Dependence in the Dutch market explained using a skewed t-copula in pair-copula models.

Bachelor thesis

Econometrics and Operational Research 2014

Tim Grevelink 348967

June 30, 2014

Abstract

The three-dimensional pair-copulas model is used during this research. Different from other research is the bivariate skewed t-copula, which is attempted to use in the pair-copula construction. First the residuals are computed with the use of an AR-GJR-GARCH model followed by Kolmogorov-Smirnov goodness of fit tests. After that the copula construction is explained, the bootstrap version of the Anderson-Darling test is used as goodness of fit test for the copula models. The limitations of the skew t-copula in a pair-copula model are described. All constructed models did not describe the data correctly, according to the bootstrap version of the Anderson-Darling test. Besides, the value at risk outcomes were not always in favour of the comparison model (the multivariate Student's t-copula).

Contents

1 Introduction

Modelling the dependence structure among variables is essential in financial markets. Investors and portfolio managers try to analyse their potential returns with the best proven methods. Explaining the dependence of several variables is an important aspect to compute reliable risk measures.

Some authors argue that the lineair correlation coefficient is the most used measure to test dependence (Aas (2004)). However, a pitfall of this measure is that it captures only lineair dependence of the variables. One of the ways to address this problem is by the use of copula functions. Copulas were first introduced by Sklar (1959) and are specifically designed to model non-lineair dependence structures. This is done by taking marginal distributions as arguments in the multivariate joint density function. One of the great advantages of the copulas is that the marginal distributions can be tested separately by performing different goodness-of-fit tests, helping us to find a better joint-density function. There are many different copula models that differ with regard to the distribution functions used. For example, the Gaussian copula is a standard market copula, which uses the normal distribution for all variables (Li (2000)).

The pair-copula model is a specific kind of copula, based on several bivariate copulas (Aas et al. (2009)). This kind of modelling helps us to built multivariate copula models. The number of variables involved is dependent on the complexity of the pair-copula model. As the pair-copula model uses different distributions, joint extreme events have more chance of occurring. This has the added advantage that joint-extreme events can be explained more easily. One of the leading scholars Aas et al. (2009) uses four bivariate copulas, namely: the Gaussian, Student, Clayton and Gumbell copulas. These differ with regard to dependence structures. Previous research shows that the three dimensional pair-copula model, consisting of three bivariate Student's t-copula's, is the best fitted model (Grevelink et al. (2014)). However, as such fit is dependent on the data, the question rises whether this pair-copula model also shows a significant fit with other data.

This research focuses on the fit of this Student's t-copula in other settings. More specifically, this research tests the Student's t-copula together with the Gaussian copula using other data (AEX index, AMX index and Netherlands benchmark 10 year bond index). We go one step further by introducing a different kind of Student's t-copula, namely: *extended skewed t-copula*(EST) Arellano-Valle and Genton (2010). This special Student's t-copula has a skewness- and a shape parameter, in contrast to the 'normal' Student's t-copula. In order to be able to determine whether the model shows a better fit, we need a point of reference. We compare the pair-copula model with is the multivariate Student's t-copula. The main reason to choose the multivariate Student's t-copula is because research shows that market returns are often better described using fatter tailed distributions (Alexander (2008)).

This research starts from two different hypotheses. The first hypothesis is that the pair-copula model using the EST copula has a better fit than the multivariate Student's t-copula. This is tested with a bootstrap version of the Anderson-Darling(AD) goodness-of-fit test (Anderson and Darling (1954); Genest et al. (2009)). The second hypothesis focuses on the value at risk(VaR). It states that the pair-copula model using the bivariate

EST copula, has better outcomes of the VaR than the multivariate Student's t-copula model.

It is impossible for us to compare the bivariate EST copula with the other copulas, as no results are generated. We suspect this is due to increased calculation time, as a result of the numerical methods that were used to compute the inverse of the EST distribution. We found results for the other models we tested. Although the bootstrap of the AD test shows that none of the models has a significant fit. However, there are minor indications of a better fit for both the pair-copula model with three bivariate Student's t-copulas and the multivariate Student's t-copula model than the other tested models. This is not the case when comparing the VaR outcomes. The pair-copula model with three bivariate Gaussian copulas performs best for the 95%-VaR.

The outline of this paper is as follows. The next section describes the data. Then the methods used are discussed. We specifically will elaborate on the extended skewed t-copula. Furthermore, goodness-of-fit tests are performed to search for both significantly fitted AR-GJR-GARCH models and copula models. The method section is followed by the results, conclusion and discussion of the results.

2 Data

During this research there are three assets chosen to describe the underlying dependence. Because of the focus on the Dutch market, the following assets are selected: the AEX index, AMX index and Netherlands benchmark 10-year bond index. The dataset is based on 6198 daily observations in the period of 1 May 1989 until 1 May 2014. Figure 1 shows the price indices of the chosen assets.

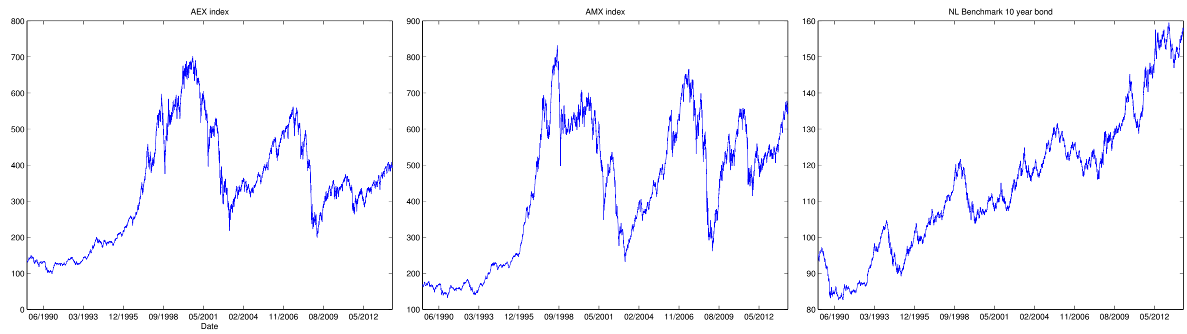

Figure 1: The price indices of AEX, AMX and 10-year bond Benchmark in the portfolio during the period 01/05/1989 - 01/05/2014.

One of the properties that is required during this research, is the stationarity of the time-series. This is a requirement of the AR-GJR-GARCH model which is described in Section 3.1. The log-price indices are tested for a unit root with the Augmented Dickey Fuller test, see Table 1. The null hypotheses, of a unit root, are not rejected. Therefore, we conclude that the log-price indices are non-stationary, which is suspected from Figure 1. To solve this problem the log-returns are used in the AR-GJR-GARCH model where the unit-root is excluded. Equation (1) shows the construction of the log-returns of Figure 2.

$$r_{it} = 100 \cdot (\log(p_{i,t}) - \log(p_{i,t-1})) \quad (1)$$

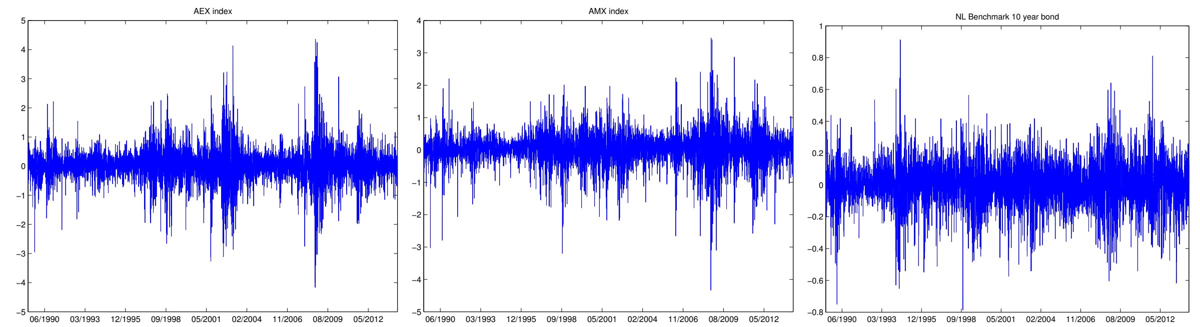

Figure 2: The log-returns of the portfolio during the period 01/05/1989 - 01/05/2014.

log(price index$_{i,t}$)	Test statistic	P-value
AEX	1.042	0.922
AMX	-0.879	0.330
Benchmark	-0.968	0.298

Table 1: Augmented Dickey Fuller test statistics of the log-price indices.

Table 2 presents the statistical properties of the first difference log-returns. The Jacque-Bera statistics indicate that all the assets are non-normally distributed. The statistics are way larger than 5.98, which is the critical value. In Figure 4 it can be seen that the assets are non-normal distributed. There is much more weight in the peaks and also in the tails compared to a normal distribution. A Student's t-distribution is possibly more useful because of the higher kurtosis. The negative skewness indicates for more negative log-returns than positive log-returns, which can supports the choice for the AR-GJR-GARCH model (Section 3.1).

	AEX	AMX	Benchmark
Mean	0.008	0.010	0.004
Median	0.026	0.040	0.008
Maximum	4.355	3.462	0.913
Minimum	-4.165	-4.335	-0.786
Std. dev.	0.592	0.523	0.141
Skewness	-0.163	-0.582	-0.229
Kurtosis	9.502	8.226	5.097
Jarque-Bera	10946	7402	1190
P-value	0.001	0.001	0.001

Table 2: Descriptive statistics of the log-returns with sample size: 6197

Another property is that the assets should be independent over time. Therefore, the log-returns are checked for autocorrelation. Figure 3 shows the autocorrelation of the first difference log-returns. The autocorrelation is not frequently outside the boundaries of the 95% confidence intervals, which is an indication for almost no autocorrelation. Different tests for autocorrelation are explained further in the next section.

The correlations between the different assets, based on Spearman's rank, are shown in

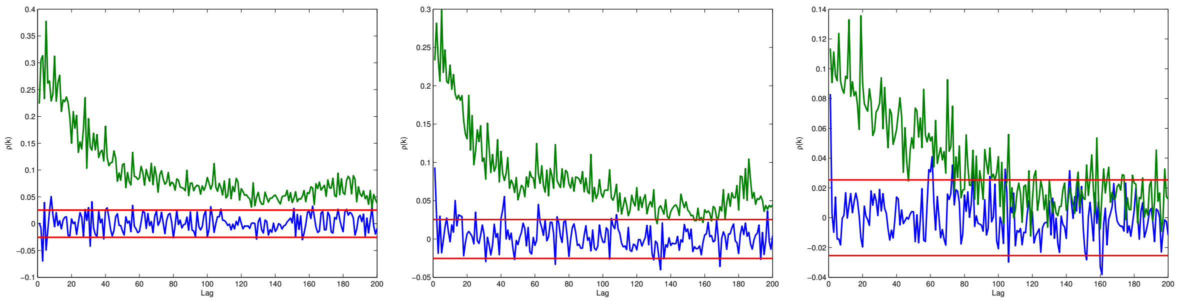

Figure 3: Autocorrelation of log-returns and squared log-returns for the period 01/05/1989 - 01/05/2014.

Figure 5. It can be seen that the AEX and the AMX are highly correlated. When there is a negative event in one of the cases, it is likely that the other is also negative and vice-versa. Joint-extreme events can be modelled extremely well using copula functions. However, the other two dependencies are unremarkable. In these cases it is not clear which copula is useful to choose.

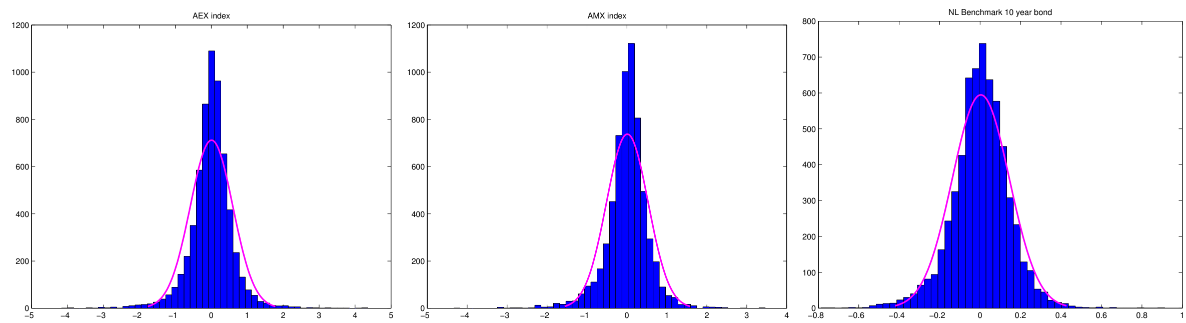

Figure 4: Histograms of the log-returns for the period 01/05/1989 - 01/05/2014.

3 Methodology

This research continues on Grevelink et al. (2014). There are some changes, like the used bivariate copulas and change of comparison model. The AR-GJR-GARCH model is explained in the next section. After that the pair-copula with the used bivariate copulas are introduced. Then the multivariate Student's t copula is described just like the parameter estimations. The Anderson-Darling test is used to find the significantly fitted models and also value at risk is used to some of the constructed models.

3.1 AR-GJR-GARCH

An assumption of the pair-copula model is that the observations of the returns are independent over time. This is the case if there is no serial correlation in the conditional mean and conditional variance. Therefore, the AR-GJR-GARCH model is constructed (Equation (2)).

$$r_{i,t} = \phi_{0,i} + \phi_{1,i} \cdot r_{i,t-1} + \epsilon_{i,t} \tag{2}$$

$$\sigma_{i,t}^2 = \omega_i + \beta_i \cdot \sigma_{i,t-1}^2 + \alpha_i \cdot \epsilon_{i,t-1}^2 + \gamma_i \cdot \epsilon_{i,t-1}^2 \cdot I \tag{3}$$

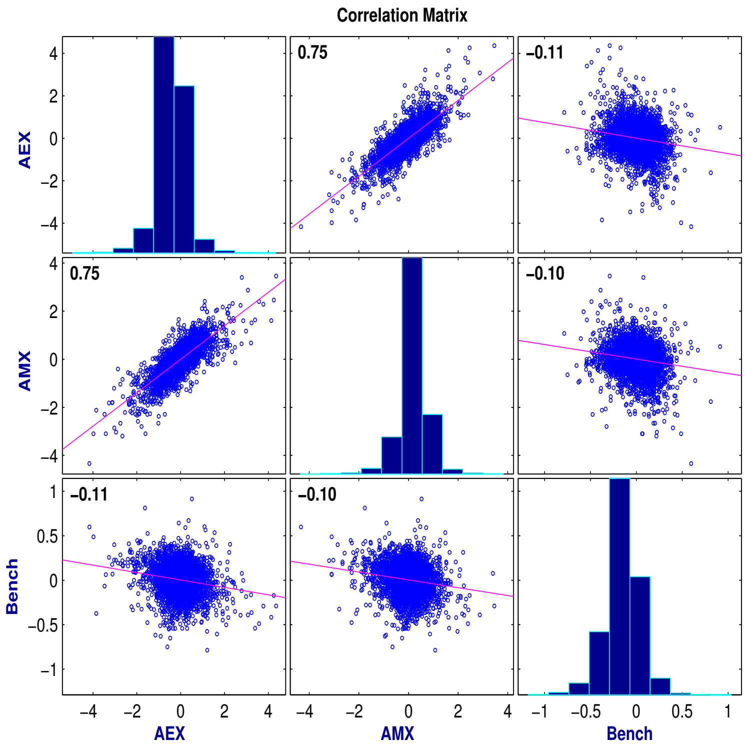

Figure 5: Correlations based on Spearman's rank between the different assets in the period 01/05/1989 - 01/05/2014.

$$\epsilon_{i,t} = \sigma_{i,t} \cdot \eta_{i,t} \qquad \eta_{i,t} \sim N(0,1) \tag{4}$$

$$I = \begin{cases} 1 & \text{if } \epsilon_{i,t-1} < 0 \\ 0 & \text{if } \epsilon_{i,t-1} \geq 0 \end{cases} \tag{5}$$

In Grevelink et al. (2014) is found that a pair-copula with bivariate Student's t-copulas was significantly fitted in all cases. It is therefore likely that models with Student's t-distributed errors fits better than the models with normal distributed errors. Therefore, the models are tested with normal- and Student's t-distributed errors in this research. The models that fit best, according to the goodness of fit tests (Section 3.7) are used during the remainder of this research.

The assumption of the pair-copula model implies that there is no autocorrelation in the standardized residuals and also not in the squared standardized residuals of the used model. This is checked with the Q-statistic Ljung and Box (1979) and the ARCH Lagrange Multiplier(LM) test Engle (1982). The Q-statistic tests whether the standardized residuals contain significant autocorrelations. The null hypothesis of the ARCH LM-test states that the standardized residuals are homoscedastic. If there is still autocorrelation, a lag will be added. This is done until there is no autocorrelation left. The standardized residuals of the AR(k)-GJR-GARCH(m,1) model are used in the further research, where k is chosen first to be the best fitting AR model. Than m are the amount of lags that are used to take away the autocorrelation.

After the correct AR-GJR-GARCH models are found, the errors are used in the copula construction. The marginal distributions of the variables should be used in a copula model. However, if they are unknown, as in this case, the data can be transformed without any influence on the dependence, see Genest and Favre (2007). The normalized ranks of the standardized residuals are used to estimate the parameters of the pair-copula models, see (6).

$$U_i = R_i/(n+1) \tag{6}$$

where R_i are the ranks of the standardized residuals of asset i.

3.2 Pair-copula decomposition

The pair-copula model is described in Aas et al. (2009). It uses several bivariate copulas to describe the dependence structure among multiple variables. One of the possible joint density functions f for the pair-copula decomposition, with three variables, is:

$$\begin{aligned} f(x_1,x_2,x_3) = f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3)\cdot c_{12}\{F_1(x_1),F_2(x_2)\}\cdot c_{23}\{F_2(x_2),F_3(x_3)\} \\ \cdot\, c_{13|2}\{F(x_1|x_2),F(x_3|x_2)\} \end{aligned} \tag{7}$$

It seems obvious that there are another two possible pair-copula decompositions, because of the last conditional copula. There are three bivariate copulas c_{12}, c_{23} and $c_{13|2}$ used in this model. To construct the last conditional copula, there is also a conditional distribution needed. This is the major difference of the pair-copula model in comparison with other copula models. The conditional distribution is in this research denoted as the h-function (Equation (8)). The h-function is one of the characteristics of the pair-copula model and is not needed in every copula model. The conditioning with the h-function is also one of the difficulties of this model and is therefore explained a bit further.

$$h(x_1,x_2,\Theta) = F(x_1|x_2) = \frac{\partial C_{x_1,x_2}(x_1,x_2,\Theta)}{\partial x_2} \tag{8}$$

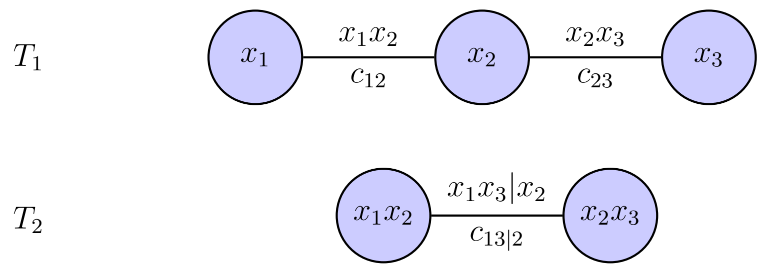

Figure 6: Graphical model with dependent variable x_2. Tree T_1 shows the unconditional variables and tree T_2 shows the conditional variables.

In Figure 6, the choice of the dependent variables is shown graphically. x_2 is the variable that has the strongest dependence with the other two assets. Therefore, that is the variable to condition on. To select the dependent variables x_2, the Spearman's rho and Kendall's tau are computed, as can be seen in Equation (9) and (11). The variable that has the strongest dependence according these measurements is chosen to be the dependent variable. These two measurements are based on the normalized ranked residuals, which are computed earlier. Spearman's rho:

$$\rho_n = \frac{\sum_{i=1}^{n}(R_i-\bar{R})(S_i-\bar{S})}{\sqrt{\sum_{i=1}^{n}(R_i-\bar{R})^2\sum_{i=1}^{n}(S_i-\bar{S})^2}} \in [-1,1] \tag{9}$$

where

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} R_i = \frac{n+1}{2} = \frac{1}{n}\sum_{i=1}^{n} S_i = \bar{S} \tag{10}$$

(R_i, S_i) is a ranked pair of (X_1, X_2), which are two of the three assets we use.

Kendall's tau:

$$\tau_n = \frac{P_n - Q_n}{\binom{n}{2}} = \frac{4}{n(n-1)} P_n - 1 \tag{11}$$

with P_n and Q_n the number of concordant and discordant pairs, respectively. A concordant pair (X_i, Y_i), (X_j, Y_j) means that if $X_i < X_j$ than also $Y_i < Y_j$ or vice versa. A discordant pair means that if $X_i < X_j$ than $Y_i > Y_j$ and also vice versa.

Spearman's rho is defined as the Pearson correlation with ranked variables. Therefore, it is not only for lineair dependence structures. Kendall's tau is a measure for difference in probability between concordant and discordant pairs. The two measures are different, but lead in most cases to corresponding conclusions.

3.3 Bivariate copulas

To construct a pair-copula model there are bivariate copulas that have to be chosen. In Grevelink et al. (2014) the Gaussian, Student's t-, Clayton and Clayton survival copulas were used, just like in Aas et al. (2009). It was clear that the best choice was the bivariate Student's t-copula. Therefore, the Student's t- and also a skewed t-copula will be used in this research. The skewed t-copula is based on a skewed t-distribution. The skewness factor is a major difference of the skewed t-copula compared to the Student's t-copula. The Gaussian copula is the third bivariate copula that will be used.

Tail dependence is a measure for joint extreme co-movements, which differs per copula. The Gaussian copula is used when there is no tail dependence between the variables. The Student's t-copula covers both lower- and upper-tail dependence. Because of the non-normal distributed variables, this is probably a better choice. The skewed t-copula is a variant of the Student's t-copula, but has also a skewness parameter. This results in a Student's t-copula, which is not totally symmetric. Figure 7 shows some skewed t-copulas. When $\lambda = 0$ the Student's t- and the skewed t-copula are the same. In the case of a positive λ the negative events occur in less extreme cases. The peak shifts more to the middle and is flattend compared to the Student's t-copula. The same occurs with the positive events and a negative λ. The skewed t-copula is discussed below, while the other used bivariate copulas are described in Appendix A.

3.3.1 Skewed t-copula

In Grevelink et al. (2014) the pair-copula with bivariate Student's t-copulas, performed best. Therefore, a variant of this Student's t-copula is probably even better. Multiple authors describe skewed t-distributions, but none of them was suitable for the pair-copula model (Hansen (1994), Azzalini and Capitanio (2003)). The pair-copula model requires that the distribution that will be implemented has a conditional distribution. This was one of the pitfalls during this research.

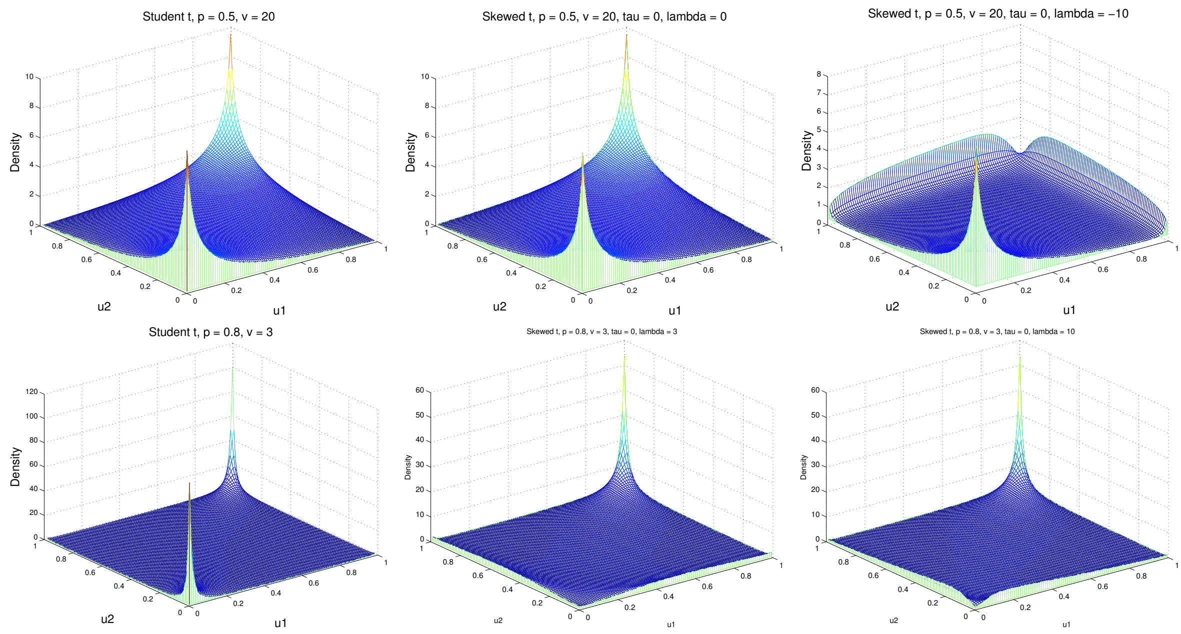

Figure 7: Simulation with the density functions of the bivariate Student's t- and the bivariate extended skewed t-copula(EST) with different chosen parameters.

As mentioned before, the extended skewed t-distribution of Arellano-Valle and Genton (2010) is used during this research. All the other skewed t-distributions are not closed under conditioning. Although the extended skewed t-distribution is suitable to use, this skewed t-distribution has not previously been used in the pair-copula model. Both the h-function and the h-inverse function should be formulated. Also the copula function of this distribution should first be constructed.

In Arellano-Valle and Genton (2010) there are described some properties of the extended skewed t-distribution(EST). This is a Student's t-distribution with two extra parameters, of which one is the skewness parameter. Below are the properties described that are needed for the construction of the pair-copula model.

The extended skewed t-distribution Before the EST copula is used, the distribution has to be implemented. For the bivariate copula the standardized univariate functions are used. Therefore, all the formulas in this section are described in the standard univariate case. The more general multivariate EST functions are stated in Appendix B.
$Y \sim EST_1(\lambda, \nu, \tau)$ and $y \in \mathbb{R}^p$ if the density function is:

$$\frac{1}{T_1(\tau/\sqrt{1+\lambda^2}; \nu} t_1(y;\nu) T_1 \left\{ (\lambda y + \tau)\sqrt{\frac{\nu+1}{\nu+y^2}}; \nu+1 \right\} \tag{12}$$

with

$$t_1(y;\nu) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)} \left(1+\frac{y^2}{\nu}\right)^{-\frac{\nu+1}{2}} \tag{13}$$

which denotes the standard univariate density function of the Student's t-distribution. $\lambda \in \mathbb{R}^p$ is the skewness parameter and $\tau \in \mathbb{R}$ is the extension parameter. $T_1(y;\nu)$ denotes the univariate standard Student's t cumulative distribution function with $\nu > 0$ degrees of freedom.

The density function above gives a better understanding of the distribution. However, the cumulative distribution function of the EST distribution is of greater importance to construct the pair-copula model. This CDF is given as follows.

$$P(Y \leq y) = \frac{1}{T_1(\bar{\tau};\nu)} T_2\left(\begin{pmatrix} z \\ \bar{\tau} \end{pmatrix}; \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}, \nu+1\right) \tag{14}$$

where $\bar{\tau} = \frac{\tau}{\sqrt{1+\lambda^2}}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.
$T_2(x;\Omega,\nu)$ denotes the multivariate 2-dimensional centred Student's t-cumulative distribution function, with 2×2 positive definite dispersion matrix Ω and ν degrees of freedom.

Besides the cumulative distribution function, also the inverse function is needed for the copula construction of this distribution. We are unaware of the exact representation of the inverse of this function. This inverse is not found in the literature and therefore we approximate it numerically. All parameters (λ, ν and ρ) and input vector a should be known in this case.

$$\min_{x\in\mathbb{R}} \{a - EST_1(x;\lambda,\nu,\tau) = 0\} \tag{15}$$

where a is the known input vector with probabilities and x the vector (with the inverse values) that should be found. This function is minimized with respect to x, which we find the corresponding x-values and denote the values of the inverse of the EST distribution. One of the drawbacks of this method is that computation can be slow. However, this is a method that is often used in this kind of problems.

Extended skewed t-copula There are different methods to construct the copula density function out of a distribution. The EST copula density is computed by dividing the bivariate EST probability density function by its univariate density functions. This method is described in Demarta and McNeil (2005) and results in the bivariate EST copula (Equation (16)).

$$\begin{aligned}
c(u_1,u_2;\rho_{12},\lambda_{12},\nu_{12},\tau_{12}) &:= \frac{\partial^2}{\partial u_1 \partial u_2} C(u_1,u_2;\rho_{12},\lambda_{12},\nu_{12},\tau_{12}) = \\
&= \frac{EST_2(EST_1^{-1}(u_1), EST_1^{-1}(u_2))}{EST_1(EST_1^{-1}(u_1))EST_1(EST_1^{-1}(u_2))} = \\
&= \frac{(T_1(\tau_{12}/\sqrt{1+\lambda_{12}^2};\nu_{12}))^2}{T_1(\tau_{12}/\sqrt{1+\rho_{12}\lambda_{12}^2+\lambda_{12}^2};\nu_{12})} c_{Student}(x_1,x_2;\rho_{12},\nu_{12}) \\
&\cdot \frac{T_1\left((\lambda_{12}x_1+\lambda_{12}x_2+\tau)(\frac{\nu_{12}+2}{\nu_{12}+Q(z)})^{1/2};\nu_{12}+2\right)}{T_1\left((\lambda_{12}x_1+\tau_{12})(\frac{\nu_{12}+1}{\nu_{12}+x_1^2})^{1/2};\nu_{12}+1\right) T_1\left((\lambda_{12}x_2+\tau)(\frac{\nu_{12}+1}{\nu_{12}+x_2^2})^{1/2};\nu_{12}+1\right)}
\end{aligned} \tag{16}$$

where $c_{Student}(x_1,x_2;\rho,\nu)$ is stated in Equation (36) and $Q(z) = x^T R^{-1} x$ with R the 2×2 correlation matrix of x_1 and x_2. With $x_1 = EST^{-1}(u_1;\rho_{12},\nu_{12},\tau_{12},\lambda_{12})$ and $x_2 = EST^{-1}(u_2;\rho_{12},\nu_{12},\tau_{12},\lambda_{12})$.

h-function EST The h-function is used to construct conditional copulas. This is one of the main differences of the pair-copula model in comparison with other copula models.

To construct this h-function, we used some properties of the EST distribution (Arellano-Valle and Genton (2010)). These properties are mentioned at the end of this section.

$$h(u_1, u_2, \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12}) =$$
$$EST_1\left\{ \frac{EST_1^{-1}(u_1; \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12}) - \rho_{12} EST_1^{-1}(u_2; \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12})}{\sqrt{\frac{\left(\nu_{12} + \left(EST_1^{-1}(u_2; \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12})\right)^2\right)(1-\rho_{12}^2)}{\nu_{12}+1}}}; \rho_{12}, \lambda_{1\cdot 2}, \nu_{12}+1, \tau_{1\cdot 2}^* \right\} \tag{17}$$

with $EST(x; \nu, \tau, \lambda)$ is the standard univariate cumulative extended skewed t-distribution distribution function with ν degrees of freedom, τ the extension parameter and λ the skewness parameter. Below is the h-inverse function given, which is used to simulate with the copulas.

$$h^{-1}(u_1, u_2, \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12}) =$$
$$EST_1\left\{ EST_1^{-1}(u_1; \lambda_{1\cdot 2}, \nu_{12}+1, \tau_{1\cdot 2}^*) \cdot \sqrt{\frac{\left(\nu_{12} + \left(EST_1^{-1}(u_2; \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12})\right)^2\right)(1-\rho_{12}^2)}{\nu_{12}+1}} \right.$$
$$\left. + \rho_{12} EST_1^{-1}(u_2; \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12}); \rho_{12}, \lambda_{12}, \nu_{12}, \tau_{12} \right\} \tag{18}$$

The derivation of the following properties can be seen in Appendix B. These are the results that are used to construct both the h-function and h-inverse function of the EST distribution.

$$E(X_1 | X_2 = x_2) = \rho_{12} x_2 \tag{19}$$
$$\Omega_{X_1|X_2=x_2} = \frac{\nu_{12} + x_2^2}{\nu_{12}+1}(1-\rho_{12}^2) \tag{20}$$
$$\nu_{X_1|X_2=x_2} = \nu_{12} + 1 \tag{21}$$
$$\tau_{X_1|X_2=x_2} = \tau_{1\cdot 2}^* = \left\{\frac{\nu + x_2^2}{\nu+1}\right\}^{-1/2} \cdot ((\lambda_{12}\rho_{12} + \lambda_{12})x_2 + \tau_{12}) \tag{22}$$
$$\lambda_{1\cdot 2} = (1-\rho_{12}^2)^{1/2} \cdot \lambda_{12} \tag{23}$$

3.4 Multivariate Student's t-copula

Because of the outcomes in Grevelink et al. (2014) it is relevant to compare the pair-copula models with the multivariate Student's t-copula (Equation (24)), see Aas (2004). It can be seen that in this case in contrast to the pair-copula construction, all the variables have the same degrees of freedom.

$$C_\nu(u_1, \ldots, u_n; \Sigma) = \int_0^{t_\nu^{-1}(u_n)} \cdots \int_0^{t_\nu^{-1}(u_1)} \kappa |\Sigma|^{-1/2} \left(1 + \nu^{-1} \mathbf{x}' \Sigma^{-1} \mathbf{x}\right)^{-(\nu+n)/2} dx_1 \ldots dx_n \tag{24}$$

Where $|\Sigma|$ is the determinant of the correlation matrix and

$$\kappa = \Gamma\left(\frac{\nu}{2}\right)^{-1} \Gamma\left(\frac{\nu+n}{2}\right) (\nu\pi)^{-n/2} \tag{25}$$

To estimate the parameters of this copula there is a method described in Aas (2004) where first $\mathbf{R}$ is computed with Kendall's tau, where after $\hat{\nu}$ is estimated, see Equation (27).

$$\hat{\mathbf{R}}_{ij} = \sin(\frac{\pi}{2}\hat{\rho}_\tau) \tag{26}$$

$$\hat{\nu} = \arg\max_\nu \left\{ \sum_{i=1}^{n} \log(c(\mathbf{U}_i, \nu, \hat{\mathbf{R}})) \right\} \tag{27}$$

The multivariate Student's t copula density can be written as:

$$c(\mathbf{u}) = \frac{\Gamma(\frac{\nu+d}{2})\Gamma(\frac{\nu}{2})^{d-1}\left(1+\frac{x'\mathbf{R}^{-1}x}{\nu}\right)^{\frac{\nu+d}{2}}}{|\mathbf{R}|^{1/2}\Gamma(\frac{\nu+1}{2})^d \prod_{j=1}^{d}(1+\frac{x_j^2}{\nu})^{-\frac{\nu+1}{2}}} \tag{28}$$

The statistical properties of the assets in Section 2 rejected normality. Market returns are often better described using a fatter tailed distribution, according to Alexander (2008)). Possibly a significantly fitted model is much easier to find with this multivariate Student's t-copula.

3.5 Parameter estimation of the pair-copula model

After the bivariate copulas are chosen, the parameters have to be estimated. This is done with the maximum pseudo-likelihood method, because of the unknown marginal distributions. With the normalized ranks, as constructed in (6), the parameters of the copula are estimated. The following log-likelihood has to be maximized.

$$L(\theta) = \sum_{i=1}^{n} \log c_\theta \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \tag{29}$$

The three-dimensional case can be seen in Equation (30) as stated in Aas et al. (2009). Here the $x_{i,t}$ are the normalized ranks of the i-th asset. The third copula in (30) is dependent on the other parameters. The algorithm of the construction of this log-likelihood function is shown in Aas et al. (2009).

$$L = \sum_{t=1}^{T} \big(\log c_{12}(x_{1,t}, x_{2,t}, \Theta_{11}) + \log c_{23}(x_{2,t}, x_{3,t}, \Theta_{12}) + \log c_{13|2}(v_{1,t}, v_{2,t}, \Theta_{21}) \big) \tag{30}$$

with $v_{1,t} = h(x_{1,t}, x_{2,t}, \Theta_{11})$ and $v_{2,t} = h(x_{3,t}, x_{2,t}, \Theta_{12})$

In this research there are three-bivariate copulas that can be used. Therefore there are $3^3 = 27$ possibilities for the pair-copula model. The choice of the used copula influences the amount of parameters that have to be estimated. Every bivariate Gaussian copula adds one parameter to the pair-copula model. For the bivariate Student's t-copula these are two parameters and in the case of the bivariate EST copula these are three parameters. First we assume that the extension parameter $\tau = 0$ is fixed (Arellano-Valle and Genton (2010)). This simplifies the maximum likelihood procedure, after which we can try to find a higher log-likelihood by changing this value of τ.

3.6 Simulation

After constructing a pair-copula model, there will be done tests for the correctness of this model. To perform a goodness of fit test (Section 3.7) and also for the calculation of value at risk (Section 3.8) there are simulations needed of the bivariate copulas. First u_1, u_2 and u_3 are generated from an independent uniform random distribution. These variables are transformed with the bivariate copula that is used for that particular pair. This transformation uses both the h-function and the h-inverse function to compute variables with a dependence described by a specific pair-copula model. The algorithm for this simulation is extensively described in Aas et al. (2009)).

3.7 Goodness-of-fit

For the AR-GJR-GARCH models there are two tests performed. The first one is the Kolmogorov-Smirnov test (Massey Jr (1951)). This test is based on the difference between the distributions. Under the null hypothesis the cumulative distribution of the sample is the same as the tested one ($= F_0$). The observed cumulative distribution ($= S_N(x) = k/N$ where k is the total amount of observations smaller or equal to x) is compared with the F_0, with the Kolmogorov test statistic $d = \sup_x |F_0(x) - S_N(x)|$. In the case of the Student's t-distributed errors, F_0 will be chosen to be a Student's t cumulative distribution function with estimated parameters.

Anderson and Darling (1954) describes the Anderson-Darling(AD) goodness-of-fit test. This test is also applied to the AR-GJR-GARCH models. In that case the comparison model is again the same as above described. Using the probability integral transform(PIT) method, as introduced in Rosenblatt (1952), this test is applied to find significantly fitted pair-copula models. PIT is used to construct i.i.d. U(0,1) variables of the data. The variables are made independent, which can be seen as the opposite of the method used for simulations.

Let U_1, U_2 and U_3 denote the normalized ranks of the standardized residuals with marginal distributions $F(u_i)$ and conditional distributions $F(u_i|u_1, \ldots, u_{i-1})$ for $i = 1, 2, 3$. The PIT of U_1, U_2 and U_3 are defined by $R(U_1)$, $R(U_2)$and $R(U_3)$, respectively. Moreover, $R(U_i)$ is given by:

$$
\begin{aligned}
R(U_1)&=F(u_1)\\
R(U_2)&=F(u_2|u_1)\\
R(U_3)&=F(u_3|u_1, u_2)
\end{aligned}
$$

With this transformed data, the AD test can be performed. The null hypothesis of this test is that if the variables $R(U_i)$ are independent uniformly distributed, then the pair-copula model is correctly chosen. When $S = \sum_{i=1}^{n} \{\phi^{-1}(R(U_i)\}^2$ is computed, under the null S should come from a chi-square distribution with 3 degrees of freedom. S is a vector of T values, which is the number of observations.

$$
\text{AD} = -T - \frac{1}{T} \sum_{j=1}^{T} (2j-1) \cdot [\ln(F_0(S_j)) + \ln(1 - F_0(S_{T-j+1}))] \tag{31}
$$

where $j = 1, \ldots, T$ and $S_1 \leq \cdots \leq S_T$. F_0 is the cumulative distribution function of a

χ^2-distributed variable.

The AD statistic will be computed for every used model. However, the critical values that are computed are incorrect, during the different transformations that are made before using the data. To solve this issue, a bootstrap version is applied. Dobrić and Schmid (2007) and Genest et al. (2009) describe this procedure of goodness of fit testing. In the last paper is mentioned that this bootstrap version gives accurate outcomes.

1. Estimate the parameters θ of the copula function from the pseudo-sample: U_1, U_2, U_3.
2. Generate X_1^*, X_2^*, X_3^* from the copula function with the estimated parameter as described in Section 3.5. And compute their normalized rank vectors U_1^*, U_2^*, U_3^*
3. Estimate the parameters θ^* from U_1^*, U_2^*, U_3^* as above.
4. Compute the $\hat{S}^*$ with the normalized rank vectors U_1^*, U_2^*, U_3^* and estimated parameters θ^*. For this construction of $\hat{S}^*$ the PIT method of above is used.
5. Compute the AD* statistic of $\hat{S}^*$.
6. Repeat steps (2), (3), (4) and (5) K times, with W the number of bootstrap repetitions.

There are W AD* test statistics computed in this way. The critical value is determined as the (0.95)-quantile of the W AD* test statistics. In Grevelink et al. (2014) this method is also used. There it became clear that is can take much calculation time to compute this test. Therefore it is important to choose models that seem to be the most interesting, e.g. comparing three bivariate Student's t-copulas with three bivariate skewed t-copulas.

3.8 Value at risk

The value at risk is measured for a portfolio of equally weighted assets. The general formula is given in Equation (32). In this case the VaR is calculated by $\text{VaR}(1-q,1) = u + r_t + s_q \cdot \sigma_{t+1}$, where s_q is the q-th quantile of the simulated portfolio return. For a period of the last 500 days of the dataset there will be made one-day ahead forecasts with a few models that seem to be interesting. In the previous research Grevelink et al. (2014) it is shown that it takes lengthy calculations to get the VaR estimates. Therefore, the amount of estimates is not further increased.

$$\text{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} : F_X(x) > \alpha\} \tag{32}$$

The forecasts are based on an expanding window for the AR-GJR-GARCH model. Simulation of the η's is used to get the predicted standardized residuals, which will be used to get the one-day ahead forecasted log-returns. After calculating the value at risk values, we compare those with the actual returns. In this way the correct conditional coverage is computed, which is an evaluation of the VaR estimates.

4 Results

In this section the results are described of the applied methodology to the data from Section 2. There are some choices made which will be briefly explained, but are discussed more extensively later. First the models that are used to get the standardized residuals will be shown. Second, the dependent variable is chosen. The estimation of the parameters is the most complicated part and will get some more attention. Finally, the value at risk estimates are shown.

4.1 Residuals

As mentioned in Section 3.1, there should be no autocorrelation left in the residuals to use the pair-copula model. The models that do not contain any autocorrelation are stated in Table 3. The best fitted AR model at the AMX was an AR(4) model, which also was needed to reduce the autocorrelation. The autocorrelation of the first five lags of the models from Table 3 are shown in Table 10, which can be seen in appendix C. It can be seen that those lags are, according to the Q-statistic, insignificant. This implies that there is no autocorrelation left in the residuals. Therefore, the residuals can be used in the pair-copula model.

Sample	Model
AEX	AR(1)-GJR-GARCH(2,1)
AMX	AR(4)-GJR-GARCH(1,1)
Benchmark	AR(1)-GJR-GARH(1,1)

Table 3: Used model for the period 01/05/1989 - 01/05/2014.

Goodness of fit The models above are used with Student's t- and normal distributed errors. Subsequently those errors are tested to be Student's t and normal distributed respectively. This is done with the use of the Anderson-Darling(AD) and the Kolmogorov-Smirnov(KS) tests. In Table 4 the KS statistics of the Student's t-distributed errors are significantly fitted at a 5% significance level. The models with normal distributed errors are incorrect. The AD-test confirms these outcomes, except for the AMX. Nevertheless the model is used for the AMX in this research.

With the outcomes of the goodness of fit tests it is made clear that the models with Student's t-distributed errors are correctly chosen. This can also be seen in a QQ-plot where the quantiles of the residuals are plotted against a Student's t-distribution with estimated parameters, just like the model under the null hypothesis (Figure 8). They differ, but not enough to reject the null.

4.2 Dependent variable

Before the errors are used it is needed to find the dependent variable. In Figure 5 are the Spearman's rho correlations shown for the period 01/05/1989 - 01/05/2014. It can be seen that the dependence between the AEX and the other variables is the largest. The AEX is therefore the first variable to condition on.

Asset	Chosen error distribution	AD statistic	P-Value AD	KS statistic	P-value KS
AEX	Normal	97.5245	0.0005	0.1789	0.0000
	Student's t	1.6463	0.1451	0.0111	0.4234
AMX	Normal	87.3006	0.0005	0.1975	0.0000
	Student's t	3.5998	0.0137	0.0146	0.1386
Benchmark	Normal	23.0368	0.0005	0.3693	0.0000
	Student's t	1.2604	0.2456	0.0097	0.6030

Table 4: Goodness of fit outcomes of the Anderson-Darling test and the Kolmogorov-Smirnov test for the specified models.

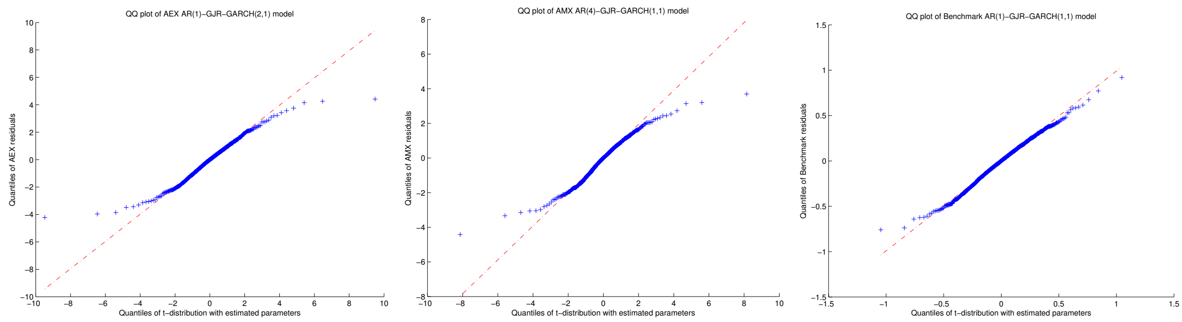

Figure 8: QQ plot: the residuals of the chosen AR-GJR-GARCH models with Student's t-distributed errors against a Student's t distribution.

	AEX	AMX	Benchmark
AEX	1	**0.7510**	**-0.1052**
AMX	**0.7510**	1	-0.1020
Benchmark	**-0.1052**	-0.1020	1

Table 5: Spearman's rho correlation of the residuals for the period 01/05/1989 - 01/05/2014.

	AEX	AMX	Benchmark
AEX	1	**0.5705**	**-0.0711**
AMX	**0.5705**	1	-0.0685
Benchmark	**-0.0711**	-0.0685	1

Table 6: Kendall's tau correlation of the residuals for the period 01/05/1989 - 01/05/2014.

We notice that the two measurements result to similar outcomes. Spearman's rho has more extreme outcomes as Kendall's tau. So there is a difference in calculation, but they lead to the same conclusions in this case.

4.3 Pair-copula model compared with the multivariate Student's t-copula

The pair-copula is constructed from multiple bivariate copulas. The choice of these bivariate copulas is very important for the outcomes of this pair-copula model. The scatter plots from Figure 9 gives an indication for the used bivariate copulas. The first scatterplot of the AEX against the AMX seems to have more events in the tails. Therefore, a Student's t-copula and possibly the skewed t-copula can perform well in this case. However, the other two scatter plots are not so obvious to link to a certain copula.

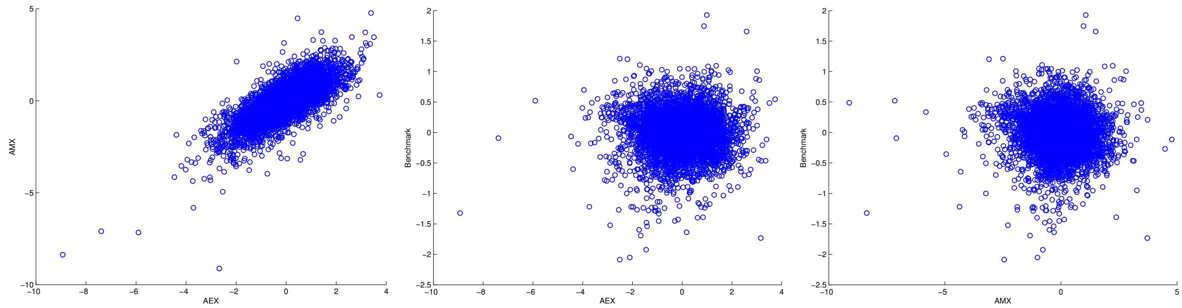

Figure 9: Scatterplot of the standardized residuals for the period 01/05/1989 - 01/05/2014.

In Table 7 are the results of the parameter estimations and also the goodness of fit test. Student's t stands for the multivariate Student's t copula model. S-S-S is the three-dimensional pair-copula model where for all pairs a bivariate Student's t-copula is chosen and the G means a bivariate Gaussian copula. 1 is the AEX, 2 the AMX and 3 is the Benchmark.

No model shows a significant fit according to the bootstrap version of the Anderson-Darling goodness of fit test. Because of those insignificant models, we could not give a clear interpretation to the estimated parameters. However, we try to give an explanation for the differences in estimated parameters. The parameters between the AEX and the AMX seem to be very similar in all different models. The ρ of about 0.76 means that the dependence between those assets is very large, what is obvious because they act in the same market. Between the AEX and the Benchmark this dependence is smaller in all cases, just like the dependence between the AMX and the Benchmark. We noticed that the dependence between the AEX and the Benchmark seemed to be better-described using thinner tailed copulas compared to the dependence between the AMX and the Benchmark. This is unexpected, because of the scatterplots in Figure 9 seem pretty similar. The Spearman's rho and Kendall's tau of those pairs did not differ so much, which could cause the difference in parameter estimations when conditioning on the AEX. Again, it must be said, that the models are not significantly fitted, so there is no clear interpretation.

4.4 Extended skewed t-copula

As can be seen in Table 7, the EST copula is not used. This is caused by some difficulties of the estimations for the pair-copula. An important disadvantage of this EST copula is that we are unaware of the inverse of this distribution. Therefore, it is tried to obtain the inverse by a numerical search, which is much slower than an exact formulation. This topic

Model	Parameter	Estimates	Log-likelihood	AD statistic	Critical Value	P-value
G-G-G	ρ_{12}	0.756	2635.877	23.873	0.838	0.000
	ρ_{13}	-0.032				
	$\rho_{23\|1}$	-0.051				
S-S-S	ρ_{12}	0.761	2787.841	2.998	0.760	0.000
	ρ_{13}	-0.050				
	$\rho_{23\|1}$	-0.047				
	ν_{12}	8.686				
	ν_{13}	5.445				
	$\nu_{23\|1}$	26.458				
G-G-S	ρ_{12}	0.788	3081.756	25.370	0.8788	0.000
	ρ_{13}	-0.151				
	$\rho_{23\|1}$	-0.023				
	$\nu_{23\|1}$	45.1596				
Student's t	ρ_{12}	0.758	2764.804	2.618	0.816	0.000
	ρ_{13}	-0.044				
	ρ_{23}	-0.065				
	ν	8.361				

Table 7: Estimated parameters, test statistics and p-values of the Anderson Darling goodness-of-fit test statistic for the three-dimensional pair-copula decomposition for the period 01/05/1989 - 01/05/2014.

is extensively discussed in Section 5. In Table 8 we can see the difference in computation times of Figure 7 in Section 3.3.1. It is obvious that the computation of the EST copula takes extraordinarily more time. This could also happen in the case of the estimations.

Model	Computation time (in seconds)
Student's t-copula, $\rho = 0.8$, $\nu = 3$	3.2
Student's t-copula, $\rho = 0.5$, $\nu = 20$	2.8
Skewed t-copula, $\rho = 0.5$, $\nu = 20$, $\tau = 0$, $\lambda = 0$	359.5
Skewed t-copula, $\rho = 0.5$, $\nu = 20$, $\tau = 0$, $\lambda = -10$	370.5
Skewed t-copula, $\rho = 0.8$, $\nu = 3$, $\tau = 0$, $\lambda = 3$	386.0
Skewed t-copula, $\rho = 0.8$, $\nu = 3$, $\tau = 0$, $\lambda = 10$	388.5

Table 8: Computation times of simulations in Figure 7

Besides the numerical method that is used, also the amount of parameters is a big disadvantage of the EST copula. In the case of three bivariate EST copulas there are nine parameters that have to be estimated, what takes much longer time then six parameters of the three-dimensional S-S-S pair-copula. The pair-copula with bivariate skewed t-copulas is therefore even slower. However, we tried to get estimates using the method described in Section 3.5. Unfortunately, this method, with a fixed $\tau = 0$ did not solve the problem. In the case of only one bivariate EST copula with two bivariate Gaussian copulas, there are only five parameters that have to be estimated. Nevertheless is the computation time much too high.

The maximum likelihood estimates were not found, what could be caused by different problems. In Appendix A it can be seen that the formulas of the bivariate Gaussian- and Student's t-copula are a lot simpler than that of the bivariate EST copula. Not only the number of parameters, but also the efficiency of programming just like the difficulty of the formulas can cause for longer computation times. However, it is obvious that the construction of a new bivariate copula is not an easy problem. It is therefore clear that the pair-copula model is not so easy to use and can get very long computation times, especially in the absence of an exact formulation of the distribution and also in cases of multiple parameters. For this reason it is of great importance to use the most efficient programming code but definitely also to use sufficiently fast computers.

4.5 Value at risk

Value at risk estimates are computed using the S-S-S and G-G-G pair-copula models. Those are compared with the VaR estimates of the multivariate Student's t-copula. Every next day is forecasted in the period 11/05/12 until 01/05/14 using the η's from the copula models. These η's are simulated 500 times. Thereafter the conditional variances are computed using Equation (3). The portfolio return is the mean of those assets. The value at risk is the $(1-q)$-th quatile of the portfolio returns. Figure 10 shows the VaR outcomes for $q = 0.10$, $q = 0.05$ and $q = 0.01$, respectively.

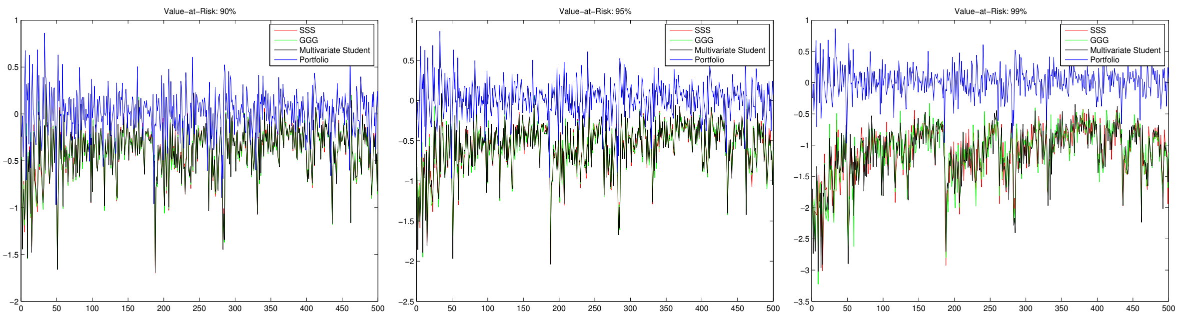

Figure 10: Value at risk of the S-S-S, G-G-G and the multivariate Student's t-copula for the period 11/05/12 until 01/05/14.

There can not be seen major difference in the figures above. However, the conditional coverage test of Table 9 can be interpreted more easily. If the p-values are below the q-values, the null hypothesis is rejected. There are two measures that are rejected, namely both the 95% VaR of the S-S-S and the multivariate Student's t-copula models. It is obvious that the G-G-G pair-copula model performs better, especially because of our hypotheses that were focused on better performing Student's t-copulas.

5 Discussions

This research started with two hypotheses. In the first hypothesis we state that the pair-copula model using the bivariate EST copula has a better fit according to the bootstrap version of the AD test compared to the multivariate Student's t-copula model. The second hypothesis focuses on the idea that the pair-copula model using the bivariate EST copula generates better outcomes of the value at risk compared to the multivariate Student's t-copula model. The main problem in this research is that although we were

q	Model	Likelihood Ratio$_{cc}$	P-value
0.01	S-S-S	4.814	0.090
	G-G-G	4.814	0.090
	Student's t	2.353	0.308
0.05	S-S-S	6.018	0.049
	G-G-G	4.8843	0.087
	Student's t	8.7373	0.013
0.10	S-S-S	2.121	0.346
	G-G-G	1.087	0.581
	Student's t	1.308	0.520

Table 9: Correct Conditional Coverage test for the VaR-breaks.

able to implement the bivariate EST copula into the pair-copula model, we did not find estimates for this pair-copula model.

There can be three different reasons for this problem. The first possibility is that there could be an error in the programming of this EST copula. The second reason for this problem could be caused by the numerical method used to compute the inverse of the EST distribution. In this case, numerical methods are much slower than exact formulations, as they are an approximation of an exact formulation. Explicit availability of the h-functions and their inverse is very important for the efficiency of the estimation procedures (Aas et al. (2009)). The third, and final possibility is that the computational power of the computer used is not sufficient for such a complex model. We suspect that the second possibility is the most likely as the function 'fzero' of Matlab is known that in some cases can be slow to use. This could be attributed to the fact that the function 'fzero', which we use to compute the maximum likelihood estimates, is not capable of calculating with matrices as it computes each scalar separately. If this is the case, there are two possible solutions: the first one is to find a faster function to calculate the inverse numerically. Secondly there is the option that we try to find another method to compute the inverse. Ideally, the exact formulation for the inverse of the EST distribution is found.

The research shows some interesting findings with respect to the AR-GJR-GARCH model. The model with Student's t-distributed errors gives a better fit for the assets than the model with normal distributed errors. This is of major importance for the calculation of reliable results from the copula models. Especially when computing the value at risk the fit of the AR-GJR-GARCH model is important. This verifies the claim that market returns are often better described by fatter tailed distributions (Alexander (2008)).

Another point we want to emphasize is related to the practical applicability of the bootstrap version of the AD test. We already know that the computation time of this test is significantly longer compared to the maximum likelihood estimation (Grevelink et al. (2014)). Therefore it is likely that even if we had found a result from this EST copula, we would not have been able to compute this bootstrap due to time limitations. For example, it takes about nine minutes to estimate the pair-copula model with three bivariate Student's t-copulas. This results in a calculation time of almost 20 hours to calculate the critical value of the AD test. Due to time constraints we did only a 1000 simulations of this bootstrap. However, increasing the number of simulations would improve the accuracy of the critical value. The most obvious response to this is to increase the number of simulations. Yet it might be more valuable to focus on finding different goodness of fit tests for pair-copula models as the problem of the long calculation time remains.

If we look more closely to the results we find no significantly fitted model. This implies that the pair-copula with three bivariate Student's t-copulas is not always the best-fitted model. For further research we suggest to use data with more joint extreme events. This increases the possibility that a Student's t-copula will fit the data, as this copula captures more of this dependence. In line with this reasoning we also expect that the EST-copula to better fit this kind of data. This increases the accuracy of a comparison between both copulas.

The final point we want to make is related to the topic of value at risk. We tested the copula models for value at risk, although the AD test turned out non significant, we compared the different models. The results of the comparison show that the Student's t-copulas did not perform best in all three cases. In the case of 95%-VaR the pair-copula with three bivariate Gaussian copulas was the only correct model to use, according to the conditional coverage test. Furthermore, while estimating the value at risk we encountered the same problems with the calculation time as in the bootstrap. The lengthy calculation time is caused by the fact the parameters have to be calculated for each forecasted day individually. Regarding the number of days forecasted the same reasoning as in the bootstrap could be applied, as an increase in the number of days that are forecasted improves the reliability of the value at risk outcomes.

6 Conclusions

This research started from the question whether the EST copula is a better model compared to the Student's t-copula. We started with a discussion of the AR-GJR-GARCH model. This model is used to compute the standardized residuals. It appears that this model gives a better fit for the assets when the errors are Student's t-distributed in comparison to normal distributed errors. We were not able to test the hypotheses with regard to a better fit according to the AD test of the pair-copula model using the bivariate EST copula compared to the multivariate Student's t-copula. Furthermore, we could not test whether a pair-copula model using the bivariate EST copula has better outcomes for value at risk than a multivariate Student's t-copula. However, there were some other interesting results. First it appears that none of the models has a significant fit. From this we conclude that in order to be able to make a better comparison between the Student's t-copulas, more attention should be paid to the selection of data. The second point is related to the applicability of the EST copula, as it appears that the calculation time is very long as a result of the numerical method used. Therefore, it is important to keep in mind the resources available when you choose between different copulas. The last point is about the outcomes of the VaR estimates. In the case of the 95%-VaR are both Student's t-copula models rejected in contrast to the G-G-G pair-copula model, while we expected that the Student's t-copula models would perform best, according to our hypotheses. The reliability of those outcomes can be increased using more days to forecast.

A Copulas

A.0.1 Gaussian copula

The bivariate copula density of the Gaussian copula is (Aas et al. (2009)):

$$c(u_1, u_2) = \frac{1}{\sqrt{1-\rho_{12}^2}} \cdot \exp\left\{-\frac{\rho_{12}^2(x_1^2+x_2^2) - 2\rho_{12}x_1x_2}{2(1-\rho_{12}^2)}\right\} \tag{33}$$

The h-function, in Equation (34), represents the conditional distribution function, which is used by estimating the parameters. This is discussed in Section 3.5.

$$h(u_1, u_2, \rho_{12}) = \Phi\left(\frac{\Phi^{-1}(u_1) - \rho_{12}\Phi^{-1}(u_2)}{\sqrt{1-\rho_1 2^2}}\right) \tag{34}$$

$$h_{12}^{-1}(u_1, u_2, \rho_{12}) = \Phi\left\{\Phi^{-1}(u_1)\sqrt{1-\rho_{12}^2} + \rho_{12}\Phi^{-1}(u_2)\right\} \tag{35}$$

A.0.2 Student's t-copula

The density of this copula is:

$$c(u_1, u_2) = \\ = \frac{\Gamma(\frac{\nu_{12}+2}{2})/\Gamma(\frac{\nu_{12}}{2})}{\nu_{12}\pi \mathrm{d}t(x_1, \nu_{12})\mathrm{d}t(x_2, \nu_{12})\sqrt{1-\gamma_{12}^2}} \cdot \left\{1 + \frac{x_1^2 + x_2^2 - 2\gamma_{12}x_1x_2}{\nu_{12}(1-\gamma_{12}^2)}\right\}^{-\frac{\nu_{12}+1}{2}} \tag{36}$$

$$h(u_1, u_2, \gamma_{12}, v_{12}) = t_{v_{12}+1}\left\{\frac{t_{v_{12}}^{-1}(u_1) - \gamma_{12}t_{v_{12}}^{-1}(u_2)}{\sqrt{\frac{\left(v_{12}+\left(t_{v_{12}}^{-1}(u_2)\right)^2\right)(1-\gamma_{12}^2)}{v_{12}+1}}}\right\} \tag{37}$$

$$h_{12}^{-1}(u_1, u_2, \gamma_{12}, v_{12}) = t_{v_{12}} \cdot \left\{t_{v_{12}+1}^{-1}(u_1)\sqrt{\frac{\left(v_{12}+\left(t_{v_{12}}^{-1}(u_2)\right)^2\right)(1-\gamma_{12}^2)}{v_{12}+1}} + \gamma_{12}t_{v_{12}}^{-1}(u_2)\right\} \tag{38}$$

B Skewed t-distribution

$Y \sim EST_p(\xi, \Omega, \lambda, \nu, \tau)$ and $y \in \mathbb{R}^p$ if the density function is:

$$\frac{1}{T_1(\tau/\sqrt{1+\lambda^T\bar{\Omega}\lambda}; \nu}t_p(y;\xi,\Omega,\nu)T_1\left\{(\lambda^T z + \tau)(\frac{\nu+p}{\nu+Q(z)})^{1/2}; \nu+p\right\} \tag{39}$$

with

$$t_p(y;\xi,\Omega,\nu) = \frac{\Gamma((\nu+p)/2)}{|\Omega|^{1/2}(\nu\pi)^{p/2}\Gamma(\nu/2)}\left(1+\frac{Q(z)}{\nu}\right)^{-(\nu+p)/2} \tag{40}$$

which denotes the density function of the usual p-dimensional Student's t-distribution. $z = \omega^{-1}(y-\xi)$, $Q(z) = z^T\bar{\Omega}^{-1}z$, $\lambda \in \mathbb{R}^p$ is the shape parameter and $\tau \in \mathbb{R}$ is the extension parameter and $\xi \in \mathbb{R}^p$ is the location parameter. Then there are the positive definite $p \times p$ dispersion matrix Ω, $p \times p$ scale matrix $\omega = diag(\Omega)^{1/2}$ and $p \times p$ correlation matrix $R = \bar{\Omega} = \omega^{-1}\Omega\omega^{-1}$. Degrees of freedom $\nu > 0$ and $T_1(x;\nu)$ denotes the univariate standard Student's t cumulative distribution function with $\nu > 0$ degrees of freedom.
The cumulative distribution function is as follows:

$$P(Y \leq y) = \frac{1}{T_1(\bar{\tau};\nu)} T_{p+1}\left(\begin{pmatrix} z \\ \bar{\tau} \end{pmatrix}; \begin{pmatrix} \bar{\Omega} & -\delta \\ -\delta^T & 1 \end{pmatrix}, \nu+1\right) \tag{41}$$

where $\bar{\tau} = (1+\lambda^T\bar{\Omega}\lambda)^{-1/2} \cdot \tau$ and $\delta = \frac{\bar{\Omega}\lambda}{\sqrt{1+\lambda^T\bar{\Omega}\lambda}}$ and $T_r(x;\Omega,\nu)$ is the multivariate r-dimensional centered Student's t cumulative distribution function, with $r \times r$ positive definite dispersion matrix Ω, ν degrees of freedom, ,

Before the skewed t-distribution can be used in a pair-copula construction, the h-function has to be made. This is done with the use of the conditional distribution. Below is the conditional distribution described. This is used to construct the h-function.

$$(X_1|X_2 = x_2) \sim EST_{p_1}(\xi_{1\cdot 2}, \alpha_{Q_2}\Omega_{11\cdot 2}, \lambda_{1\cdot 2}, \nu_{1\cdot 2}, \tau^*_{1\cdot 2}) \tag{42}$$

where $dim(X_1) = p_1$, $\xi_{1\cdot 2} = \xi_1 + \Omega_{12}\Omega_{22}^{-1}(x_2 - \xi_2)$, $\alpha_{Q_2} = (\nu + Q_2(z_2))/(\nu + p_2)$, $Q_2(z_2) = z_2^T\bar{\Omega}_{22}^{-1}z_2$, $z_2 = \omega_2^{-1}(x_2 - \xi_2)$, $\Omega_{11\cdot 2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$, $\lambda_{1\cdot 2} = \omega_{1\cdot 2}\omega_1^{-1}\lambda_1$, $\omega_{1\cdot 2} = diag(\Omega_{11cdot2})^{1/2}$, $\omega_1 = diag(\Omega_{11})^{1/2}$, $\nu_{1\cdot 2} = \nu + p_2$, $\tau^*_{1\cdot 2} = \alpha_{Q_2}^{-1/2}\tau_{1\cdot 2}$, $\tau_{1\cdot 2} = (\lambda_1^T\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \lambda_2^T)z_2 + \tau$.
In Aas et al. (2009) the standard univariate distributions were used to construct the h-functions. The standard conditional extended skewed t-distribution has the following properties:

$$\Omega = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix} \tag{43}$$

Here are some derivations of the conditional properties that are used to construct the h-function.

$$E(X_1|X_2 = x_2) = \xi_{1\cdot 2} = \xi_1 + \Omega_{12}\Omega_{22}^{-1}(x_2 - \xi_2) = \Omega_{12}\Omega_{22}^{-1}x_2 = \rho_{12}x_2 \tag{44}$$

$$\begin{aligned} \Omega_{X_1|X_2=x_2} &= \alpha_{Q_2}\Omega_{11\cdot 1} \\ &= \frac{\nu + Q_2(z_2)}{\nu + p_2}\left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right) \\ &= \frac{\nu + z_2^T\bar{\Omega}_{22}^{-1}z_2}{\nu + p_2}\left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right) \\ &= \frac{\nu + (\omega_2^{-1}(x_2 - \xi_2))^T\bar{\Omega}_{22}^{-1}(\omega_2^{-1}(x_2 - \xi_2))}{\nu + p_2}\left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right) \\ &= \frac{\nu + (\omega_2^{-1}(x_2 - \xi_2))^T\bar{\Omega}_{22}^{-1}(\omega_2^{-1}(x_2 - \xi_2))}{\nu + p_2}\left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right) \\ &= \frac{\nu + (\omega_2^{-1}(x_2 - 0))^T R_{22}^{-1}(\omega_2^{-1}(x_2 - 0))}{\nu + p_2}\left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right) \\ &= \frac{\nu + x_2^T\Omega_{22}^{-1}x_2}{\nu + p_2}\left(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}\right) \\ &= \frac{\nu + x_2^T x_2}{\nu + p_2}(1 - \rho_{12}^2) \end{aligned} \tag{45}$$

$$\nu_{X_1|X_2=x_2} = \nu_{1\cdot 2} = \nu + p_2 = \nu + 1 \tag{46}$$

These three conditional properties are exactly the same as that of a conditional Student's t-distribution. The h-functions will therefore be similar, except for the extra parameters.

$$\begin{aligned}
&\tau_{X_1|X_2=x_2} = \tau^*_{1\cdot 2} = \alpha_{Q_2}^{-1/2}\tau_{1\cdot 2} \\
&= \left\{\frac{\nu + Q_2(z_2)}{\nu + p_2}\right\}^{-1/2} \cdot \big((\lambda_1^T \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \lambda_2^T)z_2 + \tau\big) \\
&= \left\{\frac{\nu + z_2^T \bar{\Omega}_{22}^{-1} z_2}{\nu + p_2}\right\}^{-1/2} \cdot \big((\lambda_1^T \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \lambda_2^T)\omega_2^{-1}(x_2 - \xi_2) + \tau\big) \\
&= \left\{\frac{\nu + (\omega_2^{-1}(x_2 - \xi_2))^T \bar{\Omega}_{22}^{-1}(\omega_2^{-1}(x_2 - \xi_2))}{\nu + p_2}\right\}^{-1/2} \cdot \big((\lambda_1^T \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \lambda_2^T)\omega_2^{-1}(x_2 - \xi_2) + \tau\big) \\
&= \left\{\frac{\nu + (\omega_2^{-1}(x_2 - 0))^T R_{22}^{-1}(\omega_2^{-1}(x_2 - 0))}{\nu + p_2}\right\}^{-1/2} \cdot \big((\lambda_1^T R_{12} R_{22}^{-1} + \lambda_2^T)\omega_2^{-1}(x_2 - 0) + \tau\big) \\
&= \left\{\frac{\nu + (\omega_2^{-1}x_2)^T(\omega_2^{-1}x_2)}{\nu + p_2}\right\}^{-1/2} \cdot \big((\lambda_1^T R_{12} + \lambda_2^T)\omega_2^{-1}x_2 + \tau\big) \\
&= \left\{\frac{\nu + x_2^T x_2}{\nu + p_2}\right\}^{-1/2} \cdot \big((\lambda_1^T \rho_{12} + \lambda_2^T)x_2 + \tau\big)
\end{aligned} \tag{47}$$

$$\begin{aligned}
\lambda_{1\cdot 2} &= \omega_{1\cdot 2}\omega_1^{-1}\lambda_1 \\
&= diag(\Omega_{11\cdot 2})^{1/2} \cdot (diag(\Omega_{11})^{1/2})^{-1} \cdot \lambda_1 \\
&= diag(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{1/2} \cdot diag(\Omega_{11})^{-1/2} \cdot \lambda_1 \\
&= (1 - \rho_{12}^2)^{1/2} \cdot \lambda_1
\end{aligned} \tag{48}$$

C Autocorrelation

Below are the outcomes of the autocorrelations of the used AR-GJR-GARCH models.

	Standardized residuals				Squared standardized residuals			
Asset	AC	PAC	Q-Stat	P-value	AC	PAC	Q-Stat	P-value
	0.009	0.009	0.557	0.455	0.002	0.002	0.030	0.862
	0.005	0.005	0.721	0.697	0.006	0.006	0.238	0.888
AEX	-0.015	-0.016	2.197	0.532	0.007	0.007	0.571	0.903
	0.026	0.027	6.520	0.164	-0.006	-0.006	0.773	0.942
	-0.007	-0.008	6.854	0.232	-0.012	-0.012	1.655	0.894
	0.005	0.005	0.1833	0.669	-0.004	-0.004	0.079	0.779
	0.02	0.02	2.5994	0.273	0.01	0.01	0.6559	0.720
AMX	0.004	0.004	2.7136	0.438	-0.013	-0.013	1.6967	0.638
	0.011	0.01	3.4493	0.486	-0.009	-0.009	2.187	0.701
	-0.004	-0.004	3.5519	0.616	-0.004	-0.004	2.2729	0.810
	0.013	0.013	1.0749	0.300	0.015	0.015	1.346	0.246
	0.006	0.006	1.330	0.514	0.022	0.021	4.238	0.120
Benchmark	-0.005	-0.005	1.478	0.687	0.005	0.005	4.401	0.221
	0.026	0.026	5.630	0.229	0.005	0.004	4.547	0.337
	-0.012	-0.012	6.478	0.262	-0.005	-0.006	4.722	0.451

Table 10: The Ljung and Box statistic for the period 01/05/1989 - 01/05/2014 of the specified models with Student's t-distributed errors.

References

K. Aas. Modelling the dependence structure of financial assets: A survey of four copulas. *Working paper*, 2004.

K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44:182–198, 2009.

Carol Alexander, editor. *Market Risk Analysis. Practical Financial Econometrics*, volume II. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2008.

T.W. Anderson and D.A. Darling. A test of goodness of fit. *Journal of the American Statistical Association*, 49:765–769, 1954.

Reinaldo B Arellano-Valle and Marc G Genton. Multivariate extended skew-t distributions and related families. *Metron*, 68(3):201–234, 2010.

Adelchi Azzalini and Antonella Capitanio. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):367–389, 2003.

Stefano Demarta and Alexander J McNeil. The t copula and related copulas. *International statistical review*, 73(1):111–129, 2005.

J. Dobrić and F. Schmid. A goodness of fit test for copulas based on rosenblatts transformation. *Computational Statistics & Data Analysis*, 51:4633 4642, 2007.

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

C. Genest and A.-C Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of hydrologic engineering*, 12:347–368, 2007.

Christian Genest, Bruno Rémillard, and David Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, 44(2): 199–213, 2009.

T. Grevelink, S. Hui, X. Ji, and S. Li. Comparing pair-copula models with the multivariate gaussian copula model for financial assets. *Working paper*, 2014.

Bruce E Hansen. Autoregressive conditional density estimation. *International Economic Review*, pages 705–730, 1994.

D.X. Li. On default correlation: A copula function approach. *The Journal of Fixed Income*, 9:4354, 2000.

Greta M Ljung and George EP Box. The likelihood function of stationary autoregressive-moving average models. *Biometrika*, 66(2):265–270, 1979.

Frank J Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951.

M. Rosenblatt. Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 27:832–837, 1952.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statique de L'Iniversité de Paris*, 8:229–231, 1959.