Project 1: A Stock Comparison of Shell and Vestas under the COVID-19 crisis and the 2021-2023 Energy Crisis.

ORIE 5640
Spring 2023
October 20, 2023

Mads Berg, mpb99
Nicolaj Hans Nielsen, nhn6

1 Introduction

The energy markets have been through an intense period with volatility, tightness, and record prices for gas and electricity over the last three years. We have seen two major events the COVID-19 crisis and a global energy crisis with gas shortages which was further accelerated by the invasion of Ukraine. In this project, we will compare the impact of the two periods of crisis on two different stocks in the energy sector. Specifically, we will compare a big oil company, Shell, and a producer of wind turbines, Vestas, which is an enabler of the future of green energy. We compare the log returns from March 1st, 2020 to March 1st, 2021 with the period from March 1st, 2021 to March 1st, 2023; the former we call the *COVID period* as COVID put a huge distress on the markets and the latter the *energy period* is where the energy crisis hit the energy markets - especially in Europe. We will initially describe the data in the periods, then consider univariate models within each period and subsequently model their joint distribution using multivariate Gaussian and t-distributions.

2 Data

We use the **R** package `quantmod` to access adjusted stock prices directly from Yahoo Finance where we use the symbol `SHELL` for Shell and `VWDRY` for Vestas. To get some context about the period the data was collected in, we will give some brief highlights of some energy and macroeconomic events in the periods. The data starts from March 1., 2020 when COVID lockdowns led to a cool-down of the economy. When economic activity decreases, the demand for energy commodities falls drastically as e.g. factories run at lower capacity, there is less transportation of goods, etc. Followed by the general increase in stock prices during the lockdown there is a sudden fall at the beginning of 2021 for Vestas. It is reported as a market correction or burst of the *green bubble*, [1] as green companies were reportedly priced too high in the market. Throughout 2021 there was an economic rebound with an unexpected surge in demand for energy commodities in Europe which turned into a regular crisis starting at the end of 2021 [2] [3]. This also spilled over to related commodity markets as metals and related construction commodities that are essential input prices for Vestas. On the 24th of Feb., 2022 Russia invaded Ukraine which worsened the energy crisis in Europe as Russia is the main gas supplier. The events are quite visible in the price series in fig. 1.

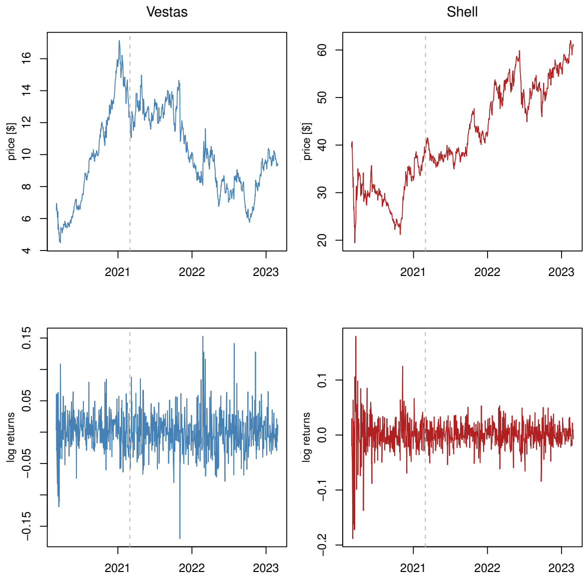

Figure 1 – Prices and log returns with a line to separate the periods of study.

3 Results

In fig. 1, we see that the price of Vestas initially has a positive drift followed by a negative drift whereas shell seems to have a net positive drift in the period. The time series of log returns seem to exhibit volatility clustering which is a characteristic property of stock returns, [4]. In table 1, we see that the means of the log returns are small but during COVID, Shell had negative average returns while Vestas had positive returns on average. This subsequently flipped during the energy crisis. We see that the standard deviation of Shell was reduced by half in the energy period while it remained approximately the same in each period for Vestas. In the COVID period, Shell was moderately negatively skewed but became more symmetrical in the energy crisis. Vestas is fairly symmetrical but drifted somewhat from positive to a slight negative skewness between the periods. All of the log returns have tails that are heavier than normal but Shell had considerably high kurtosis during the COVID period.

	COVID		Energy Crisis	
	Shell	Vestas	Shell	Vestas
mean	-0.00020	0.00259	0.00091	-0.00055
std	0.04137	0.03138	0.01960	0.03262
skewness	-0.74773	-0.34414	-0.28308	0.47340
kurtosis	8.59593	5.11773	4.06355	6.41717

Table 1 – Summary statistics for each period.

In fig. 2, we can assess the dependence of the log returns in the period. We notice a couple of days with a strong tail dependence and a positive correlation coefficient as the returns have the same sign on the day of the tail events. Using the color scale, we see that the tail events happened mostly during the first part of the COVID period. There are also a couple of tail events in the energy period but it seems that the tail dependency is not as strong. The empirical correlation coefficient during the COVID period is $\hat{\rho}_{co} = 0.46$ while it is only $\hat{\rho}_{en} = 0.15$ during the period of energy crisis.

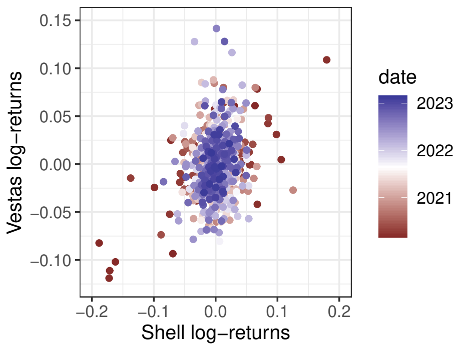

Figure 2 – Joint distribution of log returns.

Since the kurtosis is larger than 3, we will compare the log returns for each stock with the quantiles of t-distributions with different tail indices. We graphically determined which integer tail index seemed most appropriate for each company in each period to get a rough idea about the distribution. In the upper left of fig. 3, we see that the quantiles of Shell log returns during COVID seem to align somewhat with a t-distribution with a tail index of 3 with some deviations in the left tail. Likewise, Vestas seems to have a tail index of around 4 in the period with slight deviations in the left tail. In the energy period, we see that a symmetric t-distribution with an index around 7 seems appropriate for Shell which is quite an increase from the tail index of 3 in the COVID period. This underpins that the tails are really heavy in the COVID period. For Vestas in the energy period, it seems that the tail index is around 5 but the data towards each end of the graph seem to deviate slightly more. Because the data deviates slightly from symmetric t-dist. and because we have some moderate skewness estimates, we could try to fit a skewed t-distribution, see e.g. [4] for an introduction.

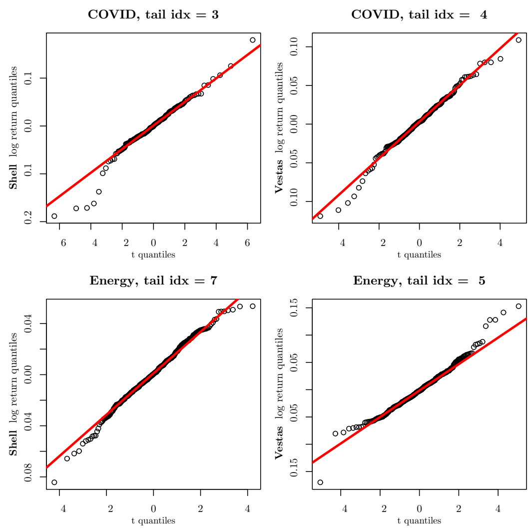

Figure 3 – QQ-plots of log returns for t-distributions with a suitable tail index for each company in each period.

3.1 Univariate Distributions

We will now fit t-distributions to the log-returns using maximum likelihood estimation.

	μ	σ	ν	AIC	BIC
Shell	0.0007	0.0454	2.8328	-949.03	-938.47
	$\pm$0.0037	$\pm$0.0173	$\pm$1.1554		
Vestas	0.0030	0.0323	4.0773	-1039.92	-1029.35
	$\pm$0.0033	$\pm$0.0062	$\pm$2.1452		
Shell	0.0011	0.0197	7.4390	-2542.33	-2529.66
	$\pm$0.0015	$\pm$0.0017	$\pm$4.7344		
Vestas	-0.0018	0.0326	4.7895	-2066.37	-2053.70
	$\pm$0.0024	$\pm$0.0037	$\pm$1.9343		

Table 2 – Parameters of the fitted univariate symmetric t-distributions. The upper estimates are for the COVID period, and the lower is for the energy period.

We will use the package `rugarch` and the function `fitdist` with the argument `dist='std'` to fit a model of the type $Y \sim t_\nu\left(\mu, \sigma^2\right)$ where ν is the tail index, μ is the location parameter, and σ^2 is the variance. Note that σ^2 is only defined for $\nu > 2$, and μ is the mean for $\nu > 1$, [4]. In table 2, the parameters of the fitted distributions are given with a CLT 95% approximate confidence interval. We see that all of the means are statistically indistinguishable from 0 as 0 is included in all of the confidence intervals. All of the standard deviations are significant and for Shell, we see that the estimated standard deviation is halved from the COVID to the energy crisis period. The tail index of Shell is 2.8328 during COVID which is small and the confidence region goes below 2 which we do not accept by assumption. It is an artifact in the construction of the symmetric CLT interval and we will ignore the issue here. During the energy crisis, Shell has a much larger tail index, 7.4390, but a much larger uncertainty associated with the estimate ± 4.7344. As mentioned, we speculated that we could fit skewed t-distributions as well. To do that, we use `fitdist` with `dist='sstd'`. In all combinations of companies and periods, the skewness parameters are just above 1, however, 1 is included in the confidence region which means we cannot reject that the data is from a symmetric t. Worth mentioning is the skewed t-distribution for Vestas in the energy period where we found AIC $= -2067$ and BIC $= -2050$ hence if we compare with the symmetric-t in table 2, we see that the AIC favors the skewed t. However, the BIC favors the symmetric t and the skewness parameter is insignificant, $\xi = 1.1125 \pm 0.14253$, hence we stick with the symmetric t.

To be able to compare with the multivariate model, we calculate the AIC and BIC by taking the product of the marginal symmetric t-densities, finding the log-likelihood, and accounting for the sum of all of the parameters. For the COVID period, we obtain $\text{AIC}_C = -1989$ and $\text{BIC}_C = -1968$, and for the energy period, we get $\text{AIC}_E = -4609$ and $\text{BIC}_E = -4583$ which we will use to compare with the multivariate distribution.

3.2 Multivariate Distributions

In fig. 2, we see that there is some tail dependency and covariance. Therefore, we will fit multivariate Gaussian distributions and multivariate t-distribution. Let $\mathbf{Y} = [r_{Shell}, r_{Vestas}]^\intercal$ be a vector of the random variables that represents the log returns for each company. Then the multivariate Gaussian is $\mathbf{Y} \sim \mathcal{N}(\mu_N, \mathbf{\Sigma}_N)$ while the t-distribution is of type $\mathbf{Y} \sim t_\nu(\mu_N, \mathbf{\Lambda})$ where the scale matrix, Λ can be written in terms of the covariance matrix by $\mathbf{\Sigma} = \frac{\nu}{\nu-2}\mathbf{\Lambda}$. We refer to [4] for a rigors introduction in which they also propose a multivariate skewed t-distribution. Here each component would have a shape parameter $\alpha = (\alpha_1, \alpha_2)$ that describes the skewness. We use base-R functions to fit the Gaussian, for the symmetric t-dist. we initially used `cov.trob` from `MASS` but switched to use our own likelihood function with `dmt` from `mnormt` to obtain the Hessian. To fit the skewed multivariate t-dist., we use `mst.prelimFit` from the `sn` package. In table 3, we have provided the AIC and BIC for all of the models. We see that in the COVID period, both t-distributions seem to be better than the Gaussian, and the AIC and BIC both favor the symmetric t-distribution. For the energy period, we see that both information criteria suggest that we should choose the symmetric t-dist. for this period. Indeed the AIC and BIC of the multivariate t-dist., are also better than using the product of the univariate t-distributions. Interestingly, the AIC and BIC for the combined univariate distributions are better than the multivariate Gaussian in both periods which again underpins the need for a model to describe the heavy tails. Given the information criteria, we will fit and provide the exact estimates for the symmetric t-distribution for each of the periods below.

	Gaussian	Symm. t	Skewed t
AIC_C	-1955	-2033	-2001
BIC_C	-1934	-2012	-1969
AIC_E	-4555	-4621	-4548
BIC_E	-4530	-4596	-4510

Table 3 – AIC and BIC for all of the multivariate models. The *upper* part is for the COVID period while the *lower* is for the energy period.

We use the technique introduced in the lecture notes and [4] to ensure that $\mathbf{\Lambda}$ is symmetric and positive definitene using a Cholesky factorization. We do that within the likelihood function and use `optim` with `"L-BFGS-B"` to obtain a hessian matrix. Thereby, we can provide the ML estimates along with an indistinguishable CLT 95% confidence interval and calculate AIC and BIC. Below, we use subscript C for COVID and E for the estimates in the energy period.

$$\mu_C = \begin{bmatrix} 0.0006 \pm 0.00385 \\ 0.0035 \pm 0.00326 \end{bmatrix}, \quad \nu_C = 3.8613 \pm 1.3672$$

$$\mathbf{\Lambda}_C = \begin{bmatrix} 0.000722 & 0.000168 \\ 0.000168 & 0.000513 \end{bmatrix} \pm \begin{bmatrix} 1.24 \cdot 10^{-5} & 1.18 \cdot 10^{-5} \\ 1.18 \cdot 10^{-5} & 1.83 \cdot 10^{-5} \end{bmatrix}$$

$$\mu_E = \begin{bmatrix} 0.00120 \pm 0.00161 \\ -0.00166 \pm 0.00254 \end{bmatrix}, \quad \nu_E = 5.94128 \pm 2.11302$$

$$\mathbf{\Lambda}_E = \begin{bmatrix} 2.71 \cdot 10^{-4} & 7.35 \cdot 10^{-5} \\ 7.35 \cdot 10^{-5} & 6.78 \cdot 10^{-4} \end{bmatrix} \pm \begin{bmatrix} 2.00 \cdot 10^{-6} & 3.56 \cdot 10^{-6} \\ 3.56e \cdot 10^{-6} & 1.15e \cdot 10^{-5} \end{bmatrix}$$

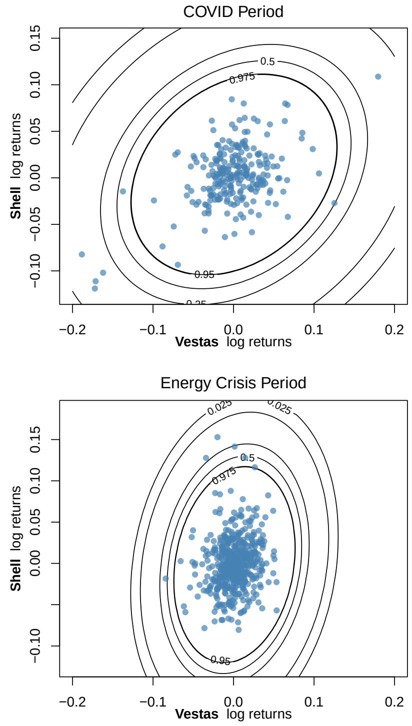

Figure 4 – Contours of the fitted t-distribution for each period.

The location parameter for Shell in the COVID period and for Vestas in the Energy period is statistically indistinguishable from 0. However, the remaining of all parameters are significant which we can determine by using the confidence intervals. The tail index appears much lower in the COVID period than in the energy period which aligns closely with the observations from the univariate case. We can graphically see that in fig. 4 by comparing the contours and the distance between them. In the COVID period, one can barely see the two outer contours; 0.025 and 0.05, while they are easily observable for the energy period. Given that we know that the scale matrices and the tail indices are significant, we use the face values of the ML estimates of λ and ν_λ to find $\mathbf{\Sigma}$ through $\mathbf{\Sigma} = \frac{\nu}{\nu-2}\mathbf{\Lambda}$. With that we find the correlation coefficient in each period, $\rho_C = 0.276$ and $\rho_E = 0.171$. We see that the correlation was relatively low in each period but considerably lower in the energy period. We see that graphically with the contours being more tilted for the COVID period than the energy period in fig. 4.

In essence, the COVID period saw heavier tails, more volatility, and a stronger correlation between Shell and Vestas than the energy period.

4 Conclusion

We investigated how two different companies, Shell and Vestas, in the energy sector reacted in periods of two major events; the COVID crisis and the recent energy crisis. We fitted univariate and multivariate models and found that symmetric multivariate t-distributions could best describe the log returns in both periods based on AIC and BIC information criteria. When we considered the marginal t-distributions, we saw that the tail index and variance for Shell were much larger during COVID than during the energy period while the estimates did not change much for Vestas in the periods. When we fitted the multivariate t-distributions, we saw that during COVID the tail index was lower and the companies had a stronger dependence with ML estimates of the tail index at 3.8613 and a correlation coefficient of 0.276. The energy crisis did not give rise to as extreme events and thus we observed lighter tails with the ML estimate of the tail index at 5.9412 and correlation coefficient at only 0.171. Thus it seems that the stocks of the companies were under more distress during COVID than during the energy crisis.

References

[1] (2021) Green shoots' salad days have come and gone. [Online]. Available: https://www.reuters.com/breakingviews/green-shoots-salad-days-have-come-gone-2021-05-06/

[2] (2021) Global energy crisis. [Online]. Available: https://www.iea.org/topics/global-energy-crisis

[3] L. T. Ha, "Storm after the gloomy days: Influences of covid-19 pandemic on volatility of the energy market," *Resources Policy*, vol. 79, p. 102921, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301420722003658

[4] D. Ruppert, *Statistics and finance: An introduction.* Springer, 2004, vol. 27.