1 Introduction and Summary of Results

1.1 Objective and Data description

Our dataset contains historical daily log returns data for 20 US stocks ranging from 02-01-1990 to 14-12-2020. Given the freedom to choose whichever two stocks we like, we have formed a portfolio consisting of Microsoft and JP Morgan stock. In turn, we were tasked to estimate one bivariate and one univariate conditional volatility model to produce two different series of portfolio volatility estimates using individual stock weights of our choice. Lastly, we have used the aforementioned volatility estimates to compute and evaluate two series of 5% VaR estimates for our portfolio.

1.2 Summary of Results

1.2.1 Univariate volatility model

In this section, we evaluate the 5% VaR estimates obtained using our selected Univariate conditional volatility model.

We define R as the matrix of stock returns and $\boldsymbol{w}$ as the vector of portfolio weights.

According to the analysis that we have conducted, the univariate model which best fits the data is a GJR-GARCH(1,1,1) model with t-distributed residuals, specified as follows:

$$\sigma_{p,t}^2 = \omega + \left[\alpha + \gamma * I_{\{r_{p,t-1}<0\}} \right] r_{p,t-1}^2 + \beta * \sigma_{p,t-1}^2$$

where $\sigma_{p,t}^2$ is the portfolio variance at time t, $\boldsymbol{r}_p = R\boldsymbol{w}$ is the vector of portfolio returns, and hence $r_{p,t}$ is the return of the portfolio at time t, with $r_{p,t} \mid \mathcal{F}_t \sim t(\nu)$.

Our estimated model coefficients are: $\widehat{\omega} = 2.96 \times 10^{-6}$, $\widehat{\alpha} = 0.0596$, $\hat{\gamma} = 0.0625$, $\hat{\beta} = 0.9089$ and $\hat{\nu} = 5.33$.

Using the above model, we obtain our 5% VaR estimate for day t using the following formula:

$$VaR_t^{5\%} = -\sigma_{p,t} \times \Phi^{-1}(0.05) \times \sqrt{\frac{\hat{\nu}-2}{\hat{\nu}}}$$

where $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution function (of a standardised version) of a t-distribution with degrees of freedom $\hat{\nu}$, as estimated above. Note also that we assume a portfolio value of 1 throughout our analysis to simplify our VaR calculations.

We first compute the hit sequence of our backtest, which indicates periods when violations have occurred. We define a violation at time t as follows:

$$V_t = \begin{cases} 1, & if\ LogReturn_t < -VaR_t \\ 0, & otherwise \end{cases}$$

Our GJR-GARCH(1,1,1) model yields a Violation Ratio of 0.8487 which suggests that although this value lies within the acceptable range of $[0.8\,, 1.2]$, our model is over forecasting risk.

We then proceed to further examine our VaR series estimates using statistical tests, namely, the Conditional and Unconditional Coverage tests. In both tests, our null hypothesis is:

$$H_0 : Violations\ are\ IID\ over\ time, and\ occur\ with\ probability\ 5\%$$

The alternative hypotheses for the Unconditional and Conditional coverage tests respectively are:

$$H_1 : Violations\ do\ not\ occur\ with\ probability\ 5\%$$

$$H_1 : Violations\ are\ not\ IID\ over\ time\ (i.e.\ violations\ cluster)$$

For the Unconditional coverage test, the MLE $\hat{p}$ value for the true probability of a violation is 0.0424, and conducting a likelihood ratio test using the χ_1^2 distribution, we get a P-value of 0.0017. Thus, we reject H_0 at the 5% significance level and conclude that violations do not occur at a 5% frequency, and could be in fact occurring at a lower rate.

With regards to the Conditional coverage test, a likelihood ratio test using the χ_1^2 distribution, yields the following results.

		Test statistic	P-Value	p00	p10	p01	p11
1	DCC-GARCH(1,1)	5.2962	0.0214	0.9588	0.9305	0.0412	0.0695
2	t-GJR-GARCH(1,1,1)	1.1212	0.2897	0.9581	0.9456	0.0419	0.0544

Table 1: Conditional Coverage test results for the chosen univariate and bivariate models

By observing the second row of Table 1, we cannot reject H_0 at the 5% significance level and we conclude that violations are indeed independent over time.

1.2.2 Bivariate volatility model

We proceed with evaluating the 5% VaR estimates obtained using our selected bivariate conditional volatility model. According to the analysis that we have conducted, the bivariate model which best fits the data is a DCC-GARCH(1,1) model specified as follows:

$$H_t = (1 - \alpha - \beta)\overline{H} + \alpha \boldsymbol{Q}_{t-1}\boldsymbol{Q}'_{t-1} + \beta H_{t-1}$$

where H_t is the dynamic covariance matrix at time t with $h_{ij,t}$ being its $(i,j)^{th}$ element, $\overline{H}$ being the unconditional covariance matrix, and $\boldsymbol{Q}_t$ the vector of standardised returns at time t.

We also define the variance-covariance matrix at time t as:

$$\Sigma_t = D_t\ C_t\ D_t$$

such that $D_t = diag\{\ \sigma_{1,t}\ \ ,\sigma_{2,t}\ \}$ where $\sigma_{i,t}^2 = \omega + \alpha * \ r_{i,t-1}^2 + \beta * \ \sigma_{i,t-1}^2$ is the stock i conditional variance at time t $\ for\ \ i = 1\ ,2$, and C_t is the correlation matrix at time t with $C_t^{i,j} = \frac{h_{ij,t}}{\sqrt{h_{ii,t} \times h_{jj,t}}}$.

Note also that $\boldsymbol{r}_1\ and\ \boldsymbol{r}_2$ represent the log-returns vectors for Microsoft and JP Morgan respectively.

Finally, the portfolio conditional variance at time t is calculated as $\sigma_{p,t}^2 = \boldsymbol{w}'\Sigma_{\boldsymbol{t}}\boldsymbol{w}$.

We begin by mentioning that the Violation Ratio in the bivariate case is exactly the same as before, hence also over estimating risk. Using the same approach as in the univariate case, we get that the MLE $\hat{p}$ value for the true probability of a violation is 0.0424, and conducting a likelihood ratio test using the χ_1^2 distribution, yields a P-value of 0.0017. Thus, we reject H_0 at the 5% significance level and conclude that violations do not happen at a 5% frequency.

However, by observing the first row of our results in Table 1 in the previous page, we can see that our conditional coverage p-value is 0.0214, leading us to reject H_0 at the 5%

significance level and conclude that violations tend to cluster and are not independent over time.

1.3 Interpretation of Results

We begin by stating that backtesting shows that the univariate GJR-GARCH(1,1,1) model with t-distributed residuals outperforms the bivariate DCC-GARCH(1,1) model on the basis of its VaR forecasts. Indeed, the conditional coverage test reveals that in our univariate model, violations are independent over time and do not form clusters, a desirable property for a model to be used for risk management. By analysing the values of $\hat{p}_{01}$ and $\hat{p}_{11}$ in Table 1 of page 2, we observe that $\hat{p}_{11}$ departs from $\hat{p}_{01}$ to a significantly higher extent in the bivariate case. This means that in the bivariate model, the probability of a violation at time t is much higher given a violation at time t-1, causing violations to cluster.

It is likely that the superiority of the univariate model comes primarily from the fact that by comparing their portfolio volatility estimates, the univariate model is in a better position to predict extreme negative shocks in returns because of the leverage parameter γ. In fact, a likelihood test in our analysis shows that the leverage parameter is highly statistically significant at the 5% level. Hence, during high stress periods, the univariate model is producing more accurate volatility estimates, even though the structure of DCC models allows them to effectively include dynamic correlation in their estimates.

Another practical aspect of why our bivariate model produces less accurate VaR estimates could be the fact that 9 parameters have to be estimated when fitting a DCC-GARCH(1,1) model as opposed to 5 in the GJR-GARCH(1,1,1) model with t-dsitributed residuals. As a result, higher estimation error is more likely in the bivariate model and can, therefore, explain the minor differences between the two series of portfolio volatility estimates.

Finally, both models yield identical results in the unconditional coverage test, suggesting that both models are too conservative and overestimate risk in their VaR forecasts. As a result, if any of the two models were to be used in practice for risk management, it could lead to the investor being overly conservative which will render him unable to maximise the potential returns of portfolio consisting of Microsoft and JP Morgan stock.

2 Methodology

2.1 Distributional nature of portfolio returns

Using our chosen weights, we compute a series of portfolio returns, which we analyse to gain a better understanding of our data, before fitting any models.

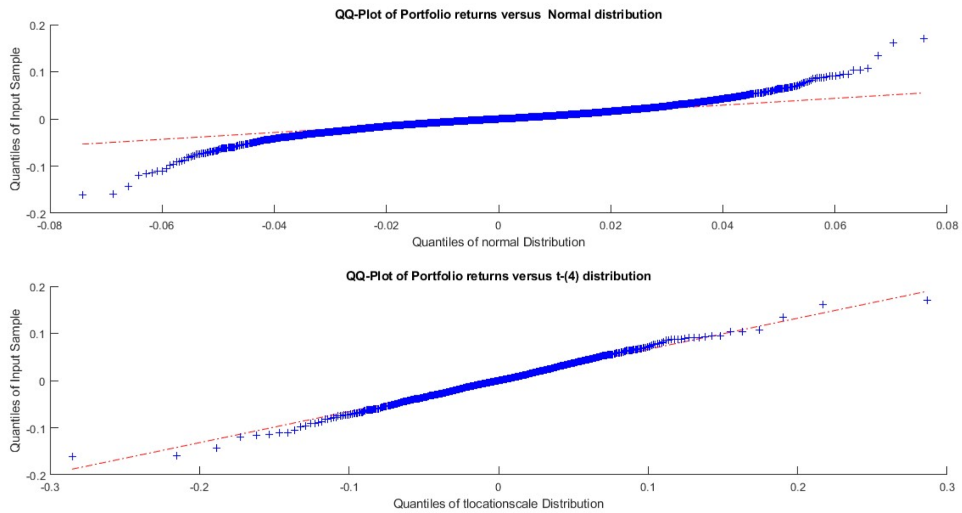

Figure 1: QQ-Plot of portfolio returns versus the Normal and t-distributions

Observing the excess kurtosis of the portfolio returns in the first plot and from the fact that departure from the red dashed line is much less significant in the second plot, we can conclude that portfolio returns are much more likely to come from a t-distribution. Indeed, a Jarque-Bera test yields a P-value of 0.001 and hence, we reject H_0 and conclude that portfolio returns are not Normally distributed. Similarly, the values for the kurtosis of the individual returns series of Microsoft and JP Morgan are much larger than three, which suggests a conclusion similar to the one mentioned above.

2.2 Determining the optimal portfolio weights

To start with, we calculate an annualised version of the vector of mean returns and an annualized covariance matrix for our two stocks. Since the number of assets we have in our portfolio is only two, determining the optimal weights using Monte Carlo simulations is a computationally feasible and reliable way to do so. In essence, we simulate and assign random weights to each of the two stocks and compute the annualised portfolio return and volatility with those particular weights. In turn, assuming a risk-free rate of 0, we calculate the annualised Sharpe ratio of the portfolio constructed using the simulated weights, and then choose the vector of weights that maximizes it. Below is a graph of the portfolio frontier generated in the simulation process, where the red marker indicates our optimal portfolio, according to our Sharpe ratio criterion.

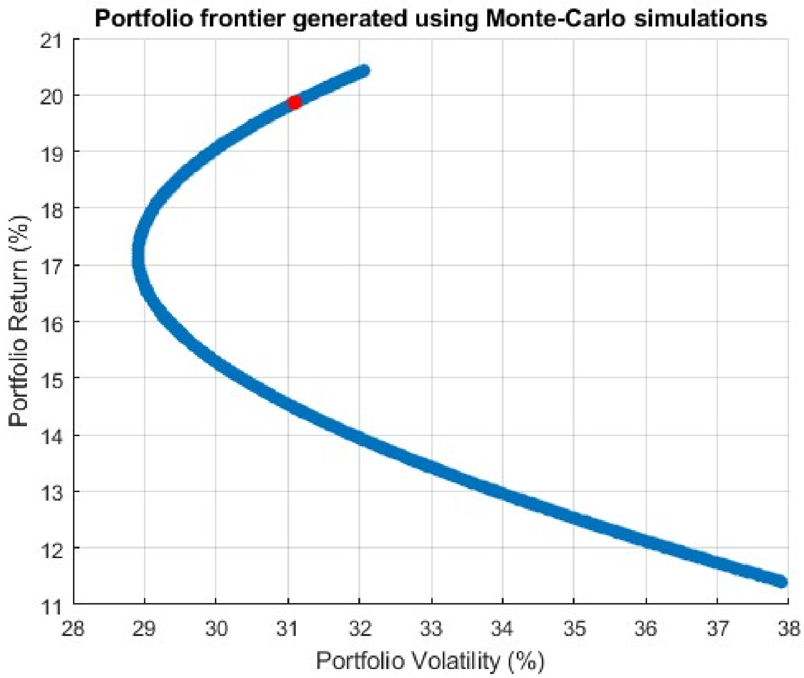

Figure 2: Portfolio frontier generated using Monte-Carlo simulations

Finally, we conclude this section by reporting that the optimal portfolio places a weight of 93.73% in Microsoft and 6.27% in JP Morgan stock.

2.3 Selecting a Bivariate volatility model

In this section, we compare and evaluate the three bivariate volatility models which were used in the analysis and provide justification to why a DCC-GARCH(1,1) model was chosen for computing the VaR estimates.

2.3.1 BEKK(1,1,2)

The first model implemented was a one-lag BEKK(1,1,2). The reason why a BEKK model was chosen is that the covariance matrix estimated is guaranteed to be positive semi-definite, which makes it a more appropriate choice compared to other multivariate GARCH models. In addition, since we only consider two stocks in our portfolio, when fitting this model, dimensionality is not a challenging issue as it would have been in a portfolio consisting of 10 different stocks, which would make accurate parameter estimation a difficult undertaking.

	Estimated parameter value	Standard errors	t-scores	P-Values
1	0.0035	0.0003	10.4852	0
2	0.0030	0.0008	3.8768	0.0001
3	-0.0007	0.0023	-0.2874	0.6131
4	0.1609	0.0217	7.4219	0.0000
5	-0.0088	0.0238	-0.3698	0.6443
6	0.0671	0.0467	1.4364	0.0754
7	0.1867	0.0275	6.7776	0.0000
8	0.9656	0.0074	131.1243	0
9	0.0016	0.0083	0.1930	0.4235
10	-0.0352	0.0188	-1.8667	0.9690
11	0.9693	0.0068	143.2063	0

Table 2: BEKK(1,1,2) parameter estimates summary

We note that by observing this table, t-tests for the individual BEKK parameters reveal that a number of them are statistically insignificant which suggests that the model is potentially overparametrised, leading to poor estimation performance on out-of-sample data.

2.3.2 O-GARCH

The second model implemented is an Orthogonal-GARCH model. Performing PCA on our returns data indicates that 71.78% of the variation in returns is explained by a factor to which both stocks are commonly exposed, this is the market factor. The remaining 28.22% of the variation in returns is attributed to the idiosyncratic risk associated with each of the two firms. Hence, we proceed and fit an O-GARCH model where we can only include a single factor.

2.3.3 DCC-GARCH(1,1)

The final model implemented during the bivariate model analysis is a DCC-GARCH(1,1) model. The main reason for choosing a DCC-type model lies within their ability to effectively incorporate time-varying correlations into their volatility estimates and capture how different sources of shocks in returns propagate within different industry sectors, allowing them to produce very accurate and reliable volatility estimates.

2.3.4 Model Comparison

In contrast to the univariate models to be discussed, since each of the bivariate models belongs to a different model class, we cannot compare them with statistical tests such as the likelihood ratio test. Thus, we compare and evaluate them on the basis of conditional volatility and correlation estimates.

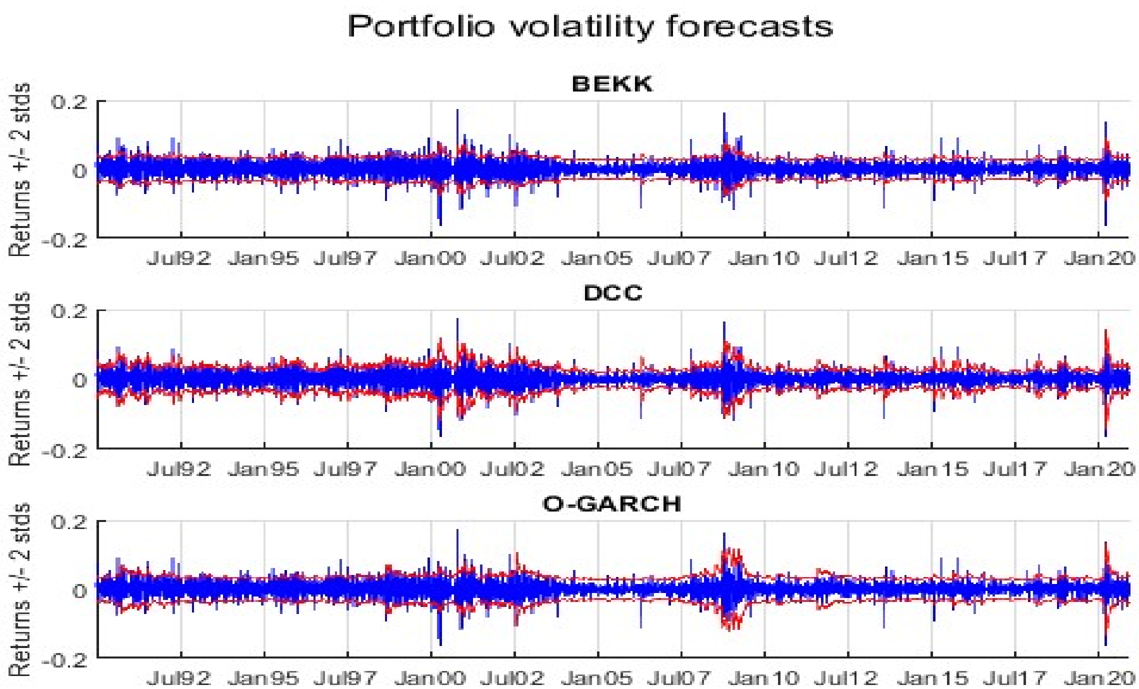

Figure 3: Comparing bivariate model portfolio volatility estimates

In figure 3, we can clearly see that amongst the three models, BEKK has the poorest portfolio volatility estimates and significantly underestimates volatility in the majority of extreme return shock events, made particularly apparent in the 2008 financial crisis period. In addition, by comparing the DCC and O-GARCH models, we obtain very similar results during times of calm markets, with the DCC model showing its superiority with more accurate estimates in extreme return events, and specifically in negative ones.

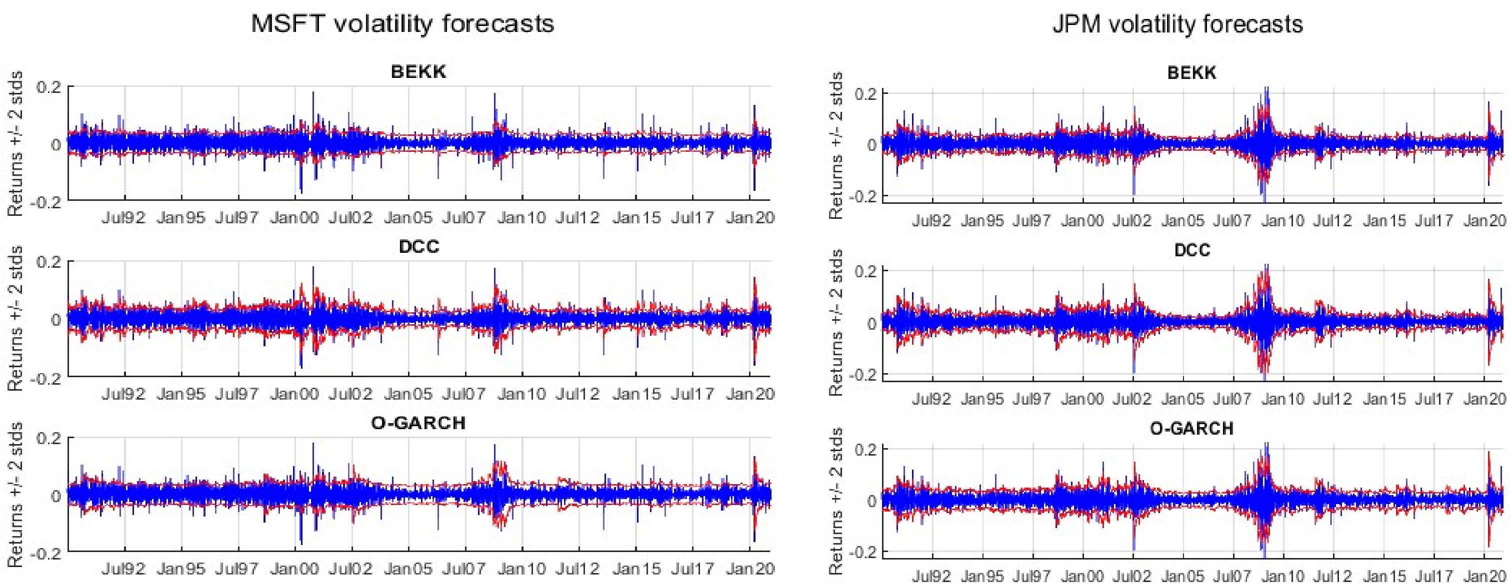

Figure 4: Bivariate volatility estimates for Microsoft and JP Morgan respectively

We compare the individual stock volatility estimates of the three bivariate models in figure 4 above. With regards to the BEKK model, again, in both plots it produces the least accurate volatility estimates suggesting that one of the other two models should be used in subsequent calculations. Comparing DCC and O-GARCH, we can see that in the right plot for JP Morgan volatility estimates, O-GARCH has the tendency to overestimate volatility in calm market periods. Finally, observing the left plot for Microsoft volatility estimates, it is clear that DCC produces more accurate estimates both in periods of calm and volatile markets.

As a final comparison to select the best model between O-GARCH and DCC, we compare the models' correlation estimates against a 100-day rolling window correlation series benchmark, as seen in figure 5.

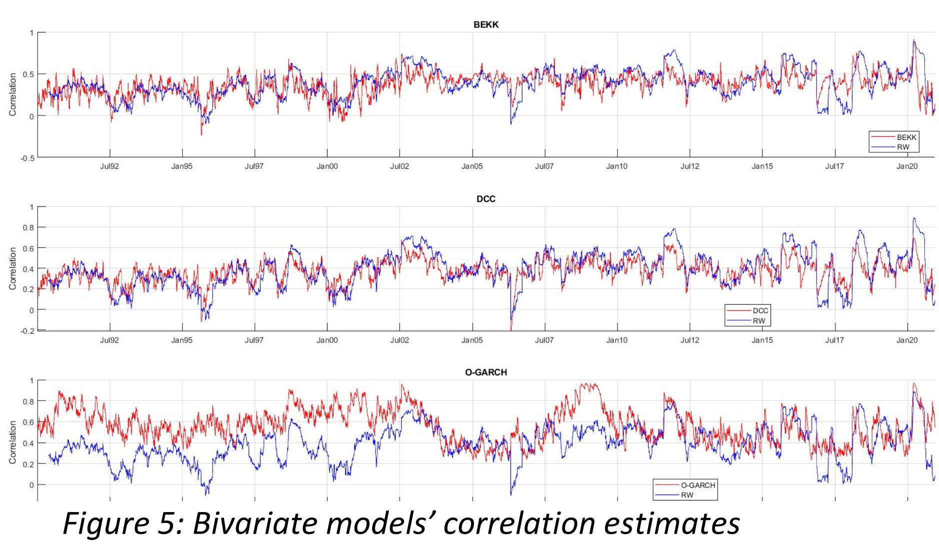

Figure 5: Bivariate models' correlation estimates

We start by noting that the O-GARCH model overestimates correlations massively until around 2003, where it improves significantly in certain periods but with the estimates being far from satisfactory. This could be attributed to the fact that we use a single factor and so the model is unable to capture the idiosyncratic risk of the two stocks, which affects their pairwise correlation to a large extent. Moving on to the DCC and BEKK models, both of them produce very reliable correlation estimates, with DCC producing more stable estimates. However, the BEKK model is overall marginally more accurate as it produces correlation estimates which are in closer agreement to the rolling window benchmark correlation estimates. We finish this section by concluding that based on the above arguments, the DCC-GARCH(1,1) model proves to be the superior bivariate volatility model.

2.4 Selecting a Univariate volatility model

GARCH-type models are known to perform better in estimating conditional volatility compared to other univariate volatility models due to their ability to effectively distinguish between shocks in returns and actual shocks in volatility, and this with only a small number of lags included. For this reason, we have decided to solely focus our univariate model analysis on GARCH-type models and their extensions. We start in this section by introducing the first model implemented, a GARCH(1,1) model with Normally distributed residuals. After fitting the model and calculating the standardised residuals using our portfolio volatility estimates, we produce two QQ-Plots to analyse the residuals. It becomes clear that the quantiles of a standardized t-distribution with five degrees of freedom are in a much closer alignment compared to the Standard Normal quantiles. Hence, we conclude that a GARCH-type model with t-distributed residuals is a more appropriate choice for our data. In turn, we fit two additional GARCH-type models, namely, a GARCH(1,1) and a GJR-GARCH(1,1,1) model with t-distributed residuals. A P-value of 4.48E-08 in a likelihood ratio test strongly suggests that the GJR-GARCH(1,1,1) model might be the superior univariate model. Note also that t-tests for the individual GJR-GARCH(1,1,1) parameters, including the leverage parameter γ, show that all five of them are highly statistically significant and leads us to the conclusion that the model is highly parsimonious.

As a final comparison for the two models, since we are unable to produce pairwise correlations between our portfolio stocks, we evaluate them by comparing their portfolio volatility estimates in figure 6 below. Even though both models produce extremely similar estimates, GJR-GARCH(1,1,1) produces more accurate estimates in extreme negative return shocks by being able to incorporate the leverage effect in its estimates. We end our

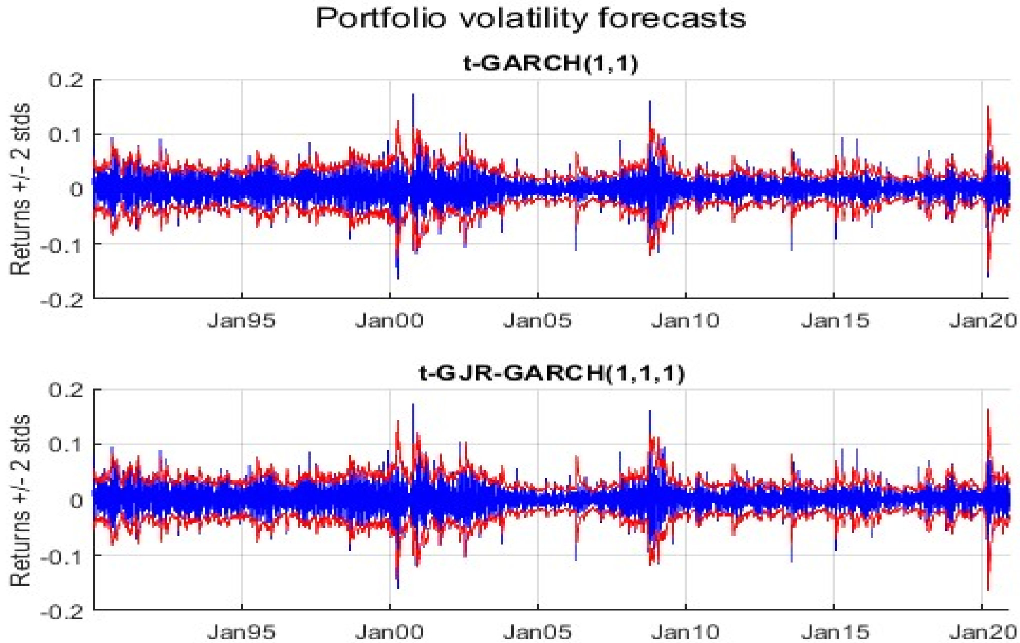

Figure 6: Portfolio volatility estimates for GARCH(1,1) and GJR-GARCH(1,1,1) models with t-distributed residuals

univariate model analysis by concluding that a GJR-GARCH(1,1,1) model is the superior model given our portfolio returns data.

2.5 Comparison of the best performing models

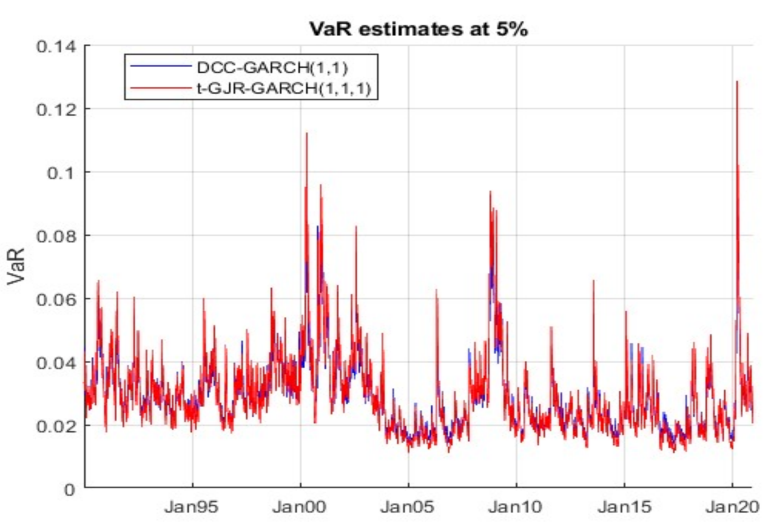

Figure 7: 5% VaR estimates of the best models

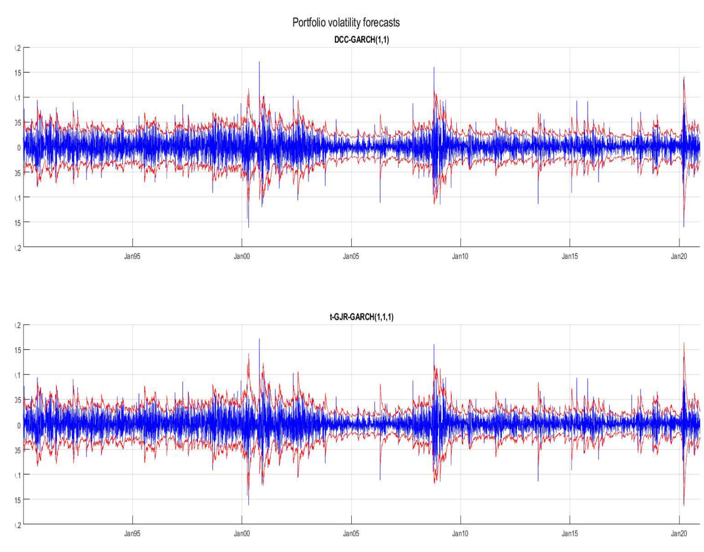

Figure 8: Portfolio volatility estimates of the best models

Observing figure 7, we can see that both GJR-GARCH(1,1,1) and DCC-GARCH(1,1) models produce almost identical 5% VaR estimates and hence we are unable to evaluate their accuracy on the basis of this plot. In figure 8, we note again the extreme similarity between their portfolio volatility estimates, with the only distinction being that GJR-GARCH(1,1,1) yields slightly more accurate results in extreme negative return events, as a result of the leverage parameter.

3 Conclusion

Even though we have ended our analysis in section 1.3 by concluding that a GJR-GARCH(1,1,1) model produces the most satisfactory results given this particular dataset and underlying stocks of choice, we now present some arguments which make us question whether our conclusion could be extended and applied in more general settings.

We re-state that the difference between the two competing models in producing reliable portfolio volatility estimates is only marginal. In addition to this, a P-value of 0.0214 in the conditional coverage test for the DCC-GARCH(1,1) model suggests that if we were to conduct this test at the 1% significance level, then we would conclude that violations do not form clusters in either of the models. To further add to the argument, we highlight that the conditional coverage test we have conducted in our backtesting analysis is only able to detect breaches in independence specified in a particular way. For example, let us assume that independence is not satisfied if the probability of a VaR violation today depends on whether or not there was a VaR violation two days ago. Then, the test we have conducted is unable to detect departures from independence in this case. Therefore, on the basis of the above argument, we must be careful to not jump to conclusions that might not hold.

Lastly, we finish off with an important concern regarding the practical usefulness of univariate models when dimensionality is significantly increased. As the number of underlying stocks in our portfolio increases, we expect that the performance of univariate models will be dramatically reduced due to the fact that their structure renders them unable to incorporate pairwise correlations between stocks in their volatility estimates. In this setting, a DCC model could prove to be a far superior model as it provides an appealing balance between computational cost and estimation precision.