ASSESSED HOMEWORK 2

LSE ID: 202252142

Topic and Stocks

This project will consist of backtesting VaR computed using three different models, Historical Simulation, GARCH and t-GARCH for three stocks over different time horizons. The chosen stocks are Bank of America (BAC), BP plc (BP) and Johnson & Johnson (JNJ). The chosen stocks are from diverse industries, banking, energy and consumer goods sectors respectively, vary in how volatile they are and how they react to market shock.

Comparing different VaR estimates with Returns

Fig.1-Returns and VaR estimates-BAC

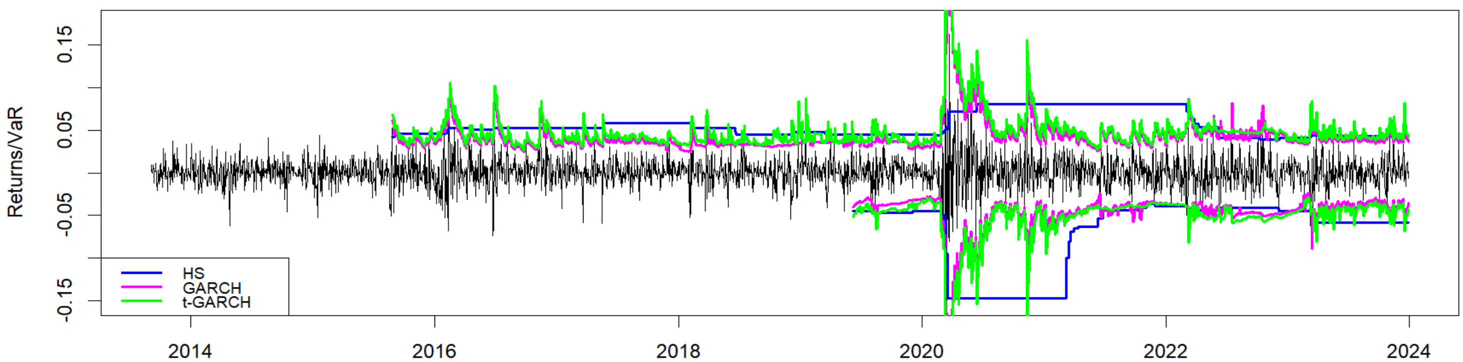


Fig.2-Returns and VaR estimates-BP

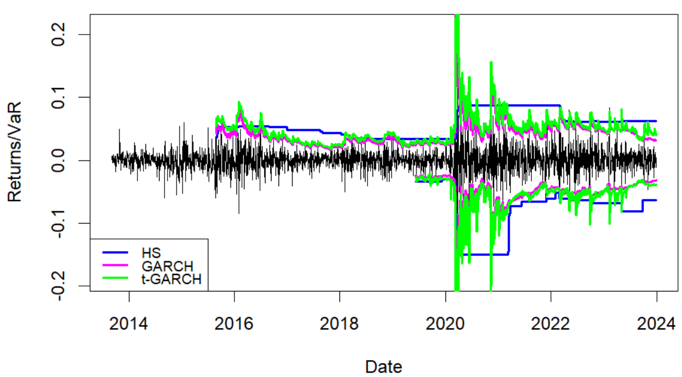
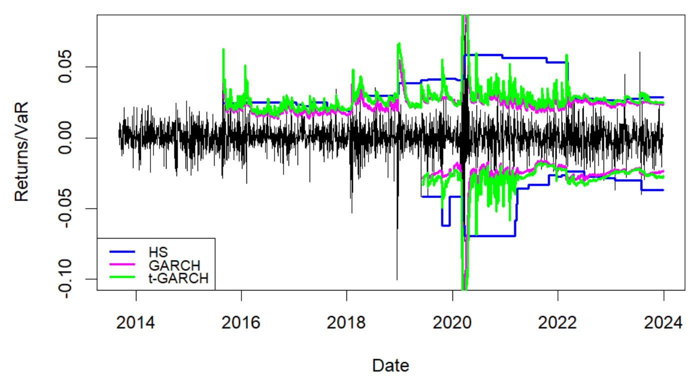


Fig.3-Returns and VaR estimates-JNJ



All three graphs show the returns for three stocks along with their forecasted VaR estimates using HS, GARCH and t-GARCH for the three stocks. While the top lines refer to the VaR estimates for 10 years, the ones at the bottom refer to VaR estimates calculated using 5 years of data. We can infer from all the graphs that HS takes a lot of time to digest shocks such as COVID-19 and remembers for a long time, causing it to overestimate risks in the following periods (till the first quarter of 2022) where the actual returns were far less volatile. This perfectly aligns with the theory as HS is largely affected by what happened before and it replicates the distribution from the past into the next period. In addition, when the time period is longer, HS reacts even more time to incorporate shocks and takes longer time to forget its effect. GARCH on the other hand, fluctuates more and is quicker in adapting itself to different periods of volatility. This again aligns with theory as GARCH is parametric and its coefficients alpha helps in quicker understanding of previous period volatility. Lastly, t-GARCH adapts the fastest to changing volatility and is working the best in stressful situations. Again owing to theory, t-GARCH uses the Student-t distribution for residuals instead of the normal distribution and the additional parameter, degrees of freedom accommodates higher kurtosis in real life financial returns.

This can be graphically represented by plotting the rolling window estimates of degrees of freedom extracted from tGARCH estimates against time, as shown in plots 3 and 4.

Fig.4-Returns & tGARCH dof-BAC

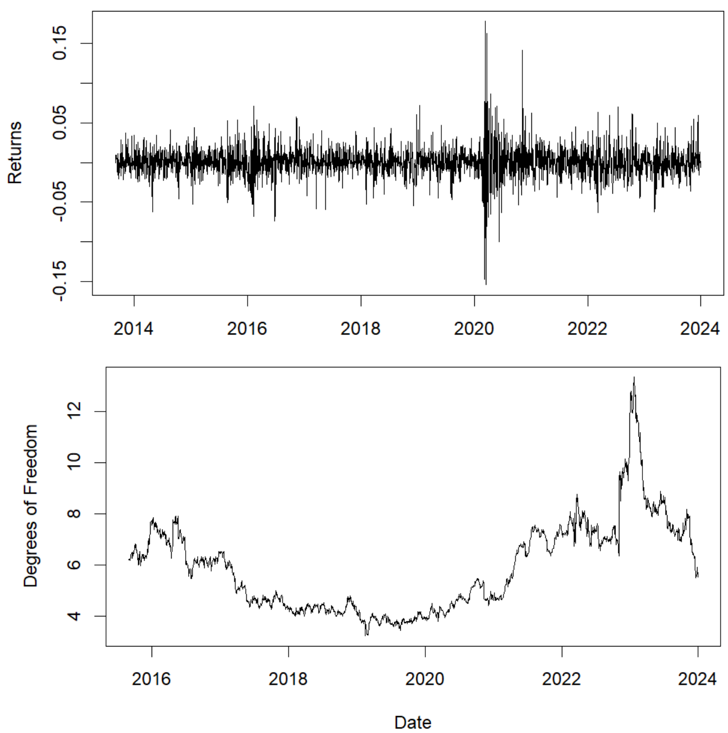
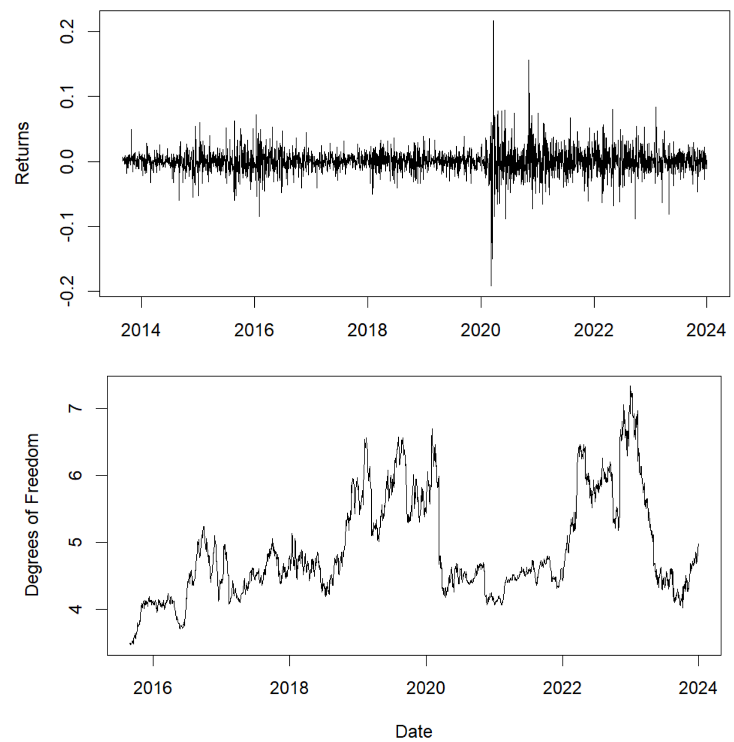


Fig.5-Returns & tGARCH dof-BP



From both plots, we can observe that during periods of stress such as COVID-19, the degrees of freedom reduce drastically to approximately 3, indicating fatter tails to accommodate extreme values. On the other hand, during stable periods with less volatile returns, the degrees of freedom parameter increase. As per the theory, the model will approach the behavior of a GARCH model with normal residuals and degrees of freedom equal to infinity when the returns require thinner tails. This kind of flexibility to change the thickness of tails according to market outcomes is not offered by standard GARCH, leading to inaccurate forecasting as compared to t-GARCH. For instance, BP faced severe financial repercussions from the 2010 Deepwater Horizon oil spill and the years between 2014 to 2016 were a financial crisis for BP. The impact of this is not only shown in the high volatility of returns during those years but there is also a drop in the degrees of freedom parameter to accommodate extreme returns.

<u>Violation Ratio</u>

Table 1 below shows the violation ratio of the three different stocks over both periods for all three methods of VaR. It is computed by taking into account the observed number of VaR violations wherein the observed loss is greater than the loss forecasted by VaR. Cells highlighted in green, orange and red represent if the model is good, acceptable, or bad respectively. We can infer that across all stocks, GARCH has constantly been a bad model by underestimating the risk (VR>1). This again validates theory and referring back to figures 1 to 3, GARCH underestimates risk during volatile periods and ignores the heavy tails present in real-world data. T-GARCH on the other hand is on the borderline of being a good model and has always been at least an acceptable model. Over the 10-year time horizon, both have a lower violation ratio, showing they are becoming more accurate. Attributing this to theory, over longer periods, GARCH and TGARCH models learn to capture the underlying volatility structure more effectively as they are fed more historical data and noise from market shocks like COVID-19 averages out, producing more persistent, stable forecasts. Coefficients of GARCH/TGARCH models (see figures 11 and 12) are better calibrated and stable in the long term leading to better forecasts. However, HS has proven to be the best method if we take the violation ratio as an indicator and it enhances by reducing the time horizon. Again, theory suggests how HS directly relies on historical data without the use of parameters and hence, shorter periods ensure that the historical returns reflect recent market conditions more closely.

Table 1- Violation Ratios

	10 years			5 years		
	HS	GARCH	t-GARCH	HS	GARCH	t-GARCH
BAC	1.190	1.381	0.905	1.13	1.478	0.957
BP	1.048	1.714	1.095	1.043	2.261	1.391
JNJ	1.1904	1.952	1.238	0.696	2.087	1.304

Fig.6- VaR violations-JNJ-10yrs

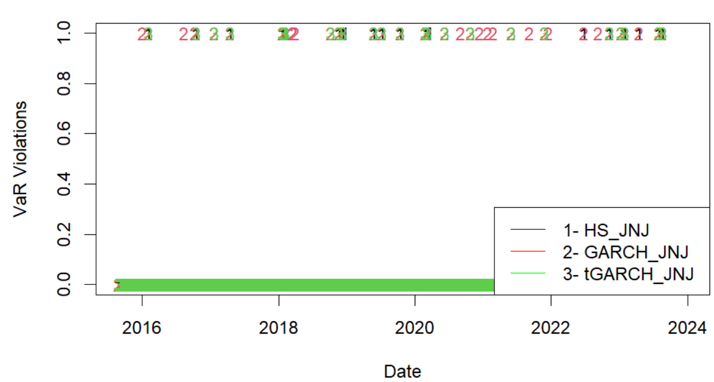
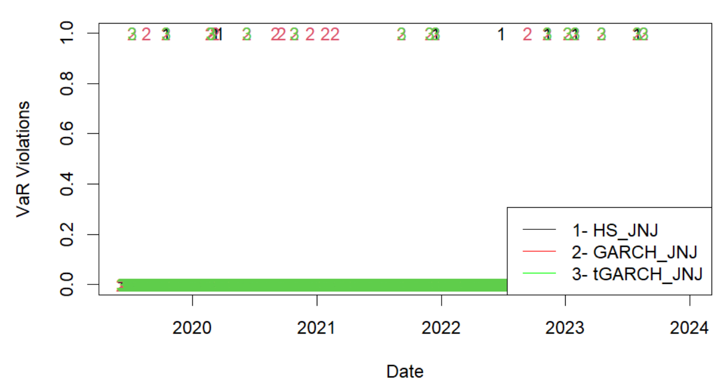


Fig.7- VaR violations-JNJ-5yrs



Following a period of high volatility in 2020, in the year 2021, Johnson & Johnson's stock price faced fluctuations owing to COVID-19 Vaccine developments. While this unexpected high volatility is covered by HS as an effect to delayed decay effect from COVID-19's effect, t-GARCH takes into account this volatility by using its degrees of freedom parameter and GARCH produces violations. During periods of stress, like COVID-19, all three models violate and cluster. As mentioned earlier in this section, we can see fewer violations of HS in the 5-year forecast.

Coverage Test

Under the Bernoulli coverage test, we see if the observed number of violations matches the expected number of violations, given by the significance level and testing window size. In our case, the expected number of violations for the 10-year VaR of all the stocks would be 21. The chi-squared critical value is 6.635. The null hypothesis states that the violations follow a Bernoulli distribution and match the expected violations.

As seen in Table 2, it is again GARCH that does not seem to be an appropriate model. For two out of the 3 stocks, the null hypothesis is rejected indicating that the number of observed observations is statistically significant from the expected number of violations. This is also evident from the number of violations as GARCH has a significantly higher number of violations as compared to other models and is evidently underestimating risk. This is again owing to the issue of its inability to capture extreme outcomes during periods of stress. HS, on the other hand, works well in keeping violations under control even though it suffers from the issue of reacting slowly.

Table 2- Coverage test estimates

	HS			GARCH			t-GARCH		
	LR	Result	No. of violations	LR	Result	No. of violations	LR	Result	No. of violations
BAC	0.73	Cannot reject the null	25	2.75	Cannot reject null	29	0.2	Cannot reject the null	19
BP	0.05	Cannot reject the null	22	8.92	H0 is rejected	36	0.19	Cannot reject the null	23
JNJ	0.73	Cannot reject the null	25	15.06	H0 is rejected	41	1.12	Cannot reject the null	26

Independence Test

By observing if the model violates on two consecutive days, this test is run on the stock of Bank of America for both periods to see if we can predict a violation for time T if there was a violation on time T-1. The null hypothesis is that the violations are randomly distributed with no dependence on each other.

Table 3- Independence test results

	10 years			5 years		
	LR Statistic	P-Value	Result	LR Statistic	P-Value	Result
HS	14.554	0.0001	Reject Null	2.244	0.1341	Fail to Reject Null
GARCH	3.416	0.0646	Fail to Reject Null	1.332	0.2485	Fail to Reject Null
t-GARCH	0.347	0.557	Fail to Reject Null	2.864	0.0906	Fail to Reject Null

From Table 3, we can see that while all models pass the independence test, HS in the 10 year period fails to reject the null. It has a high Likelihood Ratio test statistic which confirms strong evidence of dependence. This is attributed to the theory discussed in the previous sections wherein HS forecasts VaR solely on the basis of historical returns and assumes future returns to follow the distribution of the present. For instance, market shocks take a long time to diminish even if the realized returns have become stable. All these factors make the violations also to be systematic and show dependence. In addition, the 5 year forecast shows that even HS is independent and in theory, this is because as we minimize the sample size, say from 500 to 300 with 0.01 significance, the 3rd smallest value changes more often than the 5th smallest observation, thereby implying that the HS takes lesser time to observe shocks, thereby making it efficient.

The results of the independence test are also shown in Figures 8, 9 and 10. We can observe the clustering of violations in HS during periods of stress in 2020 due to COVID-19 and this clustering cannot be seen in GARCH or t-GARCH. Again, as per theory the market shock of COVID-19 persisted for a long time under HS, leading to violations one period after the other. However, GARCH and t-GARCH, show more violations as compared to HS even though they are parametric. This shows that HS would be a good model to estimate VaR in stable periods with no market shocks.

Fig.8-VaR Violations Over Time-HS-BAC

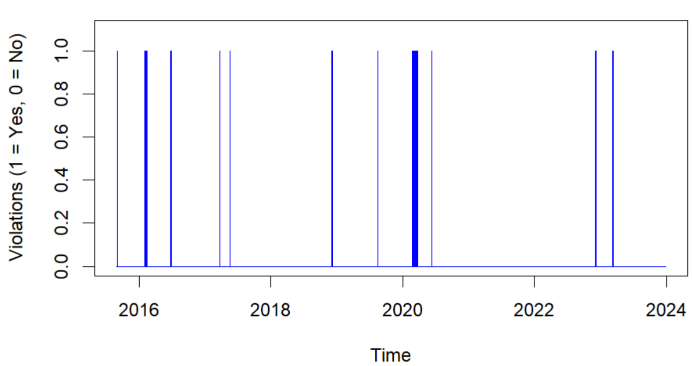


Fig.9-VaR Violations Over Time-GARCH-BAC

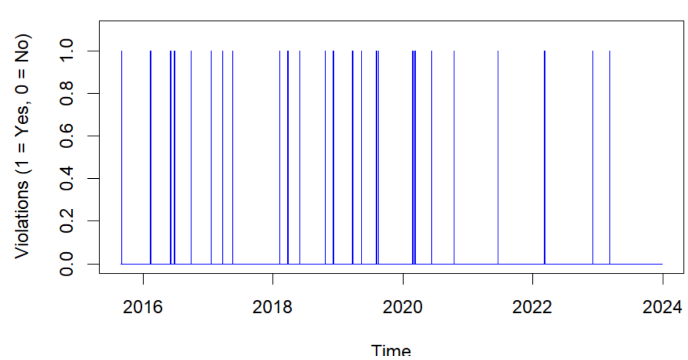


Fig.10-VaR Violations Over Time-tGARCH-BAC

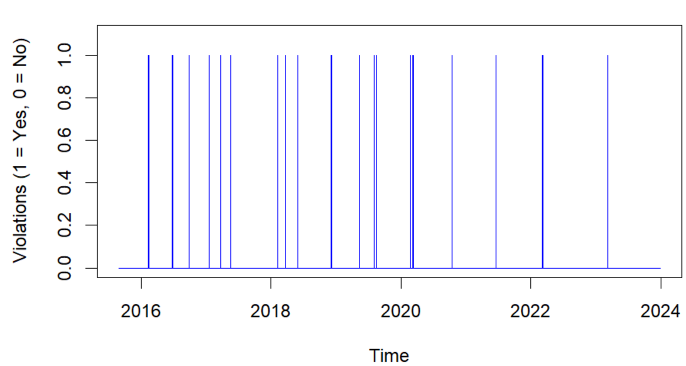
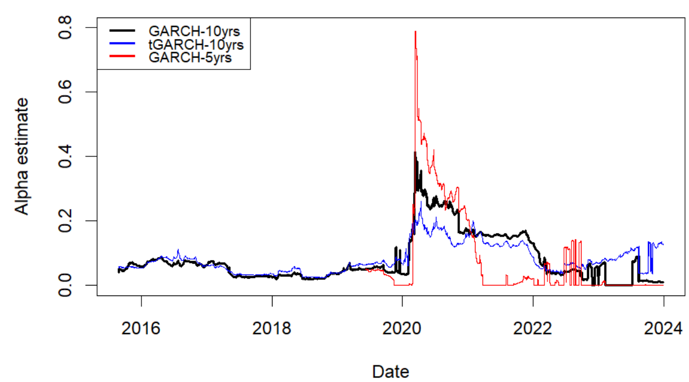


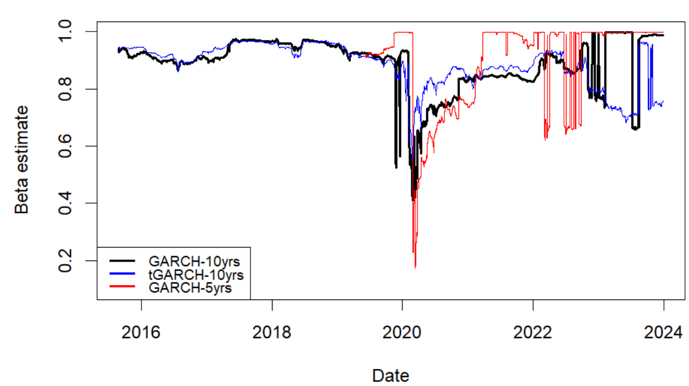
Fig.11-Alpha estimates from GARCH and tGARCH-BP



Time series analysis of parameters

From Figures 11 and 12 we can observe that the GARCH-5yrs model is highly reactive to shocks, showing sharp spikes in alpha and beta during the COVID-19 crisis, reflecting its sensitivity to short-term market disruptions. This aligns with the theory wherein the GARCH-10yrs and tGARCH-10yrs models, with longer estimation windows, smooth out short-term noise, producing more stable estimates. T-GARCH-10yrs model provides slightly smoother estimates, effectively capturing market behavior.

Fig.12-Beta estimates from GARCH and tGARCH-BP



Conclusion

GARCH violates the most owing to its inability to capture extreme observations in real-life financial data. T-GARCH on the other adapts itself effectively for varying periods of volatility owing to its degrees of freedom parameter, accommodating extreme observations that GARCH misses out on. Following periods of stress, HS overestimates risk until the effect decays making it inefficient as a risk forecast. This also makes it fail the independence test. Therefore, historical simulation can be a good model for stable periods and for stocks like Johnson & Johnson with less volatility. Its efficiency can be improved if we minimize the sample size. During crises, t-GARCH would be the best model as with a reasonable number of violations, it captures market violation in the best way possible without underestimating risk. Therefore, there is no single VaR model that is the best. It highly depends on the market situation and the nature of the volatility of the specific stock that we are looking at.