t is recommended by Paolella and Polak (2015) to circumvent potential numerical problems caused by otherwise relatively flat likelihoods.

An important property of the MGHyp distribution is that it is closed under linear operations as shown in McNeil et al. (2015, chap. 6). Thus, the returns of a portfolio consisting of our n constituents with constant portfolio weights $\mathbf{w} = (\mathrm{w}_1, \mathrm{w}_2, ..., \mathrm{w}_n)' \in \mathbb{R}^n \setminus \mathbf{0}$ are given by $r_{PF,t} = \mathbf{w}'\boldsymbol{Y_t}$ and are univariate GHyp distributed:

$$r_{PF,t}|\mathcal{F}_{t-1} \sim \mathrm{GHyp}(\mathbf{w}'\boldsymbol{\mu}, \mathbf{w}'\boldsymbol{\gamma}, \mathbf{w}'\boldsymbol{\Sigma_t}\mathbf{w}, \lambda_t, \chi_t, \psi_t). \tag{2.18}$$

This property is especially useful in portfolio optimization and risk management and will be used in section 3.2 to estimate the VaR. We decided to include this model class because of the adequate VaR forecasts it generated in Paolella and Polak (2022). Particularly, the empirical analysis in Paolella and Polak (2022) showed that the VaR estimates of this model class were superior to the ones of a CCC-GARCH structure that assumed multivariate normality.

Chapter 3

Methodology

This chapter will present the data we considered for our empirical application. Further, we will describe how the VaR forecasts were generated and shine a light on some issues we faced. Finally, the violation based likelihood ratio tests of Christoffersen (1998) as well as a loss function based test in form of the conditional predictive ability (CPA) test by Giacomini and White (2006) will be introduced.

3.1 Data

This thesis assesses models based on their ability to forecast the VaR of a long portfolio. This portfolio is equally weighted and consists of the ten large cap single stocks also used in Fortin et al. (2022). However, we consider 2767 daily returns (instead of weekly returns) which were observed from January 2, 2001 to December 30, 2011. This data is freely available on Yahoo Finance. For reasons of numerical stability we will be working with daily percentage log returns i.e.

$$r_t = 100 \cdot \log\left(\frac{P_t}{P_{t-1}}\right)$$

where P_t is the price at time t. Additionally, we require daily returns of the factors from the Carhart four-factor model for the same time frame. These returns are from Kenneth R. French's data library. Note that these factor returns are in percent but nominal and first have to be converted to percentage log returns. In the following, whenever returns are mentioned we will refer to daily percentage log returns. Furthermore, the VaR forecasts will be the one-step-ahead percentage log return VaR.

Table 3.1 – Summary statistics of daily factor, stock and portfolio percentage log returns

	Mean	Median	SD	MAD	Min	Max	Skew	Kurt	JB
Factors:									
Market	0.001	0.070	1.383	0.916	-9.376	10.75	-0.192	6.786	5337
Size	0.018	0.040	0.622	0.534	-3.864	3.710	-0.167	3.588	1501
Value	0.009	0.010	0.718	0.401	-4.458	4.736	0.284	9.827	11190
Momentum	-0.001	0.060	1.113	0.680	-8.578	6.878	-0.971	8.560	8900
Stocks:									
Boeing	0.012	0.048	2.062	1.622	-19.389	14.378	-0.331	5.764	3888
CAT	0.058	0.078	2.209	1.715	-15.686	13.735	-0.110	4.303	2145
Chevron	0.047	0.113	1.745	1.370	-13.341	18.942	0.025	12.101	16913
Coca-Cola	0.015	0.043	1.325	0.932	-10.604	12.997	0.016	9.397	10199
Exxon	0.033	0.079	1.691	1.286	-15.027	15.863	-0.019	10.747	13341
GE	-0.023	0.000	2.168	1.348	-13.684	17.984	0.040	7.746	6932
IBM	0.033	0.035	1.674	1.183	-10.668	11.354	0.275	5.991	4182
Merck	-0.016	0.024	1.938	1.324	-31.171	12.251	-1.897	30.435	108622
PG	0.029	0.033	1.238	0.888	-8.316	9.726	-0.211	6.242	4522
UTC	0.030	0.052	1.853	1.313	-30.290	12.793	-1.540	29.797	103620
Portfolio:									
Portfolio	0.022	0.082	1.300	0.891	-8.211	10.690	-0.059	7.785	7003

This table displays summary statistics of the daily factor, stock and portfolio returns and was inspired by a similar table in Fortin et al. (2022). The columns mean, median, standard deviation (SD), mean absolute deviation (MAD), minimum and maximum are in percentages. The last three columns include the skewness, the kurtosis and the Jarque-Bera test statistic. Data are daily and cover the time period from January 2, 2001 to December 30, 2011 (2767 observations).

In Table 3.1 the summary statistics for the daily factor, stocks and portfolio returns are presented. One can see that whilst the median is larger than the mean in most instances, both are close to zero for all factors and stocks in question. In addition, the mean absolute deviation (MAD) is considerably smaller than the standard deviation indicating the presence of outliers. Three of the four factors, six of the ten stocks and the portfolio are left skewed and thus have a longer left tail.

All factors, stocks and the portfolio show leptokurtic behavior i.e. their kurtosis is larger than three. This signals that the return distributions have fatter tails than a Gaussian distribution would have. Besides, we can reject the assumption of univariate normality for all of the return distributions using the extraordinarily high Jarque-Bera statistics. Additionally, we checked for multivariate normality by constructing Q-Q plots (see figure 3.1) of the robust squared Mahalanobis distances of the stock returns (factor returns) and the corresponding χ_d^2 distribution with $d = 10$ ($d = 4$) degrees of freedom. The relationship in these Q-Q plots is not linear at all indicating large multivariate outliers and multivariate non-normality.

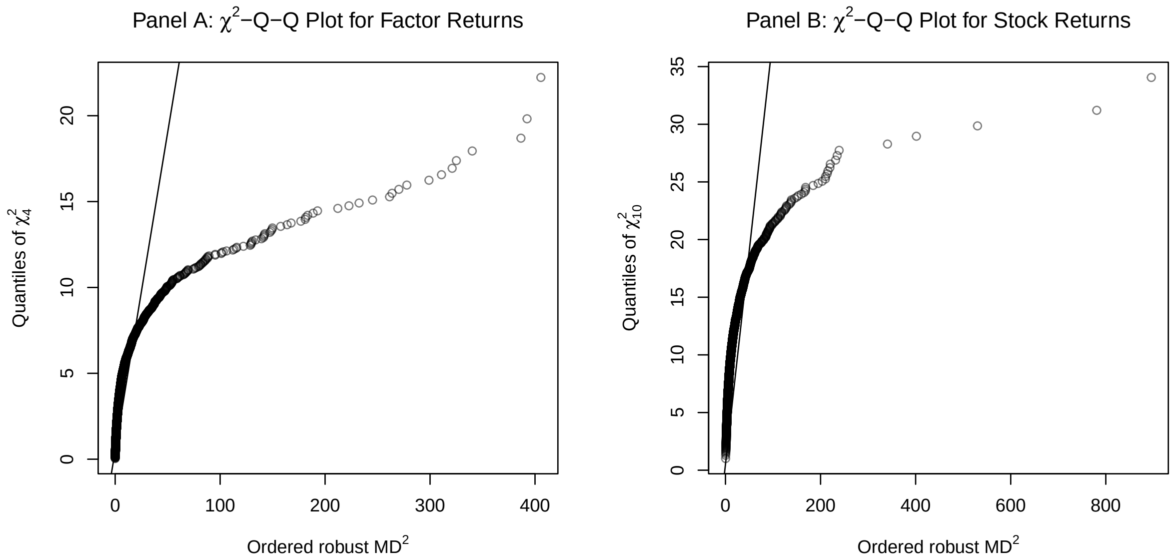

Figure 3.1 – χ^2 Q-Q plots of the robust squared Mahalanobis distance of the factor returns and the stock returns. The degrees of freedom correspond to the number of factors in Panel A and to the number of stocks in Panel B. Data are daily and cover the time period from January 2, 2001 to December 30, 2011 (2767 observations).

Moreover, the portfolio returns exhibit large volatility around their mean as shown in panel A of Figure 3.2. Panel A further shows blatant volatility clustering indicating that the conditional variance is non-constant. In addition, it is apparent from Table 3.1 that the portfolio returns are not normally distributed. Panel B graphically displays this observation and shows that the distribution of the portfolio percentage log returns is slightly left skewed and has long tails with some extreme values on either side. Interestingly, both the smallest and the largest portfolio returns were observed in October of 2008.

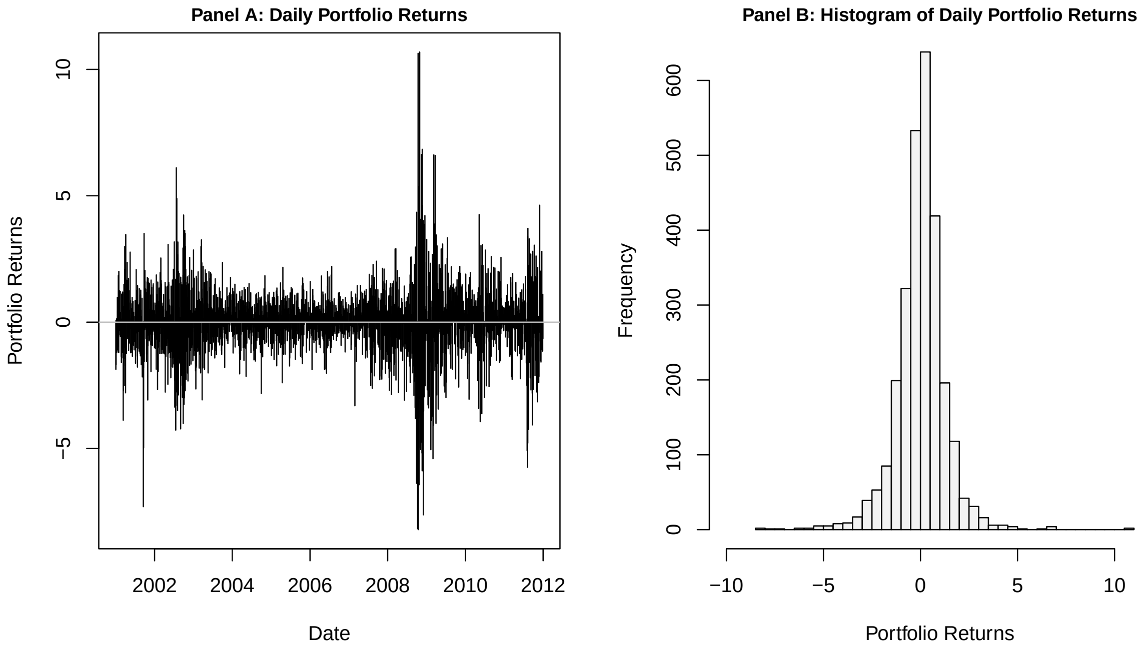

Figure 3.2 – Plot and histogram of the percentage log-returns of the equally weighted long equity portfolio. Data are daily and cover the time period from January 2, 2001 to December 30, 2011 (2767 observations).

Despite using daily and not weekly factor returns, the autocorrelation function (ACF) plots in Figure 3.3 display a similar picture as the ones in Christoffersen and Langlois (2013). It is evident that the factors exhibit stronger autocorrelation in absolute returns than in the returns themselves. In particular, panels A to D illustrate that whilst the ACF values for the returns mostly stay within the 95%-confidence bands, the absolute factor returns show significant autocorrelation for all lags considered. The same phenomenon can also be observed for the single stock returns and the portfolio returns (not displayed here) and can be seen as further justification for fitting a volatility model to the return series.

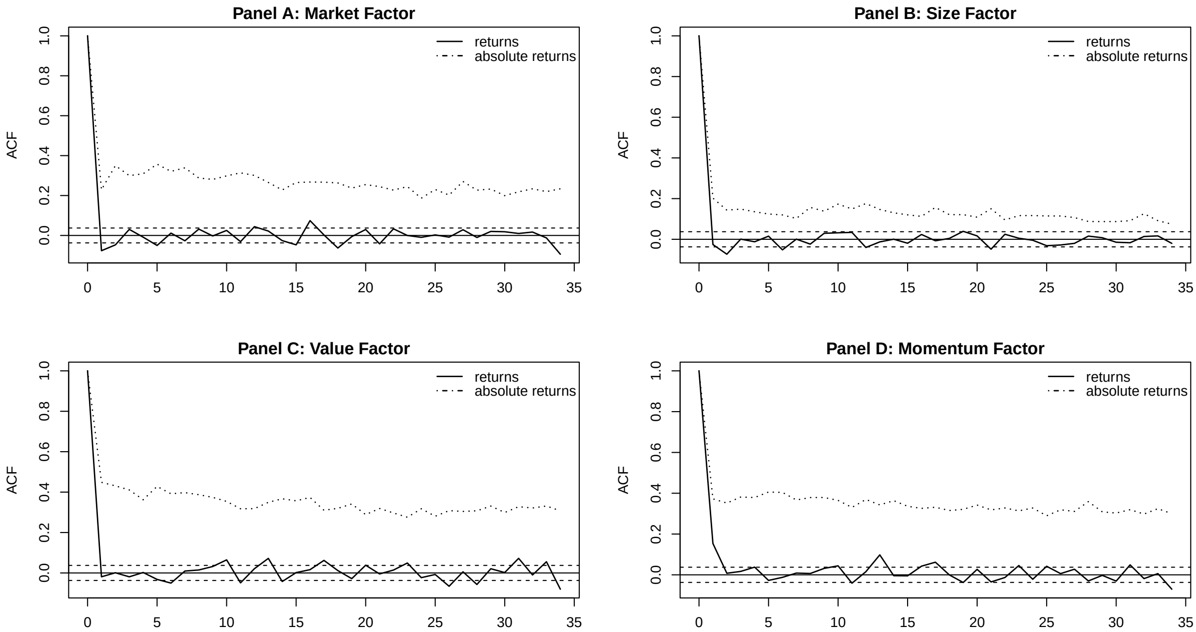

Figure 3.3 – ACF Plots of the Fama-French-Carhart Factors inspired by the ACF plots of the weekly factors in Christoffersen and Langlois (2013). The autocorrelation values are displayed using a solid line for the factor returns and a dash-dotted line for absolute factor returns. Data are daily and cover the time period from January 2, 2001 to December 30, 2011 (2767 observations).

3.2 Value at Risk Forecasts

For all models in question we assume that the conditional mean is constant over time and thus will proceed without specifying an ARMA(p,q) process for the conditional mean. Santos et al. (2013) took the same approach for the reason that the dynamic dependence of conditional means in daily portfolio returns is very weak if existent at all. Forecasting is done using a rolling window approach where the previous 1000 observations are used to predict the one step-ahead-VaR. The model parameters will be refit after every rolling window iteration as in Paolella et al. (2006). The only exception is the skewed-t copula-DCC-GARCH model where we decided to reestimate the skewed-t copula parameters only every 20 rolling windows due to the high computational burden. Addtitionally, there were some numerical issues for the skewed-t copula with skewed-t NGARCH marginals which we could not resolve and thus only the version with normal GARCH marginals will be included.

For all univariate GARCH models mentioned in section 2.1 but the MixN(k)-GARCH we can rely on the following analytical formula to estimate the portfolio VaR:

$$\widehat{\text{VaR}_t^p} = -\mu_{PF} - \sigma_{PF,t} Q_p(z_t|\mathcal{F}_{t-1}) \tag{3.1}$$

where $\sigma_{PF,t}$ is the portfolio conditional standard deviation at time t and $Q_p(z_t)$ is the p-quantile of the conditional distribution of the standardized portfolio returns, $z_t = (r_{PF,t} - \mu_{PF})/\sigma_{PF,t}$. Note that formula 3.1 includes the unconditional mean instead of the conditional one due to our mean specification . In the MixN(k)-GARCH setting the VaR forecasts are generated by the "MSGARCH" package of Ardia et al. (2019) by applying the defintion of the VaR from equation 1.1 to the predictive mixed normal distribution of the portfolio returns.

For the factor copula-DCC-(N)GARCH models in section 2.2, we proceed in the same way as Fortin et al. (2022). Hence, we start off by calculating the ordinary least squares (OLS) estimates of equation 2.10. These estimates are displayed below in table 3.2. Striking are the high Jarque-Bera statistics for the OLS residuals indicating univariate non-normality. Further analysis of the OLS residuals shows that the residuals are slightly left skewed and highly leptokurtic. Moreover, we can see in figure 3.4, which looks very similar to Panel B of figure 3.1, that the OLS residuals clearly are not multivariate normal either. Additionally, residuals of firms from similar sectors tend to be stronger correlated than the residuals of those from different sectors. Next, we follow the procedure described in section 2.2 to simulate a vector $\tilde{\boldsymbol{r}}_{F,t}$. We then simulate from the distribution of the OLS residuals $\boldsymbol{F_\varepsilon}(\cdot)$ by randomly drawing a bootstrap sample vector $\tilde{\boldsymbol{\varepsilon}}_t = [\tilde{\varepsilon}_{1,t}, \tilde{\varepsilon}_{2,t}, ..., \tilde{\varepsilon}_{10,t}]$, $t = 1, 2, ..., T$. It is important to note that, as pointed out in Christoffersen (2009), each draw of $\boldsymbol{F_\varepsilon}(\cdot)$ has to be a vector of error terms from the same day so that aforementioned dependencies between the OLS residuals of different stock returns are maintained. Further, we have to pay attention to data leakage and can therefore only use OLS residuals which have already been observed i.e. are part of the rolling window that is used for time $t = 1, 2, ..., T$. Using equation 2.10 then yields the simulated return of single stock k:

$$\tilde{r}_{k,t} = r_{f,t} + \alpha_{k,t} + \boldsymbol{\beta}_k' \tilde{\boldsymbol{r}}_{F,t} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{k,t}}. \tag{3.2}$$

This procedure is repeated $N = 200'000$ times. For each simulation we calculate the simulated return of the equal weighted portfolio. This yields a simulated portfolio return sample $\{r_{PF}^n\}_{n=1}^N$. Finally, the VaR estimate is obtained as the negative p-quantile of the simulated portfolio returns:

$$\widehat{\mathrm{VaR}_t^p} = -Q_p(\{r_{PF}^n\}_{n=1}^N).$$

Table 3.2 – Parameter Estimates for the Carhart Four-Factor Model

	Intercept	Market	Size	Value	Momentum	Adj. R2	JB
Boeing	0.004	1.041	-0.067	0.022	0.150	0.441	3586
	(0.030)	(0.038)	(0.062)	(0.081)	(0.049)		
CAT	0.045	1.201	0.203	0.095	0.094	0.555	10850
	(0.028)	(0.029)	(0.066)	(0.079)	(0.055)		
Chevron	0.044	1.042	-0.390	0.143	0.407	0.564	1606
	(0.020)	(0.074)	(0.097)	(0.085)	(0.053)		
Coca-Cola	0.011	0.604	-0.205	-0.090	0.157	0.324	6295
	(0.020)	(0.028)	(0.040)	(0.054)	(0.038)		
Exxon	0.033	1.057	-0.477	-0.029	0.435	0.592	1703
	(0.018)	(0.059)	(0.074)	(0.079)	(0.049)		
GE	-0.030	1.095	-0.220	0.159	-0.236	0.605	43274
	(0.024)	(0.070)	(0.126)	(0.087)	(0.047)		
IBM	0.031	0.854	-0.152	-0.445	-0.149	0.527	10906
	(0.022)	(0.036)	(0.046)	(0.060)	(0.033)		
Merck	-0.016	0.836	-0.376	-0.120	0.212	0.294	409268
	(0.030)	(0.043)	(0.058)	(0.088)	(0.056)		
PG	0.025	0.570	-0.291	0.060	0.193	0.343	4394
	(0.017)	(0.035)	(0.043)	(0.058)	(0.043)		
UTC	0.023	1.041	-0.093	-0.014	0.110	0.554	515050
	(0.021)	(0.037)	(0.054)	(0.051)	(0.039)		

This table was inspired by a similar table in Fortin et al. (2022) and presents the OLS parameter estimates and the corresponding Newey-West standard errors in parentheses of the Carhart four-factor model. The last two columns show the corresponding OLS adjusted R-squared and the Jarque-Bera test statistic of the OLS residuals. Data are daily and cover the time period from January 2, 2001 to December 30, 2011 (2767 observations).

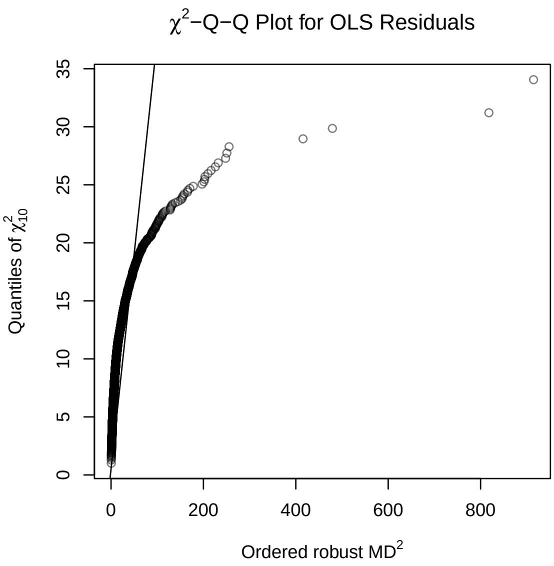

Figure 3.4 – χ^2 Q-Q plots of the robust squared Mahalanobis distance of the OLS residuals of the Carhart four-factor model. The degrees of freedom correspond to the number of stocks i.e. 10. Data are daily and cover the time period from January 2, 2001 to December 30, 2011 (2767 observations).

Lastly, in the COMFORT framework we make use of the fact that when assuming a MGHyp distribution for $\boldsymbol{Y_t}$ (i.e. the returns of the constituents) the corresponding portfolio returns $r_{PF,t} = \mathbf{w}'\boldsymbol{Y}$ are univariate GHyp distributed (see section 2.3). Thus, according to Paolella and Polak (2022) the VaR forecast can be calculated using the p-quantile function $Q_p(\cdot)$ of the corresponding conditional univariate GHyp distribution (in our case the univariate Variance Gamma distribution) of the portfolio returns:

$$\widehat{\mathrm{VaR}_t^p} = -Q_p(r_{PF,t}|\mathcal{F}_{t-1}) = -\inf_x\{x \in \mathbb{R} : \mathbb{P}(r_{PF,t} \leq x|\mathcal{F}_{t-1}) \geq p\}, \tag{3.3}$$

where $\mathcal{F}_{t-1} = \{\boldsymbol{Y_1}, ..., \boldsymbol{Y_{t-1}}\}$ is the information set at time $t-1$.

3.3 Backtesting

Backtesting is used to check whether the forecasts of the models evince some desirable properties. In doing so, we will follow the framework of Christoffersen (1998) which consists of three likelihood-ratio tests. Let I_t be the indicator variable for a VaR_t^p forecast made at time t-1, that is,

$$I_t = \mathbb{1}_{\{r_{PF,t} < -VaR_t^p\}} = \begin{cases} 1 & \text{if } r_{PF,t} < -VaR_t^p \\ 0 & \text{otherwise} \end{cases} .$$

According to Christoffersen (1998), a sequence of value at risk forecasts is then said to be efficient with respect to $\mathcal{F}_{t-1}$, the information set at time $t-1$, if

$$\mathbb{E}[I_t|\mathcal{F}_{t-1}] = \mathbb{E}[I_t|I_{t-1}, I_{t-2}, ..., I_1] = p, \quad t = 1, 2, ..., T. \tag{3.4}$$

Christoffersen (1998) shows that testing whether a sequence of value at risk forecasts is efficient is equivalent to testing that $\{I_t\}_{t=1}^T \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$. In case this property is fullfilled, the VaR forecast has correct conditional coverage. Although one could test for correct conditional coverage directly we will follow Paolella et al. (2006) and additionally provide intermediate test statistics which help to better understand deficiencies of the models.

First, we test for the correct number of exceedances under independence. The corresponding hypothesis of correct conditional coverage can be written as

$$H_0 : \mathbb{E}[I_t] = p \quad \text{versus} \quad H_A : \mathbb{E}[I_t] \neq p,$$

which is the result of applying the law of total expectation on equation 3.7. Let n_1 be the number of violations i.e. ones in the indicator sequence $\{I_t\}_{t=1}^T$ and n_0 the number of zeros in the indicator sequence. The likelihood-ratio test statistic then is

$$LR_{uc} = -2\log\left(\frac{L(p; I_1, I_2, ..., I_T)}{L(\hat{p}; I_1, I_2, ..., I_T)}\right) \overset{\text{asy}}{\sim} \chi_1^2, \tag{3.5}$$

where $\hat{p} = \hat{p}_{ML} = \frac{n_1}{n_0+n_1}$ is the maximum likelihood estimate of p and $L(\cdot)$ is the corresponding likelihood of a $\mathcal{B}in(n_0 + n_1, p)$ distribution (Christoffersen, 1998).

Through testing for unconditional coverage we verified the first part of correct conditional coverage, namely that $\{I_t\}_{t=1}^{T}$ is identically $Bernoulli(p)$ distributed for all time points. However, it might be that the coverage is satisfactory but all violations are clustered together. To detect these cases, we have to test whether the indicator sequence is also independently distributed. Christoffersen (1998) use the likelihood-ratio test of independence where independence is tested against a binary first-order Markov chain $\{I_t\}$ with transition probability matrix

$$\Pi_1 = \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix}$$

with the approximate likelihood function being

$$L(\Pi_1; I_2, ..., I_T | I_1) = (1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}^{n_{11}},$$

where $\pi_{ij} = \mathbb{P}(I_t = j | I_{t-1} = i)$ and n_{ij} is the number of transitions from state i to state j. The associated maximum likelihood estimates are $\hat{\pi}_{01} = \frac{n_{01}}{n_{00}+n_{01}}$ and $\hat{\pi}_{11} = \frac{n_{11}}{n_{10}+n_{11}}$. Under the assumption of independence, the outcome at time $t-1$ is irrelevant for time t and Π_1 becomes

$$\Pi_2 = \begin{bmatrix} 1-\pi_2 & \pi_2 \\ 1-\pi_2 & \pi_2 \end{bmatrix}$$

with maximum likelihood estimate $\hat{\pi}_2 = \frac{n_{01}+n_{11}}{n_{00}+n_{10}+n_{01}+n_{11}}$ (Christoffersen, 1998). Hence, under H_0 the likelihood is

$$L(\Pi_2; I_2, ..., I_T | I_1) = (1-\pi_2)^{(n_{00}+n_{10})} \pi_2^{(n_{01}+n_{11})}.$$

Therefore, the test statistic can be calculated as

$$LR_{ind} = -2 \log \left(\frac{L(\hat{\Pi}_2; I_2, ..., I_T | I_1)}{L(\hat{\Pi}_1; I_2, ..., I_T | I_1)} \right) \overset{\text{asy}}{\sim} \chi_1^2. \tag{3.6}$$

This statistic does not depend on the true violation probability p and thus only tests for independence and not coverage (Christoffersen, 1998).

By combining equations 3.5 and 3.6 we can directly test for conditional coverage through

$$LR_{cc} = -2\log\left(\frac{L(p; I_2, ..., I_T|I_1)}{L(\hat{\Pi}_1; I_2, ..., I_T|I_1)}\right) \overset{\text{asy}}{\sim} \chi^2_2. \tag{3.7}$$

By conditioning the test for unconditional coverage in equation 3.5 on the first observation we get that $\hat{p} = \hat{\pi}_2$. LR_{cc} can thus also be calculated as

$$LR_{cc} = LR_{uc} + LR_{ind}. \tag{3.8}$$

As mentioned above, we will make use of this identity and will calculate and analyse all three test statistics for our models.

3.4 Comparison of Predictive Ability

All the tests introduced so far only tested for a particular desirable characteristic. Based on these tests a model could either be rejected or not. However, we are also interested how the models perform relative to each other. A simple way to rank VaR estimates is by using a loss function. Like Santos et al. (2013) and Fortin et al. (2022) we will use the "tick" loss function:

$$L_V(\theta_t, r_{PF,t}) = (r_{PF,t} + \theta_t)(p - \mathbb{1}_{\{r_{PF,t} < -\theta_t\}}),$$

where θ_t is a non-random real variable. As indicated in Yu et al. (2003) this loss function is commonly used in quantile regressions and has the property that the p-quantile of our conditional portfolio return distribution at time t is

$$Q_p(r_{PF,t}|\mathcal{F}_{t-1}) = -\text{VaR}^p_t = \underset{\theta_t}{\text{argmin}}\, \mathbb{E}[L_V(\theta_t, r_{PF,t})]. \tag{3.9}$$

Moreover, it is linear and asymmetric because observations for which $\mathbb{1}_{\{r_{PF,t} < -\theta_t\}} = 1$ tend to receive a higher penalty than observations for which $\mathbb{1}_{\{r_{PF,t} < -\theta_t\}} = 0$ (Giacomini and Komunjer, 2005). Additionally, the factor $(r_{PF,t} + \theta_t)$ guarantees that, conditional on the value of the indicator function, observations for which θ_t strongly deviates from $r_{PF,t}$ are penalized more (Fortin et al., 2022). The property from equation 3.9 implies that we can rank our models based on their mean tick loss. The mean tick loss of a model is calculated as $\bar{L}_V = \frac{1}{T}\sum_{t=1}^{T} L_V(-\widehat{\text{VaR}}_t^p, r_{PF,t})$, where $\widehat{\text{VaR}}_t^p$ is the VaR_t^p estimate of the model. Whereby the lower $\bar{L}_V$ of a model, the higher its predictive accuracy and the better the model (Fortin et al., 2022). Hence, we have a framework which allows to rank the models based on differences in predictive accuracy. However, as mentioned in Bams et al. (2017), this ranking is not suited for inference.

Instead, we introduce the conditional predictive ability (CPA) test by Giacomini and White (2006) to be able to make a statement about the significance of those differences. Special about this test is that it allows for a pairwise comparison of forecasting models which are potentially misspecified (Giacomini and White, 2006). To simplify notation, let

$$\Delta L_{i,j,t} := L_{V_{i,t}} - L_{V_{j,t}}$$

be the loss differential between model i and j at time $t = 1, 2, ..., T$. The null hypothesis of equal conditional predictive ability can then be written as

$$\text{H}_0 : \mathbb{E}[\Delta L_{i,j,t} | \mathcal{F}_{t-1}] = 0 \quad \text{almost surely } \forall t$$

and entails that $\{\Delta L_{i,j.t}, \mathcal{F}_{t-1}\}_{t=1}^{T}$ is a martingale difference sequence (Giacomini and White, 2006). To test this hypothesis, Giacomini and White (2006) introduced the following Wald-type test statistic:

$$GW_{i,j} = T\bar{Z}'\widehat{\Omega}\bar{Z} \xrightarrow{d} \chi_q^2 \quad \text{as } T \to \infty \tag{3.10}$$

where $Z_t = h_{t-1}\Delta L_{i,j,t}$, $\bar{Z} = \frac{1}{T}\sum_{t=2}^{T} Z_t$ and $\widehat{\Omega} = \frac{1}{T}\sum_{t=2}^{T} Z_t Z_t'$ are all $q \times q$ matrices. h_{t-1} hereby denotes a $q \times 1$ $\mathcal{F}_{t-1}$-measurable vector. In accordance with Giacomini and White (2006) we choose $h_{t-1} = (1, \Delta L_{i,j,t})'$ and thus $q = 2$. Giacomini and White (2006) further show that the test statistic $GW_{i,j}$ follows asymptotically a χ_q^2 distribution under some regularity conditions and we thus reject H_0 at the 5% significance level if $GW_{i.j} > \chi_{2,0.95}^2$. In case of rejection, we choose model i over model j if its mean tick loss $\bar{L}_{V_i}$ is smaller than the mean tick loss $\bar{L}_{V_j}$ of model j. In other words, given a rejection, we choose model i over model j if the mean of the loss differential $\overline{\Delta L_{i,j}} := \frac{1}{T}\sum_{t=1}^{T} \Delta L_{i,j,t}$ is smaller than zero.

Chapter 4

Results

This chapter will summarise the results of our empirical analysis following the methods introduced in chapter 3. The results were obtained using Matlab (only for the COMFORT model class) and the statistical software R. Specifically, we used the "rugarch" (see Galanos (2022a)) to forecast the one-day-ahead VaR of all univariate models apart from the MixN(k)-GARCH models. For the MixN(k)-GARCH we used the "MSGARCH" package by Ardia et al. (2019). The "rmgarch" package by Galanos (2022b) was utilized for fitting a DCC-GARCH model to the factor returns. The estimation of the normal and t copula and the subsequent simulation from these copulas were done by the "copula" package of Hofert et al. (2023). For the skewed-t copula we used the R codes presented in Yoshiba (2014). The Matlab codes for the COMFORT model were provided by the supervisors of this thesis in collaboration with Prof. Dr. Pawel Polak (many thanks again to all of you who were involved). Finally, for the CPA test of Giacomini and White (2006) we translated parts of the Matlab code of the original authors into R and verified the correctness of our implementation using different loss functions and different models.

4.1 Value at Risk Backtests

The results of the backtests presented in section 3.3 are shown in table 4.1. Whilst all of the univariate models passed the likelihood ratio (LR) test of independence, only few showed adequate conditional or unconditional coverage. In particular, only the skewed-t GJR-GARCH passed the test of conditional coverage for both VaR levels. By contrast, the multivariate models seem to have no problems passing the LR tests of Christoffersen (1998) with the obvious exception being the multivariate normal (MN)-DCC-GARCH. Nevertheless, it is noticeable that using skewed-t NGARCH marginals in the framework of Fortin et al. (2022) leads to a very low percentage of violations when compared to the other factor copula models or the COMFORT models. Somewhat surprising is the revelation that even using a normal copula in combination with normal GARCH marginals for the factor returns appears to produce sound VaR estimates. Our explanation for this, which is based on figure 3.4 and table 3.2, is that a large part of the multivariate non-normality of the stock returns might be captured in the bootstrapped OLS residuals.

Table 4.1 – VaR Backtests

	p = 1%				p = 5%			
Model	%Viol.	LR_{uc}	LR_{ind}	LR_{cc}	%Viol.	LR_{uc}	LR_{ind}	LR_{cc}
Univariate:								
GARCH	2.547	0.000	0.886	0.000	6.565	0.004	0.282	0.009
EWMA	2.830	0.000	0.705	0.000	7.131	0.000	0.718	0.001
t GJR	1.811	0.002	0.277	0.005	6.848	0.001	0.622	0.003
Skewed-t GJR	1.132	0.585	0.498	0.685	6.169	0.029	0.459	0.071
Skewed-t NGARCH	1.528	0.038	0.360	0.077	6.282	0.017	0.403	0.041
MixN(2) GARCH	2.151	0.000	0.196	0.000	6.112	0.038	0.489	0.091
MixN(3) GARCH	1.585	0.023	0.342	0.048	6.565	0.004	0.375	0.010
MN-DCC-GARCH	3.339	0.000	0.983	0.000	6.904	0.000	0.873	0.002
Fortin et al.:								
Normal cop. NGARCH	0.509	0.022	0.761	0.070	4.131	0.084	0.573	0.193
Normal cop. GARCH	1.075	0.753	0.520	0.774	4.414	0.249	0.409	0.366
t cop. NGARCH	0.509	0.022	0.761	0.070	4.131	0.084	0.273	0.124
t cop. GARCH	0.962	0.872	0.565	0.837	4.414	0.249	0.409	0.366
Skewed-t cop. GARCH	0.962	0.872	0.565	0.837	4.471	0.299	0.799	0.565
COMFORT:								
MVG-CCC-GJR	0.736	0.241	0.661	0.457	4.924	0.883	0.409	0.703
MVG-CCC-GARCH	0.736	0.241	0.661	0.457	4.980	0.970	0.762	0.954

This table presents the percentage of violations and the p-values of the likelihood ratio (LR) tests for unconditional coverage, independence and conditional coverage by Christoffersen (1998) for the 1% and 5% one-day-ahead VaR. Cells for which the corresponding null hypothesis could not be rejected on the 5% significance level are highlighted in gray. The forecasts were made using a rolling window of size 1000 and cover the time period from December 28, 2004 to December 30, 2011 (1767 forecasts).

Additionally, we can see in table 4.1 that models that assume normality of the portfolio or the stock returns (i.e. GARCH, EWMA and MN-DCC-GARCH) have too many violations with none of said models showing adequate conditional or unconditional coverage. Furthermore, it is striking that for all of the univariate models as well as the multivariate normal (MN) DCC-GARCH the percentage of violations is higher than the corresponding 1% and 5% VaR level. For the COMFORT and factor copula models however, we observe (apart from one exception) fewer violations than expected. This phenomenon can be investigated closer using table 4.2. There it is evident that the number of violations of the multivariate, non MN-DCC-GARCH models tends to be closer to the expected number of exceedances for both VaR levels. The most adequate coverage is achieved by the factor copula-DCC models with GARCH marginals for the 1% VaR level and by the COMFORT models for the 5% level. Noteworthy is the factor skewed-t copula with GARCH marginals which belongs to the three models with the most appropriate coverage for both VaR percentiles.

Table 4.2 – VaR Exceedances over Time

	p = 1% (Expected: 17 Exceedances)									p = 5% (Expected: 88 Exceedances)								
Model	Total	2004	2005	2006	2007	2008	2009	2010	2011	Total	2004	2005	2006	2007	2008	2009	2010	2011
Univariate:																		
GARCH	45	0	2	4	9	13	4	7	6	116	0	14	8	22	26	15	13	18
EWMA	50	0	7	5	10	10	4	6	8	126	0	18	10	23	23	11	19	22
t GJR	32	0	2	3	7	8	3	6	3	121	0	15	9	20	26	18	17	16
Skewed-t GJR	20	0	2	2	5	4	2	3	2	109	0	14	9	19	19	16	16	16
Skewed-t NGARCH	27	0	2	4	6	6	1	5	3	111	0	14	8	19	25	13	17	15
MixN(2) GARCH	38	0	2	4	6	10	3	7	6	108	0	14	8	18	26	15	13	14
MixN(3) GARCH	28	0	1	3	3	9	2	6	4	116	0	12	8	17	39	14	11	15
MN-DCC-GARCH	59	0	8	8	19	12	2	5	5	122	0	21	12	31	26	6	12	14
Fortin et al.:																		
Normal cop. NGARCH	9	0	2	1	3	1	0	0	2	73	0	11	7	15	14	5	11	10
Normal cop. GARCH	**19**	0	2	1	6	5	0	2	3	78	0	13	7	17	16	5	10	10
t cop. NGARCH	9	0	2	1	3	1	0	0	2	73	0	12	6	15	14	5	11	10
t cop. GARCH	**17**	0	2	1	6	5	0	1	2	78	0	13	7	17	16	5	10	10
Skewed-t cop. GARCH	**17**	0	2	1	6	5	0	1	2	**79**	0	12	7	17	16	7	10	10
COMFORT:																		
MVG-CCC-GJR	13	0	0	0	2	6	0	3	2	**87**	0	1	4	15	29	13	12	13
MVG-CCC-GARCH	13	0	0	0	1	7	1	2	2	**88**	0	1	4	15	30	14	11	13

VaR exceedances observed in every year for the 1% and the 5% VaR level for all the models. For each VaR level the three closest numbers to the expected exceedences are in bold. The forecasts were made using a rolling window of size 1000 and cover the time period from December 28, 2004 to December 30, 2011 (1767 one-day-ahead VaR forecasts).

Another interesting discovery in table 4.2 is that almost all of the exceedances for the COMFORT models occurred during or after the financial crisis of 2008 despite passing the LR test of independence. Lastly, we observe that at the 1% VaR the models which assume a Gaussian distribution and do not account for the leverage effect (EWMA, GARCH and MN-DCC-GARCH) have approximately three times as many exceedances as one would expect. At the 5% VaR, this discrepancy is smaller and the choice of normality and symmetry in the GARCH structure therefore appears to have less detrimental consequences at this higher VaR level.

4.2 Conditional Predictive Ability Tests

In table 4.3 we rank the models according to their mean tick loss (see section 3.4). Interestingly, the univariate models perform well in terms of average tick loss in spite of their suboptimal backtesting results. Specifically, we see that the skewed-t GJR model had the lowest and the MN-DCC-GARCH the highest average loss for both VaR levels. Further, it is noticeable that the multivariate models achieve better ranks at the 1% level than at the 5% level. Thereby the average loss of the factor copula based models was lower at both levels than the average loss of the COMFORT models.

Table 4.3 – Average VaR Tick Loss Ranking

p = 1%		p = 5%	
Model	Mean Loss	Model	Mean Loss
Skewed-t GJR	0.03632	Skewed-t GJR	0.13456
t GJR	0.03691	t GJR	0.13662
Normal cop. GARCH	0.03692	Skewed-t NGARCH	0.13706
Skewed-t cop. GARCH	0.03702	MixN(2) GARCH	0.13716
t cop. GARCH	0.03716	Skewed-t cop. GARCH	0.13731
Skewed-t NGARCH	0.03819	t cop. GARCH	0.13749
Normal cop. NGARCH	0.03868	Normal cop. GARCH	0.13756
t cop. NGARCH	0.03898	t cop. NGARCH	0.13781
MixN(2) GARCH	0.04018	Normal cop. NGARCH	0.13791
MVG-CCC-GJR	0.04106	GARCH	0.13869
MVG-CCC-GARCH	0.04143	EWMA	0.13937
GARCH	0.04152	MVG-CCC-GJR	0.14438
EWMA	0.04238	MVG-CCC-GARCH	0.14468
MixN(3) GARCH	0.04307	MixN(3) GARCH	0.16518
MN-DCC-GARCH	0.05215	MN-DCC-GARCH	0.16950

Ranking of the average VaR tick loss (see section 3.4) of the models in ascending order for the VaR level 1% and 5%. Models starting with "MVG" refer to the COMFORT model class, models including "cop." refer to the factor copula-DCC-(N)GARCH models as specified in Fortin et al. (2022). one-day-ahead VaR forecasts).

Within the class of factor copula-DCC models, GARCH marginals achieved lower mean losses than their counterparts with skewed-t NGARCH marginals. This further reinforces our hypothesis that the bootstrapped OLS residuals account for a large part of the non-normality inherent in the stock returns. The COMFORT models however had lower average losses when GJR-GARCH processes instead of standard GARCH processes were used for the conditional variances. Finally, the MixN(2)-GARCH exhibits lower average losses than the more "jumpy" MixN(3)-GARCH which is in line with Slim et al. (2017).

We now turn our attention to the CPA test (see section 3.4) for inference. At each VaR level we conducted 105 CPA tests. Out of these 105 tests, 33 rejected the null-hypothesis of equal predictive ability on the 5% significance level at the 1% VaR level and 46 at the 5% VaR level. The p-values displayed in tables 4.4 and 4.5 are not corrected.

The p-values for the 1% VaR loss are shown in table 4.4. The first thing that strikes the eye is that the MN-DCC-GARCH is significantly outperformed by every other model. The corresponding p-values are all very close to zero. Other than that, we mostly have rejections in the univariate vs univariate or multivariate vs multivariate setting. Apart from the MN-DCC-GARCH, only the MVG-CCC-GJR and MVG-CCC-GARCH show significantly lower predictive accuracy than a univariate alternative in the form of the MixN(2)-GARCH. The majority of the rejections when comparing univariate models stem from the high predictive accuracy of the GJR-GARCH models. Specifically, the normal GARCH, EWMA and MixN(2)-GARCH are significantly outperformed by both the t GJR-GARCH and the skewed-t GJR-GARCH. For the multivariate models, we additionally discover that the factor copula-DCC models using skewed-t NGARCH marginals have significantly higher predictive ability than their counterparts with normal GARCH marginals. In addition, the normal copula is the superior choice for the skewed-t NGARCH marginals. However, for the normal GARCH marginals the t and skewed-t copula versions significantly exceed the performance of the normal copula. Thus, it is indispensable to incorporate the multivariate non-normality of the factor returns for the 1% VaR level. This result is in accordance with Christoffersen and Langlois (2013). Lastly, we find no significant improvement of using a skewed-t copula, which can account for the asymmetric tail dependence in the factor returns, over a Student t copula. Part of the reason for this might be that we had to reestimate the parameters less frequently for the skewed-t copula (see section 3.2).

Table 4.4 – CPA Test P-Values for 1% VaR Tick Loss

	EWMA	t GJR	Skewed-t GJR	Skewed-t NGARCH	MixN(2) GARCH	MixN(3) GARCH	MN-DCC GARCH	Normal cop. NGARCH	Normal cop. GARCH	t cop. NGARCH	t cop. GARCH	Skewed-t cop. GARCH	MVG-CCC GJR	MVG-CCC GARCH
GARCH	0.000←	0.010 ↑	0.015 ↑	0.077 ↑	0.019 ↑	0.267←	0.000←	0.346 ↑	0.127 ↑	0.387 ↑	0.158 ↑	0.140 ↑	0.780 ↑	0.857 ↑
EWMA		0.010 ↑	0.011 ↑	0.054 ↑	0.104 ↑	0.517←	0.000←	0.161 ↑	0.063 ↑	0.181 ↑	0.078 ↑	0.067 ↑	0.604 ↑	0.742 ↑
t GJR			0.024 ↑	0.446←	0.029←	0.074←	0.000←	0.363←	0.827←	0.298←	0.799←	0.851←	0.076←	0.070←
Skewed-t GJR				0.088←	0.023←	0.060←	0.000←	0.141←	0.834←	0.084←	0.778←	0.841←	0.083←	0.071←
Skewed-t NGARCH					0.304←	0.143←	0.000←	0.260←	0.528 ↑	0.221←	0.680 ↑	0.595 ↑	0.174←	0.129←
MixN(2) GARCH						0.066←	0.000←	0.182 ↑	0.171 ↑	0.182 ↑	0.189 ↑	0.168 ↑	0.015←	0.015←
MixN(3) GARCH							0.002←	0.296 ↑	0.193 ↑	0.315 ↑	0.216 ↑	0.200 ↑	0.555 ↑	0.685 ↑
MN-DCC GARCH								0.000 ↑	0.000 ↑	0.000 ↑	0.000 ↑	0.000 ↑	0.004 ↑	0.009 ↑
Normal cop. NGARCH									0.001 ↑	0.000←	0.004 ↑	0.000 ↑	0.182←	0.154←
Normal cop. GARCH										0.000←	0.011←	0.153←	0.111←	0.113←
t cop. NGARCH											0.001 ↑	0.000 ↑	0.183←	0.154←
t cop. GARCH												0.268 ↑	0.121←	0.123←
Skewed-t cop. GARCH													0.116←	0.117←
MVG-CCC GJR														0.492←

This table displays the p-values of the CPA test of Giacomini and White (2006) for the 1% VaR tick loss. The table is to be read column-wise. A left (up) error signifies that the row (column) outperforms the column (row). A gray cell color indicates that we reject the null hypothesis of equal predictive ability on the 5% significance level. Models starting with "MVG" refer to the COMFORT model class, models including "cop." refer to the factor copula-DCC-(N)GARCH models as specified in section 2.2.

For the p-values of the CPA tests at the 5% VaR level we refer to table 4.5. Once again, we immediately spot a plethora of rejections in columns and rows that involve the MN-DCC-GARCH. In addition, the MixN(3)-GARCH is also significantly outperformed by every other model but the MN-DCC-GARCH. For all of these rejections we obtained very small p-values. When focusing on tests that involve both a univariate and a multivariate model we notice a higher number of significant results compared to the 1% VaR level. Further, the Student t GJR-GARCH, skewed-t GJR-GARCH and the MixN(2)-GARCH all displayed significantly higher predictive ability than the the COMFORT models. Interestingly, the skewed-t GJR-GARCH, which significantly outperformed every other univariate model, did not manage to achieve significantly better VaR forecasts than the factor copula-DCC-(N)GARCH models. Only the MixN(2)-GARCH outperformed some of these factor copula based models but we fail to reject the null hypothesis for the Student t and the skewed-t Copula-DCC-GARCH model. Furthermore, when comparing the non MN-DCC-GARCH multivariate models, only two tests show significant p-values. First, in the setting with skewed-t NGARCH marginals a Student t copula yields significantly better VaR estimates than a normal copula. For the 1% VaR level it was the other way around. Second, the skewed-t copula-DCC-GARCH model again significantly outperforms the normal copula-DCC-GARCH model. Finally, at neither of the two VaR levels there were any significant differences in the predictive accuracy between the copula models and the COMFORT models.

Table 4.5 – CPA Test P-Values for 5% VaR Tick Loss

	EWMA	t GJR	Skewed-t GJR	Skewed-t NGARCH	MixN(2) GARCH	MixN(3) GARCH	MN-DCC GARCH	Normal cop. NGARCH	Normal cop. GARCH	t cop. NGARCH	t cop. GARCH	Skewed-t cop. GARCH	MVG-CCC GJR	MVG-CCC GARCH
GARCH	0.012←	0.129 ↑	0.006 ↑	0.236 ↑	0.093 ↑	0.000←	0.000←	0.109 ↑	0.141 ↑	0.106 ↑	0.140 ↑	0.179 ↑	0.142←	0.122←
EWMA		0.015 ↑	0.002 ↑	0.097 ↑	0.155 ↑	0.000←	0.000←	0.038 ↑	0.063 ↑	0.038 ↑	0.065 ↑	0.073 ↑	0.228←	0.194←
t GJR			0.010 ↑	0.174←	0.314←	0.000←	0.000←	0.082←	0.252←	0.083←	0.268←	0.328←	0.021←	0.018←
Skewed-t GJR				0.004←	0.021←	0.000←	0.000←	0.167←	0.352←	0.173←	0.371←	0.395←	0.006←	0.005←
Skewed-t NGARCH					0.774←	0.000←	0.000←	0.095←	0.434←	0.093←	0.453←	0.548←	0.066←	0.056←
MixN(2) GARCH						0.000←	0.000←	0.031←	0.049←	0.029←	0.050←	0.065←	0.028←	0.017←
MixN(3) GARCH							0.497←	0.001 ↑	0.001 ↑	0.001 ↑	0.001 ↑	0.000 ↑	0.005 ↑	0.006 ↑
MN-DCC GARCH								0.000 ↑	0.000 ↑	0.000 ↑	0.000 ↑	0.000 ↑	0.000 ↑	0.000 ↑
Normal cop. NGARCH									0.738 ↑	0.038 ↑	0.722 ↑	0.406 ↑	0.060←	0.056←
Normal cop. GARCH										0.750←	0.140 ↑	0.016 ↑	0.069←	0.061←
t cop. NGARCH											0.766 ↑	0.488 ↑	0.058←	0.054←
t cop. GARCH												0.077 ↑	0.067←	0.060←
Skewed-t cop. GARCH													0.066←	0.058←
MVG-CCC GJR														0.877←

This table displays the p-values of the CPA test of Giacomini and White (2006) for the 5% VaR tick loss. The table is to be read column-wise. A left (up) error signifies that the row (column) outperforms the column (row). A gray cell color indicates that we reject the null hypothesis of equal predictive ability on the 5% significance level. Models starting with "MVG" refer to the COMFORT model class, models including "cop." refer to the factor copula-DCC-(N)GARCH models as specified in section 2.2.

Chapter 5

Conclusion

We assessed the quality of VaR forecasts stemming from different univariate and multivariate models. We found that the majority of the univariate models produced inadequate VaR estimates that exhibited too many, albeit independent, violations. The only real exception was the GJR-GARCH with skewed-t innovations which managed to pass the likelihood ratio test of conditional coverage of Christoffersen (1998) at both VaR levels. Our multivariate models however all displayed adequate coverage with independently occurring VaR exceedances apart from our multivariate normal DCC-GARCH benchmark. This coincides with the discoveries that Santos et al. (2013) made in their Monte Carlo experiments. When comparing the VaR forecasts using conditional predictive ability (CPA) tests by Giacomini and White (2006) we found that the the MN-DCC-GARCH model is significantly outperformed by every other model. This is not surprising given the apparent multivariate non-normality of the stock returns and is in line with Paolella and Polak (2022). Other than that, most rejections of the null hypothesis of equal predictive ability that occurred in these bivariate CPA tests involved two univariate or two multivariate models. Lastly, only few significant results were observed in multivariate vs univariate tests and no general, significant outperformance in terms of predictive ability could be detected at at any of the two VaR levels we considered.

One interesting revelation we made is that using a higher data frequency (i.e. daily instead of weekly returns) makes the factor copula-DCC-NGARCH models by Fortin et al. (2022) feasible for forecasting the VaR of a portfolio of stocks. This is consistent with Kole et al. (2017) who found that the data frequency is of more importance for aggregated VaR forecasts than the model choice itself. Moreover, we found that replacing the skewed-t NGARCH marginals used in Fortin et al. (2022) by normal GARCH marginals for the factor returns increased the predictive accuracy of the forecasts and yielded more adequate unconditional coverage. One possible explanation we found was based on the observation that the OLS residuals of the Carhart four-factor model are highly non-normal. Thus, the majority of the multivariate non-normality of the stock returns might be captured by the OLS residuals and not by the factors themselves. Nevertheless, we discovered that a skewed-t copula with normal GARCH marginals for the factor returns yields better results in the backtests and the CPA tests than the normal copula and the Student t copula with the same marginals. This is in accordance with Christoffersen and Langlois (2013).

One possible explanation as to why our results do not entirely agree with Santos et al. (2013) is that contrary to this thesis, they did not include univariate models that can capture skewness. Additionally, the porfolios they analysed consisted of substantially more constituents. Hence, for further research, it would be interesting to see how the factor copula-DCC-GARCH model class by Fortin et al. (2022) performs in a higher dimensional portfolio setting. Using a much larger portfolio would also highlight the advantage of the dimensionality reduction inherent in this model class since the only really time-consuming parts are the fitting of the DCC-GARCH structure (in particular the univariate GARCH models) and the fitting of the copula. But both of these computationally expensive tasks only depend on the number of factors and would therefore be unaffected by an increase in the size of the portfolio. However, one potential drawback of this model specification is that the Carhart factors are based on the CRSP value-weighted index. Therefore, there may be some problems when considering a realistic and internationally diversified portfolio.

Appendices

Appendix A

Proofs

In the following, we will assume that the cdf of the random variable in question is invertible. For a proof that uses the concept of the generalized inverse of the distribution function we refer to McNeil et al. (2015, chap. 7).

A.1 Probability Integral Transform

For the sake of simplicity, we will only consider the case where X is a continuous random variable with an invertible cdf $F_X(\cdot)$. Then, the random variable $Y := F_X(X)$ follows a standard uniform distribution (i.e. $Y \sim \mathcal{U}(0,1)$).

Proof. For all $y \in [0,1]$ we have that

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(F_X(X) \leq y) \\ &= \mathbb{P}(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) \\ &= y \end{aligned}$$

This is just the cdf of a standard uniform distribution and hence $Y = F_X(X) \sim \mathcal{U}(0,1)$. $\square$

A.2 Quantile Transform

Let $U \sim \mathcal{U}(0,1)$ and let X for simplicity's sake once again be a continuous random variables with an invertible cdf $F_X(\cdot)$. Then, $F_X^{-1}(U)$ has the same distribution as X where $F_X^{-1}(\cdot)$ is the inverse of the cdf $F_X(\cdot)$ which in our framework the same as the quantile function.

Proof. We want to find some invertible function $T : [0,1] \to \mathbb{R}$ such that $T(U) \stackrel{d}{=} X$. Thus, for all $x \in \mathbb{R}$ we want

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq x) = \mathbb{P}(T(U) \leq x) \\ &= \mathbb{P}(U \leq T^{-1}(x)) = F_U(T^{-1}(x)) \\ &= T^{-1}(x) \end{aligned}$$

where $F_U(\cdot)$ is the cdf of a standard uniform variable. This is equivalent to saying that $T(x) = F_X^{-1}(x)$ i.e. our desired transformation function is the quantile function. □

The practical implication of this statement is that we can generate random samples from some distribution $F_X(\cdot)$ by first sampling from a standard uniform distribution and then applying the transformation $T(\cdot) = F_X^{-1}(\cdot)$ to these standard uniform samples.

Appendix B

Copulas

Both the Gaussian and the Student t copula belong to the subclass of implicit copulas meaning that they are obtained through a known cdf via Sklar's theorem. Other copula classes are the explicit copulas and the fundamental copulas. For a more thorough presentation of copulas and their applications in quantitative risk management we refer to McNeil et al. (2015, chap. 7).

B.1 Gaussian Copula

Let $\boldsymbol{X} \sim \mathcal{N}_n(\mathbf{0}, \boldsymbol{R})$ be multivariate normal distributed with a $n \times n$ correlation matrix $\boldsymbol{R}$. The corresponding Gaussian copula is then defined as

$$\boldsymbol{C}_{\boldsymbol{R}}^{\text{Gauss}}(\boldsymbol{u}) = \boldsymbol{\Phi}_{\boldsymbol{R}}(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)) = \mathbb{P}\left(\Phi(X_1) \leq u_1, ..., \Phi(X_n) \leq u_n\right),$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution, that is,

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du,$$

and $\boldsymbol{\Phi}_{\boldsymbol{R}}(\cdot)$ is the joint distribution function of $\boldsymbol{X}$.

B.2 Student t Copula

Let $\boldsymbol{X} \sim t_n(\nu, \boldsymbol{0}, \boldsymbol{R})$ be multivariate Student t distributed with a $n \times n$ correlation matrix $\boldsymbol{R}$ and $\nu > 0$ degrees of freedom. The corresponding Student t copula is

$$\boldsymbol{C}^t_{\nu,\boldsymbol{R}}(\boldsymbol{u}) = \boldsymbol{t}_{\nu,\boldsymbol{R}}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_n)) = \mathbb{P}\left(t_\nu(X_1) \leq u_1, ..., t_\nu(X_n) \leq u_n\right),$$

where $t_\nu(\cdot)$ denotes the cdf of a univariate Student t distribution with $\nu > 0$ degrees of freedom, that is,

$$t_\nu(x) = \int_{-\infty}^{x} \frac{\Gamma(\frac{\nu+1}{2})\nu^{\frac{\nu}{2}}}{\sqrt{\pi}\,\Gamma(\frac{\nu}{2})}(\nu + u^2)^{-\frac{\nu+1}{2}} du,$$

and $\boldsymbol{t}_{\nu,\boldsymbol{R}}(\cdot)$ is the joint distribution of $\boldsymbol{X}$.

Appendix C

Distributions

C.1 Finite Normal Mixtures

Let $X \sim \text{Mix}_k\text{N}(p_1, ..., p_k, \mu_1, ..., \mu_k, \sigma_1^2, ..., \sigma_k^2)$ be a random variable that follows a normal mixture distribution with k components, where k is finite. Each of these k components is a Gaussian random variable on its own i.e. $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Moreover, every X_i exhibits a pdf

$$f_{X_i}(x; \mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2}(x-\mu_i)^2\right), \quad i = 1, ..., k.$$

The pdf of the random variable X is then given by

$$f_X(x; \boldsymbol{\mu}, \boldsymbol{\sigma^2}, \boldsymbol{p}) = \sum_{i=1}^{k} p_i f_{X_i}(x), \quad \sum_{i=1}^{k} p_i = 1, \quad p_i \in (0,1) \quad \forall i,$$

where $\boldsymbol{\mu} = (\mu_1, ..., \mu_k)'$, $\boldsymbol{\sigma^2} = (\sigma_1^2, ..., \sigma_k^2)'$ and $\boldsymbol{p} = (p_1, ..., p_k)'$ is the vector of mixing weights. The corresponding cdf of X is given by

$$F_X(x; \boldsymbol{\mu}, \boldsymbol{\sigma^2}, \boldsymbol{p}) = \sum_{i=1}^{k} p_i F_{X_i}(x), \quad \sum_{i=1}^{k} p_i = 1, \quad p_i \in (0,1) \quad \forall i,$$

where $F_{X_i}(x) = \int_{-\infty}^{x} f_{X_i}(u)du$ is the cdf of the i'th component.

Bibliography

G. P. Aielli. Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299, 2013.

C. Alexander and E. Lazar. Normal mixture garch (1, 1): Applications to exchange rate modelling. *Journal of Applied Econometrics*, 21(3):307–336, 2006.

D. Ardia, K. Bluteau, K. Boudt, L. Catania, and D.-A. Trottier. Markov-switching garch models in r: The msgarch package. *Journal of Statistical Software*, 91:1–38, 2019.

D. Bams, G. Blanchard, and T. Lehnert. Volatility measures and value-at-risk. *International Journal of Forecasting*, 33(4):848–863, 2017.

O. Barndorff-Nielsen, P. Blaesild, and V. Seshadri. Multivariate distributions with generalized inverse gaussian marginals, and associated poisson mixtures. *The Canadian Journal of Statistics*, 20:109–120, 1992.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

T. Bollerslev. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *The Review of Economics and Statistics*, 72(3):498–505, 1990.

T. Bollerslev. Glossary to arch (garch). *CREATES Research paper*, 49, 2008.

M. M. Carhart. On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82, 1997.

P. Christoffersen. Value–at–risk models. In T. Mikosch, J.-P. Kreiß, R. A. Davis, and T. G. Andersen, editors, *Handbook of financial time series*, pages 753–766. Springer Berlin, Heidelberg, 1 edition, 2009.

P. Christoffersen. *Elements of Financial Risk Management*. Academic press, 2011.

P. Christoffersen and H. Langlois. The joint dynamics of equity market factors. *Journal of Financial and Quantitative Analysis*, 48(5):1371–1404, 2013.

P. F. Christoffersen. Evaluating interval forecasts. *International Economic Review*, 39(4):841–862, 1998.

S. Demarta and A. J. McNeil. The t copula and related copulas. *International Statistical Review*, 73(1):111–129, 2005.

C. Diks and H. Fang. Comparing density forecasts in a risk management context. *International Journal of Forecasting*, 36(2):531–551, 2020.

P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj. An academic response to basel 3.5. *Risks*, 2(1):25–48, 2014.

R. Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

R. F. Engle and V. K. Ng. Measuring and testing the impact of news on volatility. *The Journal of Finance*, 48(5):1749–1778, 1993.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.

A.-P. Fortin, J.-G. Simonato, and G. Dionne. Forecasting expected shortfall: Should we use a multivariate model for stock market factors? *International Journal of Forecasting*, 39, 01 2022.

A. Galanos. *rugarch: Univariate GARCH Models*, 2022a. R package version 1.4-9.

A. Galanos. *rmgarch: Multivariate GARCH models*, 2022b. R package version 1.3-9.

R. Giacomini and I. Komunjer. Evaluation and combination of conditional quantile forecasts. *Journal of Business & Economic Statistics*, 23(4):416–431, 2005.

R. Giacomini and H. White. Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578, 2006.

L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801, 1993.

M. Haas, S. Mittnik, and M. S. Paolella. Mixed normal conditional heteroskedasticity. *Journal of Financial Econometrics*, 2(2):211–250, 2004.

B. E. Hansen. Autoregressive conditional density estimation. *International Economic Review*, 35(3): 705–730, 1994.

A. Heinen and A. Valdesogo. *Copula-Based Volatility Models*, chapter Twelve, pages 293–316. John Wiley Sons, Ltd, 2012.

M. Hofert, I. Kojadinovic, M. Maechler, and J. Yan. *copula: Multivariate Dependence with Copulas*, 2023. R package version 1.1-2.

J.P. Morgan. *RiskMetrics Technical Document.* Morgan Guaranty Trust Company, New York, N.Y, 4th ed. edition, 1996.

E. Kole, T. Markwat, A. Opschoor, and D. Van Dijk. Forecasting value-at-risk under temporal and portfolio aggregation. *Journal of Financial Econometrics*, 15(4):649–677, 2017.

A. Lucas and X. Zhang. Score-driven exponentially weighted moving averages and value-at-risk forecasting. *International Journal of Forecasting*, 32(2):293–302, 2016.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk management: Concepts, Techniques and Tools – Revised Edition.* Princeton University Press, 2015.

M. S. Paolella. *Linear Models and Time-Series Analysis: Regression, ANOVA, ARMA and GARCH.* John Wiley & Sons, 2018.

M. S. Paolella and P. Polak. Comfort: A common market factor non-gaussian returns model. *Journal of Econometrics*, 187(2):593–605, 2015.

M. S. Paolella and P. Polak. Density and risk prediction with non-gaussian comfort models. *Swiss Finance Institute Research Paper*, (22-88), 2022.

M. S. Paolella, K. Kuester, and S. Mittnik. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89, 2006.

A. A. P. Santos, F. J. Nogales, and E. Ruiz. Comparing univariate and multivariate models to forecast portfolio value-at-risk. *Journal of Financial Econometrics*, 11(2):400–441, 2013.

S. Slim, Y. Koubaa, and A. BenSaïda. Value-at-risk under lévy garch models: Evidence from global stock markets. *Journal of International Financial Markets, Institutions and Money*, 46:30–53, 2017.

T. Yoshiba. Maximum likelihood estimation of skew-t copula. *Bank of Japan, The Institute of Statistical Mathematics*, 2014.

K. Yu, Z. Lu, and J. Stander. Quantile regression: Applications and current research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(3):331–350, 2003.

Eidesstattliche Erklärung

Der/Die Verfasser/in erklärt an Eides statt, dass er/sie die vorliegende Arbeit selbständig, ohne fremde Hilfe und ohne Benutzung anderer als die angegebenen Hilfsmittel angefertigt hat. Die aus fremden Quellen (einschliesslich elektronischer Quellen) direkt oder indirekt übernommenen Gedanken sind ausnahmslos als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden.

Uster, 15.02.2023 .... J. Schlegel ....

Ort, Datum Unterschrift des/der Verfassers/in