Erasmus University Rotterdam

Department of Econometrics

Master Thesis: Impact of parameter uncertainty on Value-at-Risk estimates when few data is available

E.T. Goedegebure
374875

Supervisors:	C. Zhou, H.J.W.G. Kole and A. van Oord
Master Programme:	Econometrics & Management Science
Specialization:	Quantitative Finance
Course:	Master Thesis

Rotterdam 2016

Abstract

In this thesis I investigated the impact of parameter uncertainty on the 99.5 percent one year Value-at-Risk (VaR), in particular the impact of having only a few years of available data. In a simulation study where the underlying data generation process is known, I applied estimation methods that do and do not take into account the effect of parameter uncertainty to show if, and by how much these estimation methods differ from the known VaR. The data generating processes I used vary from a Gaussian copula to a Clayton copula. Within the simulation study I differentiated between the impact of parameter uncertainty on the VaR for the correlation, variances and Clayton copula parameter. Finally, I applied the different estimation methods on the return series used for the Solvency II calibration of the Solvency II Capital Requirement for equity investments. I found that including parameter uncertainty results in a smaller probability of underestimating the true 99.5 percent one year VaR and less economic impact of underestimating the true 99.5 percent one year VaR as compared to VaR estimates obtained using Maximum Likelihood Estimation. Furthermore, incorporating parameter uncertainty, within the context of Solvency II equity module, resulted in a more strict Solvency II Capital Requirement.

Keywords: Value-at-Risk, Parameter Uncertainty, Simulation Study, Solvency II

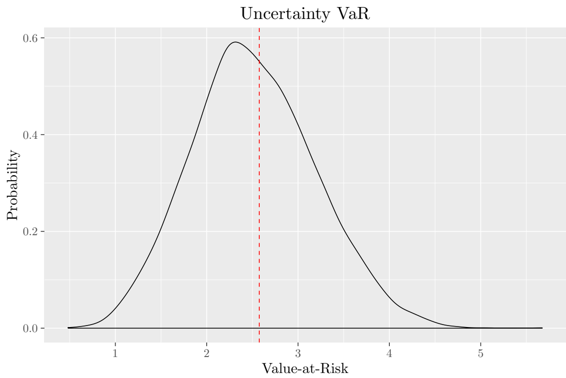
Contents

1 Introduction

Capital requirements in general are based on VaR and typically calibrated on few available data (e.g. Solvency II and IAIS standards). In these circumstances parameter uncertainty and sample size are particular relevant because it influences the amount of capital investors and insurers have to hold. An inaccurate capital requirement can therefore distort the balance between investing capital for profitability on the one hand, and maintaining enough capital to account for possible capital shocks on the other hand.

Estimating a VaR with just a few observations, results in an uncertain the VaR estimate. For example, figure 1 shows a density plot of VaR estimates obtained in a simulation study. Each VaR is estimated by assuming normally distributed returns and applying Maximum Likelihood Estimation (MLE). The figure indicates that the VaR estimates, under a correct distributional assumption, already vary considerably around the true VaR.

Figure 1: Density plot of VaR estimates



Note: VaR estimates are obtained by applying Maximum Likelihood Estimation on 5000 simulated return series with 10 observations. The returns follow a standard normal distribution and the vertical dotted line corresponds with the true VaR.

Previous research suggests that parameter uncertainty plays a key role in this variation. Dowd (2000b) shows that little uncertainty in the parameters causes the VaR confidence intervals to widen rapidly and the VaR estimator to become imprecise. Another important result of Dowd's (2000a) research is that a small data sample results in VaR estimates with fairly wide confidence intervals. Aussenegg (2006), among others, includes parameter uncertainty in a GARCH volatility environment using a Bayesian framework and finds that the proposed Bayesian approach provides a more adequate VaR framework with less uncertainty in VaR estimates compared to frequentist estimation methods.

Although this research gives some evidence that parameter uncertainty impacts VaR variation, some remarks can still be made. Dowd (2000b) demonstrates the impact of parameter uncertainty by inserting sampled parameters in a functional expression of the VaR. However, in practice real data is available to estimate the VaR. Using a Bayesian framework to determine the impact of parameter uncertainty on the VaR would be more realistic, because both data and distributional assumption can be combined. Aussenegg (2006) applies a Bayesian framework to estimate VaR and shows that this results in less uncertain VaR estimates as compared to frequentist methods. However, these results are obtained by using a large amount of data. Furthermore, both studies only take in consideration the impact of parameter uncertainty on uncertainty in the VaR estimate and not the impact on the accuracy of the VaR estimate. Therefore, my research will investigate the impact of incorporating parameter uncertainty on the VaR, when few data is available, compared to not incorporating this uncertainty.

In a simulation study I apply estimation methods that either do (e.g. Bayesian methods) or do not (e.g. Maximum Likelihood Estimation) take into account the effect of parameter uncertainty, to show which is the best performing method. When applying different estimation methods to analyze empirical data, it is difficult to assess which method performs best, since the true data generating process and VaR are unknown. Investigating the impact of parameter uncertainty in a simulation study allows me to compare the estimates of the VaR to the true VaR of the data generating process (DGP). This provides insight in the bias and variation of the different estimation methods. The data generating processes that I use are a Gaussian and a Clayton copula. I derive the estimates of the VaR that take into account parameter uncertainty in two different ways: 1) by taking the empirical VaR from the posterior predictive return distribution or, 2) by computing the VaR using the mean of the sampled parameters from the posterior distributions of these parameters. Within the simulation study I differentiate between the impact of parameter uncertainty on the VaR for the correlation, variances and Clayton parameter.

Once the simulation study results are obtained I will use the results to investigate the impact of incorporating parameter uncertainty on Solvency II Capital Requirement for equity risk. There are three reasons for choosing this empirical example. First, the Solvency II Capital Requirement for equity risk corresponds to a 99.5 percent one year VaR of an insurance company's own funds. As a consequence of this time horizon, few historical annual returns are available to calibrate this VaR. Secondly, Solvency II calibration has circumvented this issue by using a one year rolling-window in order to obtain annual returns on a daily frequency. However, Mittnik (2011) has shown that using such an annualization procedure results in high autocorrelation between the data and VaR estimates which are as a result highly erratic. Therefore, the current accuracy of the Solvency II Capital Requirement is questionable. Thirdly, the Solvency II calibration procedure does not look at the impact of incorporating parameter uncertainty on the Solvency II Capital Requirement.

The results of my research show that incorporating uncertainty in both the correlation parameter and the variance parameters results in almost no underestimation of the true VaR. However, the economic impact of overestimation is rather large. With respect to the impact of incorporating parameter uncertainty on the Solvency II Capital Requirement for equity risk I find that the current capital requirement is insufficient if parameter uncertainty is incorporated.

My paper contributes to the literature in two ways: the simulation study results give insight in the magnitude of under- and overestimation for estimation methods that do and do not incorporate parameter uncertainty. This can be used to justify the use of a Bayesian approach over a frequentist approach, or vice versa, when estimating a VaR with data scarcity. Secondly, it extents the current knowledge on the optimal design of insurance regulation. Because of a better understanding of the impact of parameter uncertainty on capital requirements.

The remainder of the thesis is structured as follows: first, I describe the standard model of Solvency II. Secondly, I explain the simulation study and the parameter estimation methods I used, to be followed by a discussion of the results. Thirdly, I estimate the capital requirement for the equity module of Solvency II. Finally, I answer my research question and discuss the practical relevance of my research.

Part I
Theoretical Part

2 The Standard Model under Solvency II

Solvency II regulation is a three pillar framework. The first pillar focuses on capital requirements. The second pillar bundles relevant activities for supervising risk management processes, while the third pillar covers transparency issues. Under the first pillar two capital requirements were established: the Minimum Capital Requirement and the a Solvency II Capital Requirement (SCR). The former defines a threshold at which insurance companies will no longer be permitted to trade. The latter corresponds to the VaR of an insurance companies own funds subject to a confidence level of 99.5 percent over a one-year period.

The details of the Solvency Capital Requirement are laid down in EIOPA's Report on the fifth Quantitative Impact Study (QIS5) for Solvency II and the Delegated Regulation (EU) 2015/35. The reports determine that the Solvency Capital Requirement is, in most cases, calculated by making use of a standard model.[i] The standard model is designed to follow a modular based approach as exhibited in figure 18, and is divided into six different risk modules, including market, health, default, life and non-life risk as well as intangibles. As shown in figure 18 Appendix E each of the modules can be further divided into different sub-modules. The total SCR is then calculated by following a bottom up approach. Firstly, the gross SCR's of the individual sub-models are calculated and aggregated to compute the SCR of the six different risk modules. Secondly, the Basic Solvency Capital Requirement is computed by using the so-called "square-root formula" given by:

$$BCSR = \sqrt{\sum_{ij} Corr_{ij} \times SCR_i \times SCR_j} + SCR_{intangibles} \tag{1}$$

where SCR_i represents the risk module's capital requirement and $Corr_{ij}$ the entries of a predefined correlation matrix $Corr$ between the different risk modules. Finally, the total SCR will be obtained by adding the operational risk SCR and the adjustments for loss absorbency of technical provisions and deferred taxes to the BSCR.

One of the sub-modules that comprises the market module is the equity module. The SCR of the equity module under Solvency II is calculated by aggregating the SCR of two different classes of equities. The first class contains equities that are listed in regulated markets in countries that are a member of the European Economic Area or the Organization for Economic Co-operation and Development (OECD). The second class comprises of equities only listed in emerging markets, non-listed equities, hedge funds and other investments that are not included elsewhere in the market risk module. The total Solvency Capital Requirement of the equity module is then given by:

$$SCR_{eq} = \sqrt{SCR_{Other}^2 + SCR_{Global}^2 + 2 * \rho_{tail} * SCR_{Other} * SCR_{Global}} \tag{2}$$

The SCR's of the individual equity classes are obtained by multiplying the total market value of each equity class times a shock. According to the 2009 EU Directive, this shock should correspond with a one-year 99.5 percent, which leads to the following expression:

$$\begin{aligned} SCR_i &= VaR_i * MV_{eq,i} \\ &= VaR_i * MV_{eq} * w_i \end{aligned} \tag{3}$$

[i]It should be noted that insurance companies can replace the standard model with an internal model, when it can be shown to be better able to fulfill the directive requirements given a insurance companies risk profile.

where $i \in$ [global, other], $MV_{eq,i}$ the market value of an equity class, VaR_i the Value-at-Risk, MV_{eq} the total market portfolio and w_i the equity class weight in the total market portfolio.

The QIS5 calibration papers also determined the calibration of the VaR of the global and other equity type. The VaR was calculated by adding two parts: the first is the standard capital stress and the second part an adjustment term that has been set in place to mitigate potential pro-cyclical effects of adverse capital market developments.

The standard capital stress (SCS) of the global and other equities were calibrated by making use of a number of benchmark time series. The SCS of the global equities was based on the MSCI World Price index and the SCS for the other equities based on the LPX 50 Total Return index, the HFRX Hedge Fund Total Return index, the MSCI BRIC Price index and the S&P GSCI Commodities Total Return index. One of the problems concerning the use of these indexes was the relative scarcity of data. Only around 8 and 40 years of data were available to calculate the SCS's for the indexes, which gave too little non-overlapping returns to construct a one-year ahead VaR using the HS method. To solve this problem the QIS5 based the calibration on a rolling window of daily measured annual returns for the longest available sample period.

The SCS for the global equities was determined by applying a number of parameter and non-parametric VaR techniques on the MSCI World Price Index. For comparison reasons QIS5 also considered some other time series and data frequencies.

The results showed some variety, depending on the data frequency and the VaR estimation technique used. The QIS5 calibration noticed that there was a large difference between the empirical and normality VaR estimates. However, this could be explained by the non-normality of the indexes. In the end, taking the SCS's of all methods into account, the QIS5 calibrated the SCS for the global class to 45 percent.

The SCS for the the other equities was computed by using only empirical VaR techniques. The QIS5 calibrations observed that there was a wide variety between the different classes of other equities. Furthermore, QIS5 mentioned that the private equity shock was somewhat overstated and the shocks to the hedge fund equities slightly understated. Based on these observation the QIS5 calibration sets the SCS of equity type other equal to 55 percent.

To allow for diversification effects between the global and other equities the QIS5 calibration also calibrated the tail correlation between the two equity classes. The QIS5 calibrations determined the tail correlation between the two equity classes to be equal to 0.75. It should be noted that this calibration was rather arbitrary. In QIS5 the tail correlation between the MSCI World Price Index and the indexes from the other groups were calculated. The QIS5 then randomly picked a correlation value that was somewhat close to the obtained tail correlations.

Eventually the VaR was determined in a political decision. The final VaR estimates were set equal to 39 percent for the global equities and 49 percent for the other equities, this value included the 9 percent adjustment term derived from the symmetric adjustment mechanism. This results in the following total Solvency Capital Requirement of the equity module:

$$SCR_{eq} = MV_{eq} * \sqrt{w_1^2 VaR_1^2 + w_2^2 VaR_2^2 + 2 * 0.75 w_1 w_2 VaR_1 VaR_2} \tag{4}$$

where type 1 and 2 correspond to global and other equity class respectively.

2.1 Research on Solvency II

The design of the new Solvency II regulation has been subject to a lot of research in the last couple of years. One of the major discussions focusses on the option to choose between an internal model or a regulatory standard model (Eling, 2014). Liebwein (2006), argues that the internal risk models are more able to quantify a company's risk position because they model the individual company's situation (risk concentration, diversification; specific investment strategies). Therefore, they reflect that insurance companies risks more accurately than a standard approach can do. Gatzer and Martin (2012) compare the standard model of Solvency II and an internal approach in quantifying the credit and market risk for a non-life insurer. The authors' major result shows that even though the standard approach is easier to use, the insurance company's risk situation is inadequately reflected by the predefined scenario's; leading to an under- and overestimation of the underlying risks.

A different aspect of the Solvency II regulation is the aggregation of the standard formula. Christian et al. (2012), examines the aggregation formula used to derive the life underwriting risk in the Solvency II standard model. The authors use a stochastic model for an internal approach to determine that the correlation matrix used in the life underwriting risk module under Solvency II is not appropriate. The authors find that this leads to an overestimation of the risk for the underlying German data. Further critical notes can be found in Sandstrom (2007) and Pfeifer and Strassburger (2008). Sandstrom (2007) shows that the standard formula needs to be recalibrated in case the underlying risks are skewed (instead of the risk being symmetric and normally distributed) to ensure consistency. Pfeifer and Stressburger (2008) show that if the individual risks have skewed distribution, then the SCR requirements are likely to be under- and overestimated.

The equity module is criticized in the papers written by Mittnik (2011) and Eling and Pankoke (2014). Mittnik (2011), analyzes the QIS5 calibration procedure and finds that the annualization procedure, of transforming daily return data into annual returns, tends to induce spurious dependence patterns, both over time and across assets, which are not present in the observed data. The author concludes that this leads to highly unreliable and erratic VaR and correlation estimations. In Eling and Pankoke (2014), the authors show that the SCR is backward looking and far away from a realistic calibration. The SCR is highly sensitive to the considered data period and underlying definition of return. Furthermore, the aggregation formula underestimates the true risk of the capital requirements due to the fixed correlations used in the correlation matrix.

3 Methods

In this section I discuss the concept of VaR and some general methods used to estimate this VaR. Then I elaborate on the simulation study set-up and the methods used to incorporate parameter uncertainty in both the Clayton and Gaussian data generating processes.

3.1 Value-at-Risk

The VaR measure has become an important tool in quantifying market risk of a portfolio and can be defined as the worst expected loss over a given time interval, under normal market conditions, at a given confidence level η, for an investment portfolio (Jorion, 2006). That is,

$$\eta = Pr(r_t < VaR_t | I_{t-1}) \tag{5}$$

where r_t is a return, η the confidence level and I_{t-1} the information set at time $t-1$. The VaR is a quantile in the conditional one-step-ahead forecast distribution of the returns.

In the literature numerous estimation methods for VaR are available, which can be separated in three categories:

- **Parametric**: includes methods that make fully parametric distributional and model form assumptions on the asset returns, e.g. RiskMetrics and GARCH models.
- **Non − Parametric**: includes the non-parametric methods by which no parameters have to be estimated, e.g. historical simulation.
- **Semi − Parametric**: includes methods which make some assumption on either the error distributions, its extremes, or the model dynamics; e.g. Extreme Value Theory, quantile regression (CAViaR).

The purpose of VaR estimation generally determines what method to choose. For example, the first category of estimation methods is relatively easy to use. However, the problem is that the data generating process is not known and can therefore lead to under- or overestimation of the VaR. The second category has the advantage that it does not assume a data generating process and is also relative easy to use. However, they do implicitly assume that past returns have the same distribution as future returns. Moreover, in case a very strict confidence level is needed, sufficient data should be available to construct the historical distribution of the asset returns. The third category suffers from the same problems as the first category though to a lesser extent. Again assumptions on the error terms may be needed and/or parameters need to be estimated.

The literature often considers a fourth category which has some overlap with the above mentioned categories or is within one of the categories. This category includes methods which follow an algorithmic approach to calculate the VaR. Often, a high number of paths of future asset returns are simulated using the current values as starting point; then the VaR is calculated in a similar way as the historical simulation approach. Examples of these methods are Monte Carlo Simulation and Filtered Historical Simulation. The advantage of using this type of method is that non-linear portfolios can also be considered. A disadvantage, however, can be that some of those methods are computational burdensome (Kuester et al., 2006 and Szular, 2014).

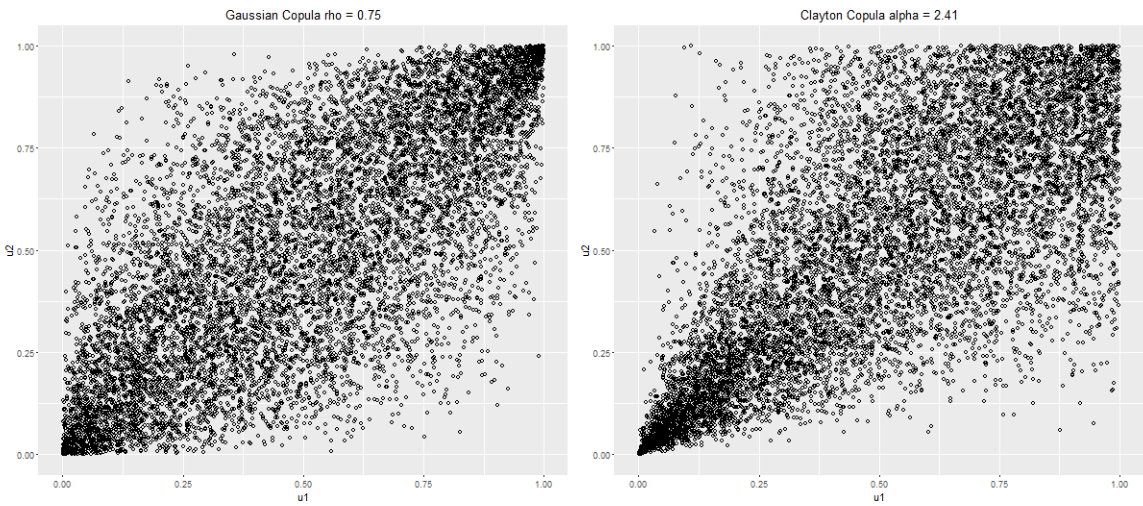
3.2 Simulation Study

In the simulation study I consider a portfolio $R_p = wR_1 + (1-w)R_2$, constructed from $R = (R_1, R_2)$ a two dimensional return vector where R_1 and R_2 equal the asset types global and other respectively. I assume i.i.d normally distributed marginals, $p(R_1) \sim N(\mu_1, \sigma_1^2)$ and $p(R_2) \sim N(\mu_2, \sigma_2^2)$. I also impose two dependency structures between the marginals: the first is a Gaussian Copula and the second a Clayton copula with parameter α.

It is important to mention that copula models do not determine the marginals for R_1 and R_2, but only account for the dependency between R_1 and R_2. The figure below shows for example the difference between dependency structure imposed by the copulas. The Gaussian copula imposes a symmetric dependence with no tail correlation between the returns, while the Clayton copula imposes asymmetric dependence with left tail correlation.

Figure 2: Copula Scatterplots

(a) (b)



Note: The left figure shows a scatterplot of returns under the Gaussian Copula with correlation parameter value equal to 0.75. The right figure shows returns under the Clayton copula with alpha equal to 2.41.

The bivariate distribution that is constructed by combining a copula model and marginals is in most cases unknown. However, combining a bivariate Gaussian copula with normal distributed marginals results in a bivariate normal distribution. This implies the following:

$$R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix}\right) \tag{6}$$

where ρ is the correlation parameter between the assets. Because R follows a bivariate normal distribution R_p is also normally distributed. Consequently, the true VaR can be obtained analytically and equals:

$$VaR_{(1-\eta)} = q_{(1-\eta)}^{-1}\sqrt{w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho\sigma_1\sigma_2} - (w\mu_1 + (1-w)\mu_2) \tag{7}$$

where $q_{(1-\eta)}$ is the $(1-\eta)$ quantile of a standard normal distribution.

In case of the Clayton copula the bivariate distribution of the returns is of a not known form and therefore there is no analytical expression for the true VaR of the portfolio. Instead, the true VaR can be estimated by means of simulation. First a sufficient amount of returns are simulated from the Clayton copula. Then the portfolio returns are calculated and the empirical VaR is computed by taking the $(1-\eta)$ quantile of the portfolio returns.

In the simulation study I consider four different types of simulations. The table below pictures these simulations and indicates that I simulate both DGP two times. Subsequently, I estimate the VaR using a Gaussian copula and a Clayton copula model specification for each DGP. By considering a Clayton copula on a Gaussian DGP and a Gaussian copula on a Clayton DGP I also determine the impact of model misspecification.

Table 1: Simulations

DGP \ Model	Gaussian	Clayton
Gaussian	✓	×
Clayton	×	✓

Note: Four simulations are considered. The ✓ indicates a correct model specification and × an incorrect model specification.

The VaR will be computed by using three different methods. The MLE method computes the VaR using parameter estimates obtained by means of Maximum Likelihood Estimation. The Bayesian I and Bayesian II method calculates the VaR by accounting for parameter uncertainty in two different ways: 1) either by taking the VaR of the posterior predictive return distribution given the simulated returns from the DGP or, 2) by taking the average of the VaR estimates from (7) using the sampled parameters from the posterior distribution. Within the simulation I differentiate between the impact of parameter uncertainty on the VaR for the correlation, the variances and the Clayton copula parameter.

3.3 Estimation Methods for Gaussian Copula Model

In the following section I elaborate on the methods used to calculate the VaR in case of the Gaussian copula model specification.

3.3.1 Uncertainty in Correlation Parameter

To estimate the correlation parameter, I use three methods: two that include parameter uncertainty and one that does not.

Method MLE: this method estimates the correlation parameter by performing maximum likelihood estimation. I consider T observations from $R \sim N(\mu, \Sigma)$, which gives the following likelihood function:

$$
\begin{aligned}
p(\mathbf{R}|\mu,\Sigma) &= \left(\frac{1}{\sqrt{(2\pi)^2|\Sigma|}}\right)^T \exp\{-\frac{1}{2}\sum_{i=1}^{T}(R_i-\mu)'\Sigma^{-1}(R_i-\mu)\} \\
&= (\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}})^T \\
&\times \exp\{-\sum_{i=1}^{T}\frac{\sigma_2^2(R_{i1}-\mu_1)^2+\sigma_1^2(R_{i2}-\mu_2)^2-2\rho\sigma_1\sigma_2(R_{i1}-\mu_1)(R_{i2}-\mu_2)}{2\sigma_1^2\sigma_2^2(1-\rho^2)}\}
\end{aligned} \tag{8}
$$

where $\mathbf{R}$ is a matrix $(R_1, R_2, \ldots, R_T)$ containing the observations, R_i is the ith observation, ρ the correlation parameter, μ_1, μ_2 and σ_1^2, σ_2^2 the mean and variance of R_1 and R_2 respectively. The correlation parameter estimate is then obtained by numerically optimizing the expression in (8).

To incorporate parameter uncertainty I consider a Bayesian approach. This approach offers the ability to incorporate parameter uncertainty by specifying a prior distribution $p(\theta)$ on the parameter θ of interest. The prior distribution is then updated by including the data and making use of Bayes' rule. This framework can be pictured as follows:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta) \tag{9}$$

where $p(\theta)$ is the prior parameter distribution which specifies the subjective believes about the parameter vector, $p(y|\theta)$ the likelihood function given a parameter vector, $p(y)$ the marginal likelihood and $p(\theta|y)$ the posterior distribution which includes the subjective believes and the data.

The main advantage of the Bayesian framework is that it assumes that the "true" parameter is unknown and follows a certain distribution. This offers the ability to include subjective knowledge in the distribution and makes it possible to make inferences on the posterior distribution in case of data scarcity (Fosdick and Raftery, 2012). However, the choice of the prior distribution has a large impact on the posterior distribution and therefore a suitable prior should be chosen. In case of the correlation parameter ρ, I consider an uniformly distributed prior:

$$\rho \sim U[\rho_l, \rho_u] \tag{10}$$

where ρ_l and ρ_u are the uniform distribution's parameters. There are two reasons for adopting this prior distribution. Firstly, the probability on a certain ρ value is equally likely therefore it does not include any information on ρ. Moreover, the uniform distribution is very suitable to restrict ρ values. Secondly, Fosdick and Raftery (2012) found that uniform priors outperformed other prior specification in case the true ρ was not extreme and the variances were known.

To construct the posterior distribution $p(\rho|\mathbf{R})$; the prior distribution and the likelihood function need to be combined as in (9). This results in the following posterior distribution:

$$\begin{aligned} p(\rho|\mathbf{R},\theta) &\propto p(\rho)p(\mathbf{R}|\rho,\theta) \\ &\propto \frac{1}{\rho_u-\rho_l}\left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\right)^T \\ &\times \exp\{-\sum_{i=1}^{T}\frac{\sigma_2^2(R_{i1}-\mu_1)^2+\sigma_1^2(R_{i2}-\mu_2)^2-2\rho\sigma_1\sigma_2(R_{i1}-\mu_1)(R_{i2}-\mu_2)}{2\sigma_1^2\sigma_2^2(1-\rho^2)}\} \end{aligned} \tag{11}$$

where $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$.

In my research I impose the general restriction $\rho \in$ [0,0.995] on the prior and set ρ_l equal to 0 and ρ_u equal to 0.995.

Method Bayesian I: this method uses the posterior predictive distribution of portfolio returns. The posterior predictive distribution of the portfolio returns $p(R_{p,T+1}|\mathbf{R},\theta)$ can be obtained as follows:

$$\begin{aligned} p(R_{p,T+1}|R_p,\theta) &= \int_{\rho_l}^{\rho_u} p(R_{p,T+1},\rho|\mathbf{R},\theta)d\rho \\ &= \int_{\rho_l}^{\rho_u} p(R_{p,T+1}|\rho,\mathbf{R},\theta) \\ &\times p(\rho|\mathbf{R},\theta)d\rho \\ &= \int_{\rho_l}^{\rho_u} p(R_{p,T+1}|\rho,\mathbf{R},\theta)p(\rho|\mathbf{R},\theta)d\rho \end{aligned} \tag{12}$$

Filling in the corresponding distributions gives the following expression:

$$\begin{aligned} p(R_{p,T+1}|\mathbf{R},\theta) \propto \int_{\rho_l}^{\rho_u} \frac{1}{\rho_u-\rho_l}&\left(\frac{1}{2\pi\sqrt{w^2\sigma_1^2+(1-w)^2\sigma_2^2+2\rho w(1-w)\sigma_1\sigma_2}}\right)^{T+1} \\ \times \exp\{-&\sum_{i=1}^{T+1}\frac{(R_{p,i}-\gamma'\mu)^2}{2(w^2\sigma_1^2+(1-w)^2\sigma_2^2+2\rho w(1-w)\sigma_1\sigma_2)}\}d\rho \end{aligned} \tag{13}$$

You can observe that $p(R_{p,T+1}|\mathbf{R},\theta)$ is not analytically tractable and therefore cannot be computed easily. To solve this problem I estimate the VaR numerically and adopt the following grid algorithm:

Step 1: Create a grid $[R^0_{p,T+1},\ R^1_{p,T+1}\ \dots,\ R^G_{p,T+1}]$ where $R^0_{p,T+1}$ is the lower bound (LB) of the grid, $R^i_{p,T+1} = R^0_{p,T+1}$ + i*0.01, $R^G_{p,T+1}$ the upper bound (UB) of the grid and G = 100 * (UB - LB).

Step 2: Numerically estimate $p(R^i_{p,T+1}|\mathbf{R},\theta)$ for i = $1, 2, \dots, G$.

Step 4: Numerically estimate $c = \int_{LB}^{UB} p(R_{p,T+1}|\mathbf{R})dR_{T+1}$ using the grid values, function values from step 2 and the trapezoid rule. Divide $p(R^i_{p,T+1}|\mathbf{R},\theta)$ by c for i = $1, 2, \ldots, G$.

Step 5: Define the empirical CDF as $P(R_{p,T+1}) = G^{-1}\sum_{j=1}^{G} I[R_{p,T+1} < R^j_{p,T+1}]$.

Step 6: Then $VaR_{1-\eta}$ = -1*$P(R^k_{p,T+1})$ such that $\sum_{j=1}^{k} P(R^j_{p,T+1}) = \eta$.

Method Bayesian II: This method estimates the VaR by sampling ρ from the posterior distribution in (11). Again this distribution is not analytically tractable therefore I apply an Accept-Reject (AR) simulation to numerically estimate the posterior distribution. The basic idea behind the AR algorithm is that values from the density function $f(.)$ can be simulated if it is possible to simulate values from another proposal density function $g(.)$ and if a number c can be found such that $f(Y) \leq cg(Y)$, $c \geq 1$ for all Y in support of $f(Y)$. The simulated value Y is then accepted if $u \leq \frac{f(Y)}{cg(Y)}$, where u is a randomly drawn value from $U(0,1)$ (Greenberg, 2008). I consider a proposal function $g(\rho)$ which is uniformly distribution on the interval [0, 0.995]. The value c should be equal to $p(\rho_{mode}|\mathbf{R},\theta)$ divided by $g(\rho)$ in order for the constraint to hold. The mode of $p(\rho|\mathbf{R},\theta)$ will be obtained by means of numerical optimization. The AR algorithm can be summarized as follows:

Step 1: Set m = 1

Step 2: Draw $u_{1,m}$ from $U(\rho_l, \rho_u)$ and $u_{2,m}$ from $U(0,1)$.

Step 3: if $u_{2,m} \leq \frac{p(u_{1,m}|\mathbf{R},\theta)}{cg(u_{1,m})}$ set ρ_m equal to $u_{1,m}$ and compute VaR_m. Otherwise reject and go to step 1.

Step 4: Set m = m +1 and go to step 2.

Step 5: Repeat 20.000 iterations.

Step 6: The VaR equals $\frac{1}{20000}\sum_{i=1}^{20000} VaR_i$.

To get a first impression of the impact of parameter uncertainty I construct a density plot of the VaR estimates using density kernel estimation. The kernel that I consider is a standard normal Gaussian kernel and the bandwidth is chosen by means of the Silverman's Rule of Thumb. To measure the performance of the three estimation methods I consider statistical and economic impact measures. The statistical measures I consider are the relative VaR computed as the mean of the VaR estimates divided by the true VaR, the standard deviation of the relative VaR and Root Means Squared Error (RSME) of the VaR estimates. The economic impact measures determine the impact of misspecifying the VaR. I consider the probability of underestimating the true VaR and probability of underestimating the true VaR times the average additional percentage underestimation. Both measures are also constructed for overestimating the true VaR.

3.3.2 Uncertainty in Variances

This section discusses the methods that are used to determine the impact of uncertainty in the variance of R_1 and R_2, while ρ is fixed.

Method MLE: This method uses the sample variances to estimate the variance parameters.

To include uncertainty in the variance I use an inverse gamma prior distributions: $p(\sigma_i^2|\nu_i,\beta_i) \sim \frac{\beta_i^{\nu_1}}{\Gamma(\nu_i)}\sigma_i^{2(-\nu_i-1)}\exp\left(-\frac{\beta_i}{\sigma_i^2}\right)$ for $i=1,2$ and where $\sigma_i^2>0$ and $F(.)$ denotes the gamma function. I set $(\nu_1,\beta_1)=$ (1e-100, 1e-100) and $(\nu_2,\beta_2)=$ (1e-100, 1e-100). By choosing low values for (ν_i,β_i) the inverse gamma distribution becomes almost non-informative, but remains a proper distribution.

Combining the prior with the likelihood function in (8) results in the following posterior distribution:

$$
\begin{aligned}
p(\sigma_1^2,\sigma_2^2|\mathbf{R},\phi) &\propto p(\sigma_1^2)p(\sigma_2^2)p(\mathbf{R}|\phi)\\
&\propto \sigma_1^{-2(\nu_1+1)}\sigma_2^{-2(\nu_2+1)}\left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\right)^T\\
&\times\exp\{-\sum_{i=1}^{T}\frac{\sigma_2^2(R_{i1}-\mu_1)^2+\sigma_1^2(R_{i2}-\mu_2)^2-2\rho\sigma_1\sigma_2(R_{i1}-\mu_1)(R_{i2}-\mu_2)}{2\sigma_1^2\sigma_2^2(1-\rho^2)}\}\\
&\times\exp\{-\frac{\beta_1}{\sigma_1^2}-\frac{\beta_2}{\sigma_2^2}\}
\end{aligned}
\tag{14}
$$

Method Bayesian I: This method computes the VaR by using the posterior predictive distribution of portfolio returns $p(R_{p,T+1}|\mathbf{R},\theta)$, which is defined as follows:

$$
\begin{aligned}
p(R_{p,T+1}|\mathbf{R},\phi) &= \int_0^\infty\int_0^\infty p(R_{p,T+1},\sigma_1^2,\sigma_1^2|\mathbf{R},\phi)d\sigma_1^2 d\sigma_2^2\\
&= \int_0^\infty\int_0^\infty p(R_{p,T+1}|\sigma_1^2,\sigma_1^2,\mathbf{R},\phi)p(\sigma_1^2,\sigma_2^2|\mathbf{R},\phi)d\sigma_1^2 d\sigma_2^2
\end{aligned}
\tag{15}
$$

filling in the corresponding distributions gives the following expression:

$$
\begin{aligned}
p(R_{p,T+1}|\mathbf{R},\phi) = \int_0^\infty\int_0^\infty &\frac{1}{\sigma_1^{2(\nu_1+1)}\sigma_2^{2(\nu_2+1)}}\left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\right)^{T+1}\\
&\times\exp\{-\sum_{i=1}^{T+1}\frac{(R_{p,i}-\gamma'\mu)}{2(w^2\sigma_1^2+(1-w)^2\sigma_2^2+2\rho w(1-w)\sigma_1\sigma_2)}\}\\
&\times\exp\{-\frac{\beta_1}{\sigma_1^2}-\frac{\beta_2}{\sigma_2^2}\}d\sigma_1^2 d\sigma_2^2
\end{aligned}
\tag{16}
$$

This expression does not have an analytical solution. Therefore I adopt the same algorithm to obtain the posterior predictive distribution as in case of the correlation parameter. Only now the predictive posterior distribution is changed to the expression (16)

Method Bayesian II: This method estimates the parameter by sampling σ_1^2 and σ_2^2 from their posterior distribution. The posterior distribution does not look to be of any known form therefore I resort to a Metropolis Hastings algorithm based on a Markov Chain. The idea is to simulate a value from a known candidate density function, which has a similar shape to the unknown density. The value is then either accepted or rejected with a certain probability α_{accept} (Greenberg, 2008). To determine a correct candidate density function I separate the joint distributions in two conditional distributions: $p(\sigma_1^2|\mathbf{R},\phi)$ and $p(\sigma_2^2|\sigma_1^2,\mathbf{R},\phi)$. In Appendix A Derivation II you can observe that $\sigma_1^2|\mathbf{R},\phi$ follows an Inverse Gamma distribution and $\sigma_2^2|\sigma_1^2,\mathbf{R},\phi$ an unknown distribution which looks similar to an Inverse Gamma. Therefore, my candidate function will be an Inverse Gamma distribution. This results in the following Metropolis Hastings algorithm:

Step 1: Specify starting values $\sigma_{1,m}^2, \sigma_{2,m}^2$ and $m = 1$.

Step 2: Simulate $\sigma_{1,*}^2$ from $g(\sigma_1^2) \sim IG\left(\alpha_{ig} = \frac{T}{2} + \nu_1, \beta_{ig} = \sum_{i=1}^{T} \frac{(R_{i1}-\mu_1)^2}{2(1-\rho^2)} + \beta_1\right)$ and $U \sim U(0,1)$

Set $\sigma_{1,m+1}^2 = \sigma_{1,*}^2$ if $U \leq \alpha_{accept}$, otherwise set $\sigma_{1,m+1}^2 = \sigma_{1,m}^2$

where $\alpha_{accept} = \min\left(\frac{p(\sigma_{1,*}^2|\sigma_{2,m}^2,\mathbf{R},\phi)g(\sigma_{1,*}^2)}{p(\sigma_{1,m}^2|\sigma_{2,m}^2,\mathbf{R},\phi)g(\sigma_{1,m}^2)}\right)$

Step 3: Simulate $\sigma_{2,*}^2$ from $g(\sigma_2^2) \sim IG\left(\alpha_{ig} = \frac{T}{2} + \nu_2, \beta_{ig} = \sum_{i=1}^{T} \frac{(R_{i2}-\mu_2)^2}{2(1-\rho^2)} + \beta_2\right)$ and $U \sim U(0,1)$

Set $\sigma_{2,m+1}^2 = \sigma_{2,*}^2$ if $U \leq \alpha_{accept}$, otherwise set $\sigma_{2,m+1}^2 = \sigma_{2,m}^2$

where $\alpha_{accept} = \min\left(\frac{p(\sigma_{2,*}^2|\sigma_{1,m}^2,\mathbf{R},\phi)g(\sigma_{1,*}^2)}{p(\sigma_{2,m}^2|\sigma_{1,m}^2,\mathbf{R},\phi)g(\sigma_{2,m}^2)}\right)$

Step 4: Compute VaR_m using the expression in 7.

Step 5: Set $m = m + 1$ and go to step 2.

Step 6: Repeat m^* iterations.

Step 7: The VaR equals $\frac{1}{m*}\sum_{i=1}^{m*} VaR_i$.

To check whether the Markov Chain has converged and to determine an appropriate burn in sample I use mean reverting plots. I also inspect the lag function and determine a thinner such that the first autocorrelation of the draws is below 0.05. This is stipulated as best practice according to Greenberg (2008). I stop the algorithm if in total 20.000 samples will be acquired.

To measure the performance of the three estimation methods I use the same methods as in case of the correlation parameter.

3.3.3 Joint Uncertainty

This section discusses the methods used to include uncertainty in the covariance matrix Σ.

Method MLE: this method estimates the covariance matrix by maximizing the likelihood function in (8) of which the solution equals the sample covariance. The VaR is then computed by filling in the parameter estimates from the sample covariance in equation (7).

To simplify later calculation when parameter uncertainty is accounted for I define $\hat{R} = (R_1, R_2, \ldots, R_T)'$ a matrix containing T observation. Then $\hat{R}$ follows a matricvariate normal distribution, $\hat{R} \sim MN(0_{2T}, I_T \otimes \Sigma)$, where I_T is a $T \times T$ identity matrix and 0_{2T} a vector containing $2 * T$ zeros. The distribution of $\hat{R}$ is then given as follows:

$$p(\hat{R}|\Sigma \otimes I_T) = (\frac{1}{2\pi})^{(2T)}|\Sigma|^{-T/2} \quad \times \exp(-1/2tr[\Sigma^{-1}\hat{R}'\hat{R}]) \tag{17}$$

where tr is the trace operator. To include joint uncertainty I choose a non-informative prior $p(\Sigma) \propto |\Sigma|^{-3/2}$. Combining this prior with the likelihood using Bayes' theorem yields the following posterior distribution:

$$p(\Sigma|\hat{R}, \mu) \propto |\Sigma|^{-(T+3)/2} \exp\{-\frac{1}{2}tr[\Sigma^{-1}\hat{R}'\hat{R}]\} \tag{18}$$

You can observe that this posterior distribution is proportional to an Inverted Wishart distribution with parameters $\hat{R}'\hat{R}$ and T degrees of freedom.

Method Bayesian I: this method computes the VaR by making use of posterior predictive distribution:

$$\begin{aligned} p(R_{T+1}|\hat{R}) &= \int p(R_{T+1}, \Sigma|\hat{R})d\Sigma \\ &= \int p(R_{T+1}|\Sigma, \hat{R})p(\Sigma|\hat{R})d\Sigma \end{aligned} \tag{19}$$

This expression is derived in Derivation III Appendix B and is identified as a multivariate Student's distribution:

$$R_{T+1}|\hat{R} \sim MT(0_2, A, v) \tag{20}$$

where $A = \frac{\hat{R}'\hat{R}}{T-1}$ is the scale matrix, 0_2 the location parameter and $v = T - 1$ the degrees of freedom. The portfolio return prediction $R_{p,t+1}$ is a linear combination of R_{t+1}, according to Appendix C, that follows a univariate distribution with location parameter 0, scale parameter $\gamma' A \gamma$ and v degrees of freedom. The corresponding VaR can be obtained by taking the $(1 - \eta)$ quantile of this distribution.

Method Bayesian II: this method computes the VaR by sampling covariance matrices from posterior distributions as follows:

Step 1: Set m = 1

Step 2: Draw Σ_m from an Inverted Wishart distribution with parameters: $\hat{R}'\hat{R}$ and T degrees of freedom.

Step 3: Compute VaR_m using equation (7).

Step 4: Set m = m +1 and go to step 2.

Step 5: Repeat 20.000 iterations.

Step 6: The VaR equals $\frac{1}{20000}\sum_{i=1}^{20000} VaR_i$.

To measure the performance of the three estimation methods I use the same methods as in case of the correlation parameter.

3.4 Estimation Methods for Clayton Copula Model

This section gives an outline on the copula framework and discusses the Clayton copula model. It also elaborates on the methods used to estimate the Clayton parameter and a simulation that calculates the VaR.

3.4.1 Copulas

A p dimensional copula $C(u_1, u_2, \ldots, u_p)$ is a multivariate distribution defined on the unit cube $[0,1]^p$, where the marginal distributions are uniformly distributed on the interval (0,1). More precisely:

$$C(u_1, u_2, \ldots, u_p) = Pr(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_p \leq u_p) \tag{21}$$

where $U_i \sim$ for $i = 1, 2, \ldots, p$. Furthermore, it can been shown (see Sklar, 1959) that given a fixed set of continuous marginal distributions, distinct copulas define distinct joint densities function. Thus, given any joint distribution $F(x_1, x_2, \ldots, x_p)$ with continuous marginals, there is an unique copula function C:$[0,1]^p \to [0,1]$ of F such that:

$$F(x_1, x_2, \ldots, x_p) = C(F_1(x_1), F_2(x_2), \ldots, F_p(x_p)) \tag{22}$$

for all $(x_1, x_2, \ldots x_p) \in R^p$ and where $F_i(x_i) = Pr(X_i \leq x_i)$, $x_i \in R$ for $i = 1, 2, \ldots, p$. The associated copula density function is given by:

$$c(x_1, x_2, \ldots, x_p) = \frac{\partial^p C(x_1, x_2, \ldots, x_p)}{\partial x_1, \partial x_2, \ldots, \partial x_p} \tag{23}$$

and the joint density will be

$$f(x_1, x_2, \ldots, x_p) = \prod_j^p f_j(x_j) c(x_1, x_2, \ldots, x_p) \tag{24}$$

In my simulation study I use a bivariate Clayton copula which is part of the Archimedean copula family. Archimedean copulas are constructed by making use of a generator function $\Psi(u)$ that is strictly convex and monotonic decreasing with $\Psi(1) = 0$ and $\lim_{u \to 0} \Psi(u) = \infty$. The corresponding Archimedean copula is then given as:

$$C(u_1, u_2, \ldots, u_p) = \Psi^{-1}(\Psi(u_1) + \Psi(u_2) + \cdots + \Psi(u_p)) \tag{25}$$

in case of the Clayton copula the generator function equals $\Psi(u) = \alpha^{-1}(u^{-\alpha} - 1)$ and the inverse of the generator function equals $\Psi^{-1}(x) = (\alpha x + 1)^{-1/\alpha}$. This results in the following copula function:

$$C(u_1, u_2, \ldots, u_p; \alpha) = (u_1^{-\alpha} + u_2^{-\alpha} + \cdots + u_p^{-\alpha} + 1)^{-1/\alpha} \tag{26}$$

and a density function equal to:

$$c(u_1, u_2, \ldots, u_p) = (1 - p + \sum_{i=1}^p u_i^{-\alpha})^{-p-(1/\alpha)} \prod_{j=1}^p (u_j^{\alpha-1}((j-1)\alpha + 1)) \tag{27}$$

There were a number of reasons to consider the Clayton copula. Firstly, the Clayton copula has positive lower tail dependence with $l^L = 2^{-1/\alpha}$ and no upper tail dependence. Lower positive tail dependence is a stylized fact of multivariate returns. Secondly, the Clayton copula only has one parameter that needs to be estimated. This is relevant because only little data is available therefore too many parameters would result in high estimation uncertainty. Thirdly, simulating from the Clayton copula is straightforward in comparison to other Archimedean copulas (Fantazzini, 2008). This will be useful when the VaR needs to be simulated.

For the marginals of R_1 and R_2 I consider the same distributions as under the Gaussian model.

3.4.2 Uncertainty in the Clayton Copula parameter

In this section I first discuss the simulation used to obtain a VaR from a Clayton copula model. Then I elaborate on the three methods used to estimate the Copula parameter.

There is no analytical expression for the VaR of the Clayton copula, therefore I use the following simulation to obtain an estimate. First, I simulate 20.000 returns from the Clayton copula with parameter α using the conditional approach.[ii]. Then the portfolio returns are computed, ordered and the empirical VaR is estimated by picking the $(1-\eta)$ quantile of the empirical portfolio returns. For computational conveniences I estimate the empirical VaR for α values on the grid $[0.001, 0, 002, \ldots, 50]$ beforehand. This offers the ability to pick the closest empirical VaR to an estimated α. In case an estimated α is outside the grid I simulate the VaR individually.

Method Kendall's Tau: this method uses the relationship $\tau = \frac{\alpha}{\alpha+2}$, where τ corresponds to Kendall's tau, to estimate the alpha parameter.

Method MLE: this method estimates α by performing maximum likelihood estimation on the likelihood function, which is derived in Appendix C. Because the marginals' parameters are known, expression (49) is further simplified and α can be obtained by solving the following expression:

$$\begin{aligned}\hat{\alpha} = ArgMax_{\alpha}(-2-1/\alpha)\sum_{t=1}^{T}\ln\left(-1+u_{1t}^{-\alpha}+u_{2t}^{-\alpha}\right)\\ +(-\alpha-1)\sum_{t=1}^{T}(\ln u_{1t}+\ln u_{2t})+T\ln(\alpha+1)\end{aligned} \tag{28}$$

where $u_{1t} = \Phi_1(R_{1t};\theta_1)$, $u_{2t} = \Phi_2(R_{2t};\theta_2)$, $\theta_1 = (\mu_1, \sigma_1^2)$ and $\theta_2 = (\mu_2, \sigma_2^2)$

To incorporate parameter uncertainty I consider an uninformative Jeffrey's prior $\alpha \propto \alpha^{-1}$. Combining this prior with the joint density in (47) gives the following posterior distribution:

$$\begin{aligned}p(\alpha|\mathbf{R},\theta_1,\theta_2) &\propto p(\mathbf{R}|\alpha,\theta_1,\theta_2)p(\alpha)\\ &\propto \alpha^{-1}\prod_{t=1}^{T}\left(-1+u_{1t}^{-\alpha}+u_{2t}^{-\alpha}\right)^{(-2-1/\alpha)}\\ &\times u_{1t}^{(-\alpha-1)}u_{2t}^{(-\alpha-1)}(\alpha+1)(\sqrt{2\pi}\sigma_1)^{-1}\exp\left(-\frac{(R_{1t}-\mu_1)^2}{2\sigma_1^2}\right)\\ &\times(\sqrt{2\pi}\sigma_2)^{-1}\exp\left(-\frac{(R_{2t}-\mu_2)^2}{2\sigma_2^2}\right)\end{aligned} \tag{29}$$

[ii] For a full explanation on the conditional approach I refer to Luciano et al. (2004)

Because the marginals' parameters are known the expression can be further simplified to:

$$
\begin{aligned}
p(\alpha|\mathbf{R}, \theta_1, \theta_2) \propto \alpha^{-1} \prod_{t=1}^{T} & \left(-1 + u_{1t}^{-\alpha} + u_{2t}^{-\alpha}\right)^{(-2-1/\alpha)} \\
& \times u_{1t}^{(-\alpha-1)} u_{2t}^{(-\alpha-1)} (\alpha + 1)
\end{aligned} \tag{30}
$$

Method Bayesian I: this method computes the VaR by using the posterior predictive distribution of returns $p(R_{T+1}|\alpha, \mathbf{R}, \theta_1, \theta_2)$ which is defined as follows:

$$
\begin{aligned}
p(R_{T+1}|\mathbf{R}, \theta_1, \theta_2) &\propto \int_0^\infty p(R_{T+1}, \alpha|\theta_1, \theta_2) d\alpha \\
&\propto \int_0^\infty p(R_{T+1}|\mathbf{R}, \theta_1, \theta_2, \alpha) p(\alpha|\mathbf{R}, \theta_1, \theta_2) d\alpha
\end{aligned} \tag{31}
$$

filling in the corresponding distributions gives the following expression:

$$
\begin{aligned}
p(R_{T+1}|\mathbf{R}, \theta_1, \theta_2) \propto \alpha^{-1} \prod_{t=1}^{T+1} & \left(-1 + u_{1t}^{-\alpha} + u_{2t}^{-\alpha}\right)^{(-2-1/\alpha)} \\
& \times u_{1t}^{(-\alpha-1)} u_{2t}^{(-\alpha-1)} (\alpha + 1)
\end{aligned} \tag{32}
$$

This expression cannot be solved analytically, therefore I adopt the same algorithm to obtain the posterior predictive distribution as in case of the correlation parameter. However, the problem is that the grid should be extended to a two-dimensional grid because the return vector has two dimensions. This resulted in an algorithm that was computational too burdensome and therefore I will not consider this method.

Method Bayesian II: this method computes the VaR as follows:

Step 1: Set m = 1

Step 2: Draw α_m from the posterior distribution in 30.

Step 3: Pick the VaR_m corresponding to α_m from the table, or simulate.

Step 4: Set m = m +1 and go to step 2.

Step 5: Repeat 20.000 iterations.

Step 6: The VaR equals $\frac{1}{20000} \sum_{i=1}^{20000} VaR_i$.

Because the posterior distribution is not analytical tractable I consider a Markov Chain based slice sampler proposed by (Neal, 2003) to obtain α estimates. The reason for this choice is that (Silva and Lopes, 2008) showed that the sampler, under similar model specifications, had good properties such as: low autocorrelations and a fast convergence rate. As an additional check I also inspect the mean reverting plots and the autocorrelation function just as in case of the Metropolis Hastings algorithm.

4 Results

In the following section I elaborate on the simulation study set-up and the obtained results.

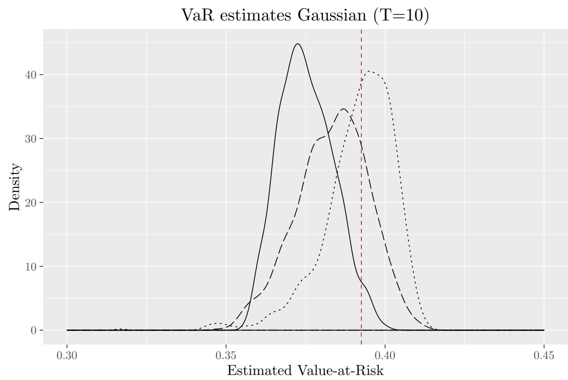
The simulation study uses parameter values that correspond with the calibrated parameter values in Solvency II. In total, 1000 paths were generated with lengths $T = 10$ and $T = 25$, μ_1 and μ_2 were set equal to 0, σ_1^2 and σ_1^2 equal 0.0229 and 0.0361, ρ equal to 0.75, α equal to 2.41, η equal to 0.05, and $m*$ equal to 20.000.

Several parameters need to be determined for the estimation methods. For the Bayesian I method, I find grid bound values equal to (-1,1) and (-0.6,0.6) for the correlation parameter and variance parameters respectively. For the Metropolis Hastings algorithm, used in the Bayesian II method, I initialize the algorithm with the simulation study variances, then I determine a burn-in sample of 10.000 and a thinner of every 8^{th} draw to get the first autocorrelation below 0.05 and I set $w*$ in such a way that 20.000 correct draws are obtained. For the Neil Slice sampling I use the same burn-in sample and thinner.

4.1 Correlation Uncertainty

Figure 3 shows the density plots of the VaR estimates when using the MLE method, Bayesian I method, and Bayesian II method under a Gaussian DGP. The dotted line corresponds with the MLE method, the solid line with the Bayesian I method, and the dashed line with the Bayesian II method. We notice that the VaR estimates of all three estimation methods have a considerable amount of variation. This variation is larger when uncertainty is incorporated in the correlation parameter. Furthermore, the amount of variation in the VaR estimates appears to be larger for the Bayesian II method than for the Bayesian I method. We also observe that the VaR density plots of the Bayesian methods are positioned more distant from the true VaR than the VaR density of the MLE method.

Figure 3: Density plots of the VaR estimates (correlation uncertainty).



Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method under a Gaussian DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of the Bayesian I method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

Table 2 shows the economic impact measures of misspecifying the VaR when the MLE method, Bayesian I method, and Bayesian II method are used with respect to a Gaussian DGP and a Clayton DGP. In case of a Gaussian DGP, we notice that the probability of underestimation equals 0.456 for the MLE method, 0.967 for the Bayesian I method, and 0.782 for the Bayesian II method. The impact of underestimating the true VaR equals 0.4 percent for the MLE method, 1.73 percent for the Bayesian I, and 1.0 percent for Bayesian II method. Hence, incorporating parameter uncertainty results in more underestimation of the VaR compared to the MLE method.

Table 2: The economic impact of misspecifying the VaR (correlation uncertainty).

T = 10	Gaussian			Clayton		
Economic Impact.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
Probability Underestimation	0.456	0.967	0.782	0.108	0.379	0.266
Probability times Underestimation	0.4	1.73	1.0	0.13	0.19	0.25
Probability times Overestimation	0.379	0.008	0.12	1.93	0.531	1.1
T = 25						
Probability Underestimation	0.45	0.902	0.72	0.04	0.294	0.093
Probability times Underestimation	0.27	1.4	0.57	0.02	0.247	0.059
Probability times Overestimation	0.262	0.030	0.109	1.8	0.819	1.4

Note: this table reports the probability of underestimating the true VaR, the probability of underestimating the true VaR times the average percentage of underestimation, and the probability of overestimation times the average percentage of overestimation. These measures are reported for VaR estimates obtained using the MLE method, the Bayesian I method, and the Bayesian II method. The table shows the measures both for 10 and 25 observations under a Gaussian DGP and a Clayton DGP

Moreover, from table 2 we can deduce that the MLE method, the Bayesian I method, and the Bayesian II method overestimate the true VaR on average with 0.379, 0.008 and 0.12 percent respectively. These values are relatively low compared to the impact of underestimating the true VaR for each method. Increasing the number of observations to 25 results in better economic impact performance for all methods.

Table 3 reports the statistical measures of the VaR estimates. We observe that, under the Gaussian DGP, the VaR estimates of the MLE method have a negative bias of 0.2 percent, the Bayesian I method a negative bias of 4.4 percent, and the Bayesian II method a negative bias of 2 percent. Incorporating parameter uncertainty therefore results in more negative bias. An explanation for the increase in negative bias is that the uniform prior distribution does include some information concerning ρ. The prior distribution assigns equal probability to each possible ρ, but this implies that the chance of a value below 0.75 is larger than a value above 0.75. This extra information causes the posterior distribution of ρ to shift more towards the left and therefore results in lower VaR.

From table 3 we can also infer that the Bayesian II method has the highest standard deviation of relative VaR, followed by the Bayesian I method, and the MLE method. Because the VaR estimates of the MLE method have the lowest bias and uncertainty this method is also the best performing in terms of RMSE. Increasing the number of observations results in a better performance for all estimation methods in terms of biasedness, and RMSE.

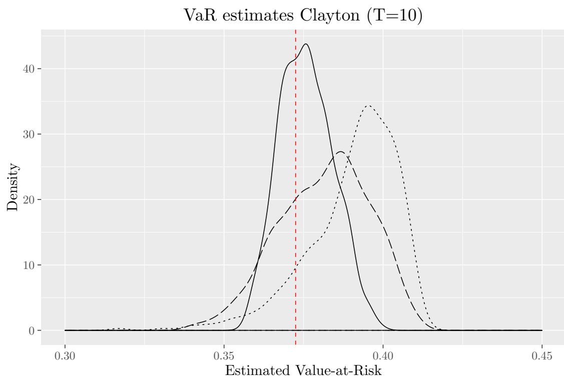
Table 3: The statistical impact of misspecifying the VaR (correlation uncertainty).

T = 10	Gaussian			Clayton		
Statistical Measures.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
Relative VaR	0.998 (0.028)	0.956 (0.022)	0.98 (0.029)	1.05 (0.038)	1.009 (0.022)	1.024 (0.038)
RMSE VaR	0.011	0.019	0.014	0.023	0.009	0.017
T = 25						
Relative VaR	1.00 (0.017)	0.963 (0.029)	0.99 (0.019)	1.05 (0.024)	1.015 (0.031)	1.037 (0.027)
RMSE VaR	0.007	0.019	0.009	0.020	0.013	0.017

Note: The MLE method, the Bayesian I method and the Bayesian II method are used under a Gaussian and Clayton DGP for estimating the VaR. Reported in the table is the relative VaR, the standard deviation of the relative VaR (between brackets) and the Root-Mean-Square Error.

Figure 4 shows the VaR estimate density plots of the estimation methods with respect to the Clayton DGP. From the figure we can infer that, compared to the Gaussian DGP, less probability mass is located to the left of the true VaR for all VaR densities. The amount of variation in the VaR estimates appears to increase slightly when using the MLE method and the Bayesian II method.

Figure 4: Density plots of the VaR estimates (correlation uncertainty and Clayton DGP).



Note: VaR estimates are estimated by the MLE method, the Bayesian I method, and the Bayesian II method under a Clayton DGP. The dotted line is the VaR density when the MLE method is used, the solid line the VaR density when the Bayesian I method is used and the dashed line the VaR density when the Bayesian II method is used. The dotted vertical line is the true VaR.

In case of the Clayton DGP, table 2 indicates that the MLE method is still the best performing method in terms of underestimation. Furthermore, we observe that the economic impact of underestimating the true VaR decreases for all three estimation methods compared to the Gaussian DGP. In terms of overestimation we notice an opposite effect, namely an increase for all methods. Expanding the sample dimension leads to more underestimation in the VaR estimates for the Bayesian I method, which is a different outcome than under the Gaussian DGP.

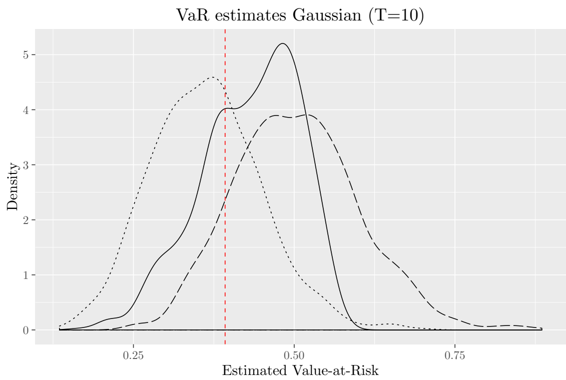
The statistical measures in table 3 show that the VaR estimates of all three method become positively biased in case of a Clayton DGP. As a result the Bayesian I method becomes the best performing method, instead of the MLE method. Furthermore, we notice that the uncertainty in the VaR estimates increases for the MLE method and Bayesian II method. Increasing the number of observations to 25 results in less uncertain VaR estimates for the latter two methods and in more uncertain VaR estimates for the Bayesian I method.

Overall, we observe that incorporating parameter uncertainty using a uniform prior results in VaR estimates that underestimate the true VaR. It is also clear that the MLE method is the best performing method in terms of RMSE, probability of underestimation and economic impact of underestimation. With respect to the Clayton DGP, the MLE method performs best in terms of probability of underestimation and the economic impact of underestimation. The Bayesian methods performs better in statistical terms. However, this is mainly due to a negative change in the true VaR.

4.2 Variance Uncertainty

Figure 5 shows the density plots of the VaR estimates obtained using the MLE method (dotted line), the Bayesian I method (solid line), and the Bayesian II method (dashed line) with respect to a Gaussian DGP. The Bayesian VaR density plots indicate that a majority of the VaR estimates overestimate the true VaR. Furthermore, the width of these density plots is larger compared to the VaR density plot of the MLE method. This demonstrates that the VaR estimates obtained using the Bayesian methods have more variation. We also observe that the probability mass of the MLE VaR density inclines slightly to the left of the true VaR, which suggests a negative bias in the VaR estimates.

Figure 5: Density plots of the VaR estimates (variance uncertainty).



Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method under a Gaussian DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of the Bayesian I method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

Table 4 shows the economic impact of misspecifying the VaR when the MLE method, the Bayesian I method, and Bayesian II method are used under a Gaussian DGP and a Clayton DGP.

Table 4: The economic impact of misspecifying the VaR (variance uncertainty).

T = 10	Gaussian			Clayton		
Economic Impact.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
Probability Underestimation	0.65	0.303	0.119	0.59	0.229	0.09
Probability times Underestimation	5.12	1.7	0.5	4.1	1.3	0.3
Probability times Overestimation	2.2	6	12	2.6	6.6	12.9
T = 25						
Probability Underestimation	0.692	0.401	0.014	0.57	0.298	0.011
Probability times Underestimation	3.6	1.6	0.3	2.6	1	0.02
Probability times Overestimation	1.1	2.9	12.9	1.7	3.8	14.6

Note: This table reports the probability of underestimating the true VaR, the probability of underestimating the true VaR times the average percentage of underestimation, and the probability of overestimation times the average percentage of overestimation. These measures are reported for VaR estimates obtained using the MLE method, the Bayesian I method, and the Bayesian II method. The table shows the measures both for 10 and 25 observations under a Gaussian DGP and a Clayton DGP

In case of the Gaussian DGP, we observe that incorporating parameter uncertainty using the

Bayesian II method results in the lowest economic impact of underestimation. The results indicate that the probability of underestimating the true VaR equals 0.65 for the MLE method, 0.30 for the Bayesian I method and 0.119 for the Bayesian II method. Furthermore, the impact of underestimation equals 5 percent for the MLE method, 1.7 percent for the Bayesian I method, and 0.5 percent for the Bayesian II method.

Moreover, table 4 shows that the MLE method is the best performing method in terms of overestimation. On average this method overestimates the true VaR with an additional 2.2 percent, while the Bayesian I method and Bayesian II method overestimate the VaR with an additional 5 and 12 percent respectively. Increasing the number of observations to 25 generates better results for all methods, excluding the Bayesian II method with respect to the impact of overestimation.

Table 5 shows the statistical measures of the VaR estimates. We infer that using the MLE method results in the lowest bias in the VaR estimates. We notice that the VaR estimates of the MLE method have a 7.6 percent negative bias, the VaR estimates of the Bayesian I method a 9.9 percent positive bias, and the VaR estimates of the Bayesian II method a 28.4 percent bias. Furthermore, we observe that variation in VaR estimates is lowest using the Bayesian I method. Looking at the RSME, we observe that the Bayesian I method estimates the VaR most accurately and the Bayesian II method the VaR the least. Expanding the sample dimension lowers the biasedness, the variation and RSME for both the MLE method and the Bayesian I method. The VaR estimates of the Bayesian II method become less uncertain, however the positive bias increases slightly.

Table 5: The statistical impact of misspecifying the VaR (variance uncertainty).

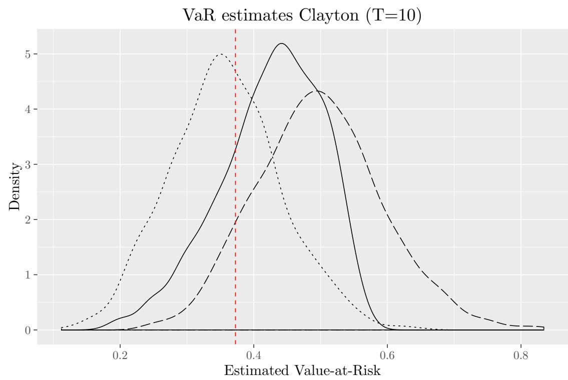
T = 10	Gaussian			Clayton		
Statistical Measures.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
Relative VaR	0.924(0.22)	1.099(0.194)	1.284(0.247)	0.958(0.22)	1.146(0.204)	1.337(0.254)
RSME VaR	0.091	0.083	0.149	0.0834	0.095	0.158
T = 25						
Relative VaR	0.937(0.132)	1.034(0.135)	1.328(0.16)	0.975(0.141)	1.075(0.142)	1.391(0.171)
RSME VaR	0.055	0.055	0.145	0.053	0.063	0.158

Note: The MLE method, the Bayesian I method and the Bayesian II method are used under a Gaussian and Clayton DGP for estimating the VaR. Reported in the table is the relative VaR, the standard deviation of the relative VaR (between brackets) and the Root-Mean-Square Error.

We notice that, in case of uncertainty in the variances, the Bayesian methods overestimate the true VaR, while the opposite effect occurs when uncertainty in the correlation is considered. Furthermore, the impact of under- and overestimation is larger. A possible explanation of the increasing impact could be that two parameters are estimated instead of one.

Figure 6 shows the VaR density plots of the three estimation methods in case of a Clayton DGP. We observe that the VaR density plots are similar compared to the Gaussian DGP. The only difference is that the VaR densities of the MLE method and the Bayesian I method become more peaked.

Figure 6: Density plots of the VaR estimates (variance uncertainty and Clayton DGP).



Note: VaR estimates are estimated by the MLE method, the Bayesian I method, and the Bayesian II method under a Clayton DGP. The dotted line is the VaR density when the MLE method is used, the solid line the VaR density when the Bayesian I method is used and the dashed line the VaR density when the Bayesian II method is used. The dotted vertical line is the true VaR.

With respect to the Clayton DGP, table 4 indicates that the Bayesian I method still performs the best in terms of underestimation, while the impact of overestimating the true VaR remains lowest for the MLE method. Overall, we observe a decrease in the impact of underestimation and an increase of overestimation in the VaR estimates for all methods.

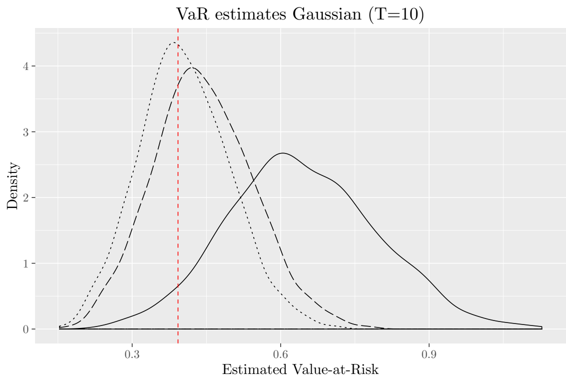
Table 5 shows that, in case of the Clayton DGP, the VaR estimates of the Bayesian methods become more negatively biased compared to the Gaussian DGP. We also observe that the variation in the VaR estimates increases slightly for all estimation methods. Expanding the sample dimension results in better performance of the VaR estimates in terms of biasedness, variation and the RSME, using the MLE method and the Bayesian I method. The VaR estimates of the Bayesian II method become more positively biased.

Overall, we observe that incorporating parameter uncertainty using the Bayesian II method results in VaR estimates that hardly underestimate the true VaR, but do have a large economic impact of overestimation. The MLE method underestimates the VaR most often in comparison to the methods which incorporate parameter uncertainty, but has the lowest impact of overestimation. The Bayesian I method performs good in terms of both the economic impact of underestimation and the economic impact of overestimation. Furthermore, its VaR estimates are the most accurate in statistical terms.

4.3 Joint Uncertainty

Figure 7 shows the VaR density plots of the MLE method (dotted line), the Bayesian I method (solid line), and the Bayesian II method (dashed line) under a Gaussian DGP. We notice that the VaR estimates of the MLE method and Bayesian II method center partially to the right of the true VaR, while the VaR estimates of the Bayesian I method center almost entirely to the right of the true VaR. The shape of the VaR density plots of the MLE method and the Bayesian I method appears to be equal. This can be explained by the fact that the MLE estimates converge to an Inverted Wishart distribution if the number of observations increases. The variation in the VaR estimates is the highest for the Bayesian I method, while the variation in the VaR estimates of the MLE method and Bayesian II method appears to be equal.

Figure 7: Density plots of the VaR estimates (joint uncertainty).



Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method under a Gaussian DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of the Bayesian I method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

Table 6 shows the economic impact measures of misspecifying the VaR for the MLE method, the Bayesian I method, and Bayesian II method under a Gaussian and Clayton DGP.

Table 6: Economic impact of misspecifying the VaR (joint uncertainty).

T = 10	Gaussian			Clayton		
Economic Impact.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
Probability Underestimation	0.471	0.041	0.313	0.365	0.027	0.235
Probability times Underestimation	3.0	0.2	1.9	2.0	0.1	1.2
Probability times Overestimation	4.3	25.4	7.0	5.0	26.5	8.2
T = 25						
Probability Underestimation	0.508	0.049	0.401	0.364	0.024	0.285
Probability times Underestimation	2.1	0.14	1.6	1.32	0.046	0.096
Probability times Overestimation	2.4	12.4	3.2	3.4	13.9	4.3

Note: This table reports the probability of underestimating the true VaR, the probability of underestimating the true VaR times the average percentage of underestimation, and the probability of overestimation times the average percentage of overestimation. These measures are reported for VaR estimates obtained using the MLE method, the Bayesian I method, and the Bayesian II method. The table shows the measures both for 10 and 25 observations under a Gaussian DGP and a Clayton DGP.

From the table we can infer that, in case of the Gaussian DGP, the probability of underestimating the VaR equals 0.471, 0.041, and 0.313 for the MLE, Bayesian I, and Bayesian II method respectively. The economic impact of underestimation equals 3 percent for the MLE method, 0.2 percent for the Bayesian I method, and 1.9 for the Bayesian II method. Looking at the impact of overestimation we notice that the Bayesian I overestimates the true VaR on average with 25.4 percent, the Bayesian II method with 7 percent, and MLE method with 4.3 percent. This outcome indicates that incorporating parameter uncertainty results in a VaR estimate that overestimates the true VaR, especially in case of the Bayesian I method. However, the economic impact of underestimation is small. Increasing the number of observations results in convergence of performance between the Bayesian II method and MLE method. Furthermore, all estimation methods perform better in terms of over- and underestimation.

Table 7 shows the statistical measures of the VaR estimates. We observe that the VaR estimates of the MLE method, the Bayesian I method, and the Bayesian II method have a positive bias of 3.3 percent, 64.4 percent, and 13 percent respectively. The amount of variation in the VaR estimates for all three methods is considerable. In terms of RSME the MLE method estimates the VaR most accurate, while the Bayesian I method estimates the VaR least accurate.

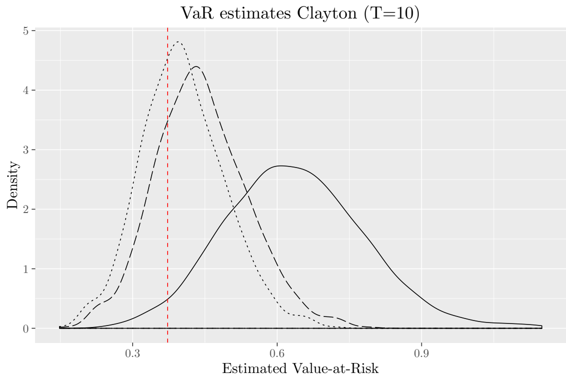
Table 7: The statistical impact of misspecifying the VaR (joint uncertainty).

T = 10	Gaussian			Clayton		
Statistical Measures.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
Relative VaR	1.033 (0.233)	1.641(0.382)	1.131 (0.255)	1.086 (0.235)	1.710(0.387)	1.188 (0.257)
RSME VaR	0.095	0.293	0.11	0.095	0.302	0.118
T = 25						
Relative VaR	1.009 (0.147)	1.313(0.192)	1.042 (0.152)	1.055 (0.147)	1.373(0.191)	1.090 (0.152)
RMSE VaR	0.058	0.158	0.062	0.058	0.155	0.066
True parameters:	VaR = 0.393	$\rho = 0.75$	$\sigma_1^2 \backslash \sigma_2^2 = 0.0229 \backslash 0.0361$	VaR = 0.373	$\rho = 0.727$	

Note: The MLE method, the Bayesian I method and the Bayesian II method are used under a Gaussian and Clayton DGP for estimating the VaR. Reported in the table is the relative VaR, the standard deviation of the relative VaR (between brackets) and the Root-Mean-Square Error.

Figure 8 shows the VaR density plots of the MLE method (dotted line), the Bayesian I method (solid line), and the Bayesian II method (dashed line) under a Clayton DGP. The figure shows no large differences compared to the VaR density plots in figure 7.

Figure 8: Density plots of the VaR estimates (joint uncertainty and Clayton DGP).



Note: VaR estimates are estimated by the MLE method, the Bayesian I method, and the Bayesian II method under a Clayton DGP. The dotted line is the VaR density when the MLE method is used, the solid line the VaR density when the Bayesian I method is used and the dashed line the VaR density when the Bayesian II method is used. The dotted vertical line is the true VaR.

In case of the Clayton DGP, table 6 indicates that the Bayesian I method remains the best performing method in terms of underestimation and the MLE method the best in terms of overestimation. In general, we observe that the impact of underestimation increases, but overestimation decreases for all estimation methods. Increasing the number of observations results in better performance for all three estimation methods.

Table 7 also shows that estimating the VaR, with respect to a Clayton DGP, results in an increase of biasedness for all estimation methods compared to the Gaussian DGP. The amount of variation does decrease slightly for all three estimation methods. In case of the Bayesian methods this decrease is insufficient to balance the increase of biasedness and therefore the RSME increases.

Overall, we observe that including parameter uncertainty results in VaR estimates that overestimate the true VaR, especially in case of the Bayesian I method. However, incorporating parameter uncertainty results in lower economic impact of underestimation. The MLE method performs well both in terms of statistical accuracy and the economic impact of overestimation.

I also look at the impact portfolio weights have on the VaR estimates for all three methods. For this purpose, I have used the Gaussian DGP with T = 10 and vary weights on the interval $(0, 1, \ldots, 1)$ and, subsequently, computed for each portfolio the VaR estimates' under- and overestimation.

Table 8 shows the impact of underestimating the true VaR for the MLE method, the Bayesian I method, and Bayesian II method. Reported are the portfolio weights, the probability of underestimation times the average percentage of underestimation and the relative increase in underestimation for each portfolio weight step.

Table 8: The economic impact of underestimating the VaR when portfolio weights vary.

Economic Impact	Underestimation			Relative Underestimation		
Weights.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
w = 0	0.040	0.0048	0.016			
w = 0.1	0.038	0.0047	0.015	0.05	0.021	0.063
w = 0.2	0.037	0.0045	0.015	0.026	0.043	0
w = 0.3	0.035	0.0044	0.014	0.054	0.022	0.067
w = 0.4	0.034	0.0042	0.013	0.029	0.045	0.071
w = 0.5	0.033	0.0040	0.012	0.029	0.048	0.077
w = 0.6	0.032	0.0040	0.011	0.030	0	0.083
w = 0.7	0.031	0.0039	0.011	0.031	0.025	0
w = 0.8	0.030	.0038	0.011	0.032	0.026	0
w = 0.9	0.030	0.0037	0.011	0	0.026	0
w = 1	0.030	0.0037	0.011	0	0	0

Note: The MLE method, the Bayesian I method and the Bayesian II are used for estimating the VaR. The reported measures are the portfolio weights, the probability of underestimation times the average percentage of underestimation, and the relative increase in underestimation for each portfolio weight step.

Table 8 shows that underestimation decreases when the weight on asset type global increases, regardless of the chosen method. The relative decrease in underestimation ranges from 0 to 5 percent for the MLE method, 0 to 4.8 percent for the Bayesian I method and 0 to 8 percent for the Bayesian II method. It is important to mention that the economic impact of underestimation increase more strongly on the weight interval 0 to 0.6 for the Bayesian I and Bayesian II method than on the interval 0.6 to 1.

Table 9 shows the economic impact of overestimating the true VaR for the MLE method, the Bayesian I method and Bayesian II method. We observe that, if the weights on asset type global increases with 0.1, the relative overestimation decreases for all estimation methods. Moreover, the decrease is the largest on the weight interval 0.1 to 0.4, regardless of the chosen method.

Table 9: The economic impact of overestimating the VaR when portfolio weights vary.

Economic Impact	Overestimation			Relative Overestimation		
Weights.	MLE	Bay. I	Bay. II	MLE	Bay. I	Bay. II
w = 0	0.051	0.23	0.083			
w = 0.1	0.047	0.22	0.080	0.06	0.043	0.036
w = 0.2	0.047	0.21	0.077	0	0.045	0.0375
w = 0.3	0.046	0.21	0.075	0.021	0	0.026
w = 0.4	0.044	0.20	0.073	0.043	0.048	0.027
w = 0.5	0.043	0.197	0.071	0.023	0.015	0.027
w = 0.6	0.043	0.193	0.070	0	0.020	0.014
w = 0.7	0.042	0.190	0.069	0.023	0.016	0.014
w = 0.8	0.042	0.188	0.069	0	0.011	0
w = 0.9	0.042	0.188	0.069	0	0	0
w = 1	0.042	0.189	0.069	0	-0.005	0

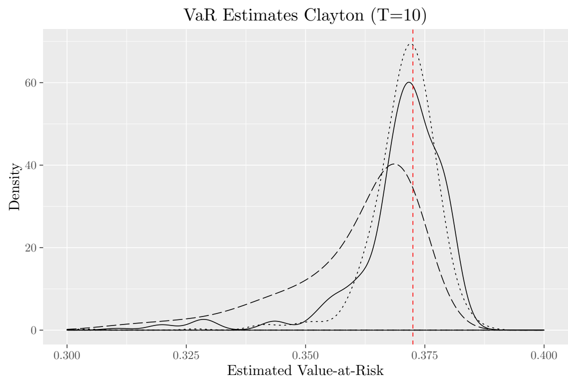
Note: The MLE method, the Bayesian I method and the Bayesian II are used under for estimating the VaR. The reported measures are the portfolio weights, the probability of overestimation times the average percentage of overestimation, and the relative increase in overestimation for each portfolio weight step.

4.4 Clayton Parameter Uncertainty

The bandwidth obtained using the Silverman's Rule of Thumb results in VaR density plots that are undersmoothed. To solve this problem I apply different scaling factors on the bandwidth and I find that scaling the bandwidth with 2 results in VaR density plots that are not under- or oversmoothed.

The figure below shows the obtained VaR density plots of the MLE method (dotted line), Kendall's method (solid line), and the Bayesian II method (dashed line) under the Clayton DGP. We notice that the VaR density of the Bayesian II method centers more to the left of the true VaR, while the VaR density of Kendall's method and MLE method position around the true VaR. The variation in the VaR estimates is the highest for the Bayesian II method and Kendall's method.

Figure 9: Density plots of the VaR estimates (Clayton parameter uncertainty).



Note: VaR estimates are estimated by the MLE method, Kendall's method, and the Bayesian II method under a Clayton DGP. The dotted line is the VaR density when the MLE method is used, the solid line the VaR density when Kendall's method is used and the dashed line the VaR density when the Bayesian II method is used. The dotted vertical line is the true VaR.

Table 10 reports the economic impact of misspecifying the VaR for the MLE method, Kendall's method, and Bayesian II method with respect to a Gaussian DGP and a Clayton DGP. From the table we can infer that, in case of the Clayton DGP, the probability of underestimating the true VaR equals 0.594, 0.599, and 0.901 for the MLE, Kendall's, and Bayesian II method respectively. The economic impact of the underestimation equals on average 0.40 percent for the MLE, 0.62 percent Kendall's method, and 1.09 percent for Bayesian II method. Thus, in terms of underestimation the MLE method is the best performing method.

Table 10: The economic impact of misspecifying the VaR (Clayton parameter uncertainty).

T = 10	Gaussian			Clayton		
Economic Impact.	MLE	Kendall	Bay. II	MLE	Kendall	Bay. II
Probability Underestimation	1	1	1	0.594	0.599	0.901
Probability times Underestimation	2.5	2.4	4.6	0.403	0.62	1.09
Probability times Overestimation	0	0	0	0.185	0.177	0.007
T = 25						
Probability Underestimation	1	1	1	0.598	0.664	0.887
Probability times Underestimation	2.4	2.3	3.1	0.086	0.19	0.22
Probability times Overestimation	0	0	0	0.14	0.12	0.005

Note: This table reports the probability of underestimating the true VaR, the probability of underestimating the true VaR times the average percentage of underestimation, and the probability of overestimation times the average percentage of overestimation. These measures are reported for VaR estimates obtained using the MLE method, Kendall's method, and the Bayesian II method. The table shows the measures both for 10 and 25 observations under a Gaussian DGP and a Clayton DGP.

Moreover, table 10 also shows that the Bayesian II method is the best performing in terms of overestimation. This method overestimates the true VaR on average with 0.007 percent, while the MLE method and Kendall's method overestimate with 0.185 and 0.177 percent. Expanding the sample dimension results in less under- and overestimation for all estimation methods.

Table 11 shows the statistical measures of the VaR estimates. In case of the Clayton DGP we can infer that the VaR estimates using the MLE method, Kendall's method, and the Bayesian II method have a negative bias of 0.7, 1.2, and 3.2 percent respectively. We also notice that the Bayesian II and Kendall's method have the largest variation in the VaR estimates with a standard deviation equal to 0.037 and 0.04 respectively. Overall, looking at the RMSE, the VaR estimates obtained using the MLE method are the most accurate.

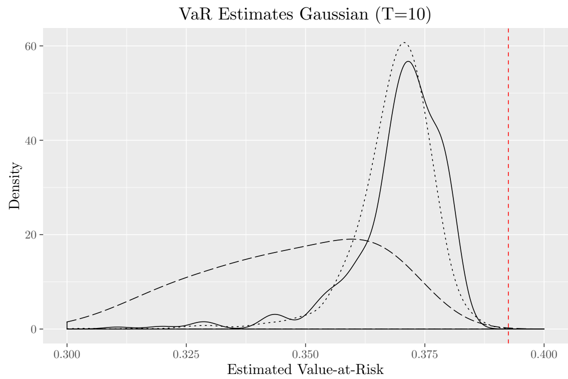
Table 11: The statistical impact of misspecifying the VaR (Clayton parameter uncertainty).

T = 10	Gaussian			Clayton		
Statistical Measures.	MLE	Kendall's	Bay. II	MLE	Kendall's	Bay. II
Relative VaR	0.937(0.028)	0.939(0.035)	0.884(0.046)	0.993(0.027)	0.988(0.04)	0.968(0.037)
RSME VaR	0.027	0.028	0.049	0.01	0.016	0.0184
T = 25						
Relative VaR	0.938(0.016)	0.941(0.015)	0.920(0.0252)	0.996(0.015)	0.992(0.017)	0.993(0.009)
RSME VaR	0.025	0.024	0.032	3.31e-5	4.48e-5	4.3e-3
True parameters:	VaR = 0.393	$\rho = 0.75$		VaR = 0.373	$\rho = 0.727$	

Note: The MLE method, Kendall's method and the Bayesian II method are used under a Gaussian and Clayton DGP for estimating the VaR. Reported in the table is the relative VaR, the standard deviation of the relative VaR (between brackets) and the Root-Mean-Square Error.

Figure 10 shows the VaR estimate density plots of the MLE method (dotted line), Kendall's method (solid line), and the Bayesian II method (dashed line) with respect to a Gaussian DGP. We notice that the VaR density plots of all estimation methods are centered left of the true VaR. Furthermore, the amount of variation in the Bayesian II VaR density plot is considerable larger compared to the Clayton DGP.

Figure 10: Density plots of the VaR estimates (Clayton uncertainty and Gaussian DGP).



Note: VaR estimates are estimated by using the MLE method, Kendall's method, and the Bayesian II method under a Gaussian DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of Kendall's method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

From table 10 we can infer that the probability of underestimating the true VaR becomes equal to 1 for all estimation methods. This can be explained by the increase in the true VaR in case of the Gaussian DGP. We also notice that the economic impact of this underestimation is the largest for the Bayesian II method, followed by the MLE method. This indicates that Kendall's method becomes the best performing method. Increasing the number of observations results in a lower impact of underestimation for all three method.

The reported statistical measures in table 11 indicate that the negative bias increases for all estimation methods. The amount of variation in the VaR estimates increases strongly for the Bayesian II method and decreases for the MLE method and Kendalls method. Looking at the RSMW we notice that the MLE method estimate the VaR most accurately. Expanding the sample dimension results in a better performance for all methods.

Overall, we observe that the MLE method is the best estimation method in terms of statistical accuracy and impact of underestimation.

Finally, I also compared the performance of the MLE method and Bayesian II method in estimating the dependency parameters of the copulas. Table 2 and 6 show that both methods have a lower probability of underestimating the true VaR in case of the correlation parameter. The impact of underestimation is also slightly lower. In terms of overestimation we notice that both the MLE and Bayesian II method are better able to estimate the Clayton parameter than the correlation parameter.

Part II

Empirical Part

5 Empirical example: Solvency II

In this section I use the results obtained from the simulation study to investigate the impact of incorporating parameter uncertainty on the Solvency II Capital Requirement for equity risk. First, I elaborate on the used data and then I discuss the results.

5.1 Data

I use the same Indexes as in the parameter calibration procedure of Solvency II, namely: MSCI World Developed Price Equity Index and the MSCI Emerging Markets BRICS. Before any analysis can be conducted I need to convert the total returns to actual returns. For this purpose I apply the same return definition as in QIS5 where:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{33}$$

I use yearly returns of the indexes, computed by using the last trading day of each year, to estimate the one year VaR and will abstain from using the annualization rolling window as applied in the QIS5 calibration papers. The reason for this choice is that Mittnik (2011) showed that this procedure leads to an unreliable correlation parameter and VaR estimates. Table 12 shows that none of the return indexes have a normal distribution. Furthermore, the Emerging Market indexes are indeed more volatile than the World Price Index and the kurtosis of both indexes are large, which indicates fat tails.

Table 12: Summary Statistics

	Mean.	Min.	Max.	Std.	Skew.	Kurt.	JB p-value	Correlation
MSCI World PR	0.113	-0.403	0.428	0.177	-0.77	0.510	5.053	
MSCI EM	0.148	-0.532	0.79	0.344	0.311	-0.726	1.067	0.665

Note: Reported are the mean, minimum, maximum, standard deviation, skewness, kurtosis, Jarque Bera p-value and the MSCI World Price Index and the MSCI Emerging Markets BRICS Price Index. The MSCI World Price Index starts on 31-12-1970, the MSCI Emerging Markets Brics Price Index on 30-12-1988 and both end on 08-06-2008. The correlation between the Indexes is also reported.

In the empirical part I investigate if the current Solvency II Capital Requirement for equity risk is sufficient when uncertainty is accounted for. I calibrate the Solvency II Capital Requirement using three models: the Gaussian copula model with fixed variances, the Clayton copula model with fixed variances and the Gaussian copula model in which no parameters are fixed. To account for parameter uncertainty I apply the same Bayesian methods as in the simulation study. I use two different data sets to perform my analysis on. The first data set consists of the MSCI World Price Index and the MSCI EM Price Index during the period 31-12-1987 until 31-12-2015. The second data set includes the same Indexes only now the lowest 10 observations of the MSCI EM Price Index are considered. The purpose of the second data set is to account for a higher correlation in the tail of the bivariate return distribution. The first data set is called the full data set and the second the adverse market scenario data set.

5.2 Results

In order to conduct the analysis the variance parameter values need to be determined. For the Gaussian and the Clayton copula model I use the sample variances of the full data set. The same applies to the adverse market scenario. The portfolio weights are set equal to 0.75.

Table 13 shows the VaR estimates of the MLE method, the Bayesian I/Kendall's method, and the Bayesian II method using a Gaussian copula with fixed variances, a Clayton copula with fixed variances and a Gaussian copula without fixed variances. We observe that the VaR estimate, under the Gaussian copula, equals 0.525 for the MLE method, 0.531 for the Bayesian I method, and 0.52 for the Bayesian II method. Under the adverse market scenario the VaR estimate of the MLE method, the Bayesian I method, and Bayesian II method becomes 0.522, 0.499, and 0.528 respectively.

Table 13: Empirical VaR estimates

MSCI EM	Gaussian		
	MLE	Bay. I\Kendall's	Bay. II
Gaussian fixed variances			
Full data set	0.525	0.531	0.52
Adverse market scenario	0.522	0.499	0.528
Clayton fixed variances			
Full data set	0.51	0.508	0.501
Adverse market scenario	0.487	0.471	0.457
Gaussian no fixed variances			
Full data set	0.588	0.66	0.605
Adverse market scenario	0.466	0.667	0.513
Current capital requirement = 0.393			

Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method in case of the Gaussian copula model. The Bayesian I method is replaced by Kendall's method in case of the Clayton copula. The VaR estimates are estimated on the full data set and the adverse market scenario data set. The current capital requirement corresponds to the capital requirement of equity risk under the Solvency II regulation.

When using the Clayton copula, we notice that the VaR estimate of the MLE method, Kendall's method, and the Bayesian II method equals 0.51, 0.508, and 0.501 respectively. The VaR estimates of the three methods do not differ much, which is in line with the results obtained in the simulation study in case of a Clayton DGP. Under the adverse market scenario, the VaR estimate decreases to 0.487, 0.471, and 0.457 for the MLE method, Kendall's method, and Bayesian II method respectively. We observe that the difference between the VaR estimates of the three methods increases under the adverse market scenario and the VaR estimate is lowest for the Bayesian II method. The simulation study shows similar results when fewer data is considered.

We expected that the VaR estimates under adverse market scenario would be higher; because correlation tends to be higher in the tail of the returns' distribution. However, the sample correlation of the returns under the adverse market scenario turns out to be lower than the sample correlation of the returns in the full data set. One explanation could be that the chosen trading day, used to compute the yearly returns, affects the correlation. To rule out this potential influence, I therefore also compute the VaR estimates when the first day of the month June is used to construct the yearly returns.

Table 15 in Appendix D shows the VaR estimates when the first trading day of June is used to compute the yearly returns. We infer from the results that the VaR estimates of the MLE method increase, as expected, under the adverse market scenario. However, the VaR estimates of the Bayesian methods do not decrease. This difference can be explained by the decrease in number of observations in the data set. When few data is available, the amount of uncertainty in the posterior distribution of the parameters increases. As a result, the probability of having a correlation lower than the sample correlation also increases, which results in lower VaR estimates.

We also notice that the VaR estimates of all methods are lower than the VaR estimates in table 13. To further mitigate the impact of the chosen trading day on the VaR estimates I compute the VaR estimates on 12 yearly return data sets, where the yearly returns are computed by taking the first trading day of each month. Subsequently I take an equally weighted average over the VaR estimates obtained from each data set.

Table 14 shows the equally weighted combination of the VaR estimates using the MLE method, the Bayesian I method, Kendall's method and Bayesian II method. We observe that most of the VaR estimates are between 0.40 and 0.48. Only the Bayesian II method, under the Clayton model, has a VaR estimate equal to 0.362 and 0.33 for the full data set and the adverse market scenario respectively. These relative low VaR estimates can be explained by figure 10 which shows that the Bayesian II method heavily underestimates the true VaR in case of an incorrect model specification.

Table 14: Empirical equally weighted VaR estimates

MSCI EM	Gaussian MLE	Bay. I\Kendall's	Bay. II
Gaussian fixed variances			
Full data set	0.468	0.48	0.469
Adverse market scenario	0.465	0.451	0.457
Clayton fixed variances			
Full data set	0.452	0.453	0.362
Adverse market scenario	0.452	0.444	0.331
Gaussian no fixed variances			
Full data set	0.437	0.502	0.453
Adverse market scenario	0.507	0.693	0.548
Current capital requirement = 0.393			

Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method in case of the Gaussian copula model. The Bayesian I method is replaced by Kendall's method in case of the Clayton copula. The VaR estimates are an equally weighted combination of the VaR estimates obtained using the full data set and the adverse market scenario data set. In case of the Gaussian model with no fixed variances the mean of the returns is subtracted from the VaR estimates. The current capital requirement corresponds to the capital requirement under the Solvency II regulation.

I also estimate the VaR by using a Gaussian copula model without fixed parameters. Table 13 shows that the MLE method has a VaR estimate equal to 0.588, the Bayesian I method a VaR estimate equal to 0.66 and the Bayesian II method a VaR estimate equal to 0.605. Under the adverse market scenario the VaR estimates decrease for all estimation methods. An explanation for this decrease is that the variance estimates under the adverse market scenario are lower than the variance estimates under the full data set. Because the mean is set equal to zero, the VaR only depends on the estimated variances and correlation. In case the latter becomes higher while the former becomes lower, the VaR estimate can still decrease. To account for this effect, I subtract the mean of the returns from the VaR for all three methods. This results in a VaR estimate of 0.486, 0.560, and 0.503 for the MLE method, the Bayesian I method, and the Bayesian II method respectively. Under the adverse market scenario the VaR estimate equals 0.579, 0.780, and 0.622 for the MLE method, the Bayesian I method, and the Bayesian II method respectively. These results indicate that the VaR increases in an adverse market scenario.

Table 14 also shows the equally weighted VaR estimates, with the mean substracted, under the Gaussian copula model. We observe that the VaR estimate of the MLE method, the Baysian I method, and the Bayesian II method equals 0.437, 0.502, and 0.453 respectively. Under the adverse market scenario the VaR estimates of the MLE method increases to 0.507, the VaR estimate of the Bayesian I method to 0.693 and the VaR estimate of the Bayesian II method to 0.548.

Overall, we observe that almost all equally weighted VaR estimates are higher than the current Solvency II Capital Requirement of 0.393, except for the equally weighted VaR estimates of the Bayesian II method with respect to a Clayton copula model.

Based on the simulation study and the empirical example results, the best Solvency Capital Requirement estimate equals 54 percent when a portfolio weight of 0.75 on asset type global is considered. There are a number of reasons for this estimate. First, the capital requirements obtained in the full data set are too low, because the Gaussian copula model does not account for higher correlations in the tail of the Indexes. Secondly, the simulation study shows that the Bayesian I method overestimates the true capital requirement with an additional 24 percent, therefore a capital requirement of 69 percent would be too strict. Furthermore, the simulation study shows that the MLE method underestimated the true capital requirement on average with 3 percent. Thus, a Solvency II Capital Requirement higher than the VaR estimate of the MLE method would be desirable. Adding the 3 percent results in a Solvency Capital Requirement estimate of 54 percent.

6 Conclusion and Discussion

In this thesis I investigated the impact of incorporating parameter uncertainty on the 99.5 percent one year VaR. The simulation study results show that including uncertainty in the correlation parameter resulted in VaR estimates that underestimate the true VaR more often compared to the MLE VaR estimates. Furthermore, the economic impact of underestimation is larger. In case of including uncertainty in the variances I found that incorporating parameter uncertainty results in a lower probability of underestimation and a lower economic impact of underestimating the true VaR as compared to the MLE estimates. The inclusion of parameter uncertainty in the both correlation and variances results in almost no underestimation of the true VaR. However, the economic impact of overestimation is rather large. Including uncertainty in the Clayton copula parameter results in more underestimation as compared to the MLE estimates, however the economic impact was low. Overall, my conclusion is that incorporating parameter uncertainty using the posterior predictive distribution results in more variation in the VaR estimates, a lower probability of underestimation, a lower economic impact of underestimation, and a higher economic impact of overestimation compared to the VaR estimates obtained using the posterior distribution. I also found that the current Solvency Capital Requirements are too low if yearly return data is used and parameter uncertainty is accounted for.

The conclusions above are subject to some constraints. Firstly, I adopted a uniform prior on the correlation parameter that included some information. A more uninformative prior could produce different results. Secondly, I only considered uncertainty in the Clayton copula parameter, while the variance parameters where kept fixed. This makes the results obtained under the Clayton copula model less relevant in practice, because then the variances are not known.

The analysis in my thesis could be further developed in several ways. Firstly, the impact of incorporating uncertainty in both the variance parameter and the Clayton copula parameter could be determined. Secondly, the impact of incorporating parameter uncertainty could be determine while estimating the variances with monthly data.

The empirical example shows that including parameter uncertainty results in a higher capital requirement than under the current Solvency II calibration. The question that remains is whether including parameter uncertainty is necessary from a practical point of view. To answer this question I consider three stakeholders: the policyholder, the insurance company and the supervisor authority.

The policyholder interests are twofold: On the one hand the policyholder benefits from a low premium on the insurance product, while enough capital should be available to pay for any possible claims on the other hand. Therefore, the capital requirement needs to be chosen in a way that plenty of capital will be available to make a profit and maintain a low premium, but enough capital should remain to secure future obligation by the insurance company. Based on the simulation study results, the policyholder should prefer the Bayesian II method, because it strikes a good balance between unnecessarily high premiums on the one hand (low overestimation), and enough security that the insurance company fulfills future obligations on the other hand (low underestimation).

The insurance company is mostly interested in making profit for its shareholders and maintaining a competitive advantage over the other insurance companies by keeping the insurance premium low. The simulation study shows that the MLE method underestimates the true VaR most often. Applying this method would most likely result in a VaR lower than the true VaR, but since the insurance company interest is to have more capital available, a low capital requirement will be desirable.

The main interest of the supervisory authorities is to protect stability of the financial sector, to keep insurance companies financially healthy, and to make sure that the policyholders are protected in a way that insurance companies will fulfill their future obligations. The Bayesian I method offers the most certainty that the true VaR is not underestimated as has been shown in the simulation study. However, this extra security makes the capital requirement too high, which implicitly impacts the amount of premium the policyholder has to pay. The Bayesian II method offers less overestimation, but at the expense of a higher chance on underestimation. Based on the results of both methods, incorporating parameter uncertainty would be desirable. However, the question which method to use depends on the weight a supervisory authority attributes to stability of the insurance sector and financial health of the insurance companies on the one hand and keeping an affordable insurance products on the other hand.

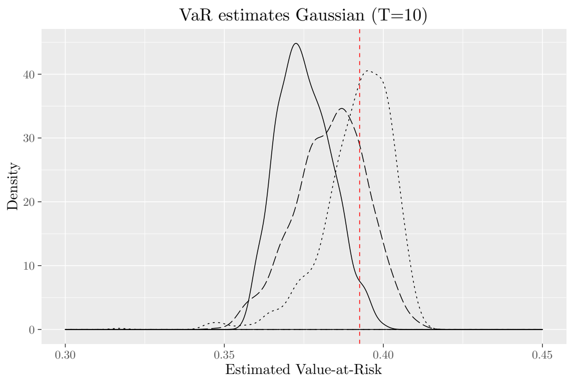
Part III
Appendix

7 Appendix A: VaR Densities

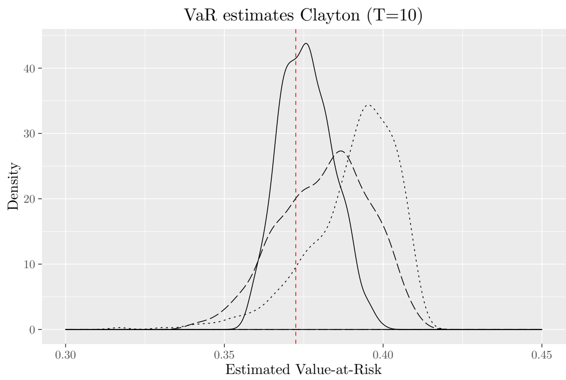
7.1 Correlation Uncertainty

Figure 11: VaR estimates density plots (correlation uncertainty).

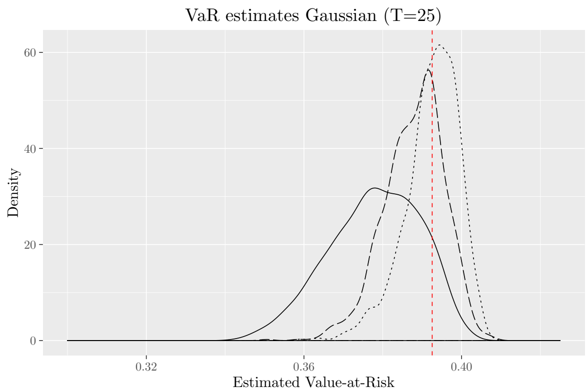
(a) VaR estimates under a Gaussian DGP



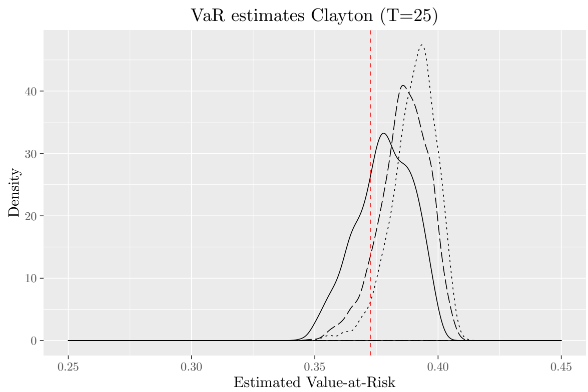
(b) VaR estimates under a Clayton DGP



(c) VaR estimates under a Gaussian DGP



(d) VaR estimates under a Clayton DGP

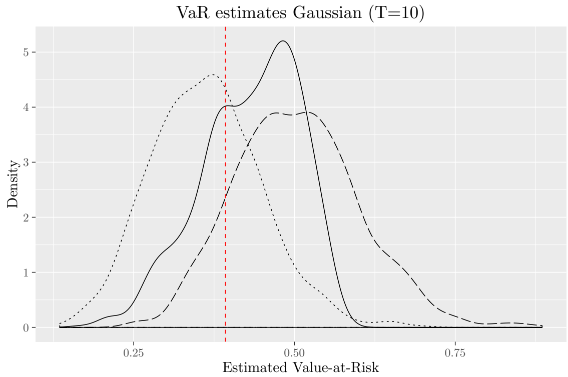


Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method under a Gaussian DGP and Clayton DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of the Bayesian I method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

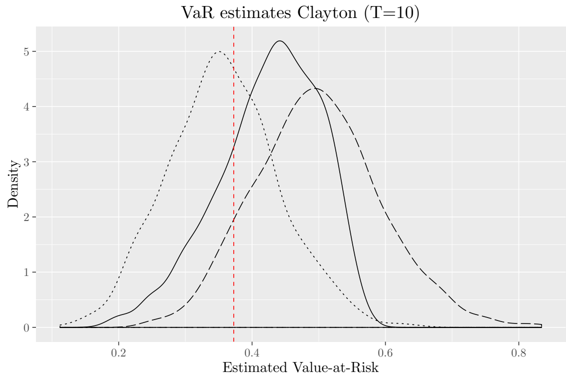
7.2 Variance Uncertainty

Figure 12: VaR estimates density plots (variance uncertainty).

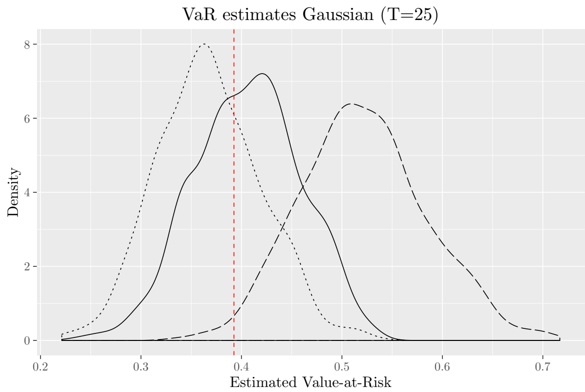
(a) VaR estimates under a Gaussian DGP



(b) VaR estimates under a Clayton DGP



(c) VaR estimates under a Gaussian DGP



(d) VaR estimates under a Clayton DGP

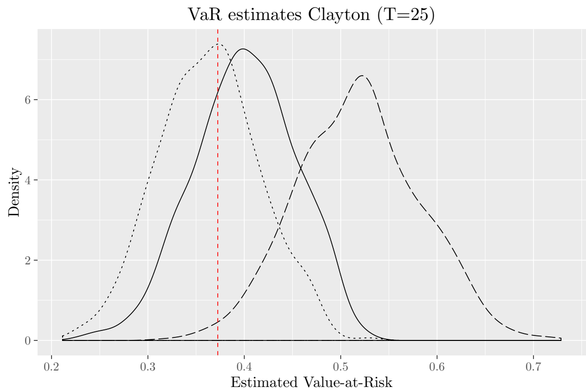
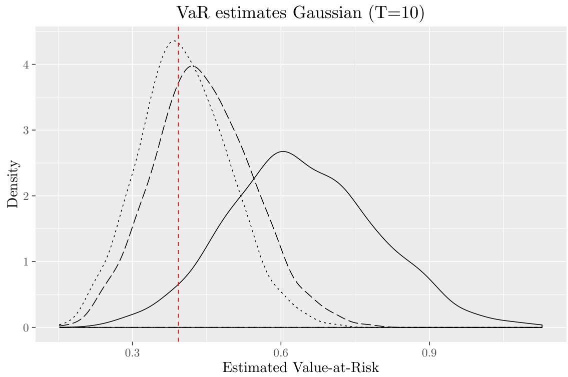


Figure 13: Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method under a Gaussian DGP and Clayton DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of the Bayesian I method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

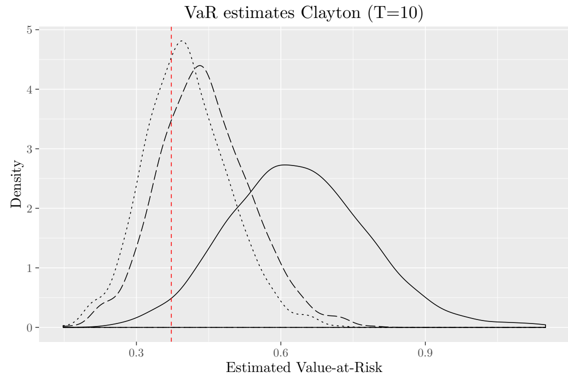
7.3 Joint Uncertainty

Figure 14: VaR estimates density plots (joint uncertainty).

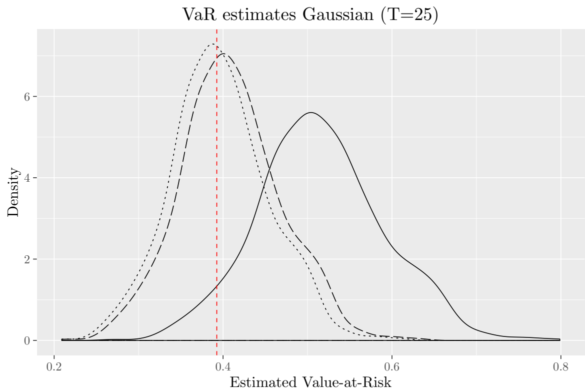
(a) VaR estimates under a Gaussian DGP



(b) VaR estimates under a Clayton DGP



(c) VaR estimates under a Gaussian DGP



(d) VaR estimates under a Clayton DGP

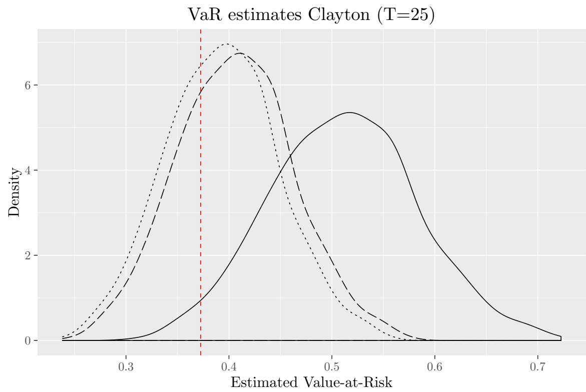
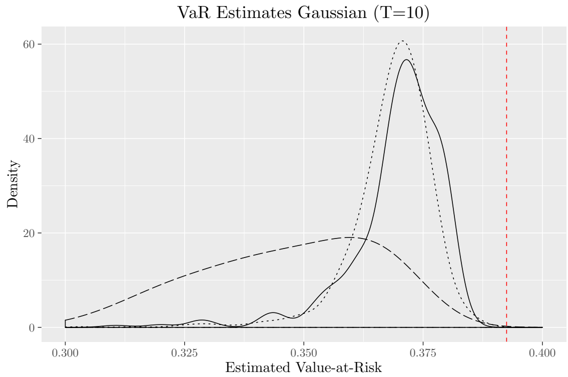


Figure 15: Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method under a Gaussian DGP and Clayton DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of the Bayesian I method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

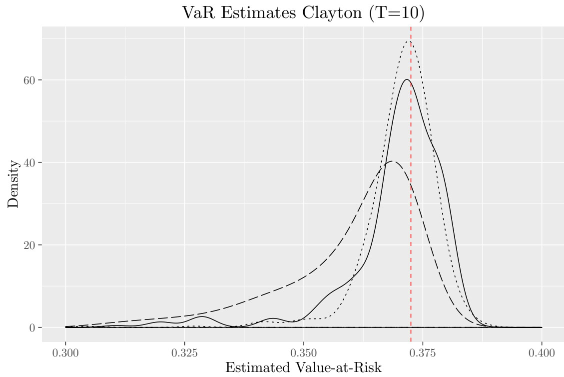
7.4 Clayton Uncertainty

Figure 16: VaR estimates density plots (Clayton parameter uncertainty).

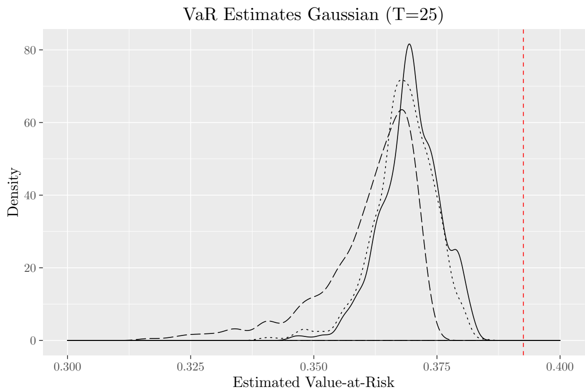
(a) VaR estimates under a Gaussian DGP



(b) VaR estimates under a Clayton DGP



(c) VaR estimates under a Gaussian DGP



(d) VaR estimates under a Clayton DGP

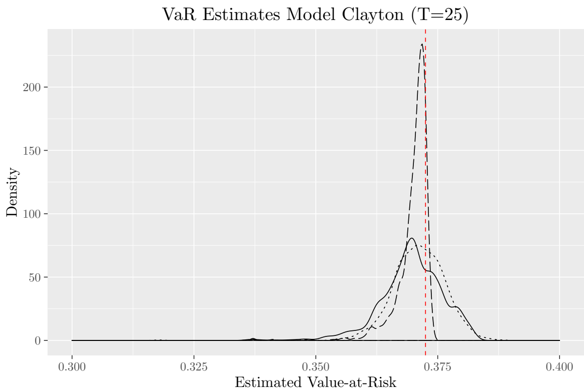


Figure 17: Note: VaR estimates are estimated by using the MLE method, Kendall's method, and the Bayesian II method under a Gaussian DGP and Clayton DGP. The dotted line is the VaR density of the MLE method, the solid line the VaR density of Kendall's method and the dashed line the VaR density of the Bayesian II method. The dotted vertical line is the true VaR.

8 Appendix B: (Predictive) Posterior Derivations

8.1 Derivation I

We know that $p(\rho|\mathbf{R_p},\theta) \propto p(\rho)p(\mathbf{R_p}|\theta,\rho)$ and that $p(R_p|\theta,\rho)$ is normally distributed with mean equal to $\gamma'\mu$ and variance $\gamma'\Sigma\gamma$. Combine both distributions to obtain:

$$\begin{aligned} p(\rho|\mathbf{R_p},\theta) \propto \frac{1}{\rho_u-\rho_l}&\left(\frac{1}{2\pi\sqrt{w^2\sigma_1^2(1-w)\sigma_2^2+2\rho w(1-w)\sigma_1\sigma_2}}\right)^T \\ &\times \exp\{-\sum_{i=1}^{T}\frac{(R_{p,i}-\gamma'\mu)^2}{2(w^2\sigma_1^2+(1-w)\sigma_2^2+2\rho w(1-w)\sigma_1\sigma_2)}\} \end{aligned} \tag{34}$$

where $\mathbf{R_p}$ includes all the portfolio return observations.

8.2 Derivation II

From expressions (16) it follows that:

$$\begin{aligned} p(\sigma_1^2,\sigma_2^2|\mathbf{R},\phi) \propto \sigma_1^{-2(\nu_1+1)}\sigma_2^{-2(\nu_2+1)}&\left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\right)^T \\ \times \exp\{-\sum_{i=1}^{T}&\frac{\sigma_2^2(R_{i1}-\mu_1)^2+\sigma_1^2(R_{i2}-\mu_2)^2-2\rho\sigma_1\sigma_2(R_{i1}-\mu_1)(R_{i2}-\mu_2)}{2\sigma_1^2\sigma_2^2(1-\rho^2)}\} \\ \times \exp\{-&\frac{\beta_1}{\sigma_1^2}-\frac{\beta_1}{\sigma_1^2}\} \end{aligned} \tag{35}$$

by re-arranging the terms we obtain:

$$\begin{aligned} p(\sigma_1^2,\sigma_2^2|\mathbf{R},\phi) \propto \sigma_1^{-(T+2+2\nu_1)}&\exp\{\sum_{i=1}^{T}-\frac{(R_{i1}-\mu_1)^2+2(1-\rho^2)\beta_1}{2(1-\rho^2)\sigma_1^2}\} \\ \times \sigma_2^{-(T+2+2\nu_2)}&\exp\{\sum_{i=1}^{T}-\frac{(R_{i2}-\mu_2)^2+2(1-\rho^2)\beta_2}{2(1-\rho^2)\sigma_2^2}+\frac{\rho(R_{i1}-\mu_1)(R_{i2}-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)}\} \end{aligned} \tag{36}$$

We know that $p(\sigma_1^2,\sigma_2^2|\mathbf{R},\phi) = p(\sigma_2^2|\sigma_1^2,\mathbf{R},\phi)p(\sigma_1^2|\mathbf{R},\phi)$ and by combining all terms of σ_1^2 we obtain:

$$\begin{aligned} p(\sigma_2^2|\sigma_1^2,\mathbf{R},\phi) \propto\ &\sigma_2^{-(T+2+2\nu_2)} \\ &\times \exp\{\sum_{i=1}^{T}-\frac{(R_{i2}-\mu_2)^2+2(1-\rho^2)\beta_2}{2(1-\rho^2)\sigma_2^2}+\frac{\rho(R_{i1}-\mu_1)(R_{i2}-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)}\} \end{aligned} \tag{37}$$

and

$$p(\sigma_1^2|\mathbf{R},\phi) \propto \sigma_1^{-(T+2+2\nu_1)}\exp-\sum_{i=1}^{T}\frac{(R_{i1}-\mu_1)^2+2(1-\rho^2)\beta_1}{2(1-\rho^2)\sigma_1^2} \tag{38}$$

where $\sigma_1^2|\mathbf{R},\phi$ is proportional to an Inverse Gamma distribution with parameters $\alpha=\frac{T}{2}+\nu_1$ and $\beta=\sum_{i=1}^{T}\frac{(R_{i1}-\mu_1)^2}{2(1-\rho^2)}+\beta_1$. It should be noted that $p(\sigma_2^2|\sigma_1^2,\mathbf{R},\phi)$ is almost an Inverse Gamma distribution, however there is an extra term $\frac{\rho(R_{i1}-\mu_1)(R_{i2}-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)}$.

8.3 Derivation III

From expression (19) it follows that:

$$p(R_{T+1}|\hat{R}) = \int p(R_{T+1}|\Sigma, \hat{R})p(\Sigma|\hat{R})d\Sigma \tag{39}$$

The density functions within the integral are:

$$p(\Sigma|\hat{R}) \propto |\Sigma|^{-(T+3)/2} \exp(-1/2tr[\Sigma^{-1}\hat{R}'\hat{R}]) \tag{40}$$

and

$$p(R_{T+1}|\Sigma, \hat{R}) \propto |\Sigma|^{-1/2} \exp(-1/2tr[\Sigma^{-1}R_{T+1}R'_{T+1}]) \tag{41}$$

Combining both expressions gives the following formula:

$$p(R_{T+1}|\Sigma, \hat{R})p(\Sigma|\hat{R}) \propto |\Sigma|^{-(T+4)/2} \exp(-1/2tr[\Sigma^{-1}(\hat{R}'\hat{R} + R_{T+1}R'_{T+1})]) \tag{42}$$

by defining $V = (\hat{R}, R_{T+1})$ I obtain $V'V = \hat{R}'\hat{R} + R_{T+1}R_{T+1}$ and (39) becomes:

$$p(R_{T+1}|\hat{R}) \propto \int |\Sigma|^{-(T+4)/2} \exp(-1/2tr[\Sigma^{-1}(V'V)])d\Sigma \tag{43}$$

It should be noted that the integral can be solved by using the Inverted Wishart integration step:

$$\int |\Sigma|^{-M/2} exp(-1/2tr[\Sigma^{-1}A])d\Sigma \propto |A|^{-1/2(M-J-1)} \tag{44}$$

where J is the row/column dimension of Σ. Applying this step to (43) and further simplifying the expression leads to the following:

$$\begin{aligned} p(R_{T+1}|\hat{R}) &\propto |V'V|^{-(T+1)/2} \\ &\propto |\hat{R}'\hat{R} + R_{T+1}R'_{T+1}|^{-(T+1/2)} \\ &\propto |\hat{R}'\hat{R}(I_2 + (\hat{R}'\hat{R})^{-1}R_{T+1}R'_{T+1})|^{-(T+1)/2} \\ &\propto |\hat{R}'\hat{R}|^{-(T+1)/2}|I_2 + (\hat{R}'\hat{R})^{-1}R_{T+1}R'_{T+1}|^{-(T+1)/2} \\ &\propto |\hat{R}'\hat{R}|^{-(T+1)/2}(1 + R'_{T+1}(\hat{R}'\hat{R})^{-1}R_{T+1})^{-(T+1)/2} \end{aligned} \tag{45}$$

By multiplying and dividing (45) by $(\frac{T-1}{T-1})^{-(T+1)/2}$ I obtain:

$$\begin{aligned} p(R_{T+1}|\hat{R}) &\propto |V'V|^{-(T+1)/2} \\ &\propto \left(\frac{|\hat{R}'\hat{R}|}{T-1}\right)^{-(T+1)/2} \left((T-1) + R'_{T+1}\left(\frac{\hat{R}'\hat{R}}{T-1}\right)^{-1} R_{T+1}\right)^{-(T+1)/2} \\ &\propto \left(\frac{|\hat{R}'\hat{R}|}{T-1}\right)^{-1/2} \left((T-1) + R'_{T+1}\left(\frac{\hat{R}'\hat{R}}{T-1}\right)^{-1} R_{T+1}\right)^{-(T+1)/2} \end{aligned} \tag{46}$$

which shows that $p(R_{T+1}|\hat{R})$ follows a Multivariate T distribution with scale parameter $A = \frac{\hat{R}'\hat{R}}{T-1}$, location parameter 0_2 and T-1 degrees of freedom.

9 Appendix C: Copula Derivations

9.1 Derivation I

Recall that the joint density of the returns can be written as:

$$f(R_1, R_2; \alpha, \theta_1, \theta_2) = c(F_1(R_1; \theta_1), F_2(R_2; \theta_2); \alpha) f_1(R_1; \theta_1) f_2(R_2; \theta_2) \tag{47}$$

where α is the Clayton parameter and $\theta_1 = (\mu_1, \sigma_1^2)$ and $\theta_2 = \mu_2, \sigma_2^2)$ the parameter vectors of the marginals. Considering $\mathbf{R} = (R_{1t}, R_{2t})_{t=1}^{T}$ the sample matrix then the likelihood function is given by:

$$\begin{aligned} f(\mathbf{R}; \alpha, \theta_1, \theta_2) = & \prod_{t=1}^{T} \left(-1 + u_{1t}^{-\alpha} + u_{2t}^{-\alpha}\right)^{(-2-1/\alpha)} u_{1t}^{(-\alpha-1)} u_{2t}^{(-\alpha-1)} (\alpha + 1) \\ & \times (\sqrt{2\pi}\sigma_1)^{-1} \exp\left(-\frac{(R_{1t} - \mu_1)^2}{2\sigma_1^2}\right) (\sqrt{2\pi}\sigma_2)^{-1} \exp\left(-\frac{(R_{2t} - \mu_2)^2}{2\sigma_2^2}\right) \end{aligned} \tag{48}$$

where $u_{1t} = F_1^{-1}(R_{1t}; \theta_1)$ and $u_{2t} = F_2^{-1}(R_{2t}; \theta_2)$. The log-likelihood of (47) then equals:

$$\begin{aligned} L(\mathbf{R}; \alpha, \theta_1, \theta_2) &= \sum_{t=1}^{T} (\ln c(F_1(R_{1t}; \theta_1), F_2(R_{2t}; \theta_2); \alpha) + \ln f_1(R_{1t}; \theta_1) + \ln f_2(R_{1t}; \theta_2)) \\ &= \sum_{t=1}^{T} (\ln c(u_{1t}, u_{2t}; \theta_1, \theta_2, \alpha) + \ln f_1(R_{1t}; \theta_1) + \ln f_2(R_{1t}; \theta_2)) \\ &= \sum_{t=1}^{T} (-2 - 1/\alpha) \ln\left(-1 + u_{1t}^{-\alpha} + u_{2t}^{-\alpha}\right) + \sum_{t=1}^{T} (-\alpha - 1)(\ln u_{1t} + \ln u_{2t}) \\ &+ T \ln(\alpha + 1) + \sum_{t=1}^{T} \left(-\ln(\sqrt{2\pi}\sigma_1) - \frac{1}{2}\frac{(R_{1t} - \mu_1)^2}{\sigma_1^2}\right) \\ &+ \sum_{t=1}^{T} \left(-\ln(\sqrt{2\pi}\sigma_2) - \frac{1}{2}\frac{(R_{2t} - \mu_2)^2}{\sigma_2^2}\right) \end{aligned} \tag{49}$$

10 Appendix D: Empirical Part Extra Table

Table 15: Empirical VaR estimates (trading day June).

MSCI EM	Gaussian MLE	Bay. I\Kendall's	Bay. II
Gaussian fixed variances			
Full data set	0.383	0.403	0.384
Adverse market scenario	0.397	0.381	0.38
Clayton fixed variances			
Full data set	0.374	0.372	0.377
Adverse market scenario	0.385	0.387	0.347
Current capital requirement = 0.393			

Note: VaR estimates are estimated by using the MLE method, the Bayesian I method, and the Bayesian II method in case of the Gaussian copula model. The Bayesian I method is replaced by Kendall's method in case of the Clayton copula. The VaR estimates are estimated on the full data set and the adverse market scenario data set. The yearly return data is computed by using the first trading day of June. The current capital requirement corresponds to the capital requirement under the Solvency II regulation.

11 Appendix E: Matricvariate Distributions

11.1 Matricvariate Normal Distribution

The probability density function of a matricvariate normally distributed $K \times J$ random matrix $\mathbf{Z}$ with mean $\mathbf{M}$ $(K \times J)$ and covariance matrix $\mathbf{S} \otimes \mathbf{L}$ $(\mathbf{S} : J \times J, \mathbf{L} : k \times k)$, that is, $\mathbf{Z} \sim MN(\mathbf{M}, \mathbf{S} \otimes \mathbf{L})$, is given by:

$$
\begin{aligned}
p(\mathbf{Z}|\mathbf{M}, \mathbf{S} \otimes \mathbf{L}) = &(\sqrt{2\pi})^{-kJ}|\mathbf{S}|^{-k/2}|\mathbf{L}|^{-J/2} \\
&\times \exp\left(-\frac{1}{2}tr[\mathbf{S}^{-1}(\mathbf{Z}-\mathbf{M})'\mathbf{L}^{-1}(\mathbf{Z}-\mathbf{M})]\right).
\end{aligned} \tag{50}
$$

This implies that vec$[\mathbf{Z}] \sim N(vec[\mathbf{M}], \mathbf{S} \otimes \mathbf{L})$

$$
\begin{aligned}
p(vec[\mathbf{Z}]|\mathbf{M}, \mathbf{S} \otimes \mathbf{L}) = &(\sqrt{2\pi})^{-kJ}|\mathbf{S} \otimes \mathbf{L}|^{-1/2} \\
&\times \exp\left(-\frac{1}{2}vec[\mathbf{Z}-\mathbf{M}]'(\mathbf{S} \otimes \mathbf{L})^{-1}vec[\mathbf{Z}-\mathbf{M}]\right),
\end{aligned} \tag{51}
$$

and shows that it is just another parameterization of a multivariate normal distribution with a restricted covariance matrix.

11.2 Inverted Wishart Distribution

The probability density function of an inverted Wishart distributed random $J \times J$ positive definite symmetric matrix $\mathbf{Z}$ with parameters $J \times J$ positive definite symmetric matrix $\mathbf{S}$ and degrees of freedom $v \geq J - 1$, that is, $\mathbf{Z} \sim IW(\mathbf{S}, v)$, is given by:

$$
p(\mathbf{Z}|\mathbf{S}, v) = c \times |\mathbf{S}|^{v/2}|\mathbf{Z}|^{-(v+J+1)/2}\exp\left(-\frac{1}{2}tr[\mathbf{Z}^{-1}\mathbf{S}]\right), \tag{52}
$$

where c is an integrating constant.

The mean of $\mathbf{Z}$ is:

$$
E[\mathbf{Z}] = \frac{1}{v - J - 1}\mathbf{S} \tag{53}
$$

If $\mathbf{Z} \sim IW(\mathbf{S}, v)$ then $\mathbf{Z}^{-1} \sim W(\mathbf{S}, v)$ (Wishart Distribution)

11.3 Multivariate Student's T Distribution

Cornish (1994), and Dunnett and Sobel (1954) defined the multivariate Student's T distribution for $\mathbf{T}$ a $p \times 1$ vector as:

$$
p(\mathbf{T}) = \frac{C_p|\mathbf{\Sigma}|^{-1/2}}{[n + (\mathbf{T}-\theta)'\mathbf{\Sigma}^{-1}(\mathbf{T}-\theta)]^{(n+p)/2}} \tag{54}
$$

where $-\infty < \mathbf{T_j} < \infty$ for $j = 1, 2 \ldots, p$, $n \geq 0$ the degrees of freedom, θ: $p \times 1$ the location parameter, $\mathbf{\Sigma}$: $p \times p$, $\mathbf{\Sigma} > 0$ the scale matrix, and where

$$
C_p = \frac{n^{n/2}\Gamma\left(\frac{n+p}{2}\right)}{\pi^{p/2}\Gamma\left(\frac{n}{2}\right)} \tag{55}
$$

The mean and covariance matrix of $\mathbf{T}$ are easily shown to be

$$
E(\mathbf{T}) = \theta, \qquad n > 1, \tag{56}
$$

$$V(\mathbf{T}) = \frac{n}{n-2}\mathbf{\Sigma}, \qquad n > 2. \tag{57}$$

In my paper I use the result that $\gamma'\mathbf{T}$ is univariate Student's T distributed if $\mathbf{T}$ is multivariate Student's T distributed. This follows from the following theorem:

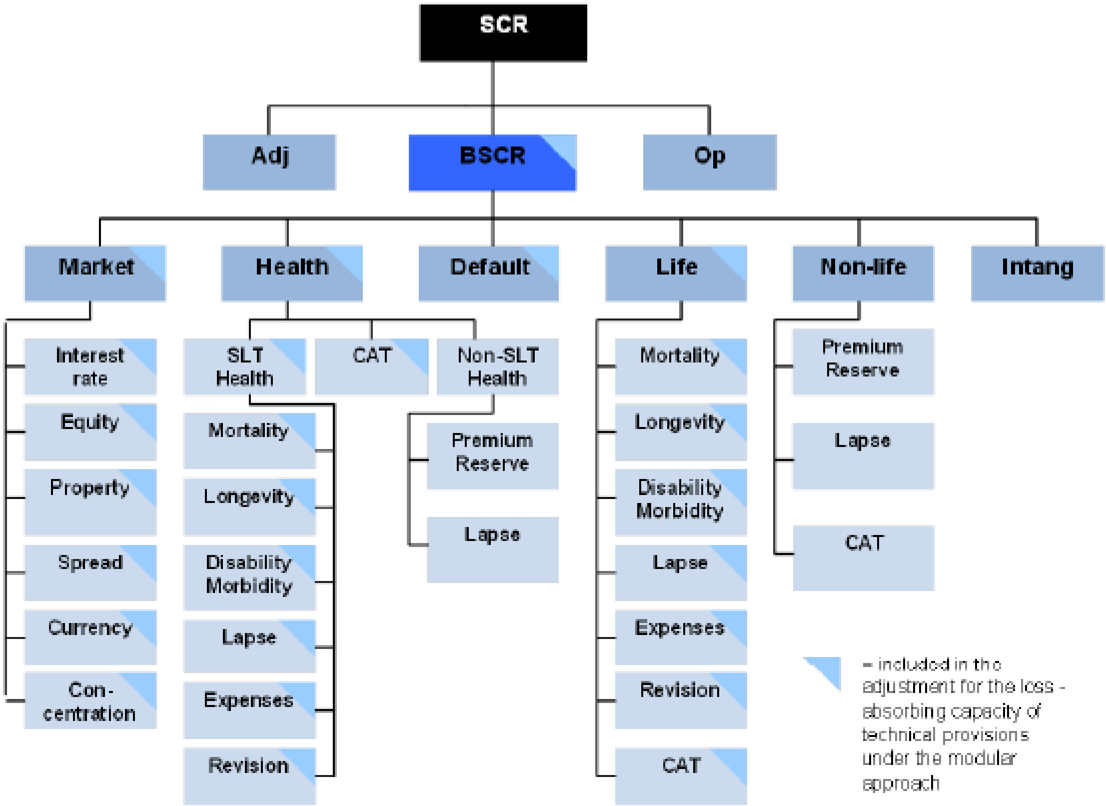
Let $\mathbf{T}$ be a multivariate Student's T distribution with density given by (54). Then, if $y = \mathbf{AT} + \mathrm{b}$, where $\mathbf{A}$: $q \times p$, b: $q \times 1$, $q \leq p$, the density of y is given by:

$$p(y) = \frac{C_p|\mathbf{\Phi}|^{-1/2}}{[n + (y - \mathbf{A}\theta - b)'\mathbf{\Phi^{-1}}(y - \mathbf{A}\theta) - b]^{(n+q)/2}}, \tag{58}$$

where $\Phi = \mathbf{A\Sigma A'}$ (James, 2005).

12 Appendix F: Figure Solvency II Standard Model

Figure 18: Solvency II Standard Model



References

[1] Eiopa report on the fifth quantitative impact study (qis5) for solvency ii. Technical report.

[2] Qis5 technical specifications. Technical report.

[3] Aussenegg, W. and Miazhynskaia, T. (2006). Uncertainty in value-at-risk estimates under parametric and non-parameteric modeling. *Financial Markets and Portfolio Management*, 20(3):243–264.

[4] Christiansen, M., Denuit, M., and Lazar, D. (2012). The solvency ii square-root formula for systematic biometric risk. *Insurance: Mathematics and Economics*, 50(2):257–265.

[5] Dowd, K. (2000a). Assessing var accuracy. *Derivatives Quarterly*, pages 61–63.

[6] Dowd, K. (2000b). Estimatin value at risk - a subjective approach. *Journal of Risk Finance*, 1(4):43–46.

[7] Eling, M. and Pankoke, D. (2014). Basis risk, procyclicality, and systemic risk in the solvency ii equity risk module. *Journal of Insurance Regulation*, 33(1).

[8] Fantazzini, F. (2008). The effects of misspecified marginals and copulas computing the value at risk: A monte carlo study. *Computational Statistics and Data analysis*, 53(6):2168–2188.

[9] Fosdick, B. and Raftery, A. (2012). Estimating the correlation in bivariate normal data with known variances and small sample sizes. Technical report, Department of Statistics: University of Washington.

[10] Gatzert, N. and Martin, M. (2012). Quantifying credit and market risk under solvency ii: Standard approach vs internal model. *Insurance: Mathemetics and Economics*, 51(3):649–666.

[11] Gokhale, D. and Press, J. (1982). Assessment of a prior distribution for the correlation coefficient in a bivariate normal distribution. *Journal of the Royal Statistical Society*, 145(2):237–249.

[12] Greenberg, E. (2008). *Introduction to Bayesian Econometrics*. Cambridge University Press.

[13] Jorion, P. (2007). *Value-at-Risk*. The McGraw-Hill Companies.

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

[15] Liebwein, P. (2006). Risk models for capital adequacty: Applications in the context of solvency ii and beyond. *The Geneva Papers on Risk and Insurance Issues and Practice*, 31(3):528–550.

[16] Mittnik, S. (2011). Solvency ii calibrations: Where curiosity meets spuriosity. Report, Munich: Center for Quantitative Risk Analysis (CEQURA, Department of Statistics, University of Munich.

[17] Pfeifer, D. and Strassburger, D. (2008). Solvency ii: Stability problems with the scr aggregation formula. *Scandinavian Actuarial Journal*, 2008:61–77.

[18] Press, S. *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*.

[19] S, C. (2014). *Handbook of Market Risk*. John Wiley and Sons.

[20] Silva, R. and Lopes, H. (2008). Copula, marginal distributions and model section: a bayesian note. *Statistics and Computing*, 18(3):313–320.

[21] Sklar, A. (1959). Fonctions de repartition a n dimensions et leaurs marges. *Publ. Inst. Statis. Univ. Paris*, (8):229–231.