Vrije Universiteit Amsterdam

E_EOR3_FTR

Financial Engineering

Assignment part 1

Authors:
Daniël van Hanswijk (2726843)
Mark de Kwaasteniet (2649271)
Martijn van Welsenes (2713315)
Sem Wierdsma (2670919)

Tutor:
Paolo Gorgi

April 30, 2021

Contents

1 Volatility in the US stock market

You work as a consultant for an investment bank. Every week you meet with the Investment Management team of the bank to decide on the strategy for the coming week. The first point of discussion is the performance of a broad market index for the US stock market: the S&P500 index. You have the historical weekly S&P500 index available since 1990. The time series of the S&P500 index is contained in the file S&P500.txt. Use this dataset to answer the following questions.

1.1 Exploratory data analysis

The weekly S&P500 index starting from 1990 is shown in Figure 1. The corresponding log-returns are obtained by

$$y_t = 100 \times \log\left(\frac{p_t}{p_{t-1}}\right), \tag{1}$$

where the log-returns have been multiplied by 100 to obtain the price variations in percentages.

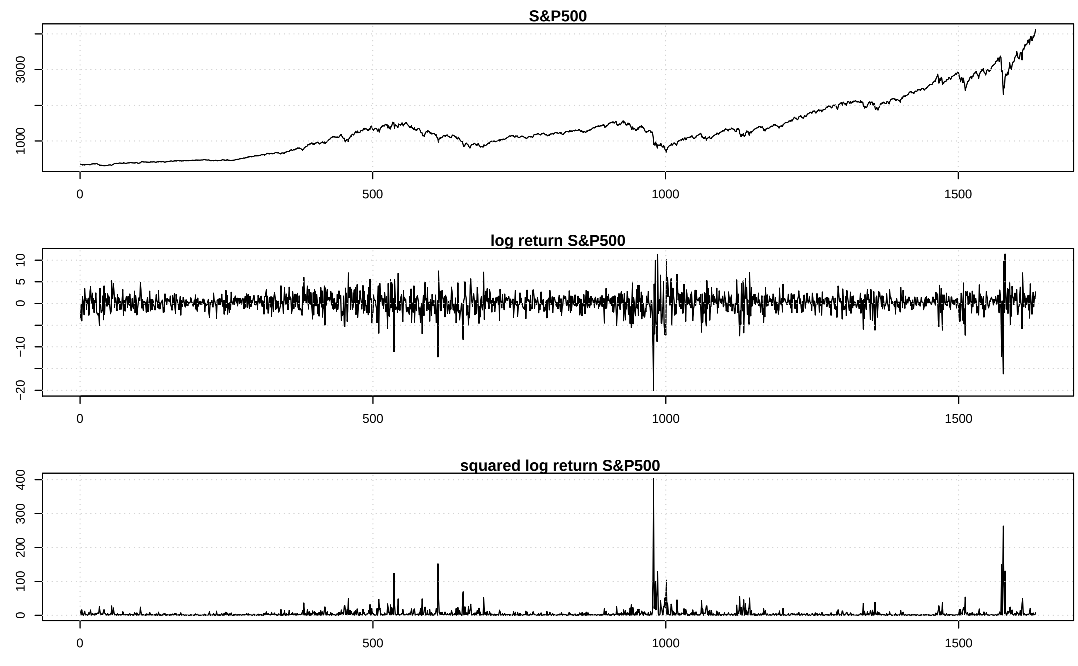

Figure 1: Weekly S&P500 index from 1990 (top panel), the corresponding log returns (middle panel) and squared log returns (bottom panel).

The middle panel of figure 1 suggest that the log-returns are stationary with a mean that is constant in time. Furthermore, the log-returns has clusters of volatility, which alternate with quiet periods where the amplitude is small. This clustering is confirmed when plotting the sample autocorrelation function of the squared log-returns (figure 2). The log-returns are barely correlated because only at the first lag the correlation is just outside the 95%-confidence interval. While the squared log-returns has significant values for the first seven lags.

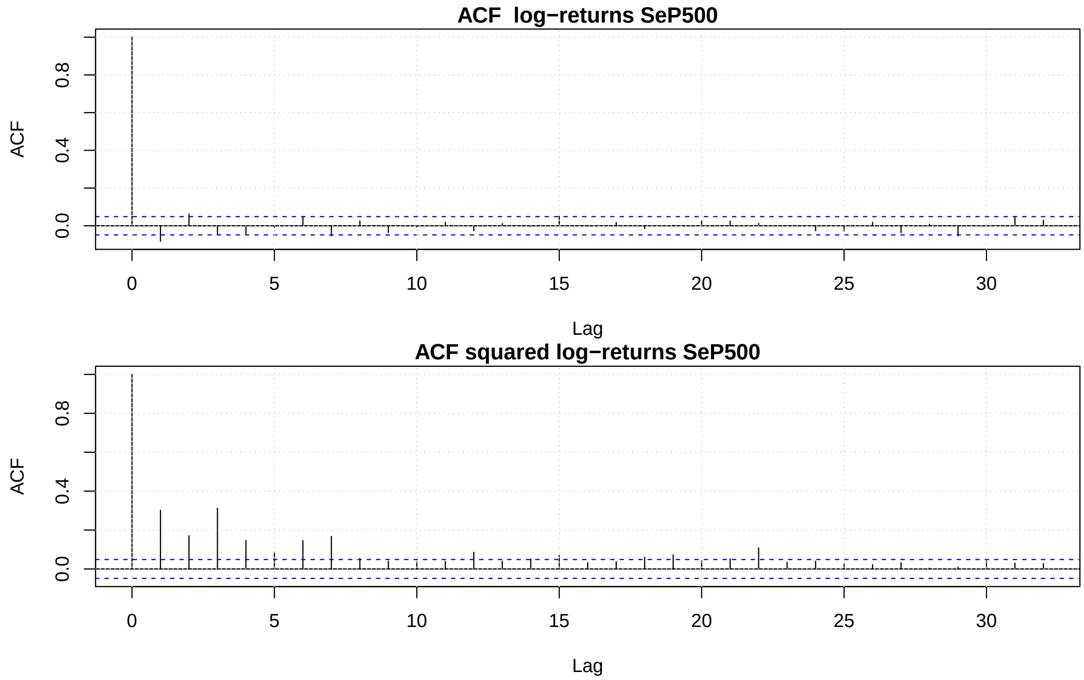

Figure 2: Sample autocorrelation function of weekly log returns of the S&P500 index (top panel) and their squares (bottom panel). The blue lines indicate the 95% confidence interval.

1.2 Next weeks investment

The suggestion of the consultant is to decrease the investment in US markets for next week because the price increase of this week indicates that we should expect negative returns. To determine if this strategy makes sense we look at the autocorrelation of the log-returns (Figure 2 top panel). In general there is a weak autocorrelation in log returns and they cannot be used in practice. This is also what we observe. Although, one could argue that correlation at one time lag is barely significant but negative. This would mean that the log-return at the next time step is more likely to have an opposite sign compared to the log-return at the current time step. Although by a very small margin in probability the price increase of this week is then followed by a price decrease next week. The consultant's strategy makes sense, but he has to be careful with the interpretation.

1.3 Estimation a GARCH(1,3) model

The data shows volatility clustering, so a possible model could be the GARCH model. Here we fit the GARCH(1,3) model to the S&P500 data. The GARCH equations are given by,

$$y_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \sum_{i=1}^{3} \beta_i \sigma_{t-i}^2, \quad \forall t \in \mathbb{Z}, \tag{2}$$

where $\omega > 0, \alpha_1 \geq 0, \beta_i \geq 0$ are the parameters to be estimated. The estimates are found by optimizing the likelihood function. For the GARCH(1,3) the simplified log-likelihood is

$$L(y_1, ..., y_T, \theta) = \sum_{t=4}^{T} -\frac{1}{2} \left(\log \sigma_t^2 + \frac{y_t^2}{\sigma_t^2} \right), \tag{3}$$

where the summations starts from $t = 4$ because the updating equation for σ_t^2 goes back three time steps. We set the first three variations equal to the sample variance so these only contribute a constant to the likelihood.

To assure that the parameters remain positive during the optimization we use the logarithm link function to give the optimization function the logarithm of the parameters and let the function calculate the exponent again. This way the optimizer can give any value to the function and the final results simply need to be exponentiated (with the link function) again to obtain the proper estimates.

For the S&P500 data the initial values $\theta_0 = (\omega_0, a, b_1, b_2, b_3) = (\omega_0, 0.01, 0.9, 0.01, 0.01)$ are used. To obtain the initial value of ω we can use the unconditional variance, because the unconditional variance of a GARCH(1,3) model is given by

$$\mathrm{Var}\left(y_t\right) = \omega / \left(1 - \alpha - \sum_{i=1}^{3} \beta_i\right). \tag{4}$$

Inverting this equation gives an initial value of $\omega_0 = \mathrm{Var}(y_t)(1 - a - \sum_i b_i)$. Furthermore, we need to initialize the updating equation with values for $\sigma_1, \sigma_2, \sigma_3$. these variances can be set to the unconditional variance of the log-returns, $\mathrm{Var}(y_t)$.

Doing the optimization gives the following parameter estimates

$$\left(\hat{\omega}_1, \hat{\alpha}_1, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3\right) \approx (0.2557, 0.2089, 0.7388, 0.0088, 0.0083). \tag{5}$$

The estimated conditional variance can now be calculated by using the found parameter estimates and plugging them into the updating equation (2). The results are shown in Figure 3.

1.4 Specification of the model

To check whether the GARCH(1,3) model gives a proper description of the data we analyse the residuals $u_t = y_t / \hat{\sigma}_t$. These are expected to be homoscedastic and Gaussian, since the residuals approximate ϵ_t.

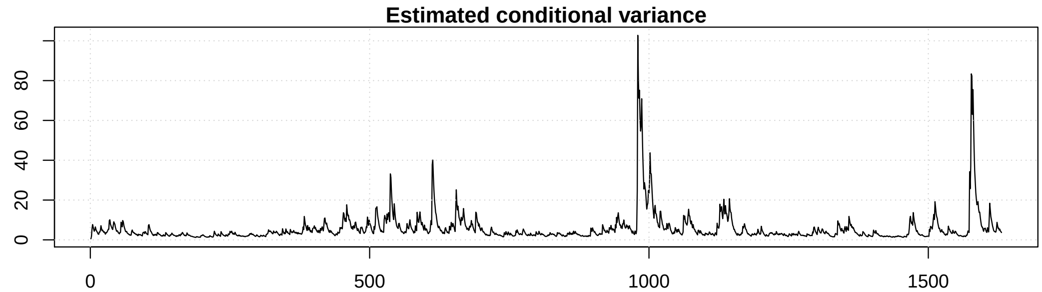

Figure 3: The filtered conditional variance of the GARCH(1,3) model with parameter estimates in equation (7)

First we check for homoscedasticity. The left panel in figure 4 shows the sample autocorrelation of the squared residuals. Most of the lags have a correlation below the 95% confidence interval, except one which is barely significant. This suggests that the residuals are homoscedastic or at least can be treated in such a way. Secondly, we check for normality. For this we use two methods. Looking at the qq-plot in Figure 4 we can see that the tails of the qq-plot bend off from the 45 deg-line. This indicates that our residuals have more extreme values than a normal distribution. Additionally, we check normality with the Jarque-Bera test. This gives a p-values which is practically zero, so we reject the null hypothesis H_0 of normality. This agrees with result from the qq-plot. Thus we conclude for our colleagues that the GARCH(1,3) model can be used but one has to be careful when doing the inference and the assumption that the residuals are normal cannot be used.

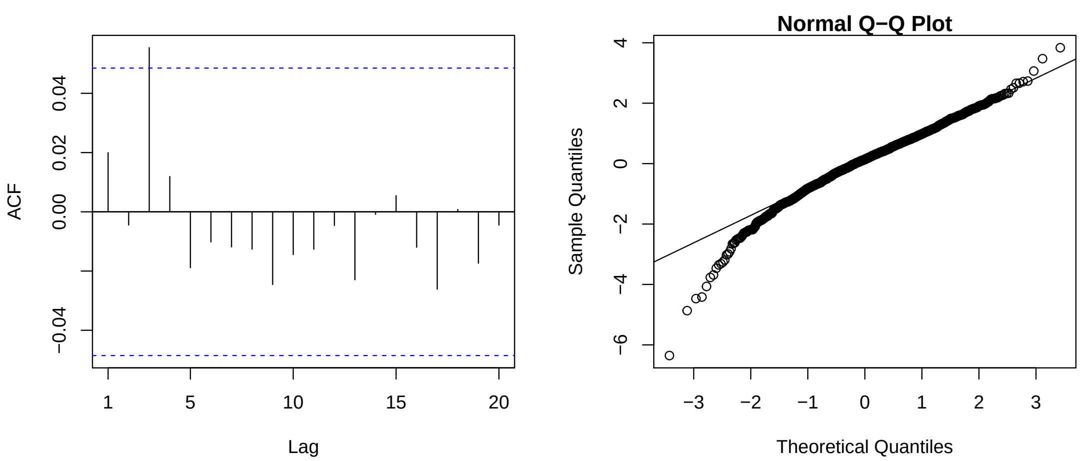

Figure 4: Left: The sample autocorrelations of the squared residuals u_t^2. Right: The qq-plot of the residuals.

1.5 Estimate the best GARCH(p,q) model

In this section we investigate what the best GARCH model is for the data. This will be done by estimating different models and comparing the AIC and BIC values of these models. In general GARCH(p,q) model is given by

$$y_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{i=1}^{q} \alpha_i y_{t-1}^2 \tag{6}$$

where $\omega > 0, \alpha_i \geq 0, \beta_i \geq 0$ are the parameters to be estimated. The estimation is similar to what was done with the GARCH(1,3) model and the likelihood function stays the same as equation (3). Except the start of the summation may vary, depending on how many steps back the updating equation needs to go, this is given by $\max(p, q)$.

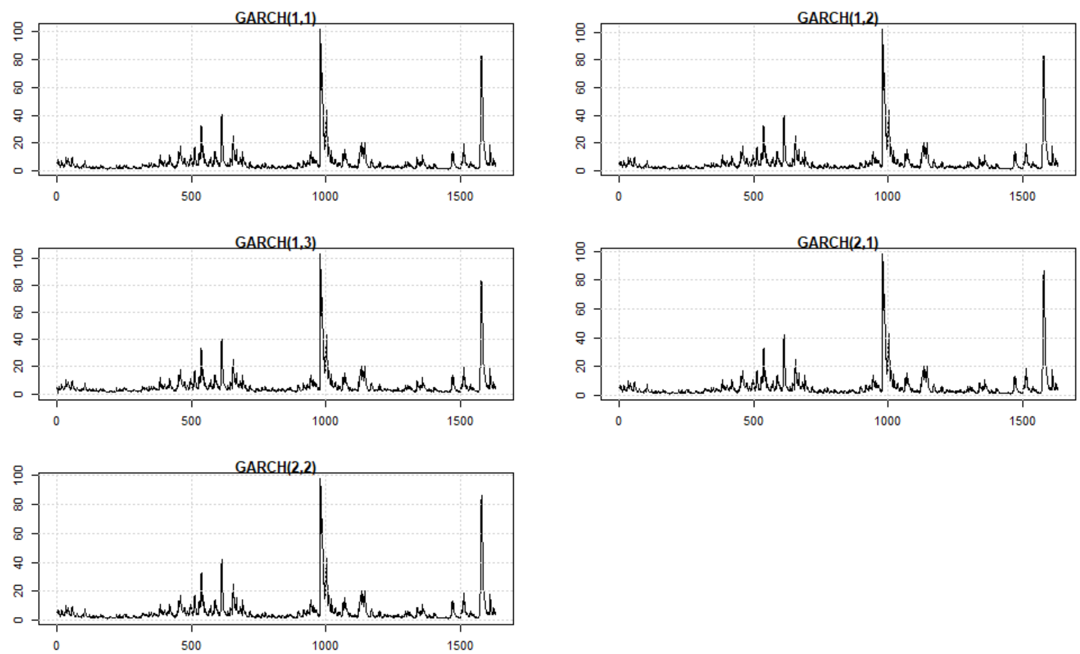

Figure 5: This five graphs show the filtered variances of five different GARCH models

There is a very small difference between the five GARCH models for the filtered variances. An example of a small difference between the five models is that the biggest peak for the GARCH(2,1) and the GARCH(2,2) is just below the value of 100 while the GARCH (1,1), GARCH(1,2) and GARCH(1,3) are just above the value of 100. The other differences between the filtered variances are so small that they are not visible.

The last step we take to compare the models with each other is to look at the AIC and BIC of the model. The model with the lowest AIC and BIC is the "best" model. We find that the GARCH(1,1) model is the best (see Table 1). Increasing p raises the information criterion and also increasing q raises the information criterion. So the GARCH(1,1) model is the model that best describes the dynamic properties of the log-returns of the S&P500 broad market index.

Model	AIC	BIC
GARCH(1,1)	6944.064	6960.255
GARCH(1,2)	6946.173	6967.761
GARCH(1,3)	6948.168	6975.153
GARCH(2,1)	6946.69	6968.278
GARCH(2,2)	6948.846	6975.831

Table 1: The AIC and BIC values of different GARCH models

1.6 Forecasting using a GARCH(1,1) model

An external consultant claims that: "Despite the recent positive performance, the US market is likely to drop more than 5% next week due to high volatility". We are about to check this statement by making a forecast while making use of a GARCH(1,1) model. The estimation of the GARCH(1,1) model has already be done in the previous section. The parameters that we found are

$$\left(\hat{\omega}, \hat{\alpha}, \hat{\beta}\right) \approx (0.25236, 0.20573, 0.75934). \tag{7}$$

Next, we have a given VaR of 5% which is defined as the value α that satisfies

$$P(y_{t+1} \leq -5|Y^t) = \alpha \tag{8}$$

Next, we use R's cumulative distribution function for the Normal distribution pnorm() to obtain the desired cumulative distribution function where $y_{t+1} = \sigma_{t+1}\epsilon_{t+1}$ and ϵ_{t+1} is assumed to be normal. The result is to 0.92%, which is very low so the consultants claim that despite the recent positive performance, the US market is likely to drop more than 5% next week due to high volatility is incorrect.

1.7 Value-at-Risk under the assumption of normal distribution

The claim of the colleague is that it may be inaccurate to obtain the Value-at-Risk under the assumption of normal distribution of the error of the GARCH model since log-returns are not normal. In general, using the normal distribution will lead to overestimation of the risk. To check if this statement is true we will take a look at the Kurtosis.

$$\text{Kurtosis} = \frac{\mathbb{E}\left(u_t^4\right)}{\mathbb{E}\left(u_t^2\right)^2} \tag{9}$$

The Kurtosis formula gives a value of 4.89. The value of 4.89 > 3 tells that the difference between the residuals is bigger than in a normal distribution. That is the reason for the fact that there will be bigger tails with more outliers. So using the normal distribution will lead to underestimation of the risk. In general have the GARCH models a Kurtosis > 3 which indicates that the log-returns got flatter tails and that is why the normal distribution underestimate the risk. So we do not agree with the statement of the colleague.

2 Portfolio management with multivariate GARCH models

An investment bank has a portfolio that invests 70% in the US stock market (S&P500) and 30% in the Hong Kong stock market (HSI). The Management of the bank is considering the possibility of changing the composition of the portfolio. As a consultant, you need to understand the dynamic relationship between the US market and the Hong Kong market. The time series of the Hong Kong market index HSI is contained in the file HSI.txt whereas the S&P500 index is contained in S&P500.txt. Use these datasets to answer the following questions.

2.1 Estimate a bivariate CCC model

To estimate a bivariate CCC model for the log-returns of the S&P 500 and the HSI, we can use the equation by equation approach. For this approach we first estimate a univariate GARCH model for each time series,

$$\{y_{i,t}\}_{t=1}^{T}, i = 1, \ldots, n \tag{10}$$

To obtain the model estimates we maximize the log likelihood function. This resulted in the following parameters.

$$\left(\hat{\omega}_1, \hat{\alpha}_1, \hat{\beta}_1\right) \approx (0.255, 0.208, 0.757) \tag{11}$$

and

$$\left(\hat{\omega}_2, \hat{\alpha}_2, \hat{\beta}_2\right) \approx (0.146, 0.0767, 0.911) \tag{12}$$

where the estimates with subscripts 1 and 2 are the GARCH model estimates of the S&P500 and HSI, respectively.

Next, we obtain the standardized errors from each of these series

$$\hat{\epsilon}_{it} = \frac{y_{it}}{\hat{\sigma}_{it}}, i = 1, \ldots, n \tag{13}$$

These standardized errors are used to obtain the correlation matrix of the residuals. This correlation is then used to obtain the conditional covariance. The conditional variances, covariance and correlation of the estimated models are shown in Figure 6. In this figure we see clearly that the S&P500 has larger variance spikes compared to the HSI, whereas the volatilities of the HSI depends more on the previous volatility. The latter clearly agrees with the estimated parameters, where β is larger for the HSI than for the S&P500.

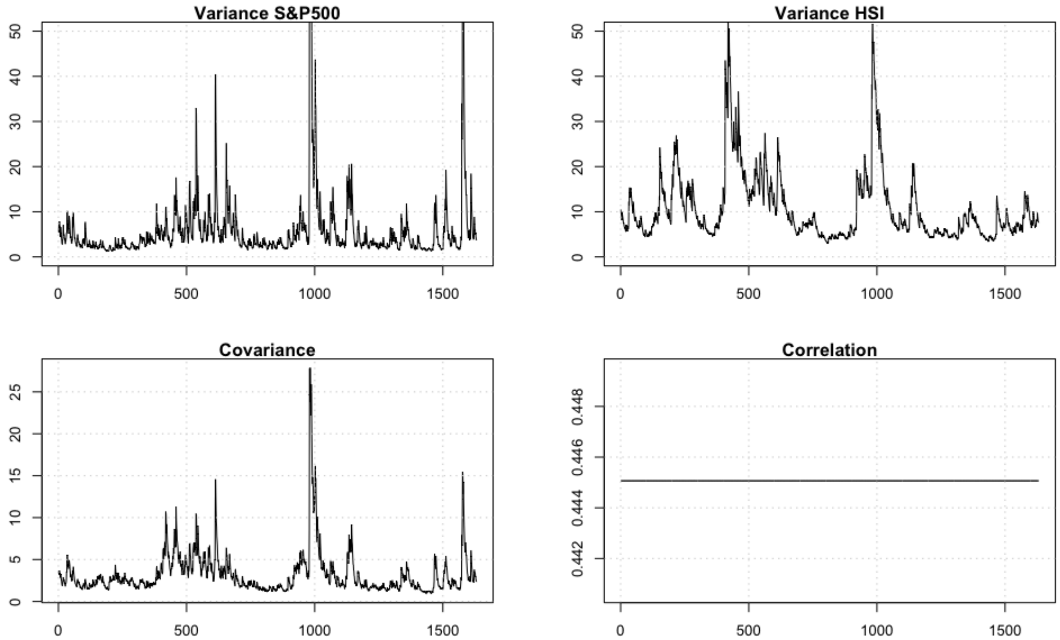

Figure 6: Plot of the estimated conditional variances, covariance and correlation.

2.2 Obtain and plot the conditional variance and the α-VaR

The bank has a portfolio of 30% HSI and 70% S&P500. Now we can obtain the conditional variance of the two assets using the following equation

$$\sigma_{p,t}^2 = k_1^2\sigma_{1,t}^2 + k_2^2\sigma_{2,t}^2 + 2k_1k_2\sigma_{12,t} \tag{14}$$

Where σ_p^2 is the conditional variance of the portfolio, k_i is the portfolio weight of asset i, σ_i is the conditional variance of asset i and σ_{12} is the conditional correlation between the two assets.

With the conditional variance of the portfolio, we can calculate the Value at Risk (VaR) for the portfolio. The α-VaR is obtained as follows:

$$\alpha\text{-VaR}_t = z_\alpha\sigma_{p,t} \tag{15}$$

where z_α is the quantile of level α of a standard normal. The conditional variance, as well as the α-VaR are shown in Figure 7. From this figure we can see that when the conditional variance is high, the 1% Value at Risk is also significantly higher. This is of course what we would expect, since this measurement is a constant times the conditional standard deviation, which is the square root of the conditional variance.

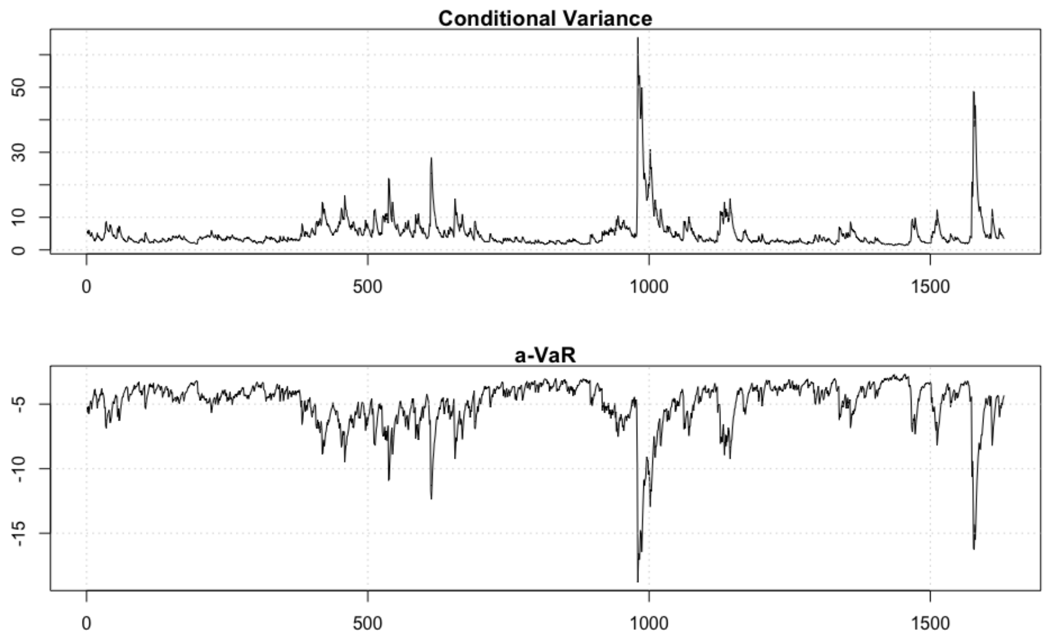

Figure 7: Conditional variance and the α-VaR at 1% level for the portfolio

2.3 Obtain the optimal portfolio weights in terms of Sharpe Ratio

For the bank it may be beneficial to have other, dynamic, weights for it's assets. One way to optimize the portfolio weights of the assets is by maximizing the Sharpe Ratio, given by the following equation

$$S_{p,t} = \frac{\mu_{p,t}}{\sigma_{p,t}} \tag{16}$$

This optimization function tries to maximize the expected return, μ_p while minimizing the volatility of the portfolio σ_p, which implies risk of the portfolio. The optimization problem of Sharpe Ratio is given by,

$$\max_{\mathbf{k}_t} \frac{\mathbf{k}_t^T \boldsymbol{\mu}_t}{\sqrt{\mathbf{k}_t \boldsymbol{\Sigma} \mathbf{k}}}, \text{ s.t. } \sum_{i=1}^{n} k_{i,t} = 1, k_{i,t} \geq 0 \tag{17}$$

Maximizing the Sharpe Ratio gives the following optimal portfolio weights, shown in Figure 8. In these figure we see that for the larger part, the portfolio should be more invested in the S&P500 than in the HSI. Especially, during the time at around 200-300. This makes sense, as we can see in Figure 6, at this time period, the conditional variance is relatively small for the S&P500, whereas it is large for the HSI. However, at around times 1000 and 1600, the conditional variance for the S&P500 has large spikes. This results in larger portfolio weights of the HSI at these times.

We can also use this optimization problem to find the optimal portfolio weights at the next time step $T + 1$. To do this we take the unconditional variances and the log-returns

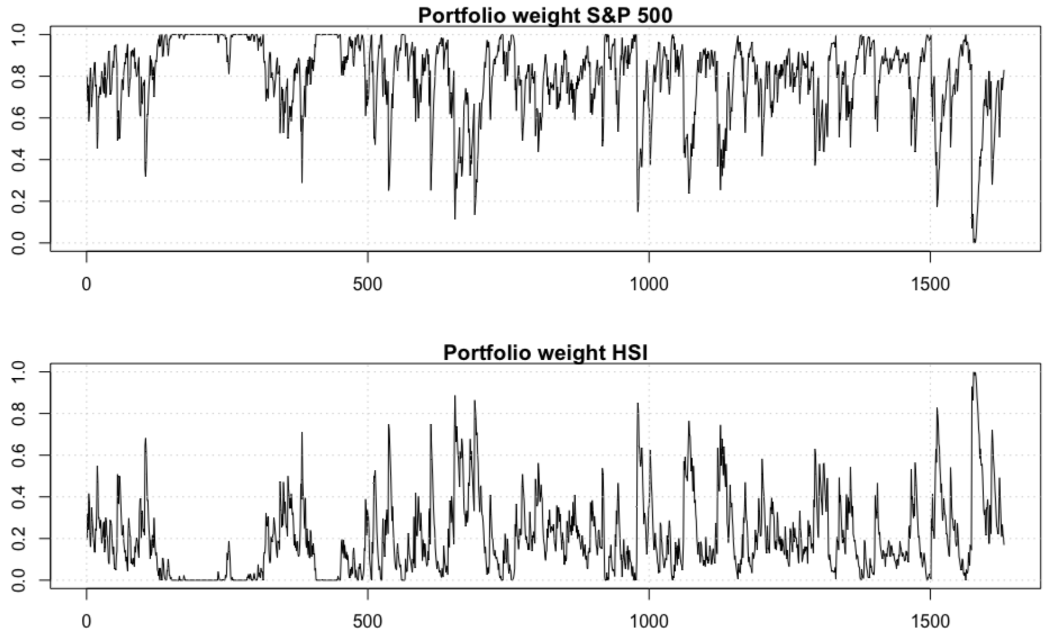

Figure 8: Optimal portfolio weights

at time T and calculate the conditional variances and conditional correlation at time $T+1$. The latter are derived as follows:

$$\sigma_{i,T+1}^2 = \omega_i + \alpha_i y_{i,T}^2 + \beta_i \sigma_{i,T}^2 \tag{18}$$

and

$$\sigma_{12,T+1} = \rho_{12}\sigma_{1,T+1}\sigma_{2,T+1} \tag{19}$$

Now that we have the expected return and the conditional covariance matrix, we can obtain the optimal portfolio weights at time $T+1$, which are

$$(k_1, k_2) \approx (0.733, 0.267) \tag{20}$$

where k_1 and k_2 are the weights of the S&P500 and HSI, respectively

2.4 Estimate a bivariate sDVECH model using covariance targeting

To estimate the scalar DVECH model for the log-returns of the S&P 500 and the HSI, we can use covariance targeting in order to estimate the parameters parameters. We start by estimating the unconditional covariance matrix, Σ, using the covariance function on the log returns of the S&P 500 and the HSI. Then we optimize the log likelihood function in order obtain the parameter estimates.

The maximization of the log likelihood function resulted in the following parameters.

$$\left(\hat{\alpha}_1, \hat{\beta}_1\right) \approx (0.076, 0.903) \tag{21}$$

$$L(y_1, \ldots, y_T, \theta) \approx 4387.5 \tag{22}$$

We will use the estimated parameters to obtain the conditional covariance matrix.

$$\boldsymbol{\Sigma}_t = \hat{\boldsymbol{\Sigma}}(1 - \alpha_1 - \beta_1) + \alpha_1 \boldsymbol{y}_{t-1} \boldsymbol{y}_{t-1}^T + \beta_1 \boldsymbol{\Sigma}_{t-1}, \tag{23}$$

When obtaining the conditional covariance matrix we will set the initial value for the conditional variance equal to the sample covariance. The conditional variances, covariance and correlation of the estimated scalar DVECH model are shown in Figure 9.

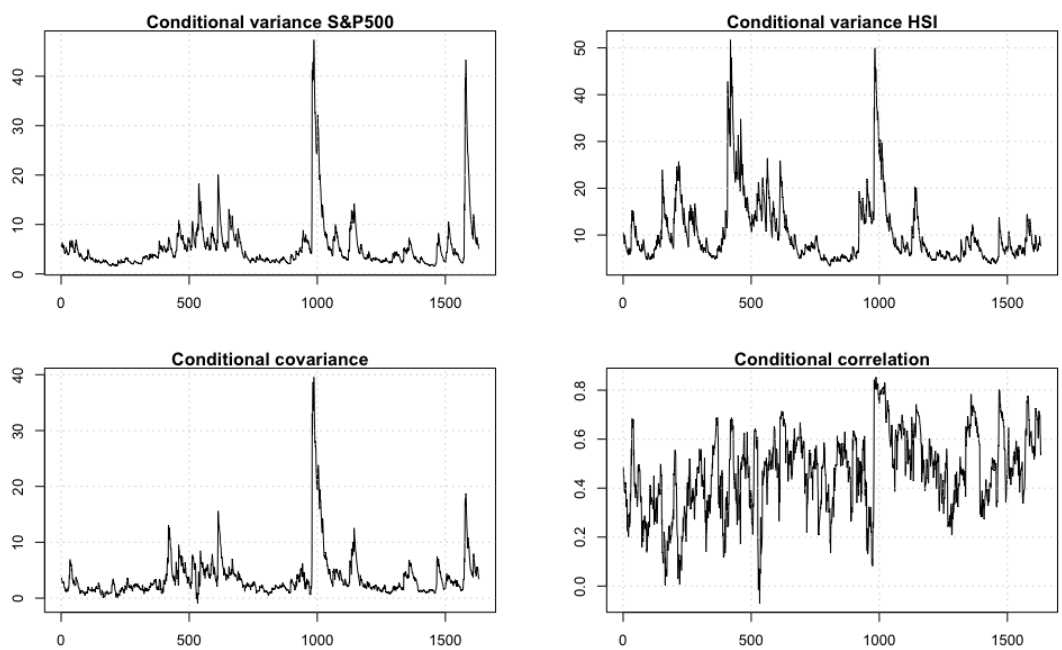

Figure 9: Plot of the estimated conditional variances, covariance and correlation of scalar DVECH model

We observe that the conditional variance of the S&P 500 has again relatively higher spikes than the conditional variance of the HSI. The parameters suggest that the conditional variances and covariance depend more on the past conditional variance in contrast to past log returns of the S&P 500 and the HSI.

The bank has a portfolio of 30% HSI and 70% S&P500. Again we obtain the conditional variance of the two assets using Equation (14) and the α-VaR using Equation (15).

Figure 10 shows the conditional variance and the α-VaR at 1% for both the estimated scalar DVECH model and the estimated CCC model in order to compare the estimation difference between the two models.

We see that both the sDVECH model and the CCC model give similar plots regarding the conditional variance and the α-VaR of the bank's portfolio. However, there are some differences. First of all, the estimates of the CCC model return higher values in absolute terms in both the plot for the conditional variance and the α-VaR. Secondly, we note that the estimates of sDVECH have less steep peaks compared to the VaR model. At last, we assume the sDVECH to be the superior model, since the CCC model assumes the correlation matrix to be constant, which is very restrictive and often not representative of practical situations.

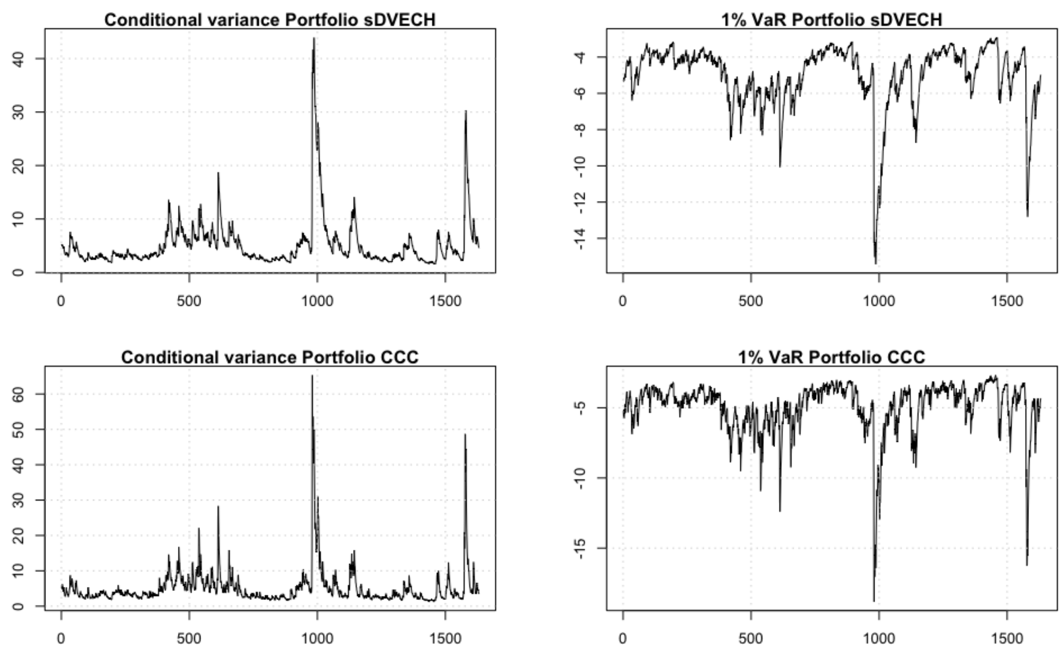

Figure 10: Plot of the estimated conditional variances, covariance and correlation of the portfolio using both a sDVECH and a CCC model.

2.5 Forecasting using the bivariate sDVECH model

In order to forecast the volatility of the bank's portfolio using the bivariate scalar DVECH model, we consider the conditional variance of the portfolio as a forecast that would be conditional on the present Y^T. We will denote the forecast of the variance of the log returns as $\sigma_T^2(h)$, where h denotes the number of forecast steps.

In our case h will be 52 since we are asked to forecast the volatility for the next 52 weeks. In order to make the forecast, we will use σ_{T+1}^2 as starting values using Equation (18). After that we will create a for loop that will compute the conditional variance vector until $h = 52$ using equation (28) from the following derivation

$$\begin{aligned}
\sigma_T^2(h) = \mathbb{E}(\sigma_{T+h}^2|Y^T) &= \mathbb{E}(\omega + \alpha y_{T+h-1}^2 + \beta\sigma_{T+h-1}^2|Y^T) && (24)\\
&= \omega + \alpha\mathbb{E}(y_{T+h-1}^2|Y^T) + \beta\mathbb{E}(\sigma_{T+h-1}^2|Y^T) && (25)\\
&= \omega + \alpha\mathbb{E}(\sigma_{T+h-1}^2|Y^T) + \beta\mathbb{E}(\sigma_{T+h-1}^2|Y^T) && (26)\\
&= \omega + (\alpha+\beta)\mathbb{E}(\sigma_{T+h-1}^2|Y^T) && (27)\\
&= \omega + (\alpha+\beta)\sigma_T^2(h-1) && (28)
\end{aligned}$$

After we obtain the conditional volatility of the S&P 500 and the HSI over the next 52 weeks, we can calculate the conditional volatility of the portfolio over the next 52 weeks using again the aforementioned portfolio weights and Equation (14). Using the conditional volatility of the portfolio, we can plot the volatility over the next 52 weeks. Figure 11 shows the corresponding plot. The plot shows an upward sloping volatility, expressed in variance, where the slope slightly decreases over time. The volatility forecast shows a volatility of 5 around week 52, which will transform into the unconditional variance when h grows infinitely large.

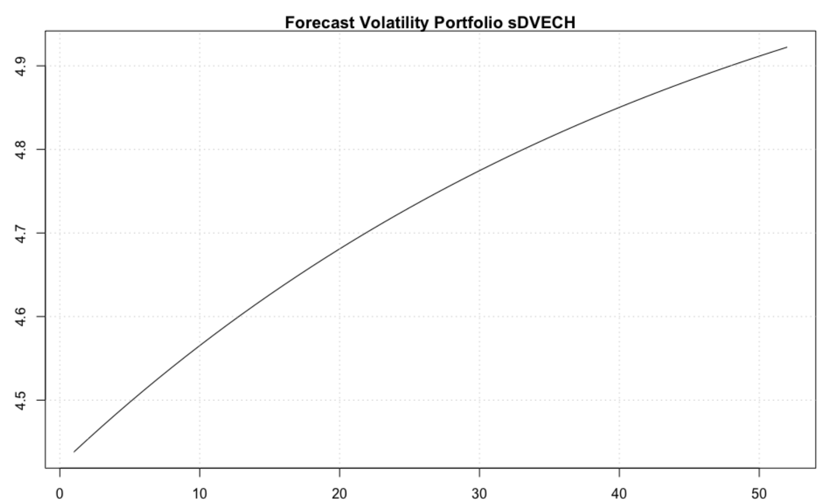

Figure 11: Volatility (variance) Forecast of the portfolio using a bivariate sDVECH model

2.6 Estimating a 3 dimensional CCC model

To estimate a 3 dimensional CCC model for the log-returns of the S&P 500, the HSI and DAX, we first estimate a univariate GARCH model for the DAX time series and use the already obtained univariate GARCH models of the the S&P 500 and the HSI.

$$\{y_{i,t}\}_{t=1}^{T}, i = 1, \ldots, n \tag{29}$$

To obtain the model estimates we maximize the log likelihood function. This resulted in the following parameters for the DAX time series.

$$\left(\hat{\omega}_3, \hat{\alpha}_3, \hat{\beta}_3\right) \approx (0.538, 0.180, 0.771) \tag{30}$$

Next, we obtain the standardized errors from the DAX time series using equation (13). These standardized errors are used to obtain the correlation matrix of the residuals. This correlation is then used to obtain the conditional covariance matrix. We plot the conditional covariances in Figure 12. When we look at the graphs we see that the three conditional covariances are very similar, however the covariance between the S&P 500 and the HSI has higher peaks compared to the other two covariances.

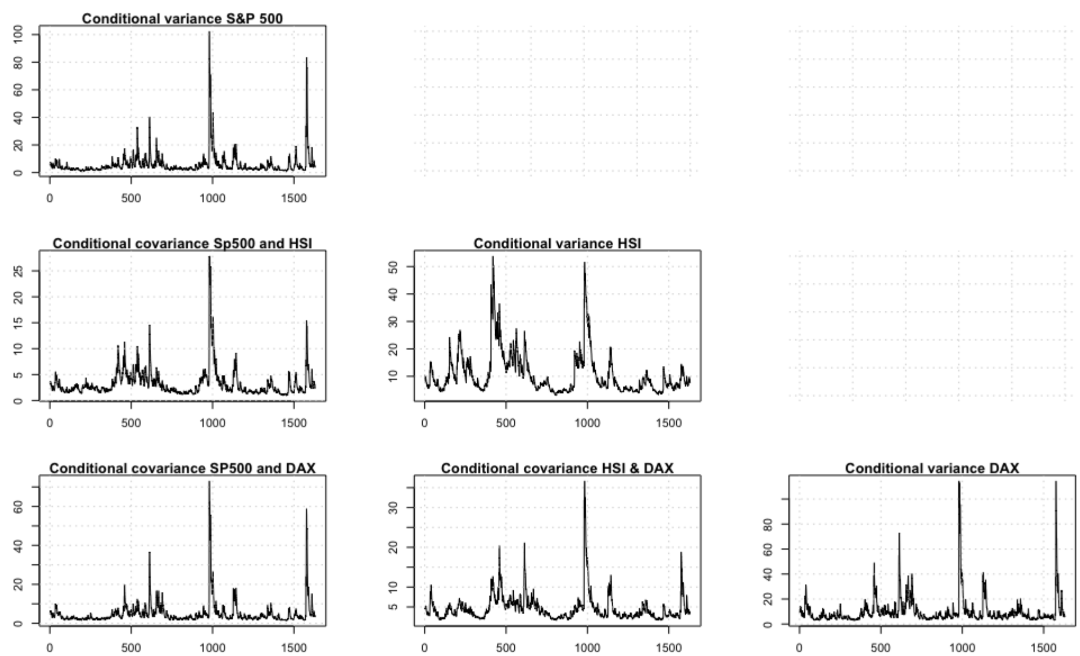

Figure 12: Conditional covariance matrix

We will use the conditional covariance matrix at time period T to calculate the optimal portfolio weights at time period $T + 1$ that maximizes the Sharpe ratio of the portfolio consisting of the three market investments.
First of all, we will calculate the conditional variances at time period $T + 1$ using equation (24). We use the conditional variances at $T + 1$ in combination with the initial correlation coefficients, computed with standardized errors, to construct the conditional covariance matrix at time period $T + 1$

$$\sigma_{i,T+1}^2 = \omega_i + \alpha_i y_{i,T}^2 + \beta_i \sigma_{i,T}^2 \tag{31}$$

With the conditional covariance matrix it is possible to construct the conditional variance of the portfolio containing the three market investments. Then we will optimize the Sharpe ratio by constructing a portfolio that maximize the expected return, μ_p, while minimizing the volatility of the portfolio, σ_p, using Equation (17). Equation (32) shows the different weights in the portfolio that maximize the Sharpe ratio.

$$(k_1, k_2, k_3) \approx (0.673, 0.247, 0.080) \tag{32}$$

Where k_1, k_2 and k_3 are the weights of the S&P 500, the HSI and the DAX in the portfolio, respectively.

We note that the weights for the maximum Sharpe ratio portfolio changed a little after including the DAX time series data into our CCC model. This could be a direct result of the large value of $\omega_3 = 0.538$, which causes high conditional variances in the DAX time series. In order to obtain the highest Sharpe ratio, the conditional variance of the portfolio should be minimal and thus the weight of the DAX in the portfolio will be small in order to obtain the highest Sharpe ratio at $T + 1$.