

Master thesis

# Pension funds, prepare for crisis!

The Asset Liability Management Including Crisis for Asset Returns (ALMICAR) model

Marco van der Lans  
304432

January 17, 2011

Supervisors: Wing Wah Tham (EUR) & Edwin Massie (AEGON)  
Co-reader: Erik Kole (EUR)



## **Abstract**

The 2008 financial crisis indicated the limitations of the Asset Liability Management models used to advise pension funds. The huge losses on the invested portfolios of pension funds were so unlikely according to the ALM model that they were deemed impossible. Combined with a very low interest rate huge drops in the coverage ratio of pension funds were realized, resulting in public disturbance. In this research a new method is proposed which includes a future crisis period within the investments. In this research the Asset Liability Management Including Crisis for Assets Returns (ALMICAR) model is proposed, which incorporates crises periods within the existing Asset Liability Management (ALM) framework. In this way the ALMICAR model allows to incorporate a crisis directly into the decision making process for future investments of pension funds instead of using separate stress scenarios to evaluate the impact of a crisis. A comparison between the ALMICAR and AEGON's ALM model is made. The research shows that the simulations for future returns made with the ALMICAR model are more realistic than the simulations from AEGON's existing ALM model. As a result, pension funds are able to make more informed decisions on their investments.

Key words: Asset Liability Management (ALM) model, crisis, skewed student-t distribution, copula, correlation measures, Asset Liability Management Including Crisis for Assets Returns (ALMICAR) model.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Data descriptives</b>	<b>7</b>
<b>3</b>	<b>Marginal distributions</b>	<b>12</b>
3.1	Introduction marginal distributions . . . . .	13
3.1.1	Normal distribution . . . . .	13
3.1.2	Student-t distribution . . . . .	13
3.1.3	Skewed normal distribution . . . . .	13
3.1.4	Skewed student-t distribution . . . . .	14
3.2	Tests for empirical fit of marginal distributions . . . . .	14
3.2.1	Kolmogorov-Smirnov . . . . .	14
3.2.2	Anderson-Darling . . . . .	15
3.3	Results . . . . .	15
3.4	Conclusion . . . . .	18
<b>4</b>	<b>Correlation measures</b>	<b>19</b>
4.1	Correlation measures . . . . .	19
4.1.1	Pearson product-moment correlation . . . . .	19
4.1.2	Kendall's tau . . . . .	19
4.1.3	Spearman's rho . . . . .	20
4.2	Simulation and tests . . . . .	21
4.2.1	Simulation of data sets . . . . .	21
4.2.2	Distance measures . . . . .	21
4.3	Results . . . . .	22
4.3.1	Estimation results . . . . .	22
4.4	Conclusion . . . . .	24
<b>5</b>	<b>Copula</b>	<b>25</b>
5.1	Different copulas . . . . .	25
5.1.1	Gaussian copula . . . . .	25
5.1.2	Student-t copula . . . . .	26
5.1.3	Gumbell copula . . . . .	26
5.1.4	Clayton copula . . . . .	26
5.1.5	Other copulae . . . . .	26
5.2	Student-t copula . . . . .	27
5.2.1	Estimation procedure . . . . .	27
5.2.2	Simulating returns from the student-t copula . . . . .	27
5.3	Tests . . . . .	28
5.3.1	Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) . . . . .	29
5.3.2	Wilcoxon signed-rank test . . . . .	29
5.3.3	One-sided t-test . . . . .	29
5.3.4	Historical simulation . . . . .	29
5.4	Results . . . . .	30
5.4.1	Asset class simulations . . . . .	30
5.4.2	Portfolio simulations . . . . .	35
5.5	Conclusion . . . . .	39
<b>6</b>	<b>Asset liability framework</b>	<b>40</b>
6.1	Asset liability management (ALM) model . . . . .	40
6.2	Value-at-Risk (VaR) and Expected Tail Loss (ETL) . . . . .	40
6.3	Tests . . . . .	41
6.3.1	Stylized portfolio . . . . .	41
6.4	Results . . . . .	42
6.4.1	Differences in outcomes of the ALM model . . . . .	42
6.4.2	Differences in outcomes of the ALM model based on VaR and ETL . . . . .	45

6.4.3	Results Wilcoxon signed-rank test . . . . .	46
6.4.4	Stylized portfolio results . . . . .	47
6.5	Conclusion . . . . .	47
<b>7</b>	<b>Conclusion</b>	<b>49</b>
<b>8</b>	<b>Appendix</b>	<b>50</b>

# 1 Introduction

The consequences of the 2008 financial crisis are still visible today. The global financial markets still struggle to find their way back to old levels and in the Netherlands the discussion on the sustainability of pensions and pension fund policies is still a hot topic. The financial crisis illustrated that many Dutch pension funds made their investments based on very positive prospects. As a consequence the investment strategy led to big losses on the investments during the 2008 financial crisis. In combination with a very low interest rate yield a huge decline in coverage ratio was observed for almost all pension funds as mentioned in Redactie [2008], ANP [2008] and ANP [2009a].

The term coverage ratio is very important for pension funds since the coverage ratio indicates whether or not a pension fund is able to pay the future pensions of the participants. The huge decline in coverage ratio is best explained by the computation of the coverage ratio, which is shown in equation 1.1:

$$\text{Coverage ratio} = \frac{\text{value of investments}}{\text{present value of future liabilities}} \quad (1.1)$$

Since the crisis led to a decline in the worth of the investments and to a "flight to quality", whereby especially the long term interest rates declined, the coverage ratio of pension funds dropped extra hard. Not only the value of the nominator decreased but also the value of the denominator increased leading to large drops in the coverage ratio of most pension funds. The big drops in coverage ratio of most pension funds led to much media attention as shown by Redactie [2008], ANP [2008] and ANP [2009a]. As a consequence the investments of pension funds became a hot topic for the Dutch public since they faced a potential decrease of their future pension payments.

The investments of a pension fund are based on a (very) long horizon since the liabilities of the pension funds are also far into the future. Many pension funds made decisions based on an ALM model that uses a multivariate normal distribution to simulate future returns. Even though this simulation method is easy to use it does not account for crisis periods enough while simulating future asset returns. The returns observed during the 2008 financial crisis were so unlikely according to the ALM model that they were deemed impossible. This is one of the reasons why ALM models have been questioned since, see also Van den Goorbergh et al. [2009].

As a result of the crisis pension funds need to be more aware of the risks of their investments. To get insight into the risks of their investments pension funds use stress scenarios in order to check what would happen if a certain scenario would occur. Breuer et al. [2009] propose to search for the worst possible stress scenario based on a set of plausible stress scenarios. Even though a set of plausible stress scenarios is selected with their method, it is still a problem with these separate stress scenarios that the probability for the stress scenario to actually occur is undefined. Further, the stress scenarios are used separately from the simulation of future returns and therefore are not directly taken into account for the asset-mix advice based on the ALM model. This results in the research question: How to implement stress scenario's within the current ALM framework in order to take crisis observations directly into account for the advisory on the asset-mix of a pension fund?

Earlier research on this topic is done by Van den Goorbergh et al. [2009]. They propose a new ALM framework in which a regime-switching model is used to account for a crisis period. A drawback of this method is that this model requires a lot of different parameters to be estimated. Given that the available data sample for Euro denominated asset classes is very small, the uncertainty on the parameter estimates becomes very high, which can lead to unreliable simulations of future asset returns. Next to that Van den Goorbergh et al. [2009] propose a whole new ALM framework, which makes it impossible to integrate this new method within AEGON's ALM model.

AEGON's ALM model simulates future interest rate yields and future inflation apart from the future asset returns. The simulation of future interest rate yields and future inflation within the AEGON ALM model has been implemented not long ago based on Vlaar [2006]. However, the simulation of future asset returns are currently still based on a multivariate normal distribution. Therefore the focus is on the implementation of stress scenarios in the simulation of future asset returns.

Since pension funds are more interested in portfolio returns than in the returns on specific asset classes, it is important to account for the dependency in returns between the different asset classes. Within the current ALM model linear correlation is used as a dependency measure. However, Lindskog [2000] shows that this measure is very sensitive to outliers in the data and therefore might be inappropriate for this research. According to Lindskog [2000] the linear correlation measure especially performs badly when a crisis period is present in the data sample. Based on Lindskog [2000] a comparison is made between three different correlation measures in order to test which measure is best to use for this research.

Besides the dependency between the asset classes a common problem is to form a joint distribution if different distributions are used as the marginal distribution for the asset classes. If all the asset classes are modeled by

the same distribution it is possible to form a joint distribution by using a multivariate distribution. However, within AEGON's ALM model there are up to ten different asset classes used which makes the assumption of using the same distribution for all the asset classes very restrictive. Therefore this research proposes to use a copula, which enables a joint distribution to be formed when the distributions used as marginal distribution for the asset classes are not the same. Further the use of a copula allows the modelling of each asset class separately based on the available data for a specific asset class to ensure a good fit.

The contribution of this research is threefold. At first, this is to best knowledge the first time that a copula is used within an ALM framework in order to take crisis periods directly into account in the advisory based on the ALM model. Further the finding of Lindskog [2000] that linear correlation should not be used as a dependency measure when extreme observations are present in the data sample is supported by the results shown in this research. At last this thesis is to my best knowledge one of the first researches that investigates the impact of using a different simulation method for future asset returns on the asset-mix advisory from an ALM model.

The remainder of this research is organized as follows: in section 2 the used data set and main data features are discussed. in section 3, based on the observed data features modeling techniques are selected, estimated and tested to find the model that best represents the empirical data of each asset class separately. Since pension funds are interested in returns on their portfolios and not in the returns on a specific asset class, dependencies between the asset classes are an important feature. This is discussed in section 4. Section 5 discusses how, based on correlation and the optimal distribution for the separate asset classes, a joint distribution is formed in order to account for crisis periods. Further, the simulations of future returns with the new and currently used method are compared with the empirical returns to test whether the ALMICAR method significantly outperforms the current ALM method. Section 6 discusses the implications of using the ALMICAR model on AEGON's advice based on the ALM framework. In section 7 the main conclusions of the research together with recommendations for further research are presented.

## 2 Data descriptives

One of the purposes of this research is to simulate realistic future returns. For this purpose it is important to analyze the data in order to find the most important data features. Based on this data analysis decisions are made on what kind of methods should be used to describe or represent data in the best possible way.

Simulated future returns are an important part of an Asset Liability Management (ALM) model. Therefore the chosen data set should be representative for the users of ALM models. Since the users of AEGON's ALM model are Dutch pension funds the focus is on European markets. Within the AEGON ALM model up to ten different asset classes are included in the analysis. Including all asset classes in the analysis would be perfect, however the history of asset classes is relative short and using constructed index values can lead to unreliable estimations. Therefore the focus in this research is on five of the most used asset classes in pension funds portfolios, namely: stocks, sovereign bonds, commodities, real estate and corporate bonds (credits). All indices are denominated in Euro, because Dutch pension funds have their liabilities in Euros and therefore only invest in European indices. This is because investing in a different currency than Euros would yield an extra currency risk for pension funds, which is undesirable. The specific indices for the different asset classes that are taken into account in this research are shown in table 2.1.

Asset class	Used Index	Data type	Daily data available
Sovereign bonds	Bank of America Merrill Lynch European Monetary Union direct government AAA index (BOFA ML EMU TR)	Total returns	yes
Corporate bonds (Credits)	Barclays Euro-Aggregate Corporate index (BC EACI)	Excess returns	no
Stocks	Financial Times Stock Exchange All-World Index (FTSE AWI TR)	Total returns	yes
Real Estate	Morgan Stanley Capital International All Country World Index (MSCI ACWI RE TR)	Total returns	yes
Commodities	Standard & Poor's Goldman Sachs Composite Index for Commodities (S&P GSCI Commodities)	Total returns	yes

Table 2.1: This table shows the used indices and data types for the different asset classes. Further the type of data is mentioned and whether or not daily data was available. All used data is collected through DataStream. For corporate bonds (credits) the excess returns are used, because the return on interest rate is not taken into account in this way. This is done since the return on interest rates is simulated with a different part of the ALM model.

As table 2.1 shows, there are only indices used and no specific companies. This is because the ALM tool is used to advise pension funds on investments in asset classes and not in specific companies. Further excess returns are used for corporate bonds (credits) in order to ensure that the returns on interest rates are not taken into account. Interest rates are simulated separately in the AEGON ALM model and therefore their returns should not be taken into account in the simulation of asset class returns. The used data is retrieved from DataStream over a period such that the returns range from 01-01-1999 till 1-4-2010. This results in a data set consisting of 136 monthly returns and 2934 daily returns. Daily data for the excess returns on credits were unavailable, which results in a daily data sample consisting of four asset classes.

The reason to consider both daily and monthly data is that the portfolio returns are forecast over a large time period ranging from one to more than twenty years. Given this long investment horizon it would be optimal to use yearly or quarterly data. However, this would result in a data sample of respectively 10 and 41 observations, which is not large enough for reliable parameter estimations. Given the long investment horizon of pension funds the simulation of future returns is often based on a quarterly or yearly basis. Since this leads to unreliable parameter estimations it would be perfect to use daily or monthly returns. However, although using daily or monthly returns would lead to more reliable parameter estimations, it is possible that daily or monthly returns exhibit different data features than quarterly or yearly returns. According to the literature, see for example Engle [1982], daily returns indeed exhibit a different data feature. While daily returns exhibit autocorrelation effects in their squared returns, this feature is not found in quarterly or yearly returns. Further Fama [1976] state that daily returns tend to depart more from normality than monthly returns. These properties can make daily data unqualified for the purpose of this research. Based on this information there is a certain trade-off between reliable parameter estimations and modeling the right data features. In order to be able to make a good choice, the data features of daily and monthly data are tested before a decision is made whether they can be used to represent the data features of a long term investment horizon.

Before testing it is important that the raw data points are transformed from indices values into continuously compounded returns. These continuously compounded returns are computed using equation 2.1:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (2.1)$$

Where  $P_t$  is today's index level,  $P_{t-1}$  is the index level one time period ago and  $R_t$  is today's continuously compounded return. This formula represents the usual way of computing continuously compounded returns, also referred to as 'log returns'.

Besides investigating differences in data frequency it is also interesting to look at different data periods for robustness. Especially since a crisis period is present in the data sample, which is also mentioned by Baily et al. [2008], Staps [2008]. To check the differences in different data periods three different data samples are used within the estimations, namely: sub sample one, with returns from 01-01-1999 till 01-01-2007, sub sample two ranging from 01-01-2007 till 01-04-2010 and the third data sample is the whole data sample.

Now that it is clear how the returns are computed and how the different data samples are selected it is important to test which data frequency should be used. In figure 2.1 the autocorrelation in the squared returns of stocks and real estate are shown for both daily and monthly returns together with the according 5% significance level.

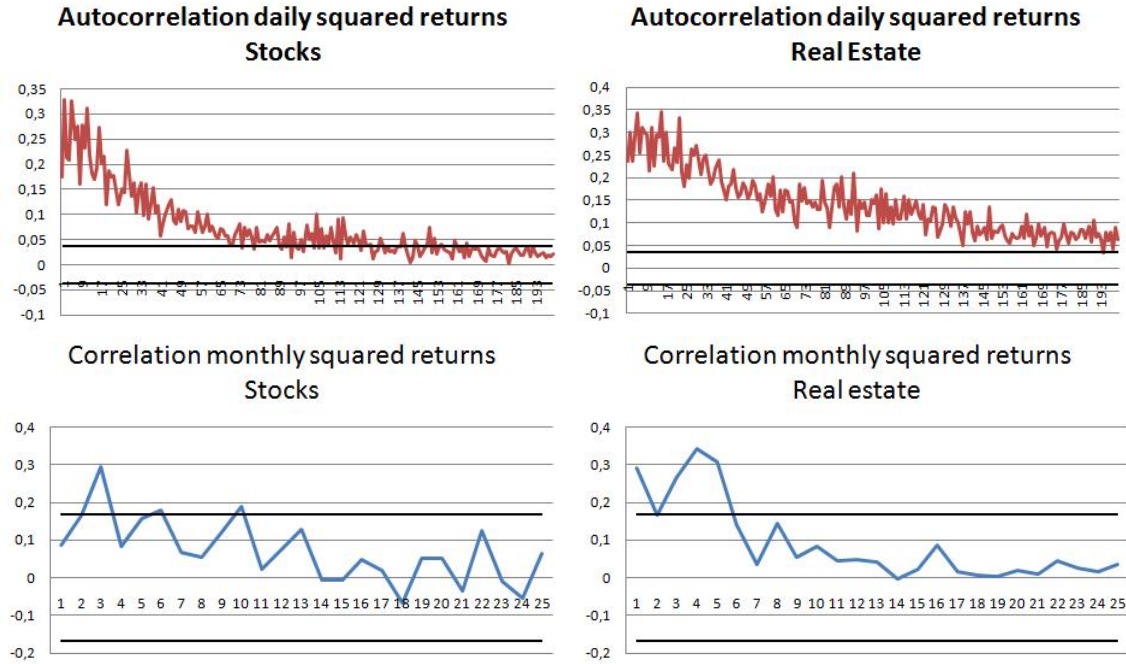


Figure 2.1: This figure shows the autocorrelation in the squared returns for stocks and real estate for daily and monthly returns. The black line in the figure represents the 95% confidence level. The figure indicates that the daily squared returns have a significant autocorrelation up to lag 75 (for stocks), while for the monthly squared returns the autocorrelation is significant for only some lags. Autocorrelation plots for the other asset classes are in the appendix.

Figure 2.1 shows that the daily squared returns have a significant autocorrelation up to lag 75 for stocks. For the daily squared returns of real estate the autocorrelation is significant for almost all lags. On the other hand, for the monthly squared returns there are only a few significant lags. Based on the significance of the autocorrelation the conclusion is that autocorrelation is indeed more present in the daily squared returns than in monthly squared returns, as described in Engle [1982]. Given that autocorrelation is not a data feature of quarterly or yearly data, the conclusion is that monthly returns do represent this data feature of quarterly and yearly data while the daily data sample does not.

Another feature described by the literature for daily returns is that they tend to depart more from normality than monthly returns. To test whether this feature is present in the empirical data the Jarque-Bera test for normality is used. This test is used to test whether a data sample is significantly different from the normal distribution. The null-hypothesis of this test is that the distribution of the returns is not significantly different from the normal distribution, while the alternative hypothesis is that the distribution of returns is different from the normal distribution. The Jarque-Bera test statistic is computed as follows:

$$Jarque\ Bera = \frac{N}{6} \times (S^2 + \frac{1}{4}(K - 3)^2) \sim \chi_2^2 \quad (2.2)$$

With  $N$  the number of observations,  $S$  the sample skewness and  $K$  the sample kurtosis. Equation 2.2 indicates that the Jarque-Bera (JB) test follows a chi-squared distribution with two degrees of freedom. Therefore the p-value of this test is computed by using this chi-squared distribution with two degrees of freedom. Based on



the p-value one can conclude whether or not the null hypothesis of the JB test is rejected. In table 2.2 the results of the Jarque-Bera test are shown.

Monthly data	Bonds	Credits	Stocks	Real Estate	Commodities	Daily data	Bonds	Stocks	Real Estate	Commodities
Sub-sample 1 (01-01-1999 till 01-01-2007)						Sub-sample 1 (01-01-1999 till 01-01-2007)				
Jarque-Bera	2.548	75.709	6.548	4.640	1.242	Jarque-Bera	116.374	379.798	161.238	180.976
p-value	0.140	0.000	0.019	0.049	0.269	p-value	0.000	0.000	0.000	0.000
Sub-sample 2 (01-01-2007 till 01-04-2010)						Sub-sample 2 (01-01-2007 till 01-04-2010)				
Jarque-Bera	4.821	31.266	0.803	3.008	17.866	Jarque-Bera	85.641	976.867	164.071	155.596
p-value	0.045	0.000	0.335	0.111	0.000	p-value	0.000	0.000	0.000	0.000
Whole sample (01-01-1999 till 01-04-2010)						Whole sample (01-01-1999 till 01-04-2010)				
Jarque-Bera	1.439	1870.755	6.642	43.007	60.672	Jarque-Bera	331.675	2079.629	1897.401	739.156
p-value	0.244	0.000	0.018	0.000	0.000	p-value	0.000	0.000	0.000	0.000

Table 2.2: This table shows the Jarque-Bera test statistic and the associating p-value for the daily and monthly data. Further the statistics are shown for both monthly and daily data divided into sub sample 1, sub sample 2 and the whole data sample.

The JB test statistic indicates a very strong rejection of the null hypothesis for the daily returns of all asset classes. Even on a 99% confidence level the null hypothesis that the daily asset returns come from a normal distribution is still rejected. Which indicates that the distribution of the returns is significantly different from the normal distribution. Noteworthy is that the Jarque-Bera test statistic for the monthly asset returns is much lower. This indicates that the rejection of the normal distribution is weaker for monthly returns than for daily returns. For monthly returns the null-hypothesis of a normal distribution is also not rejected for all asset classes. For example for the monthly bond returns the null hypothesis that the returns come from a normal distribution is not rejected by the JB test. Based on the p-value of 0.244 the null hypothesis is not even rejected when a 90% confidence interval is used. Given this result the conclusion is that the common feature of daily data departing more from normality than monthly data as described by Fama [1976] is also present in the empirical data set used for this research.

Given that the data is used to estimate parameters, which on their turn are used to simulate future returns far into the future, the conclusion is that the data features of the daily data makes daily returns less usable for this research. As a result the trade-off between capturing of the right data features and having enough observations in the data sample is in favor of the monthly returns since the monthly returns capture both requirements the best and are therefore used in this research.

The capturing of data features is important in the sense that it influences the choice of possible models used to capture the data features. Therefore table 2.3 shows some commonly used data descriptives for the continuously compounded monthly returns.

Monthly data	Bonds	Credits	Stocks	Real Estate	Commodities
Sub sample 1					
$\mu$ in %	0.361	0.040	0.412	1.013	1.096
$\sigma$ in %	1.004	0.272	4.649	4.543	6.377
$S$	-0.354	-0.677	-0.639	-0.440	-0.172
$K$	2.631	7.135	3.053	3.622	2.561
<i>Minimum</i> in %	-1.771	-1.100	-12.420	-13.240%	-14.987
<i>Maximum</i> in %	2.987	0.886	8.209	13.231%	14.657
Sub sample 2					
$\mu$ in %	0.461	-0.123	-0.194	-0.898	-0.480
$\sigma$ in %	1.163	1.546	5.352	8.340	9.785
$S$	0.499	-1.078	-0.324	-0.662	-1.139
$K$	4.377	6.757	2.753	3.226	5.352
<i>Minimum</i> in %	-1.863	-5.700	-11.962	-20.758	-34.992
<i>Maximum</i> in %	4.245	2.967	11.641	15.753	17.838
Whole sample					
$\mu$ in %	0.390	-0.008	0.234	0.451	0.632
$\sigma$ in %	1.050	0.865	4.854	5.948	7.536
$S$	-0.005	-2.137	-0.541	-0.934	-0.911
$K$	3.504	20.660	2.964	5.025	5.718
<i>Minimum</i> in %	-1.863	-5.700	-12.420	-20.758	-34.992
<i>Maximum</i> in %	4.245	2.967	11.641	15.753	17.838

Table 2.3: This table shows the descriptive statistics of the monthly continuously compounded returns over three different data samples. As stated in the table the mean, standard deviation, maximum and minimum are all in percentages. The average return for all assets is positive in sub-sample one, while the average return of most asset classes is negative in sub-sample two. This is because the first sub-sample is during a normal market period, while the second sub-sample is during a crisis period. The crisis also results in a higher standard deviation for all assets. Overall the returns on credits, real estate and commodities exhibit a strong negative skewness and a substantial excess kurtosis.

Table 2.3 shows the descriptive statistics for the the monthly continuously compounded returns over the three different sample periods. As stated in the table the mean, standard deviation, maximum and minimum are all in percentages. The results show that the average return is positive in sub-sample one, while the average return is negative in sub-sample two for all assets except for bonds. This is explained by the fact that sub-sample two is indicated as a crisis period, see also Baily et al. [2008], Staps [2008], while the first sub-sample is indicated as a normal period. The fact that the returns on bonds are not negative during the crisis period indicates a lowering of interest rates by central banks leading to a so-called "flight-to-quality". This is a phenomenon where investors sell their risky positions and invest this money in safe assets, such as sovereign bonds, resulting in a higher demand. This higher demand results in positive returns on bonds due to lower yields and an increase in the price. The difference between the normal period in sub-sample one and the crisis period in sub-sample two also has an impact on the standard deviation. The standard deviation on the returns on all asset classes is higher in the second sub-sample than in the first sub-sample. Further the difference in standard deviation between the asset classes themselves indicates that for the returns on stocks, real estate and commodities the standard deviation is much higher than for the bond returns and the excess returns on credits.

Other interesting features are skewness and kurtosis. From table 2.3 one can see that, compared with the first sub-sample, the kurtosis during the second sub-sample is higher for credits and commodities, while the kurtosis is lower for the other three asset classes. Further there is a large difference in kurtosis between the five different asset classes. Stocks and bonds have a kurtosis of around three, while the kurtosis of credits, real estate and commodities are respectively equal to 20.7, 5.0 and 5.7. The high values for the kurtosis indicate that credits, real estate and commodities exhibit a large excess kurtosis. Excess kurtosis is defined as the kurtosis above the kurtosis of the normal distribution, which is equal to three. To be specific excess kurtosis indicates that extreme returns occur more often than under the assumption of a normal distribution. Based on this data feature it is important to take other distributions than the normal distribution into account in order to capture these extreme returns. In section 3 this is discussed in further detail.

The skewness for the different asset classes shown in table 2.3 indicates a negative skewness for all asset classes over all data samples, except for bonds in sub-sample two. This positive skewness of bonds during the crisis period is, just as for the positive return on bonds over this period, explained by the "flight-to-quality". The negative skewness of the other asset classes indicates that large negative returns occur more often than large positive returns. Just as for the excess kurtosis it is important to take different distributions than the normal distribution into account in order to capture this data feature of negative skewness. This is discussed in further detail in section 3.

In the current ALM tool a multivariate normal distribution is used to generate future asset returns. However,

the analysis of the data descriptives and the results of the Jarque-Bera test indicate that the assumption of a normal distribution is strongly rejected for at least three of the five asset classes. Although normality is assumed for all asset classes in the ALM tool, the returns on sovereign bonds are generated based on interest and thereby separately from the returns on the other asset classes. Since the assumption of normality is not rejected for sovereign bonds there is no need to change the generating process for the inflation rates and return on the sovereign bond returns, which implies focusing on the return generating process for corporate bonds (credits), stocks, real estate and commodities is enough. Besides the rejection of the normal distribution, this section indicates the presence of excess kurtosis and a negative skewness in the returns of the asset classes. In section 3 it is discussed how these data features are taken into account for each specific asset class and how they are captured by the modeling of the marginal distribution for the specific asset classes.

### 3 Marginal distributions

One of the purposes of this research is to simulate realistic future returns. Based on the data analysis made in section 2 a choice is made on the possibilities for the simulation of future returns. One of the possibilities is to use a method called historical simulation. Simulations made with the historical simulation method are based on the observed returns. For historical simulation a simulated return is randomly drawn out of the observed returns. In this way it is possible to simulate future returns. However, a drawback of this method is that the possible outcomes for future returns are limited and, further, the size of the possible outcomes is limited. The size is limited since it is not possible to draw a return that is more extreme than empirically observed, while the possible outcomes are limited since it is not possible to simulate a future return different from empirically observed. Given that there are only 136 observations in the data sample, historical simulation is not suitable to simulate future returns for this research,

Another possibility is to attempt to model the marginal distribution of the observed returns. Given a set of observed returns it is possible to model these observed returns by using a marginal distribution. The marginal distribution can be any kind of known distribution. The advantage of using a marginal distribution to describe the observed returns is that it is possible to simulate future returns that are not yet empirically observed. Further, the use of a marginal distribution enables the simulation of more extreme returns than empirically observed. A problem with using a marginal distribution can be the fit with the empirically observed returns. Since any kind of distribution can be used as a marginal distribution for the observed returns it is possible that a used distribution does not have a good fit with the empirically observed returns.

Since any kind of distribution can be used as marginal distribution it is also possible to use multiple distributions and select the distribution that fits the empirically observed returns the best. The selection of distributions that are taken into account for this comparison is based on the data features described in section 2. The descriptive statistics are important because they should be captured by the possible marginal distributions.

Given the Jarque-Bera test in section 2 the conclusion is that the normal distribution is unable to capture the data features for stocks, credits, real estate and commodities. However, since the normal distribution is currently used within AEGON's ALM model this distribution is still taken into account for the comparison. Further section 2 described that most asset classes show excess kurtosis, therefore a distribution is needed in the comparison that accounts for this data feature. The distribution that has this property is the student-t distribution.

Besides excess kurtosis the continuously compounded monthly returns also exhibit a (large) negative skewness. In order to take this data feature into account, skewed distributions are used. Therefore the skewed normal and skewed student-t distributions are also added to the possible distributions. Based on the properties of the data the four different distributions that are used in order to test which distribution represent the empirical data the best are the normal distribution, the student-t distribution, the skewed normal distribution and the skewed student-t distribution. From these four possible distributions the optimal distribution is selected based on best fit to the empirically observed returns. The best fitting distribution is then selected to be the marginal distribution for the specific asset class.

Simulations from the possible distributions are based on certain parameters that are needed for the distribution. The parameters for the possible distributions are estimated based on the empirically observed returns. With a method called maximum likelihood estimation (MLE) it is possible to estimate the optimal parameters for a specific distribution based on the empirical data in order to ensure the best possible fit for that distribution. This estimation of the optimal parameter values for the different distributions is based on the standardized observed returns in order to acquire the best fit for the specific distribution. Therefore the observed returns are first transformed into standardized returns, which is shown by equation 3.1:

$$z = \frac{R - \bar{R}}{std(R)} \quad (3.1)$$

Where  $z$  represents the standardized returns,  $\bar{R}$  is the mean of the continuously compounded returns and  $std(R)$  is the standard deviation of the continuously compounded returns.

In the remainder of this section the density functions for the normal, student-t, skewed normal and skewed student-t distribution are presented together with the natural logarithm of these density function in section 3.1. The tests for the optimal distribution, given the empirical distribution, are discussed in section 3.2. Section 3.3 presents the optimal parameter results and the results for the best fitting distribution and section 3.4 contains the main conclusion.

## 3.1 Introduction marginal distributions

### 3.1.1 Normal distribution

The parameters of interest for the normal distribution are the mean  $\mu$  and standard deviation  $\sigma$ . The density function of the normal distribution is given by:

$$g(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.2)$$

The parameters  $\mu$  and  $\sigma$  can be estimated by MLE on the natural logarithm of this density function. However, in almost all programming packages the mean and standard deviation are pre-programmed functions. The estimation of the normal distribution is therefore only used to verify whether the right mean and standard deviation are used for the standardization of the returns. The parameter  $\mu$  and  $\sigma$  are estimated by performing MLE on the natural logarithm of the density function in equation 3.2 based on the "raw" continuously compounded monthly returns. As expected, the estimates of  $\mu$  and  $\sigma$  turned out to be the same as the pre-programmed ones and can be used in equation 3.1 to standardize the returns.

### 3.1.2 Student-t distribution

A problem in using the normal distribution as marginal distribution is that the kurtosis (fourth moment) is always equal to three. The consequence is that the normal distribution is unable to simulate a larger kurtosis, which is also known as "fat-tailness" in the distribution of returns. Given the data features observed in section 2 it is shown that over the whole data sample three of the four used asset classes have an empirical kurtosis substantially larger than three.

In order to capture the excess kurtosis of these asset classes, a student-t distribution is taken into account as possible marginal distribution. The student-t distribution is a flexible distribution in the sense that the shape of the distribution is determined by the degrees of freedom parameter  $\nu$ . This degrees of freedom parameter  $\nu$  is restricted between two and infinity, and for values of 100 and higher the student-t distribution becomes almost equal to the normal distribution. This makes the student-t distribution also suited as a marginal distribution for normally distributed asset classes. For smaller values for the estimated parameter  $\nu$  the student-t distribution becomes more "fat-tailed", or in other words a higher kurtosis is observed in the simulated returns.

To estimate the optimal value of  $\nu$  the density function of the student-t distribution is used. This density function is given in equation 3.3, see also Hansen [1994] :

$$g(z|\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)}\Gamma(\frac{\nu}{2})} \left(1 + \frac{z^2}{\nu-2}\right)^{-(\nu+1)/2} \quad (3.3)$$

With  $\Gamma$  the gamma distribution,  $z$  the standardized returns and  $\nu$  the degrees of freedom parameter for which it should hold that  $2 < \nu < \infty$ . Using MLE the optimal value for the degrees of freedom parameter  $\nu$  is estimated given the standardized returns  $z$ . The MLE is performed on the natural logarithm of the density function, which is shown in equation 5.4:

$$\ln(g(z|\nu)) = -\frac{1}{2} \times \ln((\nu-2) \times \pi) + \ln\left(\Gamma\left(\frac{\nu+1}{2}\right)\right) - \ln\left(\Gamma\left(\frac{\nu}{2}\right)\right) - \frac{\nu+1}{2} \ln\left(1 + \frac{z^2}{\nu-2}\right) \quad (3.4)$$

Where  $\ln$  is the natural logarithm,  $\nu$  the parameter that is optimized and  $z$  the standardized returns. The optimal degrees of freedom parameter  $\nu$  is found by maximizing the negative sum of equation 5.4 through MLE. This is represented by:

$$\text{maximize} - \sum \ln(g(z|\nu)) \quad (3.5)$$

### 3.1.3 Skewed normal distribution

Using the student-t distribution accounts for the observed kurtosis, but there is another data feature present in the empirical data shown in section 2. All asset classes show a substantial negative skewness in all three sub samples. To account for this large negative skewness the skewed normal distribution is used as a possible marginal distribution for the different asset classes.

In order to estimate the parameters of the skewed normal distribution the density function of the skewed normal distribution is needed. This density function is given by:

$$g(z|\lambda) = 2\phi(z)\Phi(\lambda z) \quad (3.6)$$

Where  $z$  are the standardized returns and  $\lambda$  is the skewness parameter. Using MLE the optimal value for  $\lambda$ , given the standardized returns  $z$ , is estimated. This is done by maximizing the negative sum of the natural logarithm of 3.6. For more information on the skewed normal distribution and the estimation procedure see for example Azzalini [2005] or the website of Azzalini.

### 3.1.4 Skewed student-t distribution

The student-t distribution is able to reproduce fat tails in an empirical distribution better than a normal distribution. However, fatter tails is not the only data feature shown by the data descriptives in section 2. Next to fat tails the asset classes also exhibit a substantial skewness. Besides capturing kurtosis and skewness apart from each other it is also possible to capture them together. Following Hansen [1994] a skewed student-t distribution is used to capture the empirical skewness as well as the empirical kurtosis. In order to estimate the parameters of the skewed student-t distribution the density function is used. This density of the skewed student-t distribution is given by:

$$g(z|\nu, \lambda) = \begin{cases} bc \left(1 + \frac{1}{\nu-2} \left(\frac{bz+a}{1-\lambda}\right)^2\right)^{-(\nu+1)/2} & z < -\frac{a}{b} \\ bc \left(1 + \frac{1}{\nu-2} \left(\frac{bz+a}{1+\lambda}\right)^2\right)^{-(\nu+1)/2} & z \geq \frac{a}{b} \end{cases} \quad (3.7)$$

Where  $\nu$  is the degrees of freedom parameter and  $\lambda$  is the skewness parameter, for which it holds that  $2 < \nu < \infty$  and  $-1 < \lambda < 1$ . Further  $z$  are the standardized returns and  $a$ ,  $b$  and  $c$  are given by:

$$a = 4\lambda c \left(\frac{\nu-2}{\nu-1}\right)$$

$$b^2 = 1 + 3\lambda^2 - a^2$$

$$c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)}$$

The by MLE estimated parameters are the skewness parameter  $\lambda$  and the degrees of freedom parameter  $\nu$ . However, MLE is performed on the natural logarithm of the density function. The by MLE optimized function is presented in Equation 3.8:

$$\ln(g(z|\nu, \lambda)) = \begin{cases} \ln(b)\ln(c) - \frac{(\nu+1)}{2} \ln\left(1 + \frac{1}{\nu-2} \left(\frac{bz+a}{1-\lambda}\right)^2\right) & z < -\frac{a}{b} \\ \ln(b)\ln(c) - \frac{(\nu+1)}{2} \ln\left(1 + \frac{1}{\nu-2} \left(\frac{bz+a}{1+\lambda}\right)^2\right) & z \geq \frac{a}{b} \end{cases} \quad (3.8)$$

Given equation 3.8 the parameters  $\nu$  and  $\lambda$  are estimated by:

$$\text{maximize} - \sum \ln(g(z|\nu, \lambda)) \quad (3.9)$$

## 3.2 Tests for empirical fit of marginal distributions

Four different parametric distributions are estimated and used to describe the data features of the four different asset classes. It is important to test which of these distributions is best to use as a marginal distribution, in order to describe the specific data features of a specific asset class. For this purpose two different "Goodness-of-Fit" (GoF) tests are used. These tests indicate numerically how good a certain distribution fits the empirical distribution of a specific asset class. Following Kole et al. [2007] the choice is to use two commonly used GoF tests, namely the Kolmogorov-Smirnov and the Anderson-Darling test. The difference between these tests is that the Anderson-Darling test puts more weight on the fit in the tails of the distribution, while the Kolmogorov-Smirnov test puts more weight to the fit of the center of the distribution.

### 3.2.1 Kolmogorov-Smirnov

The Kolmogorov-Smirnov test is used to test how well a parametric distribution fits the empirical distribution. This test is different from the Anderson-Darling test since deviations in the center of the distributions are more important. The Kolmogorov-Smirnov test statistic is computed as follows:

$$KS = \max_t |F_E(x_t) - F_H(x_t)| \quad (3.10)$$

Where  $KS$  is the Kolmogorov-Smirnov test statistic,  $F_E(x_t)$  is the value of the cumulative empirical distribution for point  $x_t$  and  $F_H(x_t)$  is the value of the cumulative hypothetical distribution at point  $x_t$ . The cumulative hypothetical distribution  $F_H$  is computed by using one of the assumed distributions together with it's estimated parameters to transform the empirical data into uniform values. Uniform values indicate that all values are between zero and one, whereby the largest observation is transformed into the highest value and the smallest observation is transformed to the lowest value. The cumulative empirical distribution  $F_E$  is computed by transforming the empirical data into uniform values by using the empirical distribution. Since all the observations are now uniform values, it is possible to compare the hypothesized distributions with empirical distribution and check the mismeasurement. By using the normal, skewed normal, student-t and skewed student-t distributions as hypothetical distributions the mismeasurement of all four parametric distributions are computed.

Besides only looking at the maximum distance between the empirical and hypothetical distribution it is also very informative to look at the average distance. Therefore a second Kolmogorov-Smirnov measure is used, which represents the average mismatch between the empirical and the hypothetical distribution. In equation 3.11 it is shown how this measure is computed.

$$KS_{average} = \int_x |F_E(x) - F_H(x)| dF_H(x) \quad (3.11)$$

### 3.2.2 Anderson-Darling

As mentioned before, the Anderson-Darling test is used to test the fit of a parametric distribution to an empirical distribution just like the Kolmogorov-Smirnov test. However, deviations in the tails of the distribution have a higher weight than deviations in the center of the distribution when compared with the Kolmogorov-Smirnov test. The value for the Anderson-Darling distance measure is computed by:

$$AD = \max_t \frac{|F_E(x_t) - F_H(x_t)|}{\sqrt{F_H(x_t)(1 - F_H(x_t))}} \quad (3.12)$$

Whereby  $AD$  is the Anderson-Darling distance measure,  $F_E(x_t)$  is the value for the cumulative empirical distribution for point  $x_t$  and  $F_H(x_t)$  represents the value for the cumulative hypothetical distribution for point  $x_t$ . Just like for the Kolmogorov-Smirnov test it holds that the cumulative hypothetical distribution  $F_H$  is computed by using one of the assumed distributions together with its estimated parameters to transform the empirical data into uniform values. The cumulative empirical distribution  $F_E$  is computed by transforming the empirical data into uniform values by using the empirical distribution. Since all the observations are now transformed into uniform values, it is possible to compare the hypothesized distributions with the empirical distribution and check the mismeasurement. By using the normal, skewed normal, student-t and skewed student-t distribution as hypothetical distributions the mismeasurement of all four distributions is computed. This Anderson-Darling test differs slightly from the original since  $|F_E(x_t) - F_H(x_t)|$  is used instead of  $(F(x_t) - F_H(x_t))^2$ . This replacement is done to reduce the influence of outliers as explained in Malevergne and Sornette [2003].

For the same reasons as for the Kolmogorov-Smirnov test it is also very informative to look at the average distance for the Anderson-Darling test. This average of the Anderson-Darling test is computed by:

$$AD_{average} = \int_x \frac{|F_E(x) - F_H(x)|}{\sqrt{F_H(x)(1 - F_H(x))}} dF_H(x) \quad (3.13)$$

Since the Anderson-Darling and Kolmogorov-Smirnov tests put more weight on different parts of the distribution they are complementary and can be used together. The optimal outcome is that both tests indicate the same distribution as the best fitting to a certain asset class. However, when both tests indicate a different distribution as the best fitting for a specific asset class, the choice for the optimal distribution is based on the purpose and the goals of the research.

## 3.3 Results

All the possible distributions are discussed and now it is interesting to check which distribution is best to use as marginal distribution to represent the asset classes credits, stocks, real estate and commodities.

This procedure starts with estimating the optimal parameters for every hypothetical distribution given the empirical data. As mentioned in section 3.1 estimating the parameters of the normal distribution leads to the same estimations for  $\mu$  and  $\sigma$  as using pre-programmed measures such as "mean" and "standard deviation". Therefore these values are only used for standardization. Based on these standardized returns the skewness and/or kurtosis parameters of the other hypothetical distributions are estimated. The results of this estimation are shown in table 3.1.

Credits			Stocks		
	$\nu$	$\lambda$		$\nu$	$\lambda$
normal	-	-	normal	-	-
skewed normal	-	-0.285	skewed normal	-	-0.588
student t	2.186	-	student t	159.3	-
skewed student t	2.172	-0.138	skewed student t	380.8	-0.336

Real estate			Commodities		
	$\nu$	$\lambda$		$\nu$	$\lambda$
normal	-	-	normal	-	-
skewed normal	-	-0.537	skewed normal	-	-0.541
student t	5.028	-	student t	7.848	-
skewed student t	5.008	-0.325	skewed student t	8.641	-0.299

Table 3.1: This table shows the parameter estimations for the four different parametric distributions. The mean and standard deviation are not shown, since the estimations are based on standardized returns. The normal distribution doesn't use a DoF parameter and also has no skewness parameter. The skewed normal distribution only has a skewness parameter  $\lambda$  and the student t distribution only had a DoF parameter  $\nu$  to correct for kurtosis. The skewed student t distribution has a  $\nu$  and  $\lambda$ . Noteworthy is the high value for  $\nu$  for stocks, which already indicates that a normal distribution might be more appropriate.

Table 3.1 shows no estimation results for the normal distribution. This is because the normal distribution has no DoF parameter  $\nu$  or skewness parameter  $\lambda$ , which has to be estimated. Given that the skewed normal distribution only accounts for skewness there is also only a value shown for the skewness parameter  $\lambda$ . For the student t distribution it holds that it only accounts for kurtosis, which is why there is only a value for the DoF parameter  $\nu$  shown. At last the skewed student t distribution takes both skewness and kurtosis into account and therefore has parameter estimates for both  $\nu$  and  $\lambda$ .

Noteworthy is the high value for  $\nu$  which is found for stocks in the estimation of the student t distribution and the skewed t distribution. This high value indicates that the (skewed) student-t distribution is very close to the normal distribution, which therefore might be more appropriate to represent the asset class stocks. The optimal distributions are tested by the GoF tests in order to show which distribution has the best fit to the empirically observed returns for each asset class.

In figure 3.1 the empirical distribution for all four asset classes are plotted together with the four different hypothesized distribution.



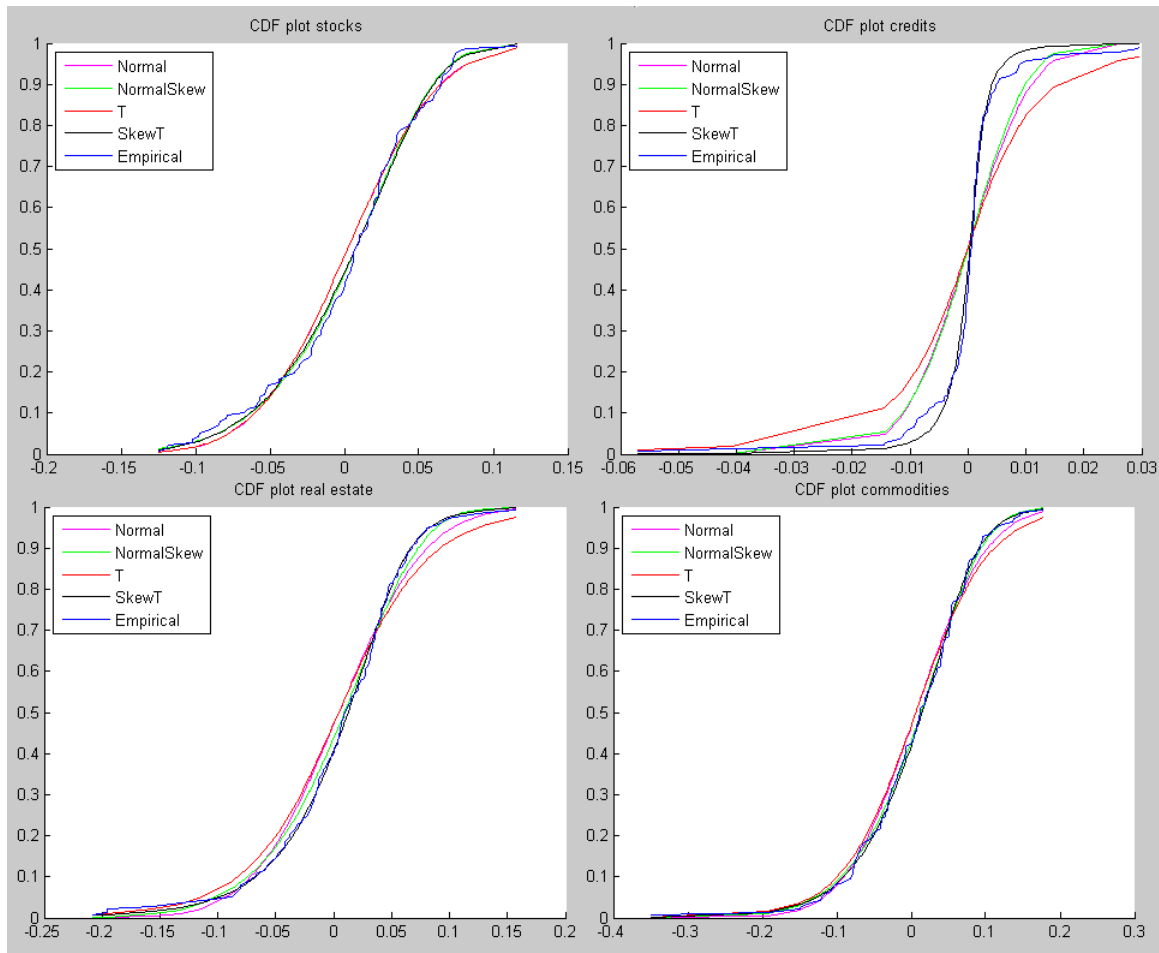


Figure 3.1: This figure shows the cumulative distribution for the empirical observations of the (skewed) normal, (skewed) student t and the empirical distribution. Based on these plots it is only possible for credits to say something about which hypothetical distribution represents the empirical distribution the best. However, for the other three asset classes it is not so clear which hypothetical distribution is best. Therefore GoF test are used.

Figure 3.1 shows the cumulative distribution for the hypothetical distributions together with the empirical distribution. Based on these figures it is possible to conclude that for credits the skewed student t distribution represents the empirical distribution the best. However, for the other three asset classes it is not possible to conclude which hypothetical distribution is best based only on these figures. So these figures give a good impression about the GoF for the hypothetical distribution, but the GoF tests are necessary to indicate which distribution has the best fit. In table 3.2 the results of these GoF tests are shown for every hypothetical distribution and for every asset class.

Credits					Stocks				
	<i>KS</i>	<i>KS<sub>average</sub></i>	<i>AD</i>	<i>AD<sub>average</sub></i>		<i>KS</i>	<i>KS<sub>average</sub></i>	<i>AD</i>	<i>AD<sub>average</sub></i>
Normal	0.210	0.116	1512	11.49	Normal	0.078	0.032	0.312	0.085
Student t	0.217	0.131	0.832	0.325	Student t	0.078	0.032	0.319	0.084
Skewed normal	0.203	0.111	83.23	0.903	Skewed normal	<b>0.039</b>	<b>0.017</b>	<b>0.134</b>	<b>0.047</b>
Skewed student t	<b>0.072</b>	<b>0.030</b>	<b>0.315</b>	<b>0.088</b>	Skewed student t	0.045	0.019	0.146	0.051

Real estate					Commodities				
	<i>KS</i>	<i>KS<sub>average</sub></i>	<i>AD</i>	<i>AD<sub>average</sub></i>		<i>KS</i>	<i>KS<sub>average</sub></i>	<i>AD</i>	<i>AD<sub>average</sub></i>
Normal	0.079	0.038	1.092	0.110	Normal	0.054	0.024	6.864	0.111
Student t	0.088	0.047	0.367	0.122	Student t	0.061	0.030	0.303	0.082
Skewed normal	0.052	0.021	0.418	0.060	Skewed normal	<b>0.031</b>	<b>0.010</b>	0.875	0.035
Skewed student t	<b>0.035</b>	<b>0.009</b>	<b>0.163</b>	<b>0.025</b>	Skewed student t	0.032	0.011	<b>0.223</b>	<b>0.029</b>

Table 3.2: This table shows the results of the GoF tests for the Kolmogorov-Smirnov test and the Anderson-Darling test for the four different parametric distributions for each asset class. Based on these values a decision is made on which parametric distribution fits the empirical distribution the best, where a lower value for the test indicates a better fit to the empirical distribution.

Table 3.2 shows the results of the Kolmogorov-Smirnov and the Anderson-Darling test for the four assumed distributions for all four asset classes. Given that the computed values of the Kolmogorov-Smirnov and the Anderson-Darling test are mismeasurements a lower value for the test statistic indicates a better fit to the empirical distribution. Therefore the lowest values are all printed in bold. Table 3.2 indicates that for credits, stocks and real estate the Kolmogorov-Smirnov and Anderson-Darling test find the same optimal assumed distribution. Based on these results the conclusion is that the skewed student t distribution is best to use as marginal distribution for credits and real estate, while for stocks the skewed normal distribution is best to use as marginal distribution. However, for commodities the Kolmogorov-Smirnov test indicates a different optimal distribution than the Anderson-Darling test. To resolve this problem a decision is made based on the research purpose. Although a good fit in the center of the distribution is wanted, the fit in the tail of the distribution is found to be slightly more important. Based on this information the result of the Anderson-Darling test is followed and therefore the skewed student-t distribution is used as marginal distribution for commodities.

### 3.4 Conclusion

The best empirical fit for a specific asset class is obtained when the empirical distribution is used. However, the use of the empirical distribution limits the number of possible outcomes when simulations are made, especially when a small data sample is used. Given that there are 136 monthly returns used, there are also only 136 different possible outcomes when simulations are based on the historical simulation. Another disadvantage is that using the empirical distribution limits the possible outcomes. This makes it impossible to draw an observation more extreme than historically observed. Since this makes the use of the empirical distribution inadequate, a marginal distribution is used to describe the observed returns. In order to use a marginal distribution a choice is made on the possible distributions that might be used as marginal distribution. Based on the possible distributions it is possible to test which of the possible distributions fits the empirically observed data the best and therefore is best to use as the marginal distribution. Before a comparison is made it is first important that the optimal parameters are estimated given the empirical data. By estimating the parameters for the possible distribution as described in section 3.1 the optimal parameters are found.

After the optimal parameters are estimated it is possible to test which of the possible distributions has the best fit to the empirically observed returns. By using the Kolmogorov-Smirnov and the Anderson-Darling tests the conclusion is that for credits, real estate and commodities the skewed student t distribution fits the empirical distribution the best, while for stocks the skewed normal distribution is best to use as marginal distribution.

Since the ALM tool is based on portfolio returns, which on their turn are based on different asset class returns it is important to find a marginal distribution that has a good fit to the empirically observed returns. Now that the optimal marginal distributions are found it is possible to simulate future returns for the asset classes separately. The problem is that when a portfolio return is simulated by simply using a weighted average of the simulated future asset returns separately, the dependencies between the asset class returns are not taken into account. Therefore, section 4 discusses different correlation measures in order to account for the dependencies between the asset class returns.

# 4 Correlation measures

In section 3 the optimal distributions are found to model the marginal distribution of the different asset classes. The interest of this research is not in the modeling of the returns of the specific asset classes but in the modeling of portfolio returns. It is possible to form portfolio returns from the specific asset class returns by using a weighted average, but in this way the dependency between the specific asset classes is not taken into account. It is important to take this dependency between asset returns into account since this dependency influences the portfolio returns.

The mostly used dependency measures are correlation measures. From these correlation measures the mostly used measure is the Pearson product-moment correlation. This correlation measure is based on a linear correlation between the different components of a portfolio. Besides the use of linear correlation another possibility is to use a correlation measure based on rank correlations. Two commonly used correlation measures which are based on rank correlation are the Spearman's rho and Kendall's tau.

Within the current ALM tool a multivariate normal distribution is used to generate portfolio returns. The dependency between the asset class returns is taken into account by using the Pearson product-moment correlation. Within this environment the Pearson product-moment correlation measure performs well especially when enough observations are used. However, the performance of this correlation measure is not robust in the sense that if the data is not captured by a normal distribution the performance becomes bad, see also Lindskog [2000].

Given the optimal distribution to describe the marginal distributions of the asset classes in section 3 it is demonstrated that it is very likely that the multivariate normal distribution is not optimal. Therefore it is possible that other correlation measures become more accurate than the Pearson product-moment correlation. Furthermore, it is described in Lindskog [2000] that the Pearson product-moment correlation measure has a bad performance especially when extreme observations are present in the data. Since the results shown in section 2 indicate that the empirical data exhibits extreme observations it is useful to test the influence of this data feature on the different correlation measures.

The outline of this section is as follows: In section 4.1 the computation of correlation is discussed for the three different correlation measures. Section 4.2 discusses the simulated data sets together with a way to test the accuracy of the correlation estimates. The results on the accuracy of the correlation measures are discussed in section 4.3. At last section 4.4 contains the main conclusions on the correlation measures.

## 4.1 Correlation measures

### 4.1.1 Pearson product-moment correlation

The Pearson's product-moment correlation measure was introduced by Karl Pearson around 1900. This correlation measure is nowadays the standard measure for measuring the dependence between two variables and is adapted in almost all computer packages. This measure is computed as follows:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (4.1)$$

Where  $X$  and  $Y$  are a sample of observations,  $\bar{X}$  and  $\bar{Y}$  are the mean of the sample  $X$  and  $Y$  and  $r$  is the Pearson product-moment correlation measure. A closer look at formula 4.1 shows that the Pearson product-moment correlation is based on some other familiar measures, namely:

$$r_{xy} = \frac{cov(X, Y)}{\sigma_X \sigma_Y} \quad (4.2)$$

Where  $\sigma_X$  and  $\sigma_Y$  are the standard deviations of sample  $X$  and  $Y$  and  $cov(X, Y)$  is the covariance between the samples  $X$  and  $Y$ .

### 4.1.2 Kendall's tau

A different measure for measuring dependence is the Kendall's tau correlation. This correlation measure is based on the concordance and discordance between two variables. Whereby a pair of variables is concordant if:  $(X_i - X_j)(Y_i - Y_j)$  is zero or positive and discordant if  $(X_i - X_j)(Y_i - Y_j)$  is negative. Given that this

measure is based on (dis)concordance of a pair of variables, this correlation measure is from the category of rank correlation measures. Using the (dis)concordance of the variables, Kendall's tau is computed by:

$$\rho_{\tau}(X, Y) = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \text{sign}((X_i - X_j)(Y_i - Y_j)) \quad (4.3)$$

where  $\rho_{\tau}(X, Y)$  is the correlation between  $X$  and  $Y$  as measured by Kendall's tau,  $N$  is the total number of observations and  $\text{sign}$  is an operator function, which is given by:

$$\text{sign}((X_i - X_j)(Y_i - Y_j)) = 1 \quad \text{if } ((X_i - X_j)(Y_i - Y_j)) \geq 0$$

$$\text{sign}((X_i - X_j)(Y_i - Y_j)) = -1 \quad \text{if } ((X_i - X_j)(Y_i - Y_j)) < 0$$

As described in McNeil et al. [2005] the Kendall's tau measure is also applicable within a copula framework (the copula is discussed in section 5). In McNeil et al. [2005] it is also shown that based on a unique copula and continuous marginal distributions for  $X$  and  $Y$  Kendall's tau rank correlation is given by:

$$\rho_{\tau}(X, Y) = 4 \int_0^1 \int_0^1 C(u, w) dC(u, w) - 1 \quad (4.4)$$

Following one of the methods described in McNeil et al. [2005] the dependence is first estimated. After the dependency is measured it is possible to start the estimation procedure of the copula. The Kendall's tau rank correlation is computed based on the (dis)concordant pairs of the marginal distributions, using equation 4.3. For more information on Kendall's tau see also Kendall [1938].

### 4.1.3 Spearman's rho

Besides Kendall's tau there is another rank correlation measure, namely Spearman's rho. Using a sample of two variables both consisting of  $N$  observation, the Spearman's rho rank correlation is computed by performing multiple steps. First, the data is sorted based on one of the two variables. Hereafter all the observations are replaced by their according rank number, whereby the lowest observation is replaced by the number one and the highest observation is replaced by the number equal to  $N$ . After the replacement of all the observations for their according rank number the difference in rank number is computed by:

$$D = X_{rank} - Y_{rank}$$

By definition, the first element in  $X_{rank}$  is always one, since the data sample is sorted on  $X$ . However, the first element of  $Y_{rank}$  can be any number, since this is the observations of  $Y$  belonging to the lowest observation of  $X$ . Based on the vector  $D$ , Spearman's rho is computed by:

$$\rho_{spearman} = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} \quad (4.5)$$

A problem in this procedure can occur when two or more observations are exactly equal. Although this is not very common in an empirical data-set, the possibility should be taken into account. The solution in the case that two or more observations are equal is that the ranks of the equal observations are summed and divided by the number of equal observations:

$$\text{Rank}_{i,i+1,\dots,N} = \frac{\sum_i^M \text{Rank}_i}{M}$$

Where  $i$  is the first observation of a sequence of equal observations and  $M$  is the number of observations in the sequence of equal observations. This results in that all equal observations are replaced by the same rank number. For the computation of the Spearman's rho with equal observations, the number of equal observations in a sequence and the number of sequences with equal observations are denoted by a vector  $t$ . For a sample  $X$  with one sequence of 2 equal observations, this results in a vector  $t_x$  containing a 2 and further only the number 1. The formula used to compute the Spearman's rho rank correlation in the case of equal observations is given by:

$$\rho_{spearman} = \frac{1 - \frac{6}{N(N^2-1)} \sum_i D_i^2 - \frac{1}{2}(T_x + T_y)}{\sqrt{(1 - T_x)(1 - T_y)}} \quad (4.6)$$

With  $D_i$  the vector with differences in rank between sample  $X$  and sample  $Y$ ,  $N$  is the total number of observations and the values for  $T_x$  and  $T_y$  are given by:

$$T_l = \frac{1}{N(N^2 - 1)} \sum_{k=1}^N t_{k,l} (t_{k,l}^2 - 1)$$

where  $l$  can be either  $X$  or  $Y$  and  $t$  represents the vector containing the number of equal observations in a sequence.

In McNeil et al. [2005] it is described that besides for the Kendall's tau measure, the Spearman's rho measure is also applicable within a copula framework (a discussion on the copula is in section 5). Based on a unique copula and continuous marginal distributions for  $X$  and  $Y$  Spearman's rho rank correlation is given by 4.7:

$$\rho_{spearman} = 12 \int_0^1 \int_0^1 u w dC(u, v) - 3 \quad (4.7)$$

However, just like for Kendall's tau the Spearman's rho is not based on the copula representation since the dependence is estimated before the parameter(s) of the copula. Therefore the dependence measure Spearman's rho is computed based on equation 4.6. For more information on Spearman's rho rank correlation see for example Spearman [1904].

## 4.2 Simulation and tests

Before a choice is made on whether the Pearson product-moment correlation, Kendall's tau or Spearman's rho is best to use to measure the dependence of the empirical data described in section 2, it is first tested whether the same results are found as in Lindskog [2000]. For this purpose different data sets are simulated with known correlation. Given that the estimates should stay efficient with slightly contaminated data it is possible to assume that the correlation matrix used for the simulations represents the real correlation matrix.

After the estimations of the correlation with the three different estimators it is possible to compute an average mismeasurement and standard deviation over the estimates. In order to test the robustness the average mismeasurement and standard deviation of the estimates are computed based on 18 different data sets in order to ensure that the choice for a specific correlation measure is not only based on a specific distribution or a specific correlation matrix.

### 4.2.1 Simulation of data sets

As mentioned before 18 different data-sets are simulated in order to compare the three different correlation measures. For the purpose of robustness it is important that different multivariate distributions are used to simulate the data-sets. This is especially driven by the results shown in Lindskog [2000] where the Pearson product-moment correlation performs worse when the tails of a distribution get heavier. For this purpose different multivariate distributions are used for the simulation of data. The used distributions are the multivariate normal distribution and the student-t distribution. Given that the fatness of the tails of the student-t distribution are dependent of the value for  $\nu$ , two different values for  $\nu$  are used, namely  $\nu = 3$  and  $\nu = 20$ .

Examining the empirical data in section 2 indicates that extreme observations are present in the data-set. As discussed in Lindskog [2000] the extreme observations can influence the correlation estimates especially for the Pearson product-moment correlation. In order to account for this data feature extreme observations are taken into account in the simulation of the data-sets. This results in simulated data-sets, with the most extreme positive and negative observation from the empirical data or without these extremes.

Further robustness is tested by accounting for different correlation matrices. In order to ensure that the results are not for a specific correlation matrix  $R$ , three different correlation matrices are used for the simulation of the data-sets. These three different correlation matrices are:

$$R_1 = \begin{bmatrix} 1 & -.23 & 0.31 & 0.30 \\ -.23 & 1 & 0.15 & 0.25 \\ 0.31 & 0.15 & 1 & -.02 \\ 0.30 & 0.25 & -.02 & 1 \end{bmatrix} \quad R_2 = \begin{bmatrix} 1 & 0.23 & -.31 & -.30 \\ 0.23 & 1 & -.15 & -.25 \\ -.31 & -.15 & 1 & 0.02 \\ -.30 & -.25 & 0.02 & 1 \end{bmatrix} \quad R_3 = \begin{bmatrix} 1 & .5 & .5 & .5 \\ .5 & 1 & .5 & .5 \\ .5 & .5 & 1 & .5 \\ .5 & .5 & .5 & 1 \end{bmatrix}$$

The values representing the correlations in the matrices are not completely random, since a restriction is that a correlation matrix has to be positive definite in order to use for data simulation.

Three different correlation matrices used for data simulations, with three multivariate distributions and either with extreme observations or without results in  $3 \times 3 \times 2 = 18$  different data-sets. Given that the empirical data consists of four different asset classes, the simulated data-sets all consist of 100 observations for four different asset classes.

### 4.2.2 Distance measures

Based on the 100 observations of each asset class an estimate for the correlation between the asset classes is made with the three different correlation measures. Given that the correlation matrix used to simulate the data-set is assumed to represent the real correlation, it is possible to measure the mismeasurement of the different correlation measures. This mismeasurement is computed by calculating the difference between the estimated correlation matrix and the real correlation matrix. Given that there are only 100 observations in the simulated

data-set, the simulated data points vary because of randomness. As a consequence the mismeasurement in the estimated correlations will also vary. In order to take the randomness effect out of the correlation estimates, the procedure is repeated 100 times. In this way an average mismeasurement and standard deviation over the estimations is computed. Based on the 100 mismeasurements, a two-sided student t-test is used to calculate whether the average mismeasurement is significantly different from zero. The t-value for this student-t test is computed by:

$$T = \frac{\bar{X} - \mu_0}{S} \sqrt{n} \quad (4.8)$$

Whereby  $\bar{X}$  is the sample difference between the real and estimated correlation,  $\mu_0$  is the assumed average difference between the real and estimated correlation (which is assumed to be zero),  $S$  is the sample standard deviation over the differences between the real and estimated correlations and  $n$  is the number of used observations.

Based on the observed value for  $T$ , an according p-value can be computed. This p-value is calculated based on the two-sided student-t distribution with 99 degrees of freedom. The number for the degrees of freedom is given as the number of used observations minus one. Since estimations are based on 100 observations, this leads to  $\nu = 99$ .

Given the computed p-value it is concluded whether there is significant evidence to reject the null hypothesis or not. The null hypothesis of this test is that there is not significant evidence to assume that there is a difference between the estimated correlation and the real correlation.

## 4.3 Results

### 4.3.1 Estimation results

In table 4.1 the estimation results for the three different correlation measures are shown, whereby correlation matrix  $R_1$  is used for data simulation. In this data generation process there are no extreme observations added and the estimations are based on 100 observations. The estimation is repeated 100 times in order to compute the shown average mismeasurement and according standard deviation (shown in brackets).

Spearman							Kendall						
	$\nu = 3$		$\nu = 20$		Normal			$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	1.534	(0.107)	0.446	(0.100)	0.434	(0.098)	$\rho_{1,2}$	0.218	(0.116)	0.330	(0.105)	0.091	(0.103)
$\rho_{1,3}$	-2.038 <sup>*</sup>	(0.099)	-1.130	(0.095)	-0.541	(0.092)	$\rho_{1,3}$	0.138	(0.106)	-0.722	(0.099)	0.134	(0.096)
$\rho_{1,4}$	-2.012	(0.103)	-1.109	(0.092)	-0.815	(0.096)	$\rho_{1,4}$	0.257	(0.110)	-0.030	(0.096)	-0.532	(0.100)
$\rho_{2,3}$	-0.890	(0.108)	-0.421	(0.102)	0.130	(0.098)	$\rho_{2,3}$	-0.494	(0.118)	0.139	(0.108)	-0.286	(0.104)
$\rho_{2,4}$	-1.749	(0.105)	-0.656	(0.098)	-0.766	(0.093)	$\rho_{2,4}$	-0.662	(0.112)	0.054	(0.102)	-0.185	(0.097)
$\rho_{3,4}$	-0.278	(0.106)	-0.205	(0.101)	0.328	(0.102)	$\rho_{3,4}$	-0.127	(0.115)	-0.194	(0.107)	-0.176	(0.107)

Pearson						
	$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	0.706	(0.211)	0.352	(0.104)	0.133	(0.097)
$\rho_{1,3}$	-0.076	(0.201)	-0.738	(0.098)	0.060	(0.091)
$\rho_{1,4}$	-0.093	(0.204)	0.009	(0.093)	-0.417	(0.092)
$\rho_{2,3}$	-0.362	(0.223)	0.116	(0.104)	-0.277	(0.098)
$\rho_{2,4}$	-0.676	(0.206)	0.077	(0.100)	-0.116	(0.093)
$\rho_{3,4}$	0.311	(0.220)	-0.162	(0.105)	-0.207	(0.101)

Table 4.1: Average mismeasurement based on correlation matrix  $R_1$  without extreme observations. The standard deviation is shown between brackets. The correlations are estimated based on 100 observations. This is repeated 100 times in order to compute an average mismeasurement and standard deviation. Further, the average mismeasurement is multiplied by 100 for presentation purposes. The <sup>\*</sup> behind a value shown for Spearman's rho indicates a rejection of the null hypothesis on the 5% significance interval. This means that Spearman's rho has one estimated correlation that is significantly different from the real correlation on the 5% significance interval. For all the other correlations of the Spearman's rho and all the other correlation estimates of Kendall's tau and the Pearson product-moment correlation the null hypothesis of no difference between the estimates and the actual correlation is not rejected. The results based on the tables for  $R_2$  and  $R_3$  can be found in the appendix since they show similar results.

Based on the values shown in table 4.1 one can see that for Spearman's rho there is one estimated correlation that is significantly different from the real correlation based on a 5% significance interval. For the correlation estimates of Kendall's tau and the Pearson correlation the conclusion is that there is not significance evidence to assume that there is a difference between the estimated and the actual correlation. This holds for all the shown correlation estimates. It is remarkable that the standard deviation in the correlation estimates made

with the Pearson measure become bigger when a distribution with fatter tails is used, which is in accordance with the result shown in Lindskog [2000].

From this table the conclusion is that for data without any extreme observations the Pearson product-moment correlation and the Kendall's tau both estimate the correlation without a significant error. The Spearman's rho has one rejection of the null hypothesis, which is only on the 5% significance level and not on the 2.5% significance level. Although Kendall's tau and Pearson correlation are preferred based on the values shown in table 4.1, it is also possible to use Spearman's rho. The results based on correlation matrix  $R_2$  and  $R_3$  can be found in the appendix, since they result in the same conclusion.

In Lindskog [2000] it is described that especially for extreme observations (fat tails and outliers) the Pearson product-moment correlation has a large mismeasurement. Therefore the estimation results for correlation matrix  $R_1$  are shown in table 4.2 when extreme observations are added in the data simulations.

Spearman							Kendall						
	$\nu = 3$		$\nu = 20$		Normal			$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	3.138*	(0.105)	2.207 <sup>^</sup>	(0.096)	2.114 <sup>^</sup>	(0.096)	$\rho_{1,2}$	1.385	(0.113)	1.916	(0.101)	1.607	(0.101)
$\rho_{1,3}$	-3.945*	(0.101)	-3.169*	(0.096)	-2.571*	(0.094)	$\rho_{1,3}$	-2.641 <sup>^</sup>	(0.108)	-2.386 <sup>^</sup>	(0.100)	-2.440 <sup>^</sup>	(0.098)
$\rho_{1,4}$	-3.929*	(0.103)	-3.091*	(0.094)	-2.743*	(0.091)	$\rho_{1,4}$	-2.700 <sup>^</sup>	(0.110)	-2.436 <sup>^</sup>	(0.099)	-2.229 <sup>^</sup>	(0.095)
$\rho_{2,3}$	-1.556	(0.105)	-1.448	(0.099)	-1.787	(0.099)	$\rho_{2,3}$	-1.002	(0.114)	-0.823	(0.104)	-1.536	(0.104)
$\rho_{2,4}$	-3.102*	(0.104)	-2.485 <sup>^</sup>	(0.097)	-2.118 <sup>^</sup>	(0.096)	$\rho_{2,4}$	-2.294 <sup>^</sup>	(0.112)	-1.871	(0.102)	-1.896	(0.100)
$\rho_{3,4}$	0.380	(0.110)	0.279	(0.104)	0.642	(0.098)	$\rho_{3,4}$	0.204	(0.120)	0.552	(0.110)	0.185	(0.103)

Pearson						
	$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	5.689*	(0.167)	6.111*	(0.099)	5.944*	(0.097)
$\rho_{1,3}$	-9.807*	(0.161)	-8.844*	(0.099)	-8.711*	(0.098)
$\rho_{1,4}$	-10.061*	(0.157)	-9.351*	(0.097)	-9.068*	(0.095)
$\rho_{2,3}$	-2.585	(0.182)	-2.135 <sup>^</sup>	(0.101)	-2.827*	(0.099)
$\rho_{2,4}$	-5.358*	(0.177)	-4.639*	(0.099)	-4.735*	(0.095)
$\rho_{3,4}$	0.339	(0.179)	0.832	(0.106)	0.435	(0.098)

Table 4.2: Average mismeasurement based on correlation matrix  $R_1$  with extreme observations. The standard deviation is shown between brackets. The correlations are estimated based on 100 observations. This is repeated 100 times in order to compute an average mismeasurement and standard deviation. Further the shown average mismeasurements are multiplied by 100 for presentation purposes (in the computations the real values are used). When a \* is shown after a value, this represent a rejection of the null hypothesis on the 1% significance interval, while a <sup>^</sup> indicates a rejection on the 5% significance interval. Based on this table it is shown that most correlation estimates made with the Pearson product-moment correlation measure are significantly different from the real correlation. For Kendall's tau the null hypothesis of correlation estimates that are equal to the actual correlation is not significantly rejected for any correlation based on the 1% significance interval. For Spearman's rho some estimated correlations are significantly different from the real correlation based on the 1% significance interval. This indicates that Spearman's rho performs better than the Pearson measure, but worse than Kendall's tau. Therefore, Kendall's tau correlation measure is preferred based on the values shown in this table. The tables for  $R_2$  and  $R_3$  show similar results and can be found in the appendix.

From table 4.2 one can see that adding extreme observations to the simulated data results in larger average mismeasurements in the correlation estimates. Where the Pearson product-moment correlation measure performed very well based on the conditions used for table 4.1, the performance is quite different. Based on table 4.2 the conclusion is that adding extreme observations to the data-set results in a significant difference between the correlation estimate of the Pearson product-moment correlation and the actual correlation for 13 from the 18 correlations based on the 1% significance interval. However, the addition of extreme observations does not only influence the estimations of the Pearson product-moment correlation. Also for the correlation estimates of the Spearman's rho the conclusion is that 8 out of 18 correlation estimates are significantly different from the actual correlation based on a 1% significance interval. Using a 5% significance interval leads to the conclusion that for 12 out of 18 correlation estimates the estimated correlation is significantly different from the actual correlation for Spearman's rho. The results for the estimates of Kendall's tau show only a slightly worse performance when extreme observations are added to the simulated data-sets. Based on a 1% significance interval the conclusion is that there is no significance evidence to assume that one of the correlation estimates is different from the actual correlation. Only when a 5% significance interval is used the conclusion is that 7 from the 18 correlation estimates are significantly different from the actual correlation.

Based on the results shown in table 4.2 the conclusion is that Kendall's tau is the best correlation estimator when extreme observations are present in the data. Further Spearman's rho also performs better than the Pearson product moment correlation measure when extreme observations are added. This indicates that rank correlation measures are better correlation estimators when extreme observations are present in a data-set. The results based on correlation matrix  $R_2$  and  $R_3$  can be found in the appendix, since the results are in accordance

with those shown in table 4.2.

## 4.4 Conclusion

In this section simulated samples are used in order to test whether it is best to the Pearson product-moment correlation, Kendall's tau or Spearman's rho as measure for the dependency between the different asset classes. Based on Lindskog [2000] the expectation is that the correlation estimates of the Pearson product-moment correlation measure will become worse when a heavier tailed distribution is used and when extreme observations enter the simulated data-sets. From the results the conclusion is that the average mismeasurement of the Pearson's product moment correlation estimates is not different when a student-t distribution is used instead of a multivariate normal distribution. However, the standard deviation over the mismeasurements does increase when more observations are in the tail of the distribution, which is in accordance with the result in Lindskog [2000]. Based on the results for a data-set without extreme observations the conclusion is that the Pearson product-moment correlation and Kendall's tau both estimate all the correlations without significant error. For the Spearman's rho there is only one correlation estimate that is significantly different from the actual correlation and therefore the Spearman's rho is also good enough to use as correlation measure when there are no extreme observations present in the data.

However, according to Lindskog [2000] the estimations of the Pearson product-moment correlation become especially worse when extreme observations are added to the data-set. Given the results based on a data-set with extreme observations this conclusion is confirmed. The correlation estimate of the Pearson product-moment correlation is significantly different from the actual correlation for almost all estimated correlation even on a 1% significance interval. On the other hand, all the correlation estimates made with Kendall's tau are not significantly different from the actual correlation based on a 2.5% significance interval.

Given that Kendall's tau and the Pearson product-moment correlation have an equal performance in a data-set without extreme observations, while the Kendall's tau performs a lot better when extreme observations are present in the data, the conclusion is that Kendall's tau is the best correlation measure. Further it is described in section 2 that in the actual data observations of multiple times the standard deviation are commonly observed. This indicates that the empirical data-set is in accordance with the simulations with extreme observations. Therefore Kendall's tau is used to measure the dependency between the different asset classes.



# 5 Copula

Now that the dependencies between the asset classes and all the marginal distribution of the different asset classes are known, a joint distribution is formed. A problem in forming the joint distribution is that the parametric distributions used as marginals for the different asset classes are not the same. Normally, when all the asset classes are modeled by the same distribution, a joint distribution is formed by using the multivariate distribution. However, this is not possible when the marginal distributions are not the same. Given that it should be possible to extend this research up to ten asset classes, it is not desirable to use a multivariate distributions. Copula theory allows the use of different distributions as marginal distributions when a joint distribution is formed.

Copula theory is based on Sklar [1959]. This paper shows that for continuous variables a decomposition is possible for the joint distribution. The joint distribution can be decomposed in to a copula (for the dependence between the variables) and the marginals. This can be represented as a continuous n-variate CDF  $H(y_1, \dots, y_n)$  with univariate marginals  $F_i(y_i)$  with  $i = 1, \dots, n$ . Whereby  $F_i(y_i) = H(\infty, \dots, y_i, \dots, \infty)$ . In Sklar [1959] it is stated that there is always a copula function  $C$  with  $[0, 1]^n$  which can be transformed to  $[0, 1]$  such that:

$$H(y_1, \dots, y_n) = C(F_1(y_1), \dots, F_n(y_n)) \quad (5.1)$$

By using the product of the copula density and the marginals the result is the following joint density function:

$$\frac{\partial H(y_1, \dots, y_n)}{\partial y_1 \dots \partial y_n} = \prod_{i=1}^n f_i(y_i) \frac{\partial C(F_1(y_1), \dots, F_n(y_n))}{\partial F_1(y_1) \dots \partial F_n(y_n)} \quad (5.2)$$

Based on this equation it is possible to define the copula as a multivariate distribution with uniform  $[0, 1]$  margins as shown by:

$$C(z_1, \dots, z_n) = H(F_1^{-1}(z_1), \dots, F_n^{-1}(z_n)) \quad (5.3)$$

With  $z_i = F_i(y_i)$  and  $i = 1, \dots, n$ , being the probability integral transformations of the marginal models. See also Cholette et al. [2008] for a more detailed description.

In this research the copula is used because it allows for the flexibility of using any possible marginal distribution together with a certain copula function in order to compute a joint distribution function. Given that the results in section 3 state that the optimal marginal distributions are not the same for every asset class it is inappropriate to use a multivariate distribution. Therefore the copula is the optimal solution in order to form a joint distribution in this research.

There are a lot of different copulas with different properties. Based on this one has to find the copula that fits the purpose of the research the best. Therefore, a few different copulas out of the broad variety of different copulas are discussed in section 5.1. Section 5.2 discusses the estimation and simulation procedure of the chosen copula, while the estimation and simulation results are presented in section 5.4. In section 5.5 the main conclusions are discussed together with an expected impact on the ALM model.

## 5.1 Different copulas

There is a broad variety of copulas. Some were developed a long time ago. While the impact of using copulas that are constructed out of smaller copulas is an upcoming interest in copula research. Since every copula has specific properties, it is important to know the properties of the different copulas in order to choose the copula that fits the research purposes the best. In this section four often used copulas are discussed together with a more complex copula.

### 5.1.1 Gaussian copula

The Gaussian copula is generally used when most marginal distribution are normally distributed, since the Gaussian copula has the same properties as the normal distribution. The parameters for estimation are  $\mu$ ,  $\sigma$  and possibly the correlation matrix  $\rho$ . Simulations from the Gaussian copula are correlated with each other with respect to the correlation matrix  $\rho$ . Within these simulations an equal dependence is assumed for simulations within the whole distribution. For a more detailed description on the estimation for the Gaussian copula and simulations from the Gaussian copula see for example McNeil et al. [2005].

### 5.1.2 Student-t copula

Besides the Gaussian copula for the normal distribution a student-t copula is most used when most marginal distributions are modeled by a student-t distribution. The difference between the student-t copula and the Gaussian copula is that the student-t copula allows for fatter tails within the simulated distribution than the Gaussian copula. The similarity between these two copulas is that they both only allow for equal dependency within the distribution. The parameter estimated for the student-t copula is the Degrees-of-Freedom (DoF) parameter  $\nu$  and possibly the correlation matrix  $\rho$ . Through the DoF parameter  $\nu$  the fatness of the tails in the distribution is controlled. A lower value for  $\nu$  indicates more observations in the tails of the distribution. For very large values of  $\nu$  (for example larger than 200) the "fatness" in the tails of the distribution is similar to simulations made with the Gaussian copula. Given this property the student-t copula is very flexible and therefore can be seen as an extension of the Gaussian copula. For more information on the student-t copula see McNeil et al. [2005].

### 5.1.3 Gumbell copula

The Gaussian and student-t copula both only allow for an equal dependence over the whole distribution. It is also possible to allow for dependence in only one side of the distribution. A copula that allows only for dependence on the upper side of the distribution of simulated future asset returns is the Gumbell copula. As a consequence a very positive simulated future returns in one asset class tends to be correlated with a large positive simulated future return in another asset class. This dependency is controlled by the estimation of the coherent parameter for the Gumbell copula. As mentioned this dependence does not hold for large negative simulated values. For more information on the Gumbell copula see Bouy  [2002].

### 5.1.4 Clayton copula

Besides only allowing for dependence in the upper side of the distribution it is also possible to only allow for dependence in the lower side of the distribution. By using a Clayton copula a large negative simulation in one asset class is correlated with a large negative simulation in another asset class, whereby the dependency is controlled by a parameter. However, large positive simulated values are uncorrelated with each other. This makes the Clayton copula a good simulator if one wants to know what the maximum loss of a certain portfolio over a certain time period might be. For a more detailed description see Cuvelier and Noirhomme-Fraiture [2005].

A drawback of the Gumbell and Clayton copula is that they allow for dependency in only one side of the distribution. Therefore these copulas are often to model the maximum or minimum. Since the scope of this research is on modeling the whole distribution of future returns the Clayton and Gumbell copulae do not fit the research purpose.

### 5.1.5 Other copulae

Using a large number of asset classes causes estimation error in the standard copulas since for all asset classes the same (copula) parameter(s) are used. A vine copula is a combination of bi-variate copula which are used to describe the dependence between two different asset classes. The advantage is that for every bi-variate copula different parameter estimations are possible, which enables a better approximation of the parameters. However, the number of estimations becomes very high for more asset classes, because this requires a lot of combinations. Further the convergence of vine copula is not always good enough and the estimations on the bi-variate copulas are not exact, which leads to an estimation error. It is possible to estimate two different tri-variate copula based on the same data-set - for an example see Guegan and Maugis [2008] - but for more asset classes the estimation errors become larger. Further the evaluation techniques of vine copulas have to be developed in order to point out the optimal combination for the vine copula and the estimation of vine copulas is computationally very intensive. To quote from Guegan and Maugis [2008]: "To compute the tails parameters of all the vines one needs  $n^2$  quotient of  $(n - 1)$ -th degree integrals for each  $\prod_{i=1}^n i!$  possible vines". This leads to an enormous computation time even for  $n = 10$  assets. Since it has to be possible to extend the research up to 15 asset classes, it is computationally too intensive to use this method. The conclusion is that even though vine copula are promising in a superior econometric fit, due to computation time, estimation error and the lack of evaluation methods vine copula are not suitable to use for the analysis of the data in this research.

Another possibility is to use mixture/vine copula in a copula-switching setting as done in Cholette et al. [2008]. Although correlations are better estimated this way, the ALM model is not updated every day or month. This makes the use of a time-varying model undesirable and therefore the method inappropriate to use because of the long term investment horizon of pension funds.

## 5.2 Student-t copula

For the purpose of this research it is important that the whole distribution is fitted and that an improvement is made to the currently used multivariate normal distribution. As mentioned, the Gumbell and Clayton copula allow for dependence in only one side of the distribution. Since a good fit over the whole distribution is needed for the purpose of this research, the Gumbell and Clayton copula are not used.

A different problem can be with the complexity of the estimation procedure for the copula. Since it should be able to extend the analysis up to more than ten asset classes, it would take a too much time to estimate the parameters of a vine copula.

The Gaussian copula is an alternative for the multivariate normal distribution. When only normal distributions are used as marginal distributions together with a Gaussian copula the resulting joint distribution would be exactly the same as the multivariate normal distribution. The advantage of using a Gaussian copula is that the marginals do not necessarily have to be a normal distribution. Since the 2008 financial crisis there is no doubt that the normal distribution is not able to capture the extreme negative return observed in the market. This makes the use of a Gaussian copula inappropriate. Therefore the best copula to use, to account for the extreme returns observed during the crisis period, is the student-t copula. Since the student-t copula account for fatter tails and allows for (symmetric) dependence over the whole distribution. Fischer et al. [2007] concludes that the student-t copula outperforms the other mentioned copulas in their research based on their empirical data sample. Based on the properties of the different copulas and the results in the research of Fischer et al. [2007], the student-t copula is chosen as the best copula for this research.

### 5.2.1 Estimation procedure

The estimation of the DoF parameter  $\nu$  for the student-t copula is based on the density of the student-t copula, which is given in equation 5.4:

$$C(u_1, u_2, \dots, u_n; \rho, \nu) = |\rho|^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2})} \left[ \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})} \right]^n \frac{(1 + \frac{\zeta' \rho^{-1} \zeta}{\nu})^{-\frac{\nu+n}{2}}}{\prod_{i=1}^n (1 + \frac{\zeta_i^2}{\nu})^{-\frac{\nu+1}{2}}} \quad (5.4)$$

Where  $n$  is the number of used asset classes,  $\nu$  is the degrees of freedom (DoF) parameter that is estimated for the student-t copula,  $\rho$  is the correlation matrix and  $\zeta$  is given by  $\zeta = (t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), \dots, t_\nu^{-1}(u_n))'$ . By performing maximum likelihood estimation (MLE) on equation 5.5 the optimal value for the DoF parameter  $\nu$  is estimated.

$$- \sum (\ln(C(u_1, u_2, \dots, u_n; \rho, \nu))) \quad (5.5)$$

This estimation procedure for calibrating t copulas is described in McNeil et al. [2005], where the DoF parameter  $\nu$  is estimated through MLE. However, in order to estimate  $\nu$  an estimate of the correlation matrix  $\rho$  is necessary as input parameter. Based on the results of section 4, the estimations made with Kendall's tau are used for the correlation matrix  $\rho$ .

In order to test that the used estimation procedure is correct a comparison is made. The results of the estimations made with this estimation procedure are compared with the results shown in the research of Kole et al. [2007]. This comparison is made by using the estimation procedure on the same data-set that is used in the research of Kole et al. [2007]<sup>1</sup>. The estimation of the correlation matrix  $\rho$  is made with Kendall's tau as discussed in section 4. Given that the estimates for the correlation matrix  $\rho$  differ slightly from the estimated values for correlation in the research of Kole et al. [2007], which is due to a different estimation method, a slight difference in the estimation for the DoF parameter  $\nu$  of the student-t copula is also expected. This is indeed the case since the estimation for  $\nu$  made with the method used in this research was very similar to the one shown in Kole et al. [2007]. The estimation comparison is shown in table 8.5 in the appendix. Based on these results it is ensured that the estimation procedure used in this research is indeed correct and can be applied on the data-set discussed in section 2.

### 5.2.2 Simulating returns from the student-t copula

Since a comparison is made between the existing multivariate normal method and the copula method, it is important that the same random values are used as input for both methods. In this way it is possible to check the difference in the transformation of the random values into actual returns between the two methods. As a consequence the randomly simulated values are treated as input for the simulation of uniform numbers from the student-t copula. Other input parameters which are used to simulate future returns with the student-t copula are the correlation matrix  $\rho$  and the estimated DoF parameter  $\nu$ .

The simulation of random values is done by using a simple function for drawing random values which is pre-programmed in almost every package. The correlation matrix  $\rho$  is estimated based on the Kendall's tau

<sup>1</sup>I want to thank Erik Kole for providing the data of his research

described in section 4 and the DoF parameter  $\nu$  is also a known estimate. Based on these estimates multiple transformations are made to transform the random values into future returns.

At first a Cholesky decomposition of the correlation matrix  $\rho$  is used to simulate correlated values for the different asset classes. Since it is only possible to perform a Cholesky decomposition on a positive definite matrix, it is first ensured that  $\rho$  is positive definite by using the algorithm described in McNeil et al. [2005].

After the Cholesky decomposition the correlated future returns are obtained by multiplying the outcome of the Cholesky decomposition with random normally distributed values. These random normally distributed values are obtained by transforming the random values. The random values are transformed into normally distributed numbers by using the inverse of the normal distribution with mean 0 and standard deviation 1. Equation 5.6 shows how the random normally distributed values are obtained:

$$Z = \text{CholeskyOutcome} * \text{NormalRandomNumbers} \quad (5.6)$$

This first part of simulating future correlated returns is the same for copula and multivariate (MV) normal methods. However, the future correlated uniform returns based on the MV normal method are the obtained values for  $Z$ , while the future correlated returns based on the copula method are obtained after performing multiple extra steps.

In order to simulate future asset returns from the copula method, the inverse of the  $\chi^2_\nu$  distribution is used. This  $\chi^2_\nu$  distribution is used to transform a randomly drawn number  $s$  into, what this research defines as:  $Chi_s$ . Hereafter the correlated, normally distributed future returns  $Z$  are transformed into  $X$  based on equation 5.7.

$$X = \frac{\sqrt{DoF}}{\sqrt{Chi_s}} * Z \quad (5.7)$$

With  $Z$  the normally distributed correlated numbers,  $DoF$  the estimated  $DoF$  for the copula and  $Chi_s$  the random number  $s$  transformed with inverse of the  $\chi^2_\nu$  distribution with  $\nu = DoF$ .

After this transformation correlated uniform student-t distributed future returns are simulated by using the CDF of the student-t distribution with  $\nu$ , the input parameter  $DoF$  and  $X$  the outcome of equation 5.7. This procedure is shown by equation 5.8:

$$U = \text{tcdf}(X, DoF) \quad (5.8)$$

where  $\text{tcdf}$  stands for the  $CDF$  of the student-t distribution,  $X$  is given in equation 5.7 and  $DoF$  is the input parameter DoF and  $U$  are the correlated student-t distributed uniform numbers.

In order to simulate actual returns, a transformation of  $U$  and  $Z$  is needed. To transform the values of  $Z$  into actual returns the mean  $\mu$  and standard deviation  $\sigma$  of the marginal distributions are used. For the transformation of  $U$  into actual returns, the estimates of the marginal distribution show in section 3 are used. In this way the simulated future returns are obtained for each specific asset class based on both methods. This immediately shows the flexibility of the copula method, since the copula method allows that for each asset class the specific marginal distribution is used in order to transform the uniform numbers  $U$  into future returns, while for the MV normal method a normal distribution is always assumed for every asset class.

### 5.3 Tests

There are quite some differences in the simulation process of future returns between the copula and MV normal method. The simulation of future returns with the copula method requires a lot of extra work and the important question is whether this results in better simulations of future asset returns. In order to test the difference between the simulations of future asset returns made with the copula and MV normal methods three requirements for future asset returns are used. These requirements are based on what one would ideally expect future asset returns to look like with respect to the historically observed returns. The three requirements are:

1. The simulation method should simulated future asset returns that are more extreme than historically observed.
2. The simulated future asset returns should have a skewness close to empirically observed.
3. The simulation method should simulate a distribution of future asset returns that replicates the distribution of the historically observed returns.

The first requirement is based on the expectation that the worst return ever might not be observed yet. When the simulation method is not able to simulate more extreme returns than empirically observed it might even be better to use historical simulation instead of this simulation method. The second requirement is based on the large negative skewness observed in the data sample, as described in section 2. The simulation method should take this skewness into account in order to simulate reliable future asset returns. Finally the simulation method should simulate a distribution that replicates the distribution of the historically observed returns, since this is a distribution observed within the market and represent the market circumstances.

In order to check whether the simulation methods meet the first requirement different distribution and histogram plots are used to show the differences in simulated future returns between the copula and MV normal method. For the second requirement it is possible to use the measure of the skewness in order to check the difference between the simulation methods, while for the third requirement different measures and tests are used to show the differences between the two simulation methods. The measures used for the third requirement are the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). These measure are used to measure the distance between the simulated and empirical returns. Based on these measure significance tests are used to conclude which simulation method meets the third requirement the best.

### 5.3.1 Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE)

Two measures to test whether simulated returns of the copula method or simulated returns of the MV normal are closer to the empirical returns is by computing Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). The RMSE is computed as shown in equation 5.9:

$$\begin{aligned} RMSE_{copula} &= \sqrt{E(Return_{copula} - Return_{empirical})^2} \\ RMSE_{MV\ normal} &= \sqrt{E(Return_{MV\ normal} - Return_{empirical})^2} \end{aligned} \quad (5.9)$$

Were  $E$  stands for the expectation of the returns of the simulation method minus the empirical returns. Given the way RMSE is computed, large deviations from the target value are heavily taken into account. When MAE is used all deviations are equally taken into account. The MAE is computed as shown by equation 5.10:

$$\begin{aligned} MAE_{copula} &= E(|Return_{copula} - Return_{empirical}|) \\ MAE_{MV\ normal} &= E(|Return_{MV\ normal} - Return_{empirical}|) \end{aligned} \quad (5.10)$$

### 5.3.2 Wilcoxon signed-rank test

The Wilcoxon signed rank test is used to test whether the differences in RMSE and MAE between the copula and MV normal method are significantly different from each other. The Wilcoxon signed-rank test has an advantage over the t-test that the Wilcoxon signed-rank test is distribution free. Another distribution free test is the Mann-Whitney-Wilcoxon test, but for that test the two used samples have to be independent. Since the same random values are used for both methods the samples are not independent, which makes the Wilcoxon signed-rank test the best test to use. The used null hypothesis of the Wilcoxon signed-rank test is that the difference in RMSE and/or MAE of the copula method and the RMSE and/or MAE of the MV normal has a median of 0. The Wilcoxon test indicates whether there is a significant difference in the median instead of the mean because the Wilcoxon test is distribution free. In order to test whether there is a significant difference in the mean a distribution has to be assumed. For more information and a more detailed description of the Wilcoxon signed-rank test see Wilcoxon [1945].

Given the null hypothesis of the Wilcoxon test it is only possible to conclude that there is a significance difference in RMSE and/or MAE. It is not possible to conclude with certainty whether it is the copula or the MV normal method that is significantly better. Therefore the Wilcoxon test is combined with the one-sided t-test to draw a certain conclusion about which simulation method is significantly better.

### 5.3.3 One-sided t-test

As a non-distribution free alternative of the Wilcoxon test it is possible to use a two-sided t-test with which one can test whether the mean of a sample is equal to mean of another sample against the alternative that the means of the samples are not the same. However, it is also possible to perform a one-sided t-test. With this test is possible to specifically test whether the mean of a sample is significantly bigger/smaller than the mean of the other sample. Therefore, the one-sided t-test is used to significantly show whether the RMSE and/or MAE of the copula method is bigger/smaller than for the MV normal method given that the Wilcoxon signed-rank test showed a significant difference.

### 5.3.4 Historical simulation

Given the data set described in section 2 it is possible to use a method called historical simulation. Historical simulation is a method which enables the user to extend the 136 observations from the data sample to 1000 observations by randomly drawing one observation out of the 136 observations and repeating this 1000 times. This means that observations simulated with historical simulation are limited to the observations in the data sample in the sense of extremes and that there are only 136 different returns in the sample of 1000 observations. This indicates the limitations of historical simulation, which is therefore not used as a possible method. However, historical simulation can be used to give an indication about the real observed returns during the in-sample period. Therefore a comparison is made between the returns observed by using historical simulation and the returns simulated by the copula and MV normal method. This comparison is based on the three requirements stated in 5.3.

## 5.4 Results

Based on the student-t copula future asset returns are simulated for each asset class, taking into account the dependency between the asset classes. Even though the main interest is in how well the simulations of a specific simulation method are compared with the empirical returns based on portfolio returns, it is interesting to look at the differences for each specific asset class. The portfolio returns can change if the weights for the specific asset classes change and thereby may result in a different outcome, while this is not possible for the specific asset classes. Therefore the returns simulated with the two simulation methods are compared with the empirical returns based on as well the separate asset classes as on an equally weighted portfolio.

### 5.4.1 Asset class simulations

As mentioned above it is interesting to see how good the simulated standardized returns for a single asset class represent the historically observed standardized returns. This gives insight in how well the standardized returns are simulated for a single asset class and whether or not the simulations of a specific simulation method meets the three requirements for all separate asset classes. In order to test the first requirement of simulating more extreme returns than historically observed 100 different paths of 1000 observations are plotted as a distribution of standardized returns for the different asset classes. Further a histogram is used to show whether the used simulation method simulate more extreme returns than historically observed and to give an indication on the distribution fit. In order to save space a distribution plot is used to show the result on the first requirement for credits and stocks, while a histogram is used to show this for real estate and commodities. The distribution plot for real estate and commodities and the histogram plot for credits and stocks can be found in the appendix.

#### **First requirement: Are the simulated returns more extreme than the empirical returns?**

In figure 5.1 100 paths of 1000 observations are plotted as a distribution of standardized returns for credits and stocks. Whereby the simulated standardized returns of the MV normal method are shown in red, the simulation from the copula method are shown in blue and the empirical standardized returns obtained by using historical simulation are shown in green.

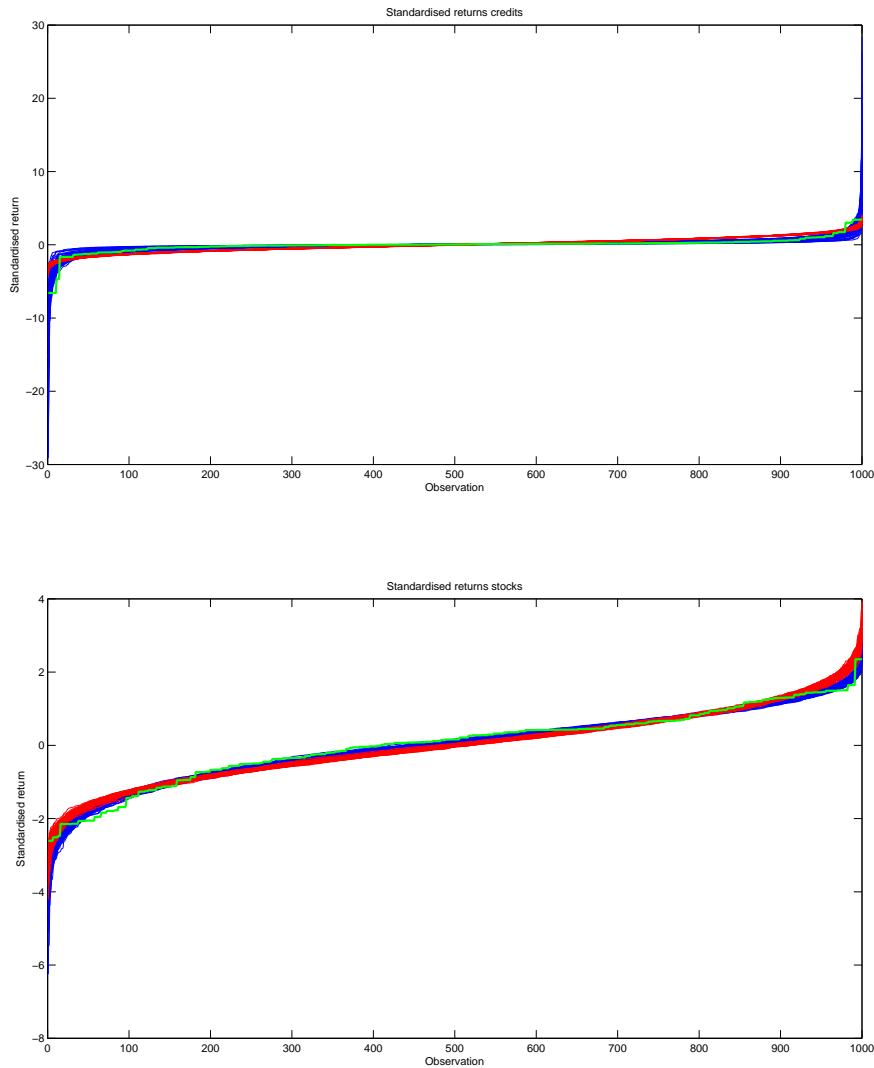


Figure 5.1: This figure shows a plot of the sorted standardized returns for the copula (red) an MV normal method (blue) together with the empirical standardized returns simulated with historical simulation (green) for credits and stocks. For the copula and MV normal method 100 paths of 1000 observations are simulated. From the figure for credits one can see that the MV normal method is not able to simulate more extreme negative returns than empirically observed, while the copula method does. Further one can see that for observation 100 till 900 the green line lies upon the blue lines, which indicates a good fit in the center of the distribution for the copula method. For stocks the MV normal method as well as the copula method simulates negative standardized returns that are more extreme than empirically observed. Further the copula method seems to fit the empirically observed standardized returns better in the center than the MV normal method.

Figure 5.1 indicates that for the asset class credits, the MV normal method is not able to simulate the empirically observed extreme negative returns observed during the crisis. The most negative value simulated by the MV normal method is equal to  $-4.298$ , while the most negative empirical standardized returns is equal to  $-6.580$ . Further the most negative standardized returns simulated by the copula method is equal to  $-29.120$ . The most negative value for the copula method might look very extreme, there is only a 0.1% chance on observing this path. Given this low probability such a extreme negative return is according expectation.

On the extreme positive side the MV normal is able to simulate standardized returns that are more positive than empirically observed, but given the negative extremes the MV normal does not meet the first requirement for the asset class credits. The copula method on the other hand simulates negative standardized returns that are far more negative than the empirical standardized returns and also on the positive side the simulated returns of the copula method are far more extreme than empirically observed. This results in the conclusion that the copula method is able to meet the first requirement for the asset class credits, while the MV normal method is not.

Based on figure 5.1 it is possible to get an indication about the fit of both simulation methods for the third requirement. For the copula method a very good fit with the empirical standardized returns in the center of the distribution (observation 100 till 900) is observed, while a large deviation is observed in the tails of the distribution. For the MV normal method a smaller error is observed in the extremes of the distribution but the center of the distribution seems not to be captured right by the MV normal method. The fit on the distribution

is discussed in more detail later in this section.

Besides a distribution plot for credits, figure 5.1 also present a distribution plot for stocks. For stocks a different conclusion is made based on the distribution plot. Where the MV normal was not able to simulate more extreme return than empirically observed on both sides of the distribution for credits, the MV normal method does meet this first requirement for stocks. Besides the MV normal method, the copula method also meets the first requirement of simulating more extreme returns for stocks. Further the indication of the fit on the whole distribution is again for the copula method since the copula method seems to have a better fit in the center of the distribution (observation 100 till 900). Although the copula method will have a larger mismeasurement especially in the left tail.

Besides plots of the simulated distribution of standardized returns a histogram is used to check whether the first requirement holds for the simulations of the copula and MV normal method for real estate and commodities. In figure 5.2 a histogram of the standardized returns of a single path for the copula and MV normal method is presented for real estate and commodities. This is a plot of a single path of 1000 simulated standardized returns for the copula (red) and MV normal method (blue) together with 1000 empirical standardized returns (green) simulated with historical simulation for real estate and commodities.

Figure 5.2 (shown on the next page) indicates that for real estate the most extreme negative standardized return of the MV normal and copula method are both more negative than empirically observed. Although the copula method has a more negative standardized return than the MV normal both methods meet the first requirement of simulating larger extremes than empirically observed for real estate. The histogram of commodities indicates a different conclusion since the MV normal method is not able to replicate the negative extremes that are empirically observed. The copula method on the other hand is able to simulate extreme negative observations that exceed the empirical ones. Therefore the conclusion is that for commodities the MV normal method does not meet the first requirement, while the copula method does.

In terms of fitting the whole distribution the histogram for real estate indicates that copula method fits the empirical standardized returns better than the MV normal method, although this is not formally tested yet. Also for commodities the copula method seems to have a better overall fit to the empirical standardized returns than the MV normal method.

Based on the four different asset classes the simulations made with the current MV normal method used in AEGON's ALM model is not able to simulate returns more extreme than empirically observed for the credits and commodities and thereby fails to meet the first requirement for these asset classes. There is a big difference with the simulations of the copula method, which are (far) more extreme than empirically observed for every asset class. The conclusion is that the copula method meets this first requirement for every asset class.



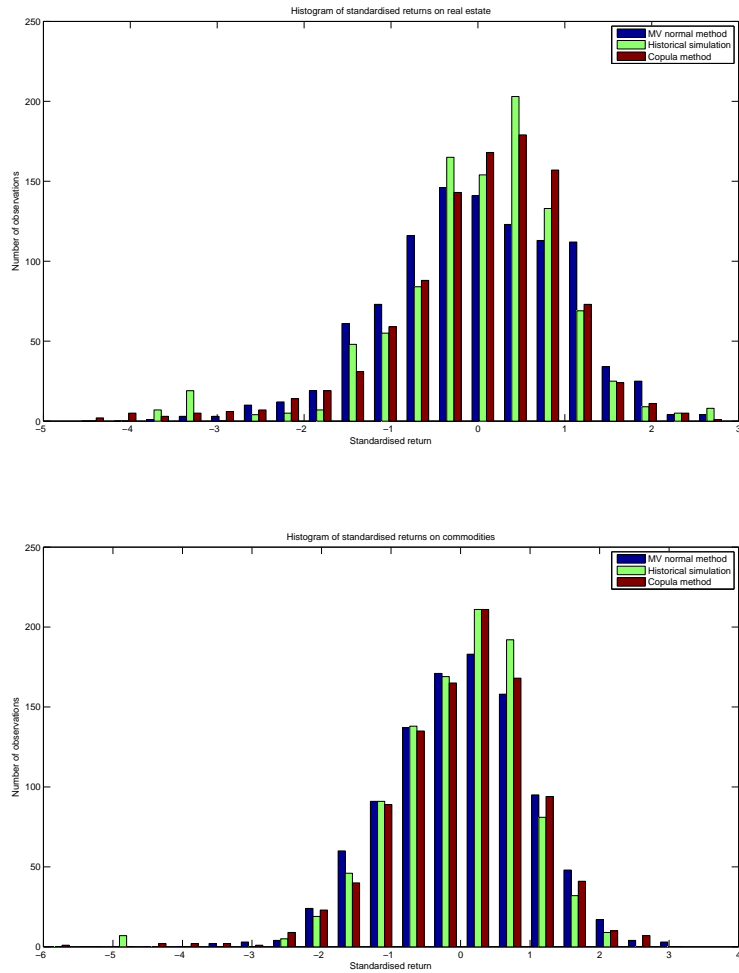


Figure 5.2: This figure shows a histogram plot of a single path of 1000 simulated standardized returns for the copula (red) and MV normal method (blue) together with 1000 empirical standardized returns (green) simulated with historical simulation for real estate and commodities. The histogram for real estate indicates that the most extreme negative standardized return of the copula method is more extreme than empirically observed which also holds for the MV normal method. However, for commodities the simulated extreme negative observations from the copula method do exceed the empirical returns, while the MV normal method does not. Based on overall fit the histograms indicate that the copula method seems to have a better fit than the MV normal method when compared with the empirical standardized returns for both real estate and commodities.

### Second requirement: Is the empirical skewness represented in the simulations?

Besides simulating more extreme returns than empirically observed it is important that the simulated returns have a skewness that meets the skewness of the empirical returns. This is important since it is shown in section 2 that a large negative skewness is present in the empirical data. When a negative skewness is present in data this results in a lot of small positive returns and some (very) extreme negative returns. Therefore it is important to capture this property in the simulation of future returns. In table 5.1 the skewness in the simulated standardized returns of the copula and MV normal method are presented together with the skewness observed in the empirical standardized returns.

Skewness standardized returns				
	Credits	Stocks	Real estate	Commodities
Copula method	-2.595	-0.577	-1.397	-0.740
MV normal method	0.000	0.004	-0.004	0.005
Historical simulation	-2.156	-0.546	-0.935	-0.869

Table 5.1: This table indicates that the skewness in the standardized simulated future returns simulated with the copula method is very close to the skewness of the empirical standardized returns for every asset class. The skewness in the standardized simulated future returns of the MV normal method is on average very close to zero, which indicates a large mismatch when compared with the empirical skewness. However, one of the properties of the normal distribution is that it has a skewness of zero, which is why this result is according expectation.

From the values shown in table 5.1 it is interesting that the simulated returns from the MV normal have a skewness very close to zero for every asset class. However, this is according to expectation since one of the properties of a normal distribution is that it has a skewness of zero. The skewness in the empirical returns range from -0.5 to -2 for the different asset classes, which indicates that there is quite a large skewness present in the data which is not captured by the normal distribution. The skewness in the simulated returns of the copula method is a bit more extreme than the empirically observed skewness for credits and real estate, while the skewness of the copula method is the same as the empirical skewness of stocks. For the skewness in commodities the skewness of the copula method is slightly smaller than empirically observed.

Based on these skewness results the conclusion is that the copula method is able to replicate the empirically observed skewness and therefore meets the second requirement of representing the empirical skewness in the simulation for every asset class. On the other hand the MV normal method is not able to account for the empirically observed skewness for any asset class and therefore does not meet the second requirement.

### Third requirement: Which simulation method fits the empirical distribution the best?

The histograms, distribution plots and the skewness results give an indication about which method meets the third requirement the best, but formal tests are applied to determine whether there is a significant difference between the two methods and which one resembles the empirical data the most.

To measure whether the copula or the MV normal method simulates standardized returns that fit the empirical standardized returns better, the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are computed as shown in section 5.3.1.

The computed values for the RMSE and the MAE for both simulation methods and for every asset class are shown in table 5.2:

Average mismeasurement per asset class for both simulation methods based on RMSE and MAE				
	RMSE copula	RMSE MV normal	MAE copula	MAE MV normal
Credits	<b>0.489</b>	0.626	<b>0.184</b>	0.451
Stocks	<b>0.140</b>	0.197	<b>0.101</b>	0.155
Real estate	<b>0.196</b>	0.251	<b>0.080</b>	0.176
Commodities	<b>0.148</b>	0.213	<b>0.077</b>	0.110

Table 5.2: This table shows the average value of RMSE and MAE for all four asset classes. RMSE and MAE are computed by comparing 100 paths of 1000 simulated standardized returns of the copula and MV normal method with 1000 standardized returns simulated with historical simulation.

From table 5.2 the conclusion is that the RMSE and the MAE of the copula method is always lower than the RMSE and MAE of the MV normal method for all asset classes. Combining this result with the distribution/histogram plots and given that the copula has a large mismeasurement in the tails of the distribution for most asset classes, the copula fits the center of the distribution very well to compensate. Given the large mismeasurement between the copula simulation and the empirical observations in the tail of the distribution, a larger value for RMSE is expected for the copula method. This is indeed the case, especially when focusing on credits one finds that the value of RMSE and MAE are larger than for the other asset classes, because of the extreme simulations. So even though the copula method is expected to have a larger mismeasurement due to meeting the first requirement, this is not the case. This indicates that the copula method fits the empirical distribution very well in the center of the distribution for every asset class. Even though the values shown in table 5.2 indicate that the copula method should be used it is first formally tested to see whether the differences in RMSE and MAE between the copula and MV normal method are significantly different from each other.

To formally test whether the differences in RMSE and MAE between the MV normal and the copula method are significant a Wilcoxon signed-rank test is performed together with two one-sided t-tests (left and right). In this way it is determined whether there is a significant difference in RMSE and MAE between the copula and MV normal method (with the Wilcoxon test). Further it is tested with the one sided t-test whether it holds that the copula method is significantly better than the MV normal or that the MV normal method is significantly better than the copula method. The results of these tests are shown in table 5.3.

P-value of significance tests for significant difference between copula and MV normal method per asset class						
	T-test right RMSE	T-test left RMSE	Wilcoxon RMSE	T-test right MAE	T-test left MAE	Wilcoxon MAE
Credits	1.000	0.000	0.000	1.000	0.000	0.000
Stocks	1.000	0.000	0.000	1.000	0.000	0.000
Real estate	1.000	0.000	0.000	1.000	0.000	0.000
Commodities	1.000	0.000	0.000	1.000	0.000	0.000

**Table 5.3:** This table shows the p-value for a t-test and a wilcoxon signed-rank test for every asset class. The tests are performed on the RMSE and the MAE of the copula method minus the RMSE and MAE of the MV normal method. Based on the p-values one can conclude that there is significant evidence to assume that the median of the RMSE and MAE of the copula is not equal to the median of the standardized returns of the MV normal method (conclusion of the wilcoxon tests). To be more specific the median of RMSE and MAE of the copula method is significantly lower than the median of the RMSE and MAE of the MV normal method (conclusion of the t-tests). This shows that the copula method simulates standardized returns that are significantly closer to the empirical ones than the simulated standardized returns of the MV normal method for every asset class.

Based on the p-values shown in table 5.3 the conclusion is that there is significant evidence to assume that the median of the RMSE and MAE of the copula is not equal to the median of the RMSE and MAE of the MV normal method for every single asset class. (conclusion of the Wilcoxon tests). Given that one of the two methods has a significantly smaller error than the other the p-values of the one-sided t-test result in the conclusion that the median of RMSE and MAE of the copula method is significantly lower than the median of the RMSE and MAE of the MV normal method for every asset class. Therefore the conclusion is that the simulated returns of the copula method are significantly closer to the empirically observed returns than the simulated returns from the MV normal method for every asset class.

Based on the three stated requirements the conclusion is that the simulations for the specific asset classes made with the copula method meets the requirements better than the simulations from the MV normal method for every single asset class. For this research the portfolio returns are most important. Based on the results for the separate asset classes the expectation is that the copula method will also meet the three requirements the best when based on the portfolio returns. However, since for portfolio returns the dependencies between the asset classes are taken into account, as long as there is more than one asset class represented in the portfolio, it is possible that the MV normal method meets the requirements better than the copula method. Therefore the results on an equally weighted portfolio are discussed in the next section.

## 5.4.2 Portfolio simulations

Although the simulation of standardized returns for the separate asset classes is a good indication on the performance of the copula and MV normal method, dependencies between the asset classes are not taken into account and therefore the portfolio simulations could yield different results than the separate asset classes. For the MV normal method the dependency in the simulations is based on the Pearson product-moment correlation, since that is the currently used correlation in the AEGON ALM model. For the copula method Kendall's tau is used to account for dependency since it is proved in section 4 that Kendall's tau generates a more accurate value for the correlation.

The portfolio returns are constructed by using an equally weighted portfolio and the portfolio returns simulated with the copula and MV normal method are compared with the empirically observed portfolio returns based on the same three requirements as for the separate asset classes, stated in section 5.3.

### **First requirement: Are the simulated returns more extreme than the empirical returns?**

The first requirement for the simulation of portfolio returns is tested by using a distribution and a histogram plot. Figure 5.3 shows 100 paths of 1000 observations for the standardized equally weighted portfolio returns simulated with the copula method (red lines), MV normal method (blue lines) and the empirical standardized portfolio returns (green line).

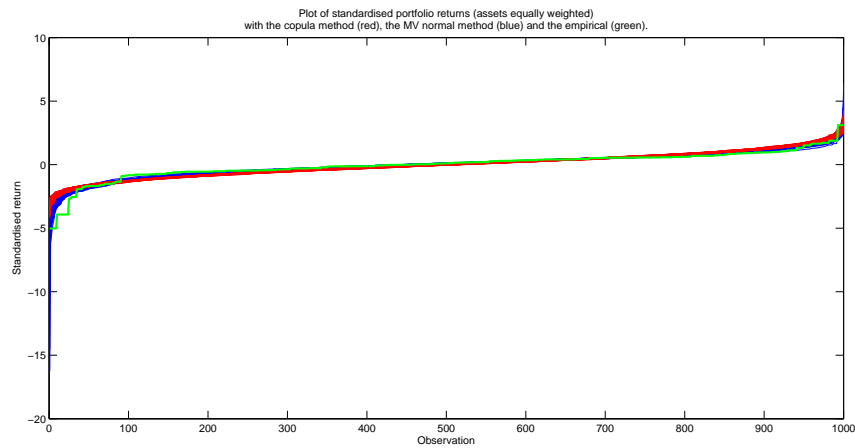


Figure 5.3: This figure shows a plot of 100 paths of 1000 simulated standardized returns with historical simulation (green line), a multivariate normal method (red lines) and a copula method (blue lines). This figure indicates that the copula method is able to simulate more extreme returns, as well as for positive and the negative returns, than historically observed. Unfortunately the simulations for portfolio returns of the MV normal method are not more extreme than historically observed in the left side of the distribution (for extreme negative returns).

Figure 5.3 indicates that for the left side of the distribution the MV normal method is unable to replicate the historically observed most extreme negative standardized portfolio return of -5.028 since the most extreme negative simulation of the MV normal method is -4.186. On the right side of the distribution the MV normal method is able to simulate a more extreme positive standardized return than the empirically observed 3.111, by simulating a value of 4.310. Nevertheless the MV normal method does not meet the first requirement of simulating more extreme returns than empirically observed based on the portfolio returns. The copula method on the other hand simulates more extreme standardized portfolio returns for both tails of the distribution than empirically observed. The most extreme negative standardized portfolio return simulated with the copula method is -16.251, which is more extreme than the -5.028 from the empirical standardized portfolio and the most positive observation is 6.346 for the copula which is also more extreme than 3.111 from the empirical distribution. Based on these results the conclusion is that the copula method meets the first requirement, while the MV normal method does not.

Further figure 5.3 indicates that the copula method has a better fit in the center of the distribution to the empirical standardized portfolio returns than the MV normal method has. In order to check this with a different graphical representation a histogram plot is shown of a single path of 1000 observations for the copula method (red), MV normal method (blue) and the empirical standardized returns (green) in figure 5.4.

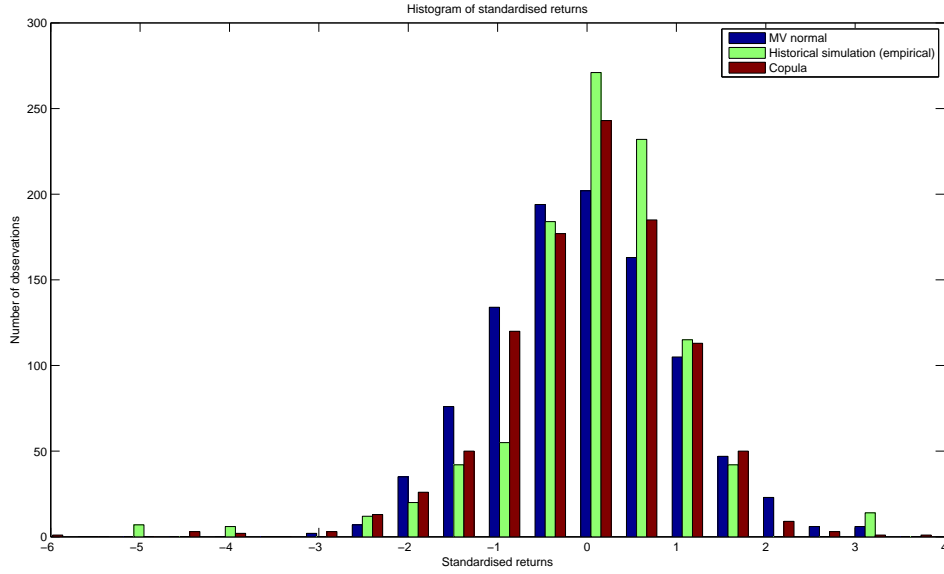


Figure 5.4: This figure shows a histogram plot of standardized portfolio returns simulated with the copula method (red) and the MV normal method (blue). Further the empirical standardized returns are plotted in the histogram (green). From the histogram the conclusion is that the MV normal method is unable to simulate more extreme portfolio returns than empirically observed, while the copula is able to simulate more extreme returns on both sides of the distribution. Further looking at the overall histogram, the standardized portfolio returns of the copula method seem closer to the empirically standardized portfolio returns than the standardized portfolio returns of the MV normal method.

From 5.4 the conclusion for the simulation of extremes is the same, since this is just a different graphical representation of the same data. Based on both graphical representations, the conclusion is that the copula method does meet the first requirement of larger extremes for the simulation of portfolio returns while the MV normal method does not meet this requirement.

**Second requirement: Is the empirical skewness represented in the simulations?**

Besides simulating more extreme standardized portfolio returns than empirically observed it is interesting to look at the skewness in the simulated portfolio returns. By checking whether skewness is present in the empirical portfolio returns and checking whether this data feature is captured by the two simulation methods one can determine whether the simulation method accounts for the empirically observed skewness. The skewness estimations on the empirical equally weighted portfolio returns and the simulated equally weighted portfolio returns with the copula and MV normal method are shown together in table 5.4.

Skewness standardized returns				
	Mean	Minimum	Maximum	Standard deviation
Copula method	-0.711	-4.859	-0.087	0.561
MV normal method	-0.004	-0.266	0.177	0.087
Historical simulation	-1.373	-2.002	-0.711	0.240

Table 5.4: Skewness is estimated based on 1000 observations. This is repeated 100 times in order to compute the maximum, minimum, mean and standard deviation of the skewness.

Table 5.4 shows that the big negative skewness that is present in the separate asset classes is also present in the empirical portfolio returns. This indicates that the simulation method should account for this empirically observed skewness. However, just like for the separate asset classes the MV normal method is not able to replicate the empirically observed skewness in the simulated portfolio returns. In line with the property of a MV normal distribution the skewness is zero. The copula method on the other hand does account for the empirically observed skewness. The copula method is also not able to fully replicate the skewness in the empirical portfolio returns, which is due to the fact that the student-t copula simulates a symmetric distribution implying a skewness of zero. The skewness in the simulated portfolio returns of the copula method stems from the skewed marginal distributions used to simulate the portfolio returns.

Based on the results shown in table 5.4 the conclusion is that the the copula method meets the second requirement of accounting for skewness in the simulation of future portfolio returns far better than the MV normal method.

### Third requirement: Which simulation method fits the empirical distribution the best?

The third requirement is on the fit to the empirical distribution. Based on figures 5.3 and 5.4 the expectation is that the copula method is better able to mimic the distribution of the empirically observed standardized portfolio returns than the MV normal method. Although, the copula method has a larger mismeasurement in the tail of the distribution. To formally measure the distance between the simulations and the empirical observations the distance measures RMSE and MAE, explained in section 5.3.1, are used. In order to compute values for the RMSE and the MAE 100 paths of 1000 future portfolio returns are simulated with the copula and MV normal methods, which are all compared with 1 path of 1000 observation of empirical portfolio returns. The results of this comparison are shown in table 5.5.

Average mismeasurement of the standardized simulated returns for both methods measured in RMSE and MAE				
	Average	Minimum	Maximum	Standard deviation
RMSE copula	<b>0.286</b>	0.224	0.461	0.033
RMSE MV normal	0.387	0.348	0.420	0.016
MAE copula	<b>0.150</b>	0.112	0.197	0.017
MAE MV normal	0.236	0.201	0.265	0.014

Table 5.5: This table shows the average, minimum and maximum value of RMSE and MAE for the copula and MV normal method. RMSE and MAE are computed by comparing 100 paths of 1000 simulated standardized returns of the copula and MV normal methods with 1000 standardized returns simulated with historical simulation.

The results shown in table 5.5 indicate that in terms of RMSE and MAE the copula method replicates the empirically observed standardized portfolio returns much better than the MV normal method. Besides a lower average for RMSE as well as MAE, the value for the maximum and minimum mismeasurement based on MAE is also lower for the copula method than for the MV normal method. The standard deviation based on the MAE measure is equal for both the MV normal and the copula methods. In terms of RMSE the minimum value is clearly lower for the copula method, while the maximum deviation is higher than for the MV normal method. The standard deviation of the RMSE of the copula method is higher than for the MV normal method. This is explained by the fact that RMSE penalizes large deviations more than MAE. Given that the copula method simulates (far) more extreme portfolio returns than empirically observed while the MV normal method does not, this explains the larger standard deviation in RMSE.

Based on the values for RMSE and MAE shown in table 5.5 the conclusion is that the copula method has a lower mismeasurement than the MV normal method. Before concluding which method should be used based on the third requirement it first should be tested whether the differences in RMSE and MAE between the copula and MV normal method are significant. For this purpose the Wilcoxon test is performed together with two times a one-sided t-test. The p-value for these test are shown in table 5.6.

P-value of significance tests for significant difference between copula and MV normal method per asset class						
	T-test right RMSE	T-test left RMSE	Wilcoxon RMSE	T-test right MAE	T-test left MAE	Wilcoxon MAE
Portfolio comparison	1.000	0.000	0.000	1.000	0.000	0.000

Table 5.6: This table shows the p-value for a t-test and a wilcoxon signed-rank test on the differences in RMSE and MAE for the simulated standardized returns on an equally weighted portfolio for the copula and MV normal methods. The tests are performed on the RMSE and the MAE of the copula method minus the RMSE and MAE of the MV normal method. Based on the p-values one can conclude that there is significant evidence to assume that the median of the RMSE and MAE of the copula method is not equal to the median of the RMSE and MAE of the MV normal method (conclusion of the wilcoxon tests). To be more specific the median of RMSE and MAE of the copula method is significantly lower than the median of the RMSE and MAE of the MV normal method (conclusion of the t-tests).

Based on the p-values of the Wilcoxon tests shown in table 5.6 the conclusion is that there is significant evidence to assume that the median of the RMSE and MAE of the copula method is not equal to the median of the RMSE and MAE of the MV normal method. More specifically, based on the outcomes of the t-tests the conclusion is that the RMSE and MAE of the copula method is significantly lower than the RMSE and MAE of the MV normal method.

Based on these results the conclusion is that the copula method simulates standardized portfolio returns that are significantly closer to the empirical portfolio returns than the simulated standardized portfolio returns of the MV normal method. Therefore the copula method meets the third requirement for the simulation of future portfolio returns significantly better than the MV normal method.

## 5.5 Conclusion

In this section it is explained why a copula is used to form a joint distribution and what the main idea behind a copula is. Based on the properties of different copulas and based on other research in the literature the student-t copula is selected as the best copula to use for the purpose of this research. After explaining the estimation and simulation procedure for the student-t copula a comparison is made between simulations made with the copula method and the MV normal method. This comparison is based on comparing the simulations of the new copula method and the simulations of the currently used MV normal method based on the same random values with the empirically observed returns.

The comparison is based on three requirements that are ideally observed in the simulation of future returns. The simulation method should simulate returns that are more extreme than historically observed. It should take the empirically large negative skewness observed in section 2 into account and the simulation method should ideally simulate future returns that represent the historically observed return distribution. This is tested for the asset class return simulations separately as well as for equally weighted portfolio returns.

The first comparison is based on the simulation of future asset returns for the asset classes separately. For the simulations of returns for the separate asset classes the MV normal method does not meet the first requirement for credits and commodities, while the copula method meets the first requirement for all the asset classes. In terms of skewness, the skewness in the simulated returns of the MV normal method is zero. Empirically a large negative skewness is observed in the returns. The copula method is able to reproduce this empirical skewness in the simulation of returns which means that the copula method also meets the second requirement better than the MV normal method for the separate asset class return simulations. Also for the third requirement it is shown that the RMSE and MAE of the copula method are significantly better for the copula method than for the MV normal method. This leads to the conclusion that for the individual asset classes it is (significantly) better to use the copula method according to all three requirements.

The second comparison is based on the simulation of future portfolio returns. Although a good simulation of the asset class returns separately is desired, it is important to look at the portfolio returns, where the dependency between the asset class returns is taken into account. The simulated returns for an equally weighted portfolio made with both simulation method is tested based on the same three requirements as for the separate asset class returns. From the comparison it is shown that for the standardized portfolio returns the copula method is able to replicate more extreme returns than empirically observed while the MV normal is not able to meet this first requirement. In terms of replicating the empirically observed skewness, the MV normal method has a skewness of zero in the simulated portfolio returns. The use of a student-t copula indicated symmetric simulations which would also result in a skewness of zero. However, given the flexibility in the marginal distributions of the copula, the use of the optimal skewed marginal distributions results in a negative skewness in the simulated portfolio returns. This skewness represents the skewness observed in the empirical portfolio return a lot better than the MV normal distribution. Therefore the copula method also meets the second requirement better than the MV normal method for the portfolio returns. Finally the fit of the simulated portfolio returns of the copula and MV normal methods to the empirical standardized returns were tested. Based on the results the conclusion is that the copula method has a significantly better fit, based on RMSE and MAE, to the empirical distribution than the MV normal method. Which results in the conclusion that the copula method is (significantly) better to use for the simulation of future portfolio returns than the MV normal method based on all three requirements.

Given that the copula method outperforms the MV normal method based on all three requirements for the separate asset classes as well as for the portfolio returns, the conclusion is that the copula method be used within the ALM model. In this way the ALM model is extended by accounting for crisis observations in the simulation of future asset returns. The ALM model with the copula method for simulating future returns is therefore called the Asset Liability Management Including Crisis for Asset Returns (ALMICAR) model.

# 6 Asset liability framework

An Asset Liability Management (ALM) model is used by pension funds and pension fund advisers to simulate future scenarios for yield and inflation curves and asset returns. The ALM model gives the users insight into the trade-off between risk and return to assess their investment decisions. AEGON currently uses an ALM model in which the simulation of future yield and inflation curves are simulated apart from the future asset return simulations. In AEGON's ALM model a new method for the simulation of future yield and inflation curves is recently implemented based on Vlaar [2006]. The simulation of future asset returns is still based on the multivariate normal distribution. However, practitioners observed during the financial crisis of 2008 that the risk of asset returns is not properly modeled by the multivariate normal distribution. This led to bigger losses on the asset returns than expected based on the ALM model. Together with a very low (uncovered) yield this led to huge drops in coverage ratio for almost all Dutch pension funds, see for example ANP [2008, 2009a,b], Redactie [2008]. Since long term investment decisions of pension funds are partially based on the outcomes of the ALM model it is important that realistic future asset returns are simulated. In section 5 it is concluded that the copula method simulates more realistic future asset returns. Given this knowledge it is interesting to check what the impact of using the copula method in the AEGON ALM model is on the asset mix advice.

## 6.1 Asset liability management (ALM) model

The AEGON ALM model is based on a combination of *MS excell*<sup>®</sup>, *VBA*<sup>®</sup> and *matlab*<sup>®</sup> programming. A new method is developed for the simulation of future asset returns given that the current ALM model simulates future yield and inflation curves separately from the future asset returns. Since only the simulation of the asset returns is changed, it is important to only modify that part of the simulations within the current AEGON ALM model. To compare the results of the currently used multivariate normal and the new copula method, new input variables are added in the model. After modifications in the VBA programming and the integration of the new copula method it is possible to choose between the two different simulation methods.

Since the conclusion in section 5 is that the copula method simulates more realistic future asset returns it is interesting to check what the impact of using the copula method is on the outcomes of the AEGON ALM model. Since coverage ratio is very important for pension funds the outcomes of the AEGON ALM model are possible coverage ratio paths. Since the mean and standard deviation are external parameters, the difference in simulations between the copula and multivariate normal method is based on differences in skewness and kurtosis. A comparison of the coverage ratio paths would therefore not lead to different outcomes. In order to check the differences between the simulations method the change in coverage ratio is used instead. The change in coverage ratio is defined as:

$$\Delta Coverage\ ratio = \ln(Coverage\ ratio_t - Coverage\ ratio_{t-1}) \quad (6.1)$$

By taking the natural logarithm after subtracting the coverage ratio one time period ago from the coverage ratio observed today, the change in coverage ratio is constructed.

Since the same random numbers are used as input for the copula and the MV normal method, the differences in changes in coverage ratio are entirely due to the way the used simulation method transforms these random numbers into future asset returns and allows for an objective comparison between the two simulation methods.

## 6.2 Value-at-Risk (VaR) and Expected Tail Loss (ETL)

Value-at-Risk (VaR) and Expected Tail Loss (ETL) are two commonly used risk measures. See for example Holton [2003], McNeil et al. [2005] and Jorion [2007] for VaR and Tasche [2002] and Rockafellara and Uryasev [2002] for ETL. Risk measures are used to give insight into the risk of a certain portfolio. The use of VaR and ETL specifically gives insight into the maximum loss that can occur in a certain portfolio over a certain time period with a certain possibility. The 'certains' in the last sentence are the input parameters on which the VaR and ETL are computed. Given the time period  $h$ , the possibility  $\alpha$  and a portfolio the value for VaR and ETL are computed.

VaR and ETL are two different risk measures that are complementary to each other. VaR is used to compute a value that represents the maximum loss for which it is  $1 - \alpha\%$  sure that it is not exceeded in the coming time period  $h$ . Within this research VaR is used to give an indication on the maximum drop in coverage ratio given a certain confidence level  $\alpha$ . The value for VaR is computed by sorting all the observed changes in coverage ratio. Based on the sorted changes in coverage ratio the  $\alpha\%$  VaR is equal to:



$$VaR_\alpha = \text{Sorted } \Delta\text{coverage ratio}(\alpha\% \times N) \quad (6.2)$$

Where  $N$  is the number of simulated changes in coverage ratio and  $\alpha$  the chosen risk level. So based on 1000 simulated changes in coverage ratio, the 5% VaR is equal to the 50th sorted observation. Given that the observations are sorted from the most negative change in coverage ratio to the most positive change in coverage ratio. In this way the 5% VaR indicates that there is a 5% chance that the value shown by VaR is violated. A problem with VaR as a risk measure is that it does show a value for which it is  $1 - \alpha\%$  sure not to be violated, but it does not take into account how large the violation is if VaR is violated.

In order to account for the size of the exceedances if VaR is violated, the risk measure ETL is used. The value computed for ETL represents the expectation of the losses given that VaR is violated. ETL is defined as the expectation of the changes in coverage ratio given that VaR is violated, as shown in equation 6.3:

$$ETL_\alpha = E(\Delta\text{coverage ratio} | \Delta\text{coverage ratio} < VaR_\alpha) \quad (6.3)$$

When ETL is computed based on the sorted changes in coverage ratio, the ETL is computed by taking the expectation over the first  $\alpha\% \times N$  observations. Based on 1000 simulated changes in coverage ratio, the 5% ETL is equal to the expectation over the first 50 observations. In this way ETL also takes the largest simulated drops in coverage into account.

Given that decision making from an ALM model is not only dependent on the correct simulation of returns but also on the used risk measure, VaR and ETL are both taken into account. The risk measures are used to check the differences in the simulations of drops in coverage ratio based on the copula and MV normal method. With these risk measures it is possible to select a certain value that represent the risk of a big drop in coverage ratio. With the values shown by these risk measures it is possible to check whether there is a difference in the risk of a large drop in coverage ratio when the future asset returns are simulated with the copula method instead of the MV normal method.

## 6.3 Tests

Although the tables and figure in section 6.4.1 give an impression on the differences in change in coverage ratio between the copula and MV normal method, it is formally tested whether these differences are significant or not. By using both VaR and ETL it is possible to test whether the risk based on the simulated changes in coverage ratio are significantly different from each other or not. This is tested by using the Wilcoxon test, discussed in 5.3.2. Further stylized portfolios are used to test whether a different stylized portfolio is selected when the copula method is used instead of the MV normal method. The selection of a stylized portfolio that meets the requirement is based on the risk measures discussed in 6.2. With these two tests it is tested whether the differences in changes in coverage ratio between the copula and MV normal method are significant and what the impact of the simulation method is on the selection of the asset-mix by pension fund.

### 6.3.1 Stylized portfolio

Pension funds usually use ALM models to select their strategic portfolio from a range of accepted portfolios. In order to test whether the copula simulation method leads to the selection of a different range of accepted portfolios, stylized portfolios are used. The stylized portfolios are five different portfolios with a different asset-mix. The asset-mix of the five stylized portfolios is chosen such that they have a decreasing risk profile. This results in stylized portfolio one being the most risky portfolio and stylized portfolio five the most risk-averse portfolio. The asset-mix of the five stylized portfolios is shown in table 6.1:

Part invested in a specific asset class for a stylized portfolio					
Portfolio \ asset class	bonds	credits	stocks	real estate	commodities
1	0.000	0.250	0.250	0.250	0.250
2	0.100	0.225	0.225	0.225	0.225
3	0.200	0.200	0.200	0.200	0.200
4	0.300	0.175	0.175	0.175	0.175
5	0.400	0.150	0.150	0.150	0.150

Table 6.1: This table shows five stylized portfolios. Where portfolio one is the riskiest investment and portfolio five the most safe investment.

Table 6.1 shows that the weight between the asset classes credits, stocks, real estate and commodities are kept constant for all portfolios, while the investment in bonds increases. The decreasing risk profile of the stylized portfolios is because an investment in bonds is seen as the safest investment. Based on all five stylized portfolios the changes in coverage ratio are computed. Given the changes in coverage ratio a selection is made on which

stylized portfolio meets some predefined requirements. The requirement is that the maximum drop in coverage ratio is below a pre-defined level based on an  $\alpha\%$  confidence level. The maximum drop in coverage ratio is hereby computed by using either the VaR or the ETL.

The requirement for the drop in coverage ratio to be below a pre-defined level is because pension funds do not want to observe a decrease in coverage ratio equal to the decrease during the 2008 financial crisis. Besides a limitation of the possible risk pension funds also want to acquire good returns. Therefore the problem is defined as finding the most risky stylized portfolio that meets the requirement that the drops in coverage ratio are smaller than some pre-defined level. This is illustrated by the following equation:

$$\begin{aligned} & \max \text{ return} \\ & \text{s.t. risk} < \text{pre - defined level} \end{aligned} \tag{6.4}$$

In this set-up it is possible to check whether there are differences in the selected stylized portfolio for the simulations based on the copula and MV normal method. As equation 6.4 shows, the constraint is that the pension fund wants to avoid drops in coverage ratio that are too large. The pre-defined level of maximum drop is set to a drop in coverage of 15%, 20% or 25% in any of the coming 15 years. The maximum drop is measured by using the measures discussed in section 6.2 based on the 2.5% and 5% confidence level.

The 2.5% and 5% confidence levels are chosen since these are used by pension funds. The widespread use of the 95% and 97.5% VaR by pension funds is because De Nederlandsche Bank (DNB) prescribes that pension funds should have portfolios that are underfunded (coverage ratio lower than 100%) only once in every 40 years. This results in the set-up in which different risk appetites for maximum drop in coverage ratio are taken into account based on the commonly used risk levels (2.5% and 5%).

## 6.4 Results

### 6.4.1 Differences in outcomes of the ALM model

Based on the result that the copula method simulates more realistic returns, it is interesting to check what the impact is of using the copula method within the ALM model on the outcomes of the ALM model.

In section 2 it is discussed that two different sub samples are used in order to test the out-of-sample performance of both methods. A problem for the out-of-sample testing is that there are no crisis observations present in sub-sample one ranging from 01-01-1999 till 01-01-2007. Since there are no crisis observations present in this data set the copula method is also not able to simulate extreme negative observations. As a consequence the copula method simulates almost the same future changes in coverage ratio as the multivariate normal method. The simulated changes in coverage ratio based on an equally weighted portfolio without bonds are shown in figure 6.1:

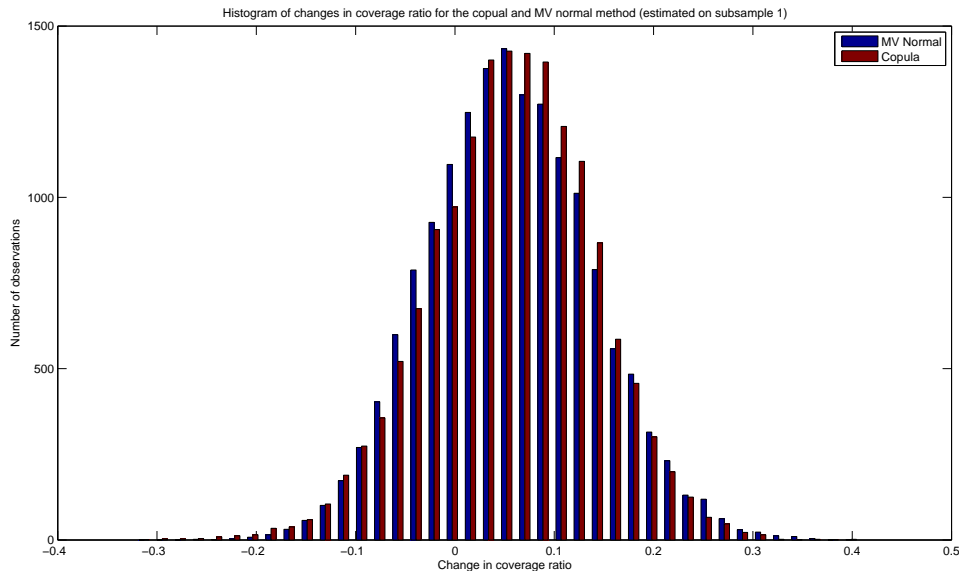


Figure 6.1: This figure shows a histogram of the changes in coverage ratio for the copula and MV normal method. From the figure one can conclude that there is not a big difference between the copula and MV normal method. This is due to the fact that there are no crashes present in the data used for parameter estimation. Using a Jarque-Bera test for all 16 years with 1000 observations to test whether or not the assumption of a normal distribution can be rejected results in 4 rejections for the copula method and no rejections for the MV normal method based on a 1% significance interval. This shows that the changes in coverage ratio simulated with the copula and MV normal methods differ only slightly.

Figure 6.1 shows a histogram of the changes in coverage ratio for the copula and MV normal method based on parameters estimated on sub-sample one. The figure indicates that there is only a slight difference between the copula and MV normal method. This results from sub-sample one lacking crisis observations for the parameter estimation.

To test whether there is a difference in the simulated distributions the Jarque-Bera is used. The Jarque-Bera test is used to test over all sixteen years with 1000 observations whether or not the assumption of a normal distribution can be rejected. This results in four rejections for the copula method and no rejections for the MV normal method based on a 1% significance interval. This result indicates that the changes in coverage ratio simulated with the copula are significantly different from a normal distribution in 25% of the times. Therefore the conclusion is that based on sub-sample one no large differences are observed between the simulated future changes in coverage ratio made with the copula and MV normal method.

When the estimations are based on the whole data sample, ranging from 01-01-1999 till 01-04-2010, including crisis observations this difference in simulated changes in coverage ratio becomes much larger for the copula and MV normal method. Given that the mean and standard deviation are the same for both methods since they are external input parameters in the Aegon ALM model, the differences in simulations between the copula and MV normal method are obtained by simulating a different shape of the future distribution. These differences in future simulations are realized by improvements on the simulated skewness and kurtosis of the distribution. In figures 6.2, 6.3 and 6.4 plots and histograms are used to show the distribution of changes in coverage ratio. Given that section 5 indicated that the copula method simulates negative skewness and excess kurtosis it is interesting to check whether this is also observed in the simulated future changes in coverage ratio.

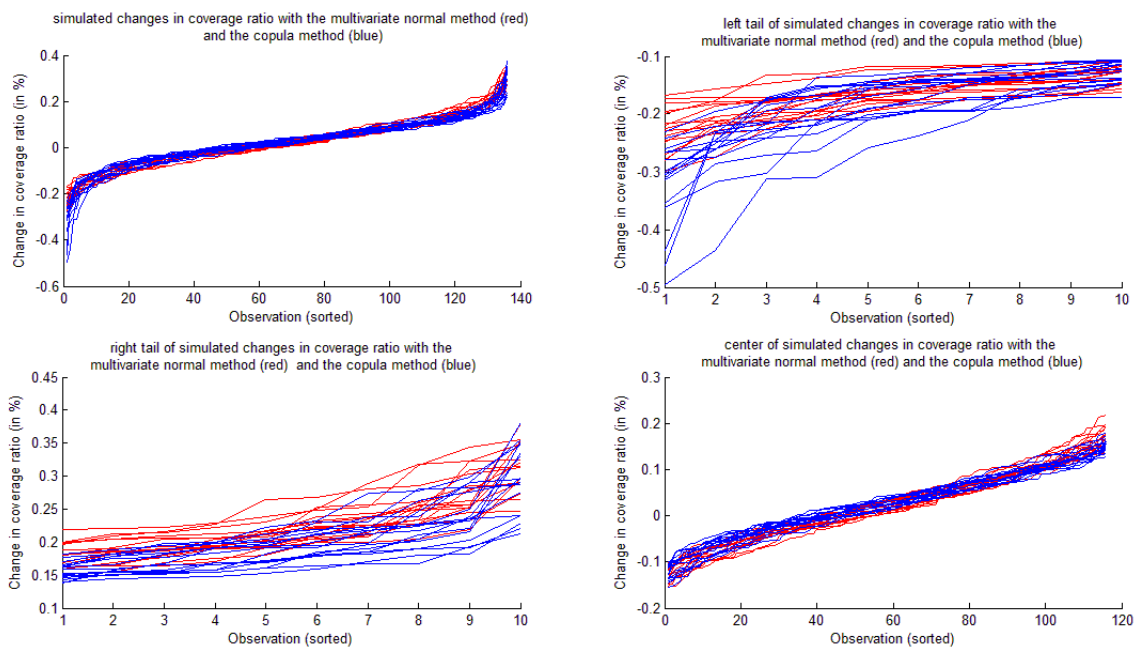


Figure 6.2: This figure shows plots for 16 years of 136 sorted observations of changes in coverage ratio simulated with the copula (blue) and MV normal method (red). In the top left corner is a plot of all observations, while the top right corner shows the left tail (10 most negative observations). The bottom left corner shows the right tail (10 most positive observations) and the bottom right plot shows the center (observation 11 till 125). From the plot in the top right corner one can see that the drops in coverage ratio simulated with the copula method are more negative than with the MV normal method. The copula method simulates drops in coverage ratio up to -50%, while the drops in coverage ratio with the MV normal method are not bigger than -30%.

Figure 6.2 shows sixteen years of 136 observations of changes in coverage ratio. There are four figures, whereby the left top corner shows the whole distribution, the right top corner shows the left tail (ten most negative observations) of the distribution, the left bottom figure shows the right tail (ten most positive observations) of the distribution and the right bottom figure shows the center (observation eleven till 126) of the distribution. These figures indicate that the copula method (blue) simulates larger negative observations (left tail figure) than the MV normal method (red). The simulations for changes in coverage ratio for the center and the right tail of the distribution are close to each other. The conclusion is that the copula method is able to simulate larger drops in coverage ratio than the MV normal method, while there are only slight differences in other parts of the distribution. Another way to show the differences in a distribution is by using a histogram. A histogram of sixteen years with 1000 observations in each year is shown in figure 6.3:

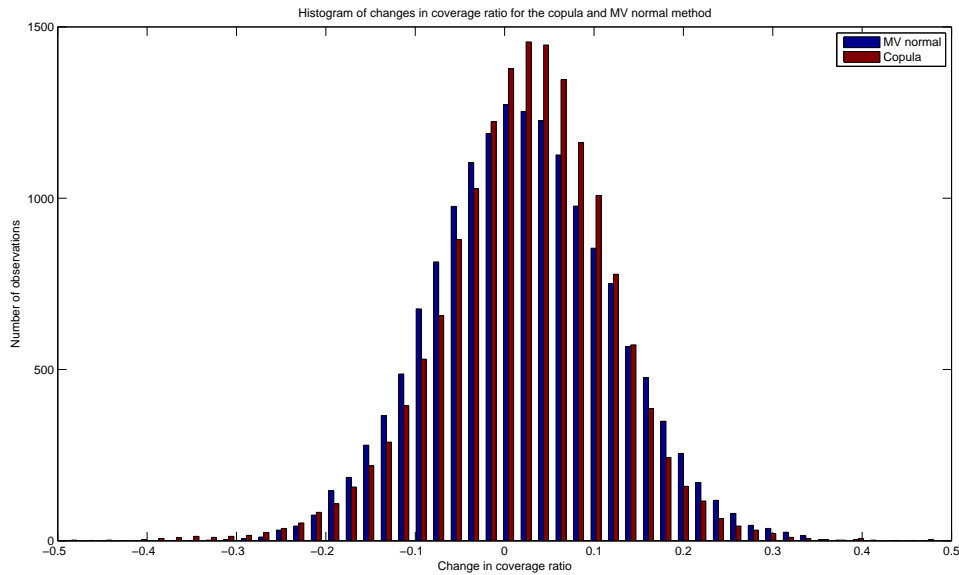


Figure 6.3: Here a histogram of changes in coverage ratio is plotted for the copula method (red) and the MV normal method (blue). The peak of the red histogram lies around 0.05, while the peak of the blue histogram lies around 0. This shows that there is skewness present in the red histogram, while the skewness of the blue histogram is zero (which is according expectation, since a MV normal method implies a skewness of zero). Further the copula method has a fatter left tail when compared with the MV normal method, while the right tail is almost the same. In contrast to figure 6.1 the null hypothesis of a normal distribution of changes in coverage ratio with the copula method is significantly rejected for all 16 years on the 1% significance interval, while for the MV normal method the null hypothesis of a normal distribution is only significantly rejected for one year on the 1% significance interval. This shows that the simulated distribution of changes in coverage ratio with the copula method is significantly different from a normal distribution, while the simulated distribution of the MV normal method is not.

Figure 6.3 illustrates the differences in the distribution of changes in coverage ratio between the copula and MV normal method. In order to test whether the negative skewness and excess kurtosis are present in the distribution of changes in coverage ratio the Jarque-Bera test, discussed in section 2, is used. The conclusion is that for fifteen of the sixteen years the distribution of changes in coverage ratio is significantly different from the normal distribution when the copula method is used, while for the MV normal distribution it is observed in only one of the sixteen years that the distribution is significantly different from the normal distribution on a 1% significance interval. The conclusion is that the simulated distributions of changes in coverage ratio with the copula method are significantly different from a normal distribution while this does not hold for the MV normal method.

Figure 6.3 indicates that the median of simulated future changes in coverage ratio for the copula method is more positive than for the MV normal method. The copula method simulates more low positive changes in coverage ratio in the region from 0 to 0.15 than the MV normal method. For the negative changes in coverage ratio ranging from 0 till -0.2 the MV normal method simulates these more than the copula method. This is observed because the copula method has to compensate somewhere in the distribution in order to simulate more extreme negative observations larger than -0.2 than the MV normal method. For the positive changes in coverage ratio larger than 0.15 the MV normal method simulates more observations than the copula method, but for observations larger than 0.35 the copula method simulates more observations than the MV normal method again, which is probably due to the higher kurtosis of the copula method. Although the whole distribution needs to be simulated well, the tail risk is of extra interest for the pension funds given the 2008 financial crisis. Figure 6.4 shows the left tail of figure 6.3 in more detail.

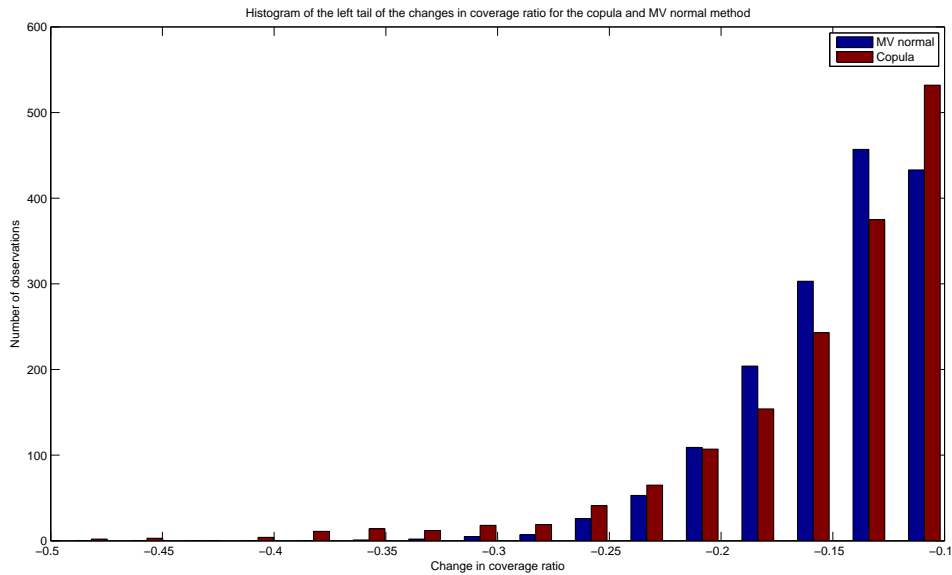


Figure 6.4: This histogram shows the left tail of the histogram in figure 6.3. This histogram shows that the copula method (red) has far more extreme drops in coverage ratio than the MV normal method (blue).

Figure 6.4 indicates that the copula method simulates larger negative changes in coverage ratio than the MV normal method. Where the most extreme negative for the MV normal method is -0.363, the copula method simulates drops in coverage up to -0.481. Further a change in coverage ratio of -0.3 is a rare event based on the MV normal method, while it is more commonly observed based on the copula method. Together with the conclusion of section 5 that the copula replicates the empirical returns better this can result in an underestimation of the risk of large drops in coverage ratio by the MV normal method underestimates.

#### 6.4.2 Differences in outcomes of the ALM model based on VaR and ETL

Given the results based on the shape of the distribution of changes in coverage ratio the expectation is that when the copula method is used to simulate future asset return, larger drops in coverage ratio are simulated. Based on figure 6.3 the expectation is that a higher VaR and ETL is observed when the copula method is used instead of the MV normal method.

Given that the copula method simulates more realistic future asset returns, the question is what the impact is on the risk of the portfolio. In tables 6.2 and 6.3 the difference in values based on VaR and ETL are shown for the five stylized portfolios of section 6.3.1 for the copula and MV normal method.

Copula (Kendall equal weights)						MV Normal (Pearson equal weights)					
VaR\year	1	2	3	4	5	VaR\year	1	2	3	4	5
0.50%	-25.33%	-36.13%	-29.34%	-24.76%	-28.40%	0.50%	-20.83%	-24.35%	-23.76%	-20.77%	-22.97%
1.00%	-23.55%	-25.08%	-24.18%	-20.88%	-23.02%	1.00%	-17.79%	-20.89%	-22.55%	-19.73%	-20.15%
2.50%	-17.94%	-18.92%	-18.49%	-16.83%	-19.08%	2.50%	-16.22%	-18.57%	-18.03%	-16.29%	-18.13%
5.00%	-14.01%	-14.36%	-14.40%	-12.99%	-13.74%	5.00%	-12.96%	-14.15%	-14.38%	-13.61%	-14.33%
10.00%	-8.93%	-9.81%	-9.91%	-9.52%	-10.34%	10.00%	-9.08%	-10.59%	-11.23%	-10.07%	-11.23%

Table 6.2: This table shows the value of the VaR of changes in coverage ratio for the multivariate normal method as well as the copula method, for different confidence levels. The results in the table show that the copula method has a more negative value for the 0.5% and 1% interval, while for the 10% interval the MV normal method has a lower value. This suggests that the 2.5% and especially the 5% observation have almost an equal value for the copula method as well as for the MV normal method.

Table 6.2 shows the differences in VaR on different confidence levels by using either the copula method or the MV normal method. The results in table 6.2 indicate a larger risk when simulations are based on the copula for the 0.5%, 1% and 2.5% VaR. Based on the 5% VaR the difference in risk between simulations made with the copula method do not look to differ from simulation made with the MV normal method. Next to differences based on VaR as risk measure, the differences based on ETL are also taken into account. Table 6.3 shows the differences in ETL on different confidence levels between the copula and MV normal method.

Copula (Kendall equal weights)						MV Normal (Pearson equal weights)					
ETL\year	1	2	3	4	5	ETL\year	1	2	3	4	5
0.50%	-27.72%	-38.50%	-35.40%	-30.95%	-30.45%	0.50%	-22.89%	-27.66%	-25.62%	-22.47%	-26.17%
1.00%	-25.83%	-34.44%	-30.51%	-26.56%	-27.46%	1.00%	-20.77%	-24.96%	-24.28%	-21.24%	-23.49%
2.50%	-22.49%	-26.46%	-24.70%	-21.61%	-23.55%	2.50%	-18.42%	-21.79%	-21.71%	-19.19%	-20.76%
5.00%	-18.97%	-21.14%	-20.52%	-18.12%	-19.64%	5.00%	-16.23%	-18.97%	-18.85%	-16.99%	-18.49%
10.00%	-14.96%	-16.44%	-16.29%	-14.54%	-15.76%	10.00%	-13.60%	-15.67%	-15.74%	-14.28%	-15.57%

Table 6.3: This table shows the ETL of changes in coverage ratio for the multivariate normal method as well as the copula method. The results in the table show that the copula method has a more negative expected change in coverage ratio than the multivariate normal method. This suggests that when the copula method is used to simulate future returns that the drops in coverage ratio are larger than when the multivariate normal method is used. Given that the multivariate normal method was unable to capture the drops in coverage ratio during the crisis, the fact that the copula method has a larger expected negative drop in coverage ratio is an improvement.

Table 6.3 indicates that using ETL instead of VaR results in different values for the risk. Where the value for the copula method was clearly more negative on the 0.5% and 1% VaR, the value of the copula method with ETL is more negative on all confidence levels ranging from 0.5% till 10% ETL. This indicates that the largest drops in coverage ratio simulated with the copula method are more extreme than when the MV normal method is used.

Differences are observed based on the values for VaR and ETL. In the next section it is formally tested whether the computed risk based on VaR and ETL is significantly higher or not if the copula method is used instead of the MV normal method.

### 6.4.3 Results Wilcoxon signed-rank test

For the five stylized portfolios, discussed in section 6.3.1m the Wilcoxon signed-rank test is used to test whether the drops in coverage ratio differ significantly in the copula method and MV normal method. The results based on various risk levels for VaR and ETL are shown in table 6.4:

P-value of the Wilcoxon test for stylized portfolios for VaR						P-value of the Wilcoxon test for stylized portfolios for ETL					
VaR\Portfolio	1	2	3	4	5	ETL\Portfolio	1	2	3	4	5
0.5%	0.000	0.000	0.000	0.000	0.000	0.5%	0.000	0.000	0.000	0.000	0.000
1.0%	0.000	0.000	0.000	0.000	0.000	1.0%	0.000	0.000	0.000	0.000	0.000
2.5%	0.000	0.000	0.000	0.000	0.000	2.5%	0.000	0.000	0.000	0.000	0.000
5.0%	0.489	1.000	0.890	0.934	0.978	5.0%	0.000	0.000	0.000	0.000	0.000
10.0%	0.000	0.000	0.000	0.000	0.000	10.0%	0.000	0.000	0.000	0.000	0.000

Table 6.4: This table shows that only on the 5% VaR the null hypothesis of equal medians for the copula and MV normal method is not rejected by the Wilcoxon signed-rank test even on a 10% confidence interval. However, for the other VaR levels the null hypothesis of equal means is rejected, even on a 1% confidence interval. This means that there is significant evidence to reject the null hypothesis of equal medians.

The results in table 6.4 indicate that based on the 5% VaR the null hypothesis of an equal median for the changes in coverage ratio simulated with the copula and MV normal method can not be rejected. This indicates there is no significant difference in the drops in coverage ratio simulated with the copula and MV normal method based on 5% VaR. For the 0.5%, 1%, 2.5% and 10% VaR there is significant evidence to reject the null hypothesis of equal medians for the drops in coverage ratio. Combining the results in the left part of table 6.4 with the results in tables 8.7, 8.8, 8.9, 8.10 and 8.11 (in the appendix) the conclusion is that the risk based on the 0.5%, 1.0% and the 2.5% VaR is significantly bigger for the copula method than for the MV normal method. On the 10% VaR level the conclusion is that the value is significantly more negative for the MV normal method.

Based on ETL the conclusion is that on every risk level the median of the copula method is significantly different from the median of the MV normal method. Combining the results in the right part of table 6.4 with tables 8.12, 8.13, 8.14, 8.15 and 8.16 (in the appendix) the conclusion is that for every risk level of ETL the median of the copula method is significantly more negative than the median of the MV normal.

This indicates that when the copula method is used to simulate future asset returns the estimated risk with VaR and ETL is significantly larger than when the MV normal method is used to simulate the future asset returns for every risk level but the 5% and 10% VaR. Based on the 5% VaR there is no significant difference in the estimated risk between future asset return simulated with the copula or MV normal method.

### 6.4.4 Stylized portfolio results

Based on the results in the left part of table 6.2 the expectation is that when the 2.5% and 5% VaR is used as risk measure the selected stylized portfolio will not differ when the copula method is used instead of the MV normal method. However, based on the results in the right part of table 6.3 the expectation is that when ETL is used as a risk measure the copula method will select different stylized portfolios than the MV normal method given all pre-defined confidence levels. The selected stylized portfolios for the different maximum drop constraints based on the 2.5% and 5% VaR and ETL are shown in table 6.5.

Advised stylized portfolio based on 2.5% VaR			Advised stylized portfolio based on 5.0% VaR		
Constraint\Method	MV normal	Copula	Constraint\Method	MV normal	Copula
<i>Drop</i> < -15%	3	4	<i>Drop</i> < -15%	1	2
<i>Drop</i> < -20%	1	1	<i>Drop</i> < -20%	1	1
<i>Drop</i> < -25%	1	1	<i>Drop</i> < -25%	1	1

Advised stylized portfolio based on 2.5% ETL			Advised stylized portfolio based on 5.0% ETL		
Constraint\Method	MV normal	Copula	Constraint\Method	MV normal	Copula
<i>Drop</i> < -15%	5	None*	<i>Drop</i> < -15%	4	5
<i>Drop</i> < -20%	2	4	<i>Drop</i> < -20%	1	2
<i>Drop</i> < -25%	1	2	<i>Drop</i> < -25%	1	1

Table 6.5: This table shows the selected stylized portfolio for the copula and MV normal method for different maximum drops in coverage ratio on the 2.5% and 5% VaR and ETL. The \* means that none of the stylized portfolio meets the requirement of a drop in coverage ratio being less than x%. The numbers of the stylized portfolios are the stylized portfolios of table 6.1. Based on this table one can see that using VaR leads to selecting a safer portfolio with the copula method instead of the MV normal method only when the drop in coverage ratio has to be less than 15%. However, when ETL is used the copula method selects a safer portfolio all the time on the 2.5% ETL. Also if the drop has to be lower than 15% and 20% with 5% ETL a safer stylized portfolio is selected with the copula method than with the MV normal method. It is interesting that using VaR instead of ETL leads to selecting riskier portfolios and the difference in portfolio selection between the copula and MV normal method is very small when based on VaR. However, when ETL is used safer stylized portfolios are selected and the copula method selects clearly more defensive stylized portfolios than the MV normal method. On other confidence levels the same results are found, these tables can be found in the appendix.

Table 6.5 indicates that if VaR is used as risk measure the copula method selects safer stylized portfolios than the MV normal method only when the maximum allowed drop is equal to 15% for the coming 15 years. There is no difference in selected stylized portfolios when the copula is used instead of the MV normal method based on the constraints of a maximum drop in coverage ratio of 20% and 25% .

When portfolio selection is based on ETL, more differences between the copula and MV normal method are observed. Based on the 2.5% ETL the conclusion is that for all levels of maximum drop a safer portfolio is selected when the copula method is used for the asset return simulations. Also when the 5% ETL is used the result is that using the copula method to simulate future asset returns leads to selecting a more defensive portfolio if the maximum allowed drop is equal to 15% and 20%.

The results in table 6.5 indicate that using VaR as a risk measure leads to selecting riskier portfolios than using ETL. This is explained by the way VaR and ETL are computed. Given that pension funds also want to take the largest simulated drop in coverage into account, it is better to use ETL as risk measure instead of VaR as shown in section 6.2. Based on the 0.5% and 1% VaR and ETL the differences in selected stylized portfolios based on simulations with the copula and MV normal method become bigger. For these confidence levels a far more defensive portfolio is selected when the future asset return are simulated with the copula method instead of the MV normal method based on both VaR and ETL. Further using ETL as risk measure on the 0.5% and 1% leads to selecting far more defensive portfolios than when the 0.5% and 1% VaR is used. Given that the 0.5% and 1% levels are currently not taken into account by pension funds, the results for these levels can be found in table 8.6 in the appendix.

## 6.5 Conclusion

Based on the Wilcoxon signed-rank test it is shown when returns are simulated with the copula method, the computed risk is significantly higher for all values of ETL ranging from 0.5% till 10% . The use of the copula method also leads to a significantly higher risk based on 0.5%, 1% and 2.5% VaR measures. For the 5% VaR there is no significant difference in risk for future returns simulated with the MV normal method and the copula method, while based on the 10% VaR the returns simulated with MV normal method lead to a significantly higher risk than with the copula method. Currently 2.5% VaR or 5% VaR is often used. Given that pension funds want to take the worst drop in coverage ratio into account in their decision taking, the use of ETL should

be preferred over the use of VaR. Given that section 5 shows that the copula method represents the historically observed portfolio returns better than the MV normal method the conclusion is that using the MV normal distribution as simulation method for future asset returns leads to an underestimation of the risk.

The results on the stylized portfolio selection lead to the conclusion that using the copula method together with ETL leads to selecting more defensive portfolio than with the MV normal method with VaR given a preset acceptance in maximum drop in coverage ratio. In section 5 it is shown that the copula method simulates more realistic future returns. The results shown in this section conclude that pension funds currently select portfolio that bear (far) more risk than wanted because the portfolio is selected based on a too optimistic simulation of future asset returns. This leads to the conclusion that using the copula method instead of the MV normal method for the simulation of future asset returns enables pension funds to select a portfolio based on a more realistic risk profile. Further the currently used VaR as risk measure does not take the most extreme simulated drops in coverage into account which leads to selecting a portfolio that bears even more risk than expected by the pension fund. By using ETL instead of VaR in combination with using the copula method to simulate future asset returns instead of the MV normal method pension funds have a more realistic insight into the risks of a selected portfolio. In this way it is possible for pension funds to avoid the huge drops in coverage ratio observed during the 2008 financial crisis or at least ensure that the advice based on the ALM model simulations enables pension funds to make well informed decisions on which portfolio is selected and whether risks should be covered or not.



## 7 Conclusion

The 2008 financial crisis indicated the limitations of the Asset Liability Management models used to advise pension funds. The huge losses on the invested portfolios of pension funds were so unlikely according to the ALM model that they were deemed impossible. Combined with a very low interest rate huge drops in the coverage ratio of pension funds were realized, resulting in public disturbance.

In order to allow pension funds to make better informed decisions whereby the risk of investments are properly taken into account, the simulation of future asset returns has to be improved. Simulating realistic future returns is important for pension funds since they base the weight invested in the different asset classes on these simulations. In this research the Asset Liability Management Including Crisis for Assets Returns (ALMICAR) model is proposed to account for a future crisis period within the simulated investments. This method uses a copula function instead of a multivariate normal distribution within the current Asset Liability Management (ALM) framework.

The copula allows one to use different distributions as a marginal distribution for the different asset classes. In this way it models the different asset classes by any distribution. In this research the fit with the empirically observed returns is tested for four different distributions. Where after a distribution is chosen to use as marginal distribution based on two different Goodness-of-Fit (GoF) tests. Based on these GoF tests the result is that the skewed student-t distribution represents the asset classes credits, real estate and commodities the best while the skewed normal distribution is best to use for stocks.

To form a joint distribution marginal distributions are needed together with an estimate of the dependency between the asset classes. Dependencies measured by the Pearson product-moment correlation, Kendall's tau and Spearman's rho were analyzed. Via simulations Kendall's tau proved to be the best fitting dependency measure.

Now that the optimal marginal distributions and the best dependency measure are known a joint distribution is formed. Given that there are different distributions used as marginal distribution the choice is to use a copula. Based on the properties of different copulas and earlier research the student-t copula is selected. By using the student-t copula within the ALM model a the ALMICAR model is obtained for the simulation of future returns.

In this research it is found that the conclusion is that the copula method performs better than the multivariate normal method for the simulation of future asset returns. The copula method simulates more realistic future asset returns than the multivariate normal method for individual asset as well as for portfolio returns based on every requirement. Given that the copula method simulates more realistic future returns it is interesting to check the impact on AEGON's advice. By checking the impact of using the copula method within the ALM model based on VaR and ETL as risk measures it is found that ETL represents the risk of a certain investment better.

The use of the copula method within the ALM model together with using ETL as risk measure results in the ALMICAR model. By using the ALMICAR model instead of the ALM model it is found that the computed risk of a certain investment is more realistic. Given that the realistic risk of the ALMICAR model is higher than for the ALM model that underestimates the risks, safer portfolio's are advised for the investments of pension funds. This allows pension funds to prevent huge drops in coverage ratio observed during the 2008 financial crisis during the next crisis.

A possible direction for further research is to check the performance of a skewed student-t copula against the performance of the student-t copula. The skewed student-t copula should allow for more skewness in the simulations and therefore might improve the fit with the empirical observations. Another possible direction is by including extra asset classes in the ALMICAR model or to check the performance of the model based on US data. Given that US data has a longer data history it is possible to have better parameter estimations. Based on a larger data sample it is interesting to check the difference between the ALMICAR model and the regime-switching method of Van den Goorbergh et al. [2009] in the simulation of future asset returns.

# 8 Appendix

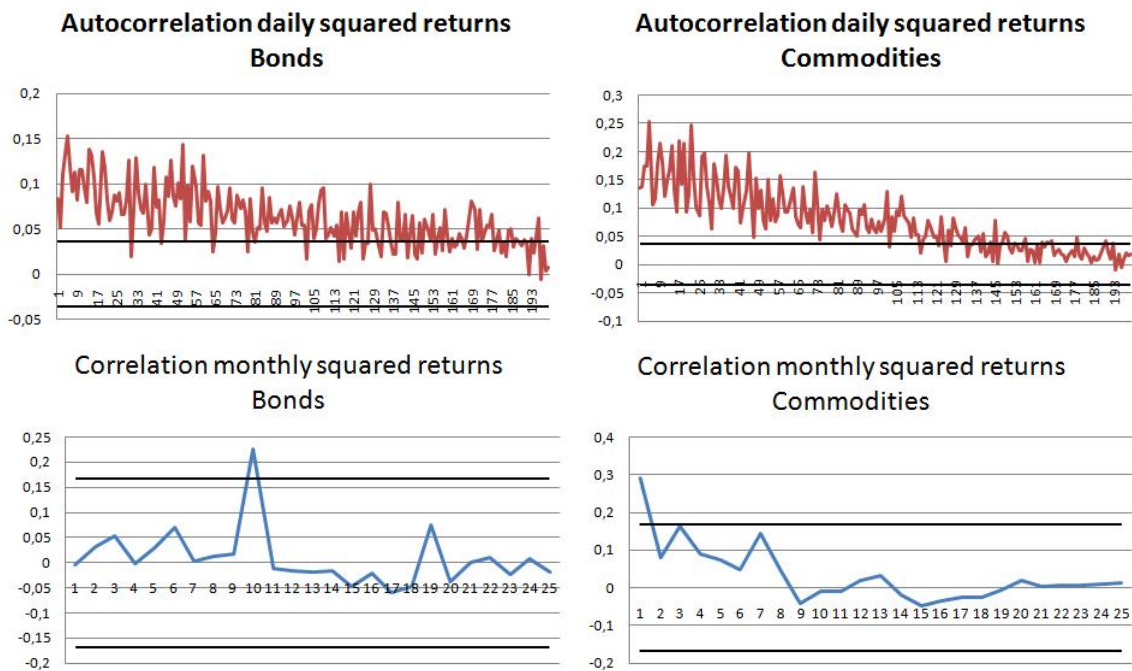


Figure 8.1: This figure shows the autocorrelation in the squared returns for bonds and commodities for daily and monthly returns. The black lines represent the 95% confidence levels. The figure indicates that the daily squared returns have a significant autocorrelation up to lag 110 (for commodities), while for the monthly squared returns the autocorrelation is significant only for the first lag.

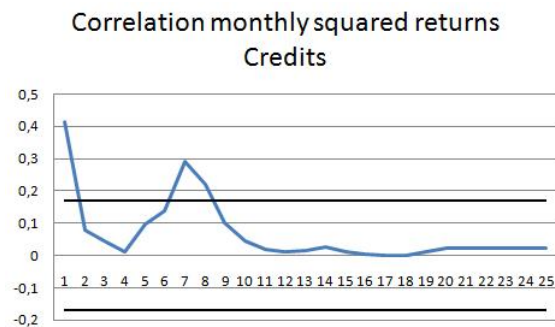


Figure 8.2: This figure indicates that there are three significant autocorrelation lags for the monthly squared returns of credits. This is in accordance with the other asset classes. Further the black lines represent the 95% confidence levels.

Spearman							Kendall						
	$\nu = 3$		$\nu = 20$		Normal			$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-3.138*	(0.104)	-2.563*	(0.098)	-2.648*	(0.099)	$\rho_{1,2}$	-2.181	(0.112)	-1.840	(0.103)	-1.706	(0.104)
$\rho_{1,3}$	3.982*	(0.104)	3.022*	(0.095)	3.353*	(0.093)	$\rho_{1,3}$	2.493 <sup>^</sup>	(0.111)	2.873*	(0.100)	2.275 <sup>^</sup>	(0.097)
$\rho_{1,4}$	4.096*	(0.100)	3.221*	(0.095)	3.087*	(0.096)	$\rho_{1,4}$	2.734 <sup>^</sup>	(0.107)	2.299 <sup>^</sup>	(0.099)	2.160 <sup>^</sup>	(0.100)
$\rho_{2,3}$	1.719	(0.108)	1.444	(0.098)	1.733	(0.100)	$\rho_{2,3}$	1.124	(0.117)	1.260	(0.103)	0.804	(0.105)
$\rho_{2,4}$	3.121*	(0.107)	2.902*	(0.097)	2.130 <sup>^</sup>	(0.095)	$\rho_{2,4}$	1.831	(0.116)	1.946	(0.102)	1.895	(0.099)
$\rho_{3,4}$	-0.461	(0.109)	0.006	(0.103)	-0.685	(0.100)	$\rho_{3,4}$	-0.138	(0.119)	-0.100	(0.109)	-0.035	(0.105)

Pearson						
	$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-6.945*	(0.177)	-6.103*	(0.100)	-5.924*	(0.100)
$\rho_{1,3}$	9.303*	(0.165)	9.181*	(0.099)	8.580*	(0.096)
$\rho_{1,4}$	9.742*	(0.162)	9.223*	(0.098)	9.280*	(0.096)
$\rho_{2,3}$	2.519	(0.193)	2.510 <sup>^</sup>	(0.102)	2.060 <sup>^</sup>	(0.099)
$\rho_{2,4}$	4.887 <sup>^</sup>	(0.187)	4.723*	(0.100)	4.759*	(0.096)
$\rho_{3,4}$	-0.123	(0.185)	-0.197	(0.105)	-0.293	(0.099)

Table 8.1: Average mismeasurement based on correlation matrix  $R_2$  with extreme observations. The standard deviation is shown between brackets. The correlations are estimated based on 100 observations. This is repeated 100 times in order to compute an average mismeasurement and standard deviation. The shown average mismeasurement is multiplied by 100 for presentation purposes (in the computations the real values are used). The \* stands for a rejection of the null hypothesis on the 1% significance interval and the <sup>^</sup> for a rejection on the 5% significance interval. The table indicates that for the Pearson product-moment correlation measure most estimated correlations are significantly different from the real correlation. Kendall's tau has the smallest average mismeasurement, with as an advantage that for all estimated correlations the null hypothesis is not rejected on the 1% significance interval. For Spearman's rho there are a lot of estimated correlation that are significantly different from the real correlation. Although, Spearman's rho has less rejections on the 5% significance interval than the Pearson product-moment correlation measure. Based on this table the Kendall's tau correlation measure is best to use.

Spearman							Kendall						
	$\nu = 3$		$\nu = 20$		Normal			$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-6.848*	(0.094)	-5.570*	(0.084)	-5.443*	(0.084)	$\rho_{1,2}$	-3.724*	(0.097)	-3.828*	(0.085)	-3.594*	(0.086)
$\rho_{1,3}$	-7.068*	(0.094)	-5.791*	(0.085)	-5.519*	(0.083)	$\rho_{1,3}$	-3.550*	(0.096)	-3.597*	(0.086)	-3.525*	(0.084)
$\rho_{1,4}$	-7.197*	(0.091)	-5.544*	(0.086)	-5.586*	(0.083)	$\rho_{1,4}$	-3.712*	(0.094)	-3.814*	(0.088)	-3.561*	(0.084)
$\rho_{2,3}$	-6.853*	(0.092)	-5.619*	(0.084)	-5.325*	(0.082)	$\rho_{2,3}$	-3.617*	(0.095)	-3.773*	(0.086)	-3.719*	(0.083)
$\rho_{2,4}$	-6.884*	(0.093)	-5.620*	(0.087)	-5.315*	(0.086)	$\rho_{2,4}$	-3.626*	(0.096)	-3.690*	(0.088)	-3.580*	(0.087)
$\rho_{3,4}$	-6.904*	(0.094)	-5.667*	(0.086)	-5.441*	(0.081)	$\rho_{3,4}$	-3.685*	(0.097)	-3.847*	(0.087)	-3.394*	(0.082)

Pearson						
	$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-13.625*	(0.149)	-13.250*	(0.088)	-13.044*	(0.089)
$\rho_{1,3}$	-14.589*	(0.140)	-14.337*	(0.091)	-14.106*	(0.089)
$\rho_{1,4}$	-16.033*	(0.142)	-15.651*	(0.092)	-15.267*	(0.090)
$\rho_{2,3}$	-7.910*	(0.154)	-8.029*	(0.086)	-8.014*	(0.083)
$\rho_{2,4}$	-10.032*	(0.155)	-9.288*	(0.088)	-9.285*	(0.087)
$\rho_{3,4}$	-10.916*	(0.154)	-10.536*	(0.089)	-10.327*	(0.085)

Table 8.2: Average mismeasurement based on correlation matrix  $R_3$  with extreme observations. The standard deviation is shown between brackets. The correlations are estimated based on 100 observations. This is repeated 100 times in order to compute an average mismeasurement and standard deviation. The shown average mismeasurement is multiplied by 100 for presentation purposes (in the computations the real values are used). The \* stands for a rejection of the null hypothesis on the 1% significance interval. The table shows that for all correlation measures for every estimated correlation the null hypothesis of no difference between the estimated correlation and the real correlation is rejected on the 1% significance interval. This indicates that all correlation measures have trouble to estimate the right correlation when some very extreme observation are added. However, based on the average mismeasurement Kendall's tau is still preferred over the Pearson product-moment correlation and Spearman's rho.

Spearman							Kendall						
	$\nu = 3$		$\nu = 20$		Normal			$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-1.251	(0.102)	-0.288	(0.096)	-0.561	(0.095)	$\rho_{1,2}$	-0.202	(0.111)	-0.223	(0.101)	-0.048	(0.100)
$\rho_{1,3}$	1.423	(0.098)	1.061	(0.096)	0.614	(0.089)	$\rho_{1,3}$	0.449	(0.105)	0.486	(0.100)	-0.068	(0.093)
$\rho_{1,4}$	2.213 <sup>^</sup>	(0.104)	1.025	(0.095)	0.683	(0.093)	$\rho_{1,4}$	-0.036	(0.111)	0.105	(0.099)	0.192	(0.097)
$\rho_{2,3}$	0.742	(0.108)	0.063	(0.102)	0.263	(0.099)	$\rho_{2,3}$	0.251	(0.118)	0.387	(0.108)	0.163	(0.104)
$\rho_{2,4}$	1.647	(0.106)	0.247	(0.096)	0.638	(0.095)	$\rho_{2,4}$	-0.014	(0.114)	-0.155	(0.101)	0.019	(0.099)
$\rho_{3,4}$	-0.231	(0.109)	-0.093	(0.104)	0.212	(0.099)	$\rho_{3,4}$	-0.099	(0.119)	-0.378	(0.111)	-0.023	(0.105)

Pearson						
	$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-0.405	(0.202)	-0.255	(0.099)	-0.020	(0.095)
$\rho_{1,3}$	0.446	(0.188)	0.482	(0.096)	-0.104	(0.088)
$\rho_{1,4}$	0.130	(0.199)	0.062	(0.096)	0.241	(0.093)
$\rho_{2,3}$	-0.284	(0.201)	0.575	(0.105)	0.072	(0.099)
$\rho_{2,4}$	0.371	(0.197)	-0.200	(0.098)	0.053	(0.094)
$\rho_{3,4}$	0.032	(0.200)	-0.381	(0.108)	0.036	(0.100)

Table 8.3: Average mismeasurement based on correlation matrix  $R_2$  without extreme observations. The standard deviation is shown between brackets. The correlations are estimated based on 100 observations. This is repeated 100 times in order to compute an average mismeasurement and standard deviation. Further, the average mismeasurement is multiplied by 100 for presentation purposes. The <sup>^</sup> behind a value shown for Spearman's rho indicates a rejection of the null hypothesis on the 5% significance interval. This means that Spearman's rho has one estimated correlation that is significantly different from the real correlation on the 5% significance interval. For all the other correlations of the Spearman's rho and all the other correlation estimates of Kendall's tau and the Pearson product-moment correlation the null hypothesis of no difference between the estimates and the actual correlation is not rejected.

Spearman							Kendall						
	$\nu = 3$		$\nu = 20$		Normal			$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-3.440 <sup>^</sup>	(0.091)	-2.140 <sup>^</sup>	(0.086)	-1.997 <sup>^</sup>	(0.078)	$\rho_{1,2}$	-0.776	(0.093)	-0.314	(0.087)	-0.229	(0.079)
$\rho_{1,3}$	-3.576 <sup>^</sup>	(0.093)	-2.219 <sup>^</sup>	(0.083)	-1.805 <sup>^</sup>	(0.082)	$\rho_{1,3}$	-0.502	(0.095)	-0.164	(0.083)	-0.153	(0.082)
$\rho_{1,4}$	-3.481 <sup>^</sup>	(0.091)	-2.163 <sup>^</sup>	(0.083)	-2.017 <sup>^</sup>	(0.080)	$\rho_{1,4}$	-0.519	(0.093)	-0.226	(0.084)	-0.265	(0.081)
$\rho_{2,3}$	-3.557 <sup>^</sup>	(0.090)	-2.330 <sup>^</sup>	(0.082)	-2.121 <sup>^</sup>	(0.082)	$\rho_{2,3}$	-0.290	(0.092)	-0.349	(0.083)	-0.176	(0.083)
$\rho_{2,4}$	-3.616 <sup>^</sup>	(0.092)	-2.073 <sup>^</sup>	(0.084)	-2.319 <sup>^</sup>	(0.082)	$\rho_{2,4}$	-0.333	(0.094)	-0.474	(0.085)	-0.442	(0.083)
$\rho_{3,4}$	-3.495 <sup>^</sup>	(0.091)	-2.245 <sup>^</sup>	(0.084)	-1.939 <sup>^</sup>	(0.082)	$\rho_{3,4}$	-0.479	(0.093)	-0.281	(0.085)	-0.423	(0.082)

Pearson						
	$\nu = 3$		$\nu = 20$		Normal	
$\rho_{1,2}$	-1.099	(0.172)	-0.402	(0.084)	-0.174	(0.074)
$\rho_{1,3}$	-1.427	(0.177)	-0.237	(0.081)	-0.265	(0.077)
$\rho_{1,4}$	-1.391	(0.176)	-0.237	(0.081)	-0.294	(0.075)
$\rho_{2,3}$	-0.190	(0.164)	-0.276	(0.080)	-0.173	(0.078)
$\rho_{2,4}$	-0.566	(0.167)	-0.468	(0.083)	-0.409	(0.078)
$\rho_{3,4}$	-0.936	(0.176)	-0.250	(0.080)	-0.367	(0.077)

Table 8.4: Average mismeasurement based on correlation matrix  $R_3$  without extreme observations. . The standard deviation is shown between brackets. The correlations are estimated based on 100 observations. This is repeated 100 times in order to compute an average mismeasurement and standard deviation. Further, the average mismeasurement is multiplied by 100. The <sup>^</sup> stands for a rejection of the null hypothesis on the 5% significance interval. The table indicates that for the Spearman's rho all the estimated correlations differ significantly from the real correlation on the 5% significance interval. All correlation estimates made with Kendall's tau and the Pearson product-moment correlation measure are not significantly different from the real correlations. Based on this table the Pearson product-moment correlation measure and Kendall's tau are both good estimators, although the standard deviation of the Pearson product-moment correlation measure becomes bigger when the tails of the distribution get fatter.

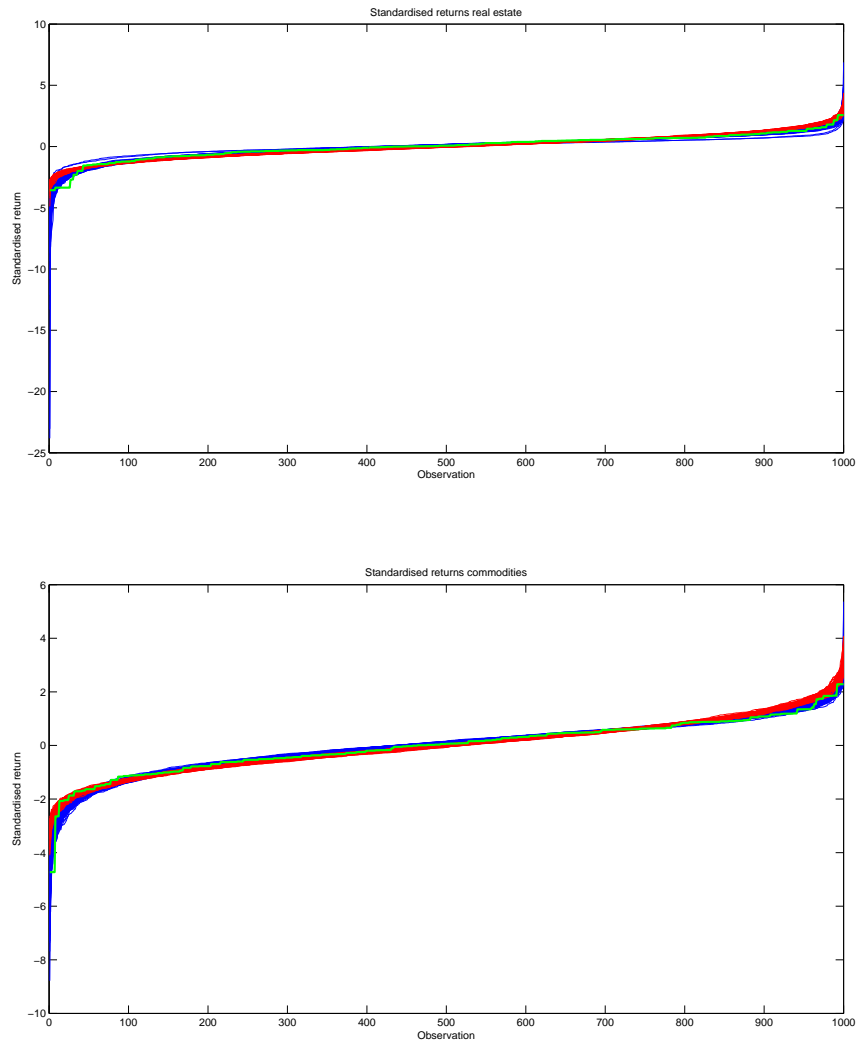


Figure 8.3: This figure shows a plot of the sorted standardized returns for the copula (red) an MV normal method (blue) together with the empirical standardized returns simulated with historical simulation (green) for real estate and commodities. For the copula and MV normal method 100 paths of 1000 observations are simulated. From the figure for real estate one can see that the MV normal method simulates negative standardized returns up to minus 5, while the empirical standardized returns using historical simulation can achieve value up to minus 3. The copula method simulates negative standardized returns up to minus 25 and thereby is far more extreme than empirically observed. However just as for credits the returns on real estate were also very extreme during the crisis, which explains the extreme simulations of standardized returns. Further one can see that for the center of the distribution (observation 100 till 900) the copula method fit the empirical standardized returns better than the MV normal just as for credits and stocks. For commodities the MV normal method is not able to simulate standardized returns that are as negative as empirically observed, just like for credits. The copula method on the other hand does simulate negative standardized returns that are more negative than empirically observed. In the center of the distribution (observation 100 till 900) the copula and MV normal method both seem to have a good fit with the empirical standardized return, although the copula method seems to have a better fit.

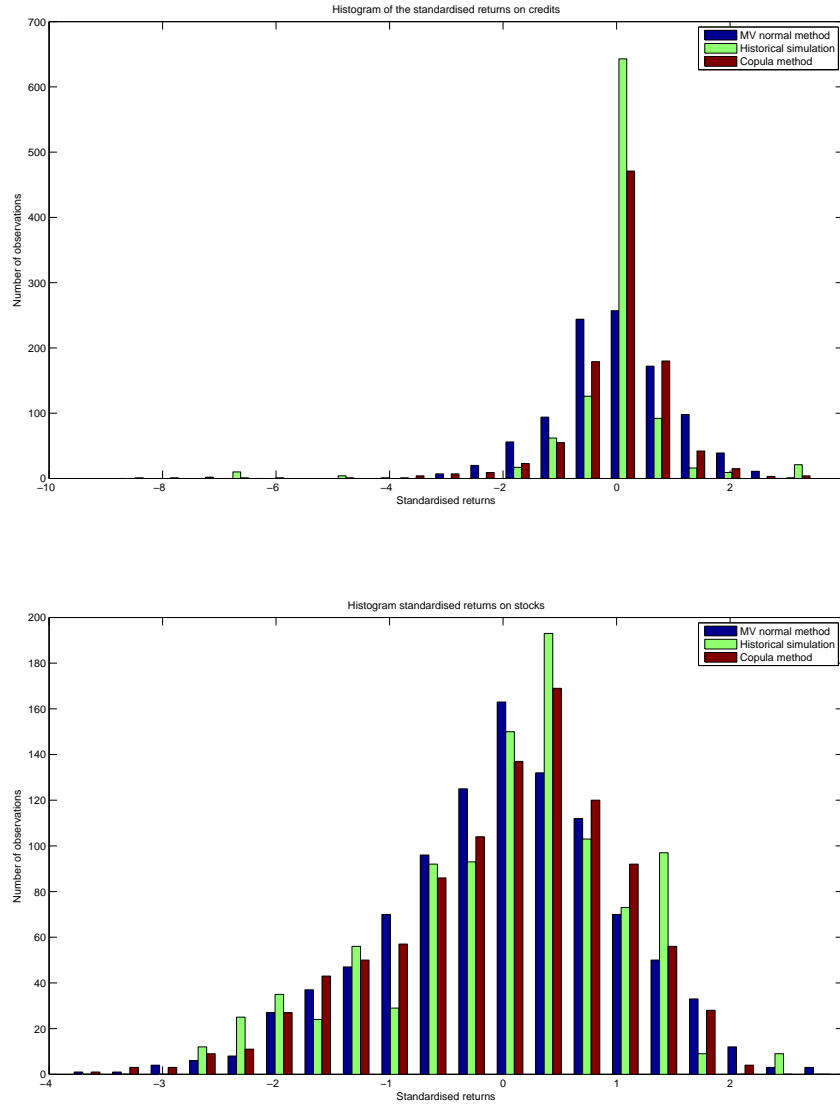


Figure 8.4: This figure shows a histogram plot of a single path of 1000 simulated standardized returns for the copula (red) and MV normal method (blue) together with 1000 empirical standardized returns (green) simulated with historical simulation. From the histogram for credits one can see that the most extreme negative standardized return of the copula method is now around minus nine, while the empirically observed is around minus seven and for the MV normal method is around minus four. Further the copula seems to have a better fit than the MV normal method. For stocks the extreme observations are according expectation when compared with the empirical ones. Overall the copula method ,again, seems to have a better fit than the MV normal method.

Comparison estimation results		
parameter\method	This research estimations	Paper Kole et al. [2007] research estimations
$\rho_{s,b}$	-0.179	-0.195
$\rho_{s,r}$	0.481	0.471
$\rho_{b,r}$	-0.081	-0.074
$\nu$	12.0	12.1

Table 8.5: This table shows the estimation results based on the same data sample as used in Kole et al. [2007]. The result indicates that due to a slight difference in the correlation estimates, a slight difference is observed in the estimated value for  $\nu$ . Given these results it is shown that the estimation procedure is correct.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 3 (20% bonds)															
VaR\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-0.93%	-5.85%	-3.36%	-1.22%	-4.76%	-6.60%	-4.35%	-0.97%	-7.06%	-5.11%	-6.65%	-2.90%	-4.13%	-1.24%	-2.27%
1.0%	-1.03%	-1.72%	-0.86%	-0.38%	-1.00%	-2.97%	-3.91%	-1.49%	-4.34%	-3.34%	-3.53%	-1.57%	-4.13%	-1.00%	-0.92%
2.5%	-1.26%	-1.43%	-1.16%	-0.32%	-0.56%	-1.11%	-0.94%	-0.12%	-1.00%	-1.11%	-1.16%	-1.65%	-2.66%	-0.31%	-0.58%
5.0%	0.16%	-1.09%	0.29%	0.20%	0.12%	-0.70%	-0.19%	0.09%	-0.22%	0.85%	0.07%	-0.29%	-0.38%	0.82%	0.99%
10.0%	0.57%	0.20%	0.36%	0.51%	0.80%	0.81%	0.47%	0.56%	0.99%	1.07%	0.68%	0.71%	0.26%	1.26%	1.22%

Table 8.9: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Advised stylized portfolio based on 0.5% VaR			Advised stylized portfolio based on 1.0% VaR		
Constraint\Method	MV normal	Copula	Constraint\Method	MV normal	Copula
$Drop < -15\%$	None*	None*	$Drop < -15\%$	None*	None*
$Drop < -20\%$	4	5	$Drop < -20\%$	2	4
$Drop < -25\%$	2	3	$Drop < -25\%$	1	2

Advised stylized portfolio based on 0.5% ETL			Advised stylized portfolio based on 1.0% ETL		
Constraint\Method	MV normal	Copula	Constraint\Method	MV normal	Copula
$Drop < -15\%$	None*	None*	$Drop < -15\%$	None*	None*
$Drop < -20\%$	None*	None*	$Drop < -20\%$	5	None*
$Drop < -25\%$	3	5	$Drop < -25\%$	2	4

Table 8.6: This table shows the selected stylized portfolio for the copula and MV normal method for different maximum drops in coverage ratio on the 0.5% and 1% VaR and ETL. The \* means that none of the stylized portfolio meets the requirement of a drop in coverage ratio being less than x%. The numbers of the stylized portfolio's are the stylized portfolio's of table 6.1. Based on this table one can see that using VaR leads to selecting no portfolio with the copula method and the MV normal method for drops less than 15% for 0.5% and 1% VaR, for drops less than 20% and 25% a safer portfolio is selected with the copula method than with the MV normal method. When ETL drops lower than 15% and 20% lead to none of the stylized portfolio's being selected with as well the copula method as with the MV normal method (except for 1% ETL with the MV normal method). It is interesting that using VaR instead of ETL leads to selecting riskier portfolio. When ETL is used safer stylized portfolio's are selected and the copula method select clearly more defensive stylized portfolio's than the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 1 (0% bonds)															
VaR\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-1.24%	-8.06%	-4.42%	-1.45%	-5.73%	-9.18%	-6.49%	-2.20%	-9.83%	-6.47%	-8.08%	-3.89%	-5.56%	-2.03%	-2.20%
1.0%	-1.35%	-1.87%	-2.00%	-0.36%	-1.85%	-3.19%	-4.81%	-1.75%	-6.04%	-3.72%	-4.63%	-2.21%	-4.63%	-1.33%	-1.74%
2.5%	-1.64%	-1.51%	-1.53%	-0.44%	-0.61%	-2.17%	-1.46%	-0.27%	-1.06%	-1.49%	-1.81%	-2.19%	-3.30%	-0.75%	-0.46%
5.0%	0.21%	-1.24%	0.35%	0.19%	0.09%	-1.02%	-0.28%	0.08%	0.03%	1.46%	0.69%	-0.30%	-0.60%	1.42%	0.93%
10.0%	0.73%	0.02%	0.48%	0.78%	1.02%	1.02%	0.40%	0.58%	1.39%	1.09%	0.80%	0.71%	0.39%	1.83%	1.62%

Table 8.7: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 2 (10% bonds)															
VaR\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-1.08%	-6.93%	-3.88%	-1.01%	-5.18%	-7.91%	-5.38%	-1.58%	-8.34%	-5.77%	-7.48%	-3.45%	-4.60%	-1.60%	-2.42%
1.0%	-1.19%	-1.79%	-1.48%	-0.47%	-1.24%	-2.89%	-4.34%	-1.64%	-5.02%	-3.57%	-4.06%	-1.68%	-3.91%	-1.22%	-0.75%
2.5%	-1.45%	-1.50%	-1.39%	-0.42%	-0.53%	-1.44%	-1.28%	-0.07%	-0.95%	-1.17%	-1.48%	-2.03%	-2.85%	-0.53%	-0.66%
5.0%	0.19%	-1.27%	0.23%	0.12%	-0.13%	-0.99%	-0.25%	-0.05%	-0.07%	1.14%	0.36%	-0.43%	-0.51%	1.10%	0.85%
10.0%	0.65%	-0.09%	0.45%	0.65%	0.87%	0.94%	0.47%	0.55%	1.03%	1.21%	0.65%	0.68%	0.26%	1.79%	1.68%

Table 8.8: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 4 (30% bonds)															
VaR\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-0.79%	-4.82%	-3.05%	-1.41%	-4.07%	-5.18%	-3.33%	-0.52%	-5.69%	-4.49%	-5.54%	-2.53%	-3.68%	-0.92%	-1.92%
1.0%	-0.88%	-1.59%	-0.71%	0.08%	-1.81%	-2.36%	-3.56%	-1.50%	-3.81%	-3.09%	-2.90%	-1.33%	-3.36%	-1.65%	-1.29%
2.5%	-1.09%	-1.14%	-0.93%	-0.14%	-0.67%	-1.31%	-0.98%	-0.17%	-0.94%	-0.67%	-1.02%	-1.50%	-2.19%	-0.41%	-0.63%
5.0%	0.14%	-0.94%	0.08%	0.08%	0.25%	-0.63%	-0.13%	0.10%	-0.17%	0.78%	0.10%	-0.37%	-0.36%	0.46%	0.86%
10.0%	0.50%	0.16%	0.40%	0.32%	0.57%	0.81%	0.72%	0.51%	0.87%	0.84%	0.73%	0.67%	0.27%	0.90%	1.14%

Table 8.10: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 5 (40% bonds)															
VaR\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-0.66%	-3.85%	-2.65%	-1.58%	-3.69%	-4.50%	-2.56%	-0.39%	-4.90%	-3.90%	-4.25%	-2.17%	-3.13%	-0.49%	-1.30%
1.0%	-0.74%	-1.15%	-0.44%	0.24%	-1.77%	-1.83%	-2.90%	-1.22%	-3.30%	-2.43%	-2.54%	-1.37%	-3.13%	-1.57%	-1.58%
2.5%	-0.91%	-0.88%	-0.71%	0.10%	-0.49%	-1.23%	-1.21%	-0.25%	-0.70%	-0.41%	-1.04%	-1.65%	-1.46%	-0.30%	-0.55%
5.0%	0.12%	-0.68%	-0.07%	-0.07%	0.11%	-0.60%	-0.14%	0.00%	0.11%	0.63%	0.14%	-0.12%	-0.23%	0.34%	0.48%
10.0%	0.42%	0.08%	0.45%	0.19%	0.30%	0.65%	0.57%	0.44%	0.77%	0.69%	0.46%	0.46%	0.05%	0.64%	0.96%

Table 8.11: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 1 (0% bonds)															
ETL\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-7.16%	-8.22%	-5.30%	-10.9%	-13.2%	-11.8%	-10.1%	-3.22%	-9.73%	-9.99%	-12.3%	-12.2%	-6.36%	-11.1%	-4.83%
1.0%	-3.64%	-6.03%	-3.75%	-5.83%	-8.84%	-8.59%	-7.60%	-2.46%	-8.43%	-7.06%	-8.63%	-7.49%	-5.81%	-6.60%	-3.50%
2.5%	-2.03%	-3.69%	-2.44%	-2.55%	-4.40%	-5.19%	-4.27%	-1.54%	-5.04%	-3.88%	-5.23%	-4.45%	-4.94%	-3.14%	-1.76%
5.0%	-1.51%	-2.51%	-1.29%	-1.30%	-2.28%	-3.52%	-2.76%	-0.62%	-2.72%	-1.65%	-2.93%	-2.75%	-3.22%	-1.61%	-0.54%
10.0%	-0.40%	-1.41%	-0.45%	-0.54%	-0.97%	-1.58%	-1.11%	-0.19%	-1.06%	-0.22%	-1.19%	-1.09%	-1.42%	-0.05%	0.34%

Table 8.12: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 2 (10% bonds)															
ETL\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-6.18%	-7.19%	-4.61%	-9.60%	-11.4%	-10.3%	-8.74%	-2.79%	-8.42%	-8.65%	-10.67%	-10.5%	-5.37%	-9.26%	-4.33%
1.0%	-3.14%	-5.26%	-3.17%	-5.06%	-7.60%	-7.52%	-6.58%	-2.13%	-7.32%	-6.12%	-7.51%	-6.44%	-5.02%	-5.58%	-3.08%
2.5%	-1.76%	-3.20%	-2.11%	-2.16%	-3.78%	-4.51%	-3.75%	-1.36%	-4.38%	-3.38%	-4.55%	-3.86%	-4.32%	-2.68%	-1.58%
5.0%	-1.32%	-2.20%	-1.12%	-1.12%	-1.99%	-3.07%	-2.41%	-0.59%	-2.35%	-1.46%	-2.57%	-2.37%	-2.85%	-1.36%	-0.51%
10.0%	-0.35%	-1.24%	-0.40%	-0.46%	-0.84%	-1.38%	-0.95%	-0.19%	-0.93%	-0.18%	-1.01%	-0.94%	-1.27%	-0.03%	0.31%

Table 8.13: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 3 (20% bonds)															
ETL\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-5.28%	-6.24%	-3.95%	-8.49%	-9.75%	-8.80%	-7.40%	-2.38%	-7.14%	-7.40%	-8.94%	-9.01%	-4.53%	-7.70%	-3.76%
1.0%	-2.69%	-4.53%	-2.65%	-4.35%	-6.47%	-6.48%	-5.63%	-1.82%	-6.27%	-5.23%	-6.42%	-5.54%	-4.30%	-4.64%	-2.74%
2.5%	-1.52%	-2.73%	-1.79%	-1.83%	-3.20%	-3.88%	-3.28%	-1.23%	-3.75%	-2.88%	-3.88%	-3.33%	-3.71%	-2.27%	-1.43%
5.0%	-1.14%	-1.88%	-0.98%	-0.97%	-1.71%	-2.66%	-2.09%	-0.60%	-2.01%	-1.24%	-2.23%	-2.03%	-2.46%	-1.13%	-0.49%
10.0%	-0.29%	-1.06%	-0.35%	-0.38%	-0.72%	-1.20%	-0.83%	-0.22%	-0.79%	-0.12%	-0.86%	-0.83%	-1.12%	-0.02%	0.25%

Table 8.14: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.



Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 4 (30% bonds)															
ETL\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-4.44%	-5.34%	-3.33%	-7.36%	-8.19%	-7.41%	-6.10%	-1.99%	-5.88%	-6.24%	-7.37%	-7.71%	-3.74%	-6.31%	-3.13%
1.0%	-2.26%	-3.85%	-2.23%	-3.74%	-5.52%	-5.42%	-4.72%	-1.54%	-5.27%	-4.38%	-5.37%	-4.67%	-3.55%	-3.81%	-2.42%
2.5%	-1.28%	-2.28%	-1.49%	-1.53%	-2.68%	-3.29%	-2.84%	-1.08%	-3.15%	-2.36%	-3.20%	-2.86%	-3.12%	-1.89%	-1.29%
5.0%	-0.97%	-1.57%	-0.87%	-0.81%	-1.46%	-2.27%	-1.78%	-0.61%	-1.66%	-1.01%	-1.87%	-1.74%	-2.10%	-0.92%	-0.47%
10.0%	-0.25%	-0.90%	-0.31%	-0.32%	-0.62%	-1.01%	-0.71%	-0.23%	-0.65%	-0.08%	-0.73%	-0.74%	-0.96%	-0.01%	0.22%

Table 8.15: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

Differences in drops in coverage ratio between the copula and the MV normal method for stylized portfolio 5 (40% bonds)															
ETL\Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.5%	-3.67%	-4.51%	-2.74%	-6.25%	-6.79%	-6.23%	-5.09%	-1.48%	-4.88%	-5.15%	-5.75%	-6.37%	-2.96%	-5.00%	-2.45%
1.0%	-1.87%	-3.20%	-1.83%	-3.16%	-4.64%	-4.45%	-3.91%	-1.12%	-4.36%	-3.55%	-4.27%	-3.86%	-2.83%	-3.08%	-2.05%
2.5%	-1.06%	-1.85%	-1.23%	-1.24%	-2.20%	-2.74%	-2.43%	-0.87%	-2.59%	-1.91%	-2.52%	-2.41%	-2.56%	-1.63%	-1.15%
5.0%	-0.81%	-1.28%	-0.75%	-0.64%	-1.21%	-1.91%	-1.52%	-0.54%	-1.32%	-0.78%	-1.48%	-1.51%	-1.75%	-0.75%	-0.49%
10.0%	-0.20%	-0.75%	-0.24%	-0.26%	-0.52%	-0.82%	-0.59%	-0.19%	-0.51%	-0.03%	-0.58%	-0.66%	-0.81%	0.00%	0.14%

Table 8.16: Differences are computed as value of copula method minus the value of the MV normal method. This means that a negative value in the table represent a bigger drop in coverage ratio for the copula method than for the MV normal method.

# Bibliography

- ANP. Dekking 'enkele' pensioenfondsen te laag. Internet, available at: <http://www.nrc.nl/economie/article1874032.ece/www.hoi-online.nl>, April 2008.
- ANP. Dekkingsgraad abp daalt tot 90 procent. Intenet, available at: [http://www.volkskrant.nl/economie/article1139941.ece/Dekkingsgraad ABP daalt tot 90 procent](http://www.volkskrant.nl/economie/article1139941.ece/Dekkingsgraad%20ABP%20daalt%20tot%2090%20procent), January 2009a.
- ANP. Dekkingsgraad pensioenfondsen verder verbeterd. Internet, available at: [http://www.nrc.nl/economie/article2444572.ece/Dekkingsgraad pensioenfondsen verder verbeterd](http://www.nrc.nl/economie/article2444572.ece/Dekkingsgraad%20pensioenfondsen%20verder%20verbeterd), December 2009b.
- A. (site) Azzalini. <http://azzalini.stat.unipd.it/sn/>.
- Adelchi Azzalini. The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, 32:159–188, 2005.
- Martin Neil Baily, Robert E. Litan, and Matthew S. Johnson. *The Origins of the Financial Crisis*. The Brookings institution, fixing finance series - paper 3 edition, November 2008.
- Eric Bouyé. Multivariate extremes at work for portfolio risk measurement. *Finance*, Volume 23:125–144, 2002.
- Thomas Breuer, Martin Jandacka, Klaus Rheinberger, and Martin Summer. *How to find plausible, severe, and useful stress scenarios*. Öesterreichische Nationalbank, 2009.
- Loran Cholette, Andréas Heinen, and Alfonso Valdesogo. Modeling international financial returns with a multivariate regime switching copula. March 2008.
- Etienne Cuvelier and Monique Noirhomme-Fraiture. Clayton copula and mixture decomposition. *Applied Stochastic Models and Data Analysis (ASMDA 2005)*, pages 699–708, May 2005.
- RF Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1008, July 1982.
- E.F. Fama. *Foundations of finance*. Basic Books, 1976.
- M. Fischer, C. Köck, S. Schlüter, and F. Weigert. An empirical analysis of multivariate copula models. was submitted to *Quantative Finance*, 2007.
- Dominique Guegan and Pierre André Maugis. Note on new prospects on vines. *CES Working papers*, 2008.
- Bruce E. Hansen. Autoregressive conditional density estimation. *International Economic Review*, 35(3), August 1994.
- Glyn Holton. *Value-at-Risk: Theory and Practice*. Academic Press, 2003.
- Philippe Jorion. *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, 2007.
- M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, 1938.
- Erik Kole, Kees Koedijk, and Marno Verbeek. Selecting copulas for risk management. *Journal of Banking & Finance*, Volume 31:2405–2423, August 2007.
- Filip Lindskog. Linear correlation estimation. available on: <http://www.math.ethz.ch/lindskog>, December 2000.
- Y. Malevergne and D. Sornette. Testing the gaussian copula hypothesis for financial assets dependences. *Quantative Finance*, 3(4):231–250, 2003.
- Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantative risk management*. Princeton University Press, 2005.
- Financiële Redactie. Miljardenverlies pensioenfondsen. Internet, available at: [http://www.nrc.nl/economie/article1875464.ece/Miljardenverlies pensioenfondsen](http://www.nrc.nl/economie/article1875464.ece/Miljardenverlies%20pensioenfondsen), April 2008.

- Tyrrell R. Rockafellara and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26:1443–1471, July 2002.
- A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231, 1959.
- C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, January 1904.
- Freek Staps. Wall street in crisis na val zakenbanken. Internet, available at: [http://www.nrc.nl/economie/article1983919.ece/Einde Lehman in zicht, Merrill overgenomen, September 2008](http://www.nrc.nl/economie/article1983919.ece/Einde_Lehman_in_zicht,_Merrill_overgenomen,_September_2008).
- Dirk Tasche. Expected shortfall and beyond. *Journal of Banking & Finance*, 26:1487–1503, July 2002.
- Rob Van den Goorbergh, Onno W. Steenbok, R. Molenaar, and Peter Vlaar. Risk models after the credit crisis. Working Paper Series Available at SSRN: <http://ssrn.com/abstract=1521061>, December 2009.
- Peter J.G. Vlaar. Term structure modeling for pension funds: What to do in practice? Working Paper Series Bank of the Netherlands - Econometrics Available at SSRN: <http://ssrn.com/abstract=965363>, December 2006.
- Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, December 1945.