FM320 Course Project

Write Up

Candidate Number: 40946

Date Submitted: 26/01/2020

Time Submitted: 10:00 PM

1. Introduction and Summary

1.1 Objective

Given a dataset containing 7537 daily adjusted log returns for 20 individual stocks, I tasked myself with producing and evaluating an appropriate conditional volatility model to compute VaR and ES risk measures at the 5% confidence level, for a portfolio consisting of Microsoft and JP Morgan stock. In doing so, we made no initial assumptions on whether a univariate or bivariate model was most appropriate, and thus we also had to evaluate which of these incorporated the relevant risk factors and best fit the security data.

1.2 Summary of Results

1.2.1 Univariate Summary

Our chosen univariate conditional volatility model is univariate GJR-GARCH(1,1,1);

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

This is because, when we conducted a likelihood ratio test for the leverage effect, we found it to be highly significant, with a P-Value of $1.412 * 10^{-13} < 0.05$.

Using the above model, we obtain the following estimates for the coefficients:

$$\widehat{\omega} = 4.044 * 10^{-6}, \qquad \hat{\alpha} = 0.0346\,, \qquad \widehat{\beta}\; = 0.9295, \qquad \hat{\gamma} = 0.055$$

With T-test P-values of 0.0478, 0.0042, 0, 0.0077 respectively, and thus all significant at the 5% level.

Producing $VaR_{5\%}$ and $ES_{5\%}$ estimates, using this GJR-GARCH specification, we then conducted both conditional and unconditional coverage statistical tests against the hypotheses:

H_0: $Violations\ are\ I.I.D.\ and\ occur\ with\ probability$ 5%

vs

H_1: $Violations\ do\ not\ occur\ with\ probability$ 5% (Unconditional Coverage test)

or

$H1$: $Violations\ are\ not\ I.I.D.\ over\ time$ (Conditional Coverage test)

Obtaining a $\hat{p} = 0.0385$ and conducting a likelihood ratio test using a χ_1^2 distribution, we obtain a P-Value = $1.767 * 10^{-6}$ for the unconditional coverage test. We therefore reject H_0 at the 5% level, concluding there is highly significant evidence that our model does not produce violations with probability 5%.

Nevertheless, using a similar method, we obtain a P-Value = 0.9606 for the conditional coverage test, leading us unable to reject H_0 at the 5% level, and conclude there is no significant evidence that violations are not I.I.D.

1.2.2 Bivariate Summary

Our chosen bivariate conditional volatility model is DCC-GJR-GARCH(1,1,1);

$$R_t = (1 - \alpha - \beta\,)\bar{R} + \alpha\boldsymbol{Z}_{t-1}\boldsymbol{Z}_{t-1}^{\prime} + \beta R_{t-1}$$

$$\Sigma_t = D_t R_t D_t$$

$$\sigma_{p,t}^2 = \acute{\boldsymbol{w}}_p \Sigma_t \boldsymbol{w}_p$$

s.t. $D_t^{i,j} = \begin{cases} \sigma_{i,t} \; if \; i = j \\ 0 \; otherwise \end{cases}$, $R_t^{i,j} = corr\left(LogReturns_i, LogReturns_j\right)$

where $\sigma_{i,t}$ is estimated for each security, i, using the GJR-GARCH(1,1,1) specification outlined in 1.2.1. and $\boldsymbol{w}_p$ the vector of weights.

Similarly to the work done in 1.2.1, we find the leverage effect significant at the 5% level, with the leverage parameters having T-test P-Values of 0.0281 and $3.44 * 10^{-8}$ for MSFT and JPM respectfully.

Producing $VaR_{5\%}$ and $ES_{5\%}$ estimates, using this DCC-GJR-GARCH(1,1,1) specification, we then conducted both conditional and unconditional coverage statistical tests against the hypotheses outlined in 1.2.1.

Our DCC-GJR-GARCH(1,1,1) unconditional coverage test values are identical to that of the univariate GJR-GARCH(1,1,1), with a $\hat{p} = 0.0385$ and thus P-Value = $1.767 * 10^{-6}$. We therefore again reject H_0 at the 5% level, concluding there is highly significant evidence that our model does not produce violations with probability 5%.

However, we obtain a P-Value = 0.7955 for the conditional coverage test on our bivariate model, leading us to reject H_0 at the 5% level, and conclude that violations may not be I.I.D.

1.3 Interpretation of Results

- As the leverage effect has been shown to be highly significant at the 5% level, for both portfolio and individual stock returns, both the GJR-GARCH(1,1,1) and DCC-GJR-GARCH(1,1,1) models are a reasonable choice to estimate conditional volatility, as they allow us to differentiate between the distributions of both positive and negative returns through the inclusion of the leverage parameter, γ. Furthermore, all four parameter estimates for GJR-GARCH(1,1,1) are significant at the 5% level, leading to increased confidence in this model.

- We note that GJR-GARCH(1,1,1) is univariate, and thus is unable to capture how shocks propagate within and between sectors, in addition to the correlated returns of underlying securities. However, whilst this may constitute a significant flaw in modelling portfolio returns when we have a large number of securities, as our portfolio only consists of two stocks in separate industries, this appears to be a minor factor in the accuracy of our estimates of σ_t^2. This statement can be made, due to the almost identical σ_t^2 estimates produced by GJR-GARCH(1,1,1) and the DCC-GJR-GARCH(1,1,1) model, the latter of which, in theory, is able to capture these effects through the incorporation of dynamic correlation within the model. As mentioned in 2.4.1, however, it is the univariate model which seems to have the edge in predicting the magnitude of shocks, with its estimates of $|2\sigma_t^2|$ being larger and more inline with the realised values of returns during stress events, than that of our bivariate model, indicating its marginal superiority.

- Whilst, as noted in our prior comment, the greater accuracy of our univariate over our bivariate model in predicting σ_t^2 during stress events may indicate a superior model, it could alternatively be a result of GJR-GARCH(1,1,1) being overfit. We note, as we estimated the parameters for both our models using our entire dataset, both are likely overfit as a result and can be expected to perform less well on out of sample data. However, given univariate GJR-GARCH(1,1,1) utilises only 4 parameters, in contrast to DCC-GJR-GARCH(1,1,1) which incorporates 11, it conversely appears more likely that our bivariate mode has a greater probability of being overfit. Furthermore, DCC-GJR-GARCH(1,1,1) also incorporates a greater amount of model risk, through its utilisation of both a Dynamic Conditional Correlation and two GJR-GARCH(1,1,1) models to estimate the same process as a single univariate model. Combined with the high dimensionality of DCC-GJR-GARCH(1,1,1) increasing its parameter error also, this indicates, at least in theory, that our bivariate estimates have the potential of being both more overfit on in-sample data and less accurate when modelling out of sample data, than univariate GJR-GARCH(1,1,1), pointing towards the latter's theoretical superiority.

- The unconditional coverage test revealed both our univariate and bivariate models to be overly conservative, producing volatility estimates which over-forecast risk, thus if either model were employed to hedge a portfolio of Microsoft and JP Morgan stock, it would hinder the investor from maximising their potential returns. In addition to this, the scale of the over-estimation is extreme, with our sample estimate of p (identical for both models) differing from its true value by 23% , and with a near zero P-value. Thus, neither may be the best model to employee in practice from a risk management prospective, with a model which neither under nor over forecasts risk being preferred.

- Ideally, the probability of a violation occurring at time t, should be independent of whether a violation occurred at time $t-1$, otherwise our risk model is failing to incorporate the tendency of violations to cluster together, and can thus be improved. The conditional coverage test shows however, that at the 5% level, there is no such evidence of violation clustering in the Hit Sequence produced by our univariate GJR-GARCH(1,1,1) model. In contrast, we cannot deduce the same for bivariate DCC-GJR-GARCH(1,1,1), and in fact, we even observe a difference in $\hat{p}_{11}$and $\hat{p}_{01}$ of 7.81%, both of which indicate the potential presence of violation clustering. Therefore, we can conclude that in this aspect, our univariate model is superior to multivariate DCC-GJR-GARCH.

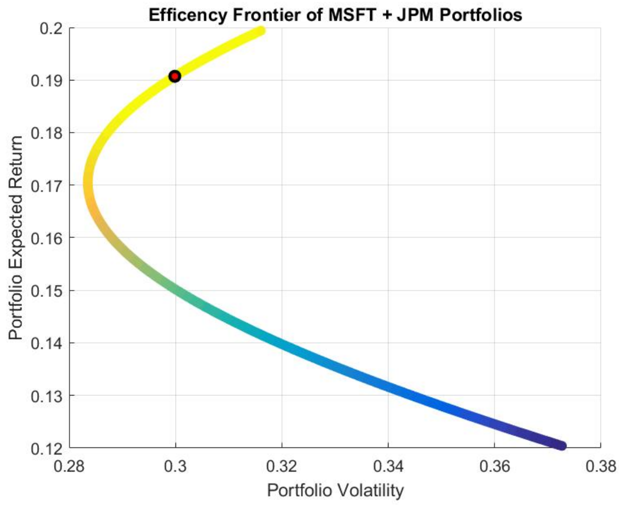
2. Methodology and Analysis

I shall now begin to describe the methods and reasoning I went through in order to deduce the results in 1.3 and my conclusion in 1.4.

2.1 Calculating Optimal Portfolio Weights

In order to produce a model for conditional portfolio volatility, $\sigma_{p,t}$, we first need to produce portfolio weights. Quadratic optimisation provides one method for this, however we note such a process yields an extreme optimal weight for both securities (especially MSFT) at 36.6 and 4.53 for MSFT and JPM respectfully, suggesting the use of an extreme amount of leverage, and thus another method should be used.

As the number of assets is small (2), it is feasible to use Monte-Carlo simulation to randomly assign weights and then select those corresponding to the greatest Sharpe Ratio. – Note for a greater number of assets, such a method becomes very computationally intensive and potentially non-feasible.



Assuming a risk-free rate of 0, we obtain optimal weights 0.8899 and 0.1101, for MSFT & JPM respectfully, corresponding to a Sharpe Ratio of 0.19. We find that these weights are much more reasonable and apply them to our data to calculate log-portfolio returns and $\sigma^2_{p,t}$, in future analysis.

2.2 Multivariate Conditional Volatility Models

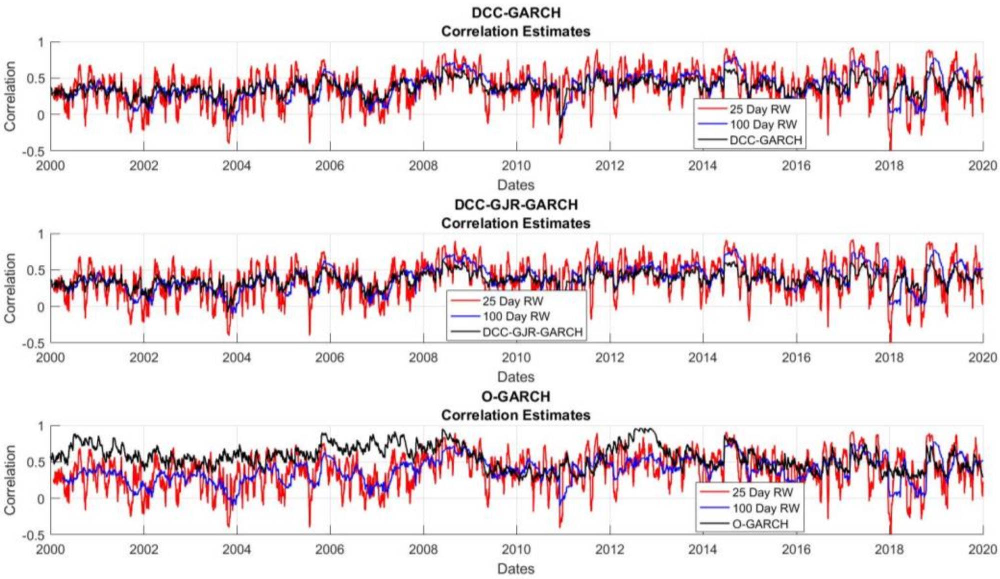
2.2.1 The Models

As we are dealing with a portfolio of 2 stocks, multivariate models allow us to model how these underlying securities interact with each over and influence overall portfolio volatility. We decided to produce two DCC models (as the number of stocks is small and therefore the issue of dimensionality in parameter estimation will be minimal), and an Orthogonal-GARCH model on our data, with the DCC models differing in their methodology for estimating univariate $\sigma^2_{i,t}$ – one using a standard GARCH(1,1) approach and the other a GJR-GARCH(1,1,1) model.

When we performed Principle Components Analysis on or data, we saw it seemed to indicate only a single common factor – The Market, explaining 71% of the variation, whilst the second portfolio represented the remaining idiosyncratic risks associated with JPM & MSFT stocks. This was to be expected, as with only two stocks, PCA can only represent a single commonality between them. Thus, when producing our O-GARCH model, we included only a single factor.

2.2.2 Comparison & Evaluation of Multivariate Conditional Correlation Estimates

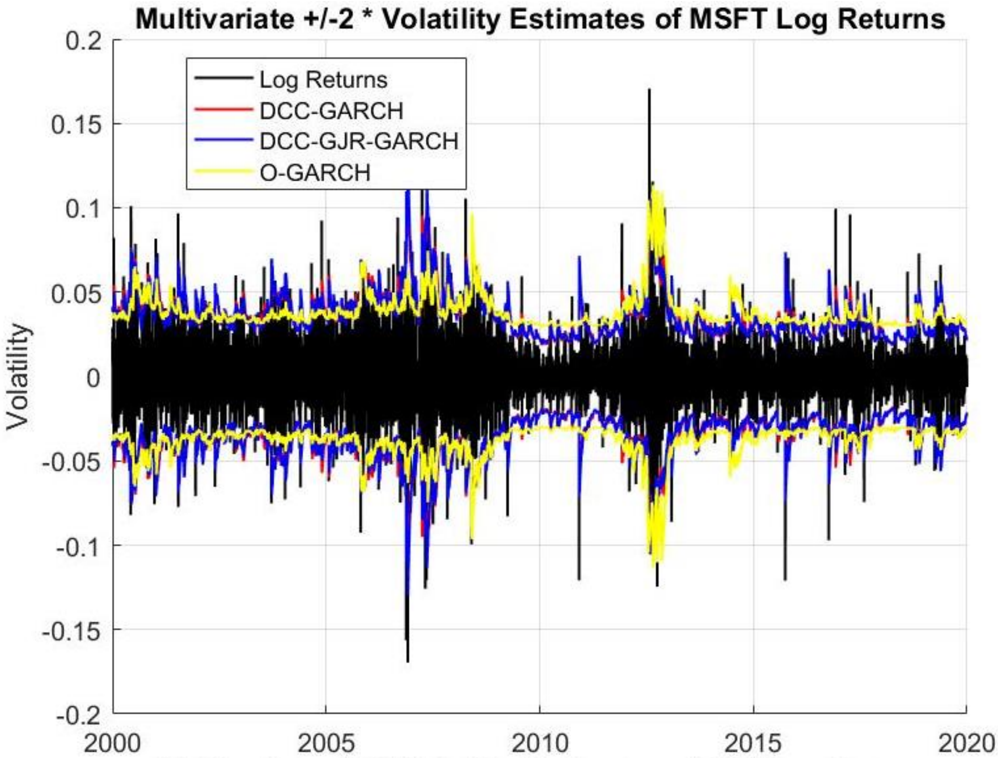
As the supposed benefit of a multivariate model, over a simplier univariate model, is its ability to interpret dynamic correlations of stocks and/or factors into their conditional volatility estimates, this will be the first metric we evaluate them by, with 100 Day rolling window correlation estimates forming our benchmark.

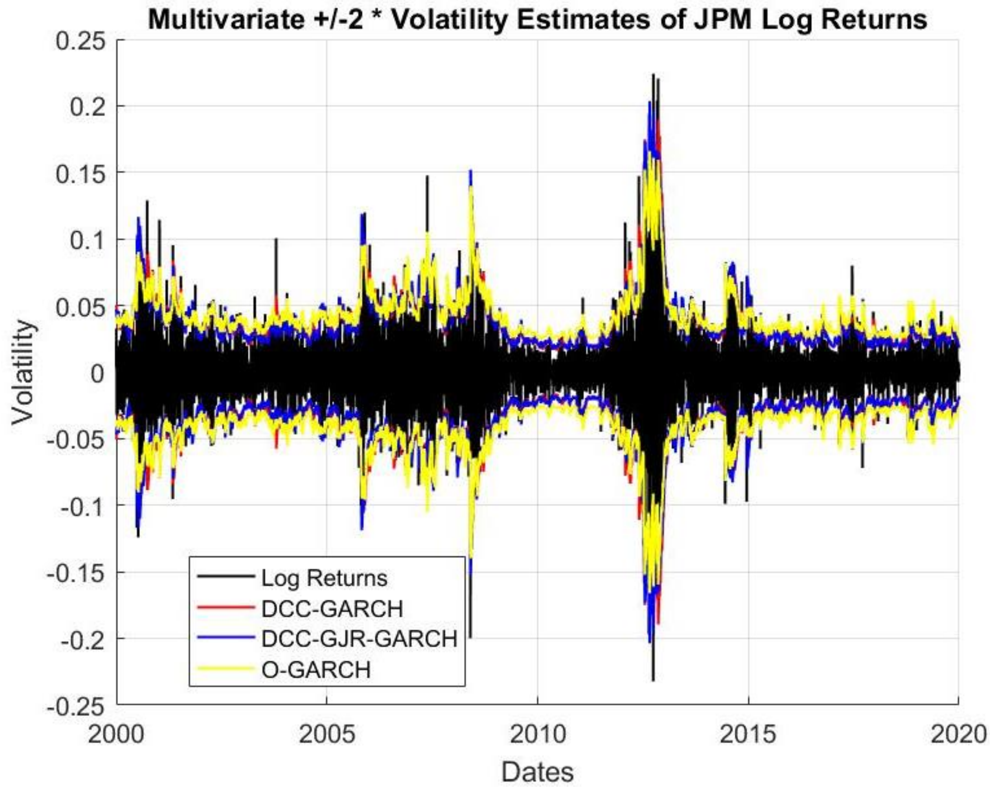


We note that the DCC models clearly provide much better conditional correlation estimates than that of the O-GARCH model, which persistently overestimates correlation relative to the baseline, only producing accurate esimates during and briefly after crisis periods. In fact, with the exception of the years 2009-2013, O-GARCH correlation estimates seem to simply follow a ranging pattern, with no clear underlying trends, likely due to the models failure to capture the idiosyncratic risks of the underlying JPM & MSFT stocks, which are responsible for 29% of the variation, and inpractice has a significant impact on pairwise correlation. This same reasoning can be used to explain why both our DCC models produce significantly better estimates, relative to the baseline, as they incorporate these idiosyncratic risk factors into their conditional correlation estimates through their univariate $\sigma_{i,t}^2$ models.

2.2.3 Comparison & Evaluation of Multivariate Conditional Volatility Estimates

We observe that the O-GARCH volatility estimates for MSFT are incredibly poor, massively overestimating $\sigma^2_{MSFT,t}$ durring calm markets, forming channel like estimates. Furthermore, this lack of sensitivity to returns, persists even throughout stress events, when the market factor should theoretically cause an appropriately large increase in $\sigma^2_{MSFT,t}$ estimates. However, as the graph to the right shows, there are countless volatility breaches of the model, due to a failure to incorporate indiosyncratic risk into the model, leading to volatility being greatly underestimated during shocks.

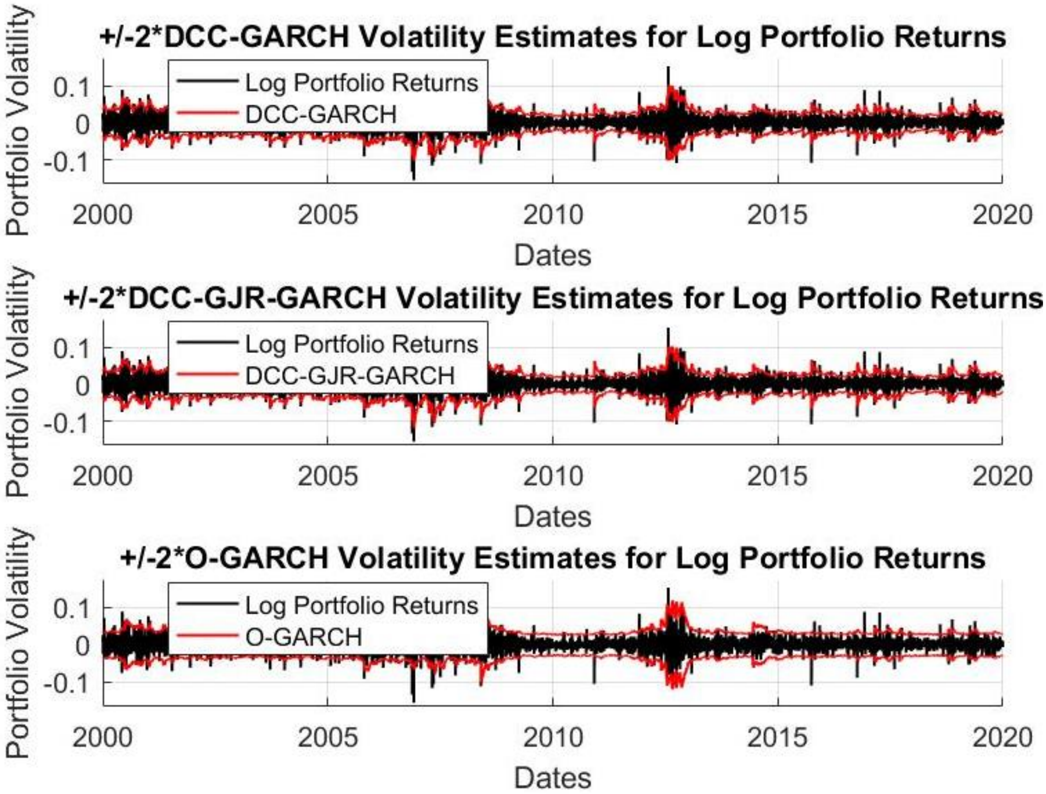




Whilst this channel like behaviour does not seem apparent in O-GARCH estimates for $\sigma^2_{JPM,t}$, we still see a slight overestimation of volatility during calm markets and underestimation throughout shocks. This analysis proves O-GARCH to be a poor model for single stock volatility, something which greatly affects estimates of $\sigma^2_{P,t}$, which are equally poor and resemble the channel like behaviour of $\sigma^2_{MSFT,t}$, likely a product of the large portfolio weighting applied to this security. Thus, without repeating the same comments made about MSFT, we conclude this to be a very poor bivariate model.

In contrast, we see that our DCC models provide very good estimates, for both stock and portfolio volatilities, neither over nor underestimating $\sigma^2_{i,t}$ consistently. For this reason, as well as earlier analysis made on correlations in 2.2.2, it is reasonable to conclude that in this case, DCC type models are superior.

Whilst the distinction between the estimates of DCC-GARCH(1,1) and DCC-GJR-GARCH(1,1,1) are not immediately apparent, DCC-GJR-GARCH(1,1,1) does seem to produce volatility estimates which fit the obvserved results more accurately. This is primarily a result of a smaller parameter estimate $\widehat{\omega}$ for the DCC-GJR-GARCH(1,1,1) model than that of DCC-GARCH(1,1) and thus a smaller volatility lower bound.



Furthermore, unlike both DCC-GARCH(1,1) and Orthogonal-GARCH, DCC-GJR-GARCH(1,1,1) is able to incorporate the leverage effect into its $\sigma^2_{P,t}$ estimates, leading us to conclude it to be the superior multivariate model.

2.3 Univariate Conditional Volatility Models

2.3.1 The Models

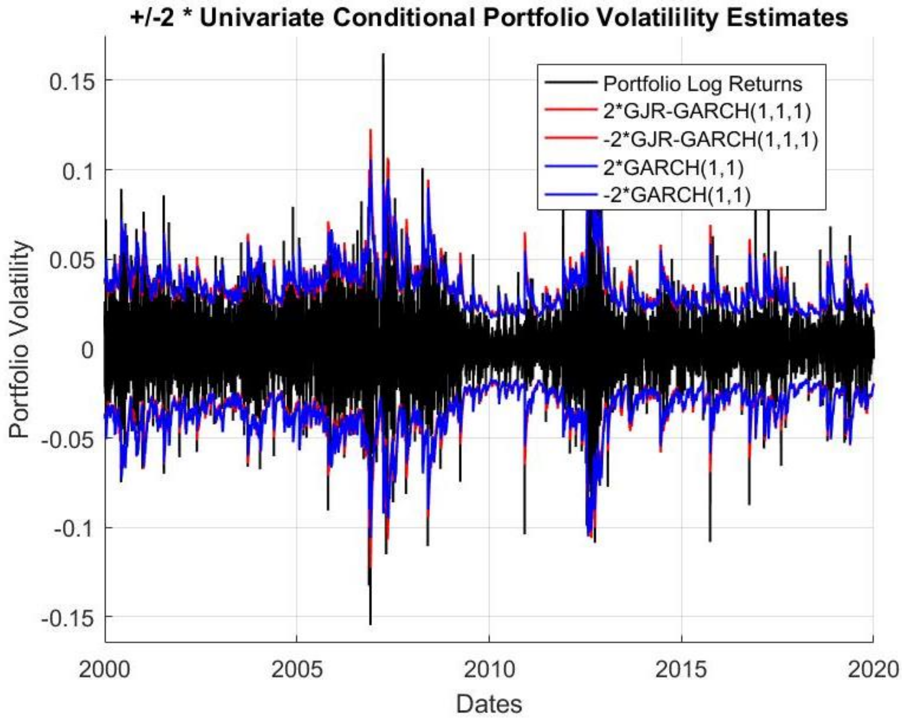
Incomparison to the multivariate models in 2.2, the benefit of using a univariate model is found in their simplicity, posessing a fixed number of parameters even as the size of the portfolio increases, thus not running into the issue of dimensionality and severe parameter error. This is in contrast to the DCC models mentioned earlier, where parameters are a quadratic function of the number of stocks. In addition to this, these models also do not ommit idiosyncratic risk factors in the same way simplified multivariate models, such as O-GARCH, do. Therefore their $\sigma^2_{P,t}$ forcasts can prove more accurate, despite not modelling pairwise correlations. We decide to produce two Garch type models – GARCH(1,1) and GJR-GARCH(1,1,1), for estimating $\sigma^2_{P,t}$.

2.3.2 Comparison & Evaluation of Univariate Conditional Volatility Estimates

We observe that, whilst $\sigma^2_{P,t}$ estimates are extremely similar for the two univariate models, GJR-GARCH(1,1,1) produces greater and more accurate estimates during stress events, when returns are negative. Therefore it is reasonable to conclude, that GJR-GARCH(1,1,1) would serve as the better model for hedging positions during crises.

Furthermore, conducting a log-likelihood ratio test on the two models, reveals the leverage effect captured by GJR-GARCH(1,1,1), to be extremely significant, with a P-Value = $1.412 * 10^{-13} < 0.05$.

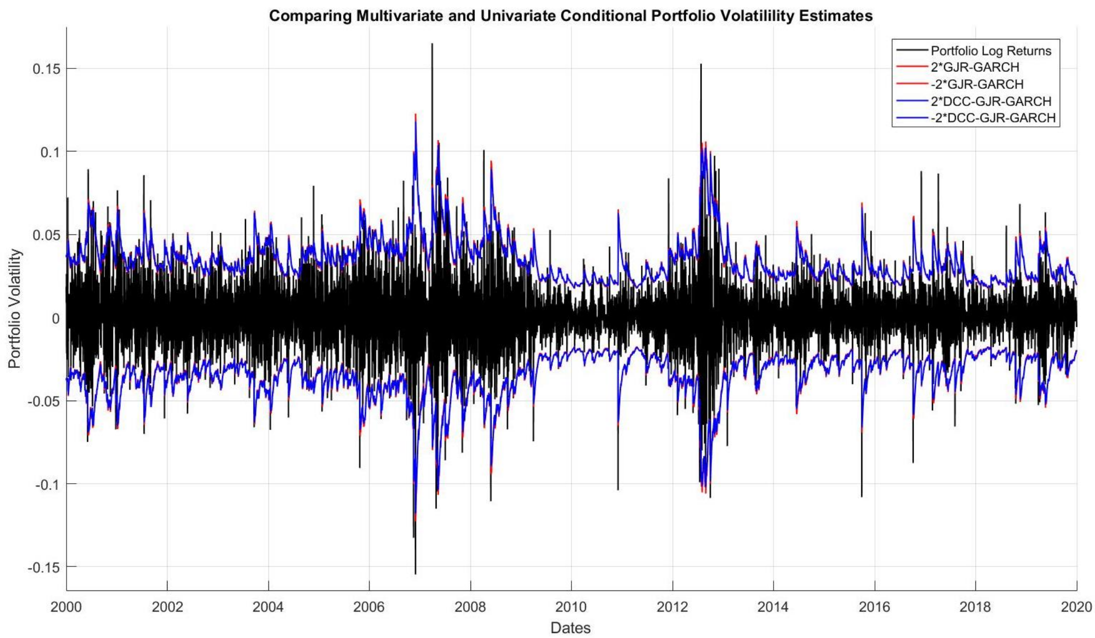
Given these factors, it is reasonable to conclude GJR-GARCH(1,1,1) to be the superior univariate model, for this portfolio.



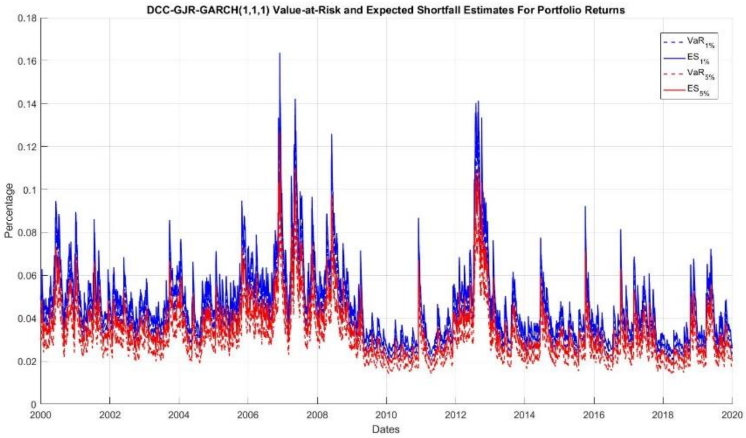
2.4 Comparing Univariate & Multivariate Conditional Volatility Models

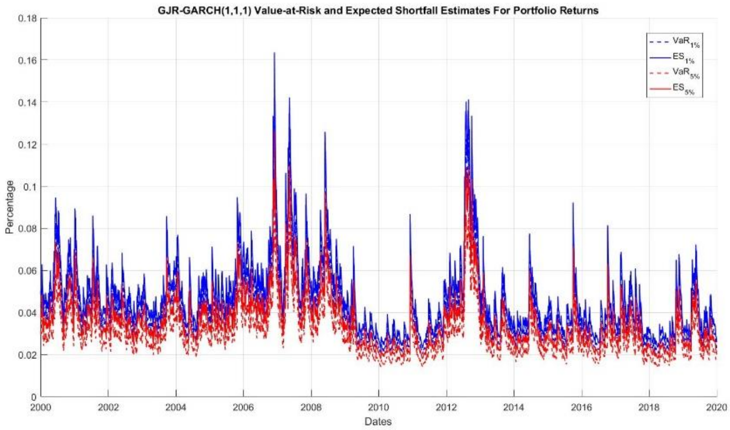
2.4.1 Comparing Conditional Volatility, VaR & ES Estimates

Looking at the volatility estimates produced by both GJR-GARCH(1,1,1) and DCC-GJR-GARCH(1,1,1), we note their extreme similarity, differing only slightly during extreme events, with the univariate model yielding marginally greater and more accurate $\sigma^2_{P,t}$ estimates.



This extreme similarity is further highlighted, when looking at conditional VaR and ES plots. Thus to determine if either model is superior, we must evaluate their statistical characteristics.





2.4.2 Comparing VaR & ES Backtests

To perform statistical tests upon our conditional VaR & ES estimates, we must first compute the Hit Sequence: The sequence of violation indicators, $V_t\ , s.t.\ V_t = \begin{cases} 1\ if\ LogReturn_t < -VaR_p \\ 0\ otherwise \end{cases}$

Having computed this, we then perform Unconditional and Conditional Coverage tests, with respect to the hypothese in 1.2.1. The subsequent analysis which follows is idential to that in 1.3 and is thus omitted.

3. Evaluation and Conclusion

3.1.

Given the remarks made in 1.3, it seems rather simple to conclude univariate GJR-GARCH(1,1,1) to be our model of choice for any subsequent analysis conducted on a MSFT-JPM portfolio. However, we note that despite its lower probability of being overfit and lower estimation error (through both lower dimensionality/parameter and model errors), GJR-GARCH(1,1,1) still produces only marginally better volatility estimates that DCC-GJR-GARCH(1,1,1) and only then during stress-events. Otherwise the two models are indistinguishable from each over in terms of output, and thus this these theoretical pros don’t seem to add much benefit in practice.

Furthermore, both models are deemed overly conservative when we backtest VaR and ES, and whilst GJR-GARCH(1,1,1) produces a statistically I.I.D Hit-Sequence, when DCC-GJR-GARCH(1,1,1) fails too, it does so only marginally at the 5% level, with us unable to label it highly statistically significant (i.e. a P-Value < 0.01). Therefore it seems unlikely that GJR-GARCH(1,1,1) is an optimal model, with one which is able to produce both an I.I.D Hit-Sequence, without over/under-forecasting risk in a backtest and maintaining the capacity to incorporate the highly statistically significant leverage effect, being preferred.

Given that we tested 5 different possible models however (GARCH(1,1), GJR-GARCH(1,1,1), DCC-GARCH(1,1,1), DCC-GJR-GARCH(1,1,1) & Orthogonal-GARCH), it may be the case that such a model is exceptionally hard to produce, and may not in fact even yield our desired results for different combinations of weights or securities. If this is indeed true, it may be best to use a univariate approach with a GJR-GARCH(1,1,1) model.

3.2 Higher Dimensionality

As the number of underlying securities rises, so two will the dimensionality of DCC-GJR-GARCH(1,1,1), and thus its estimation error, which is when the output of our univariate model may start to become distinguishable as being more accurate. Yet even in such a scenario, one could debate it is exactly then that the necessity to incorporate the effects of sector specific shocks increases, and thus the benefits of a dynamic conditional correlation model may become more useful. Therefore, we are unable to extrapolate our findings for portfolios of greater dimensions and recommend both univariate and multivariate models be estimated and evaluated, by performing a similar methodology to that laid out in 2., in order to find and select the most optimal model.