
AIX-MARSEILLE SCHOOL OF ECONOMICS

ECONOMETRICS BIG DATA & STATISTICS

# Nonlinear and multivariate Time Series

**Teacher:** Gilles DUFRENOT **Student**: Hadare IDRISSOU

April 2022

# Subject 1

## Part 1

The objective of this work is to study the correlation between various assets and the market.

We use the CAPM database. The assets in question are Oracle, Microsoft, General Electric and Ford. USTB3M is the risk-free rate and the asset S&P500 is the market index.

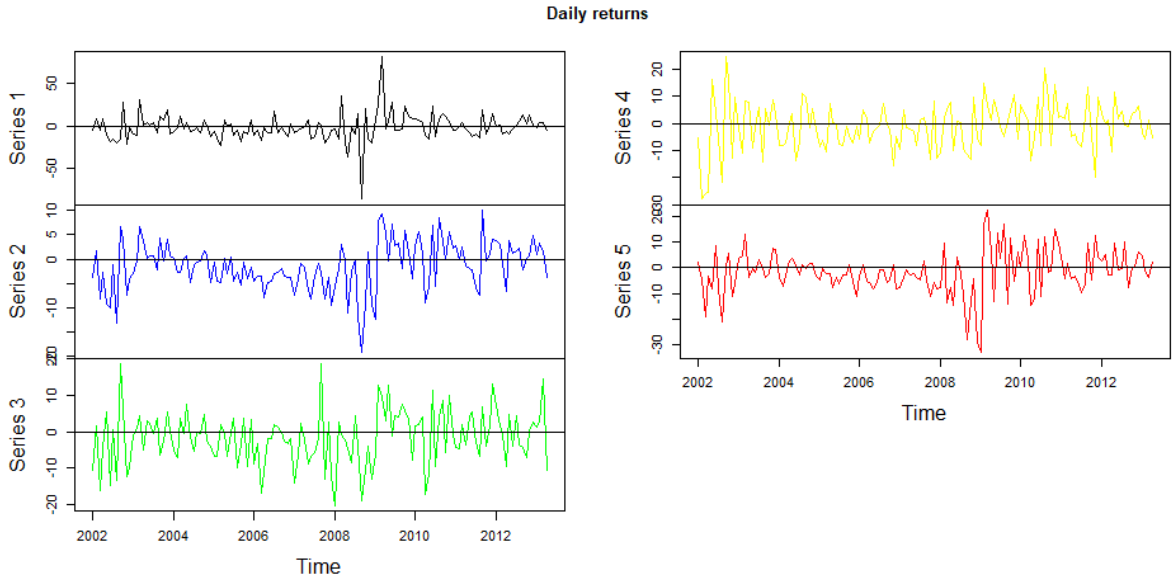
### ❖ Preprocessing of the data

In this subsection, we import the database and calculate the excess returns of the assets and the market. To do this, we differentiate the assets to get them as rates and then remove the risk-free rate from them to get the excess returns.

### ❖ Preliminary analysis

A set of analyses to learn more about the distribution of our financial series and the potential laws they follow.

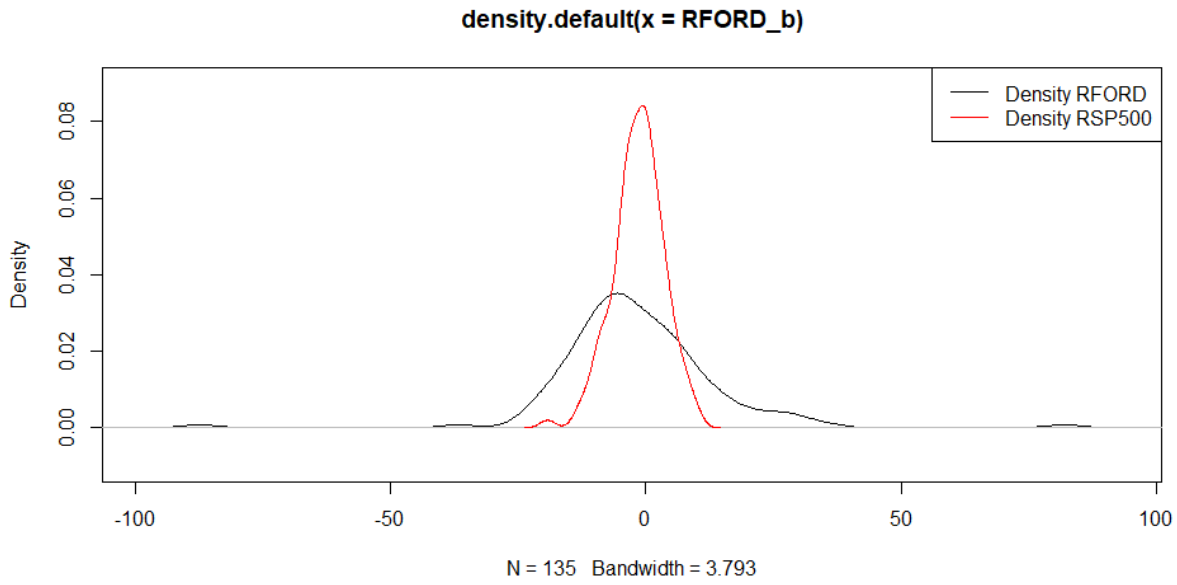
- Daily returns



Here we have plotted the evolution of the returns of our assets over time. We notice that they all oscillate on average around 0 with periods of very high volatility on both sides. The Ford, S&P500, Microsoft and General Electric series experienced a period of high volatility in 2008, which may be due to the financial crisis of that year. Only the Oracle series experienced instability in the early 2002s.

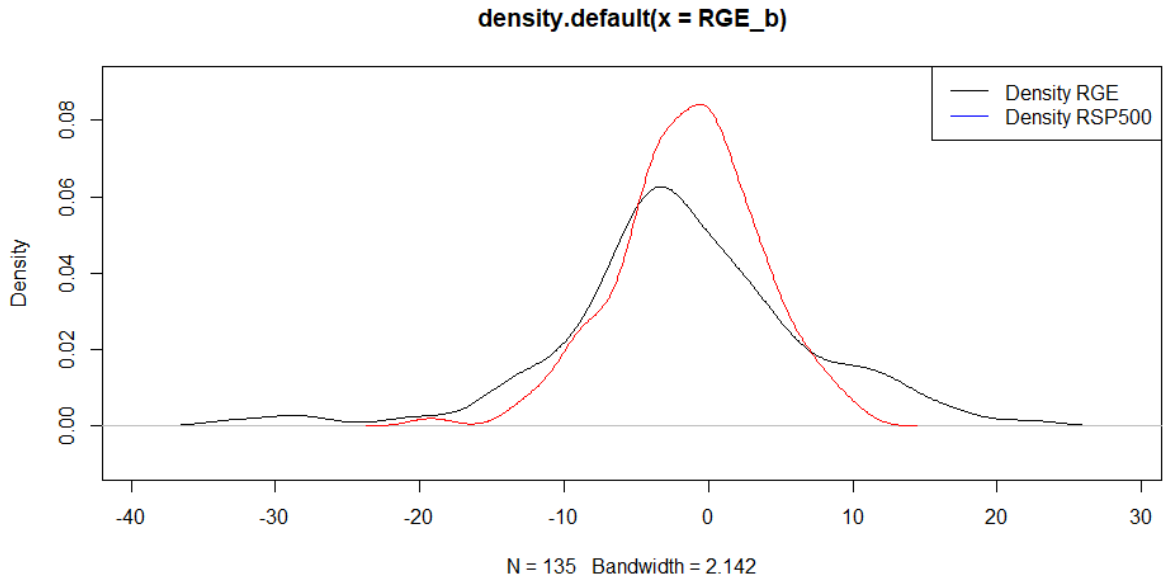
- Densities

➢ FORD



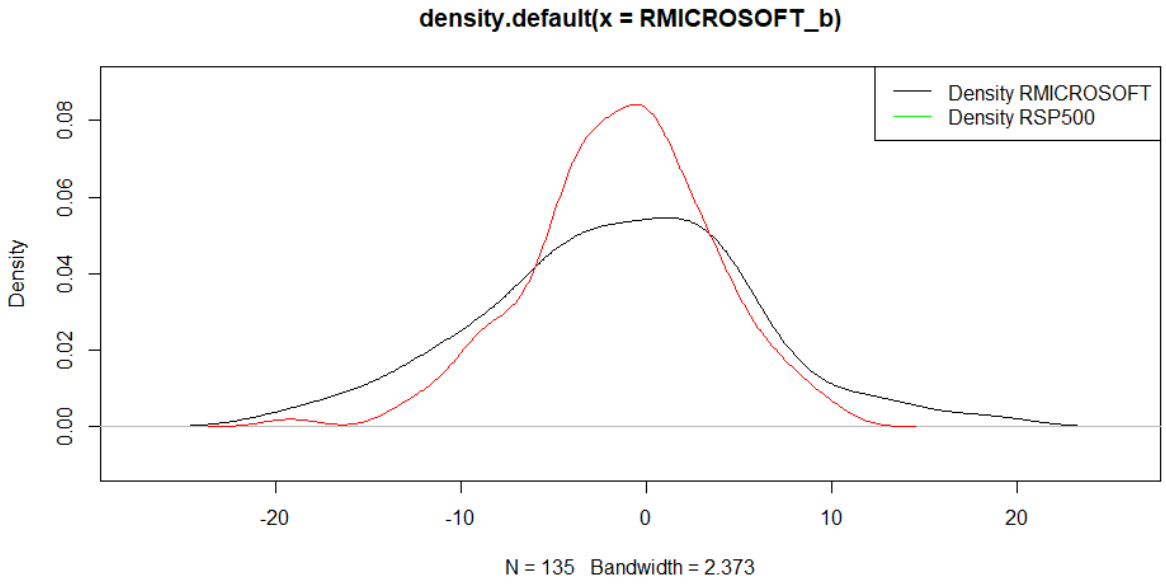

We notice that the market density is less flattened than that of the FORD asset but slightly more symmetrical. Enough observations below zero and many observations at the level of the tails of the two series thus presence of heavy tails

➢ General Electric



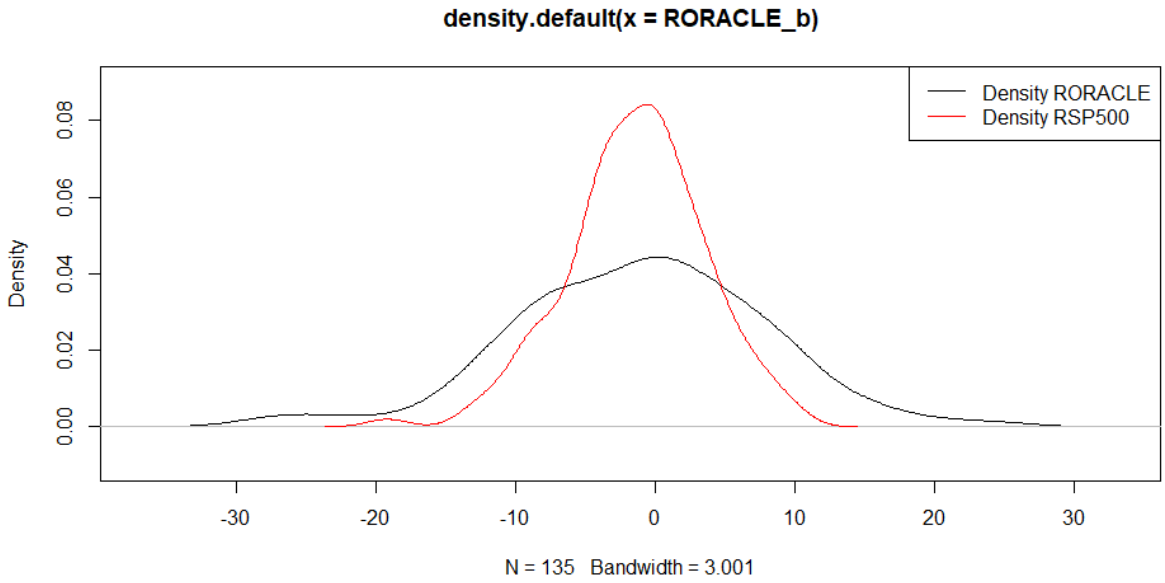

Here, the two series are almost identical with also a presence of heavy tails.

- MICROSOFT




The density of the MICROSOFT series is similar to a Gumbel distribution with also heavy tails.

- ORACLE



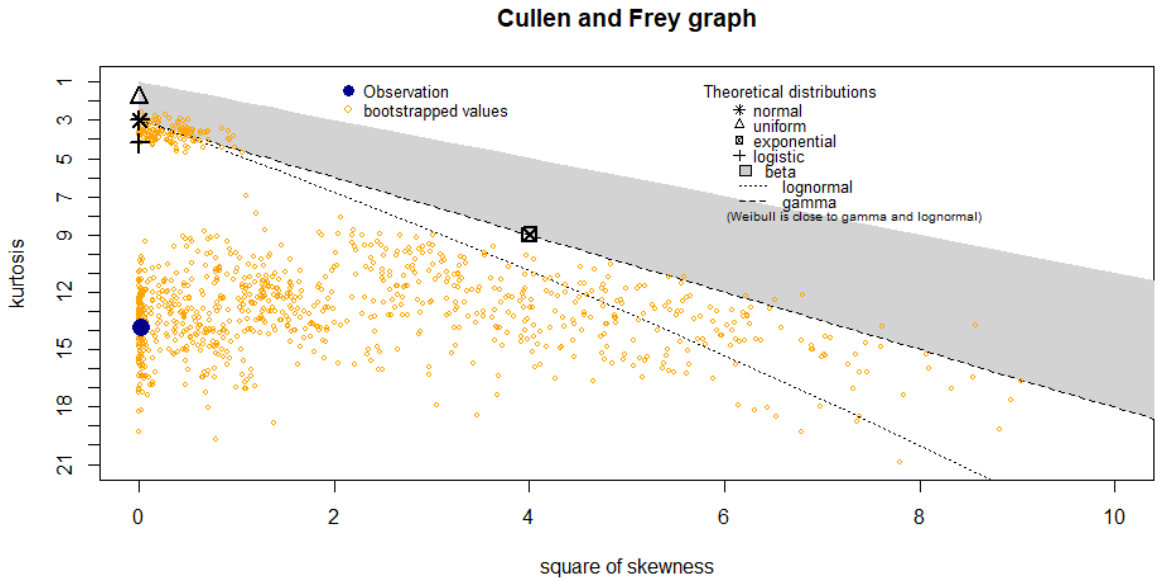

Similar to the previous distribution, the tails are a little less heavy for the ORACLE series distribution

- Distributions

In this sub-section, we try by looking at skewness and kutiosis to see which probability distribution our series are most similar to

- FORD

```
summary statistics
------
min:  -87.25039   max:  82.00396
median:  -3.103151
mean:  -1.586154
estimated sd:  15.8674
estimated skewness:  0.149895
estimated kurtosis:  13.83972
```

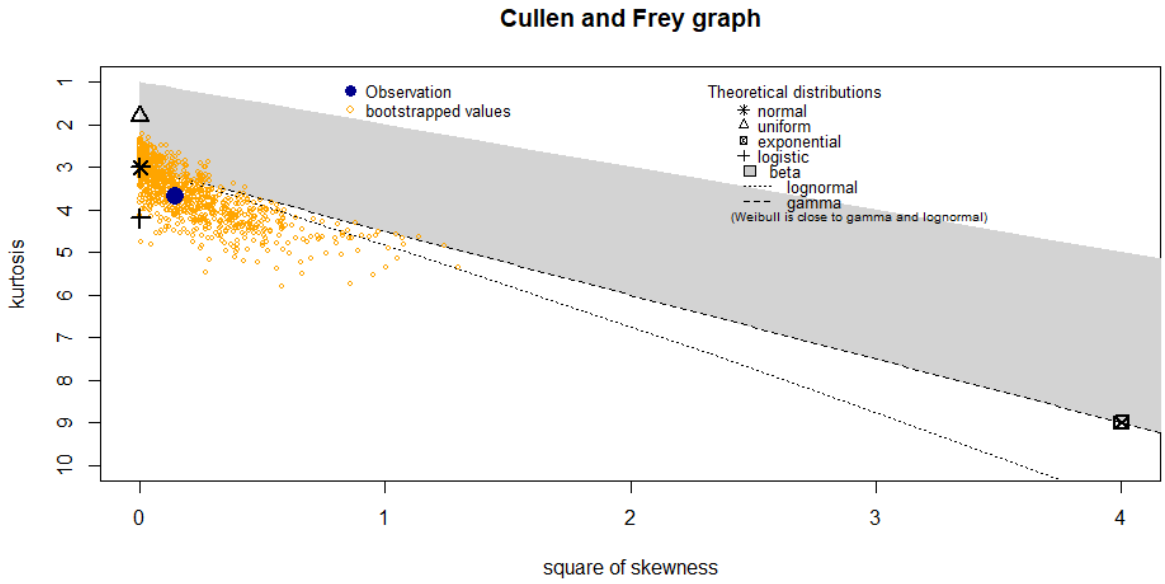



The blue dot represents the skewness and kurtiosis of our Ford series, with an estimated skewness of 0.14 and an estimated kurtiosis of 13.83, the distribution is quite far from all the others but not so far from the one represented by a cross, thus from the logistic distribution

- S&P500

```
------
min:  -19.23397   max:  10.04554
median:  -0.9157733
mean:  -1.388684
estimated sd:  4.938581
estimated skewness:  -0.3821603
estimated kurtosis:  3.686569
```

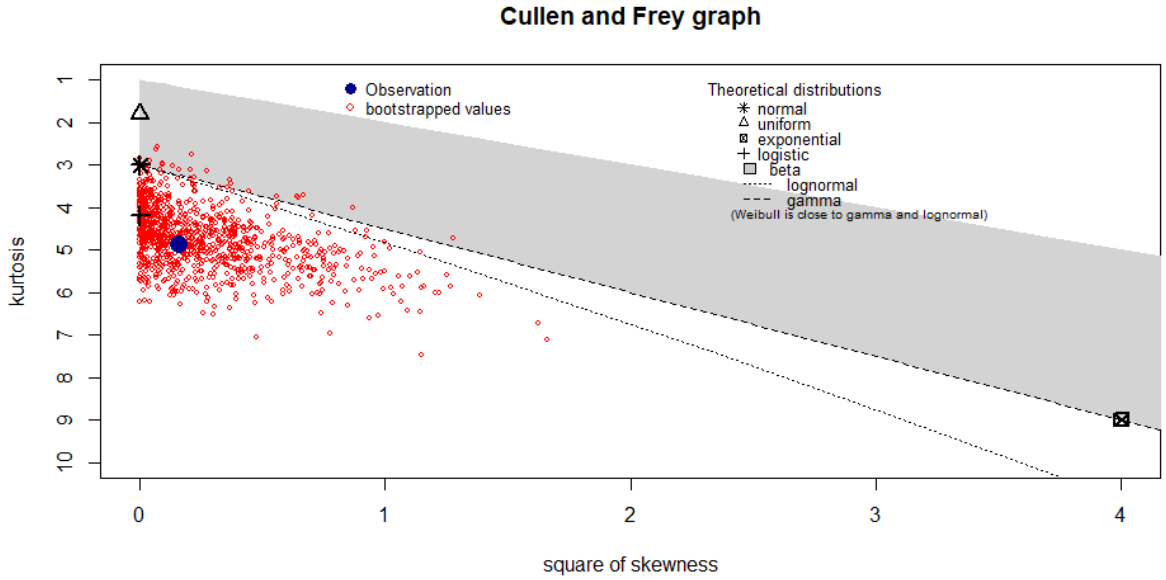
Notr




The distribution is between a normal distribution and a logistic distribution

- General Electric

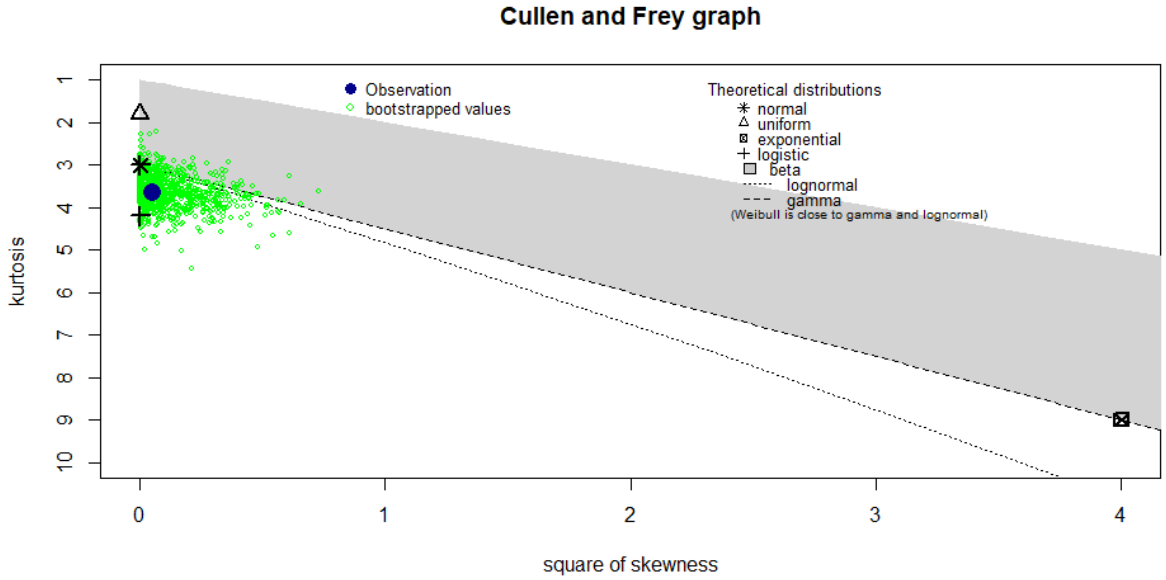
```
------
min:   -32.82703   max:   22.27357
median:   -2.451675
mean:   -1.750484
estimated sd:   8.581131
estimated skewness:   -0.4009055
estimated kurtosis:   4.881196
```




Here, we would conclude on a logistic law

- ORACLE

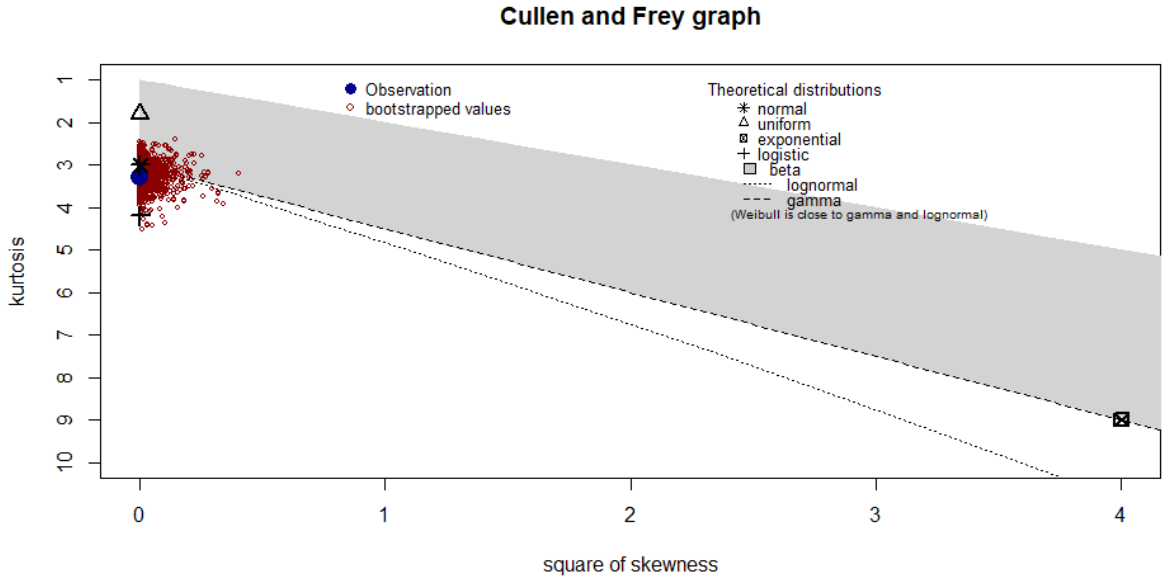
```
------
min:  -27.93159   max:  24.41257
median:  -0.9102991
mean:  -1.142808
estimated sd:  8.961123
estimated skewness:  -0.225446
estimated kurtosis:  3.658715
```




Between a normal and logistic law also

- MICROSOFT

```
------
min:  -20.20953   max:  18.49192
median:  -0.9737291
mean:  -1.503149
estimated sd:  7.266745
estimated skewness:  0.001495808
estimated kurtosis:  3.281596
```

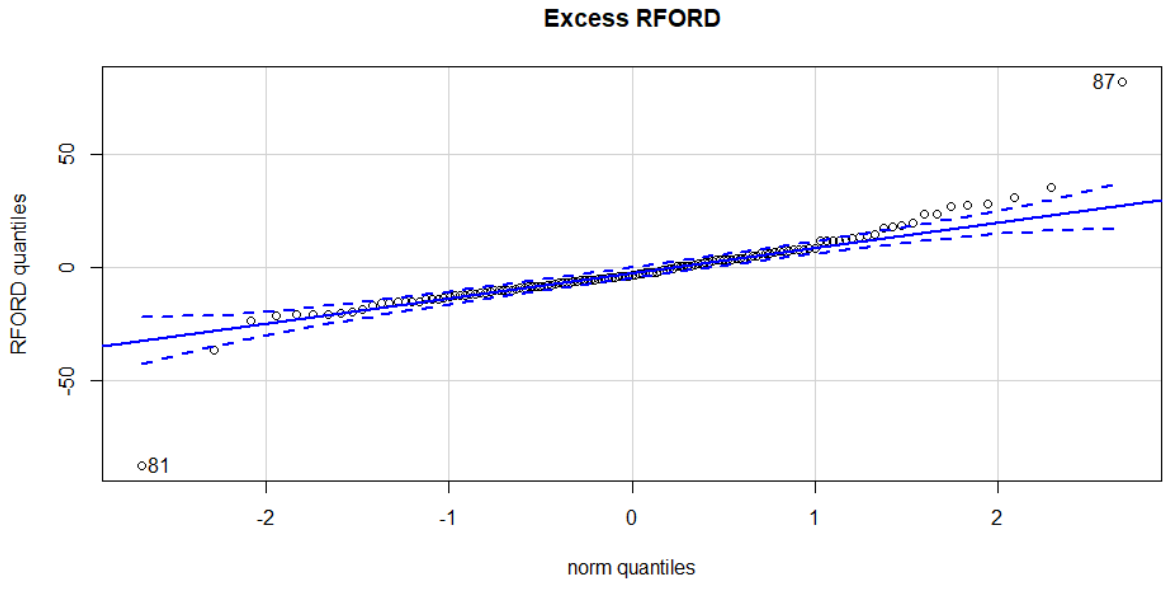



Here, we would conclude on a normal distribution

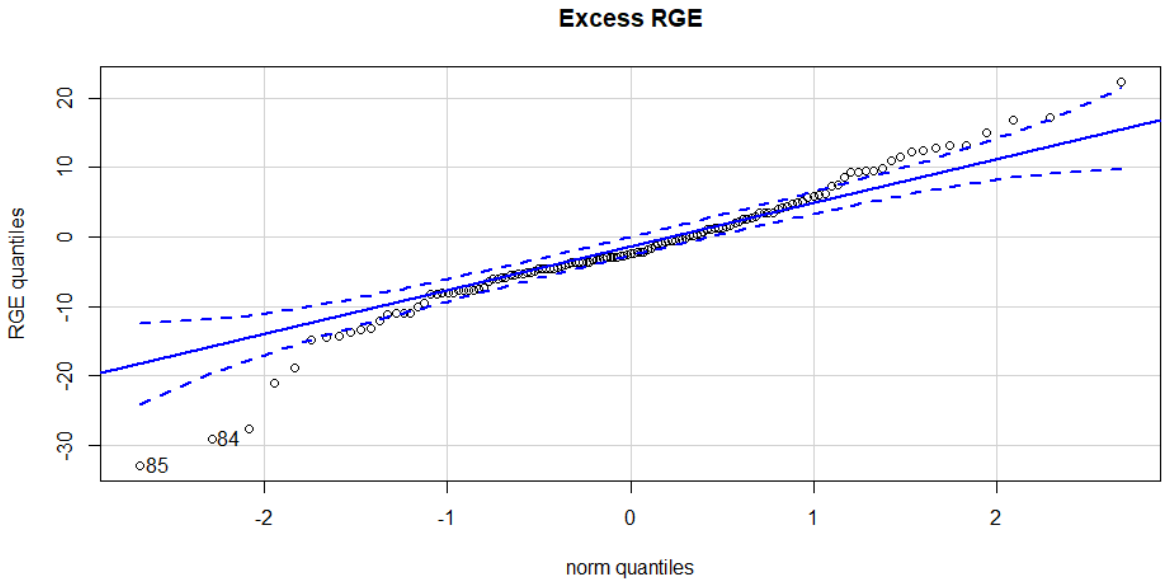
- Univariate QQ-PLOTS

The objective of the qq-plot is to check if the series has a normal distribution, if so, the points are close to the straight line and are at least in the confidence interval delimited by the dotted lines

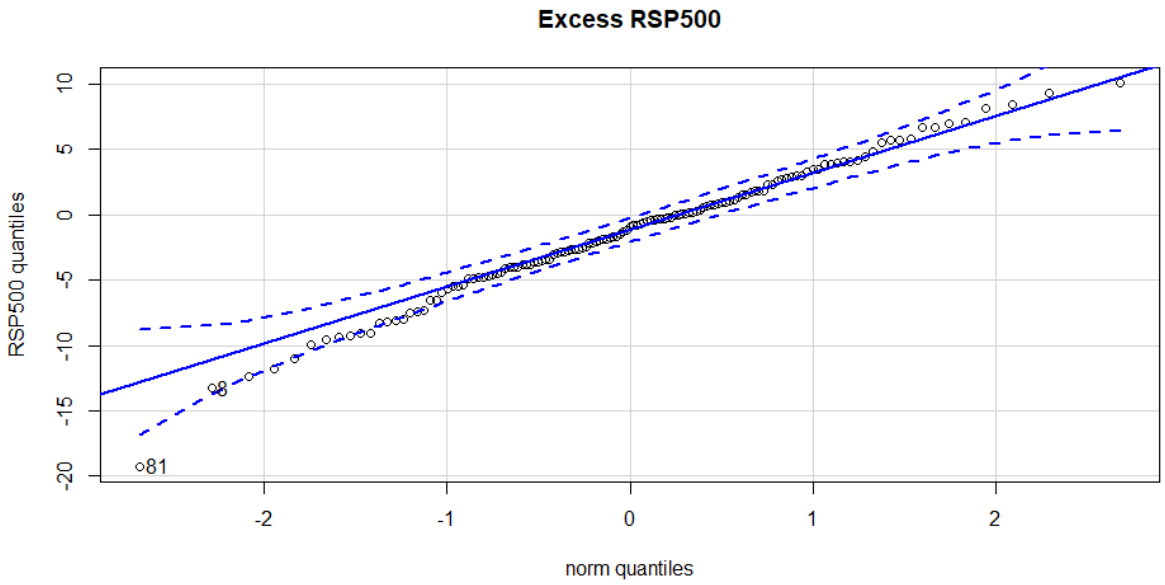
- FORD



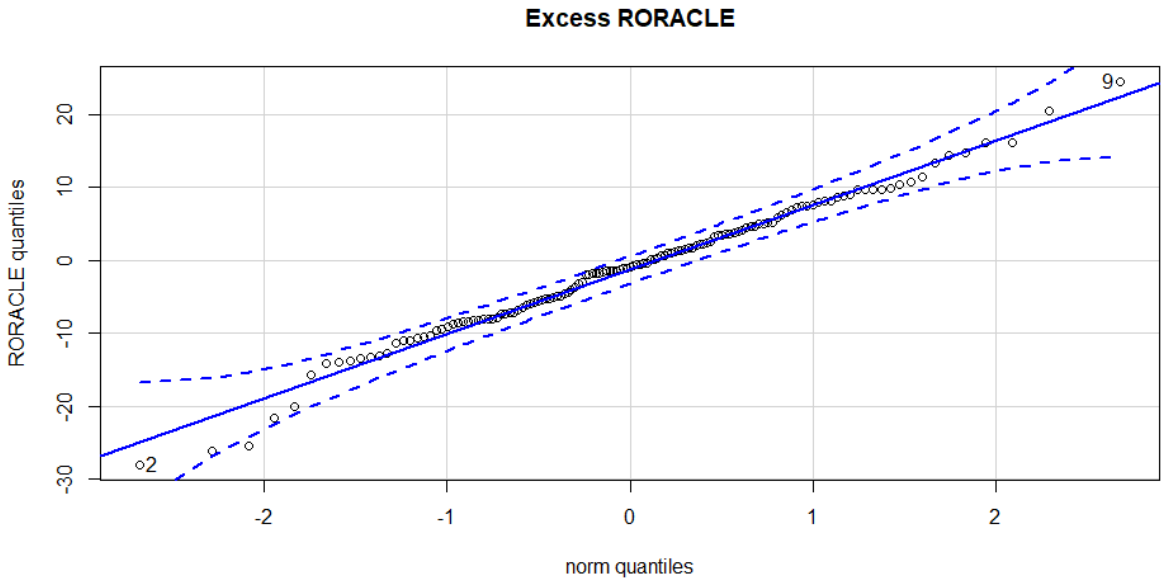
- General Electric



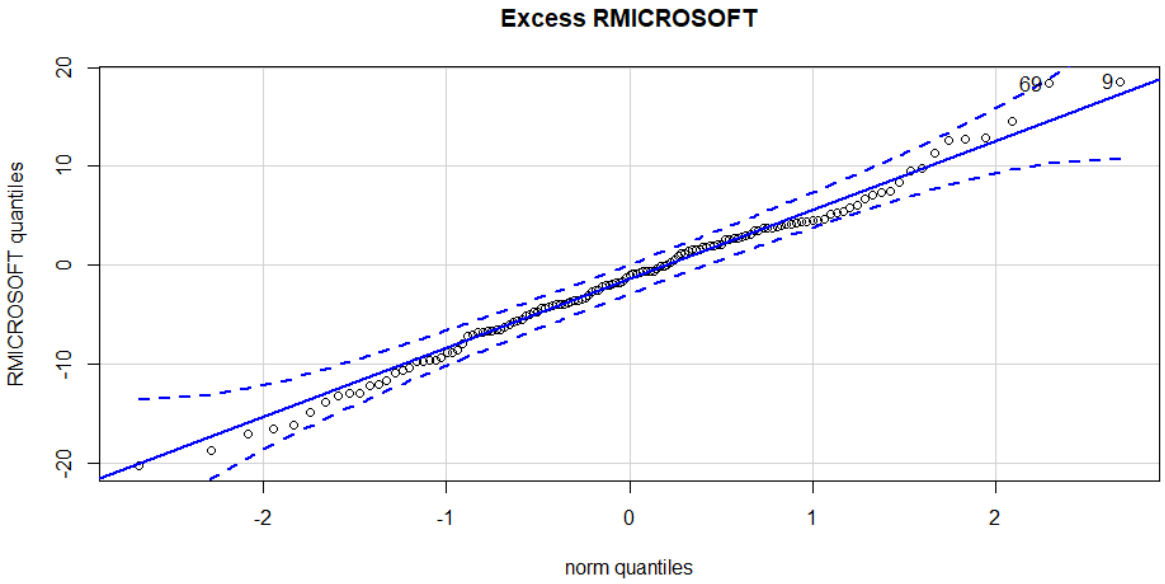

- S&P500




- ORACLE




- MICROSOFT

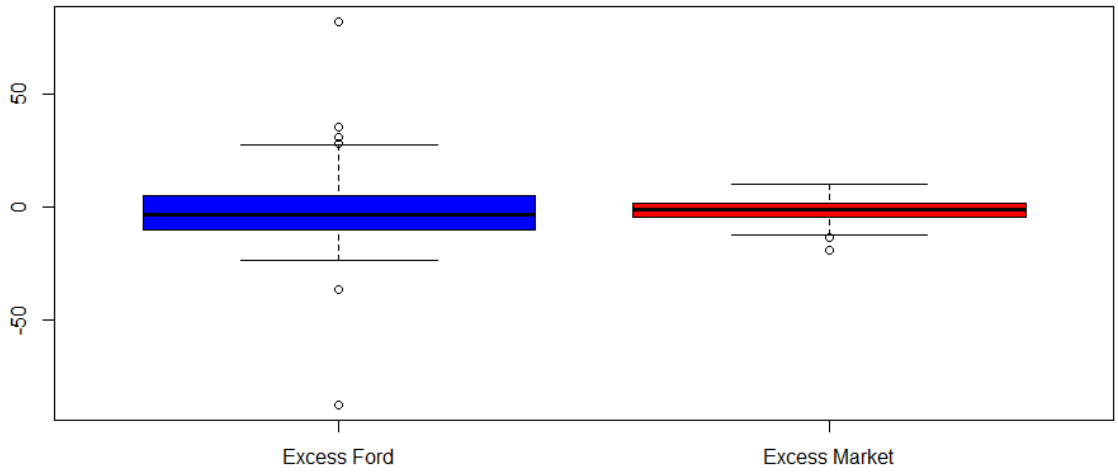



Overall, the series seems to follow a Gaussian distribution except for some outliers

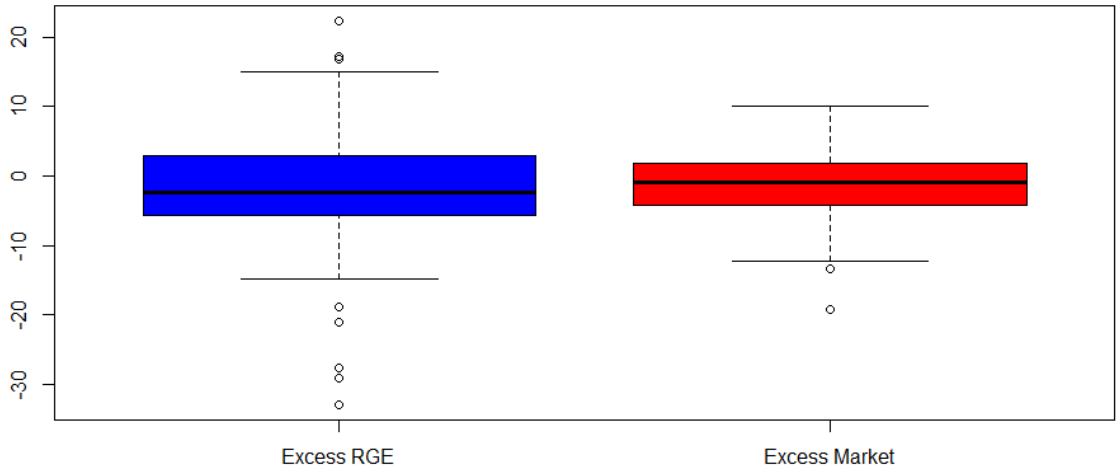
- Boxplot

We compare the moustache boxes of the series with the market index

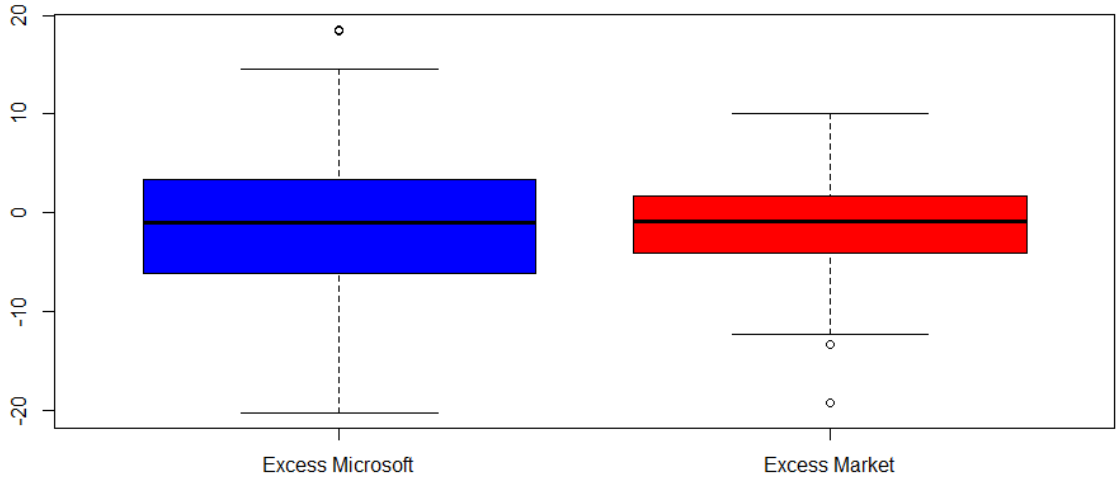
- FORD



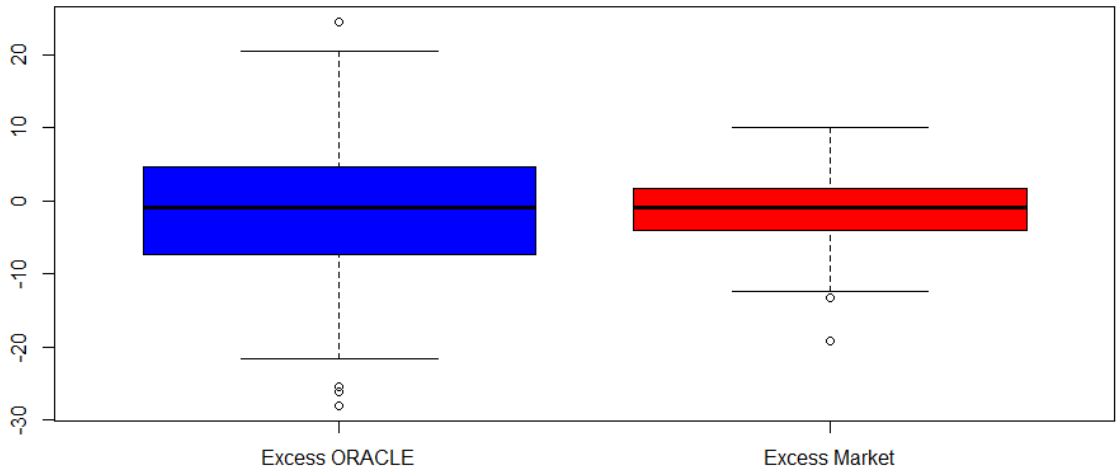

- General Electric




- Microsoft

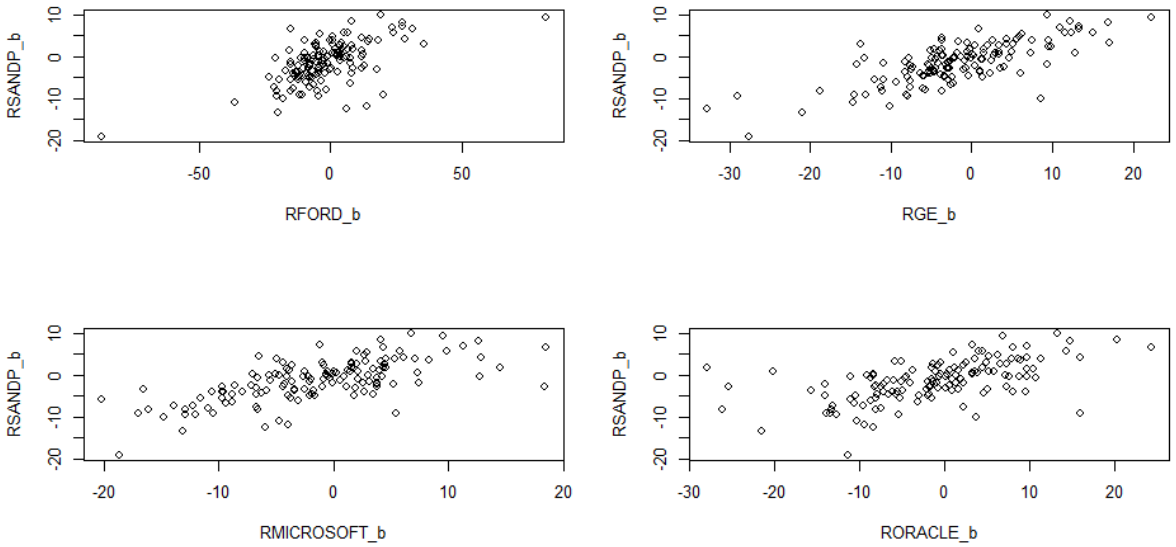



- ORACLE



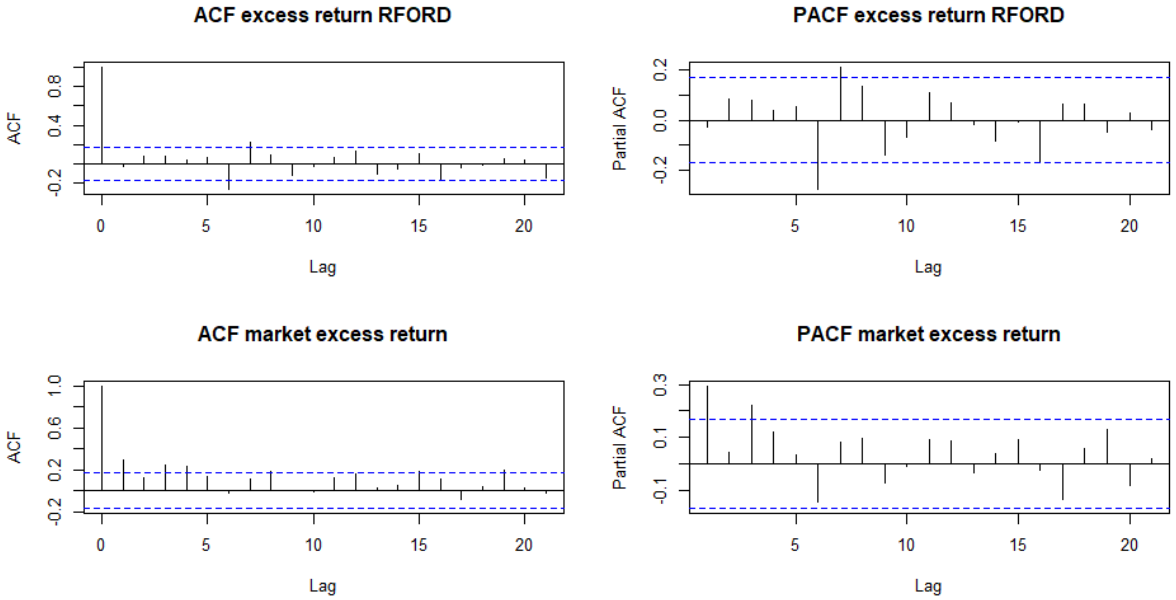

There are some outliers for each of the assets and fairly large quartiles except for the FORD asset, the averages are almost the same as the market average

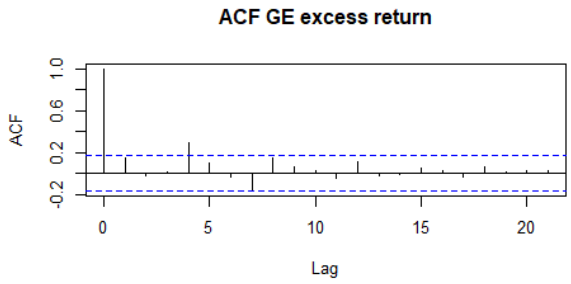
- Scatter Plot

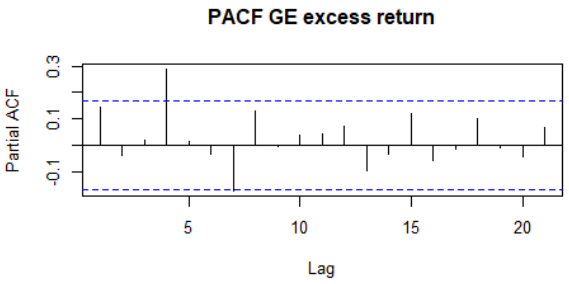


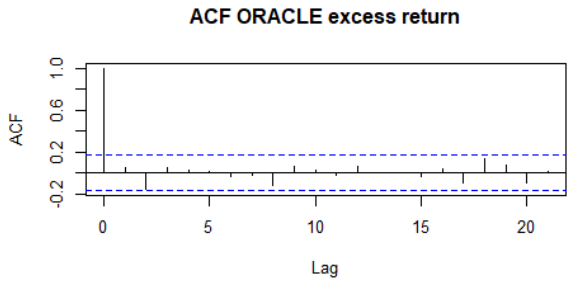

The relationship between the series and the market seems to be linear and increasing according to the previous figures

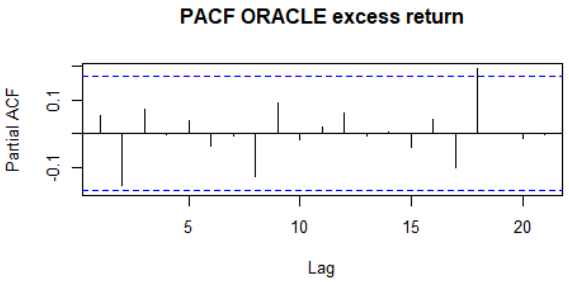
- Autocorrelation

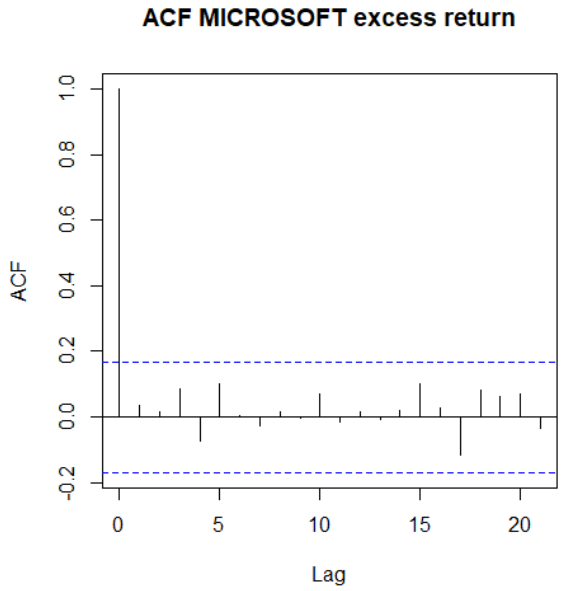


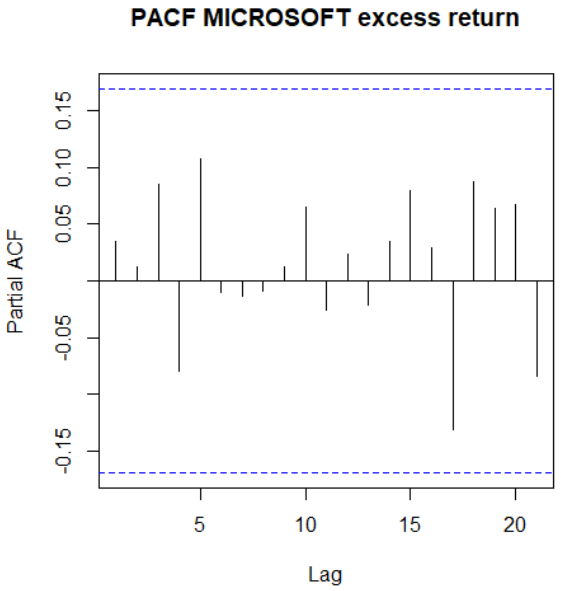



The ACF measures the autocorrelation between a series and itself lagged by various periods while the PACF does the same but removing the effect of intervening variables or periods. Significant periods are those where the line extends beyond the dotted line.

- ❖ Elliptical copulas

- FORD

```
          [,1]      [,2]
[1,] 1.0000000 0.5352405
[2,] 0.5352405 1.0000000
```

- GE

```
          [,1]      [,2]
[1,] 1.0000000 0.6199834
[2,] 0.6199834 1.0000000
```

- ORACLE

```
          [,1]      [,2]
[1,] 1.0000000 0.6130768
[2,] 0.6130768 1.0000000
```

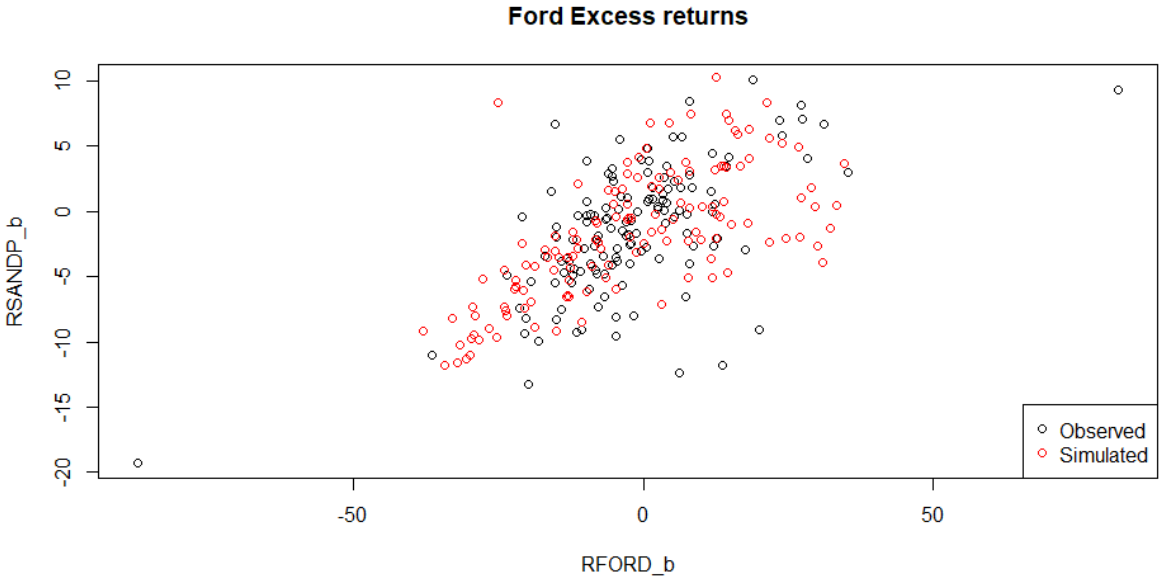
- MICROSOFT

```
          [,1]      [,2]
[1,] 1.0000000 0.6482685
[2,] 0.6482685 1.0000000
```

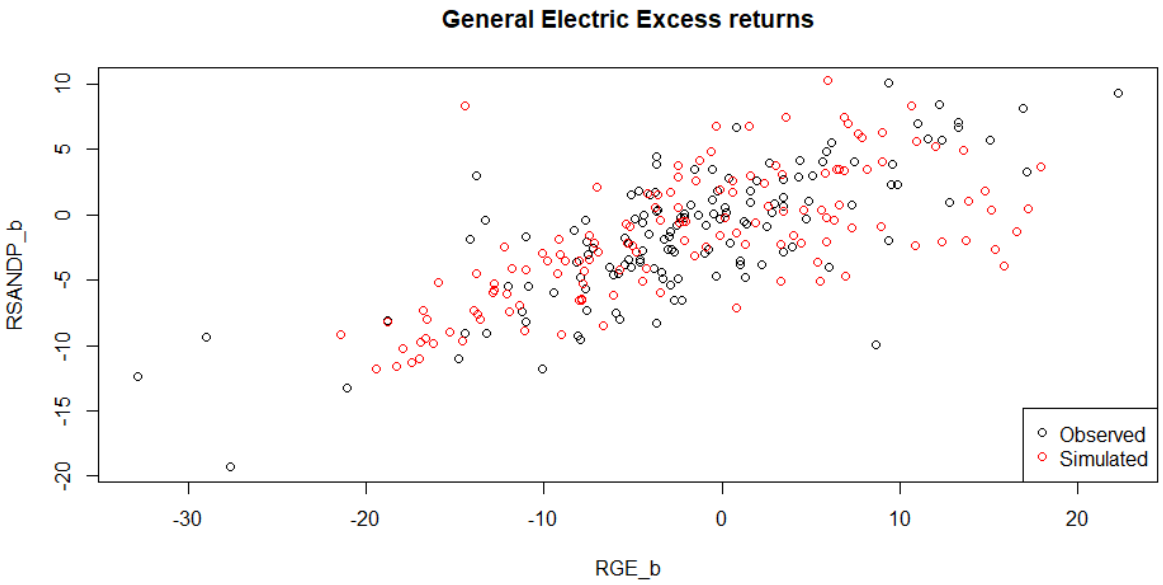
- ARCHIMEDIAN COPULAS (CLAYTON)

Copulas are used as objects to capture the correlation between several variables

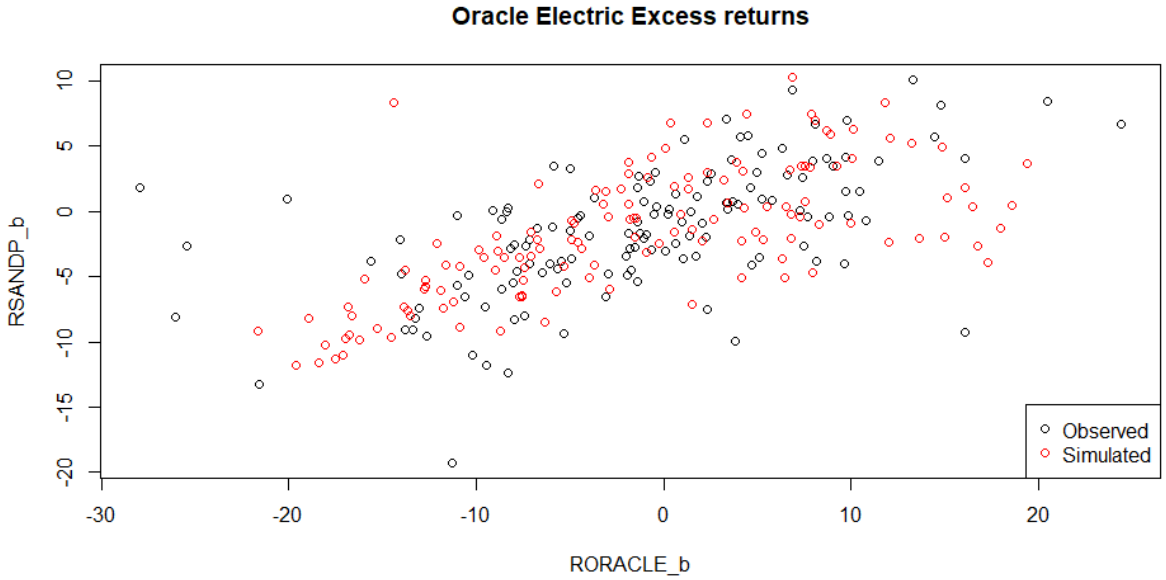
- FORD


Ford Excess returns
RSANDP_b
RFORD_b
Observed
Simulated

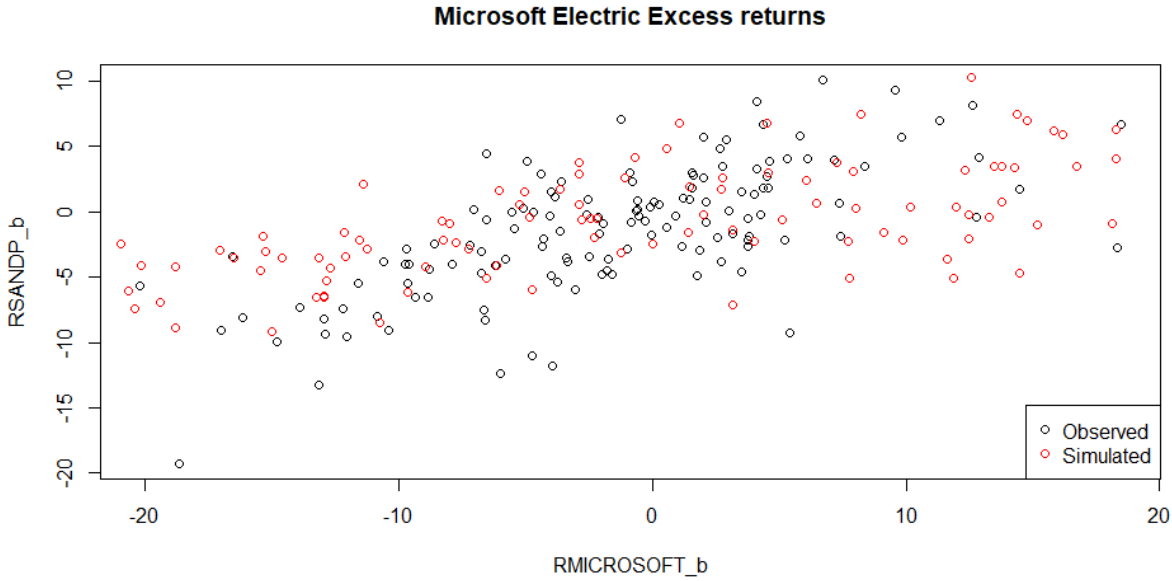
- GE


General Electric Excess returns
RSANDP_b
RGE_b
Observed
Simulated

- ORACLE




- MICROSOFT

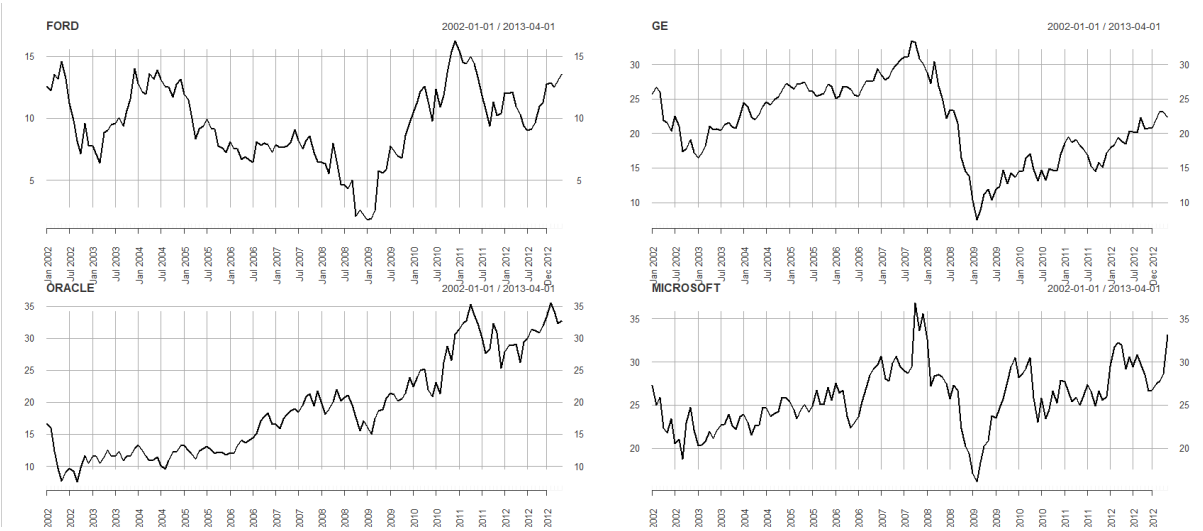



The copulas are represented by the red dots and define quite well the probably linear relationships between the series and the market.

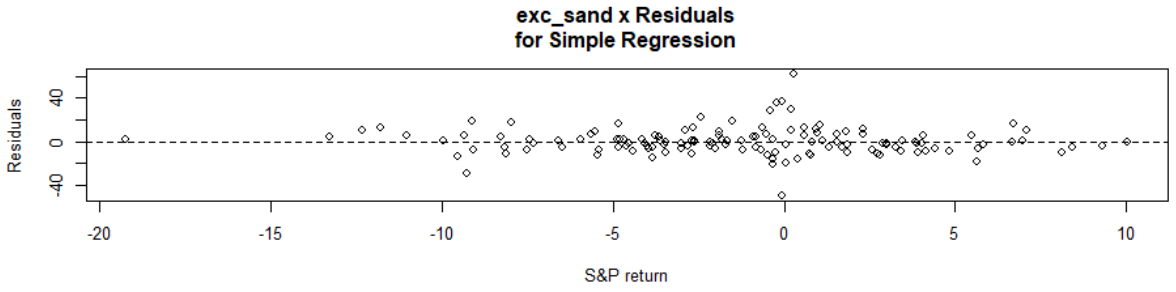
Conclusion: Changes in our assets are reflected in the same direction in the market but with different intensities.

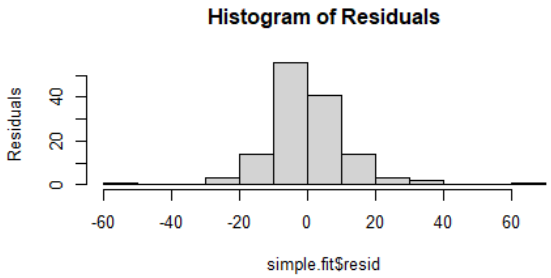
Part 2

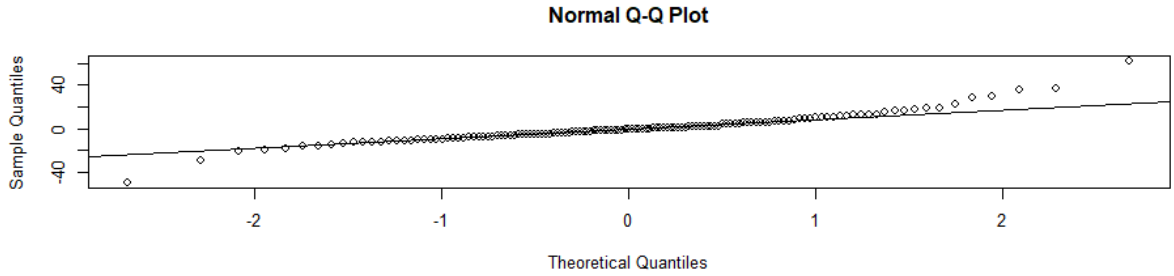
- Evolution graphs

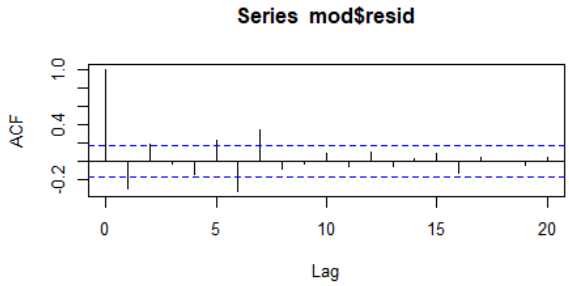



- FORD






Normality test of errors

```
Title:
 Jarque - Bera Normalality Test

Test Results:
  STATISTIC:
    X-squared: 215.6021
  P VALUE:
    Asymptotic p Value: < 2.2e-16

Description:
 Fri Apr 01 11:11:44 2022 by user: hadar
```

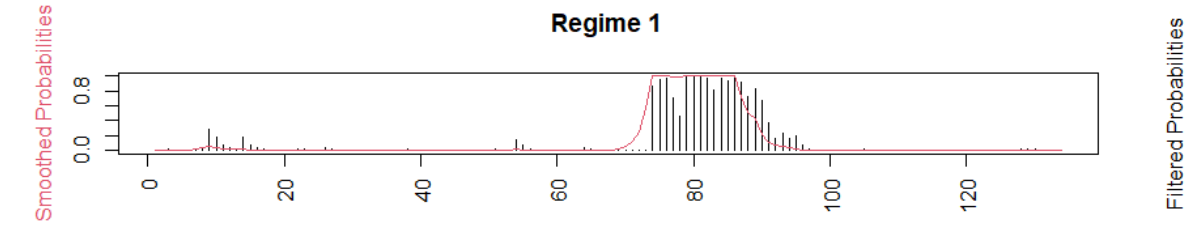
```
Call:
ar(x = simple.fit$resid, aic = TRUE, order.max = 4, method = "ols")

Coefficients:
      1       2       3       4
-0.2755  0.1404  0.0120 -0.1801

Intercept: -0.221 (1.021)

Order selected 4  sigma^2 estimated as  136.6
```

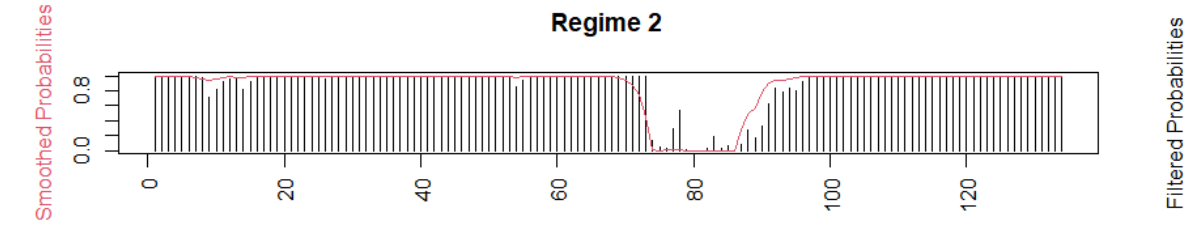
- Markov-Switching model

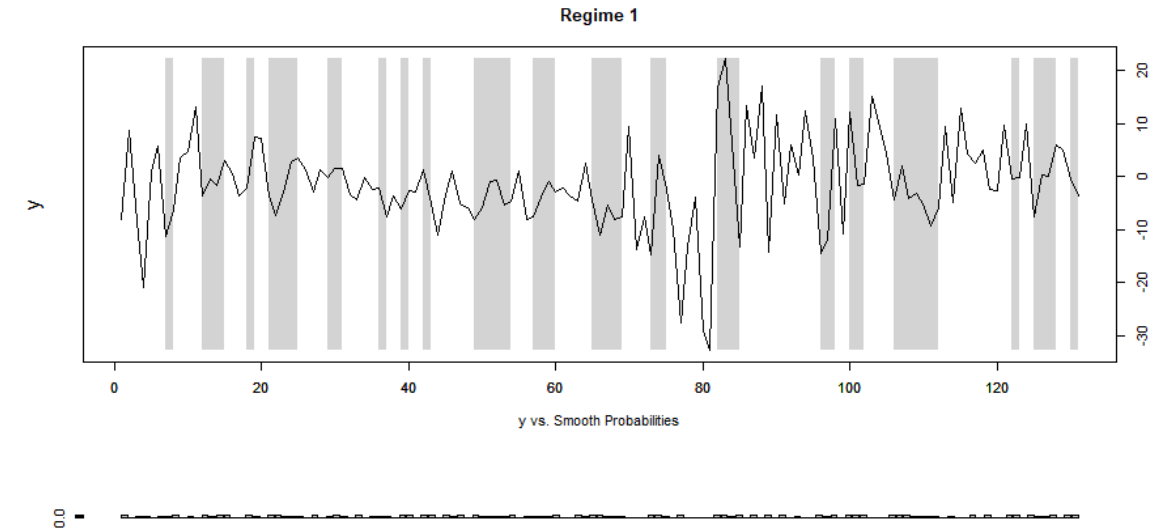
```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(TRUE, TRUE, TRUE, TRUE, TRUE,
    TRUE, TRUE), p = 4, control = list(parallel = FALSE))

       AIC      BIC    logLik
  978.7577 1071.762 -477.3788

Coefficients:

Regime 1
---------
               Estimate Std. Error t value  Pr(>|t|)
(Intercept)(S)  -0.5284     3.6784 -0.1436    0.8858
x(S)             5.7907     0.3704 15.6336 < 2.2e-16 ***
y_1(S)          -0.6849     0.0797 -8.5935 < 2.2e-16 ***
y_2(S)           0.0224     0.0978  0.2290    0.8189
y_3(S)          -1.1355     0.1714 -6.6249 3.475e-11 ***
y_4(S)          -1.0353     0.1226 -8.4445 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.25786
Multiple R-squared: 0.9523

Standardized Residuals:
          Min            Q1           Med            Q3           Max
-1.196195e+01 -4.952315e-01  1.011144e-04  5.039269e-01  1.340590e+01

Regime 2
---------
               Estimate Std. Error t value Pr(>|t|)
(Intercept)(S)  -0.1713     0.8293 -0.2066   0.8363
x(S)             1.6226     0.1892  8.5761   <2e-16 ***
y_1(S)          -0.0421     0.0722 -0.5831   0.5598
y_2(S)          -0.0051     0.0666 -0.0766   0.9389
y_3(S)           0.0499     0.0514  0.9708   0.3316
y_4(S)          -0.0201     0.0614 -0.3274   0.7434
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.313167
Multiple R-squared: 0.4303

Standardized Residuals:
        Min          Q1         Med          Q3         Max
-26.1537573  -5.2362118  -0.3786235   4.0589118  21.4927421

Transition probabilities:
          Regime 1   Regime 2
Regime 1 0.7982842 0.02273582
Regime 2 0.2017158 0.97726418
```

The probability of switching from regime 1 to regime 2 is 2.27%, that of switching from regime 2 to regime 1 is 20%. The probability of remaining in regime 1 is 79.8% and the probability of remaining in regime 2 is 97.7%.

- Probabilities

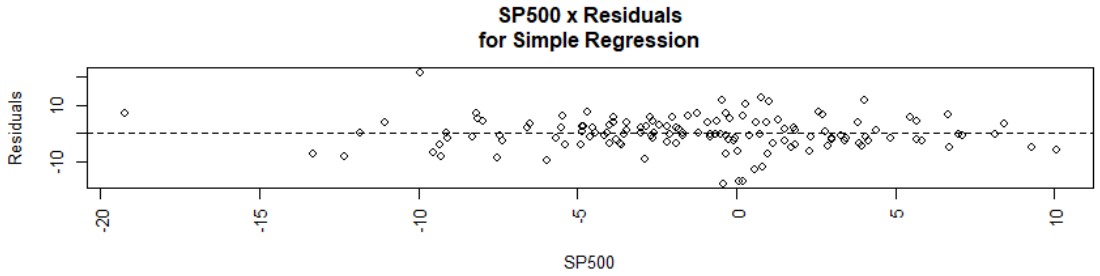


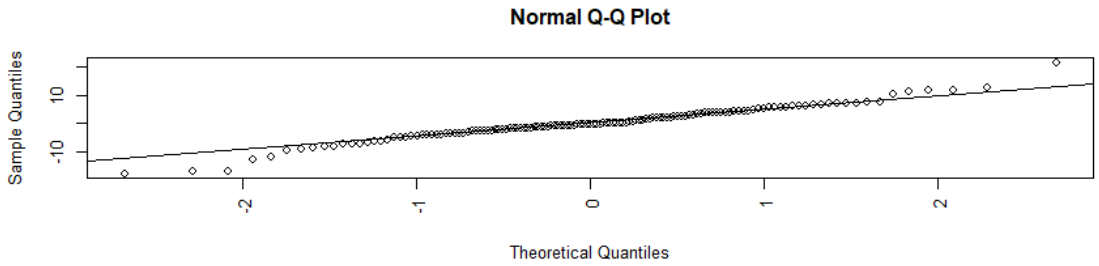



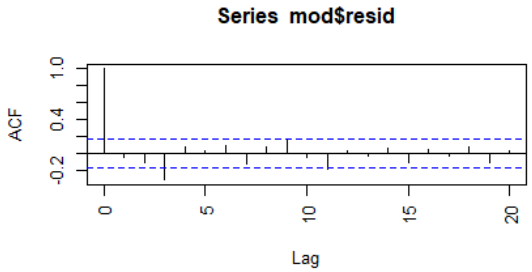

- Comments: A linear model was estimated and the residuals recovered. By testing the normality of the residuals, we noticed that they did not follow a normal distribution. The transition probabilities from one regime to the other are given in the model results and the graph represents the two regimes.

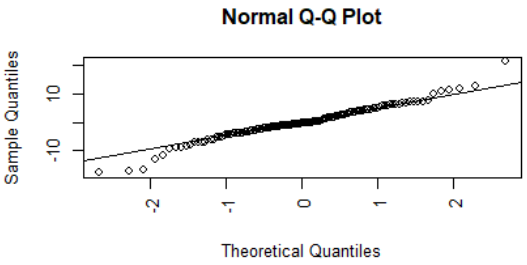
- GE

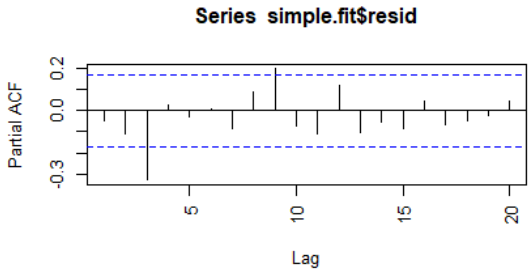

SP500 x Residuals
for Simple Regression
Residuals
10
-10
-20
-15
-10
-5
0
5
10
SP500


Histogram of Residuals
Residuals
0
20
40
-20
-10
0
10
20
mod$resid


Normal Q-Q Plot
Sample Quantiles
10
-10
-2
-1
0
1
2
Theoretical Quantiles


Series mod$resid
ACF
1.0
0.4
-0.2
0
5
10
15
20
Lag

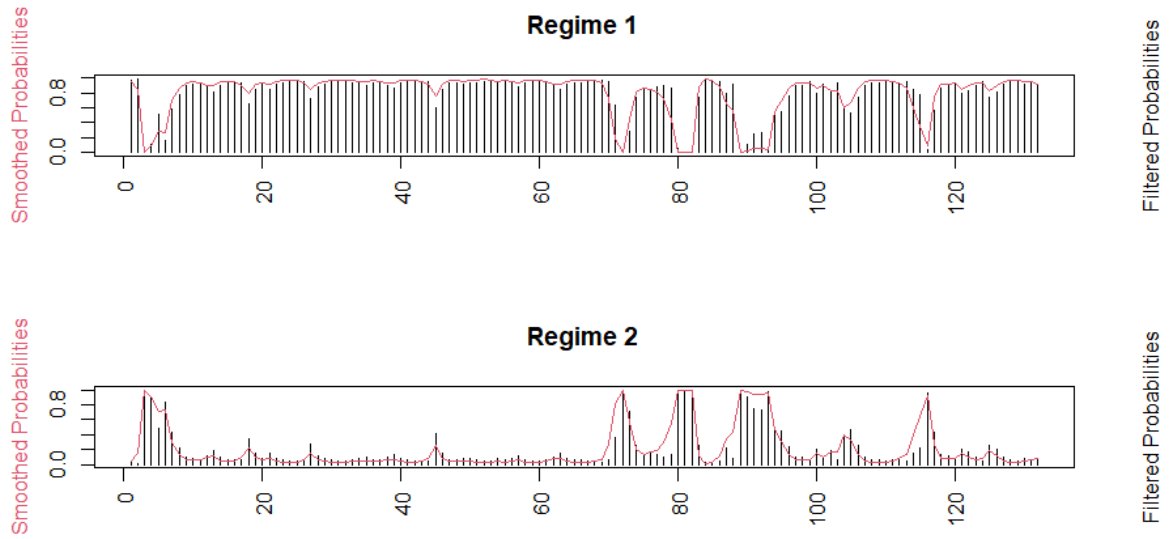

Normal Q-Q Plot
Sample Quantiles
10
-10
-2
-1
0
1
2
Theoretical Quantiles

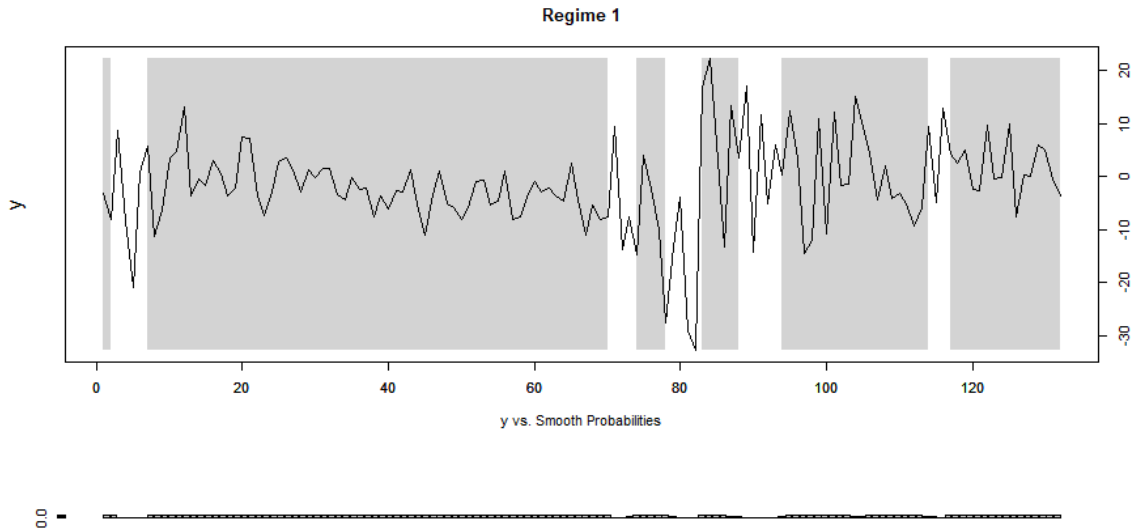




```
Call:
ar(x = simple.fit$resid, aic = TRUE)

Coefficients:
      1         2         3
-0.0871   -0.1255   -0.3272

Order selected 3  sigma^2 estimated as  29.41
```

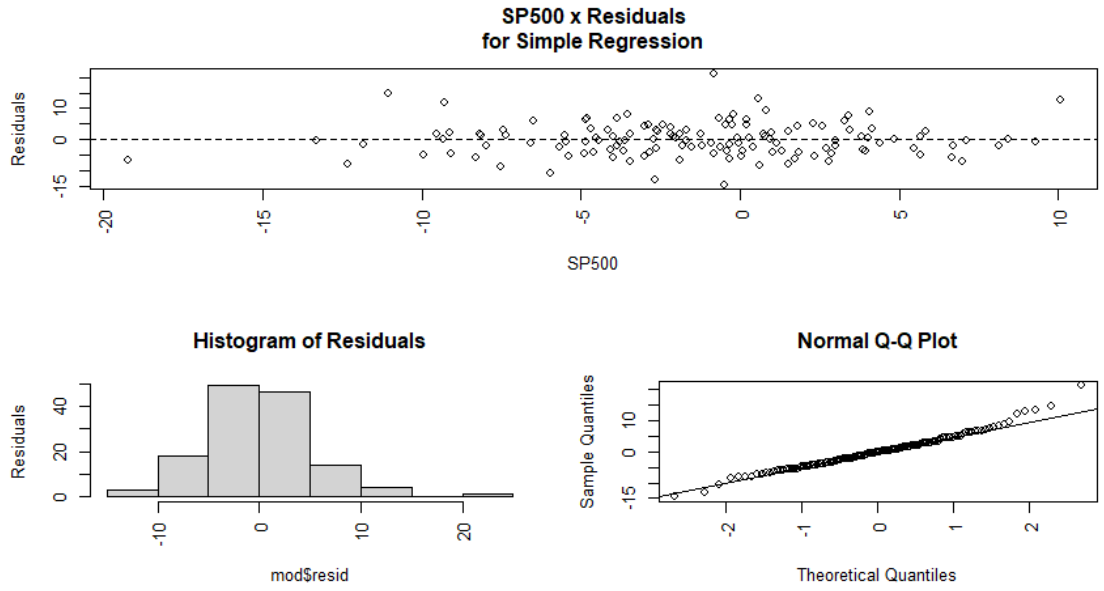
- Markov Switching Model

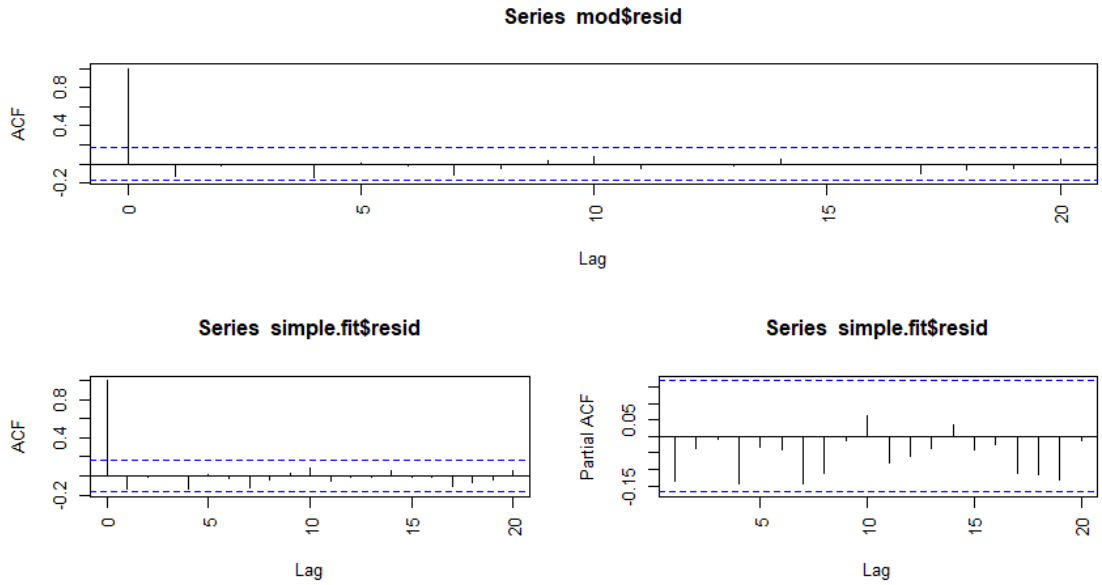
```
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(TRUE, TRUE, TRUE, TRUE, TRUE,
    TRUE), p = 3, control = list(parallel = FALSE))

       AIC      BIC    logLik
  816.8074 894.4635 -398.4037

Coefficients:

Regime 1
---------
               Estimate Std. Error t value  Pr(>|t|)
(Intercept)(S)  -0.1296     0.4658 -0.2782  0.780859
x(S)             1.2849     0.0983 13.0712 < 2.2e-16 ***
y_1(S)          -0.0557     0.0671 -0.8301  0.406482
y_2(S)          -0.1592     0.0569 -2.7979  0.005144 **
y_3(S)          -0.1023     0.0560 -1.8268  0.067730 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.649928
Multiple R-squared:  0.73

Standardized Residuals:
       Min         Q1        Med         Q3        Max
-7.8971169 -2.0744438 -0.3982204  2.2706641  7.2861852

Regime 2
---------
               Estimate Std. Error t value  Pr(>|t|)
(Intercept)(S)  -0.5119     2.1447 -0.2387    0.8113
x(S)             1.4950     0.3753  3.9835 6.791e-05 ***
y_1(S)          -0.0567     0.1752 -0.3236    0.7462
y_2(S)           0.3083     0.2046  1.5068    0.1319
y_3(S)          -0.1306     0.1994 -0.6550    0.5125
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.845752
Multiple R-squared: 0.5115

Standardized Residuals:
        Min          Q1         Med          Q3         Max
-16.0103968  -0.6972950   0.3554669   1.3383775  22.0176642

Transition probabilities:
           Regime 1  Regime 2
Regime 1 0.91748397 0.3233573
Regime 2 0.08251603 0.6766427
```



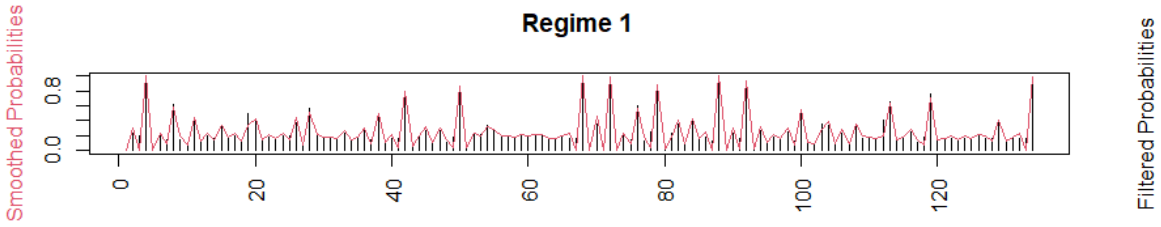

Regime 1
y
y vs. Smooth Probabilities
0.0

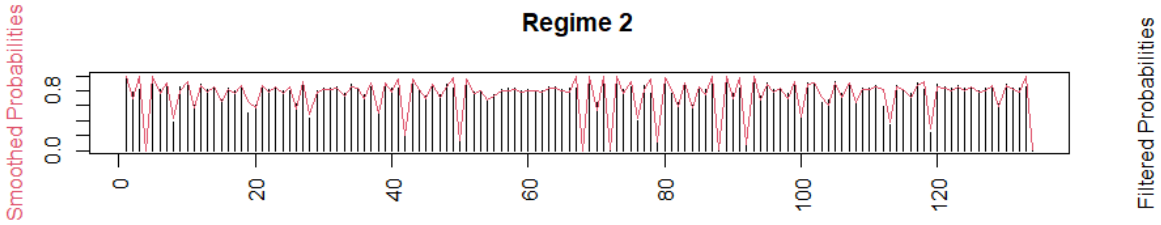
- MICROSOFT

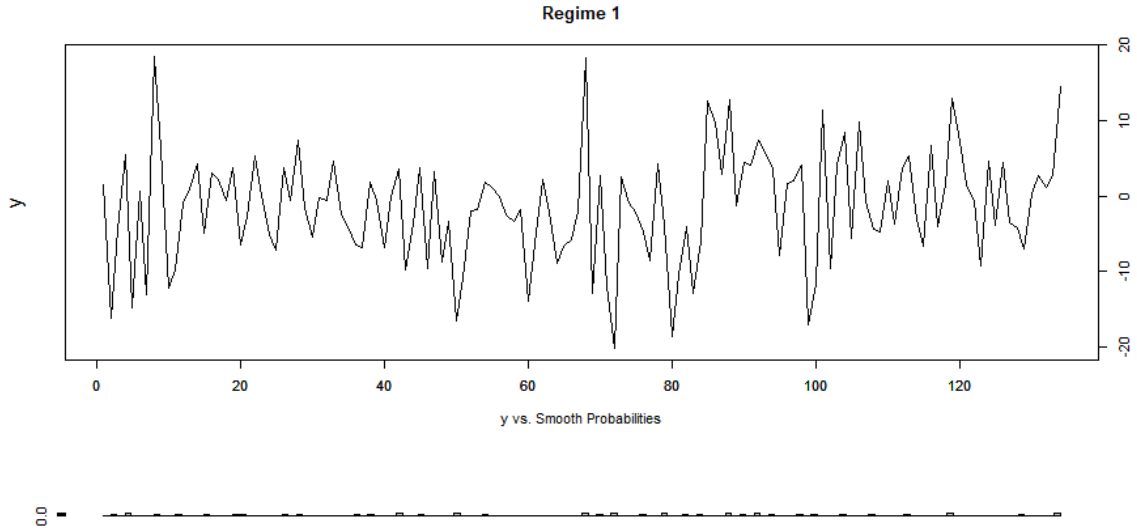

SP500 x Residuals
for Simple Regression
Residuals
SP500
Histogram of Residuals
Residuals
mod$resid
Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles



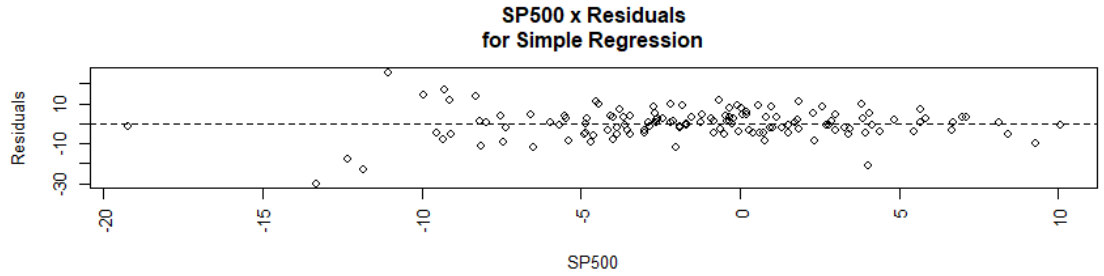

```
Order selected 1  sigma^2 estimated as  27.88
> mod_capm.mswm = msmFit(mod,k=2,p=1,sw=c(TRUE,TRUE,TRUE,TRUE),control=list(parallel=FALSE))
>
> summary(mod_capm.mswm)
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(TRUE, TRUE, TRUE, TRUE), p = 1,
    control = list(parallel = FALSE))

       AIC      BIC    logLik
  824.6276 871.4016 -406.3138

Coefficients:

Regime 1
---------
              Estimate Std. Error t value Pr(>|t|)
(Intercept)(S)   2.2704     2.4441  0.9289  0.35294
x(S)             0.8160     0.4317  1.8902  0.05873 .
y_1(S)           0.2310     0.3406  0.6782  0.49764
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.67771
Multiple R-squared:  0.25

Standardized Residuals:
        Min          Q1         Med          Q3         Max
-15.0132741  -2.1230993  -0.6774310   0.4459311  18.7524369

Regime 2
---------
              Estimate Std. Error t value Pr(>|t|)
(Intercept)(S)  -0.9055     0.5139 -1.7620  0.07807 .
x(S)             1.0677     0.1138  9.3822  < 2e-16 ***
y_1(S)          -0.1477     0.0767 -1.9257  0.05414 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.761988
Multiple R-squared: 0.6613

Standardized Residuals:
       Min         Q1        Med         Q3        Max
-7.3651745 -2.4027027  0.3649136  2.6099678  6.7126450

Transition probabilities:
           Regime 1  Regime 2
Regime 1 0.0002298648 0.3362424
Regime 2 0.9997701352 0.6637576
```

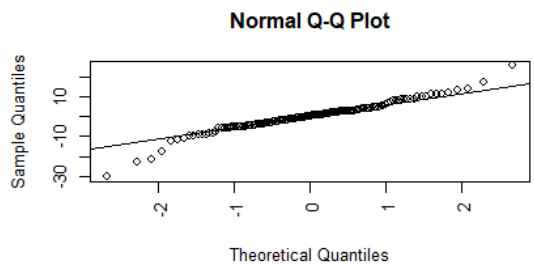

Regime 1
Smoothed Probabilities
Filtered Probabilities
0.0
0.8
0
20
40
60
80
100
120

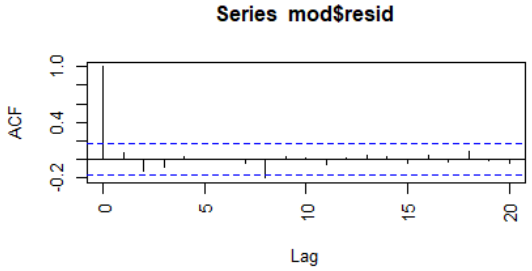

Regime 2
Smoothed Probabilities
Filtered Probabilities
0.0
0.8
0
20
40
60
80
100
120

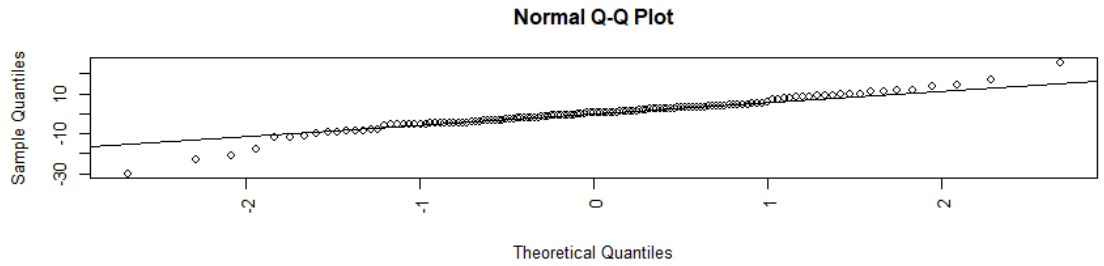

Regime 1
y
20
10
0
-10
-20
0
20
40
60
80
100
120
y vs. Smooth Probabilities
0.0

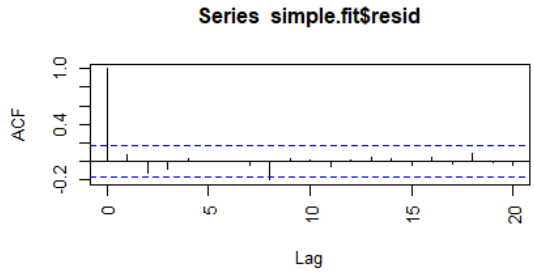
- ORACLE

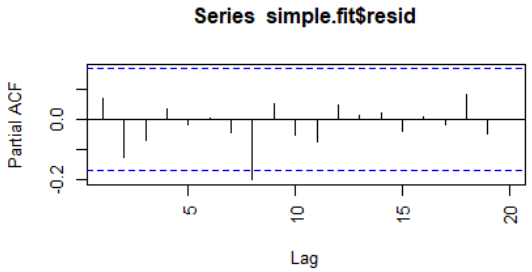

SP500 x Residuals
for Simple Regression
Residuals
SP500


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

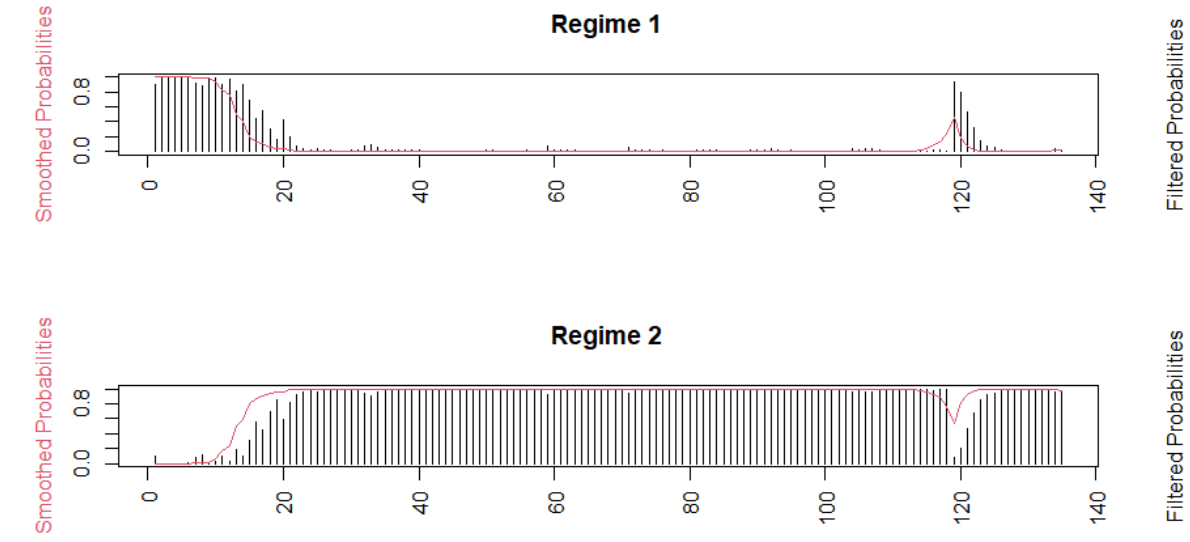

Series mod$resid
ACF
Lag

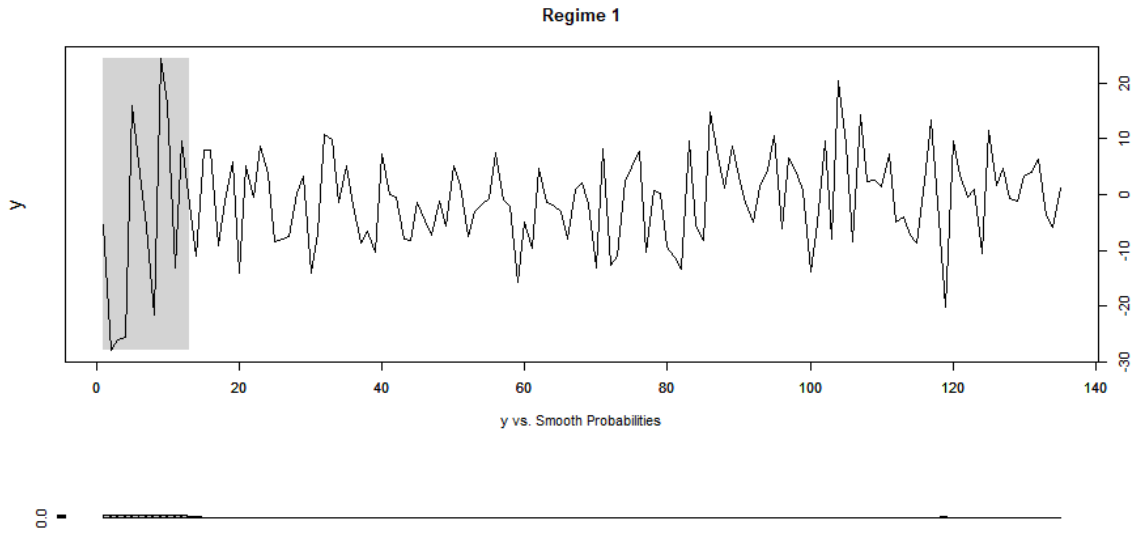

Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles


Series simple.fit$resid
ACF
Lag


Series simple.fit$resid
Partial ACF
Lag

```
Order selected 0  sigma^2 estimated as  53.03
> mod_capm.mswm = msmFit(mod,k=2,p=0,sw=c(TRUE,TRUE,TRUE),control=list(par
>
> summary(mod_capm.mswm)
Markov Switching Model

Call: msmFit(object = mod, k = 2, sw = c(TRUE, TRUE, TRUE), p = 0,
    control = list(parallel = FALSE))

       AIC      BIC    logLik
  886.7355 917.9777 -439.3678

Coefficients:

Regime 1
---------
              Estimate Std. Error t value Pr(>|t|)
(Intercept)(S)  -1.7861     5.2969 -0.3372   0.7360
x(S)             0.9645     0.7417  1.3004   0.1935

Residual standard error: 15.58766
Multiple R-squared: 0.1136

Standardized Residuals:
         Min           Q1          Med           Q3          Max
-27.91812681  -0.09000611   0.08059302   0.27821387  26.79483595

Regime 2
---------
              Estimate Std. Error t value Pr(>|t|)
(Intercept)(S)   0.5450     0.6601  0.8256    0.409
x(S)             1.0567     0.1094  9.6590   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.412026
Multiple R-squared: 0.4645

Standardized Residuals:
         Min           Q1          Med           Q3          Max
-15.95721054  -3.44093481   0.01774218   2.88584010  11.65867890

Transition probabilities:
         Regime 1   Regime 2
Regime 1 0.961071 0.01282027
Regime 2 0.038929 0.98717973
```






Conclusion: We follow the same process for the other series and determine the probabilities of moving from one regime to another and visualise the regimes in question