On the use of Value-at-Risk based models for the Fixed Income market as a risk measure for Central Counterparty clearing

Oskar Kallur

May 23, 2016

Supervisor at KTH: Boualem Djehiche

Supervisors at NASDAQ: Johan L. Olsson & David Sjöblom

Abstract

In this thesis the use of VaR based models are investigated for the purpose of setting margin requirements for Fixed Income portfolios. VaR based models has become one of the standard ways for Central Counterparties to determine the margin requirements for different types of portfolios. However there are a lot of different ways to implement a VaR based model in practice, especially for Fixed Income portfolios. The models presented in this thesis are based on Filtered Historical Simulation (FHS). Furthermore a model that combines FHS with a Student's t copula to model the correlation between instruments in a portfolio is presented. All models are backtested using historical data dating from 1998 to 2016. The FHS models seems to produce reasonably accurate VaR estimates. However there are other market related properties that must be fulfilled for a model to be used to set margin requirements. These properties are investigated and discussed.

Användningen av Value-at-Risk baserade modeller för Fixed Income marknaden som riskmått för Central Counterparty clearing.

Sammanfattning

I denna uppsats undersöks användningen av VaR baserade modeller för att sätta marginkrav för Fixed Income portföljer. VaR baserade modeller har blivit en standardmetod för Central Counterparties för att räkna ut marginkrav för olika typer av portföljer. Det finns många olika tillvägagångssätt för att räkna ut VaR i praktiken, speciellt för Fixed Income portföljer. Modellerna som presenteras i den här uppsatsen är baserade på Filterad Historisk Simulering (FHS). Dessutom presenteras en modell som kombinerar FHS med en Student's t copula för att modellera korrelationen mellan olika instrument. Alla modeller backtestas på historisk data från 1998 till 2016. Modellerna ger rimliga VaR skattningar i backtesterna. Däremot finns det andra marknadsrelaterade egenskaper som en modell måste uppfylla för att kunna användas för att sätta margin. Dessa egenskaper undersöks och diskuteras.

Contents

1 Introduction

In this thesis the use of *Value-at-Risk* (VaR) based models to determine the *Initial Margin* (IM) of a portfolio consisting of Fixed Income instruments is investigated. VaR is one of the standard measures that is used to describe the risk of a portfolio. The advantages of using VaR as a risk measure is that it is intuitively easy to understand. Determining the VaR of a portfolio requires assumptions about the loss distribution of the portfolio, and different assumptions will lead to different results. One of the most popular model families is known as *Filtered Historical Simulation* (FHS). This method of determining VaR has been widely studied academically, see for example [13] and [4], however it is not clear if these models are suitable to determine the IM in a practical implementation.

The role of a *central counterparty* (CCP) is to ensure stability and efficiency in financial markets. The CCP acts as an intermediary between counterparties in a financial transaction and thereby makes sure that the contractual obligations of the trade is fulfilled even in the case of default of one of the counterparties. Thus one important practice of the CCP is to manage the risk of financial agreements between counterparties. This is done in several ways, one of which is to require the counterparties to post margin associated with the risk of their portfolios. The margin is calculated in a monetary measure and the counterparty corresponding to each portfolio must provide collateral to cover the margin.

In the case that a counterparty defaults, the time that it takes to replace their portfolio is called the liquidation period or margin period of risk. More specifically, this is the time it takes to close out the instruments of a portfolio and to hedge the market risk. The purpose of the IM is to protect the portfolio against the market risk during the liquidation period.

One of the most important properties of the IM is that the model which is used should be transparent to the members of the CCP such that the members can understand the implications of different types of instruments on the IM of a portfolio. Furthermore the IM should be dynamic in the sense that it adapts to current market conditions. A sudden increase in volatility should be reflected in the IM. However it is also important that the IM does not overly compensate for dynamic volatility since this will result in a large difference between the IM during calm and volatile periods. The large IM:s during volatile periods will then result in even more stress in an already stressed market and thus encourage market instability, this is known as *procyclicality*. For a discussion about the procyclicality of different IM models see [12] and [10].

In a report by ESMA (European Securities and Markets Authority) [11] it is stated that: *According to the Regulatory Technical Standards, a CCP shall ensure that its policy for selecting and revising the confidence interval, the liquidation period and the look-back period deliver forward looking, stable and prudent margin requirements that limit procyclicality* *to the extent that the soundness and financial security of the CCP is not negatively affected. This shall include avoiding when possible disruptive or big step changes in margin requirements and establishing transparent and predictable procedures for adjusting margin requirements in response to changing market conditions.*

Since the purpose of the IM is to protect the portfolio during the liquidation period the VaR value should be determined on a time period corresponding the length of the liquidation period. However, the available data can be used to determine intraday changes in value of a portfolio and thus more data points will be available to determine one day VaR values. For the purpose of investigating the different models this will increase the significance of the results, and thus the one day VaR will be modeled in this thesis.

The amount of models that are possible to use to determine the VaR of a portfolio are immense. Thus only a few selected models are investigated in this thesis. The choice to investigate FHS models were made since they are widely used and have properties that are interesting from a clearing perspective. Furthermore the accuracy of the available data and the implications of using different valuation techniques are not explored.

It is the purpose of this thesis to investigate different models for determining the IM of a portfolio using VaR based models. The best model to use might not be the one that produces the best results in classical VaR backtests. It is more important that the model is transparent and robust. Furthermore it is beneficial if the model has parameters that intuitively describes different kinds of portfolio risk. These parameters can then be calibrated to account for concerns about the market. Additionally the models should comply with regulatory restrictions. Thus the implementation of the model might not be a VaR model by definition. However the model should produce VaR values that are reasonable in backtesting when the parameters are tuned to the available data.

In Sections 2 - 5 the necessary background knowledge and theory is presented and in Section 6 the available data is presented. In Section 7 FHS models with different volatility processes are backtested on portfolios consisting of a single instrument. In Section 8 the models are extended to a portfolio level. In Section 9 possible adjustments of the model is discussed to account for regulatory restrictions and market concerns and in Section 10 the results are discussed.

2 Value-at-Risk

VaR is a risk measure that measures the risk of an investment at a specified confidence level using the distribution of the potential return (loss) of the investment. If the distribution of the potential loss L of an investment X is known, the VaR at confidence level p is

$$\mathrm{VaR}_p(X) = F_L^{-1}(1-p) \tag{1}$$

where F_L^{-1} is the quantile function of L. The interpretation of VaR is quite intuitive and this is one of the strengths of using VaR as a risk measure. The formula (1) can be interpreted as that the probability of a loss exceeding $\mathrm{VaR}_p(X)$ is p.

There are some theoretical shortcomings of using VaR as a risk measure. Perhaps the most important is that VaR disregards the distribution of losses greater than the VaR value, this leads to a problem usually referred to as *hiding risk in the tail*.

Another shortcoming of VaR is that it is not *subadditive*. A risk measure ρ is subadditive if

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2) \tag{2}$$

where X_1 and X_2 denotes two different investments. This property represents the fact that diversification should be rewarded and it is especially important from a clearing point of view. Since VaR is not subadditive it is theoretically possible that the VaR of a portfolio is greater than the sum of the VaR of the individual instruments.

To account for these shortcomings, another risk measure known as *Expected Shortfall* (ES) has been argued to be an improvement of VaR. The ES at confidence level p of investment X is defined as

$$\mathrm{ES}_p(X) = \frac{1}{p}\int_0^p \mathrm{VaR}_u(X)du. \tag{3}$$

Intuitively, ES is the average VaR below the threshold p. Thus it resolves the problem of hiding risk in the tail. In addition ES is subadditive.

2.1 Estimation

There are several different approaches to estimating the VaR of a portfolio including both parametric and non-parametric approaches. An example of a parametric approach is the Var-Covar method where the returns are assumed to have a multivariate normal distribution. The VaR can then be determined analytically from the quantile using the properties of an elliptical distribution. This goes against many of the widely accepted stylized facts about financial returns such as volatility clustering and fat tailedness.

An example of a non-parametric approach is the Historical Simulation where the returns are assumed to follow the distribution of past returns. The VaR/ES is then constructed by taking the empirical quantile of a sample of past returns. Here no assumptions are made about the specific distribution of the returns. The empirical quantile estimate $\widehat{F}_{n,X}^{-1}$ for the sample $(X_1, ..., X_n)$ of independent and identically distributed (*IID*) random variables is

$$\widehat{F}_{n,X}^{-1}(p) = X_{[n(1-p)]+1,n} \tag{4}$$

where $(X_{1,n}, ..., X_{n,n})$ is ordered such that $X_{1,n} \geq \cdots \geq X_{n,n}$ and the hard brackets denotes the floor function. Thus the empirical VaR estimate for a sample of historical ordered losses $(L_{1,n}, ..., L_{n,n})$ is

$$\widehat{\mathrm{VaR}}_p = \widehat{F}_{n,L}^{-1}(1-p) = L_{[np]+1,n}. \tag{5}$$

The corresponding empirical ES estimate is acquired by inserting (4) into (3) such that

$$\widehat{\mathrm{ES}}_p = \frac{1}{p}\int_0^p L_{[nu]+1,n} du = \frac{1}{p}\left(\sum_{k=1}^{[np]} \frac{L_{k,n}}{n} + \left(p - \frac{[np]}{n}\right) L_{[np]+1,n}\right). \tag{6}$$

Note that when $[np]$ is an integer the empirical ES estimate will simply be the average of the $[np]$ largest losses.

The drawback of HS is that the volatility is assumed to be constant. For financial loss data volatility clustering can often be observed, thus the assumption that the losses are *IID* does not seem to be reasonable.

3 Filtered Historical Simulation

3.1 Filtering

The assumption that the volatility is constant in a sample of financial returns can often be rejected simply by inspection. The returns often shows properties such as volatility clustering. When working with VaR models it is convenient to model the losses (negative returns) and this approach will be used in this thesis. The FHS model accounts for time dependent volatility by *filtering* each loss in a sample with a volatility process. The word filtering is used by convention and refers to that the losses in a sample are scaled with some volatility process to obtain a sample of *filtered losses*. Let l_t and σ_t be the loss and volatility process of an investment at time t respectively. The losses are then assumed to follow the relation

$$l_t = \sigma_t z_t \tag{7}$$

where z_t is the filtered loss at time t. The set of filtered losses $\{z_t\}$ are assumed to be *IID* random variables with zero mean. The volatility process can be specified in different ways but it should be predictable given the information $\mathcal{F}_{t-1}$. Thus for a sample of losses $(L_1, ..., L_{T-1})$ and a defined volatility process σ_t the filtered losses $(Z_1, ..., Z_{T-1})$ can be obtained by filtering each loss with the volatility process. Since the filtered losses are assumed to be *IID* the VaR at time T given the information $\mathcal{F}_{T-1}$ is

$$\mathrm{VaR}_p^T = \sigma_T F_{z_T}^{-1}(1-p). \tag{8}$$

The quantile of z_T is then estimated by the empirical quantile of the filtered losses.

In the FHS model one assumes that the losses of an investment follow a time series process such that the volatility of the losses is allowed to be time dependent. In this thesis the GARCH (Generalized Autoregressive Conditional Heteroscedasticity) and EWMA (Exponentially Weighted Moving Average) processes are used to filter the samples. These are two relatively simple models that have been widely studied for the purpose of modeling financial returns. Both of these models have properties that will make the volatility process react appropriately to changing market conditions.

3.2 GARCH

Let $\{z_t\}$ be a set of *IID* random variables with zero mean and variance 1. Then the GARCH(1,1) process is defined as

$$\begin{aligned} l_t &= \sigma_t z_t \\ \sigma_t^2 &= \omega + \alpha l_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{9}$$

For $\omega > 0$ and $\alpha + \beta < 1$ the GARCH(1,1) process is stationary. In this thesis only the GARCH(1,1) model will be studied and it will be referred to as the GARCH process.

The parameters (ω, α, β) can be estimated using the maximum likelihood method. To use this method distributional assumptions of $\{z_t\}$ must be made. In the case that $\{z_t\}$ is assumed to be normally distributed the Gaussian likelihood function of the sample $(L_1, .., L_n)$ is given by

$$L_n(\omega, \alpha, \beta; L_1, ..., L_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{L_t^2}{2\sigma_t^2}\right) \tag{10}$$

with σ_t^2 given as in (9). Maximising (10) is equivalent to maximising the Gaussian log-likelihood given by

$$l_n(\omega, \alpha, \beta; Z_1, ..., Z_n) = -\frac{1}{2}\sum_{t=1}^{n}(\log(2\pi) + \log(\sigma_t^2) + \frac{Z_t^2}{2\sigma_t^2}). \tag{11}$$

It is possible to assume different distributions of $\{z_t\}$ in which case other likelihood functions are acquired. All estimation of parameters of the GARCH model used in this thesis were performed using the rugarch package in R [5] and $\{z_t\}$ are assumed to have a standard normal distribution. Note that the assumption of normality is only used in the estimation of the parameters.

3.3 EWMA

Let $\{z_t\}$ be a set of *IID* random variables with zero mean. Then the EWMA process σ_t^2 is defined as

$$\begin{aligned} l_t &= \sigma_t z_t \\ \sigma_t^2 &= (1-\lambda)l_{t-1}^2 + \lambda\sigma_{t-1}^2, \qquad \lambda \in (0,1) \end{aligned} \tag{12}$$

Note that the EWMA process can be seen as a GARCH(1,1) process with $(\omega, \alpha, \beta) = (0, 1-\lambda, \lambda)$, however with these parameters the GARCH process is not stationary (in fact the EWMA process is a form of Integrated-GARCH process). The EWMA process was popularised by RiskMetrics [7]. The EWMA process can be viewed as a weighted average of the sample variance. To demonstrate this consider the sample $(L_1, ..., L_{t-1})$. The EWMA volatility estimate can then be expressed as

$$\sigma_t^2 = \sum_{i=1}^{t-1} \alpha_i L_{t-i}^2, \qquad \alpha_i = (1-\lambda)\lambda^{i-1} \tag{13}$$

With these weights equation (13) can be rewritten as

$$
\begin{aligned}
\sigma_t^2 &= \sum_{i=2}^{t-1}(1-\lambda)\lambda^{i-1}L_{t-i}^2 + (1-\lambda)L_{t-1}^2 \\
&= \lambda\sum_{i=1}^{t-2}(1-\lambda)\lambda^{i-1}L_{t-1-i}^2 + (1-\lambda)L_{t-1}^2 \\
&= \lambda\sigma_{t-1}^2 + (1-\lambda)L_{t-1}^2.
\end{aligned} \tag{14}
$$

Typical values of λ is in the range $(0.90, 0.99)$. For these values the weights α_i will be greater for small i, thus more weight are assigned to recent observations. The advantage of using the EWMA process as a volatility estimate in the filtering is that there is only one parameter. Conventionally no estimation is done when using the EWMA process, instead the value of λ is fixed to a value which is assumed to describe the properties of the sample. Thus no estimation is required and no assumptions about the distribution of $\{z_t\}$ needs to be made. From this point onward EWMA 95 and EWMA 99 will refer to an EWMA process with λ equal to 0.95 and 0.99 respectively.

3.4 Kupiec unconditional coverage test

The Kupiec test was introduced by Kupiec in '95 and has become a standard way of backtesting VaR models [8]. Let $\{I_t\}$ be a sequence of indicator variables defined as 1 if the VaR is lower than the loss at time t (if a breach has occurred) such that

$$
I_t = \begin{cases} 1 & \text{if} \quad \mathrm{VaR}_p^t < L_t \\ 0 & \text{if} \quad \text{otherwise} \end{cases} \tag{15}
$$

By the definition of the VaR the probability of a breach should be p such that $\mathbb{E}[I_t] = p$. Thus we want to test the null hypothesis $\mathbb{E}[I_t] = p$ against the hypothesis $\mathbb{E}[I_t] \neq p$. The likelihood of the null hypothesis is

$$
L(p; I_1, ..., I_n) = (1-p)^{n_0} p^{n_1} \tag{16}
$$

where n_1 is the amount of breaches and n_0 is $n - n_1$. The likelihood under the alternative hypothesis is

$$
L(\pi; I_1, ..., I_n) = (1-\pi)^{n_0} \pi^{n_1}. \tag{17}
$$

The standard likelihood ratio then becomes

$$LR = -2\log\left(\frac{L(p; I_1, ..., I_n)}{L(\widehat{\pi}; I_1, ..., I_n)}\right) \tag{18}$$

where $\widehat{\pi} = \frac{n_1}{n}$ is the maximum likelihood estimate of π. The LR is assumed to be a realisation of a χ^2 random variable with 1 degree of freedom and thus a p-value corresponding to the rejection of the null hypothesis can be determined.

4 Copulas

Copulas are a tool for specifying the joint distribution of variables which have known marginal distribution. Let $(X_1, ..., X_n)$ be a set of random variables with known marginal distributions $(F_{X_1}, ..., F_{X_n})$ then the purpose of a copula is to specify the joint distribution $F_{X_1,...,X_n}$. Copulas are widely used to model correlations within the financial industry. One of the benefits of using a copula is that it is easy to simulate values from a joint distribution when the marginals are known. There are a lot of different copulas with different properties and the key to using copulas is to find one that fits well with the data. The theory of copulas rely on the probability and quantile transforms;

Definition 4.1
If X is a random variable with the continuous marginal distribution F then $F(X)$ is uniformly distributed on the interval $(0, 1)$.

If U is uniformly distributed on $(0, 1)$ and G is any distribution function, then $G^{-1}(U)$ has distribution function G.

Using these properties it is possible to specify a joint distribution of $(X_1, ..., X_n)$ with a certain dependence structure by specifying a distribution of a set of dependent uniform $(0, 1)$ variables $(U_1, ..., U_n)$. Since $X_i \overset{d}{=} F^{-1}_{X_i}(U_i), \quad i \in \{1, ..., n\}$ the marginals will be the same for $(X_1, ..., X_n)$ and $(F^{-1}_{X_1}(U_1), ..., F^{-1}_{X_n}(U_n))$ but the correlation will be "inherited" from the uniform variables. In a Student's t copula the uniform variables are distributed as

$$\begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix} \overset{d}{=} \begin{pmatrix} t_\nu^{-1}(Z_1) \\ t_\nu^{-1}(Z_2) \\ \vdots \\ t_\nu^{-1}(Z_n) \end{pmatrix} \tag{19}$$

where $(Z_1, ..., Z_n)^T$ have a multivariate Student's t distribution with ν degrees of freedom and correlation matrix R and t_ν^{-1} denotes the quantile function of the Student's t distribution with ν degrees of freedom. The density function $f_{\nu,R}$ of the multivariate Student's t distribution with ν degrees of freedom and correlation matrix R is given by

$$f_{\nu,R}(\mathbf{x}) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\nu^{n/2}\pi^{n/2}}|R|^{-1/2}\left(1+\frac{1}{\nu}\mathbf{x}^T R^{-1}\mathbf{x}\right)^{-(\nu+n)/2}. \tag{20}$$

Let $F_{U_1,...,U_n}$ denote the joint distribution function of $(U_1,...,U_n)$ then

$$\begin{aligned} F_{U_1,...,U_n}(F_{X_1}(x_1),...,F_{X_n}(x_n)) &= \mathbb{P}(U_1 \le F_{X_1}(x_1),...,U_n \le F_{X_n}(x_n)) \\ &= \mathbb{P}(F_{X_1}^{-1}(U_1) \le x_1,...,F_{X_n}^{-1}(U_n) \le x_n) \\ &= F_{X_1,...,X_n}(x_1,...,x_n) \end{aligned} \tag{21}$$

The distribution function $C = F_{U_1,...,U_n}$ of a random vector $(U_1,...,U_n)$ is called a copula. Thus the Student's t copula with ν degrees of freedom and correlation matrix R is

$$C^t_{\nu,R}(u_1,...,u_n) = t_{\nu,R}(t_\nu^{-1}(u_1),...,t_\nu^{-1}(u_n)). \tag{22}$$

As seen in equation (19) simulating from a Student's t copula is easily done by simulating from a multivariate Student's t distribution and then transforming the resulting variables appropriately.

The density of the Student's t copula can be derived from (22) and (20) to be

$$c(u_1,...,u_n) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)\Gamma\left(\frac{\nu}{2}\right)^{n-1}\left(1+\frac{\mathbf{x}^T R^{-1}\mathbf{x}}{\nu}\right)^{-\frac{\nu+n}{2}}}{|R|^{1/2}\Gamma\left(\frac{\nu+1}{2}\right)^n \prod_{j=1}^n\left(1+\frac{x_j^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \tag{23}$$

where $\mathbf{x} = (t_\nu^{-1}(u_1),...,t_\nu^{-1}(u_n))$. For a more thorough analysis of the properties of the Student's t copula see for example [9].

4.1 Kendall's Tau

Kendall's tau is a measure of correlation that is convenient when working with copulas. It is defined for the vector (X_1, X_2) as

$$\tau(X_1,X_2) = \mathbb{P}((X_1-X_1')(X_2-X_2')>0) - \mathbb{P}((X_1-X_1')(X_2-X_2')<0) \tag{24}$$

where (X_1', X_2') is an independent copy of (X_1, X_2). Given a sample $\{\mathbf{X_1}, ..., \mathbf{X_n}\}$, $\quad \mathbf{X_i} = (X_{i,1}, X_{i,2})$ of *IID* random variables the sample estimate of Kendall's tau is

$$\widehat{\tau} = \binom{n}{2}^{-1} \sum_{j<k} \operatorname{sign}\left((X_{j,1} - X_{k,1})(X_{j,2} - X_{k,2})\right) \tag{25}$$

where $\operatorname{sign}(x)$ is equal to 0 for $x = 0$ and $\frac{x}{|x|}$ otherwise. Furthermore if the $\mathbf{X_k}$ are elliptically distributed then the linear correlation parameter can be estimated as

$$\widehat{\rho} = \sin\left(\frac{\pi}{2}\widehat{\tau}\right). \tag{26}$$

The advantage of using Kendall's tau is to measure correlation is that it is invariant to strictly increasing transformations such that if f and g are continuous and strictly increasing functions then $\tau(f(X_1), g(X_2)) = \tau(X_1, X_2)$. Thus if (U_1, U_2) has a Student's t copula such that

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} t_\nu(Z_1) \\ t_\nu(Z_2) \end{pmatrix} \tag{27}$$

where (Z_1, Z_2) has a multivariate t distribution then $\tau(U_1, U_2) = \tau(Z_1, Z_2)$. The correlation between Z_1 and Z_2 is defined in the correlation matrix of the multivariate t distribution, and since the multivariate t is an elliptic distribution the correlation used in the correlation matrix can be estimated through (25) and (26).

4.2 Pseudo maximum likelihood estimation of the t copula

Here the pseudo maximum likelihood estimation of the parameters (ν, R) of a Student's t copula is explained. Given a set of observations $(\mathbf{u_1}, ..., \mathbf{u_m})$ where $\mathbf{u_k} = (u_{1,k}, ..., u_{n,k})$, $\quad k \in \{1, ...m\}$ the correlation parameters in the correlation matrix R can be estimated through using (26). The resulting estimate $\widehat{R}$ is not guaranteed to be positive definite, however the matrix may be adjusted using some method to find a positive definite matrix that is close to the estimated. Then the only parameter that needs to be estimated is the degrees of freedom ν. This is done by maximum likelihood such that

$$\widehat{\nu} = \operatorname{argmax}_\nu \left(\sum_{i=1}^{m} \log(c(\mathbf{u_k}, \nu, \widehat{R})) \right) \tag{28}$$

where c is defined in (23).

This method of estimating the parameters of a Student's t copula is convenient since it is fast, even for high dimensions. A pure maximum likelihood estimate where ν and R are maximised simultaneously is a lot slower and not realistic in high dimensions.

5 Fixed income

In this section the notation and assumptions used to value certain interest rate derivatives in the fixed income market will be presented. Here fixed income instruments refers to instruments which are dependent on the interest rate structure of the market. The valuation is done by replicating the cash flows of the corresponding instruments using simple discounting techniques. There is no correct way of determining the present value of fixed income instruments, different models will yield different values, it should be mentioned however that more sophisticated models exist (see for example [14]). For the purpose of this model the interesting part is the change in values corresponding to changing market conditions and not the values themselves. The values will be used to construct a set of characteristic instruments which are supposed to describe some of the common instruments cleared at NASDAQ.

5.1 Discounting cash flows

The value of having one SEK today is not always the same as having one SEK tomorrow or in one year. The valuation of future *cash flows* is known as *discounting*. Here a cash flow is a known amount of money to be exchanged at a specified future time, and discounting refers to the relationship between the value of money at two different times.

Definition 5.1
The discount function $d(t,T)$ determines the value at time t of having one SEK at time T. It holds for all times t that $d(t,t) = 1$.

Thus the discount function can be used to determine the present value of future cash flows. A zero coupon bond is an example of a contract where a fixed cash flow is to be exchanged at a future date. The holder of a zero coupon bond is guaranteed to receive the *face value* N at a future time T. The value Π_{zcb} at time t of a zero coupon bond with maturity T and face value N can be determined using the discount function to be

$$\Pi_{zcb}(t; T, N) = d(t,T)N. \tag{29}$$

In this type of contract there is a risk that the issuer of the bond defaults and thus is not able to pay out the money it owes. This type of risk is known as *credit risk*. The markets view of the risk of a zero coupon bond can be determined from the market price of the bond, if the bond is considered risky the market price will be lower than if the bond is considered safe. From equation (29) it can be seen that this also reflects directly on the discount function. Thus different discount functions are used to discount different types of instruments. To derive the *discount curve* $d(0,T), \quad T \geq 0$ (the value today of having one SEK at time T) a set of instruments with different maturities are assumed to have the same underlying risk. The discount curve is then bootstrapped using some method to determine the continuous discount curve. An example of this is the treasury curve, here a set of bonds issued by the government are assumed to have the same risk. The discount curve is then determined using the market prices of these bonds.

The notation $\Pi(t)$ is used in this thesis to denote the net present value of cash flows to be received in the future. From this point onward contract specific parameters such as the maturity T and face value N in the zero coupon bond case will not be explicitly stated as arguments to simplify the notation.

5.2 Compounding

Compounding determines how interest rates are accumulated over time. The compounding types that will be used in this thesis are simple and yearly compounding. Simple compounding means that interest are accumulated as a linear function of time such that the interest gathered between time T_1 and T_2 are

$$r_s(T_1, T_2)(T_2 - T_1), \qquad T_2 > T_1 \tag{30}$$

where r_s is the simple interest rate. The relation between a simple rate and the corresponding discount function is

$$1 + r_s(T_1, T_2) = \frac{1}{d(T_1, T_2)}. \tag{31}$$

Yearly compounding means that interest are gathered yearly such that the interest gathered between time T_1 and T_2 is

$$(1 + r_y(T_1, T_2))^{T_2 - T_1} - 1, \qquad T_2 > T_1 \tag{32}$$

where T_1 and T_2 are measured in years and r_y is the yearly compounded interest rate. The relation between a yearly compounded rate and the corresponding discount function is

$$(1 + r_y(T_1, T_2))^{T_2 - T_1} = \frac{1}{d(T_1, T_2)}. \tag{33}$$

5.3 LIBOR type forward rates

LIBOR is short for London Interbank Offered Rate and it is the average rate that the leading London banks use to borrow money from each other. The concept of this rate has extended to other countries and currencies. In the Swedish market the STIBOR (Stockholm Interbank Offered Rate) is used as a reference rate for different types of interest rate derivatives. LIBOR type rates are simple rates and the present rates for different maturities can directly be observed in the market.

Since the STIBOR is associated with cash flows between banks the rate can be used to determine the discount function corresponding to a loan between banks. Using prices of interest rate derivatives with the STIBOR as a reference rate the discount function corresponding to the risk between the Stockholm banks that are used to determine the STIBOR can be determined. The cash flows of some contracts are determined by what the STIBOR rate will be at a future time i.e. the forward STIBOR rate. The forward STIBOR rate is not known at the inception of a contract and thus in valuation it has to be estimated.

The implied forward STIBOR rate contracted at time t for the period between T_1 and T_2 is defined as

$$L(t;T_1,T_2) = \frac{d(t,T_1) - d(t,T_2)}{(T_2 - T_1)d(t,T_2)}. \tag{34}$$

5.4 Forward rate agreements

A Forward Rate Agreement (FRA) is an agreement between two counterparties to exchange an amount corresponding to a fictitious loan initiated at a future date T_1 with maturity at T_2 ($T_1 < T_2$). The rate of the loan is the difference between a *fixed rate* and the future value of a *floating rate* that corresponds to the forward rate between time T_1 and T_2. The fixed rate is agreed upon between the counterparties at the start of the contract.

The FRA contracts cleared in SEK at NASDAQ uses the 3MSTIBOR as the floating rate (the "3M" in the name implies that it is a three month forward rate). The present value Π_{FRA} of a FRA contract with fixed rate R and notional N is

$$\Pi_{FRA}(t) = SN\big(R - L(t,T_1,T_2)\big)(T_2 - T_1)d(t,T_2) \tag{35}$$

where S is the side of the contract (1 it the FRA is bought and -1 if the FRA is sold). Using equation (34) this can be rewritten as

$$\Pi_{FRA}(t) = SN\big((R(T_2 - T_1) - 1)d(t, T_2) - d(t, T_1)\big). \tag{36}$$

5.5 Interest rate swaps

An interest rate swap can be seen as a chain of FRAs. A fixed rate is agreed upon and the difference between the fixed and a floating rate is exchanged a number of times until expiry of the swap. The swaps treated in this report will have a constant fixed rate throughout the lifespan of the swap and the floating rate will again be the 3MSTIBOR. The fixed rate is exchanged once a year and the floating rate is exchanged quarterly. The present value Π_{swap} of an interest rate swap with fixed rate R and notional N where the floating payments are exchanged at $(T_1, ..., T_n)$ and the fixed payments are exchanged at $(S_1, ..., S_m)$ is

$$\begin{aligned}\Pi_{swap}(t) = SN\Bigg(&\sum_{j=1}^{m} R(S_j - S_{j-1})d(t, S_j) \\ -R_0(T_1 - t)d(t, T_1) - &\sum_{i=2}^{n} L(t, T_{i-1}, T_i)(T_i - T_{i-1}))d(t, T_i)\Bigg)\end{aligned} \tag{37}$$

where $S_0 = t$. Here R_0 is the first floating rate, this is known at time t and thus does not need to be estimated. From this points onwards the contract specific parameters will not be stated to simplify notation. By inserting (34) in (37) we get

$$\begin{aligned}\Pi_{swap}(t) = SN\Bigg(&\sum_{j=1}^{m} R(S_j - S_{j-1}))d(t, S_j) \\ &- R_0(T_1 - t)d(t, T_1) - \sum_{i=2}^{n}(d(t, T_{i-1}) - d(t, T_i))\Bigg)\end{aligned} \tag{38}$$

and notice that the sum of the floating rates is a telescopic sum and therefore cancels out. The final equation therefore becomes

$$\Pi_{swap}(t) = SN\Bigg(\sum_{j=1}^{m} R(S_j - S_{j-1})d(t, S_j) - (1 + R_0(T_1 - t))d(t, T_1) - d(t, T_n)\Bigg) \tag{39}$$

In the case that $T_1 - t$ is equivalent to three months, i.e. that there is three months until the first floating payment, the value of R_0 is equivalent to $L(t; t, T_1)$. Since the fixed payments are yearly we get that $S_j - S_{j-1} = 1$ for $j \in (1, ..., m)$. Using this the equation can be simplified further to get

$$\Pi_{swap}(t) = SN\left(\sum_{j=1}^{m} Rd(t, S_j) - 1 + d(t, T_n)\right). \tag{40}$$

5.6 Bond Forwards

A bond forward contract is an agreement where the holder of the contract is obligated to buy a *coupon bond* at a future settlement date. The price of the coupon bond to be bought is agreed upon at the inception of the contract. The underlying coupon bond is exchanged at the settlement time.

A coupon bond is similar to a zero coupon bond in that at the maturity T the holder of the contract is guaranteed the cash flow corresponding to the notional N of the bond. In a coupon bond the holder is also guaranteed coupon payments at certain times during the lifespan of the bond. The coupon is often quoted as a percentage of the notional. The value of a coupon bond with notional N and constant coupon c paid out at times $(T_1, ..., T_n = T)$ is

$$\Pi_{cb}(t) = N\left(\sum_{i=1}^{n} cd(t, T_i) + d(t, T_n)\right) \tag{41}$$

If coupon payments are made yearly for n years such that $T_i - T_{i-1} = 1$ then the value of the bond can be expressed using the yearly compounded interest rate corresponding to the discount function as

$$\Pi_{cb}(t) = N\left(\sum_{i=1}^{n} \frac{c}{(1 + r_i)^i} + \frac{1}{(1 + r_i)^n}\right) \tag{42}$$

with the notation $r_i = r(t, T_i)$.

The *yield* of a coupon bond is acquired by replacing r_i in equation (42) with a constant value y. Thus the yield y of a coupon bond is the value that solves the equation

$$\Pi_{cb}(t) = N\left(\sum_{i=1}^{n} \frac{c}{(1 + y)^i} + \frac{1}{(1 + y)^n}\right). \tag{43}$$

The sum in (43) can be rewritten using the properties of a geometric sum to get

$$\begin{aligned}\Pi_{cb}(t) &= N\left(\frac{c}{1+y}\left(\frac{1-(1+y)^{-n}}{y(1+y)^{-1}}\right)+\frac{1}{(1+y)^n}\right)\\ &= \frac{N}{(1+y)^n}\left(\frac{c}{y}\left((1+y)^n-1\right)+1\right)\end{aligned} \tag{44}$$

The price of the bond forward contract is quoted in the yield of the underlying coupon bond at the settlement time T_s. The amount paid at settlement for the underlying coupon bond with a fixed yield R is therefore

$$\frac{N}{(1+R)^n}\left(\frac{c}{R}\left((1+R)^n-1\right)+1\right) \tag{45}$$

and the market value of the bond at time T_s is obtained using the usual discounting of cash flows as

$$\left(\sum_{i=1}^{n} cd(T_s,T_i)+d(T_s,T_n)\right) \tag{46}$$

where $(T_1,..,T_n)$ is the times at which coupons are paid. The final value Π_{bf} of a bond forward contract X with contracted rate R, nominal value N and constant coupon c with coupons paid out yearly can now be computed using (45) and (46) to be

$$\Pi_{bf}(t) = SN\left(\left(\sum_{i=1}^{n} cd(t,T_i)+d(t,T_n)\right)-\frac{1}{(1+R)^n}\left(\frac{c}{R}\left((1+R)^n-1\right)+1\right)d(t,T_s)\right). \tag{47}$$

6 Data

The available data are historical discount curves corresponding to the risk of STIBOR SWAPs, Swedish Government Bonds and Statshypotek Bonds[1]. The curves corresponding to each set of instruments will hereby be referred to as the swap curve, treasury curve and mortgage curve respectively. The curves were bootstrapped using historical prices

[1]The Statshypotek Bonds are issued by Statshypotek AB and are covered by mortgage lending.

and cubic splines. The data ranges from 1998 to 2016 with a total of 4510 data points (approximately 250 data points per year). Thus for each date in the data set there are three curves corresponding to the interest rate term structure of that date. The longest maturity available for the swap, treasury and mortgage curve were 10, 10.5 and 2.5 years respectively. The curves corresponding to the last date of the data set are shown in Figure 1. The plots of the historical data are shown in Figures 20, 21 and 22. The swap curve corresponds to the risk of the STIBOR and the FRAs and swaps will be valued against this curve. The bond forward contracts are valued against the curve corresponding to the underlying deliverable bond such that bond forwards on treasury bonds are valued on the treasury curve and bond forwards on mortgage bonds are valued against the mortgage curve.

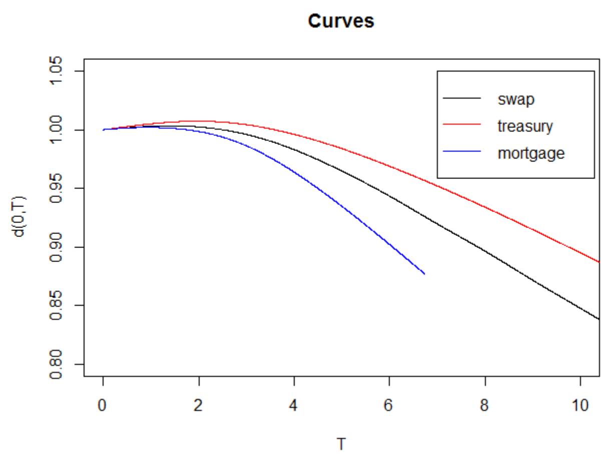

Figure 1: Swap, mortgage and treasury curve at 2015-12-30. Note that the curves are larger than 1 for short maturities, this is a consequence of negative interest rates.

6.1 Characteristic instruments

To investigate the FHS model on a realistic portfolio a set of characteristic instruments were defined. The fixed rates of the instruments were set such that the value of the instrument at the last corresponding curve were close to zero and the notional value were set to 10^6. For the Bond Forward contracts the coupon of the underlying bond were set to 0.06. The characteristic instruments were;

1. Sell side FRA with 3 months to settlement (3x3 FRA).
2. Sell side swap with 6 months to maturity (6M SWAP).

3. Sell side swap with 10 years to maturity (10Y SWAP).
4. Buy side Treasury Bond Forward with 6 months to settlement, the underlying bond has a 10 year maturity (0.5x10Y T-BF).
5. Buy side Mortgage Bond Forward with 6 months to settlement, the underlying bond has a 2 year maturity (0.5x2Y M-BF).

The value of each instrument defined above depends on their corresponding discount curve at a set of maturities $(T_1, ..., T_m)$. Let $d_i^k, k \in \{1, ..., n+1\}, i \in \{1, ..., m\}$ denote the value of $d(0, T_i)$ for the curve corresponding to date k. The relative change in the discount factor with maturity T_i between two days were then calculated as

$$\delta_i^k = \frac{d_i^{k+1} - d_i^k}{d_i^k}, \quad i \in \{1, ..., m\}, \quad k \in \{1, ..., n\}. \tag{48}$$

These changes in the discount factors were then applied to the last curve such that

$$\bar{d}_i^k = d_i^{m+1}(1 + \delta_i^k). \tag{49}$$

Further let $\Pi_p = \Pi_p(T_1, ..., T_m; d_1, ..., d_m)$ denote the value of the characteristic instrument number p as defined previously. The scenarios for each instrument were created by valuing the instrument using the discount factors $\bar{d}_i^k$ such that

$$\Pi_p^k = \Pi_p^k(T_1, ..., T_m; \bar{d}_1^k, ..., \bar{d}_m^k), \qquad k \in \{1, ..., n\}. \tag{50}$$

Thus a sample of size n for each instrument were created by valuing the instrument using discount factors created by multiplying historical changes in the discount factors with the latest discount factor. A sample of absolute losses for each instrument can now be determined as

$$L_p^k = \Pi_p^k - \widehat{\Pi}_p \tag{51}$$

where $\widehat{\Pi}_p$ denotes the value of instrument p at the last curve. The losses acquired using this model is shown in Figure 2. In the plot the loss for each date corresponds to how the value of the instrument would change if the historic relative discount factors of that date were applied to the last curve. Here it can clearly be seen that the volatility of the instruments seems to be time dependent. For example the fluctuations during the financial crisis of '08 can be seen to be a larger than "normal", especially for the instruments with short maturity. Furthermore one can observe large spikes in all of the samples. It is interesting to try to connect those spikes to global events, for example there is a large spike in the loss samples for the 3x3 FRA and the 6M SWAP at 2001-09-17, this spike is likely related to the 9/11 attacks in New York. It is probably possible to explain a lot of the other fluctuations observed in the samples. In Figure 23 the filtering of the 3x3 FRA loss sample using an EWMA 95 volatility process is illustrated and the properties of the distribution of the filtered losses are shown.

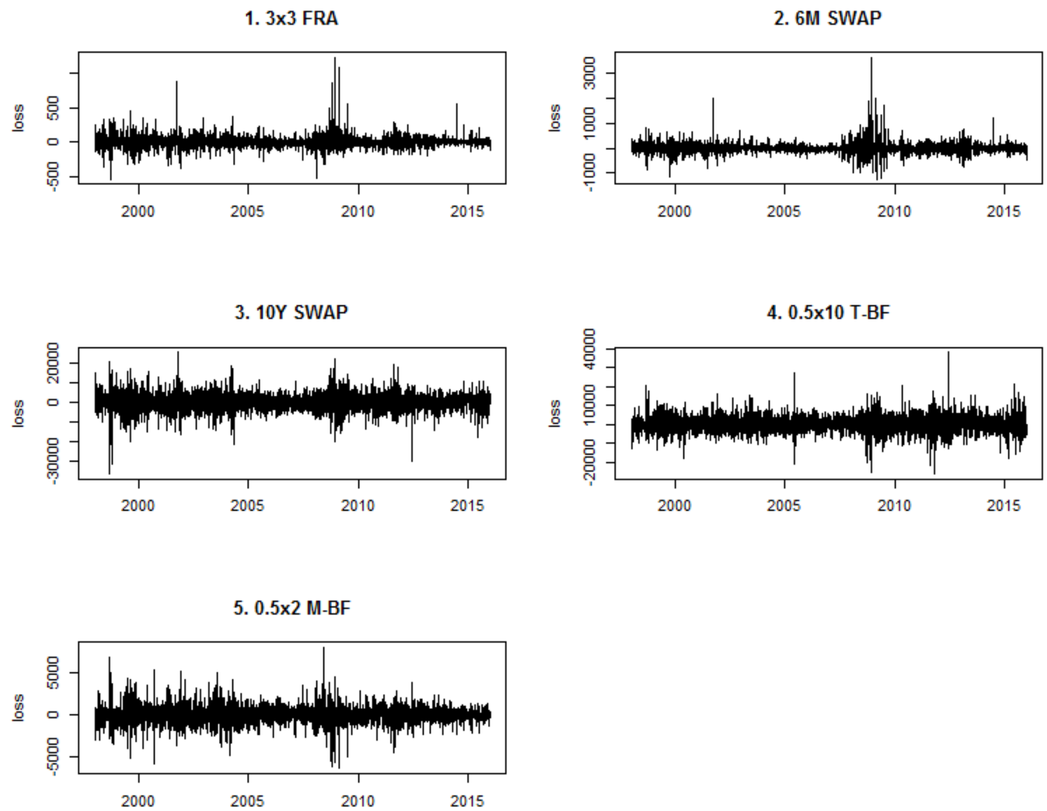

Figure 2: The losses of each characteristic instruments when historical relative changes in the discount factors were applied to the last discount curve. The x-axis corresponds to the date at which the historical relative changes occured.

6.2 Independence of the filtered losses

In this section the assumption that the filtered losses are *IID* is investigated using the properties of stationary time series. For a more information about time series and stationarity see for example [2]. The *autocorrelation function* (ACF) can be used to determine if there is any correlation between the lags of a sample. The ACF ρ of a stationary time series $\{X_t\}$ is defined as

$$\rho(h) = \frac{\mathrm{Cov}(X_t, X_{t+h})}{\mathrm{Var}(X_t)} \tag{52}$$

where h is the lag. If $\{X_t\}$ is *IID* the ACF will be 1 for $h = 0$ and 0 for $h \neq 0$. The sample ACF can be used to investigate whether a sample appears to be *IID* by plotting it, loosely speaking the assumption that the sample is *IID* can not be rejected if the sample ACF is reasonably close to 0 for $h \neq 0$. The sample ACF of the characteristic instruments are shown in Figure 24. Here it can be seen that the values are reasonable close to zero, however the results are not perfect. For the 3x3 FRA there seems to be a weak positive autocorrelation for the first ≈ 20 lags and for the 6M SWAP and the BF contracts there seems to be a relevant autocorrelation for the first lag. To remedy this it is possible to include an ARMA mean process in the filtering. However this might lead to over fitting historical data. In this thesis the losses for all instruments are assumed to have zero mean as shown in equation (7).

The ACF of the squared sample can be used to investigate if there appears to be any volatility clustering since large values for lags $h \neq 0$ implies that there are correlation between X_t^2 and X_{t+h}^2. The ACF of the squared sample for the characteristic instruments are shown in Figure 25. In Figures 26, 27 and 28 the ACF of the squared filtered losses obtained by filtering using GARCH, EWMA 95 and EWMA 99 volatility processes are presented. Here it can be seen that the volatility clustering is reduced significantly by filtering the losses. The GARCH and EWMA 95 processes seems to produce similar results while the EWMA 99 process seems to be a worse fit with the data (larger values of the ACF for $h \neq 0$).

7 Single instrument model

7.1 Backtesting

In this section the focus is to backtest different VaR models on portfolios consisting of a single instrument. In an implementation it might be better to use ES as the risk measure however the results from backtesting VaR are more intuitive, thus it will be easier to interpret the different models with VaR as the risk measure. To backtest the VaR methodologies described in Section 3 a lookback length l were selected such that the VaR at time t were determined by the loss sample $(L_{t-l-1}, ..., L_{t-1})$. Since the loss sample for each instrument is of size $n = 4509$ the amount of VaR values that can be calculated with the lookback length l is then $n-l$. To determine the one day VaR at time $t \in \{l+1, ..., n\}$ using the FHS model with EWMA and GARCH volatility estimates the following steps were made

- Filter the loss sample $(L_{t-l-1}, ..., L_{t-1})$ using a GARCH or EWMA filter to obtain a set of filtered losses $(Z_1, ..., Z_l)$.
- Order the filtered losses such that $Z_{1,l} \geq \cdots \geq Z_{l,l}$.
- Determine the empirical quantile estimate of the filtered losses using (4).
- The VaR at time t is then $\mathrm{VaR}_p^t = \sigma_t \widehat{F}_{l,Z}^{-1}(1-p) = \sigma_t Z_{[np]+1,n}$

In the HS approach no filtering were made. The results of performing this backtest with $l = 2500$ (10 year lookback period) and $p = 0.01$ are shown in Figures 3, 4, 5, 6 and 7. In Figure 8 the comparison between VaR and ES for the EWMA 95 model and characteristic instrument one is shown. It is obvious from the figures that the HS method does not react appropriately to increases in volatility, this is partly because of the long lokback sample. The GARCH and EWMA 95 models performed very similar for all portfolios based on the figures. The procyclicality of the FHS models can also be seen as when the portfolios start to show volatile behaviour the VaR increases drastically. In Table 7.1 the breaches of each model for the different instruments and the corresponding Kupiec test p-value are shown. Here it can be seen that the only model that can be rejected at any relevant confidence level is the HS model. Since there are 2009 VaR predictions in each sample and the VaR is calculated with $p = 0.01$ the expected number of breaches is 20. Even though the EWMA 95, EWMA 99 and GARCH models seems to produce good results it should be noted that these models generally exceeds the expected amount of breaches in this backtest. The parameter λ of the EWMA model also seems to control the procyclicality in the sense that an increased λ will lead to slower changes in the VaR.

	HS		EWMA 95		EWMA 99		GARCH	
	#	p-value	#	p-value	#	p-value	#	p-value
3x3 FRA	28	0.09412	25	0.28888	21	0.83949	22	0.67316
6M SWAP	39	0.00017	24	0.39496	23	0.52361	23	0.52361
10Y SWAP	21	0.83948	22	0.67316	22	0.67316	23	0.52361
0.5x10Y T-BF	41	0.00004	22	0.67316	26	0.20490	25	0.28888
0.5x2Y M-BF	8	0.00203	19	0.80518	19	0.80518	19	0.80518

Table 1: The # denotes the amount of breaches and the p-value is the Kupiec p-value.

7.2 Discussion

The EWMA model is a relatively simple model for estimating the volatility. However the difference between the EWMA model with $\lambda = 0.95$ and the GARCH model seems to

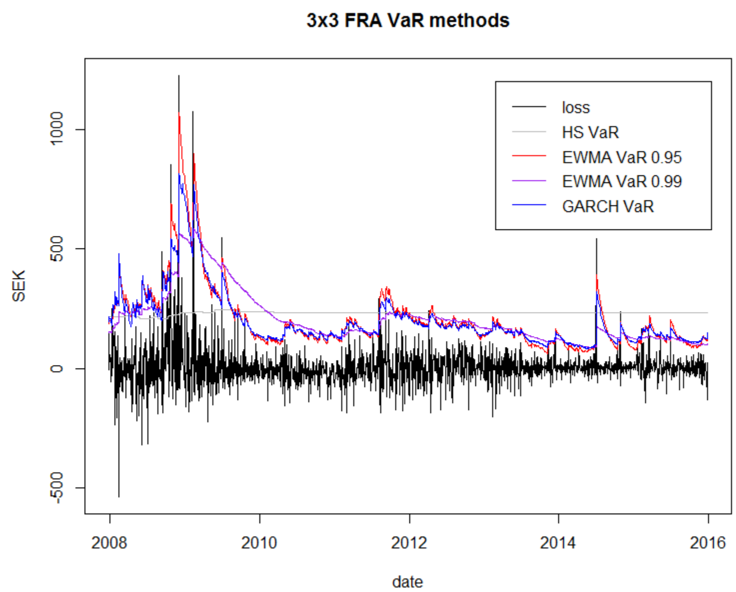

Figure 3: Comparison of different VaR models for characteristic instrument 1.

be small for all the characteristic instruments. Furthermore the GARCH model is a lot more computationally intensive than the EWMA model. Another benefit of the EWMA model is that it will produce transparent and intuitive results for all samples since the value of lambda is defined beforehand. The main drawback of the FHS models seems to be that extreme profits and losses will drastically increase the VaR. In some cases the VaR increases by as much as 400% from one day to the next, this is not reasonable and should be remedied in an implementation.

There are a lot of other models in the GARCH family such as EGARCH and GJR-GARCH and it is also possible to use an ARMA process to describe the mean of the losses in combination with these volatility processes. The use of other GARCH type models can be motivated by that the standard GARCH model does not capture asymmetry, a large loss affects the volatility estimate exactly the same as a large return. These models however introduce additional parameters to the filtering which further complicates the model. The parameters also becomes more difficult to estimate and problems such as convergence issues and over fitting might appear. It is also possible to use a more sophisticated estimate of the quantile such as one relying on Extreme Value Theory (EVT).

In this thesis the filtering is applied to the absolute losses of the instruments. Another approach is to apply the filtering on the discount curves before the valuation is done to produce a set of filtered returns of the discount factors. An example of this would be to

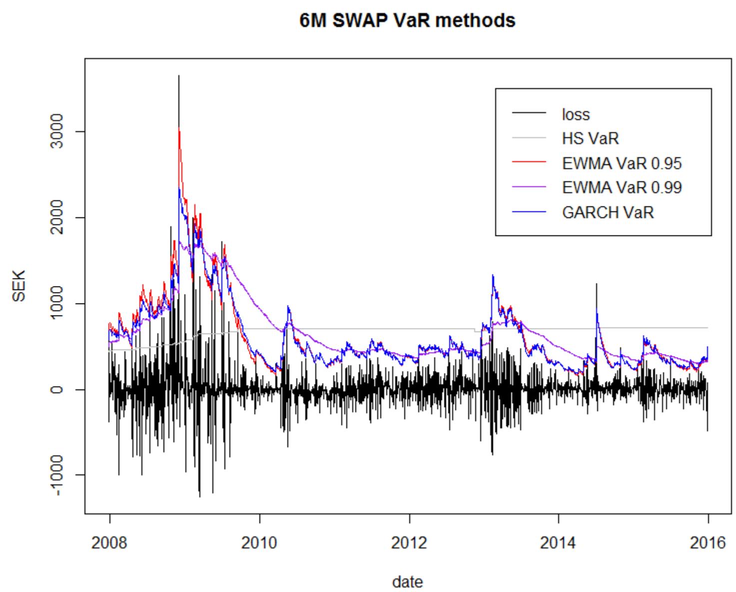

Figure 4: Comparison of different VaR models for characteristic instrument 2.

filter the relative changes in the discount factors as defined in equation (48) such that for a sample of $m+1$ historical curves and a set of fixed maturities $(T_1, ..., T_n)$ the delta is defined as $\delta_i^k = \sigma_i^k z_i^k, \quad k \in \{1, ..., n\}, \quad i \in \{1, ..., m\}$. The filtered relative changes are then applied to the last discount factors using the volatility estimates at time $m+1$ to produce a set of m scenarios where the portfolio is valued using the discount factors

$$\bar{d}_i^k = d_i^{m+1}(1 + \frac{\sigma_i^{m+1}}{\sigma_i^k} z_i^k). \tag{53}$$

If a portfolio depends on the discount function at a lot of maturities filtering must then be done for each maturity which can be very computationally intensive. It is also possible to fix a set of maturities for a given curve, filter the discount factors at these maturities, apply them on the latest curve and then use some interpolation technique to produce a discount function for each historic scenario $i \in \{1, ..., m\}$.

Furthermore the choice of how the returns are modeled is also interesting. It is possible to model the absolute returns, relative returns or log returns when filtering the changes in discount factors. When filtering the losses of the instruments this is not as straightforward since the instruments can have zero and negative values, in this thesis the absolute losses were therefore used. It is also possible to filter the spot rates corresponding to the discount

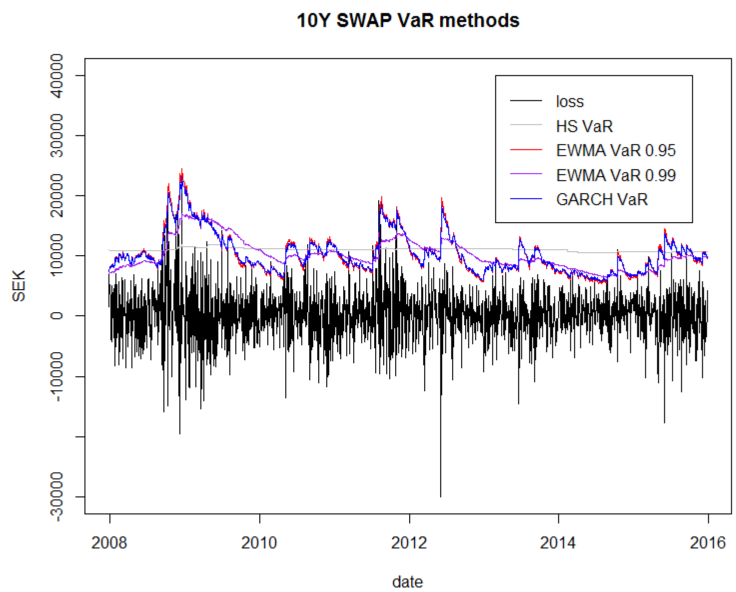

Figure 5: Comparison of different VaR models for characteristic instrument 3.

factors since they have a one to one relationship.

A summary of different modeling choices one has to make when implementing a FHS model is presented below. Not all of these modeling choices were investigated in this thesis.

- What type of volatility estimate to use (EWMA, GARCH, EGARCH, GJR-GARCH etc.) and how to determine the corresponding parameters.
- Filter the losses or the discount factors/spot rates.
- Length of lookback period.
- VaR or ES and the corresponding value of p.
- How to estimate the quantiles.

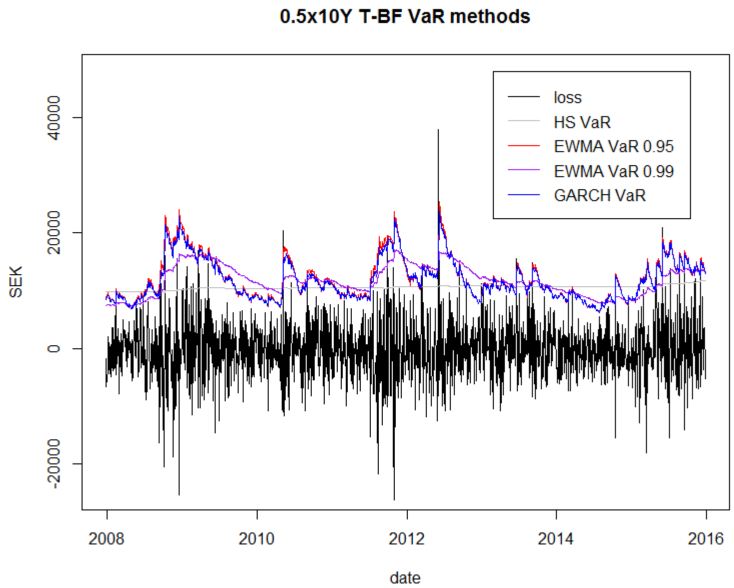

Figure 6: Comparison of different VaR models for characteristic instrument 4.

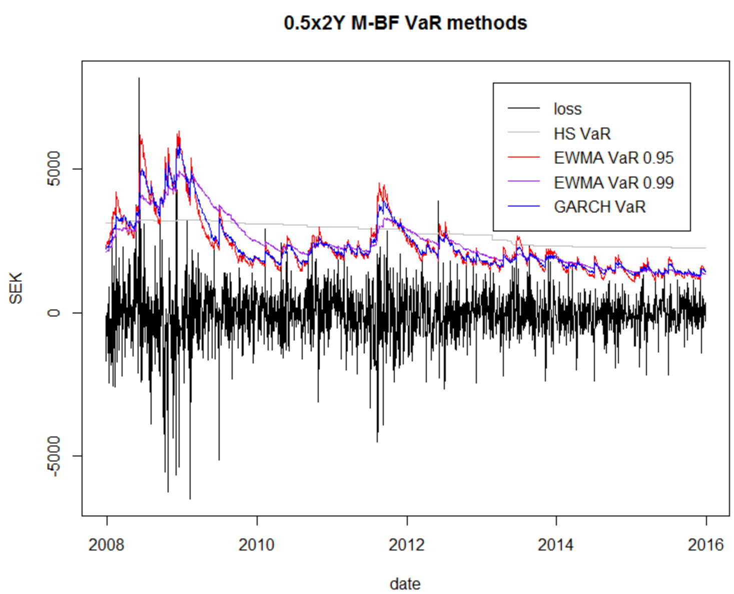

Figure 7: Comparison of different VaR models for characteristic instrument 5.

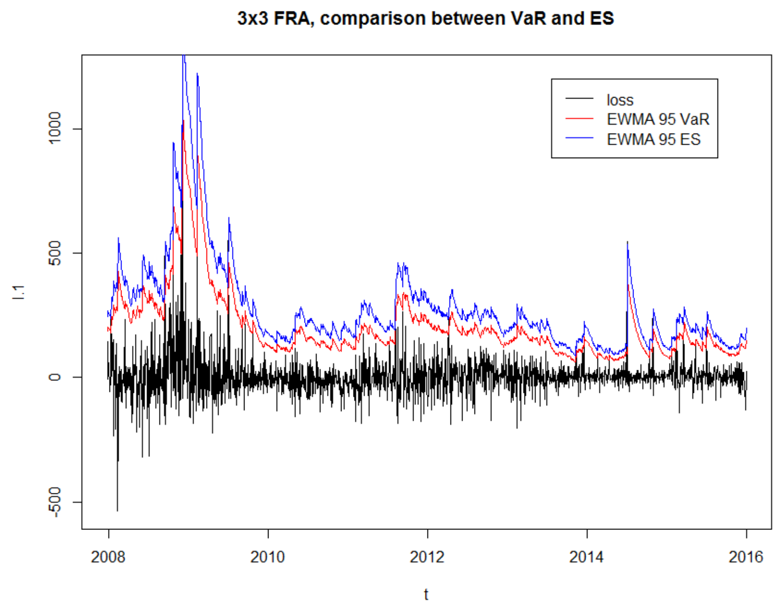

Figure 8: Comparison between VaR and ES with EWMA 95 filter for characteristic instrument 1.

8 Correlation model

When determining the IM of a portfolio a very important aspect is how to treat correlation between instruments. If two instruments seems to be correlated it could be reasonable to lower the IM of a portfolio consisting of these two instruments. To illustrate this consider a portfolio consisting of the characteristic instruments 3 and 4 as defined previously (Sell position in a 10Y swap and Buy position in 0.5x10Y T-BF). The aggregated losses of these instruments are shown in Figure 9. The portfolio consisting of instrument 3 + instrument 4 seems to be well hedged in the sense that the average intraday returns and losses are reduced significantly in size as a result of the correlation. If one calculates the VaR for each instrument separately and aggregates the VaR values the portfolio VaR will be a lot higher than necessary. Another approach is to aggregate the losses of each instruments and then determine the VaR of the aggregated portfolio. This approach will be referred to as aggregated VaR. The differences between these two approaches are shown in Figure 10. It is clear from the figure that it is reasonable to "give" correlation between instruments when determining the IM. However the approach where one simply aggregates the losses of the whole portfolio and then determines the VaR could be dangerous to use from the perspective of a CCP since there is no control of any correlation and a sudden "correlation break" where the correlation between some instruments change drastically will have severe effects on the produced VaR values. In a practical setting it is convenient to have control of the correlation between the instruments for example by having a parameter that describes it. In a practical implementation it is then possible to tweak these parameters, the resulting model will then not be a VaR measure but the result might be useful to determine the IM. It is the purpose of this section to describe a model that parametrises the correlation between the instruments.

In this section only the EWMA 95 model will be used to estimate the volatility since it produced reasonable results in the single instrument case and since it is significantly faster than the GARCH model.

8.1 Model

The model used to test the copula approach to determine the correlation between instruments is to filter the loss sample for each instrument and then assume that the filtered losses has a Student's t copula distribution such that

$$\begin{aligned} l_{1,t} &= \sigma_{1,t} z_{1,t} \\ &\vdots \\ l_{m,t} &= \sigma_{m,t} z_{m,t} \end{aligned} \tag{54}$$

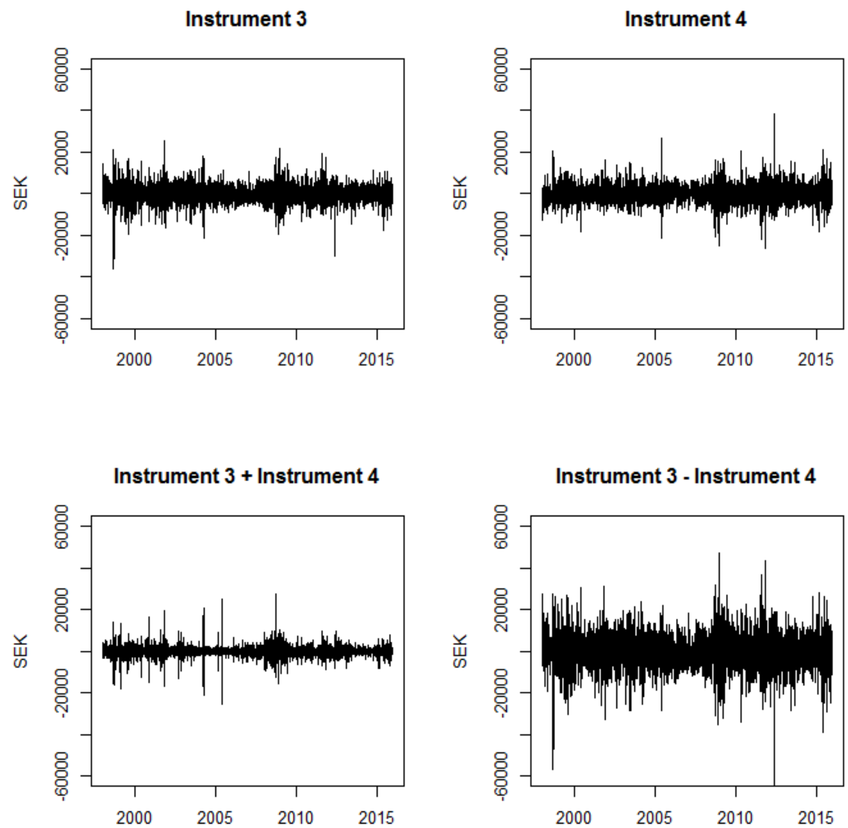

Figure 9: Upper plots shows the losses of characteristic instruments 3 and 4. The lower left plot shows the losses of instrument 3 + instrument 4 (hedged portfolio) and the lower right shows the losses of instrument 3 - instrument 4 (directed portfolio).

where $\sigma_{i,j}$ is an EWMA 95 process corresponding to the losses of instrument i and $z_{i,t}, \quad i \in \{1, ..., m\}$ are assumed to have a t copula. The empirical quantile were used to convert the uniform copula variables to the simulated filtered losses. The quantile of the joint distribution specified by the t copula is not available, to determine the VaR a Monte Carlo simulation method were used. To determine the VaR at time t each loss sample were first filtered using the EWMA 95 process, then the copula were fitted to the sample of filtered losses. The copula were then used to simulate a set of filtered losses which were then multiplied with their corresponding predicted volatility and aggregated to create a sample of simulated portfolio losses. The VaR were then determined by taking the appropriate quantile of the simulated sample.

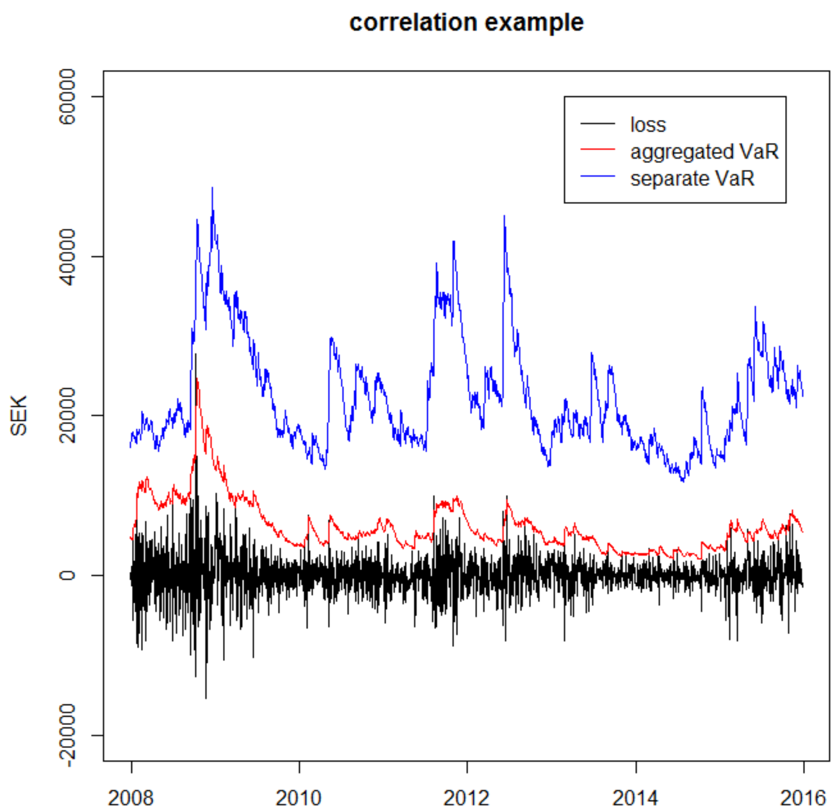

Figure 10: The portfolio consisting of characteristic instrument 3 + 4. The red line shows the EWMA 95 VaR of the aggregated portfolio losses and the blue line shows the aggregated EWMA 95 VaR of each instrument separately.

8.2 Correlation model example 1

To illustrate the correlation model consider again the portfolio consisting of characteristic instruments 3 and 4. Both of these instruments are exposed to long term interest rate structure, however they are valued against different curves. In Figure 9 it can be seen that these instruments are strongly correlated which implies that the long term movements of the swap and treasury curves are correlated. To measure this correlation the loss sample of each instrument were filtered using an EWMA 95 process and then the Kendall's tau were estimated for each t using the filtered losses in a one year lookback window. The result is shown in Figure 11. Here it can clearly be seen that the correlation is very strong between the instruments. However during the financial crisis of '08 there is a correlation break where the estimate of Kendall's tau suddenly increases to approximately -0.5 which is almost half of the "normal" value. The result of the correlation break is also evident in the loss sample as the average returns and losses starts to increase in size.

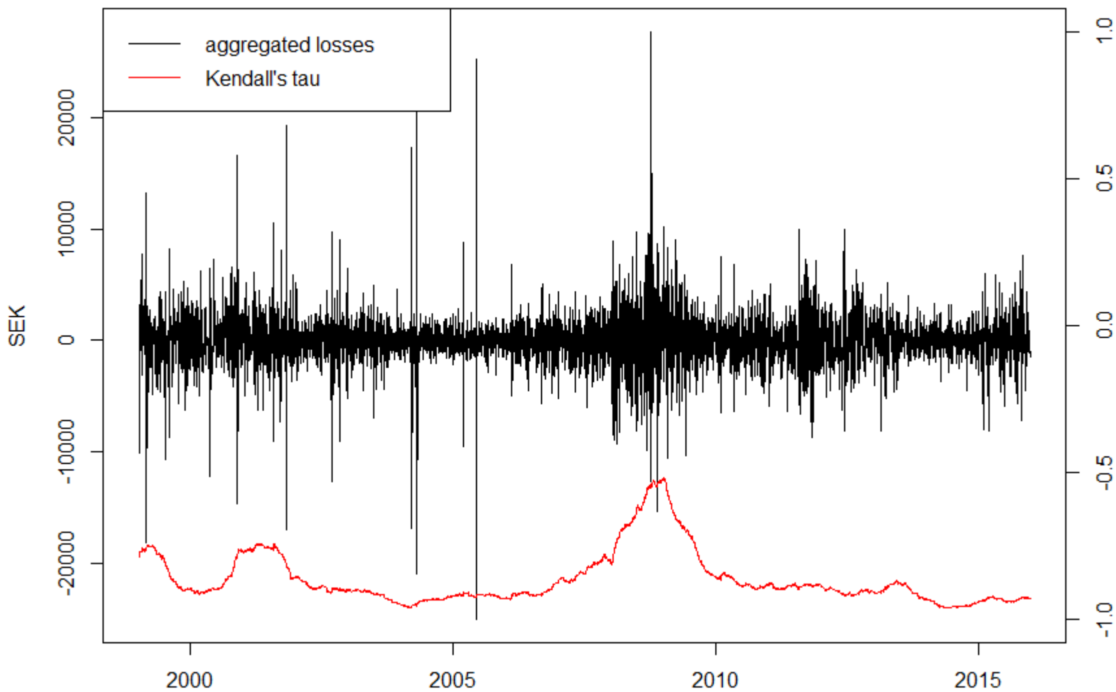

Figure 11: The aggregated losses of the portfolio as well as the one year lookback estimated Kendall's tau of the filtered losses. The values of Kendall's tau can be read using the values on the right axis.

In Figure 12 the filtered returns of the instruments are compared with those simulated from a Student's t copula estimated from the returns. The plots looks similar which indicates that the t copula is a good fit with the data. In Figure 13 the copula model is compared to the aggregated model using the EWMA 95 volatility filtering. The backtesting is once again done with a 10 year lookback and $p = 0.01$. The copula parameters were estimated every 50 time steps to increase computation speed since the estimated parameters were assumed to change slowly. The filtered losses were simulated 10000 times for each time step using the copula model. The amount of breaches for the copula and aggregated VaR were 23 and 22 respectively (expected amount of breaches is 20).

It should be noted that there are no exponential weighing for the correlation matrix since the correlation is assumed to be constant using the copula model. Thus it will react slow to potential correlation breaks. For example to determine the correlation matrix at Jan 1'st 2009 the losses during 2008 will only be 10% of the sample since the lookback period is 10 years.

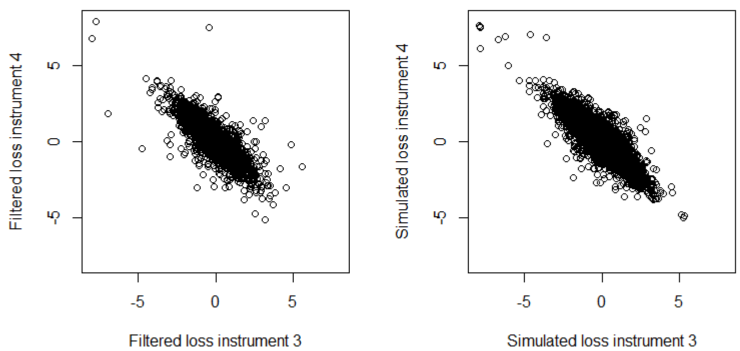

Figure 12: The left plot shows the filtered historical losses of instrument 3 and 4 and the right plot shows the filtered losses simulated from a t copula.

8.3 Correlation model example 2

The portfolio investigated here will consist of;

- 40 units of characteristic instrument 1,
- -20 units of characteristic instrument 2,
- 1 units of characteristic instrument 3,
- 1 units of characteristic instrument 4,
- 4 units of characteristic instrument 5.

Here the weights of each instrument were selected such that the average losses and returns were of similar size. Figure 14 shows the filtered losses of each instrument plotted against each other. Here we see that the instruments valued on the swap curve with short maturity

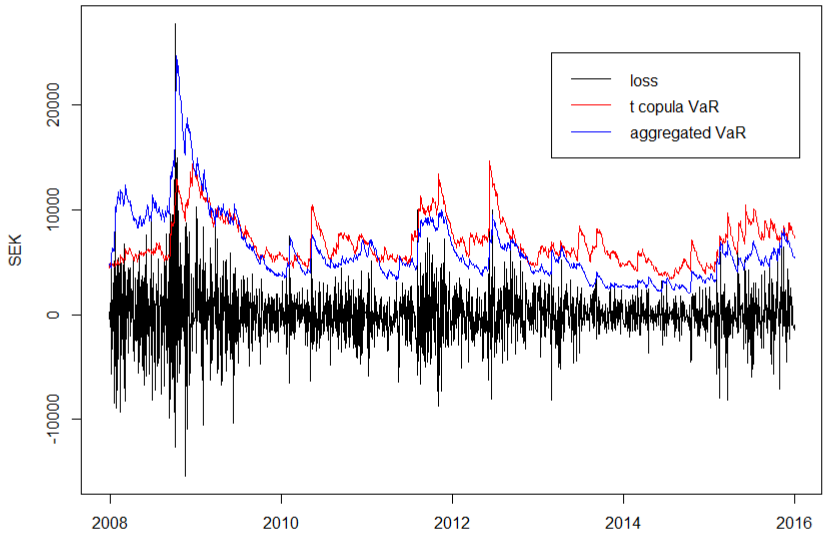

Figure 13: The red line is the VaR obtained by simulating from a t copula. The blue line is the VaR obtained by aggregating the losses of the instruments and then filtering. For both methods the EWMA 95 volatility estimates were used.

(1 and 2) seems to be strongly correlated with each other but not with the other instruments. Furthermore all other instruments (3, 4 and 5) show signs of correlation with each other. In Figure 15 the t copula VaR is compared to the aggregated VaR for the portfolio. All filtering is once again done with the EWMA 95 model. The amount of breaches is 12 for the t copula VaR and 16 for the aggregated VaR. This should be compared with the expected amount of breaches which is 20.

Note that the 3x4 simulated correlation in Figure 14 does not seem to be as accurate as the one in Figure 9. This is because the degree of freedom parameter of the copula is optimized to be a good fit with all of the pairs and not only with respect to the 3x4 pair. Thus when the dimension of the copula is increased the model accuracy of individual pairs are decreased.

8.4 Correlation model example 3

In the example portfolio discussed in the previous section some of the instruments showed strong negative correlation which indicates that the portfolio is well hedged. In this example

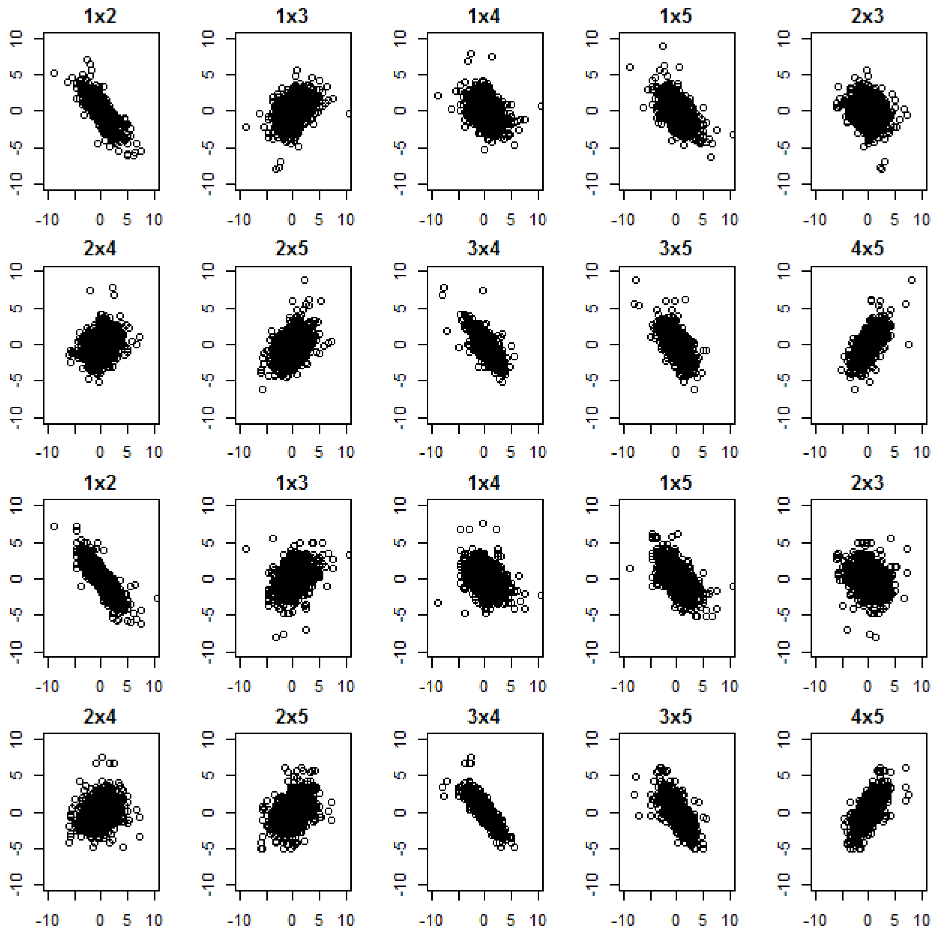

Figure 14: Filtered losses of each characteristic instrument against each other. The 10 upper plots are the real values and the 10 lower plots are the corresponding simulated values.

the sign of some instruments will be changed to illustrate what happens when the portfolio is directed. The portfolio is now consisting of;

- 40 units of characteristic instrument 1,
- 20 units of characteristic instrument 2,
- -1 units of characteristic instrument 3,
- 1 units of characteristic instrument 4,
- 4 units of characteristic instrument 5.

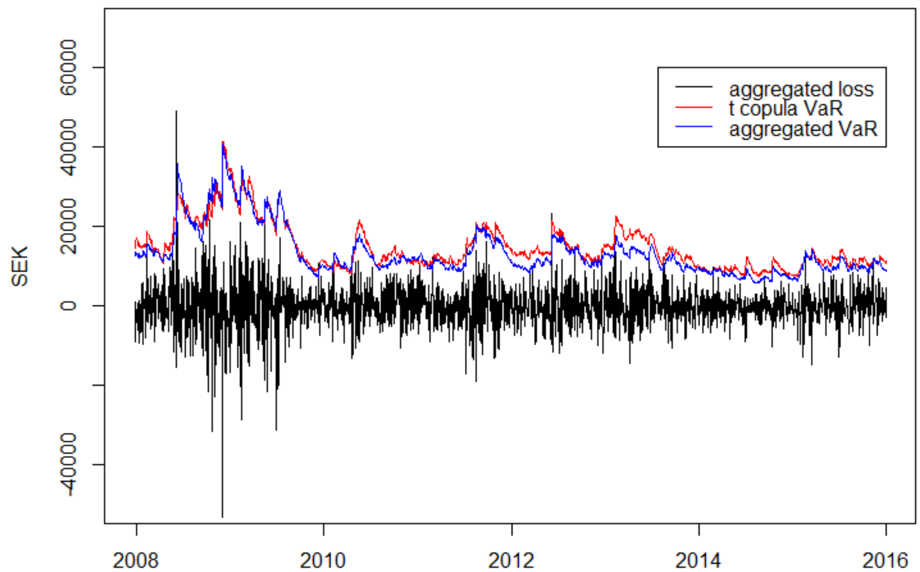

Figure 15: Copula VaR compared with aggregated VaR.

In Figure 16 the same results as in 15 is shown but for the new portfolio. The amount of breaches is 27 for the t copula VaR and 26 for the aggregated VaR. This should be compared with the expected amount of breaches which is 20.

8.5 Correlation model discussion

The FHS approach to determining the VaR of a portfolio seems to produce reasonable results for the purpose of determining the IM of a portfolio. The benefits of the model is that the results are intuitive and relatively simple to understand. However if one were to simply aggregate the losses of an entire portfolio and then use FHS there is no protection against correlation breaks between instruments which could have a great impact on the potential losses of the portfolio. Furthermore there are no parameters or results of the aggregated model that one can interpret to understand the correlation effects on the VaR of a portfolio. As a CCP it is essential that these properties are well understood. The lack of control over the correlation is a major drawback of the aggregated VaR model. The advantage of the copula approach is that correlation can be given between instruments or

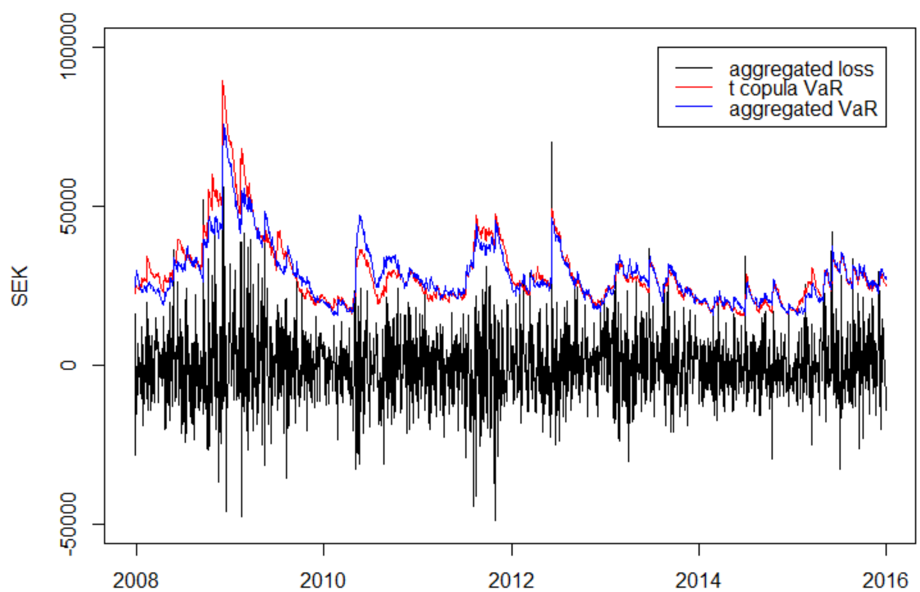

Figure 16: Different copula models compared with the aggregated VaR.

risk groups in a controlled setting. For example it is possible to fix the correlation matrix to a stressed correlation matrix, such as one with the worst possible correlations for each instrument based on historical data.

It is also possible to combine the aggregated and copula models by dividing instruments into *risk groups*. Then the losses of the instruments in a certain risk group is aggregated before the copula model is applied. The advantage of this would be that the dimension of the copula can be decreased. Then full correlation is given within the same risk group and correlation determined by the correlation matrix of the copula is used to give correlation between the risk groups. It seems reasonable to group instruments that are valued on the same curve and are exposed to the same maturities, such as characteristic instruments 1 and 2 defined previously. Both these instruments are exposed to short maturities on the swap curve, and a correlation break between those instruments would loosely speaking imply unrealistic changes in the curvature of the swap curve. It is likely that the dimension of the copula could be reduced significantly without loosing control of the correlation between instruments by grouping instruments that have similar exposure on the same curve.

This method of dividing a portfolio into risk groups are also possible to use for the aggregated model, then the losses of instruments within each risk group are aggregated and a VaR is determined for each risk group. The VaR values for the risk groups are then aggregated to determine the portfolio VaR. Using this approach no correlation will be given between instruments in different risk groups and full correlation will be given between risk groups.

Since different instruments can have different correlation structure the copula model used were restricted to elliptic copulas. Thus the copulas considered were Gaussian (Normal) and Student's t copulas. The reason that the Student's t copula was selected over the Gaussian copula was made partly by observing the plots of filtered losses simulated from the corresponding copula where the Student' t copula seemed to be a better fit with the data, especially in the relevant areas corresponding to extreme losses. Furthermore the results of the copula model in the backtesting were reasonable.

In this thesis the backtesting only tests the copula approach in max 5 dimensions. The implications of increasing the dimension even further are not investigated here. Although the pseudo MLE estimation procedure for the parameters should perform reasonable as the dimensions increase. The estimation comes down to estimating Kendall's tau for each pair of variables which can be done $\mathcal{O}(n \log n)$ [3] and finally estimating the degrees of freedom which is a simple one dimensional optimization problem. A problem that will arise by increasing the dimension is that the degree of freedom parameter ν of the Student's t copula will limit the dependence structure between individual variables since different pairs will likely have different optimal values of ν. A topic for further research could be to modify the t copula as specified in this thesis to have individual degree of freedom parameters for each dimension. It could also be interesting to investigate the use of vine copulas. Furthermore the amount of simulations required to get stable VaR estimates using the copula model were not explored.

The advantage of the copula approach is that correlation can be given between instruments in a controlled setting. Even though the model uses parameters to describe the correlation it produces results that are similar to the aggregated VaR model. Contrary to the aggregate model the copula model is easily manipulated to account for scenarios such as correlation breaks. The parameters of the copula model are fairly intuitive in the sense that the implications on the VaR of increasing/decreasing a parameter is easily understood. However, a drawback of using the copula approach is that it will reduce the transparency and intuition of the model for people who are not familiar with the concept of copulas.

9 Possible model adjustments

9.1 Procyclicality

The volatility estimate is the largest source of procyclicality in the model. Since all the volatility models discussed are proportional to the square of the previous loss a large loss will result in a significant increase in the predicted volatility. It can be seen in Figures 3, 4, 5, 6 and 7 that the EWMA 99 model appears to be less procyclical than the GARCH and EWMA 95 models. However the EWMA 99 model seems to be worse than the other models in reducing the volatility clustering of the losses as can be seen in Section 6.2. It is possible to decrease the procyclicality by selecting a larger λ for the EWMA model but in doing so the model will be a worse fit with the data and it might be wiser to decrease the procyclicality by other methods.

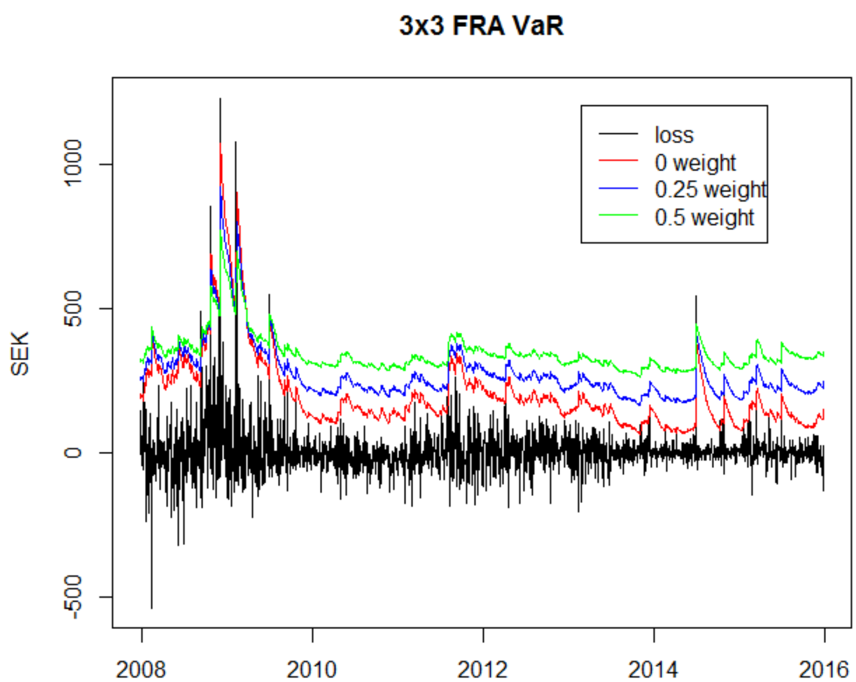

Figure 17: Result of weighing the predicted volatility with a stressed volatility to decrease the procyclicality for characteristic instrument 1.

There are a lot of possible ad-hoc approaches that can be used to decrease the procyclicality of the model. In the ESMA report [11] three options on how the IM model can be adjusted is presented;

- *The CCP shall apply a margin buffer at least equal to 25% of the calculated margins which it allows to be temporarily exhausted in periods, where calculated margin requirements are rising significantly.*
- *The CCP shall assign at least 25% weight to stressed observations in the look-back period calculated in accordance with Article 26.*
- *The CCP shall ensure that its margin requirements are not lower than those that would be calculated using volatility estimated over a 10 year historical look-back period.*

The first option is straightforward but the other two options can be interpreted in different ways. One way to interpret the second option is to weigh the prediction volatility used to determine the VaR with a volatility estimate from a period of stress. Note that only the prediction volatility is weighted and that the volatility used to filter the losses stay the same. Here the prediction volatility refers to σ_T in equation (8). In Figure 17 the results of this for characteristic instrument 1 is shown. Here the prediction volatility is weighted as

$$\sigma_t = (1 - w)\sigma_t^e + w\sigma^s \tag{55}$$

where $w \in [0, 1]$ is the weight, σ_t^e is the standard EWMA prediction volatility and σ_t^s is the unweighted sample standard deviation estimate during the stress period of '08. The results of applying this technique to the copula model is shown in Figure 18 (the portfolio is the one described in Section 8.2). These results shows that this type of model adjustment can be used to decrease the procyclicality of the VaR estimate.

The third option can be interpreted as that the VaR should be set as $\max(\text{VaR}_{HS}, \text{VaR}_{FHS})$ where VaR_{HS} is the ordinary 10 year HS VaR and VaR_{FHS} is determined by one of the FHS VaR models. Consider the maximum of the different FHS VaR models and the HS VaR in Figures 3, 4, 5, 6 and 7.

Furthermore some protection should be provided against sudden extreme losses or returns since these will drastically increase the predicted volatility and thus the VaR estimates. It is not reasonable for the IM of a portfolio to increase by 400% between two days which is the case for some backtests. A limit on the maximum allowed increase of the volatility between two days is one simple way to account for this.

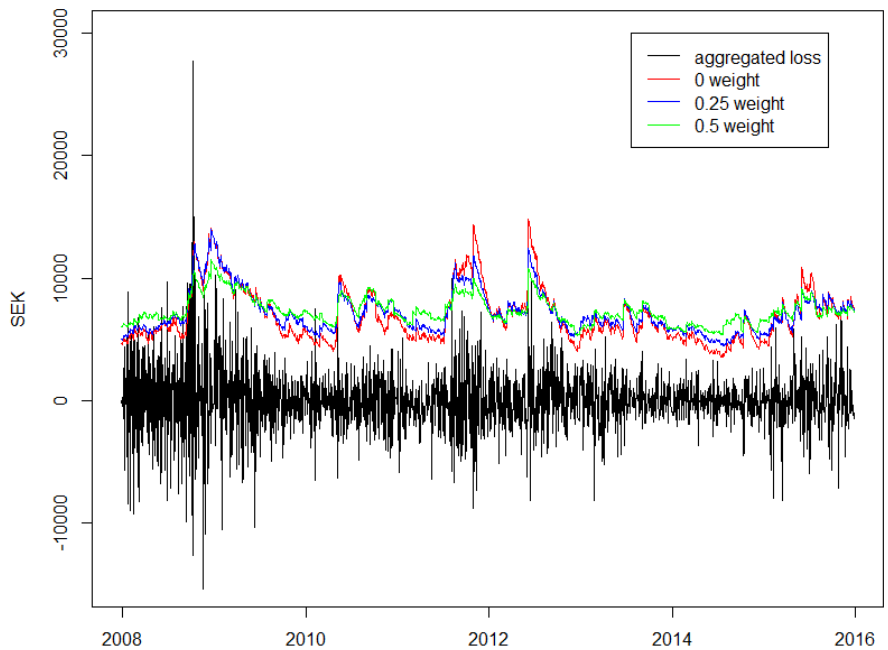

Figure 18: Result of weighing the predicted volatility with a stressed volatility to decrease the procyclicality for the portfolio described in Section 8.2.

9.2 Correlation

Consider Figure 11 where the one year lookback estimate of Kendall's tau is shown. In Figure 19 the results of fixing the copula parameters for all VaR estimates to those estimated using losses during '08 for the portfolio defined in Section 8.2 are shown. The result suggests that the stressed copula parameters increases the VaR of the portfolio which is expected. This reflects the fact that the correlation given between instruments can be manipulated in an intuitive way using the copula model to protect a portfolio against potential correlation breaks.

It should be noted that using correlation from periods of stress will not always increase the VaR. In the example presented here the VaR is increased since the correlation parameter were significantly higher during the stress period.

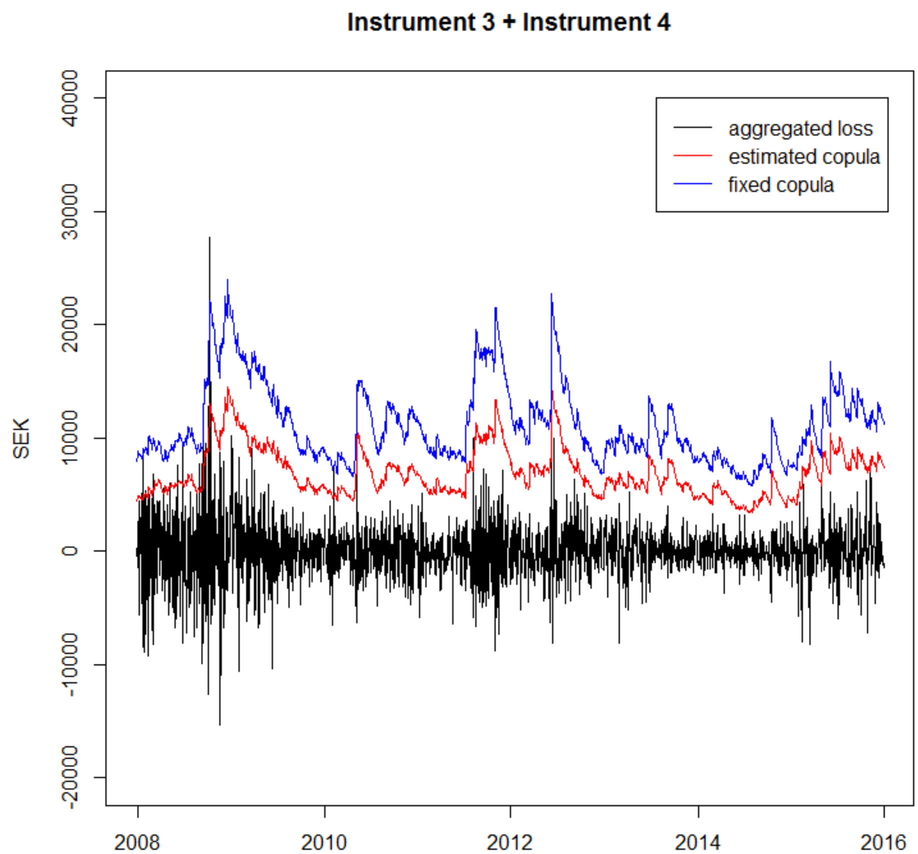

Figure 19: Comparison between using estimated copula parameters (red line) and fixed parameters estimated during '08 (blue line).

10 Conclusions

In this thesis the problems of using a VaR based model for determining the IM of a Fixed Income portfolio are investigated. Some important issues are discussed and explored. Furthermore some models are defined and backtested using historical data.

There are a lot of different issues one has to deal with when implementing a VaR model for determining the IM of Fixed Income portfolios. The main problems one has to face when implementing FHS model for fixed income can be summarized as;

- How to bootstrap the curves.
- What model should be used to determine the value of the instruments.
- What variable should be filtered (sample of losses or a selected risk factor, eg. discount factors or spot rates).
- What time series should be used to perform the filtering.
- How to model the effects of correlation within a portfolio.
- How the model should be adjusted to account for regulations and market concerns.

In this thesis some models has been investigated and tested to illustrate their properties. The bootstrapping of the curves and effects of using different valuation models has not been investigated. The main conclusions one can draw from the results is that filtering seems to be necessary to account for volatility clustering. Both the GARCH and EWMA models seems to be reasonable models to capture this phenomenon. These models performed reasonable in the backtests and are easily adjusted to account for procyclicality. However, extreme losses and returns during calm periods will drastically increase the VaR estimate the following day and appropriate measures should be taken to prevent extreme VaR shocks. The main issue of using FHS models to determine the IM of a portfolio is to achieve an appropriate trade off between reacting to changing volatility and procyclicality, and care is needed to calibrate the model to accomplish this.

It is also important to understand how correlation is given between instruments and that this is done in a controlled manner. For the small size portfolio examples illustrated in this thesis the Student's t copula performed reasonable for this purpose. The main drawback of this model is that the degree of freedom parameter of the copula limits the accuracy of the fit between individuals pairs when the dimension is increased. Furthermore the model is not as transparent and intuitive as the aggregated FHS model. There are a lot of different types of copulas that could be interesting to investigate further such as the Student's t copula with individual degrees of freedom or different types of vine copulas.

The figures from the backtests comparing copula VaR with aggregated Var suggests that the dynamics of the two models seems to behave similarly. This implies that the accuracy lost from the parametrization of the correlation in the copula model is relatively small. The benefits of using the copula model is that the parameters of the copula can be interpreted and adjusted to account for market concerns. An example of this is to fix each correlation parameter in the correlation matrix of the copula to a stressed parameter.

In this thesis only five different instruments are used to backtest the different models. The only instrument classes that are considered are FRAs, swaps and bond forwards. To extend the models to general fixed income portfolios the models should be backtested using a more general set of instruments and portfolios. Additionally great care needs to be taken when selecting and calibrating the parameters of the model such that the implications of different market scenarios are well understood.

A model for determining the IM of a portfolio should be robust, transparent and it should not encourage market instability. The general concept of using FHS in combination with a copula seems to be a reasonable model for determining the IM under these restrictions. Even though one is not familiar with the concept of copulas the model provides intuitive parameters that can be tuned to account for market related concerns such as correlation breaks and procyclicality.

A Appendix

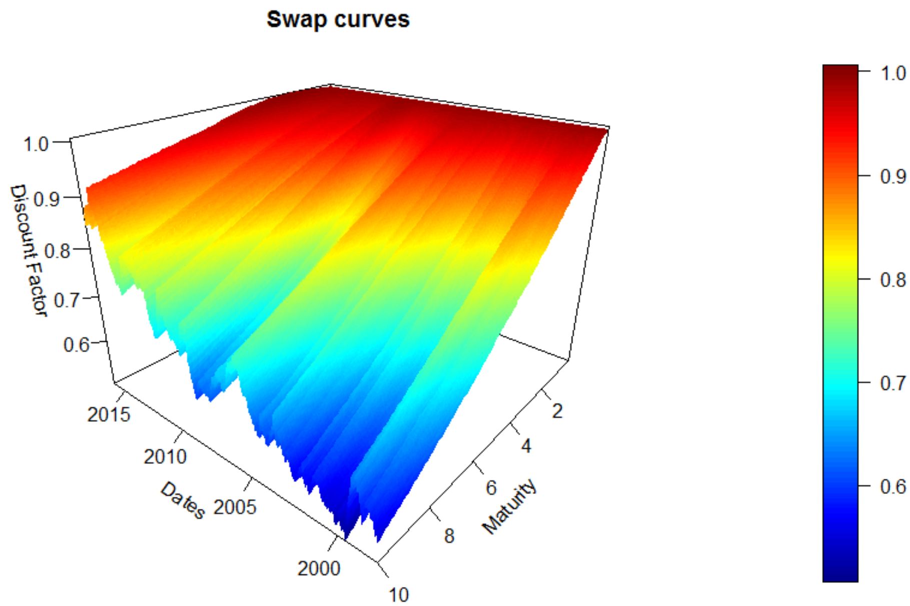

Figure 20: Historical swap curves.

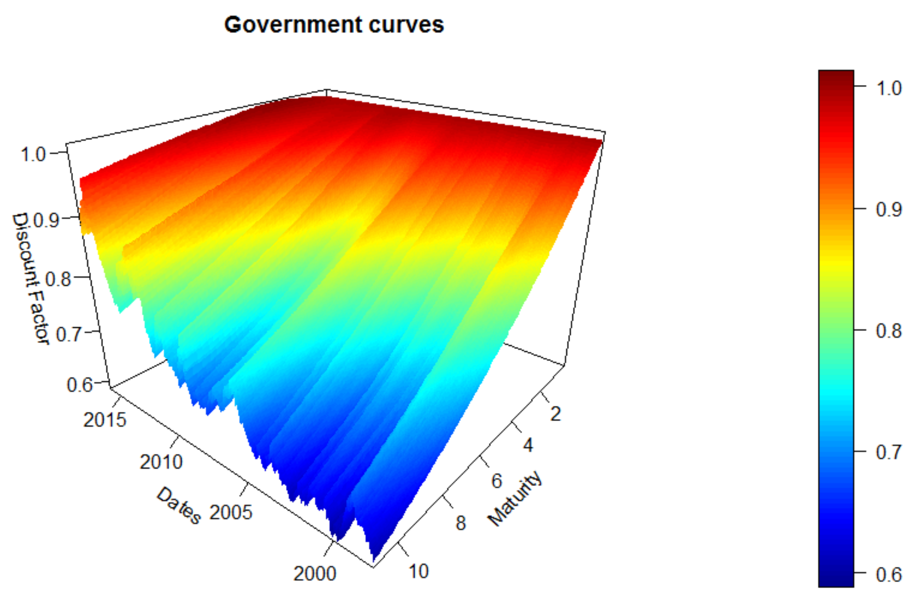

Figure 21: Historical government curves.

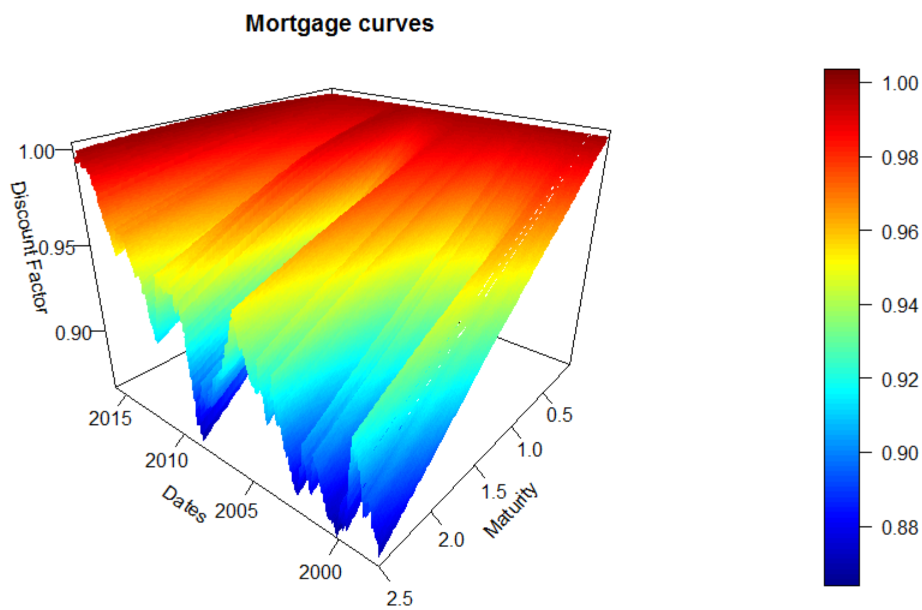

Figure 22: Historical mortgage curves.

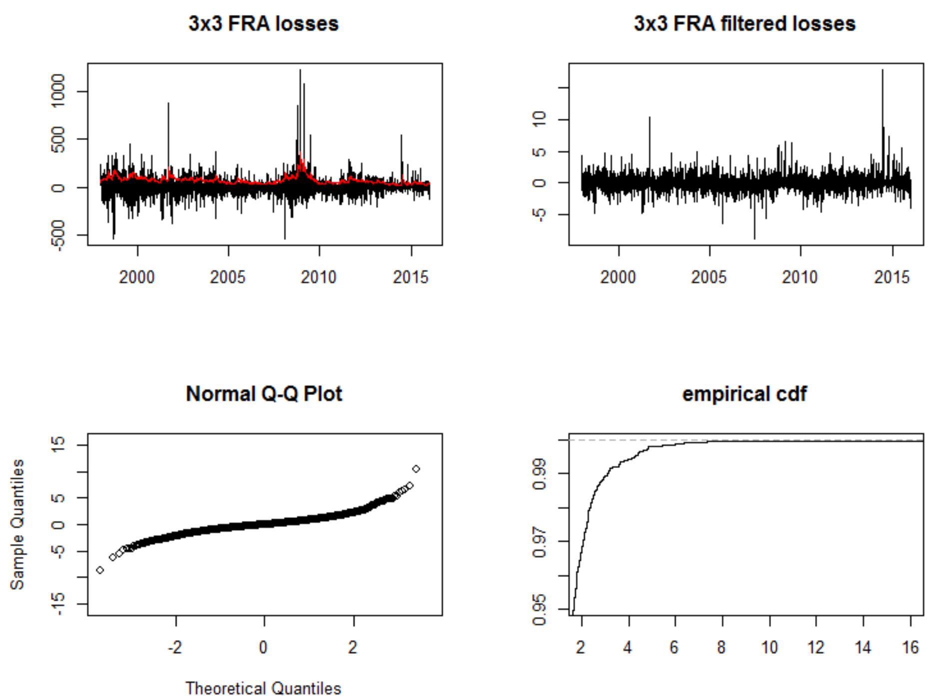

Figure 23: Properties of the 3x3 FRA sample. Upper left plot shows the losses and the red line shows the volatility process. The upper right plot shows the corresponding filtered losses. The lower left plot shows the normal QQ plot of the filtered losses and the lower right plot shows the right tail of the empirical cdf of the filtered losses. The filtering were done using an EWMA 95 process.

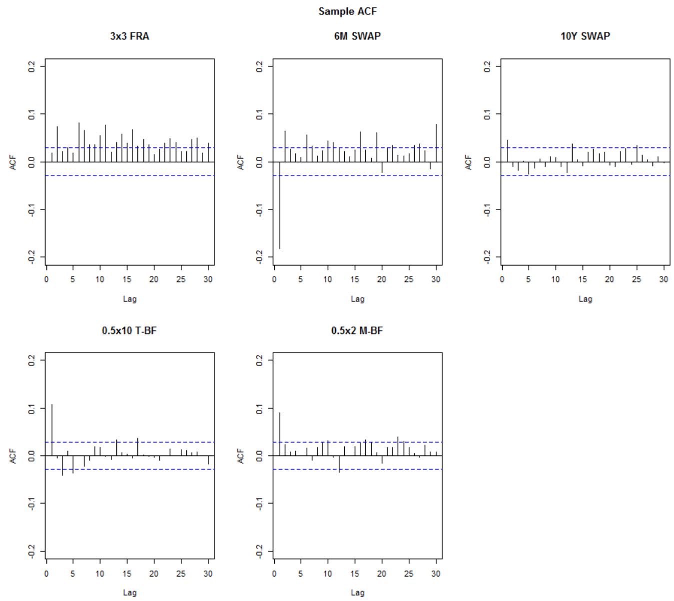

Figure 24: Sample ACF of the losses of the characteristic instruments.

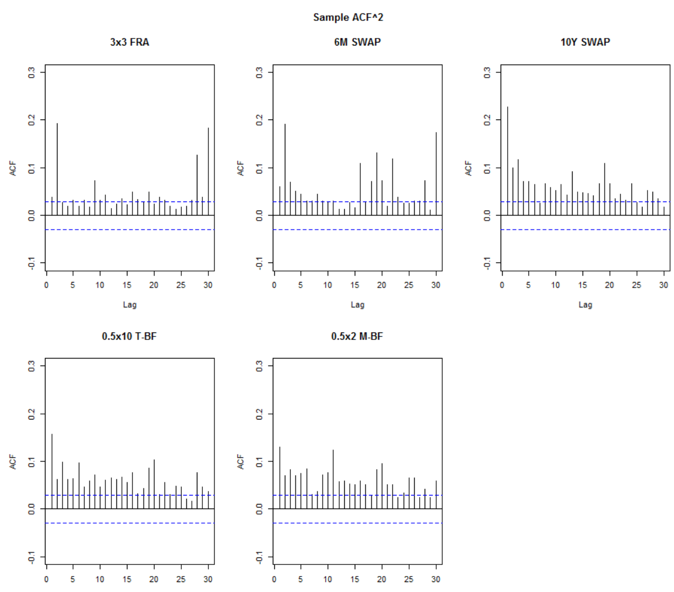

Figure 25: Squared ACF of the losses of the characteristic instruments.

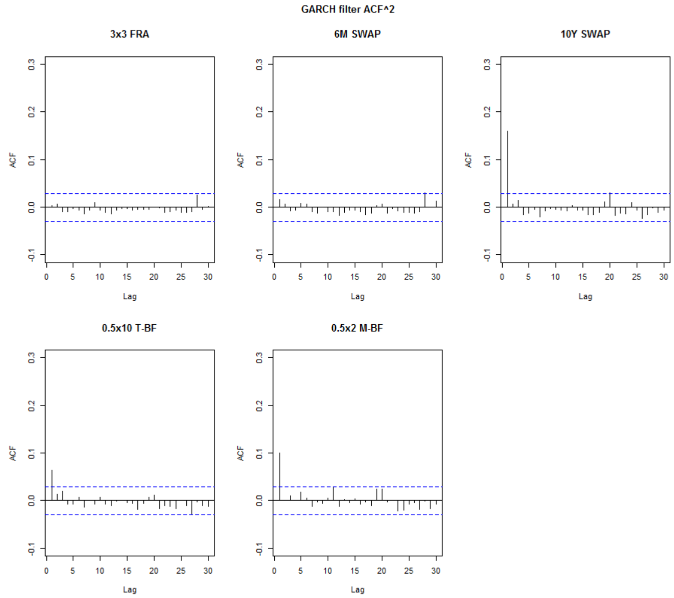

Figure 26: Squared ACF of the filtered losses of the characteristic instruments using a GARCH filter.

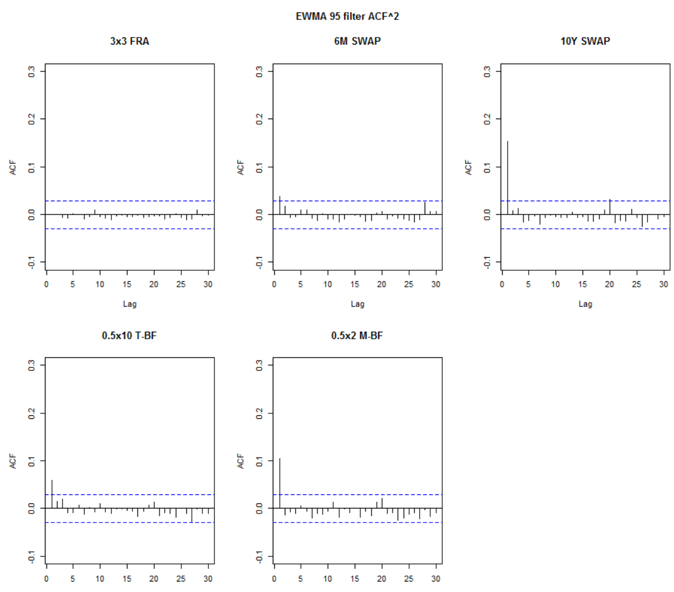

Figure 27: Squared ACF of the filtered losses of the characteristic instruments using a EWMA 95 filter.

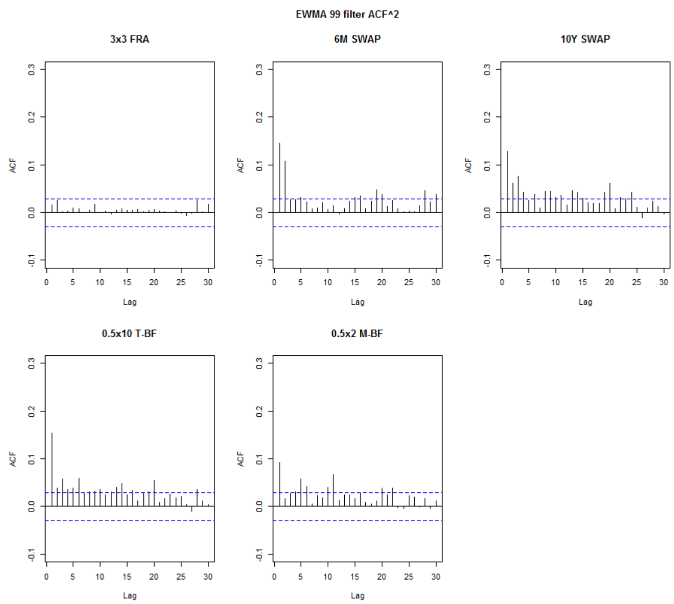

Figure 28: Squared ACF of the filtered losses of the characteristic instruments using a EWMA 99 filter.

References

[1] T. Björk. *Arbitrage Theory in Continuous Time*. Oxford Finance Series. OUP Oxford, 2004. ISBN: 9780199271269.

[2] P.J. Brockwell and R.A. Davis. *Introduction to Time Series and Forecasting*. Introduction to Time Series and Forecasting v. 1. Springer, 2002. ISBN: 9780387953519.

[3] D. Christensen. "Fast algorithms for the calculation of Kendall's tau". In: *Journal of Computational Statistics* 20 (2005).

[4] A. McNeil, R. Frey. "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach". In: *Journal of Empirical Finance* 7 (2000).

[5] A. Ghalanos. *rugarch: Univariate GARCH models.* R package version 1.3-6. 2015.

[6] H. Hult, F. Lindskog, O. Hammarlid, and C.J. Rehn. *Risk and Portfolio Analysis: Principles and Methods*. Springer Series in Operational Research and Financial Engineering. Springer Science+Business Media New York, 2012. DOI: `10.1007/978-1-4614-4103-8_1`.

[7] JPMorgan/Reuters. *RiskMetrics Technical Manual*. 1996.

[8] P. Kupiec. "Techniques for Verifying the Accuracy of Risk Measurement Models". In: *Journal of Derivatices* 3 (1995).

[9] S. Demarta, A. McNeil. "The t Copula and Related Copulas". In: *International Statistical Review* 73 (2005).

[10] P. Gurrola-Perez, D. Murphy. "Filtered historical simulation Value-at-Risk models and their competitors". In: *Bank of England Working Paper No. 525* (2015).

[11] EMIR Review Report no.2. *Review on the efficiency of margining requirements to limit procyclicality*. 2015.

[12] D. Murphy, M. Vasios, N. Vause. "An investigation into the procyclicality of risk-based initial margin models". In: *Bank of England Financial Stability Paper No. 29* (2014).

[13] J. Hull, A. White. "Incorporating volatility updating into the historical simulation method for Value-at-Risk". In: *Journal of Risk* 1 (1998).

[14] J. Hull, A. White. "LIBOR vs. OIS: The derivatives discounting dilemma". In: *Journal of Investment Management* 11 (2013).