Erasmus University Rotterdam

Master of Science in Quantitative Finance

Out of Sample Performance of Factor Copulas

Author: Frederik Campagne

Studentnumber: 303015

Supervisor: Dick van Dijk

Co-reader: Erik Cole

Abstract

This paper investigates the effects of several different copulas in combination with different data frequencies on risk forecasts, such as 1% 10-day Value at Risk, for a stock portfolio. The copulas enable a multivariate distribution to be defined as a function of marginals and a dependence structure, the copula. The marginals are estimated using the normal distribution and for the copulas the Normal, Student-t, Clayton and factor copula are chosen. The factor copula is a newly proposed copula by Oh and Patton (2012) based on a new estimation method and supposedly incorporates more characteristics than the other copulas such as asymmetric dependence and scalability. The 1% 10-day VaR is best modeled using the Student Copula as the thick tails capture the extremeties. The most accurate 1% VaR forecasts are obtained by using high frequency data in combination with the Student-t copula. The best 5% VaR forecasts are also obtained using high frequency data with a Student-t copula. The factor copula turns out to be a hard to estimate copula that depends on large simulations in order to maintain accuracy, which becomes a computational burden.

Keywords: copula, value at risk, dependence, forecast

Table of Contents

1. Introduction

The dent, that the credit crisis has left in the trust and beliefs of investors, caused a complete game change in the financial world. Whereas an investor might have been looking for high returns with a reasonable risk before the crisis; the investor now aspires to find low risk with a fair return. This change was due to the fact that, before the credit crisis, there was less reason to fear the risk of a total collapse. It was believed that if one diversified their portfolio enough, risk reduction would be adequate. However, a simple diversification is based on the fact that returns of assets are correlated and a portfolio is created by the a choice of assets that have both positive correlation and negative correlation (when assets crash, negatively correlated assets move in opposite direction). In theory, a portfolio wouldn't suddenly crash completely and you will have reduced your risk. However, among some there was an underestimation of the shift that correlations between assets can make (negative correlations turning into positive). Hence, the moment assets started crashing simultaneously, people observed that dependences between assets shifted severely and are an important aspect in portfolio estimation. Thus, there was a major shift towards a more risk aware environment. One of the unexpected phenomena was that assets that had been thought to move independently suddenly started moving together. Correlation-based models failed to incorporate the possibility of this happening, because there was often an assumption of normality and this carries the risk of neglecting tail dependence or in other words neglect the possibility that crashes (huge negative returns) are correlated between assets. Thus, in recent years, more and more extensions on existing models have been developed in order to capture this tail dependence and other desired features, such as symmetric and asymmetric dependence, in a valid model.

One of the widely discussed and used set of models is the copula, which shows great potential for capturing the above mentioned aspects that are desirable in a model for financial assets. A copula is simply another way to describe a dependency between two or more variables. A traditional way is to have a multivariate distribution with a dependence structure that is implicit, while a copula allows us to split a multivariate distribution into marginal distributions and a dependence structure. This relatively new way of describing and estimating a set of financial assets opened up a whole new world of possibilities for portfolio evaluation. Whilst the existing copulas suffer from either the curse of dimensionality (large portfolio's means an uncontrollable increase of parameters) or are simply too simplistic, Oh and Patton (2012) propose a whole new set of copulas, named the factor copula. These copulas offer the possibility of positive and negative dependence, asymmetric dependence and non-zero tail dependence. Plus, Oh and Patton (2012) claim that the factor copula can handle a high dimension of variables, up to a set of 100 variables. This makes the factor copula a very interesting copula for portfolio management as portfolio's may contain a large set of financial assets. Oh and Patton (2012) also introduce a new type of estimation in order to estimate the parameters of the newly proposed factor copula, which is similar to the simulation method of moment. The model depends on a new estimation method since there isn't any use of 'moments' in the estimation process as the estimation method would suggest, but dependence measures are used instead. The accuracy is satisfactory, especially when a large number of assets is used. Hence, it will be interesting to see whether this accuracy can be maintained throughout different scenario's and whether this leads to an improved method for dependence structures.

The importance of these models lies in the ability to create risk forecasts for the Value at Risk. Thus, a superior model would be of great value in the financial world. In the world of finance, Value at Risk (VaR) is a widely used risk measure for a portfolio of financial assets. In words, VaR is the loss we are fairly sure that will be not exceeded when holding a portfolio for a certain amount of time. Normally VaR is expressed by means of a confidence level (α%), a specified period and an estimated loss in absolute terms. Hence, one could read a 1% 2-day VaR as a loss that will not be exceeded with a confidence of 99% over the next 2 days for a specified portfolio. Hence, in 1% of the cases, it is expected that this threshold will be exceeded.

The Basel Committee incorporated VaR in combination with a capital requirement. The VaR partly determines the amount of capital an institution must have to suffice to this requirement. Hence, institutions have to comply with a daily report of the biweekly 1% VaR. Since financial institutions could therefore benefit from a low reported VaR, these are checked for violations by the committee and fined if these indeed are violated. This way, a financial institution is regulated and prevents them from taking immense risks.

This paper makes two primary contributions. Firstly, this paper will examine the out-of-sample performance and hence, the usefulness of the proposed factor copulas in terms of risk forecasts (e.g. VaR) and compares these to other copulas. In the literature, the evidence is overwhelming that out-of-sample estimates of supposedly superior models are usually quite disappointing. A simplistic benchmark, such as the Normal Copula, is usually hard to beat. In this out-of-sample study a selection of 10 stocks of the S & P 100 are used for the estimation and for the risk forecasts.

Secondly, the new dependence structure (factor copula) in Oh and Patton (2012) will be examined in more depth. This paper will examine whether different data frequencies influence the accuracy of the estimation method significantly and hence, examine whether this dependence structure is vastly applicable in the finance sector (where an investor might use weekly data). Also, in this out-of-sample study, different forecast horizons with different data frequencies will be examined in order to determine whether the factor copula outperforms in an absolute sense (in every case) and whether it could be of use when considering the biweekly 1% VaR described above. This will also include data frequencies that are higher than the forecast horizon. This could mean incorporating information/characteristics that would otherwise be lost and therefore maybe enhance the forecasts. The use of different forecast horizons and different data frequencies could prove to be useful when looking at the performance of a copula from an economic standpoint and the implications it could have on the biweekly 1% VaR of the portfolio of the 10 stocks of the S & P 100.

In this paper a selection of copulas will be used: the Normal, T, the Clayton copula and the proposed factor copulas. The various copulas are incorporated as they can capture some, but not all, of the desired features that are evident in financial time series and might serve as a good benchmark for the factor copulas. The factor copulas supposedly capture all desired features and might prove to enhance risk forecasts. Hence, if even the Normal Copula outperforms other models, will the factor copula prove to be a superior, useful tool?

The rest of this paper will be divided as follows. In the next section there will be a review of the literature. It will serve as a review for what have been the building blocks for this paper. In section 3 the methodology with all the used models will be specified. All parameters, estimation, model-specifics and forecasting will be further explained. Then, in section 4, an elaborate analysis of the data (daily data of the Standard and Poor's 100 index) will be performed. In section 5 the results that were procured from the implementation of the models and forecasts will be discussed. The main focus will be determining whether the factor copula presents estimation problems and testing the superiority or usefulness in terms of risk forecasts in comparison to other respectable models. The VaR (at a 1% and 5% level) and another measure of risk will be compared for every econometric model, data frequency and forecast horizon. This will be concluded in section 6. The paper ends with a brief discussion on questions that arise from the results and more research that can be done.

2. Literature

Researchers have been exploring the vast amount of possibilities when it comes to the copula dependence structure. Their goal has been to construct an ideal copula that will be able to capture the empirically found features of financial assets. Patton (2012) summarizes these in the ability to capture positive and negative dependence, symmetric and asymmetric dependence and the possibility of non-zero tail dependence. On top of that a copula could also possess another attribute: scalability. This would mean that a copula would be able to incorporate a large set of variables (more than 50 variables). While many capture some of these features, this paper will examine whether a copula that possesses all features will show a better fit and superior risk forecasts.

The first development of a copula dependence structure is the well known Normal Copula, which has the assumption of Gaussianity. The Normal Copula, see Li (2000) and many others, has been empirically proven to be a misspecification for financial assets, because there is an assumption of zero tail dependence and symmetric dependence between booms and crashes. The Student-t copula allows for non zero tail dependence, but still has symmetric tail dependence. The Clayton copula does account for tail dependence, but only captures the lower tail dependence. Plus, this copula is very restrictive for higher dimensions as it assumes that the dependence between all assets is equal. This does result in an easy computation as it only has one parameter. Genest and Rivest (1989) and other papers discuss the advantages and disadvantages in further detail. Demarta and McNeil (2005) discuss the t copula and a variety of extensions. Daul et al. (2003) proposes an interesting "grouped t" copula. It allows for non zero tail dependence and high dimensions up to 100 variables, but it has an assumption of equal upper and lower tail dependence, which is empirically rejected in Oh and Patton (2012) for equity returns. More complicated copulas have also been constructed such as "vine" copulas; see Aas et al. (2009). These are high dimensional copulas that are created via sequential application of bivariate copulas. However, vine copulas have had quite some critique, due to their complexity; see Acar et al. (2012). Oh and Patton (2012) propose a factor copula, which is supposed to captures all features including scalability. This is therefore the focus of this paper and will be compared in terms of risk forecasts to much simpler models such as the Normal Copula. Oh and Patton (2012) do notice that the factor copulas mainly outperform other models with high dimensions. Comparisons for higher dimensions (read: higher than a dimension of 10) are quite time-consuming and won't be part of this paper.

In this paper risk forecasts are compared between different models, data frequencies and dimensions. Markwat, Kole and Van Dijk (2010) find that, for stock and bond portfolios, higher data frequencies result in improvements for the VaR forecasts on a 1% confidence level. Also, they find that a bivariate dataset increases the forecast accuracy in terms of VaR compared to a univariate dataset. However, they find that lower data frequencies and a univariate dataset contribute to a better 5% VaR forecast. Whether this applies to datasets with higher dimensions, like in this paper, is to be seen. With higher data frequency, multiple periods ahead have to be forecasted in order to obtain the same forecast horizon. Ghysels et al. (2009) and Diebold et al. (1997) have analyzed various methods in order to maintain multiple period ahead forecasts for stock market return volatilities. They find that the direct approach gives poor results. They also look into the scaling method which also produces inaccurate results, due to the fact that scaling blows up the volatility fluctuations whereas they should be significantly reduced. However, they find that the iterating gives sufficient results. Also, the mixed data sampling method returns good results, but is only useful when looking at forecast horizons larger than two weeks. Furthermore, Santos et al (2012) conclude that when considering a VaR forecast for a portfolio that consists of a linear combination of stocks (including the S&P 100 for example), it would be useful to incorporate a multivariate model that considers the joints dynamics of the assets (e.g. a model that takes asset correlations into account). Especially the incorporation of dynamic correlations and asymmetric effects would lead to improved VaR forecasts.

This paper will compare a selection of these advanced dependence structures. The literature suggests that there is an ideal dependence structure that can capture all features of financial assets and therefore could result in improved accuracy in risk forecasting. The aim of this paper is to see if a highly complex dependence structure will actually prove to be useful in terms of measures of risk.

3. Methodology

The objective of this paper is to compare and analyze the differences in dependence structures in terms of risk forecasts, with a particular interest in the factor copula. In order to compare these type of risk forecasts we shall need to estimate the different dependence structures, which in combination with volatility forecasts will form the forecasts for the returns. From these forecasted returns the portfolio VaR can be constructed and hence, can be compared. These steps will be explained in the following sections below. To avoid ambiguity, certain terminology will be used, which will be clarified below.

Let $r_{p,t,k}$ and $r_{a,t,k}$ denote the log return for the portfolio p and for asset a, respectively. The log return is defined as following:

$$r_{a,t,k} = \ln\left(\frac{P_{a,t}}{P_{a,t-k}}\right) \text{ (1)}$$

Where P is the spot price of asset a (or portfolio p), t is the time, k denotes the frequency of the data. Here, k=5 would denote weekly data and k=1 would denote daily data. Note that the returns created here are non-overlapping, so that there isn't any structural serial correlation. The portfolio return, h-days ahead, is then obtained by:

$$r_{p,t+h} = \sum_{i=1}^{h/k} \sum_{a=1}^{A} w_a r_{a,t+i*k,k} \text{ (2)}$$

w_a is the weight that is assigned to asset a, h is the forecast horizon and A is the total number of assets. In this paper an equally weighted portfolio will be created, because the optimal diversification of portfolio weights is not the focus of this paper. The portfolio returns are not completely accurate, because we simply take the sum of the weighted returns. However, the short time intervals of the time series make the deviations negligible.

The aim is to investigate which dependence structure captures the dependency between financial assets the best and thus, investigate which structure would lead to the most accurate VaR. The question remains whether the factor copula will be the dependence structure that prevails. In the sub sections below all the estimations of the models, forecasting methods and risk forecasts will be further explained. After an explanation of the copula structure and tail dependence, this paper will first look at the traditional copulas and display the features of these models including which features it might lack. Finally, the factor copula will be explained. This way the paper builds up to the supposedly ideal copula structure.

3.1 Copula

Before diving into the estimation and forecasting of the models, a quick review is given of the copula structure so there is a full understanding of the concept. A copula is another way to describe a dependency between two or more variables. A traditional way is to have a multivariate distribution with a dependence structure that is implicit, while a copula allows us to split a multivariate distribution into marginal distributions and a dependence structure. This allows a flexible approach to multivariate dependence modeling as we can now specify different marginal distributions for every asset. The multivariate distribution can now be written as a copula with the marginal distributions as input. Let X be a vector with M variables with a joint distribution F, marginal distributions F_i, and copula C:

$$[x_1, \dots, x_m]' = X \sim F(x_1, \dots, x_m) = C(F_1(x_1), \dots, F_m(x_m)) \quad (3)$$

where $F(x_1, \dots, x_m)$ is the joint distribution function, $C(F_1(x_1) \dots F_m(x_m))$ is the copula density function and $F_1(x_1)$is the cdf of a variable x_1 . A m dimensional copula is simply a function that maps m marginals to a joint distribution, which is the copula, on the interval [0,1]. This decomposition of the copula and marginals instead of one distribution function has two main advantages. Firstly, it generally enables a multi-stage estimation, which is useful in cases of high dimensions. Secondly, it enables a researcher to specify the marginals and the copula separately. With the wide range of literature available on univariate distributions, the remaining challenge is finding a suitable dependence structure. For a further and more detailed description of the copula theorem, see Sklar (1959).

3.1.1 Tail Dependence

A common way for examining the way a copula models the dependence between two variables is examining the tail dependence. In general we speak of lower tail dependence and upper tail dependence. We define the lower tail dependence between variable X_i and X_j as follows:

$$\lambda_{ij}^l = \lim_{q \to 0} P\left(X_i < F_i^{-1}(q) | X_j < F_j^{-1}(q)\right) \quad (4)$$

Where λ_{ij}^l is the lower tail dependence coefficient, $F_i^{-1}(q)$ is the inverse disitribution function and q is the quantile. Lower tail dependence measures the probability that X_i is below a certain quantile given that X_j is also below that quantile. Hence, a copula is said to have lower tail dependence if $\lambda_{ij}^l > 0$ and the higher the coefficient, the stronger the lower tail dependence is. Also note that:

$$\lambda_{ij}^l = \lim_{q \to 0} \frac{C(q,q)}{q}$$

The numerator within the scope of the limit is equal to the copula with q as input parameter.

Similarly, the upper tail dependence is given by:

$$\lambda_{ij}^u = \lim_{q \to 1} P\left(X_i > F_i^{-1}(q) | X_j > F_j^{-1}(q)\right) \quad (5)$$

$$\lambda_{ij}^u = \lim_{q \to 0} \frac{1 - 2q + C(q,q)}{1 - q}$$

Upper tail dependence measures the probability that X_i is above a certain quantile given that X_j is also above that quantile. It is said that a copula has symmetric dependence when λ_{ij}^{u}=λ_{ij}^{l}. When the upper and lower tail dependence differ, it is said that the copula captures asymmetric tail dependence. Information about asymmetric features in time series can prove to be quite important, as some copulas impose symmetric tail dependence such as the Normal and Student-t copula.

3.1.2 Normal Copula

The first and very frequently used copula is the Normal copula. It is based on the multivariate normal distribution as the name suggests. The correlation matrix contains the copulas' parameters, which makes it a popular choice in financial analysis due to that fact that correlation plays a central role in financial analysis. However, a wide selection of papers shows that this is also due to convenience rather than accuracy. The Gaussian distribution function is defined as follows:

$$C(u_1, \dots, u_m; \Sigma) = \boldsymbol{\Phi}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_m)) \quad (6)$$

Where $\boldsymbol{\Phi}$ and Φ are the n-dimensional multivariate and univariate standard normal distribution functions, respectively, and Σ is the correlation matrix. U is a vector with uniformly distributed random variables, which are obtained by the transformation $u_i = F_i(z_i)$, where z_i are standardized residuals.

However, it is impossible to write this copula in closed form and can only be expressed as an integral. Unlike the copula distribution, the copula density function can be expressed in closed form and is therefore useful when it comes to estimating the parameter of the Gaussian copula. The Gaussian copula density is obtained by differentiating equation (6) and gives:

$$c(u_1, \dots, u_m; \Sigma) = |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\xi'(\Sigma^{-1} - \mathrm{I})\xi\right) \quad (7)$$

Where $|\Sigma|$ is the determinant of the correlation matrix, ξ is defined as $\xi = \left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_m)\right)'$ and I is the identity matrix.

Unfortunately, the Gaussian copula is not an appropriate model for modeling dependency between financial assets. It is a symmetric copula and has no tail dependence unless the correlation is 1. Also, it cannot capture asymmetric tail dependence. Nonetheless, it could serve as a fair benchmark, because in practice simplistic models do not prove to perform under par .

To give an idea of what to expect from this copula in terms of VaR forecasts (a better capture of dependence is thought to result in better forecasts), a scatter plot of a bivariate simulation is shown below. The scatter plot shows simulations from the four copulas, all with N(0,1) marginals and linear correlation of 0.5.

[Insert figure 1 here]

The Normal copula shows a symmetric dependence and absolutely no tail dependence, meaning that a large negative return for one variable does not necessarily mean a higher probability of a large return for the other variable. For financial assets, this typically leads to smaller VaR forecasts compared to other dependence structures.

3.1.3 Student-t Copula.

Just like the Gaussian copula, the student-t copula is derived implicitly from a multivariate distribution. It is based on the multivariate student-t distribution and is defined as:

$$C_v(u_1, \dots, u_m; \Sigma) = \mathbf{t_v}(\mathrm{t_v}^{-1}(u_1), \dots, \mathrm{t_v}^{-1}(u_m)) \quad (8)$$

Where $\mathbf{t_v}$ and $\mathrm{t_v}$ are the multivariate and univariate Student-t distribution functions, respectively, v is the degrees of freedom and Σ is again the correlation matrix.

Similarly to the normal copula, the Student-t copula cannot be written in closed form and the copula density function will be used for estimation. The copula density function of the Student-t copula is obtained the same way we obtain the normal copula density function and is defined as follows:

$$c_v(u_1, \dots, u_m; \Sigma) = K|\Sigma|^{-\frac{1}{2}}(1 + \nu^{-1}\xi'\Sigma^{-1}\xi)^{-(\nu+m)/2} \prod_{i=1}^{m}(1 + \nu^{-1}\xi_i^2)^{(\nu+1)/2} \quad (9)$$

$$\xi = \left(\mathrm{t_v}^{-1}(u_1), \dots, \mathrm{t_v}^{-1}(u_m)\right)' \quad (10)$$

$$K = \Gamma\left(\frac{\nu}{2}\right)^{m-1} \Gamma\left(\frac{\nu+1}{2}\right)^2 \Gamma\left(\frac{\nu+m}{2}\right) \quad (11)$$

Unlike the Gaussian copula, the Student-t copula does incorporate tail dependence. However, the Student-t copula has the assumption of equal upper and lower tail dependence and does not incorporate any asymmetric effects. The expression for the upper and lower tail dependence is given as follows:

$$\lambda_{ij} = 2 * t_{\nu+1}(-\sqrt{\tfrac{(\nu+1)(1-\rho)}{1+\rho}}) \quad \text{for } \rho > -1 \qquad (12)$$

Here, $t_{\nu+1}$ is the standard univariate t-distribution with v+1 degrees of freedom and ρ is the correlation between two variables. It can be seen that when the degrees of freedom becomes large, the coefficient goes to zero. The tail dependence becomes zero and the Student's t copula simply becomes the Normal Copula.

Lets look at the scatter plot of the Student-t copula in figure 1. Here, the degrees of freedom is 4. The Student-t copula shows a symmetric dependence and also shows tail dependence, which can really be seen by the concentration in the tails of the scatter plot. This means that the Student-t could overestimate the VaR and resulting in less violations. In terms of economic implications, this could lead to a high capital requirement for banks and diminish investment opportunities

3.1.4 Clayton copula

The Clayton copula belongs to the Archimedean copulas. The Clayton copula distribution is given by:

$$C(u_1, \dots, u_m; \delta) = (u_1^{-\delta} + \dots + u_m^{-\delta} - m + 1)^{-\frac{1}{\delta}} \quad (13)$$

Where δ is the parameter to be estimated, with restriction $\delta > 0$. The copula density function is given by:

$$c(u_1, \dots, u_m; \delta) = \left(1 - m + \textstyle\sum_{i=1}^{m} u_i^{-\delta}\right)^{-m-(\frac{1}{\delta})} \prod_{j=1}^{m}\left(u_j^{-\delta-1}((j-1)\delta + 1)\right) \quad (14)$$

One of the main features is that is has zero upper tail dependence and a positive lower tail dependence. However, it isn't very flexible, because the Clayton copula assumes equi-dependence.

Before looking at the scatter plot of the Clayton copula, it is useful to see how the tail dependence of the Clayton copula depends on the parameter δ. See the formula below:

$$\lambda^l = 2^{-1/\delta} \quad \text{if } \delta > 0 \quad (15)$$

Where λ^l is the lower tail dependence. It can be seen that for larger values of δ, the lower tail dependence goes to 1 and that for low values of δ there is little dependence between the variables. Now let's look at the scatter plot in figure 1. Here, δ is set to 1.

We see that for a δ of 1 the lower tail dependence is 0.5 and gives a fair correlation between variables. The Clayton clearly shows lower tail dependence with the high concentration in the negative tail. The positive tail is left completely random. This means that the Clayton copula could lead to overestimation of the VaR forecasts or simply a bigger VaR than other dependence structures. Depending heavily on the parameter δ we expect the Clayton copula to have a larger or smaller VaR forecast.

3.1.5 Simple Factor Copula

The last model of the series of copulas is the factor copula. This model also has the main focus as this copula incorporates more desired features than the other copulas. The factor copula is based on latent variables, which are simulated and from which the desired copula is created. A simple factor structure is composed out of M+1 latent variables and can be defined as:

$$[x_1, \dots, x_m]' = X \sim F(x_1, \dots, x_m) = C(F_1(x_1), \dots, F_m(x_m)) \quad (16)$$

$$Y_i = Z + \varepsilon_i \,, \;\; i = 1,2, \dots, M \quad (17)$$

$$Z \sim F_z(\kappa), \;\; \varepsilon_i \sim iid\; F_\varepsilon(\zeta)\,, \;\; Z \coprod \varepsilon_i \;\; \forall\, i \quad (18)$$

$$[Y_1, \dots, Y_m]' = \boldsymbol{Y} \sim \boldsymbol{F_x} = \boldsymbol{C}(D_1(\theta), \dots, D_m(\theta); \theta) \quad (19)$$

Equation (16) is simply the observed variables with the copula structure as a reminder. θ is the copula parameter vector with $\theta = (\kappa, \zeta)$ and Y_1 is a latent variable with marginal distribution $D_1(\theta)$. The copula of latent variable Y, C(θ), is then used as the copula of the observable variables X. Note that the marginal distributions of the latent variable, Y, may be different from the observed variable X, so in general $D_i \neq F_i$. Do not confuse the different marginal distributions. It is important to remember that we use the structure for the vector Y *only* for the construction of its copula. The marginal distributions are not of any interest and may be discarded after the copula is constructed. We use the obtained copula from equation (15) to create a model for the copula of X (hence, the copula in (12). The marginal distributions F_i are to be defined and estimated in a later step. The latent variables Y_i are obtained by simulating Z and ε_i from their distribution. Z is a simulated variable that represents a characteristic that might be present in a certain financial time series (a.k.a. 'factor') and ε_i are the residuals. Both these variables are to be determined by the researcher and this immediately shows the flexibility of the factor copula. It forms according to the times series of your choice.

An important feature of this factor structure is that it is generally not known in closed form. There are some cases in which is does have a closed form, but there aren't. This poses difficulties for the estimation of the parameters. The copulas specified in the sections above can be estimated by maximum likelihood, because they have a copula density in closed form. The factor copula does not and hence, does not qualify for the traditional estimation method.

3.1.6 Multi-factor Copula Model

A further extension that is suggested by Oh and Patton (2012), is the "single factor, flexible weights" factor copula (a.k.a. multi-factor copula). It allows the weights on the factors to differ across variables. This structure can be defined as:

$$Y_i = \beta_i Z + \varepsilon_i\,,\ \ i = 1,2, \dots, M \quad (20)$$

$$Z \sim F_z(\kappa),\ \ \varepsilon_i \sim iid\ F_\varepsilon(\zeta)\,,\ \ Z \perp\!\!\!\perp \varepsilon_i\ \ \forall\, i \quad (21)$$

The rest of the structure/model is left unchanged. This model leads to one advantage and one major disadvantage. One great advantage is that is increases the flexibility to model heterogeneous pair of variables due to M-1 one extra parameters and that a given pair of variables may have a stronger or weaker dependence compared to a different set of variables. Hence, this copula is no longer equidependent. However, this also results in a more difficult estimation problem, which is a disadvantage as factor copulas carry a heavy computational burden already.

Figure 1 show the scatter plot of a bivariate simulation for the factor copula. The factor copula chosen is a multi-factor copula with common factor distribution skew-t and for the idiosyncratic factor the Student-t distribution. This is supported by the fact that Oh and Patton (2012) see the most potential in these distribution in relation to stocks. Here, the skewness parameter is set to -0.5, the degrees of freedom is set to 4 and the betas are set to 1. We see the skew t-t factor copula showing asymmetric tail dependence. The negative tail shows a higher concentration than the positive tail. Note that the positive tail isn't completely random, showing the asymmetric dependence. The difference between the Clayton copula and factor copula can be seen in the shape of the scatter plot. Where the Clayton shows a certain triangular shape distribution, the factor copula still shows a more elliptical/oval or even egg shape distribution. We can expect that with an overestimation of the $\beta's$ (which implicates a high correlation between the variables) that there will be an overestimation of the VaR forecasts and vice versa. The flexibility of the factor copula makes it harder to define what to expect.

3.1.7 Marginal Distributions

With the copulas specified, only the task of modeling the marginal distributions remains. For simplification purposes, one model will be considered. This way the paper doesn't lose its focus, which is on the copulas and not the on marginal distributions. This paper deals with the modeling of financial asset (stocks). Hence, an assumption is made that the time series follows a GARCH(1,1) process with normal errors. This is a reasonable assumption as it is done frequently in the literature and also in Oh and Patton (2012). It can be defined as follows:

$$X_{it} = \mu_{it} + \sigma_{it}\eta_{it} \quad t = 1, \dots, T \qquad (22)$$

$$\sigma_{it}^2 = \omega + \gamma\sigma_{i,t-1}^2 + \alpha\sigma_{i,t-1}^2\eta_{i,t-1}^2 \qquad (23)$$

$$\boldsymbol{\eta}_t = [\eta_{1t}, \dots, \eta_{mt}] \sim iid \quad F_\eta = C(\phi, \phi, \dots, \phi) \qquad (24)$$

Where ϕ is the standard normal distribution and C is the copula that we are trying to model with the copula models specified above. ω, γ and α are the parameters to be estimated. X_{it} is a vector with returns of M different assets, σ_{it}^2 is the variance of the asset i and η_{mt} Is the standard residual of asset m. This process is estimated by a maximum likelihood and the standard residuals will be used later in the estimation of the factor copulas. This GARCH(1,1) process is estimated in similar way to the copulas with the maximum likelihood.

3.2 Estimation Methods

In this section we elaborate on the estimation of the parameters of the different copula models. The estimation of the Normal, Student-t and the Clayton are fairly straightforward, but the factor copula introduces a new type of estimation for simulation based copulas (factor copulas in this case). First, this paper will start with the explanation of the estimation method of traditional copulas and continue with the explanation of the SMM-type estimation for factor copulas.

3.2.1 Copulas

The copulas are estimated via Maximum Likelihood by which the researcher must make an assumption on the underlying multivariate distribution. In this paper, the Normal, Student-t and Clayton distribution are chosen. The decomposition plays an important role in the estimation process, as it enables estimation in different stages. This estimation in different stages is also known as the inference functions for margins (IFM) method.

The estimation starts with looking the joint density that can be decomposed into the marginal distributions and a copula density, which is obtained when differentiating equation (3):

$$f(X_1, \dots, X_m; \boldsymbol{\alpha}, \boldsymbol{\theta}) = c(F_1(X_1; \alpha_1), \dots, F_m(X_m; \alpha_m); \boldsymbol{\theta}) \prod_{\boldsymbol{i=1}}^{\boldsymbol{m}} \boldsymbol{f_i}(X_{\boldsymbol{i}}; \alpha_{\boldsymbol{i}}) \qquad (25)$$

Where $f(X_1, \dots, X_m; \theta, \alpha)$ is the joint density function, $f(X_1; \alpha)$ is the marginal density function of variable X_1 and c is the copula density function. The parameters are the parameters of the copula and the marginals, with for example $\theta = [\Sigma, \nu]$ for the Student-t and $\alpha = [\omega, \gamma, \alpha]$ for the marginals. From this, we can obtain the likelihood function. We first take the log of the function, because this simplifies the equation when we construct the likelihood function. We obtain:

$$\ln L(\ \theta, \alpha\ ; X_1, \dots, X_m | I_t) =$$
$$\sum_{i=1}^{T} \left(\ln c(F_1(X_{1t_i}; \alpha_1 | I_t), \dots, F_m(X_{mt_i}; \alpha_m | I_t); \theta | I_t) + \sum_{i=1}^{m} \ln f_i(X_{it_i}; \alpha_i\ | I_t) \right) \qquad (26)$$

Where ln is the natural log and $L(\ \theta, \alpha\ ; X_1, \dots, X_m)$ is the likelihood function, T is the time and the information set I_t is $I_t = \{t_1, \dots, t_T\}$. Due to temporal dependence, the copula becomes conditional on the information set as this changes when the rolling window changes. The IFM enables us to split this into two separate estimations. We can now first estimate the parameters of the marginal densities individually, which translates into finding $\widehat{\alpha_i}$ that maximizes the likelihood function for the marginals:

$$\max_{\alpha_i} \sum_{i=1}^{T} \ln f_i(X_{it_i}; \alpha_i | I_t) \qquad (27)$$

Where $f_i(X_{it}; \alpha_i)$ is the marginal density function to be estimated. With standard normal residuals and time depending standard deviation, this can be defined as:

$$f_i(X_i; \alpha_i | I_t) = \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{\frac{-(x_i - \mu_i)^2}{2\sigma_t^2}}\ for\ i = 1, \dots, m \qquad (28)$$

After the estimation of the marginals, the copula parameters can be estimated. This starts by the usage of the standardized residual, where we use the estimated $\widehat{\sigma_t}$ and $\hat{\mu}$ to construct:

$$\eta_{it} = \frac{x_{it} - \widehat{\widehat{\mu_i}}}{\widehat{\sigma_{it}}} \qquad (29)$$

The standardized residuals are then transformed into uniformly distributed variables by inserting the residual in the univariate distribution of the marginals:

$$u_i = \Phi(\eta_{it}) \qquad (30)$$

Where $u_i \sim U(0,1)$ for i=1,…,m. The uniformly transformed variables are then inserted into the inverse univariate distribution of the desired copula. In the case of the Clayton copula this is not necessary. The uniformly transformed variable is directly inserted.

$$\Phi^{-1}(u_i) \ \ and \ \ T_v^{-1}(u_i) \qquad (31)$$

Note that for the Normal distribution it just gives back the residual. These are then inserted into the copulas. This results in:

$$C(u_1, \dots, u_m; \theta) = \mathbf{\Phi}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_m)) \ \text{ or } \ C_v(u_1, \dots, u_m; \theta) = \mathbf{T_v}(\mathrm{T_v}^{-1}(u_1), \dots, \mathrm{T_v}^{-1}(u_m)) \quad (32)$$

Next, the parameter θ can be estimated, using the log likelihood of the copula density function, which is given as follows:

$$\max_\theta \sum_{i=1}^{T} \ln c(u_{1t_i}, \dots, u_{mt_i}; \theta | I_t) \qquad (33)$$

Where the copula density function can be either one of the three copula density functions that have been specified (Normal, Student-t or Clayton). From this we procure the desired estimate $\hat{\theta}$, which differs across the different copulas.

3.2.2 Factor Copula

The factor copula introduces a completely new method of estimation for copulas that aren't defined in a closed-form likelihood. Oh and Patton (2012) base their estimation method on the Simulated Method of Moments (SMM). Even though it is strictly not SMM, they still refer to it as SMM, because the rank statistics that they use can be seen as "moments". They extract the parameters from the minimization of the difference in dependence measures between the residuals and the copula simulations. Hence, first the used dependence measures will be specified and then the estimation process will be explained.

Oh and Patton (2012) suggest using pair-wise rank correlation and quantile dependence, which they support with their preliminary studies. These measures are "pure" measures of dependence, which means that they are only affected by changes in the copula and not by the marginal distributions. The first measure of dependence is rank correlation.

A common know type of rank correlation is the Spearman's rank correlation. Using a monotonic function, the Spearman's rank correlation assesses how well the relationship between two variables can be described. For a sample of t, the t raw scores X_{1i} and X_{2i} are converted to ranked vectors x_{1i} and x_{2i} from which rho is computed ($Spearman's\ rho\ \varrho$):

$$\varrho = \frac{\sum_i (x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)}{\sqrt{\sum_i (x_{1i}-\bar{x}_1)^2 \sum_i (x_{2i}-\bar{x}_2)^2}} \qquad (34)$$

The second measure of dependence that Oh and Patton (2012) use is the quantile dependence. This measure of dependence measures the probability that an observation of one variable is in the q-tail of that variable, given that the same observation is made for another variable. It can be defined as:

$$\tau_q = \begin{cases} \frac{1}{q}\Pr[U_1 \le q, U_2 \le q] & q \in (0,0.5] \\ \frac{1}{1-q}\Pr[U_1 > q, U_2 > q] & q \in (.5,1) \end{cases} \qquad (35)$$

Where $U_i = D_i(Y_i) \sim Unif(0,1)$ are the transformed variables of the simulated Y_i. The probabilities can be obtained by simulating Y_i, summing the occurrence of the even happening and dividing by the total number of simulated variables. Oh and Patton (2012) show with these measures of dependence that different features from financial assets might be better captured by a flexible copula that incorporates asymmetric dependence. Note that the quantile dependence converges to the upper and lower tail dependence coefficients for q ->1 and q->0, respectively.

The actual estimation of the parameters is done by using these measures of dependence. They basically compare characteristics of the residuals to the characteristics of the simulated copula and try to get it as close as possible. The task at hand is to estimate a vector of px1 copula parameters, which is based on the standardized residuals η_{it} and the simulation from the copula.

Let $\widetilde{m_s(\theta)}$ be the vector with mx1 dependence measures, which is computed by simulating S times from the supposed joint distribution $F_y(\theta)$ (with copula C(θ))and let $\widehat{m_t}$ be the dependence measures that are computed by using the residuals η_{it}. The exact method for calculating the dependence measures in the following formula is defined in Appendix A. The estimation is defined as follows:

$$\widehat{\theta_{T,S}} = argmin_{\theta \in \Theta}\, Q_{T,S}(\theta) \qquad \text{T=1,...,t} \qquad (36)$$

$$Q_{T,S}(\theta) = g_{T,S}(\theta)' W_T g_{T,S}(\theta) \qquad (37)$$

$$g_{T,S}(\theta) = \widehat{m_t} - \widetilde{m_s(\theta)} \qquad (38)$$

Where S is the amount of simulations and W is some positive definite weight matrix, which may or may not depend on the data. In this paper it will be set to the identity matrix, which is also done in Oh and Patton (2012) on certain occasions, since the use of an efficient weight matrix produces similar results in Oh and Patton (2012). Oh and Patton (2012) also show, under regularity conditions, that this estimator is consistent and asymptotically normal, which means that as the sample size and the number of simulations increase the estimates become more reliable. They find it to be a reliable and sufficient method for estimating the parameters of a factor copula. However, this estimation method does come at a price, which will be discussed later on.

3.3 Inference for Copula Models

This section covers inference on the parameters of copula-based multivariate models. The main focus will be acquiring the variance of the estimated parameters. A short explanation will be provided of which methods are used in this paper. The procedure for the estimation of the variance of the parameters will differ across the estimation methods that were used for the copulas. In this paper this can be identified as the SMM-like estimation (factor copula) and the maximum likelihood (Normal, Student-t and Clayton copula).

3.3.1 Multi-stage maximum likelihood estimation

In this paper a multi-stage estimation is used. This is useful as the marginal distribution and copulas can be estimated separately. As for one-stage MLE, under regularity conditions, the MSMLE is asymptotically normal:

$$\sqrt{T}\left(\hat{\theta}_{T,MSML} - \theta^*\right) \xrightarrow{d} N(0, V^*_{MSML}) \;\; as \;\; T \to \infty \qquad (39)$$

Where T is the length of the time series, $\hat{\theta}_{T,MSML}$ is the vector of estimated parameters, θ^* is the true vector of parameters and V^*_{MSML} is the asymptotic variance matrix of the parameters. Acquiring V^*_{MSML} using a sandwhich approach can be quite tedious (Patton, 2012). However, an alternative, less tedious, method is to use a block bootstrap for inference. This can be done as follows:

1. Use a block bootstrap (stationary bootstrap of Palitis and Romano (1994) for example) to generate B bootstrap samples with length T. (e.g., B=1000)
2. Estimate the copula using the MSMLE for every bootstrap sample.
3. The result should be 1000 estimations for every parameter. For every parameter, take the $\alpha/2$ and 1- $\alpha/2$ of this distribution $\{\hat{\theta}_i\}_{i=1}^{B}$ to obtain the 1-α confidence interval of the parameters.

3.3.2 SMM

For the factor copula used in this paper a SMM-like estimation is used. Under regularity conditions, if S/T → ∞ $as\ T \to \infty$, the SMM estimator is consistent and asymptotically normal:

$$\sqrt{T}\left(\hat{\theta}_{T,S} - \theta^*\right) \xrightarrow{d} N(0, \Omega^*) \;\; as \;\; T, S \to \infty \qquad (40)$$

$$where\ \Omega^* = (G^{*\prime}W^*G^*)^{-1}G^{*\prime}W^*\Sigma^*W^*G^*(G^{*\prime}W^*G^*)^{-1}$$

Where Σ^* is the asymptotic variance, G^* is the derivative of $g^*(\theta)$ at θ^*, $g^*(\theta) = p - lim_{T,S\to\infty} g_{T,S}(\theta)$ and W^* is the weight matrix (the asterisk indicating the true values). These components require different estimation than standard applications (e.g. GMM application). A iid bootstrap can be used to create a consistent estimator for Σ^* and the standard numerical derivative of $g_{T,S}(\theta)$ at $\hat{\theta}_{T,S}$ can be used as a consistent estimator for G^*, conditional on the fact that the step size of the numerical derivative approaches zero slower than $\frac{1}{\min(\sqrt{S},\sqrt{T})}$. An elaborate explanation of this procedure can be found in the paper: "Simulated Method of Moments Estimation for Copula-Based Multivariate Models" (Oh and Patton, 2013).

3.4 Forecasting VaR

After all the parameters of the different copulas are estimated, forecasting the VaR is the only task that remains. The Value at Risk will be denoted as $VaR_{\alpha,t,h}$, where α is the chance of observing a return lower than the VaR, t is the time and h is the forecast horizon. The measure of this risk all depends on the econometric model used and the forecast method (direct or iterated).

Because copulas don't allow an analytic approach to the calculation of the h-day ahead value at risk, the iterating and direct approach need to make use of simulated data. The process for the direct forecast is explained in the following steps:

1. After the estimation of the parameters, draw a vector of standard uniformly distributed random variables, per financial asset, from the copulas with the estimated parameters. Matlab supplies an easy to use 'copularnd' function. The user simply needs to supply the parameters and the type of copula. For the Factor copula, see appendix B.
2. Use the inverse of the marginal distributions on the standard uniformly distributed variables to simulate the standardized returns. In this case the normal distribution.

$$e_{is} = \Phi^{-1}(u_{is}) \quad (41)$$

Where i is the financial asset and S is an arbitrary amount of simulations.

3. See formula 22. The returns can now be simulated by:

$$X_{i,t+h} = \mu_{it} + \sigma_{it+h}\eta_{it+h} \quad (42)$$

Where $\sigma_{i,t+h}$ can now by procured from the univariate GARCH process by forecasting one step ahead, μ_{it} from the univariate GARCH process and $\eta_{i,t+h}$ is simply the simulated standardized return in equation 41.

Once the returns are simulated, the portfolio returns are calculated according to formula 2. Next, the returns are sorted according to value and the α% absolute value is identified, which is the $VaR_{\alpha,t,h}$.

The process of the iterated approach is done in a similar manner. The main difference is that the process includes multiple step ahead forecasts. Also, for the iterated approach, the dependence structure and its parameters are held constant. Even though dependencies are known to change over time, this assumption is plausible for a short horizon. The structure above for the direct forecast can be modified slightly in step 3 to explain the process for the iterated approach:

3. See formula 22. The returns can now be simulated by:

$$X_{i,t+j*k} = \mu_{i,t+j*k} + \sigma_{i,t+j*k}\eta_{i,t+j*k} \;\; j = 1,..,\frac{h}{k} \quad (42)$$

Where k is the data frequency and j is the iteration (set j=1 at initiation). $\sigma_{i,t+j*k}$ can now be procured from the univariate GARCH process by forecasting one step ahead, $\mu_{i,t+j*k}$ from the univariate GARCH process and $\eta_{i,t+j*k}$ is simply the simulated standardized return in equation 41. Set j=j+1 and return to step 1 with the newly simulated time series in order to estimate new parameters. Continue this process until j=h/k, where h is the forecast horizon and k is the data frequency.

Finally, the value at risk is found in the method described above for the direct forecast.

3.4 Forecasting kES

The occurrence of the financial crisis reminded us that there is always risk. This called for the need of management and measure of systemic risk. Brownlees and Engle (2011) proposed a new measure of systemic risk: Marginal Expected Shortfall (MES). It is defined as the expected return on asset I when the market return is below a certain threshold.

$$MES_{it} = E_{t-1}[\, r_{it} \mid r_{mt} < C] \qquad (43)$$

Where r_{it} is the return on a financial asset i, r_{mt} is the market return, E_{t-1} is the expected value with information up to t-1 and C is an arbitrary small number. In this paper, the information up until t is used and by the use of simulation the MES is calculated for a horizon h.

Oh and Patton (2012) used this idea to come up with another different measure of systemic risk, the kES. This is the expected return of asset i when the return of k different assets fall below a certain threshold.

$$kES_{it} = E_{t-1}[\, r_{it} \mid \textstyle\sum_{j=1}^{N} \mathbf{1}\{r_{jt} < C\} > k] \qquad (44)$$

This is the measure that will be used. Brown and Engle (2011) found a simple way of ranking these estimates in order to compare models. Oh and Patton (2012) adjusted this method accordingly to the kES. This will be presented in the next section.

3.5 Evaluating VaR and kES forecasts

In this section the VaR and kES forecasts are evaluated, starting with the VaR forecasts. VaR forecasts are evaluated according to their number of violations. A violation occurs when the actual portfolio return is below the predicted VaR. For an accurate VaR, it would be expected that the VaR would violate α * F times. As VaR is always expressed in positive values, the amount of violations is defined as:

$$V_t = \sum_{t=1}^{T} \mathbf{1}\{r_{p,t,h} < -VaR_{\alpha,t,h}\} \quad (45)$$

Where V_t is the amount of violations and T is the amount of the out-sample forecasts. These violations are then backtested in order to check whether the forecasted VaR's correspond to the value of α. This is done according to an unconditional coverage test and this will test whether the amount of violations corresponds to its theoretical value α. This is done by using the likelihood ratio:

$$LR_{uc} = \frac{(1-\pi_{exp})^{n0}\pi_{exp}^{n1}}{(1-\pi_{obs})^{n0}\pi_{obs}^{n1}} \qquad -2\ln(LR_{uc}) \sim \mathrm{X}_{v=1}^2 \quad (46)$$

Where π_{exp} is the theoretical proportion of violations, π_{obs} the observed proportion of violations, n0 is the amount of times that the return does not exceed the VaR (non-violations) and n1 is the amount of violations. The test statistic, $-2\ln(LR_{uc})$, is asymptotically chi squared distributed with one degree of freedom. As this tests whether the threshold is violated, it doesn't look at the fact that a model could underestimate the VaR, which in turn would create no violations and might come out as the best model. However, this is of concern for banks. As the capital requirement for banks is directly related to the 1% VaR, this could mean a bank would have to hold too much capital and they would be missing out on a vast amount of investment opportunities.

The kES is evaluated in a similar manner as the simple Mean Squared Error evaluation:

$$MSE_i = \frac{1}{T}\sum_{t=1}^{T}(r_{it} - kES_{it})^2 \ 1_{\{\sum_{j=1}^{N}\mathbf{1}\{r_{jt}<C\}>k\}} \quad (47)$$

$$relMSE_i = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_{it}-kES_{it}}{kES_{it}}\right)^2 \ 1_{\{\sum_{j=1}^{N}\mathbf{1}\{r_{jt}<C\}>k\}} \quad (48)$$

Where MSE is the mean squared error, relMSE is the relative mean squared error and r_{it} is the actual return of the stock. A higher MSE compared to another model means that the model overestimated or underestimated the return even more than the other model. A higher RMSE (relMSE) combined with a lower MSE compared to another model means that even though the model might have had a better estimation, the model tends to low MSE's. This can be expected with models that tend to simulate low correlation between assets. In the next section there will first be a look at the data that was used to compare the models and risk measures.

3.6 Copula expectations

Before looking at the results and picking a winner amongst the competitors, a small simulation study is used to see whether the theory matches the results in terms of performance. Thus, first the expectations are expressed in this section and in the results these are compared to see if the theory can explain the performances of the copulas. In table 1 we can see the simulation results for the VaR of the different copulas. The 1-day and 10-day VaR for three assets are estimated by using the parameters given in table 1 to simulate a time series (N=5000) and then the VaR is calculated by taking the α% absolute value. The time series are generated by using the parameters to draw 5000 simulations of standard uniformly distributed variables from each of the copulas, which then are transformed into standard normally distributed variables such as in formula 41. Then, by combining the parameter estimates of the GARCH(1,1) process and the simulated standard normal variables, the simulated returns can be calculated by formula 43.

For the table on the left the VaR is estimated with linear correlation of 0.5 between the three variables and the other parameters are the same as in figure 1. On the right we observe the VaR forecasts for a different set of parameters and hence, different linear correlations. These estimations are the result of the first in-sample, which is defined in the next section. This can be useful as it illustrates the different behavior of the copulas and their influence on the VaR. For example, it illustrates how the equidependence of the Clayton copula results in a weak lower tail dependence and thus, might not show the characteristics one would expect such as in figure 1. The parameters used are: correlations for the Normal and Student t copula are around .26, .28 and .46, degrees of freedom is 4 and for the Clayton copula alpha is .28. For the factor copula: degrees of freedom (factor)=4.47, λ=-0.12, degrees of freedom(epsilon)=2.96, Beta1=1.88, Beta2=0.85 and Beta3=0.91.

[Insert table 1 here]

Looking at table 1, for the VaR forecasts with a linear correlation of 0.5 we see that the Clayton does indeed produce the largest VaR forecasts for most of the different forecast horizons and VaR parameters. The most interesting thing to see is that the factor copula produces second or third largest 1%VaR, but produces lowest 5% VaR. This is interesting because this might be an indication of the flexible nature of the factor copula. Large negative returns are more correlated than smaller negative returns, which might also be the case with other copulas, but this effect might be stronger with the factor copula. Now if we look at table 2, the VaR of the copulas gives us a completely different image. The parameters for this table are shifted in a different direction in order to show the potential shifts in VaR forecasts between models. We see that the Clayton copula is thrown of its throne and the Student-t copula clearly leads in largest VaR forecasts. This is where the one-parameter-limitation of the Clayton copula shows. It has trouble modeling for higher dimensions than 2 and produces quite low dependence. It now produces VaR forecasts around the Normal Copula or lower. This means, while we expect a low amount of violations for the Clayton with a large parameter, we now expect a lot more. The other remarkable thing is the factor copula, that has quite low 1% VaR and 5% estimates. However, due to the variability and computational difficulty of this copula, this can change a lot. Hence, it could lead to a large or a low amount of violations. The least number of violations is expected for the Student-t copula as its estimates are quite high, resulting in less violations. To see whether this is a good result depends on whether the observed amount of violations matches the expected amount of violations. Financial institutions might have to hold too much capital as a result of a low amount of violations, which might mean losses of potential investments.

To give an idea what kind of quantiles the C values in the kES mean, table 2 shows the quantile values for different values of sigma.

[Insert table 2 here]

It can be seen that the quantiles depend very much on the volatility of an asset. With a volatility of 1% the values C=-2% and C=-3% indicate large negative return, while higher volatilities indicate this in a less severe manner. It can be said that the extremity of the quantile depends heavily on the volatility of the asset and that the values of C should be chosen accordingly.

4 Data

In the empirical analysis, data from Yahoo Finance is used. Yahoo Finance has an enormous database for the most common stocks and provides an easy access to historical prices. Plus, Yahoo gives the option for downloading stock prices daily, weekly and monthly making it easier for the researcher to create the desired dataset. It does limit itself in sense that data can only be downloaded one at a time. Yahoo doesn’t provide the option of downloading historical prices for a certain amount of tickers. Yahoo Finance can be found via yahoo.com. In this paper, 10 stocks from the s&p 100 were chosen in order to perform the empirical analysis. Oh and Patton (2012) used all 100 stocks of the s&p 100 for their analysis, but unfortunately this was impossible in this paper due to computational limitations. However, to have similar data, stocks from the s&p 100 were chosen.

In appendix C the chosen stocks are listed. The spot prices and weekly spot prices cover the sample period 30-5-1986 till 13-5-2013, containing 5796 returns. In the case of weekly returns this results into 1159 weekly returns. Furthermore, the rolling window in-sample for the daily and weekly returns that was used was 3000 and 600, respectively, which corresponds to around 12 years in-sample. The out-sample is roughly eleven years. Also, in this paper a sub period is examined, which covers the period between 2-1-2003 and 2-1-2007. This period is chosen due to the fact that it is known as a period without turmoil. The behavior/performance of the copulas might differ across different states of the economy, which tests the flexibility of a copula. Other sub periods show similarities to the complete out of sample period with crises and such.

The large in-sample is chosen so that the weekly data contains enough data for the factor copula. Oh and Patton (2012) stipulate that for the estimation of a factor copula a large data set is essential for the accuracy of the parameter estimation. They set their dataset at N=1000, whereas the weekly data in this paper has 600 observations, which should be sufficient. However, this is a moment to realize that even though the factor copula shows a great deal of flexibility, it also shows a great deal of limitations.

To get a better understanding of what type of data we’re dealing with, table 3 shows a summary of some data statistics for the 10 stocks from the s&p 100.

[Insert Table 3 here]

It can be said that for this period average return was close to zero. Certain stocks were certainly also a lot more volatile than others as the standard deviation varies from 1.71% to 3.36% approximately. This high volatility can be explained by crashes such as the dotcom crash in the early 2000’s or extremely bad quarterly results like with Apple in 2001. These events can explain large crashes that are present in the data. Note that the returns are calculated by taking the log difference and that this approximation for short term returns does not hold for large returns. Thus, for some large returns this approximation does not hold, but only happens on rare occasion. This must prove to be a difficult task for the copulas as it would be for any model to model large crashes with a large in-sample (even for a small in-sample). However, this insample is needed for the factor copula. Furthermore, we see that stocks either have a positive or a negative skewness, which concurs with the literature. We also see high kurtosis, which is to be expected as returns usually show fat tails, meaning that extreme returns are to be expected more often than under normal assumption. The financial returns stylized facts (skewness and large kurtosis) do appear to be present in the weekly returns, even though it might be somewhat weak.

The Jarque-Bera values are fairly large, which is a test-statistic for the assumption of normality. The Jarque-Bera probability (right below the Jarque-Bera values) rejects the assumption of normality for all stocks, meaning that that it might be useful , for further research, to investigate optimal marginal distributions.

Table 4 shows the correlation between the stocks chosen from the s&p 100.

[Insert Table 4 here]

The correlations for the stocks vary from around 0.16 up to around 0.61, which might prove to be a fair diversification as most stocks are not so much correlated with each other, but it doesn't single out the possibility of a major crash of such a portfolio. However, the aim of this paper doesn't lie in optimal portfolio choice.

As this paper also deals with non-linear forms of dependence, table 5 shows the quantile dependences for the entire sample period for values of q=0.05, 0.1, 0.9 and 0.95. From the tables can be deduced that the lower quantile dependences are stronger than the upper quantile dependences. This would suggest that the incorporation of asymmetric tail dependence would be justified in the factor copula and the Clayton copula. However, the table also shows that the upper quantile dependence is weaker, but positive, which might indicate that upper tail dependence is non-zero. This is not captured by the Clayton copula and therefore weakens the arguments for the use of a Clayton copula.

[Insert table 5 here]

Figures 2 and 3 show the correlations and the quantile dependences over sub periods with windows of 200 observations for different stocks for the entire sample period. These graphs show that these dependencies can be quite volatile and especially in times of crises. Correlations vary from around 0 to around 0.7 and quantile dependence from around 0.1 to 0.8. These volatile correlations are observed throughout all stocks and therefore further graphic display is unnecessary. The volatile correlations might indicate that the large in-sample requirement of the factor copula poses another restriction, as the data might suggest that dependences could vary greatly within sub periods.

[Insert figure 2&3 here]

5 Results

The data described in the previous section is used to estimate the risk forecasts. These are then evaluated for daily and weekly data, for all forecasts horizons (1, 5 and 10 days) and for all copulas of course. However, before we compare and analyze the end results of the different types of copulas, this paper first takes a look at the estimations of the copulas.

5.1 Copula Estimations

Looking at the copula estimations in more depth can be quite useful, as this can display the reliability and the goodness of fit in more detail. A reliable and well fitted copula would then indicate better predictive power.

5.1.1 Tail dependence

In this section the scatter plots of a bivariate simulation of the copulas are created once again, but now with the in-sample estimates of two stocks (American Express Inc. and Baxter International inc., which are good representatives of characteristics between the 10 stocks). The parameter estimates can be seen in table 7, which will also be discussed later on. This scatter plot will show whether the copula still show their characteristic features as they did in the methodology section. In figure 4 the scatter plots are shown of the bivariate simulations for the four copulas. This time the parameter estimates of the in-sample are used and should result in different modeled dependencies. It can be seen that the Normal copula behaves fairly the same. However, it seems that the strong characteristics of the other copulas seem to have become weaker. The lower tail dependence of the Clayton copula has become weaker, which is also the case for the symmetric dependence of the Student-t copula. The factor copula is quite vague; not showing asymmetric dependence. Figure 5 shows the same scatter plot for the standardized returns of the actual data. A quick glance might make one think to believe it is similar to the normal copula or the factor copula. However, the visualization is not clear enough to draw any conclusions. In the next section, a goodness of fit test is done, which will clarify which is the best fit for the data.

[Insert figure 4 & 5 here]

5.1.2 Goodness of Fit

In table 6 the RSME for the goodness of fit test of the copulas are given. A goodness of fit of a statistical model measures how well the model fits the data. The goodness of fit test was done by constructing the empirical copula and the fitted copulas for 2 variables. The empirical copula is based on sample order statistics, which involves ordering the observations of two variables in increasing order of magnitude. Then, the empirical copula is defined as follows:

$$C\left(\frac{i}{T},\frac{j}{T}\right) = \frac{Number\ of\ pairs\ (x,y) such\ that\ x \leq x^i\ and\ y \leq y^j}{T} \quad (49)$$

where x^i and y^j are the ordered variables. For the estimated copulas the same can be done. A time series of the same length can be created by the data generating process of section 3.6. Then, in a similar matter a fitted copula can be computed, after which the RMSE can be calculated.

It can be quickly seen that the factor copula provides the best fit. One might conclude that the conclusion is evident; the factor copula is expected to provide the best forecasting power. However, this conclusion might be premature. It all depends on the limitations in terms of time, knowledge or facilities of a researcher, portfolio manager, student, etc. One might be interested in the estimation for a large set of variables or demand that the estimation takes less than 2 minutes. The question remains how applicable the copulas are. This will be elaborated in the following section.

[Insert table 6 here]

5.1.3 Reliability of Copulas

In table 7 the parameter estimates for the in-sample are given with the matching variance, standard deviation and confidence interval with alpha=0.05. This is done for the stocks of American Express Inc. and Baxter International inc. The estimates of these two variables resemble the problems that arise from the estimation of 10 variables and hence, provide an adequate view on the accuracy of the parameter estimates. A quick glance can tell that the factor copula estimates seem to have higher variances and thus, a larger confidence interval. This seems to make the factor copula a less reliable candidate as its parameters vary quite a bit around the true parameter. This might indicate that the amount of simulations wasn't high enough as the accuracy of the parameters rely on a high amount of simulations. However, as will be explained later on, this comes at a price. It isn't clear however what impact the high variance has on the copula fit.

[Insert table 7 here]

In table 6 the goodness of fit is tested once more. This time the lower and upper bound of the confidence interval of the factor copula parameters are compared to see whether the uncertainty impacts the goodness of fit severely. Thus, we investigate how sensitive the factor copula fit is compared to its parameters. The RMSE shoots up for the lower and upper bound of the parameters. Clearly, the goodness of fit is quite sensitive to variation in parameters. An increase in certainty can be accomplished by a higher amount of simulations. However, as shown below, this comes at a cost.

In table 8 the estimation time of this factor copula is shown for different sizes of a portfolio and different size of simulation. Also, as a comparison, the estimation times for a Student-t copula are presented. Results show that large simulations for large portfolios become a time consuming operation. This makes the factor copula less applicable for certain causes. A portfolio manager interested in the VaR of a large portfolio of 50 variables might have to wait a week before the estimation would be done. However, the estimation time for the factor copula does increase at a slower pace than the Student-t copula (which seems to increase exponentially or in other words, really fast), which might mean that the factor copula is a better prospect for future use. Keep in mind, though, that external factors might influence these results such as the standard estimation methods in Matlab, which might not be the most efficient way to estimate the copulas.

[Insert table 8]

5.2 VaR & kES Evaluation

In this section the evaluations of the different horizon VaR forecasts are presented for the different copulas. From the simulation study a few expectations were derived. The Student-t would show the least amount of violations, followed by either the Normal or Clayton. The factor copula turned out to be weak on the 5% VaR, supposedly performs poorly under low frequency data and is supposed to perform better for higher dimensions. The results for a portfolio of 3 assets are presented in table 9 below. It shows the results for weekly and daily data, for the four copulas, the 1, 5 and 10-day VaR at significance levels of 1% and 5%:

[Insert table 9 here]

With a quick glance, it can be seen that the Student-t mostly dominates on terms of violations as expected. The Clayton copula underestimates the VaR as expected due to the low parameter estimation and loses to the other copulas in terms of VaR violations. The Normal Copula comes in second or third most of the times. The factor copula, being the most interesting copula, actually does not underestimate the VaR too much in all cases, which might have been expected after the pre-study. The factor copula does perform weakly for the 5% VaR and weekly data, which was to be expected. The surprising aspect of the factor copula is that is does outperform the other copulas for the 10-day VaR at 1%, which might indicate a stronger estimation for longer forecast horizons. At a 5% level, the factor copula comes in second, which isn't a bad result and again might confirm the performance strength for longer forecast horizons. Apart from the copulas, it can also be noted that the direct forecasts of the 5-day VaR slightly outperform the iterated forecasts, which is not a huge surprise as iterated forecasts are more prone to biase due to possible misspecification.

Apart from the VaR forecasts, the kES was also computed for one variable at k=1, C=-2% and C=-3% for 3 assets and at k=3, C=-2% and C=-3% for 10 assets. First, the results for 3 assets are presented.

[Insert table 10 here]

The MSE is dominated by the Clayton copula as it has the lowest MSE for every type of output. However, it performs poorly in terms of relMSE. This could mean that the Clayton copula estimation results in a small parameter, meaning a low correlation between assets. This is confirmed throughout this paper and is thus a reasonable explanation for its results, since a low correlation could mean conservative predictions and on average result in low prediction errors. For the relMSE, the Student-t copula performs really well. However, performs poorly in terms of MSE. This is probably explained by the overestimation of the kES due to the strong negative tail dependence.

When we look at the results for the estimation with 10 assets, we see some changes. In table 11 the results are presented:

[Insert table 11]

We can see that the Student-t remains the victor and shows the least amount of violations and closest to the expected amount of violations. The Clayton copula performs badly as the higher dimension is too much to handle for the Clayton copula due to the equi-dependence across all assets. It can be seen that for higher frequency data and higher dimensions, the factor copula performs badly compared to the other copulas. It underestimates the VaR by far. This might due to the estimation inaccuracy described in the section above. For the weekly data, the factor copula performs mediocre, which is somewhat of a surprise as the expectation was that weekly data estimation could result in less accuracy. However, this also means that when there is a sub-optimal solution and there are less iterations for the forecast, the weekly data might end up with less violations because the iteration will add up to less error than with daily data.

On further note, in table 12 it can be seen that the copulas perform similarly in a 10 asset scenario compared to 3 assets for evaluations of the kES.

[Insert table 12 here]

Just like the kES in combination with 3 assets, the MSE is dominated by the Clayton copula as it has the lowest MSE for every type of output. However, it performs poorly in terms of relMSE. For the relMSE, the Student-t copula performs really well again. However, performs poorly in terms of MSE. These results of the kES, with two different sets of variables, might suggest that it´s a relatively new method to measure risk and a definite way to implement the kES for out-of-sample results is yet to be found.

After evaluating the out of sample, another out of sample was introduced in the data section; a sub period. This period resembles a period without turmoil (e.g. the credit crisis) and it would be interesting to see how the copulas cope with periods of tranquility. It would be interesting as, for example, the Student-t might overestimate the VaR by a long shot and have an undesired low amount of violations. As the focus in this paper is mainly the VaR, the kES is not incorporated for the sub period. The results for the VaR estimations with 3 assets are presented below:

[Insert table 13 here]

It is important to see that the Student-t still dominates for the 1% VaR on all forecast horizons, which is the most important VaR as it is economically applicable. We do see however that other copulas dominate the 5% VaR on both daily and weekly data. The Student-t overestimates and gets a low amount of violations. In general, no preference can be seen between low frequency data or higher frequency data.

Table 14 present the results for 10 assets:

[Insert table 14 here]

The results for the 10 assets are actually comparable to the estimation with 3 assets. The Student-t mostly dominates the 1% VaR and on the 5% VaR overestimates. This leaves room for the other models to get closer to the expected 5% violations. Again, low frequency or high frequency data doesn't come out as a convincing victor.

5.3 Factor Copula: Computational Limitations

The factor copula shows a lot of potential and leaves room for even more flexibility in the dependence structure (see Oh and Patton (2012)). However, the factor copula does come with its limitations and cannot be ignored in this paper. The limitations come down to a huge computational disadvantage compared to the other copulas, which can be divided in the following sub-problems:

- The factor copula estimation process relies solely on simulation. Of course, a large enough simulation sample would potentially mitigate this fundamental problem. However, the advantage of less parameters (and computation time) disappears as the estimation time increases quite rapidly when the size of the simulation sample is increased. This increase in estimation time can be the difference between hours and days. Especially a study exploring forecasts in combination with a large out of sample and long forecast horizons, which is the case in this paper.
- The objective function of the factor copula tends to have a lot of local minima. Hence, there are occasions that a suboptimal solution is reached. This can be reduced by reducing the function tolerance (increasing accuracy), but this also is at the cost of computation time and introduces the aspect of human judgement in what is a just function tolerance.
- A higher dimension comes at a cost, which again is the computation time. This makes sense as it happens with any model. However, the benefit of the parameter reduction does not show in terms of computation time even though Oh and Patton(2012) praise the model for its use with high dimensions. Hence, the dimension remains a burden.

Due to these problems it is advised to set out a large grid of data points and try to approach the global minimum. This way, the first estimation doesn't get stuck in a local minimum and the next estimation will go a lot smoother.

The model leaves room for two possible solutions. The first is that Oh and Patton (2012) put a restriction on the common factor, which causes a reduction in the number of parameters to be estimated (see Oh and Patton (2012) for "block equidependence"). The second option would be increasing computational power. This could be in terms of using a high-end computer, but this could also mean the use of a supercomputer. An increasing amount of universities house a supercomputer for researchers to use, so this could be a more realistic solution for students and professors in the future. An out-of-the box solution would be to assemble the parts in order to build a home supercomputer, which is becoming increasingly more viable possibility. See Durham and Geweke (2012) for more information about the possibilities of enhanced programming. All in all, the vivid development of technology could mean a great lot to the field of research, which could mean the rise of the factor copula.

6 Conclusion

In this paper a newly proposed copula, the factor copula (introduced by Oh and Patton(2012), was compared to other well known copula in terms of out of sample risk forecasts. The factor copula also required a new method of estimation, which is comparable to the Simulated Method of Moments. Oh and Patton (2012) proved that this method can be a reliable method estimation. However, the requirement of immense large simulations and large datasets slows the factor copula estimation down and makes an out of sample study a challenging task. Hence, a researcher might choose to use some restrictions, which might lead to sub optimal estimations.

The copulas were compared in terms of VaR and kES. The Student-t copula proved to be superior in most cases for the 1% VaR and for the 5% VaR it differed quite a lot. In terms of high frequency versus low frequency, the low frequency data was mostly preferred. The factor copula did show a great deal of flexibility, but didn't manage to keep up accuracy. However, this might be greatly influenced by the problems discussed above.

This paper leaves a lot to discuss and suggest. To begin with, the factor copula leaves a lot to investigate. As the estimation time of the factor copula was quite a burden in this paper, it would be interesting to investigate the gain in accuracy when the dimension increases, when size of simulations increase and when the function tolerance[1] is changed. This would give more of an idea of how applicable this copula is. As far as this paper goes, estimation for more than 50 variables would be very time consuming in a large out of sample study. These suggestions above could be possible easier realized in the nearby future. Innovation keeps ascending technology to the next level, which means the curse of dimensionality could become less of a problem and more a myth from the past.

In this paper the Normal distribution was chosen for the marginals. This, which could be widely discussed, might not be the most suitable distribution for the marginals. It would be intriguing to see how different types of marginals compare across these different copulas. A usual suspect would be the Student-t distribution, which might be a better suit for financial assets.

And last, but not least, Oh and Patton (2012) also propose another extension on top of the Skew t-t factor copula used in this paper. This would leave room for even more flexibility. However, this does mean extra parameters and estimation time. Oh and Patton (2012) also set restriction on the Beta parameters in their factor copula in order to reduce this problem and make it a more applicable copula. However, technology at the moment makes it inaccessible for the vast majority of people that share these interests. Perhaps the fast and vivid changing world of technology might soon make such researches more accessible.

[1] The function tolerance determines when the change in the objective function is small enough to conclude that you have found an optimal solution.

Bibliography

[1] Oh and Patton, 2012, Modelling Dependence in High Dimensions with Factor Copulas, Duke University.

[2] Patton, 2012, Copula Methods in Forecasting Multivariate Time Series, *Handbook of Economic Forecasting*, *Volume 2.* (G. Elliott and A. Timmermann, Eds.). Springer Verlag, forthcoming.

[3] Li, 2000, On Default correlation:" A copula function approach", *Journal of Fixed Income*, 9, 43-54.

[4] Genest and Rivest, 1989, A characterization of Gumbel's family of extreme value distributions, *Statistics & Probability Letters*, 8, 207-211.

[5] Demarta and McNeil, 2005, The t copula and related copulas, *International Statistical Review*, 73, 111-129.

[6] Daul, De Giorgi, Lindskog and McNeil, 2003, The grouped t-copula with an application to credit risk, *RISK*, 16, 73-76.

[7] Aas, Czado, Frigessi and Bakken, 2009, Pair-copula constructions of multiple dependence, *Insurance: Mathematics and Economics*, 44, 182-198.

[8] Acar, Genest and Neslehova, 2012, Beyond simplified pair-copula constructions, *Journal of Multivariate Analysis*, 110, 74-90.

[9] Markwat, Kole and Van Dijk, 2010, Forecasting value-at-risk under temporal and portfolio aggregation, Erasmus University of Rotterdam.

[10] Gysels, Rubia and Valkanov, 2009, Multi-period forecasts of volatility: direct, iterated and mixed-data approaches. EFA 2009 Bergen Meetings Paper.

[11] Diebold, Hickman, Inoue and Schuermann, 1997, Converting 1-day volatility to h-day volatility: Scaling by $\sqrt{n}$ is worse than you think, University of Pennsylvania.

[12] Santos, Nogales and Ruiz, 2013, Comparing Univariate and Multivariate Models to Forecast Portfolio Value-at-Risk, *Journal of Financial Econometrics*, Volume 11, No. 2, 400-441.

[13] Sklar, 1959, Functions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229-231.

[14] Van Aalten, 2011, Multi-period value-at-risk forecasting using time-varying covariance models, Erasmus University of Rotterdam.

[15] Brownlees and Engle, 2012, Volatility, Correlation and Tails for Systemic Risk Measurement.

[16] Tham, 2011, Lecture 4 – Introduction to Quantitative Risk Management, 56-57, Erasmus University Rotterdam.

Tables and Graphs

Table 1 On the left VaR forecasts for the copulas with linear correlation of 0.5 for a set of 3 variables. Other parameters are the same as in figure 1 discussed previously. The 1-day and the 10-day VaR are estimated with daily data and thus, the 10-day VaR is estimated with the iterative approach. On the right we observe the VaR forecasts for a different set of parameters. The correlations for the Normal and Student t copula are around .26, .28 and .46, degrees of freedom is 4 and for the Clayton copula alpha is .28. For the factor copula: degrees of freedom (factor)=4.47, λ=-0.12, degrees of freedom(epsilon)=2.96, Beta1=1.88, Beta2=0.85 and Beta3=0.91. This turns out to be the estimation result of the first in-sample, which is defined in the next section.

	correlation=0.5			
	Normal	Student-t	Clayton	Factor
	1-day			
1%	0.0335	0.0341	0.0386	0.0342
5%	0.0234	0.0232	0.0252	0.0220
	10-day			
1%	0.0986	0.0943	0.0999	0.0952
5%	0.0660	0.0638	0.0642	0.0615

	Various estimation			
	Normal	Student-t	Clayton	Factor
	1-day			
1%	0.0296	0.0320	0.0302	0.0256
5%	0.0206	0.0209	0.0205	0.0175
	10-day			
1%	0.0915	0.0962	0.0811	0.0568
5%	0.0641	0.0649	0.0554	0.0386

Table 2 This table shows the kES quantiles for different values of sigma and C in equation . The quantiles are based on a large simulation and N(0,1) marginals, from which returns are simulated for a given sigma.

kES quantiles for different values of sigma				
	σ=			
Threshold	1.00%	1.50%	2.00%	2.50%
C=				
-2%	0.0225	0.0919	0.1585	0.208
-3%	0.0014	0.0225	0.0684	0.1154

Table 3 Summary statistics for the 10 s&p 100 stocks

	AXP	BA	BAC	Abbott	AIG	AME	AMGN	APACHE	Apple	BAX
				Daily returns since May 2008						
Mean	0.04%	0.03%	0.02%	0.04%	-0.01%	0.03%	0.07%	0.04%	0.06%	0.04%
Median	0	0	0	0	0	0	0	0	0	0
Maximum	18.78%	14.37%	30.20%	11.73%	50.68%	18.04%	14.06%	19.32%	28.77%	10.38%
Minimum	-31.20%	-31.20%	-34.22%	-31.20%	-60.40%	-31.20%	-31.20%	-31.20%	-73.17%	-31.20%
Std. Dev.	2.38%	2.01%	2.86%	1.71%	3.36%	1.52%	2.37%	2.43%	3.13%	1.85%
Skewness	-0.35	-0.95	-0.44	-1.27	-0.53	-1.92	-0.33	-0.33	-2.28	-2.31
Kurtosis	14.65	18.89	29.11	27.49	63.07	57.91	12.79	11.70	60.95	37.66
Jarque-Bera	32875.51	61808.32	164736.9	146420.3	871400.9	731507.5	23237.43	18388.4	815856.6	295282.6
Probability	0	0	0	0	0	0	0	0	0	0
Observations	5795	5795	5795	5795	5795	5795	5795	5795	5795	5795
				Weekly returns since May 2008						
Mean	0.20%	0.11%	0.11%	0.18%	0.05%	0.11%	0.33%	0.18%	0.30%	0.16%
Median	0.24%	0.23%	0.23%	0.23%	0.08%	0.28%	0.33%	0.13%	0.49%	0.32%
Maximum	27.51%	18.57%	55.75%	26.43%	121.40%	35.00%	24.52%	18.74%	41.56%	18.19%
Minimum	-60.52%	-60.52%	-60.52%	-60.52%	-60.52%	-60.52%	-60.52%	-60.52%	-70.69%	-60.52%
Std. Dev.	5.08%	4.77%	6.43%	3.86%	7.88%	4.09%	5.16%	5.41%	6.96%	4.20%
Skewness	-1.76	-2.59	-0.38	-3.24	3.33	-4.62	-1.39	-1.32	-1.39	-3.20
Kurtosis	24.73	31.54	25.30	57.71	67.69	74.94	21.08	16.87	19.50	44.05
Jarque-Bera	23403.21	40645.79	24044	146555.5	204233.3	254027.1	16165.5	9627.435	13524.68	83377.05
Probability	0	0	0	0	0	0	0	0	0	0
Observations	1159	1159	1159	1159	1159	1159	1159	1159	1159	1159

Table 4 Correlation values for the selected stocks for the entire sample period.

	AXP	BA	BAC	Abbott	AIG	AME	AMGN	APACHE	Apple	BAX
AXP	1.00	0.43	0.61	0.32	0.45	0.35	0.31	0.29	0.28	0.31
BA	0.43	1.00	0.36	0.30	0.31	0.33	0.27	0.30	0.23	0.31
BAC	0.61	0.36	1.00	0.27	0.50	0.30	0.24	0.28	0.25	0.27
Abbott	0.32	0.30	0.27	1.00	0.25	0.32	0.33	0.18	0.16	0.44
AIG	0.45	0.31	0.50	0.25	1.00	0.29	0.18	0.23	0.18	0.24
AME	0.35	0.33	0.30	0.32	0.29	1.00	0.23	0.28	0.17	0.31
AMGN	0.31	0.27	0.24	0.33	0.18	0.23	1.00	0.18	0.24	0.31
APACHE	0.29	0.30	0.28	0.18	0.23	0.28	0.18	1.00	0.17	0.20
Apple	0.28	0.23	0.25	0.16	0.18	0.17	0.24	0.17	1.00	0.16
BAX	0.31	0.31	0.27	0.44	0.24	0.31	0.31	0.20	0.16	1.00

Table 5 This table extends over the next two pages. From top to bottom, relatively, the quantile dependence of the sample period at q°0.05, 0.1, 0.90 and 0.95.

	AXP	BA	BAC	Abbott	AIG	AME	AMGN	APACHE	Apple	BAX
AXP	1	0.362	0.425	0.214	0.404	0.304	0.228	0.214	0.193	0.217
BA		1	0.297	0.238	0.311	0.262	0.197	0.238	0.173	0.235
BAC			1	0.169	0.473	0.255	0.190	0.224	0.166	0.197
Abbott				1	0.197	0.214	0.252	0.138	0.138	0.300
AIG					1	0.283	0.169	0.211	0.142	0.193
AME						1	0.183	0.224	0.138	0.214
AMGN							1	0.162	0.186	0.224
APACHE								1	0.121	0.190
Apple									1	0.128
BAX										1

	AXP	BA	BAC	Abbott	AIG	AME	AMGN	APACHE	Apple	BAX
AXP	1	0.378	0.471	0.283	0.435	0.305	0.302	0.252	0.257	0.293
BA		1	0.335	0.283	0.342	0.290	0.255	0.280	0.224	0.286
BAC			1	0.257	0.481	0.290	0.224	0.255	0.235	0.261
Abbott				1	0.266	0.262	0.309	0.188	0.198	0.364
AIG					1	0.318	0.248	0.266	0.221	0.280
AME						1	0.235	0.288	0.202	0.254
AMGN							1	0.224	0.248	0.286
APACHE								1	0.198	0.226
Apple									1	0.193
BAX										1

	AXP	BA	BAC	Abbott	AIG	AME	AMGN	APACHE	Apple	BAX
AXP	1	0.305	0.449	0.267	0.419	0.276	0.247	0.240	0.247	0.245
BA		1	0.254	0.216	0.290	0.242	0.211	0.254	0.192	0.235
BAC			1	0.233	0.426	0.243	0.202	0.250	0.212	0.207
Abbott				1	0.255	0.242	0.254	0.173	0.167	0.305
AIG					1	0.252	0.195	0.233	0.192	0.231
AME						1	0.178	0.200	0.185	0.204
AMGN							1	0.164	0.228	0.221
APACHE								1	0.200	0.160
Apple									1	0.190
BAX										1

	AXP	BA	BAC	Abbott	AIG	AME	AMGN	APACHE	Apple	BAX
AXP	1	0.240	0.409	0.209	0.368	0.233	0.195	0.202	0.178	0.198
BA		1	0.209	0.171	0.226	0.192	0.160	0.167	0.143	0.171
BAC			1	0.164	0.392	0.202	0.150	0.205	0.136	0.157
Abbott				1	0.157	0.160	0.216	0.129	0.091	0.254
AIG					1	0.202	0.102	0.174	0.123	0.164
AME						1	0.154	0.150	0.119	0.174
AMGN							1	0.140	0.164	0.174
APACHE								1	0.143	0.116
Apple									1	0.133
BAX										1

Table 6 RMSE's between empiral copula and copula fit. Also for the factor copula, RSME for the estimated parameters, for the upper bound of the parameters and the lower bound of the parameters.

	RSME		
	Estimation	Upper	Lower
Normal	0.19	-	-
Student-t	0.1863	-	-
Clayton	0.7028	-	-
Factor (Skewt-t)	0.1804	0.9299	1.3438

Table 7 Parameter estimates for the in-sample of two stocks (American Express Inc. and Baxter International inc.).

	Normal	Student-t		Clayton	Factor				
	P	P	v	α	v1	λ	v2	β1	β2
Estimation	0.4966	0.5106	13.1748	0.5749	4.0461	-0.0326	2.3775	1.6312	1.5359
Variance	0.0003	0.0003	8.1351	0.0023	0.0318	0.0746	0.2126	0.1408	0.0684
Std.	0.0185	0.0168	2.8522	0.0482	0.1783	0.2731	0.4611	0.3752	0.2615
Conf. int.	[0.4586; 0.5308]	[0.4779; 0.5422]	[9.4670; 20.3602]	[0.4764; 0.6711]	[4.4027; 3.6894]	[0.5136; -0.5788	[3.2996; 1.4552]	[2.3816; 0.8807]	[2.0589; 1.0128]

Table 8 Estimation times for the in-sample with the Student-t copula and Factor copula for different portfolio sizes and different simulation sizes.

	Factor			Student-t
	S=15000	S=45000	S=75000	MLE
N=2	63.734	158.019	117.825	0.963
N=3	82.035	393.275	398.833	2.203
N=10	550.644	2483.582	3154.218	1859.557

Table 9 In the table below we see the proportion of violations for the 1, 5 and 10-day VaR at a 1% and 5% level for 3 assets (Bank of America Corp., Boeing Co. and American Express Inc.) estimated with the 4 copulas. The tables are further subdivided into daily and weekly data, meaning that if the data frequency matches the VaR horizon it's a direct forecast and an iterative forecast if it doesn't, which is only the case in table 5 and 6 (the data frequency matches the forecast horizon). If a proportion is highlighted in bold it means that copula performs the best. Note that almost all violation proportions are rejected on a 1,5 and 10% level, but isn't the most relevant thing in terms of copula comparison.

	Daily Data	1-day VaR	3 assets
	1%	5%	
Normal	0.020394	0.064401	
Student-t	**0.018247**	**0.060823**	
Clayton	0.019678	0.067621	
Factor	0.022182	0.06619	

		5-day VaR	3 assets	
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	0.022931	0.065926	0.023256	0.057245
Student-t	**0.022214**	**0.063418**	**0.021467**	**0.059034**
Clayton	0.024722	0.076317	0.025045	0.06619
Factor	0.024006	0.065926	0.0322	0.060823

		10-day VaR	3 assets	
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	0.024408	0.063891	0.019713	0.055556
Student-t	0.024049	**0.059225**	**0.017921**	**0.050179**
Clayton	0.03051	0.072864	0.021505	0.069892
Factor	**0.022972**	0.063532	0.023297	0.060932

Table 10 Evaluations of the MSE and relMSE for the four copulas at k=1, C=2% and C=3%. This is done for a set of 3 assets at forecast horizons of 1 day, 5 days and 10 days and also for weekly data and daily data. The tables are further subdivided into daily and weekly data, meaning that if the data frequency matches the VaR horizon it's a direct forecast and an iterative forecast if it doesn't, which is only the case in table 8 and 9 (the data frequency matches the forecast horizon).

	Daily Data	1-day	3 assets	
	2%		3%	
	MSE	relMSE	MSE	relMSE
Normal	3.407898	5264.273	3.80952	3796.541
Student-t	3.557859	**4842.8**	4.030222	**3341.057**
Clayton	**2.677525**	7164.455	**3.021236**	4659.288
Factor	3.642793	5162.564	4.273537	4030.475

			5-day	3 assets				
	Daily Data				Weekly Data			
	2%		3%		2%		3%	
	MSE	relMSE	MSE	relMSE	MSE	relMSE	MSE	relMSE
Normal	13.11284	7383.619	14.16692	6219.431	2.754404	1323.853	3.019819	1115.717
Student-t	13.5137	**6992.852**	14.65297	**5920.129**	2.821868	**1280.562**	3.103167	**1098.519**
Clayton	**10.05721**	17887.79	**10.56609**	12853.26	**1.965069**	3453.48	**2.096726**	2440.087
Factor	12.98357	7450.839	14.01059	6248.234	2.602277	1441.143	2.851329	1603.006

			10-day	3 assets				
	Daily Data				Weekly Data			
	2%		3%		2%		3%	
	MSE	relMSE	MSE	relMSE	MSE	relMSE	MSE	relMSE
Normal	24.60727	8667.169	25.83546	8667.169	5.182429	1443.479	5.489443	1261.111
Student-t	25.30892	**8340.578**	26.63484	**8340.578**	5.303215	**1421.819**	5.63855	**1241.237**
Clayton	**19.21215**	403618.4	**19.68197**	403618.4	**3.719341**	4744.936	**3.836381**	3373.299
Factor	25.05776	9902.309	26.27326	9902.309	4.925675	1683.426	5.184519	1445.443

Table 11 In the table below we see the proportion of violations for the 1, 5 and 10-day VaR at a 1% and 5% level for 10 assets estimated with the 4 copulas. The tables are further subdivided into daily and weekly data, meaning that if the data frequency matches the VaR horizon it's a direct forecast and an iterative forecast if it doesn't, which is only the case in table 5 and 6 (the data frequency matches the forecast horizon). If a proportion is highlighted in bold it means that copula performs the best. Note that almost all violation proportions are rejected on a 1,5 and 10% level, but isn't the most relevant thing in terms of copula comparison.

	Daily Data	1-day VaR	10 assets
	1%	5%	
Normal	0.021825	0.063327	
Student-t	**0.016458**	**0.056172**	
Clayton	0.030769	0.073345	
Factor	0.026476	0.101968	

		5-day VaR	10 assets	
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	0.020781	0.062702	0.021467	0.08229
Student-t	**0.013615**	**0.055894**	**0.017889**	**0.078712**
Clayton	0.034755	0.075242	0.055456	0.139535
Factor	0.055177	0.133644	0.019678	0.094812

		10-day VaR	10 assets	
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	0.017229	0.056712	0.023297	0.057348
Student-t	**0.015793**	**0.049892**	**0.017921**	**0.055556**
Clayton	0.027997	0.075377	0.050179	0.137993
Factor	0.091888	0.178033	0.019713	0.086022

Table 12 Evaluations of the MSE and relMSE for the four copulas at k=1, C=2% and C=3%. This is done for a set of 10 assets at forecast horizons of 1 day, 5 days and 10 days and also for weekly data and daily data. The tables are further subdivided into daily and weekly data, meaning that if the data frequency matches the VaR horizon it's a direct forecast and an iterative forecast if it doesn't, which is only the case in table 8 and 9 (the data frequency matches the forecast horizon).

Daily Data 1-day VaR 10 assets

	2%		3%	
	MSE	relMSE	MSE	relMSE
Normal	1.7582	3363.5837	1.8281	2432.3192
Student-t	1.7863	3231.2469	1.8919	1733.5763
Clayton	**1.1672**	151348.1798	**1.1632**	32437.7612
Factor	2.2015	**2896.4436**	2.5765	**1503.2677**

5-day VaR 10 assets

	Daily Data 2%		Daily Data 3%		Weekly Data 2%		Weekly Data 3%	
	MSE	relMSE	MSE	relMSE	MSE	relMSE	MSE	relMSE
Normal	7.196	7448.342	7.924	5271.238	2.118	2144.943	2.373	1547.844
Student-t	7.302	7090.596	8.092	**4989.648**	2.176	**2066.336**	2.445	**1524.343**
Clayton	**5.628**	111757733.805	**5.779**	10262889.316	**1.651**	9426769.834	**1.676**	596981.756
Factor	8.160	**5245.061**	9.658	7172.159	1.975	3368.806	2.292	2183.096

10-day VaR 10 assets

	Daily Data 2%		Daily Data 3%		Weekly Data 2%		Weekly Data 3%	
	MSE	relMSE	MSE	relMSE	MSE	relMSE	MSE	relMSE
Normal	13.978	168698.083	14.780	11634.234	3.963	2828.852	4.247	2040.199
Student-t	14.023	**13464.946**	14.858	**8006.324**	4.081	**2632.823**	4.393	**1923.117**
Clayton	**11.447**	48096487.818	**11.512**	2888778.095	**3.224**	16091079.368	**3.233**	15779378.291
Factor	16.390	15597.028	18.398	12071.243	3.647	5060.749	3.866	3150.009

Table 13 In the table below we see the proportion of violations for the 1, 5 and 10-day VaR at a 1% and 5% level for 3 assets (Bank of America Corp., Boeing Co. and American Express Inc.) estimated with the 4 copulas for the subperiod from 2-1-2003 until 2-1-2007. The tables are further subdivided into daily and weekly data, meaning that if the data frequency matches the VaR horizon it's a direct forecast and an iterative forecast if it doesn't, which is only the case in table 5 and 6 (the data frequency matches the forecast horizon). If a proportion is highlighted in bold it means that copula performs the best. Note that almost all violation proportions are rejected on a 1,5 and 10% level, but isn't the most relevant thing in terms of copula comparison.

Daily Data	1-day VaR	3 assets
	1%	5%
Normal	0.012346	0.029721
Student-t	**0.009602**	0.028349
Clayton	**0.009602**	0.033379
Factor	0.017833	**0.045267**

	5-day VaR	3 assets		
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	0.015117	0.037563	0.018265	0.03653
Student-t	**0.013284**	0.036647	**0.009132**	0.038813
Clayton	0.015117	**0.040311**	0.013699	**0.043379**
Factor	0.038937	0.077874	0.02968	0.061644

	10-day VaR	3 assets		
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	**0.01056**	0.038567	0.016018	0.034325
Student-t	**0.01056**	0.029385	**0.01373**	0.029748
Clayton	0.01607	**0.039486**	0.016018	**0.048055**
Factor	0.066116	0.121212	0.025172	0.066362

Table 14 In the table below we see the proportion of violations for the 1, 5 and 10-day VaR at a 1% and 5% level for 10 assets estimated with the 4 copulas for the sub period from 2-1-2003 until 2-1-2007. The tables are further subdivided into daily and weekly data, meaning that if the data frequency matches the VaR horizon it's a direct forecast and an iterative forecast if it doesn't, which is only the case in table 5 and 6 (the data frequency matches the forecast horizon). If a proportion is highlighted in bold it means that copula performs the best. Note that almost all violation proportions are rejected on a 1,5 and 10% level, but isn't the most relevant thing in terms of copula comparison.

Daily Data	1-day VaR	10 assets
	1%	5%
Normal	0.01326	0.029721
Student-t	**0.010059**	0.02652
Clayton	0.01326	0.029264
Factor	0.017375	**0.047554**

	5-day VaR	10 assets		
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	**0.012368**	**0.033898**	0.011416	**0.045662**
Student-t	0.007787	0.027027	**0.009132**	0.041096
Clayton	0.014659	0.032066	0.031963	0.079909
Factor	0.034814	0.081081	0.018265	0.06621

	10-day VaR	10 assets		
	Daily Data		Weekly Data	
	1%	5%	1%	5%
Normal	**0.01056**	0.038567	**0.011442**	**0.032037**
Student-t	**0.01056**	0.029385	**0.011442**	0.022883
Clayton	0.01607	**0.039486**	0.022883	0.077803
Factor	0.066116	0.121212	0.016018	0.06865

Figure 1 Scatter plots from the four bivariate distributions, all with N(0,1) marginals and linear correlation of 0.5, constructed using the four different copulas. The skew t-t factor copula has a skewness parameter of -0.5 and degrees of freedom of 4. The Student-t also has 4 degrees of freedom. For the Clayton copula this means an alpha of 1. Standard normal returns/variables are constructed as in the beginning of section 3.6.

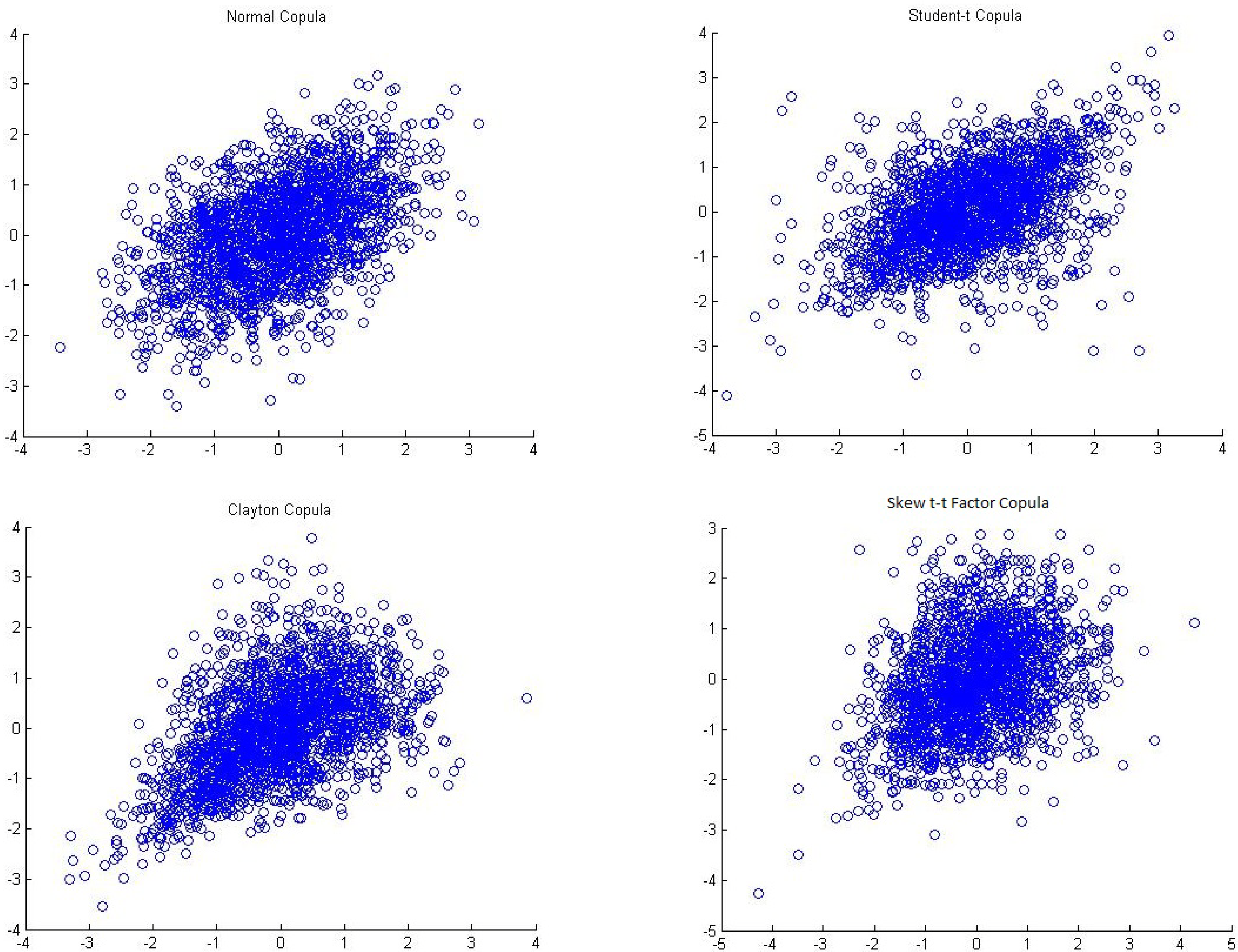
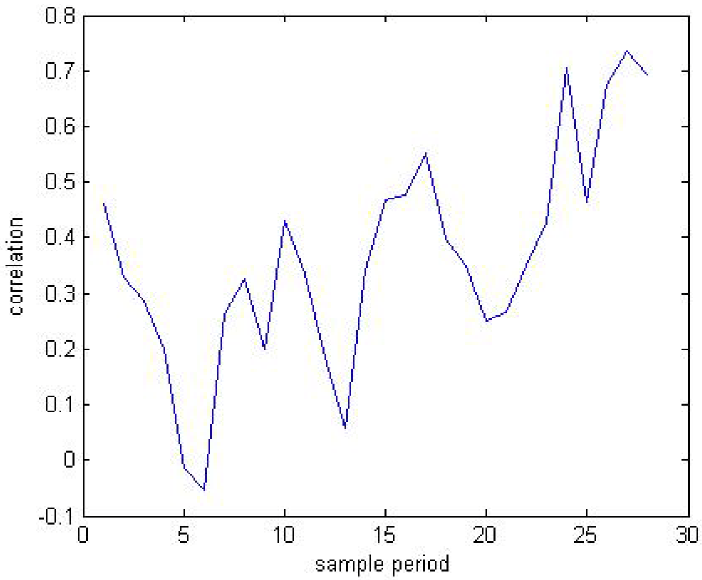


Figure 2 Correlations of subperiods of entire sample period with non overlapping windows of 200 days. In the graphs AXP vs. ABT and AXP vs. AMZN, respectively.

AXP vs ABT



AXP vs. AMZN

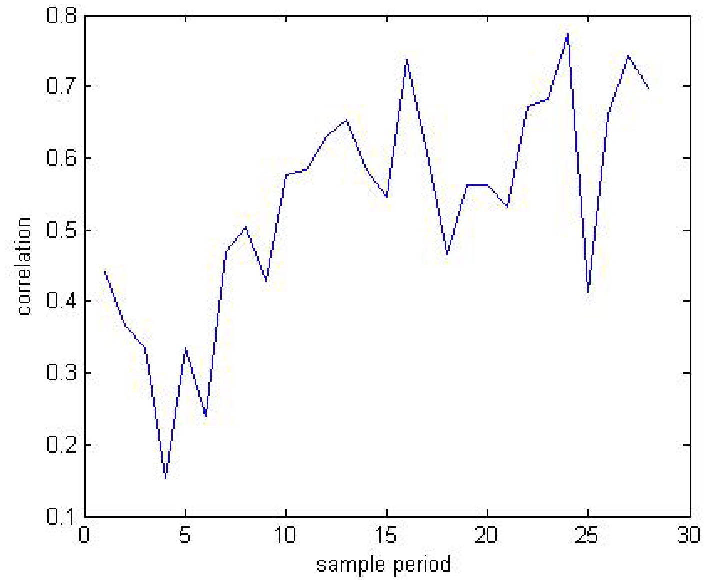


Figure 3 Quantile dependences at q=0.05 of subperiods of entire sample period with non overlapping windows of 200 days.

AXP vs. ABT

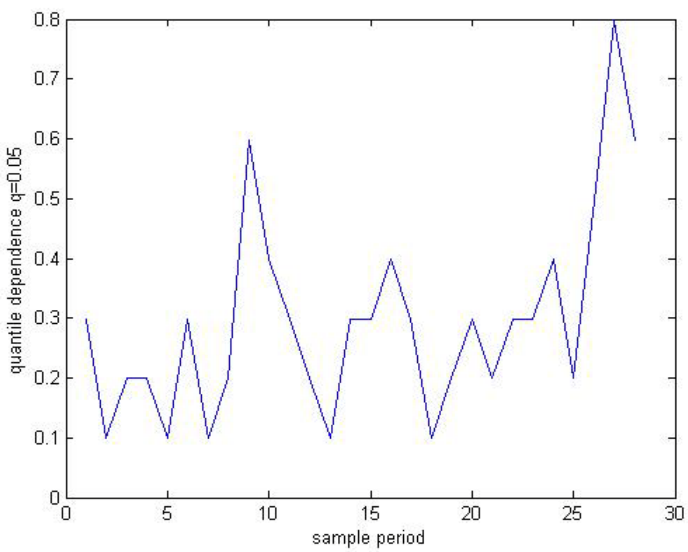


Figure 4 Scatter plots from the four bivariate distributions, with N(0,1) marginals, are constructed using the four different copulas with the in-sample parameter estimates of two stocks (American Express Inc. and Baxter International inc.). These can be seen in table 7. Standard normal returns/variables are constructed as in the beginning of section 3.6.

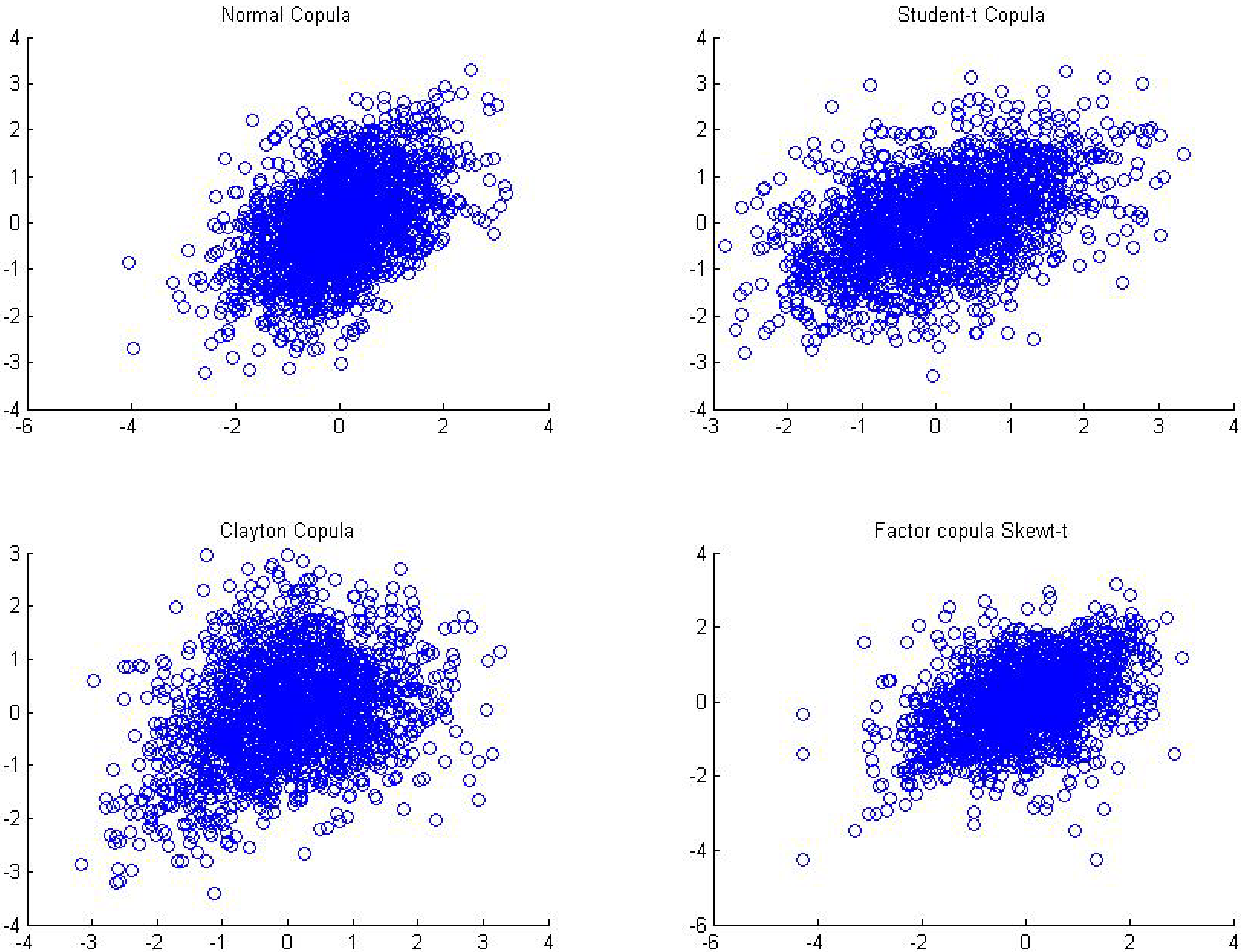
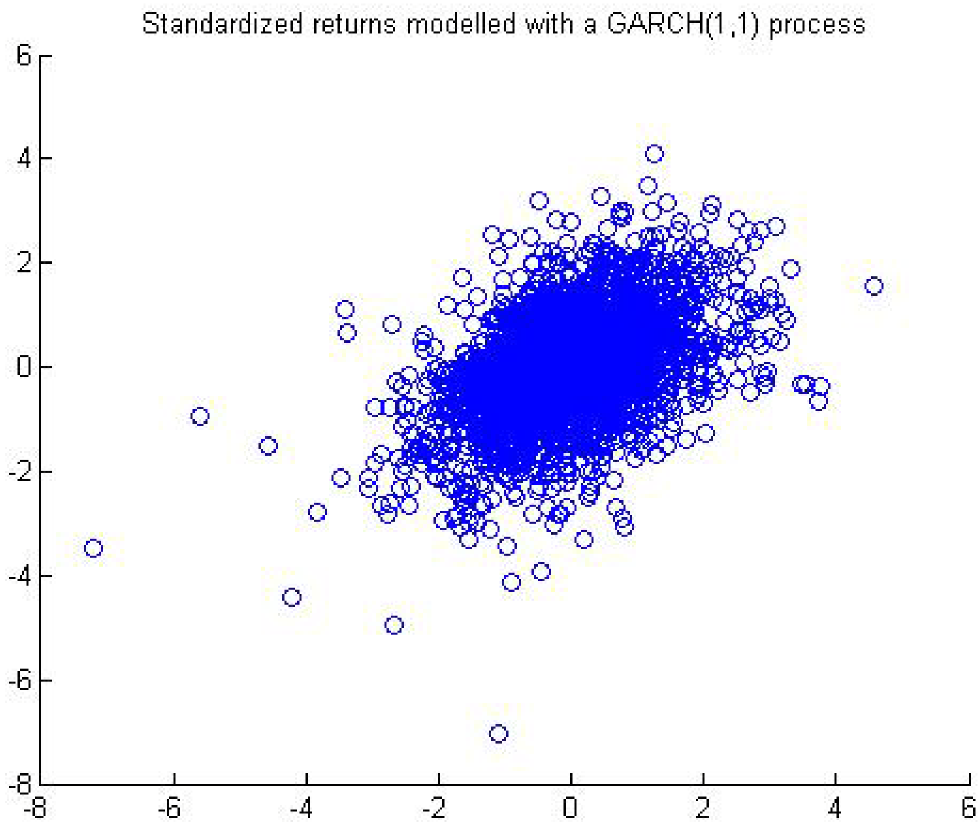


Figure 5 Scatter plot of the standardized returns (American express inc. and Baxter international inc.) modelled with a GARCH(1,1) process and standard normal returns. This is done for the in-sample.



Appendix A

The implementation of the SMM estimator of Oh and Patton (2012) requires measures of dependence. Oh and Patton (2012) choose those measures of dependence that are solely affected by the copula and not the marginal distributions, which make it 'pure' measures of dependence. Out of their preliminary studies they conclude that the pair-wire rank correlation and quantile dependence with q=[0.05 0.1 0.9 0.95] are best suited for the job, which gives 5 measures of dependence.

Let κ_{ij} denote a measure of dependence between variable I and j and let the following matrix be the 'pair-wise dependence matrix':

$$D = \begin{bmatrix} 1 & \kappa_{12} & \cdots & \kappa_{1M} \\ \vdots & & \ddots & \vdots \\ \kappa_{1M} & \kappa_{2M} & \cdots & 1 \end{bmatrix}$$

For the model that is used in this paper, the final measure that is used is the vector $[\overline{\kappa_1}, \ldots, \overline{\kappa_m}]$. This vector consists out of the averages of every row:

$$\overline{\kappa_\iota} = \frac{1}{N}\sum_{j=1}^{N} \widehat{\kappa_{\iota J}}$$

Hence, for every dependence measure, we get a vector of these measures. This totals to 5 dependence measures times the amount of measures in the vector (in other words the amount of variables), which leaves 5*M measures in total.

Appendix B

As the factor copula does not have a closed form so that random draws can be taken from it, the factor copula must be simulated. These simulations are just the simulations of Y like in expression 13. With enough simulations you create the copula density. Once simulated, a random simulated sample can be created by first drawing a large number of Y again, see expression 13. Then, these draws can be converted into standardized uniformly distributed variables by transforming the random draws in combination with the simulated copula density. Let C denote the simulations for the copula and Y_S denote the random draws.

$$u_S = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\{Y_s < C_i\}}$$

Where u_S stands for a uniform distributed random draw, N is the number simulated values for the density, 1 is an indicator function and equals 1 when the statement between the brackets is true.

Appendix C

Company	Ticker
American Express Inc.	AXP
Abbott Laboratories	ABT
Amazon.com	AMZN
American International Group Inc.	AIG
Amgen Inc.	AMGN
Apache Corp.	APA
Apple	AAPL
Bank of America Corp.	BAC
Baxter International Inc.	BAX
Boeing Co.	BA