Master 2 Financial Economics – Université Paris 1 Panthéon-Sorbonne

Empirical Application 4

Financial Econometrics

Roland BOUILLOT
(Roland.Bouillot@etu.univ-paris1.fr)

Khalil JANBEK
(Khalil.Janbek@etu.univ-paris1.fr)

Mehdi LOUAFI
(Mehdi.Louafi@etu.univ-paris1.fr)

November 2021

EMPIRICAL APPLICATION 4

In econometrics, we usually have to be careful about heteroscedasticity in cross-section analyses and autocorrelation in time series. However, Robert Engel (1982) showed that heteroscedasticity could also be found in time series. This has some important implications regarding forecasts especially in the financial sphere where analysts working on exchange rate or stock markets noticed that errors seem to happen in clusters. To address this issue, Engel suggests the idea of conditional variance, implying that the recent past might give information about the future variance. In his paper, he gives the following formula:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-p}^2 \tag{1}$$

The conditional variance is the variance of $u_t$conditional on information available at time $t-1$. The conditional variance can be expressed as follows:

$$\sigma_t^2 = E\left(u_t^2 \middle| u_{t-1}, \dots, u_{t-p}\right)$$

$$\sigma_t^2 = E_{t-1}(u_t^2) \tag{2}$$

Where E_{t-1} takes the expectation conditional on all information up to the end of period $t-1$. That is, recent disturbances in the variance should have an impact on the current variance. From equation **(1)**, we can find the aforementioned disturbance:

$$u_t = \epsilon_t[\alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-p}^2]^{1/2} \tag{3}$$

Where ϵ_t is a white noise with unit variance. Thus, this can be defined as an ARCH(p) process.

To estimate an ARCH model an extensive econometrical literature has been developed, notably around the GARCH model and its many extensions. Indeed, GARCH models are usually less restrictive regarding the conditional variance equation **(1)** and can be expressed as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \cdots + \alpha_p u_{t-p}^2 + \gamma_1 \sigma_{t-1}^2 + \cdots + \gamma_q \sigma_{t-q}^2 \tag{4}$$

This yields the GARCH(p, q) model, which expresses the conditional variance as a linear function of p lagged square disturbances and q lagged conditional variances. The standard model -that we are using in this fourth empirical application- has a $p = 1$ and $q = 1$ parameters specification, giving us the following GARCH(1,1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1 \sigma_{t-1}^2 \tag{5}$$

Now that we have succinctly explained the econometrical framework of the models we are going to use for the estimation and forecast of our financial variables, it is time to present the data that will fuel this application.

A. The data

As in the previous empirical application, we take our monthly S&P500 stock index, CPI index, WTI oil prices and Industrial Production index variables. However, estimating the GARCH models for those variables is not conclusive enough to be presented in this empirical application. Only the time series of the monthly returns of the S&P500 yields some significant results.

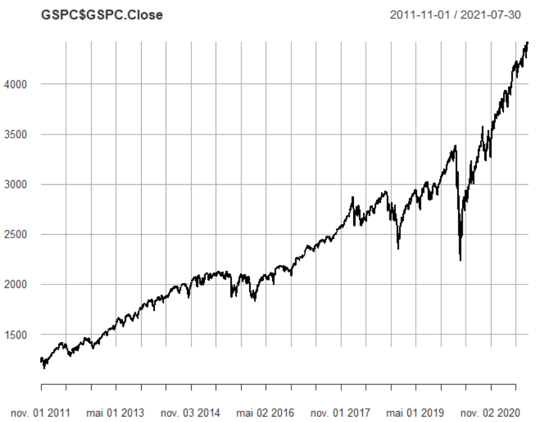
To address this data issue, we swap some variables. After reading the literature and understanding how volatility models work, we decided to stick with the S&P500 returns but on a daily basis and to introduce the daily returns of the Google (Alphabet to be precise) and Microsoft stocks.

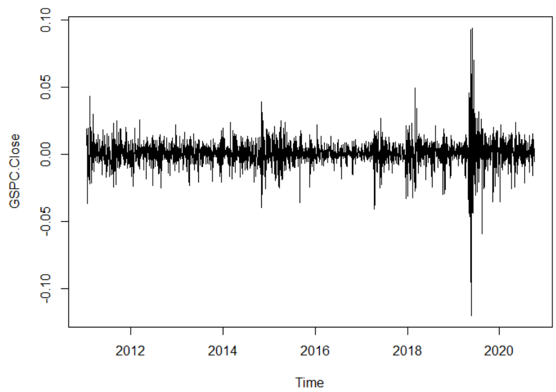
Our new dataset spans over a period starting 01-11-2011 and finishing the 30-07-2021. The data is retrieved from Yahoo Finance. We chose to work with the closing prices of business days (252 days per year) corresponding to the opening of the stock market and yield 2451 observation per variable over the whole period.

Figure 35

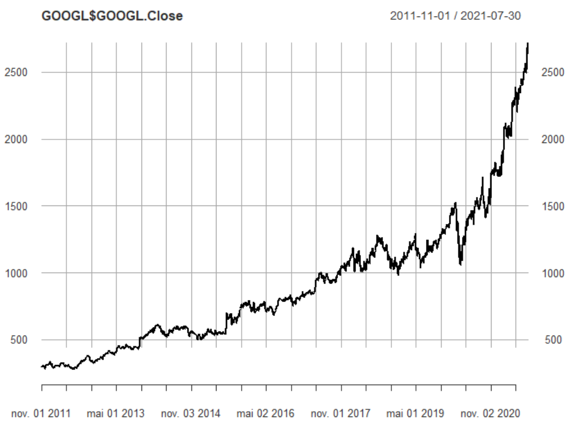
Plotted series of the S&P500 index and the Google and Microsoft stocks and their respective returns

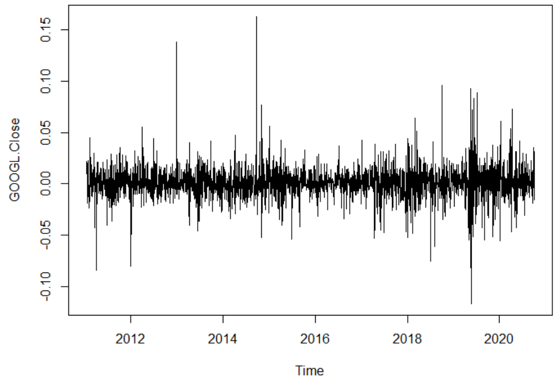
- <u>S&P500</u>



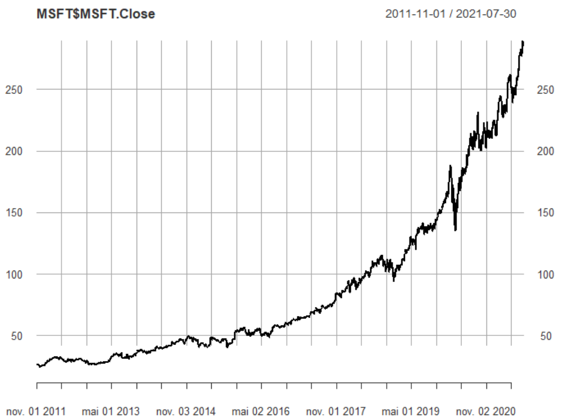


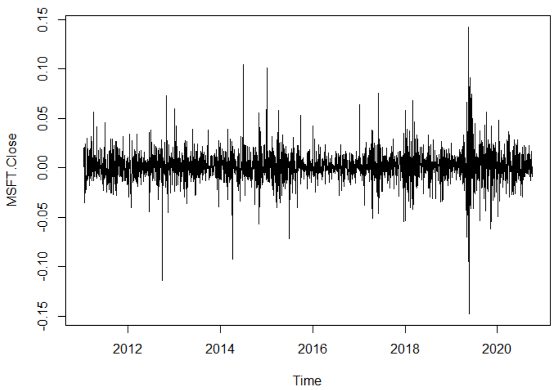
- <u>Google</u>





- <u>Microsoft</u>





B. GARCH analysis of the conditional variance

In this second section, we conduct two kind of GARCH analysis on our variables. First, we look for a univariate standard GARCH analysis of our three variables independently. Secondly, we gather the Google and Microsoft stock returns in a vector to perform a multivariate standard GARCH analysis. Yet, we first have to fit our data to comply with the restrictions of volatility models.

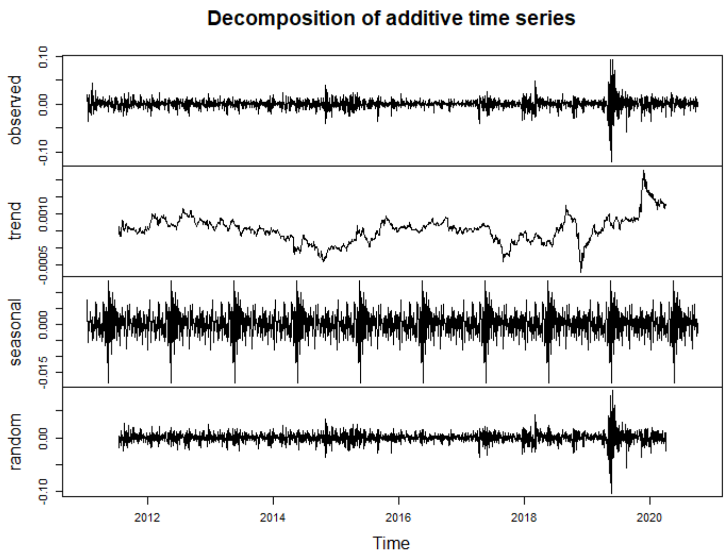
i. Fitting the data

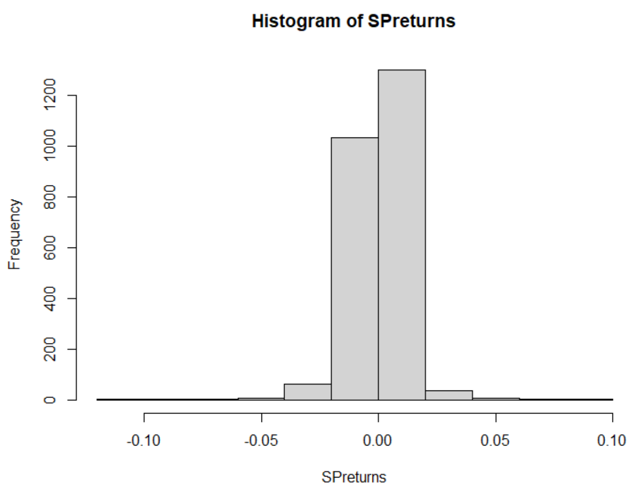
We start by identifying for deterministic trends and drifts in our data sample. We first decompose each of our time series to have an overall look of the trend and seasonal components. The plotted decompositions can be found in Figure 36.

Figure 36

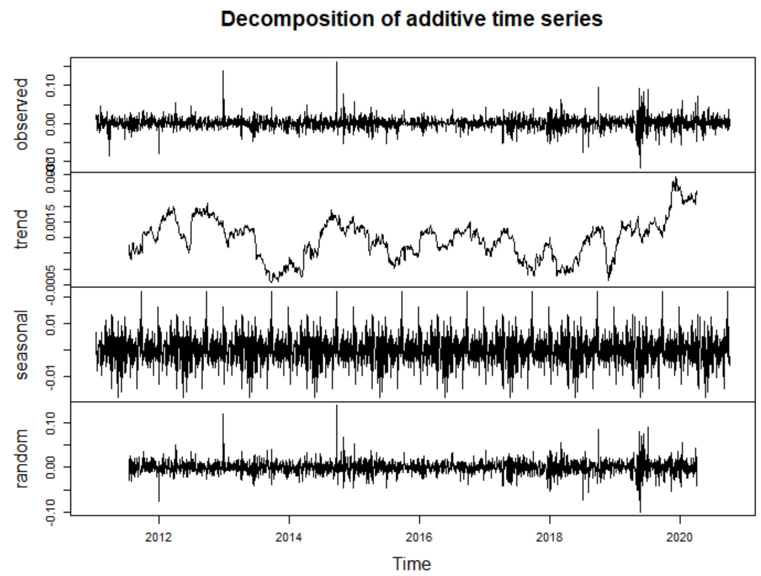
Decomposition of the time series and plotted histogram of returns

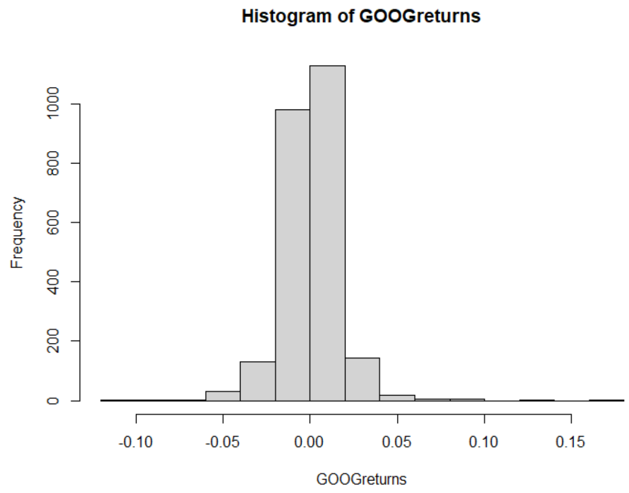
- <u>S&P500</u>



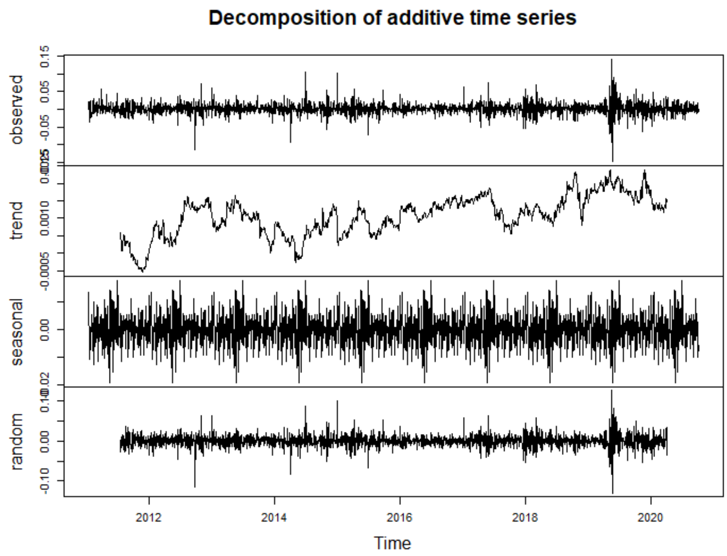


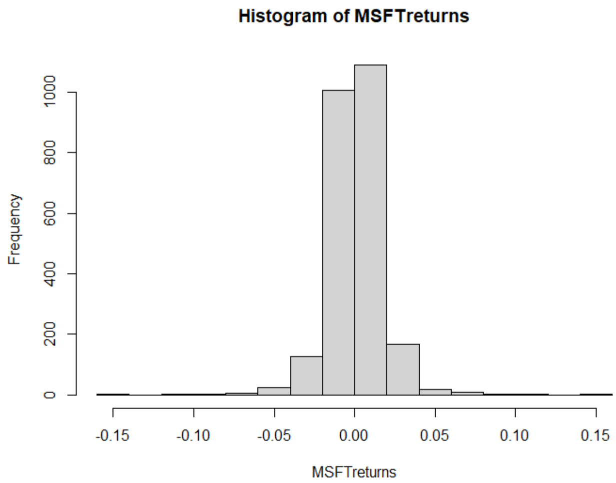
- Google





- Microsoft





The visual analysis seems to indicate that there are no clear upward or downward trends in our time series. Yet, we have to investigate further to be sure our data is good for use, especially by testing for the seasonality component. The results can be found in the following R-Code 46 (trend) and R-Code 47 (WO seasonality test) and Figure 37 (ACF/PACF).

As displayed, our time series have neither a trend nor a seasonal component. Thus, we can continue our fitting scheme and test for the presence of a stochastic trend by performing both an ADF and a PP unit root tests.

R-Code 46

Deterministic trend

- S&P500

```
> trend_SPreturns <- tslm(SPreturns~trend)
> summary(trend_SPreturns)

Call:
tslm(formula = SPreturns ~ trend)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.120498 -0.003824  0.000062  0.004649  0.093170 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.694e-04  4.241e-04   1.107    0.268
trend       8.932e-08  2.996e-07   0.298    0.766

Residual standard error: 0.01049 on 2449 degrees of freedom
Multiple R-squared:  3.63e-05,  Adjusted R-squared:  -0.000372 
F-statistic: 0.08889 on 1 and 2449 DF,  p-value: 0.7656
```

- Google

```
> trend_GOOGreturns <- tslm(GOOGreturns~trend)
> summary(trend_GOOGreturns)

Call:
tslm(formula = GOOGreturns ~ trend)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.117633 -0.007202 -0.000146  0.007870  0.161633 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 6.809e-04  6.447e-04   1.056    0.291
trend       2.904e-07  4.555e-07   0.638    0.524

Residual standard error: 0.01595 on 2449 degrees of freedom
Multiple R-squared:  0.000166,  Adjusted R-squared:  -0.0002423 
F-statistic: 0.4066 on 1 and 2449 DF,  p-value: 0.5238
```

- Microsoft

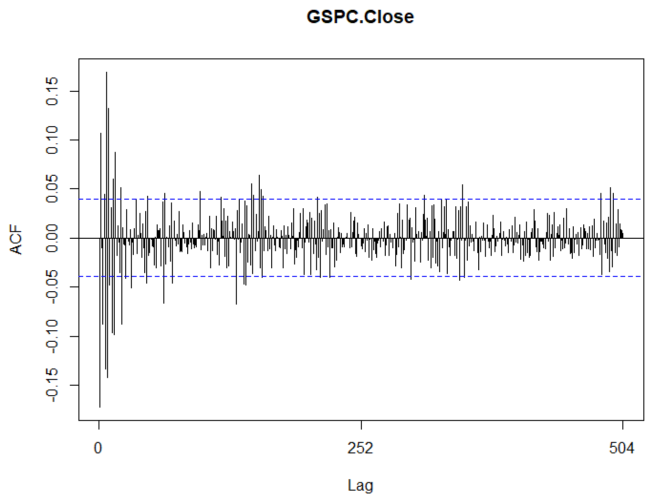
```
> trend_MSFTreturns <- tslm(MSFTreturns~trend)
> summary(trend_MSFTreturns)

Call:
tslm(formula = MSFTreturns ~ trend)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.148940 -0.007430 -0.000403  0.007845  0.140620 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.920e-04  6.533e-04   0.753    0.451
trend       5.025e-07  4.615e-07   1.089    0.276

Residual standard error: 0.01617 on 2449 degrees of freedom
Multiple R-squared:  0.0004837, Adjusted R-squared:  7.561e-05 
F-statistic: 1.185 on 1 and 2449 DF,  p-value: 0.2764
```

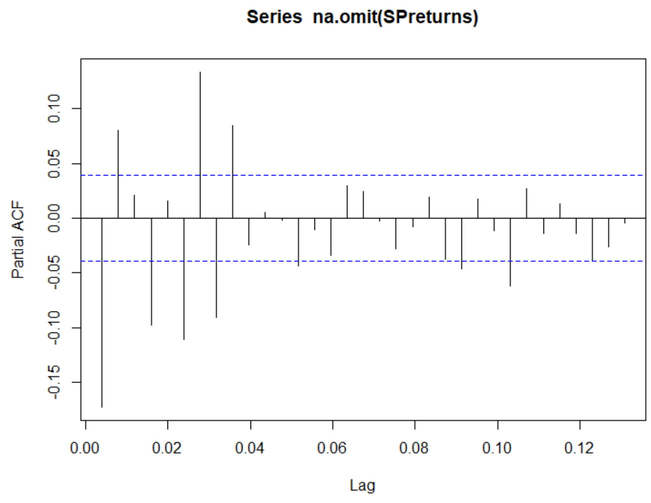
Figure 37

Seasonality component

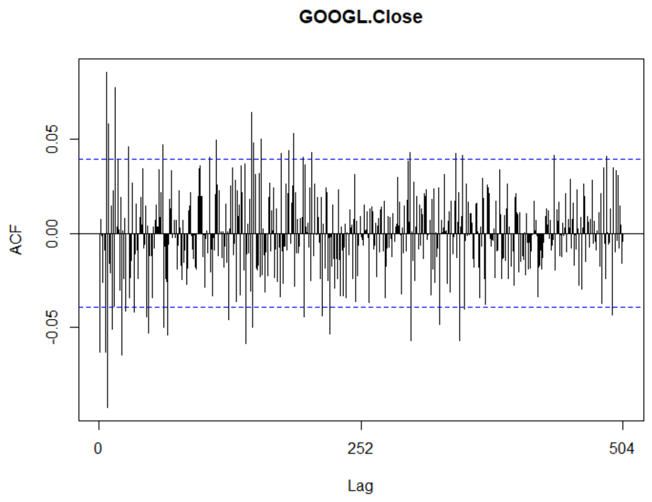
(ACF/PACF)

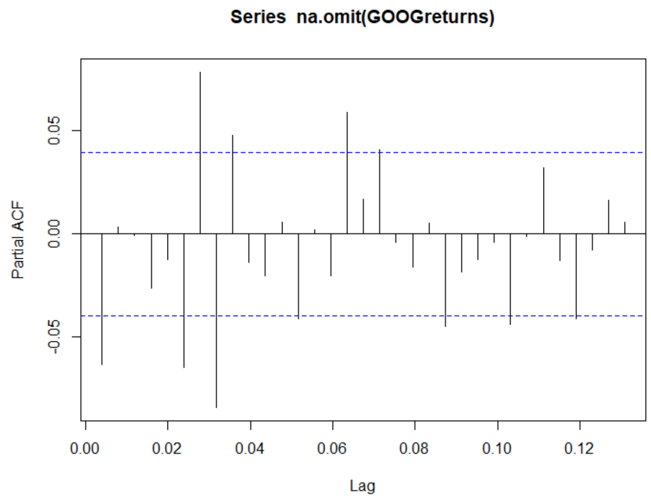
- S&P500



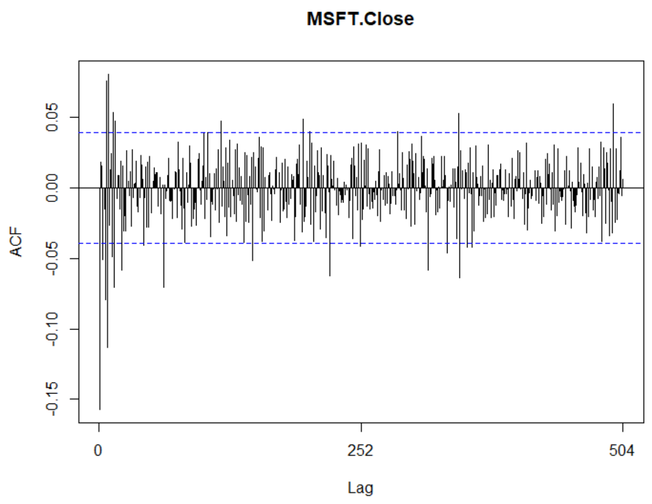


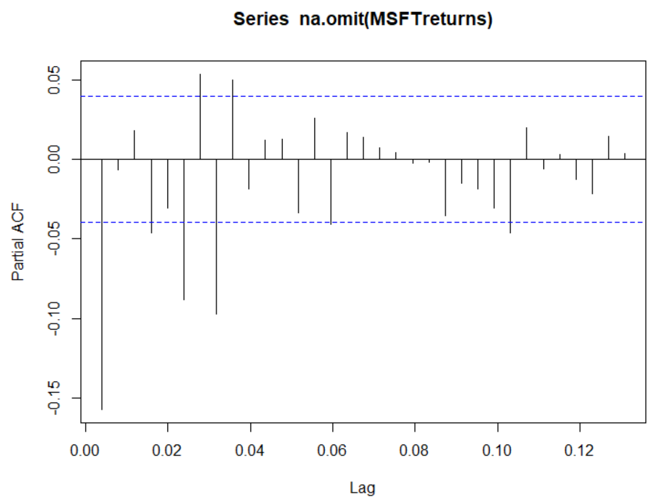
- Google





- Microsoft





R-Code 47

Seasonality component

(WO test)

- S&P500

```
> isSeasonal(SPreturns)
[1] FALSE
> summary(wo(SPreturns))
Test used:  WO

Test statistic:  0
P-value:  1 1 0.7391403

The WO - test does not identify  seasonality
```

- Google

```
> isSeasonal(GOOGreturns)
[1] FALSE
> summary(wo(GOOGreturns))
Test used:  WO

Test statistic:  0
P-value:  1 1 0.5465243

The WO - test does not identify  seasonality
```

- Microsoft

```
> isSeasonal(MSFTreturns)
[1] FALSE
> summary(wo(MSFTreturns))
Test used:  WO

Test statistic:  0
P-value:  0.01331483 0.4300111 0.2429494

The WO - test does not identify  seasonality
```

In the following R-Code 48 and 49, we present the results for the ADF and PP unit root tests for each of our variables. Stationarity is a crucial condition for the ARCH/GARCH models to perform well in both estimation and forecast.

R-Code 48

Testing for a stochastic trend
(ADF unit root test)

- <u>S&P500</u>

```
> adf.test(na.omit(SPreturns, k=1))

        Augmented Dickey-Fuller Test

data:  na.omit(SPreturns, k = 1)
Dickey-Fuller = -13.964, Lag order = 13, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(na.omit(SPreturns, k = 1)) :
  p-value smaller than printed p-value
```

- <u>Google</u>

```
> adf.test(na.omit(GOOGreturns, k=1))

        Augmented Dickey-Fuller Test

data:  na.omit(GOOGreturns, k = 1)
Dickey-Fuller = -14.19, Lag order = 13, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(na.omit(GOOGreturns, k = 1)) :
  p-value smaller than printed p-value
```

- <u>Microsoft</u>

```
> adf.test(na.omit(MSFTreturns, k=1))

        Augmented Dickey-Fuller Test

data:  na.omit(MSFTreturns, k = 1)
Dickey-Fuller = -14.029, Lag order = 13, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(na.omit(MSFTreturns, k = 1)) :
  p-value smaller than printed p-value
```

R-Code 49
Testing for a stochastic trend
(PP unit root test)

- S&P500

```
> pp.test(na.omit(SPreturns))

        Phillips-Perron Unit Root Test

data:  na.omit(SPreturns)
Dickey-Fuller Z(alpha) = -2893.2, Truncation lag parameter = 8, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In pp.test(na.omit(SPreturns)) : p-value smaller than printed p-value
```

- Google

```
> pp.test(na.omit(GOOGreturns))

        Phillips-Perron Unit Root Test

data:  na.omit(GOOGreturns)
Dickey-Fuller Z(alpha) = -2523.6, Truncation lag parameter = 8, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In pp.test(na.omit(GOOGreturns)) : p-value smaller than printed p-value
```

- Microsoft

```
> pp.test(na.omit(MSFTreturns))

        Phillips-Perron Unit Root Test

data:  na.omit(MSFTreturns)
Dickey-Fuller Z(alpha) = -2672.9, Truncation lag parameter = 8, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In pp.test(na.omit(MSFTreturns)) : p-value smaller than printed p-value
```

The presented results show that our time series do not display any stationarity issues. Our time series are stationary, which is a key condition to perform ARCH/GARCH analyses and allow us to continue our fitting scheme. Finally, we will check for ARCH effects in our data. Identifying ARCH effects is a requirement for ARCH/GARCH analyses. Without those effects, our estimations would not be significant at all and our forecasts erroneous. We run an ARCH-LM test to find out if our individual time series have such ARCH effects. We display the results in the R-Code 50 boxes below.

R-Code 50

Testing for ARCH effects

- S&P500

```
> SPreturnsArchTest <- ArchTest(SPreturns, lags=1, demean=TRUE)
> SPreturnsArchTest

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  SPreturns
Chi-squared = 665.26, df = 1, p-value < 2.2e-16
```

- Google

```
> GOOGreturnsArchTest <- ArchTest(GOOGreturns, lags=1, demean=TRUE)
> GOOGreturnsArchTest

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  GOOGreturns
Chi-squared = 42.662, df = 1, p-value = 6.505e-11
```

- Microsoft

```
> MSFTreturnsArchTest <- ArchTest(MSFTreturns, lags=1, demean=TRUE)
> MSFTreturnsArchTest

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  MSFTreturns
Chi-squared = 454.44, df = 1, p-value < 2.2e-16
```

The results of the ARCH-LM tests are quite conclusive, as each of three time series seem to have ARCH effects. Indeed, for each test we reject the null hypothesis of no ARCH effects.

Now that our data has been tested for deterministic components and stochastics trends as well as controlling for the presence of ARCH effects, we can assume that our data is good for use and ready to be implemented in a standard GARCH model.

ii. The Univariate analysis on our 3 variables

In the univariate analysis, we focus on the time series independently. We decide to present each analysis one after the other, starting with the S&P500 analysis.

a) <u>Univariate GARCH analysis of the S&P500 returns</u>

As stated in the introduction, we will carry our analysis based on a standard GARCH(1,1) model. In order to do so, we first need to define the optimal (p, q) parameters of the ARIMA model specification. As in the previous empirical application, we use the "*auto.arima*" function that yields the optimal parameters for our time series.

R-Code 50

Auto.arima result for S&P500 returns

```
> auto.arima(SPreturns, seasonal=FALSE, stationary = TRUE)
Series: SPreturns
ARIMA(4,0,5) with non-zero mean

Coefficients:
          ar1     ar2      ar3      ar4     ma1      ma2     ma3     ma4      ma5   mean
      -0.3655  0.6386  -0.2729  -0.7419  0.2457  -0.6105  0.3665  0.5727  -0.0454  6e-04
s.e.   0.1168  0.1146   0.0777   0.0753  0.1181   0.1125  0.0779  0.0763   0.0285  2e-04

sigma^2 estimated as 0.0001009:  log likelihood=7803.85
AIC=-15585.71   AICc=-15585.6   BIC=-15521.86
```

We find that the best fitting combination of p and q parameters are $p = 4$ and $q = 5$ in an ARIMA model. Those parameters are then used to calibrate our GARCH model specification. Our standard GARCH model specification has the following calibrated parameters:

```
*---------------------------------*
*       GARCH Model Spec          *
*---------------------------------*

Conditional Variance Dynamics
------------------------------------
GARCH Model             : sGARCH(1,1)
Variance Targeting      : FALSE

Conditional Mean Dynamics
------------------------------------
Mean Model              : ARFIMA(4,0,5)
Include Mean            : TRUE
GARCH-in-Mean           : FALSE

Conditional Distribution
------------------------------------
Distribution     :  norm
Includes Skew    :  FALSE
Includes Shape   :  FALSE
Includes Lambda  :  FALSE
```

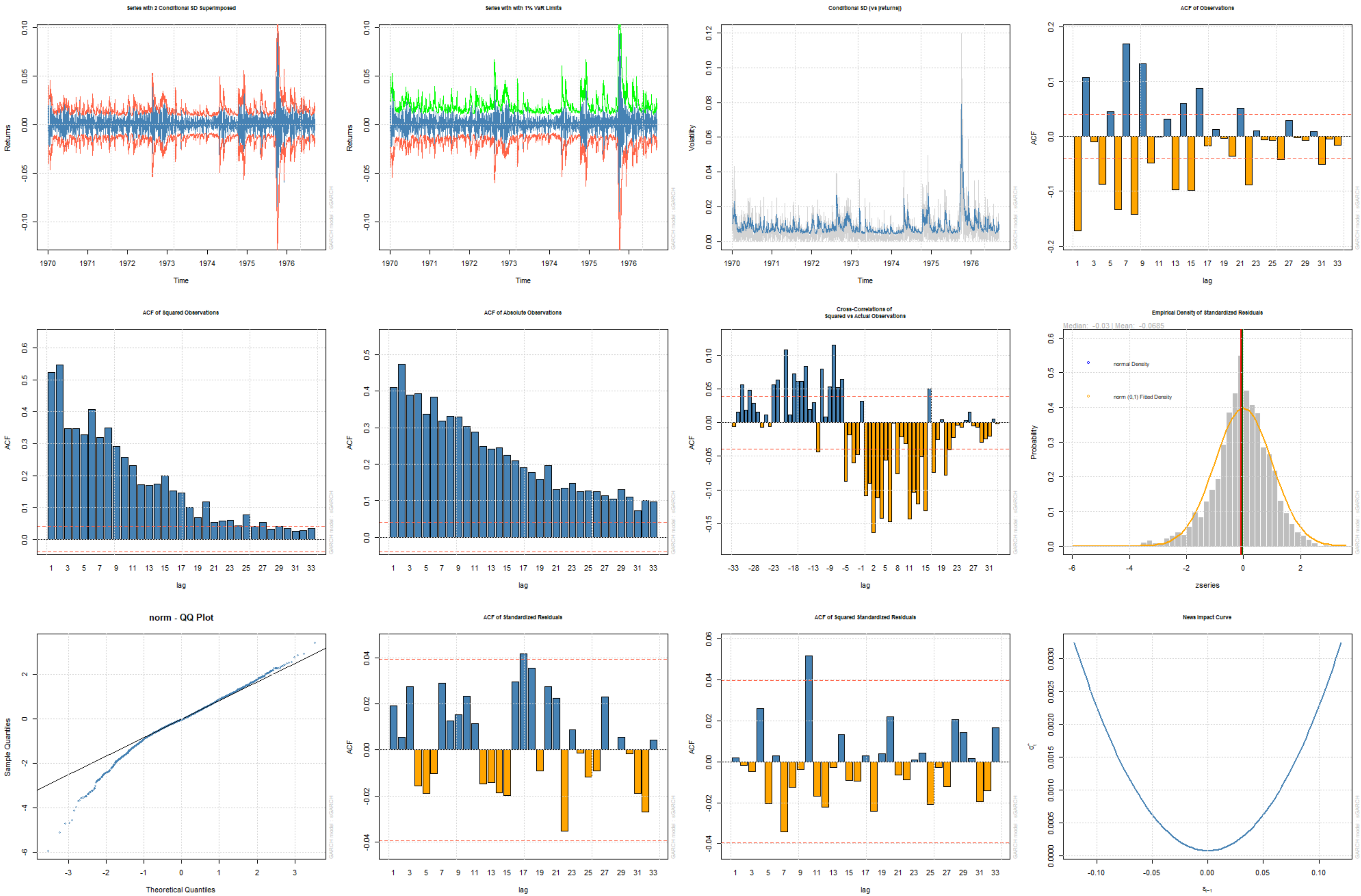
Once the specification has been set, it is now time to estimate our model! The estimation results are displayed in the following R-Code 51.

R-Code 51

Standard GARCH model for the S&P500 time series

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(4,0,5)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error     t value Pr(>|t|)
mu      0.000835    0.000064     13.0737  0.0e+00
ar1     0.152372    0.004390     34.7110  0.0e+00
ar2    -0.410549    0.002579   -159.1966  0.0e+00
ar3     0.675192    0.002752    245.3748  0.0e+00
ar4     0.385822    0.004986     77.3800  0.0e+00
ma1    -0.222411    0.001342   -165.7029  0.0e+00
ma2     0.431622    0.000029  15118.7238  0.0e+00
ma3    -0.744814    0.000038 -19537.0273  0.0e+00
ma4    -0.350124    0.000129  -2724.3523  0.0e+00
ma5    -0.018147    0.000279    -65.1315  0.0e+00
omega   0.000004    0.000001      4.2077  2.6e-05
alpha1  0.219445    0.017085     12.8440  0.0e+00
beta1   0.737263    0.013953     52.8390  0.0e+00

Robust Standard Errors:
        Estimate  Std. Error     t value Pr(>|t|)
mu      0.000835    0.000077     10.8149  0.00000
ar1     0.152372    0.005927     25.7060  0.00000
ar2    -0.410549    0.004364    -94.0772  0.00000
ar3     0.675192    0.003428    196.9794  0.00000
ar4     0.385822    0.006906     55.8694  0.00000
ma1    -0.222411    0.001437   -154.7831  0.00000
ma2     0.431622    0.000058   7459.4800  0.00000
ma3    -0.744814    0.000068 -10884.5195  0.00000
ma4    -0.350124    0.000127  -2766.4964  0.00000
ma5    -0.018147    0.000530    -34.2278  0.00000
omega   0.000004    0.000003      1.3350  0.18188
alpha1  0.219445    0.036237      6.0559  0.00000
beta1   0.737263    0.040554     18.1796  0.00000

LogLikelihood : 8379.313
```

Notice that our ω, α_1 and β_1 estimates are statistically significant. Further details of the estimation can be found in the annex part (Information Criteria, Ljung-Box, ARCH-LM, Nyblom stability, Sign bias and Pearson goodness-of-fit tests). In the next figure, we present all the plotted results of our estimation.

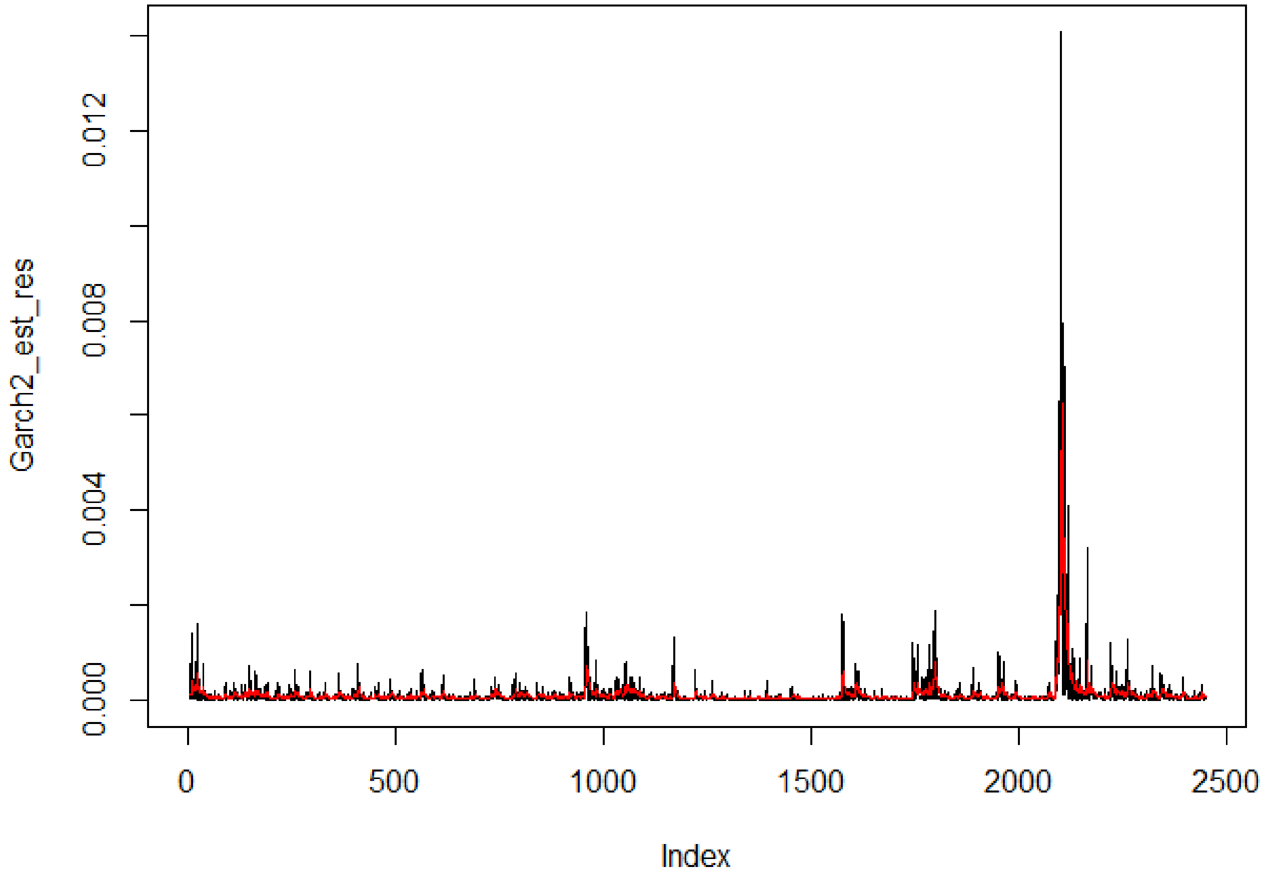


Series with 2 Conditional SD Superimposed
Returns
Time
Series with with 1% VaR Limits
Returns
Time
Conditional SD (vs |returns|)
Volatility
Time
ACF of Observations
ACF
lag
ACF of Squared Observations
ACF
lag
ACF of Absolute Observations
ACF
lag
Cross-Correlations of
Squared vs Actual Observations
ACF
lag
Empirical Density of Standardized Residuals
Median: -0.03 | Mean: -0.0685
normal Density
norm (0,1) Fitted Density
Probability
zseries
norm - QQ Plot
Sample Quantiles
Theoretical Quantiles
ACF of Standardized Residuals
ACF
lag
ACF of Squared Standardized Residuals
ACF
lag
News Impact Curve

In addition, we present the estimation of the variance for the S&P500 in Figure 38.

Figure 38

Plotted estimation of the variance for the S&P500 time series



Now that we have estimated the disturbance of the variance, we can turn to the forecast of our model in order to build over this estimation a prediction for the future variance of the S&P500 time series. In order to do so, we forecast our GARCH model to extract the Sigma values for a 20 periods horizon.

R-Code 52

Forecasted Sigma values

```
*------------------------------------*
*       GARCH Model Forecast         *
*------------------------------------*
Model: sGARCH
Horizon: 20
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=1976-09-17 02:00:00]:
         Series    Sigma
T+1   2.068e-03 0.006973
T+2  -1.182e-04 0.007141
T+3  -1.896e-04 0.007298
T+4   1.399e-03 0.007446
T+5   1.277e-03 0.007584
T+6  -3.888e-04 0.007714
T+7   4.528e-04 0.007837
T+8   1.796e-03 0.007952
T+9   4.827e-04 0.008061
T+10 -3.431e-04 0.008164
T+11  1.301e-03 0.008261
T+12  1.523e-03 0.008353
T+13 -1.830e-04 0.008440
T+14  2.580e-04 0.008522
T+15  1.809e-03 0.008600
T+16  7.983e-04 0.008674
T+17 -3.528e-04 0.008744
T+18  1.104e-03 0.008811
T+19  1.715e-03 0.008874
T+20  4.245e-05 0.008934
```

Figure 39

Plot of forecasted Sigma values

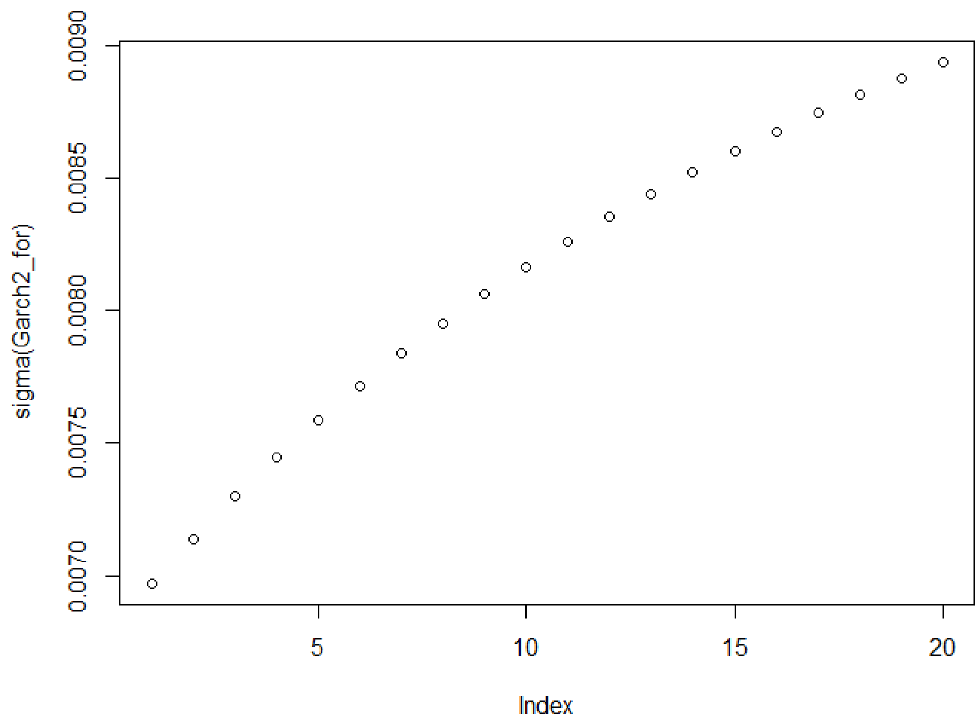


Figure 40

Forecasted variance for the S&P500 time series

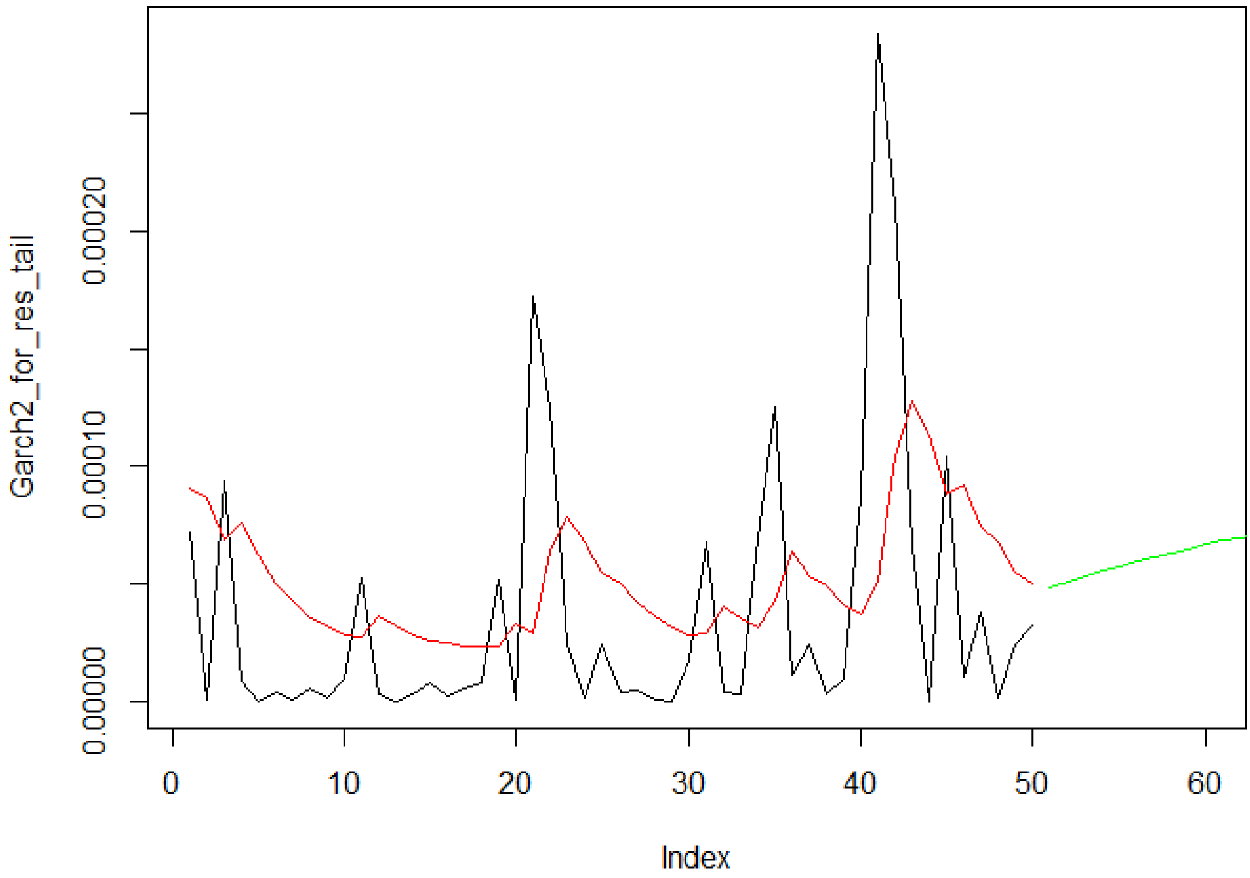


Figure 40 displays the forecasted path of the variance over a 20 period horizon (green line) taking into account the conditional estimated variance (red line). As spotted in Figure 39, the sigma displays an upward trend, that is reflected in the forecasted green line in Figure 40. From an economic perspective, this implies that the volatility of the S&P500 returns in the near future should increase.

b) Univariate GARCH analysis of the Google returns

As stated in the introduction, we will carry our analysis based on a standard GARCH(1,1) model. In order to do so, we first need to define the optimal (p, q) parameters of the ARIMA model specification. As in the previous empirical application, we use the "*auto.arima*" function that yields the optimal parameters for our time series.

R-Code 53

Auto.arima result for the Google returns

```
> auto.arima(GOOGreturns, seasonal=FALSE, stationary = TRUE)
Series: GOOGreturns
ARIMA(0,0,1) with non-zero mean

Coefficients:
         ma1    mean
      -0.063  1e-03
s.e.   0.020  3e-04

sigma^2 estimated as 0.0002536:  log likelihood=6670.25
AIC=-13334.5   AICc=-13334.49   BIC=-13317.09
```

We find that the best fitting combination of p and q parameters are $p = 0$ and $q = 1$ in an ARIMA model. Those parameters are then used to calibrate our GARCH model specification. Our standard GARCH model specification has the following calibrated parameters:

```
*---------------------------------*
*       GARCH Model Spec          *
*---------------------------------*

Conditional Variance Dynamics
------------------------------------
GARCH Model             : sGARCH(1,1)
Variance Targeting      : FALSE

Conditional Mean Dynamics
------------------------------------
Mean Model              : ARFIMA(0,0,1)
Include Mean            : TRUE
GARCH-in-Mean           : FALSE

Conditional Distribution
------------------------------------
Distribution     :  norm
Includes Skew    :  FALSE
Includes Shape   :  FALSE
Includes Lambda  :  FALSE
```

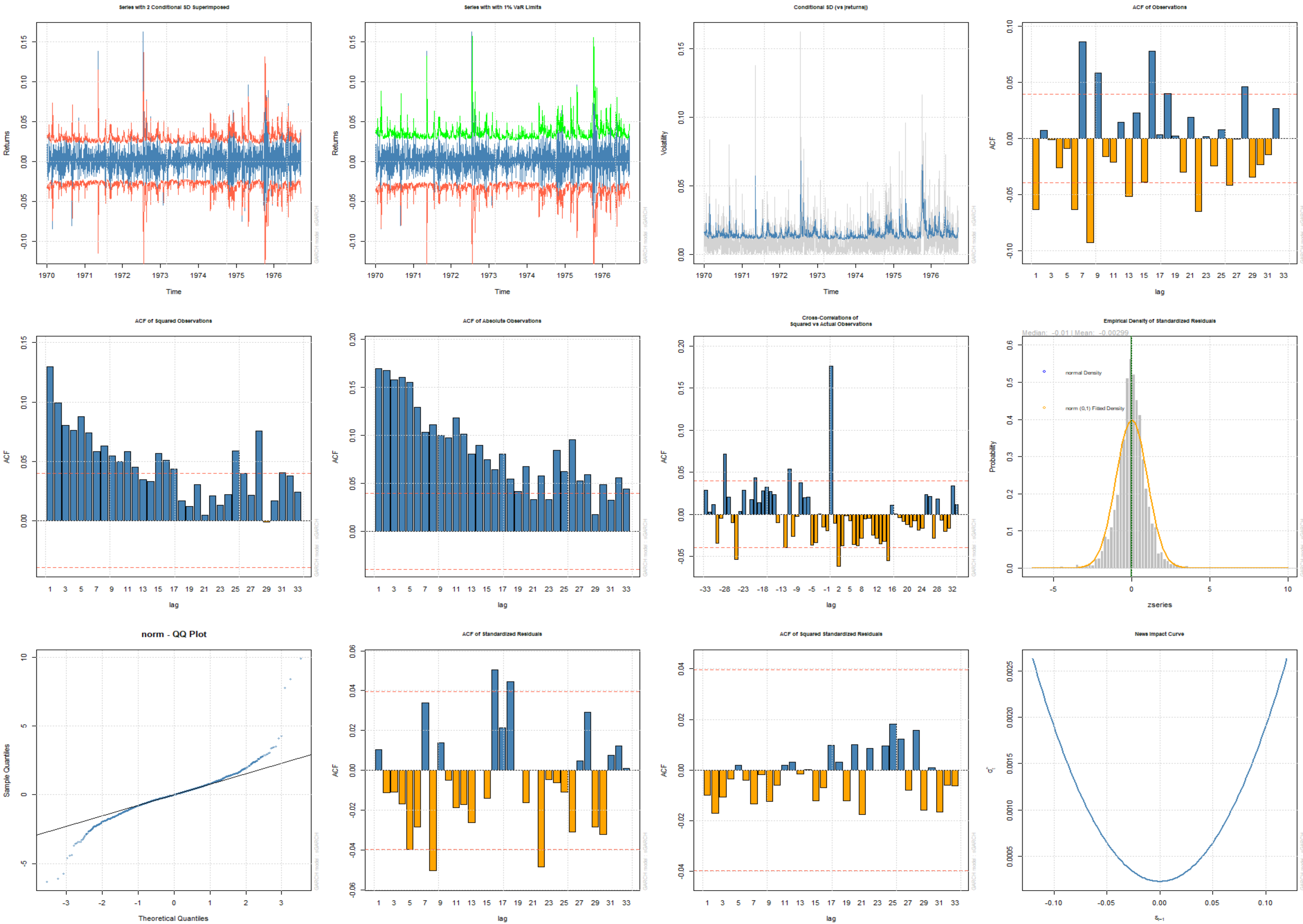
Once the specification has been set, it is now time to estimate our model! The estimation results are displayed in the following R-Code 54.

R-Code 54

Standard GARCH model for the Google returns time series

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.001084    0.000273  3.96877 0.000072
ma1    -0.018677    0.024527 -0.76148 0.446369
omega   0.000036    0.000007  5.17260 0.000000
alpha1  0.167257    0.027467  6.08937 0.000000
beta1   0.698290    0.045969 15.19031 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.001084    0.000291  3.72180 0.000198
ma1    -0.018677    0.022409 -0.83347 0.404578
omega   0.000036    0.000021  1.73628 0.082514
alpha1  0.167257    0.080864  2.06838 0.038604
beta1   0.698290    0.133607  5.22645 0.000000

LogLikelihood : 6827.718
```

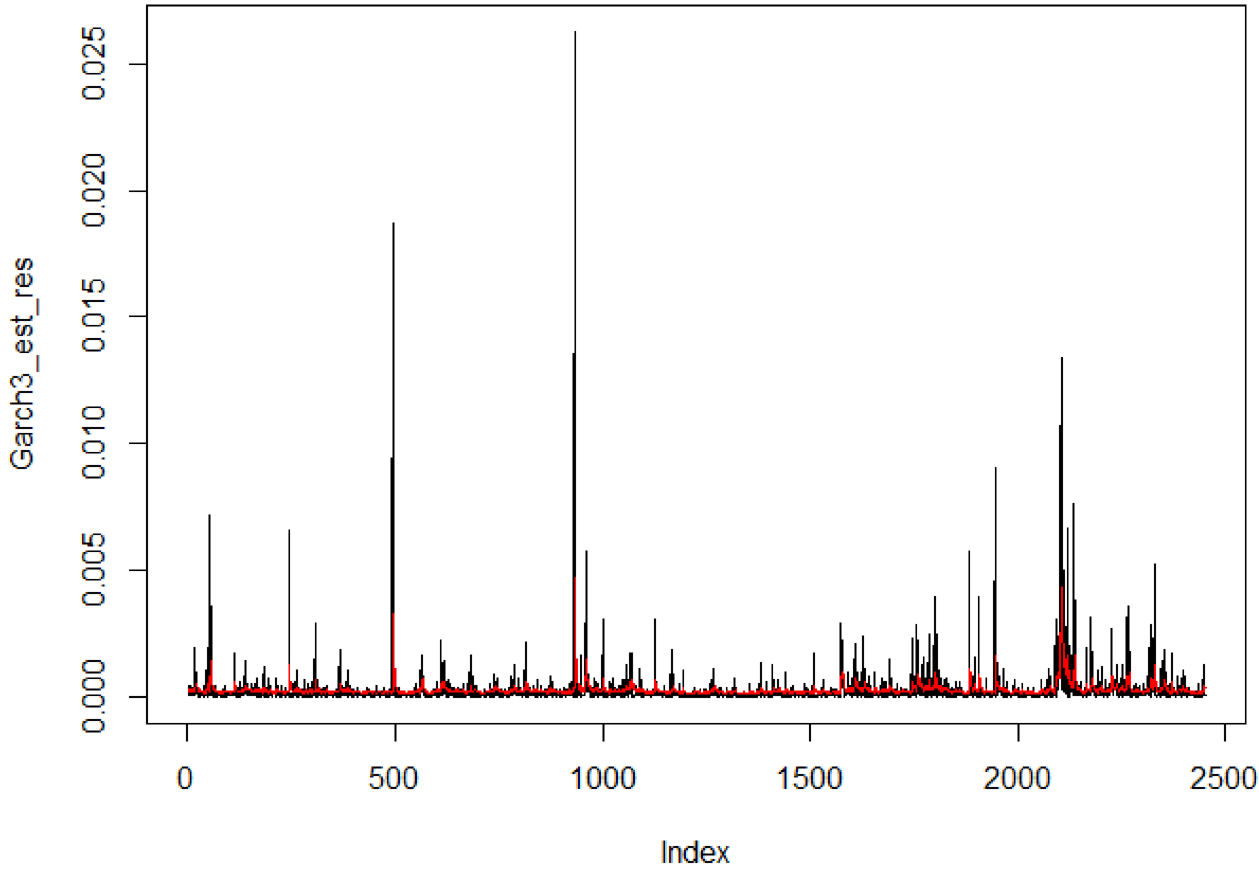
Notice that our ω, α_1 and β_1 estimates are statistically significant. Further details of the estimation can be found in the annex part (Information Criteria, Ljung-Box, ARCH-LM, Nyblom stability, Sign bias and Pearson goodness-of-fit tests). In the next figure, we present all the plotted results of our estimation.


Series with 2 Conditional SD Superimposed
Returns
Time
Series with with 1% VaR Limits
Returns
Time
Conditional SD (vs |returns|)
Volatility
Time
ACF of Observations
ACF
lag
ACF of Squared Observations
ACF
lag
ACF of Absolute Observations
ACF
lag
Cross-Correlations of
Squared vs Actual Observations
ACF
lag
Empirical Density of Standardized Residuals
Median: -0.01 | Mean: -0.00299
normal Density
norm (0,1) Fitted Density
Probability
zseries
norm - QQ Plot
Sample Quantiles
Theoretical Quantiles
ACF of Standardized Residuals
ACF
lag
ACF of Squared Standardized Residuals
ACF
lag
News Impact Curve

In addition, we present the estimation of the variance for the Google in Figure 41.

Figure 41

Plotted estimation of the variance for the Google returns time series



Now that we have estimated the disturbance of the variance, we can turn to the forecast of our model in order to build over this estimation a prediction for the future variance of the Google time series. In order to do so, we forecast our GARCH model to extract the Sigma values for a 20 periods horizon.

R-Code 55

Forecasted Sigma values

```
*------------------------------------*
*       GARCH Model Forecast         *
*------------------------------------*
Model: sGARCH
Horizon: 20
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=1976-09-17 02:00:00]:
       Series   Sigma
T+1  0.001250 0.01640
T+2  0.001084 0.01641
T+3  0.001084 0.01642
T+4  0.001084 0.01643
T+5  0.001084 0.01643
T+6  0.001084 0.01644
T+7  0.001084 0.01644
T+8  0.001084 0.01645
T+9  0.001084 0.01645
T+10 0.001084 0.01645
T+11 0.001084 0.01646
T+12 0.001084 0.01646
T+13 0.001084 0.01646
T+14 0.001084 0.01646
T+15 0.001084 0.01646
T+16 0.001084 0.01646
T+17 0.001084 0.01646
T+18 0.001084 0.01647
T+19 0.001084 0.01647
T+20 0.001084 0.01647
```

Figure 42

Plot of forecasted Sigma values

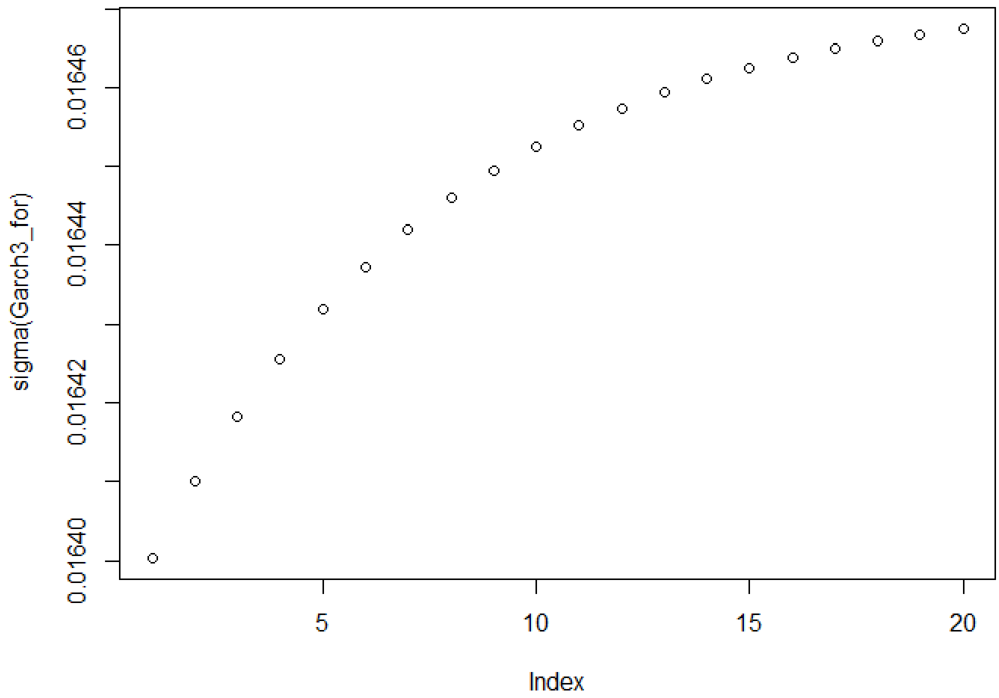


Figure 43

Forecasted variance for the Google time series

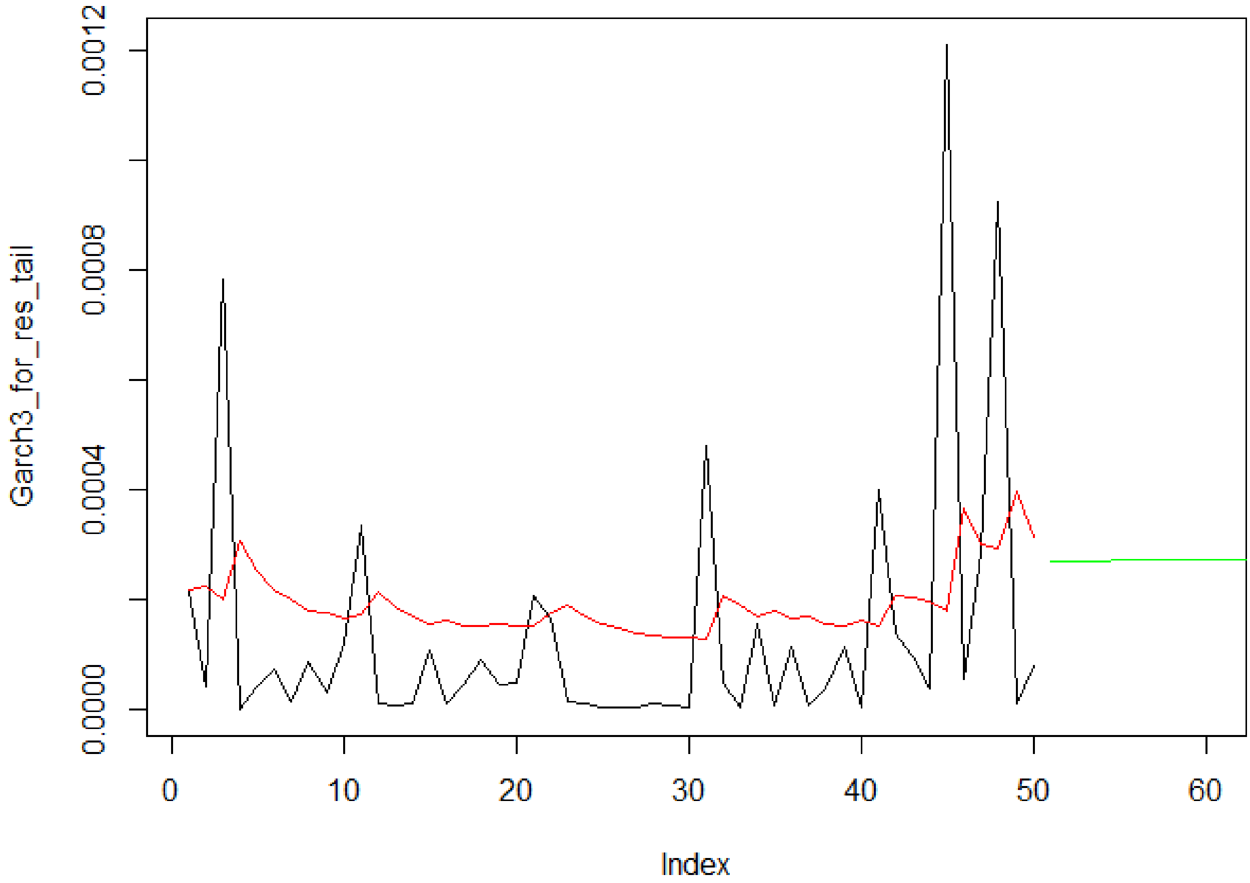


Figure 43 displays the forecasted path of the variance over a 20 period horizon (green line) taking into account the conditional estimated variance (red line). As spotted in Figure 42, the sigma displays an upward trend, that is slightly reflected in the forecasted green line in Figure 40. From an economic perspective, this implies that the volatility of the Google returns in the near future should slightly increase.

c) Univariate GARCH analysis of the Microsoft returns

As stated in the introduction, we will carry our analysis based on a standard GARCH(1,1) model. In order to do so, we first need to define the optimal (p, q) parameters of the ARIMA model specification. As in the previous empirical application, we use the "*auto.arima*" function that yields the optimal parameters for our time series.

R-Code 56

Auto.arima result for the Microsoft returns

```
> auto.arima(MSFTreturns, seasonal=FALSE, stationary = TRUE)
Series: MSFTreturns
ARIMA(4,0,4) with non-zero mean

Coefficients:
         ar1     ar2      ar3      ar4      ma1      ma2     ma3     ma4    mean
      0.1309  1.3625  -0.0233  -0.8132  -0.2379  -1.3314  0.1331  0.7381  0.0011
s.e.  0.0446  0.0408   0.0430   0.0383   0.0517   0.0497  0.0514  0.0470  0.0003

sigma^2 estimated as 0.0002503:  log likelihood=6689.32
AIC=-13358.64   AICc=-13358.55   BIC=-13300.6
```

We find that the best fitting combination of p and q parameters are $p = 4$ and $q = 4$ in an ARIMA model. Those parameters are then used to calibrate our GARCH model specification. Our standard GARCH model specification has the following calibrated parameters:

```
*---------------------------------*
*       GARCH Model Spec          *
*---------------------------------*

Conditional Variance Dynamics
------------------------------------
GARCH Model             : sGARCH(1,1)
Variance Targeting      : FALSE

Conditional Mean Dynamics
------------------------------------
Mean Model              : ARFIMA(4,0,4)
Include Mean            : TRUE
GARCH-in-Mean           : FALSE

Conditional Distribution
------------------------------------
Distribution     :  norm
Includes Skew    :  FALSE
Includes Shape   :  FALSE
Includes Lambda  :  FALSE
```

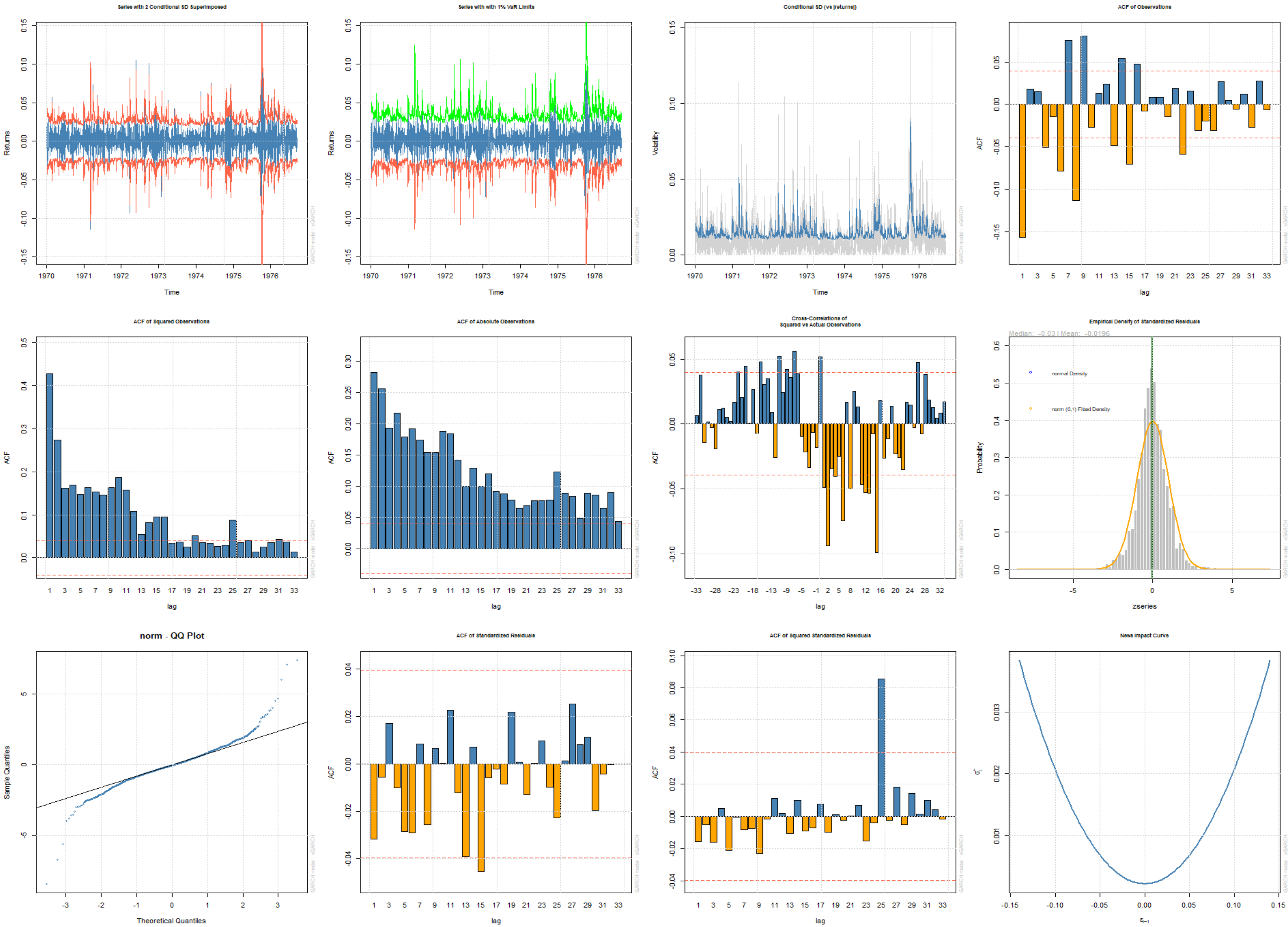
Once the specification has been set, it is now time to estimate our model! The estimation results are displayed in the following R-Code 57.

R-Code 57

Standard GARCH model for the Microsoft returns time series

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(4,0,4)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error    t value Pr(>|t|)
mu      0.001381    0.000249     5.5564        0
ar1     1.437542    0.005061   284.0534        0
ar2    -1.159857    0.001789  -648.1891        0
ar3     1.450731    0.001418  1023.1654        0
ar4    -0.961939    0.004595  -209.3604        0
ma1    -1.458540    0.000477 -3059.3800        0
ma2     1.175712    0.000111 10605.9554        0
ma3    -1.473596    0.000647 -2275.8278        0
ma4     0.973543    0.000226  4306.0332        0
omega   0.000030    0.000005     5.9488        0
alpha1  0.185144    0.027166     6.8152        0
beta1   0.699572    0.035459    19.7293        0

Robust Standard Errors:
        Estimate  Std. Error    t value Pr(>|t|)
mu      0.001381    0.000240     5.7467 0.000000
ar1     1.437542    0.009475   151.7246 0.000000
ar2    -1.159857    0.002227  -520.8805 0.000000
ar3     1.450731    0.002559   566.8427 0.000000
ar4    -0.961939    0.010334   -93.0872 0.000000
ma1    -1.458540    0.000885 -1648.9494 0.000000
ma2     1.175712    0.000562  2091.9356 0.000000
ma3    -1.473596    0.001318 -1118.1143 0.000000
ma4     0.973543    0.000340  2860.5561 0.000000
omega   0.000030    0.000011     2.6273 0.008606
alpha1  0.185144    0.046415     3.9889 0.000066
beta1   0.699572    0.063953    10.9388 0.000000

LogLikelihood : 6913.537
```

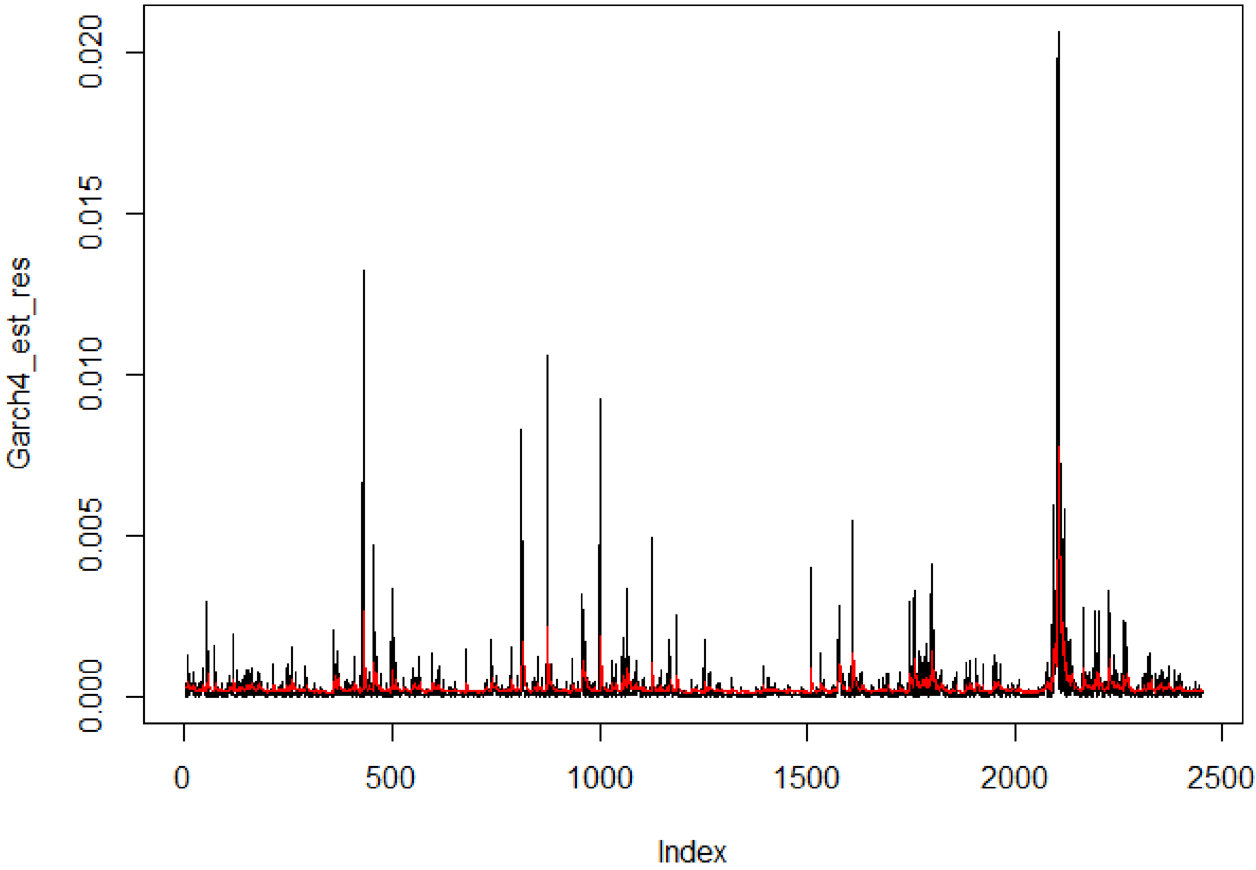
Notice that our ω, α_1 and β_1 estimates are statistically significant. Further details of the estimation can be found in the annex part (Information Criteria, Ljung-Box, ARCH-LM, Nyblom stability, Sign bias and Pearson goodness-of-fit tests). In the next figure, we present all the plotted results of our estimation.


Series with 2 Conditional SD Superimposed
Returns
Time
Series with with 1% VaR Limits
Conditional SD (vs |returns|)
Volatility
ACF of Observations
ACF
lag
ACF of Squared Observations
ACF of Absolute Observations
Cross-Correlations of
Squared vs Actual Observations
Empirical Density of Standardized Residuals
Median: -0.03 | Mean: -0.0196
normal Density
norm (0,1) Fitted Density
Probability
zseries
norm - QQ Plot
Sample Quantiles
Theoretical Quantiles
ACF of Standardized Residuals
ACF of Squared Standardized Residuals
News Impact Curve

In addition, we present the estimation of the variance for the Microsoft in Figure 44.

Figure 44

Plotted estimation of the variance for the Microsoft returns time series



Now that we have estimated the disturbance of the variance, we can turn to the forecast of our model in order to build over this estimation a prediction for the future variance of the Microsoft time series. In order to do so, we forecast our GARCH model to extract the Sigma values for a 20 period horizon.

R-Code 58

Forecasted Sigma values

```
*------------------------------------*
*       GARCH Model Forecast         *
*------------------------------------*
Model: sGARCH
Horizon: 20
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=1976-09-17 02:00:00]:
        Series   Sigma
T+1  -0.0013464 0.01185
T+2  -0.0003431 0.01242
T+3   0.0048464 0.01290
T+4   0.0016429 0.01332
T+5  -0.0021395 0.01367
T+6   0.0027021 0.01398
T+7   0.0044099 0.01425
T+8  -0.0011564 0.01448
T+9  -0.0004767 0.01468
T+10  0.0047768 0.01486
T+11  0.0018224 0.01501
T+12 -0.0021773 0.01515
T+13  0.0024672 0.01527
T+14  0.0044433 0.01537
T+15 -0.0010634 0.01546
T+16 -0.0006862 0.01554
T+17  0.0046421 0.01561
T+18  0.0019747 0.01567
T+19 -0.0021956 0.01573
T+20  0.0022704 0.01578
```

Figure 45

Plot of forecasted Sigma values

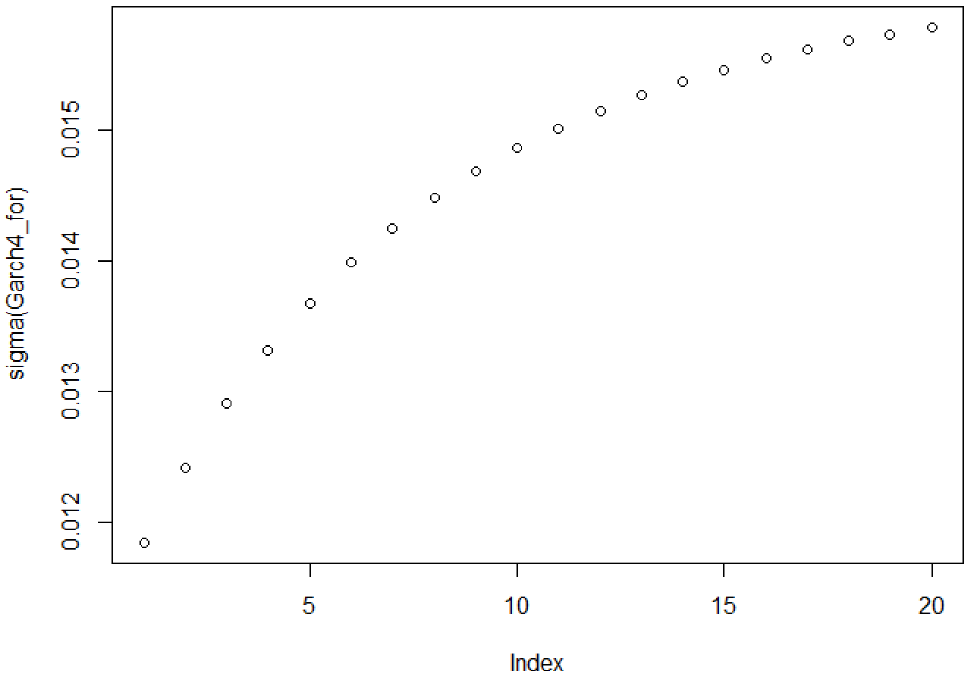


Figure 46

Forecasted variance for the Microsoft time series

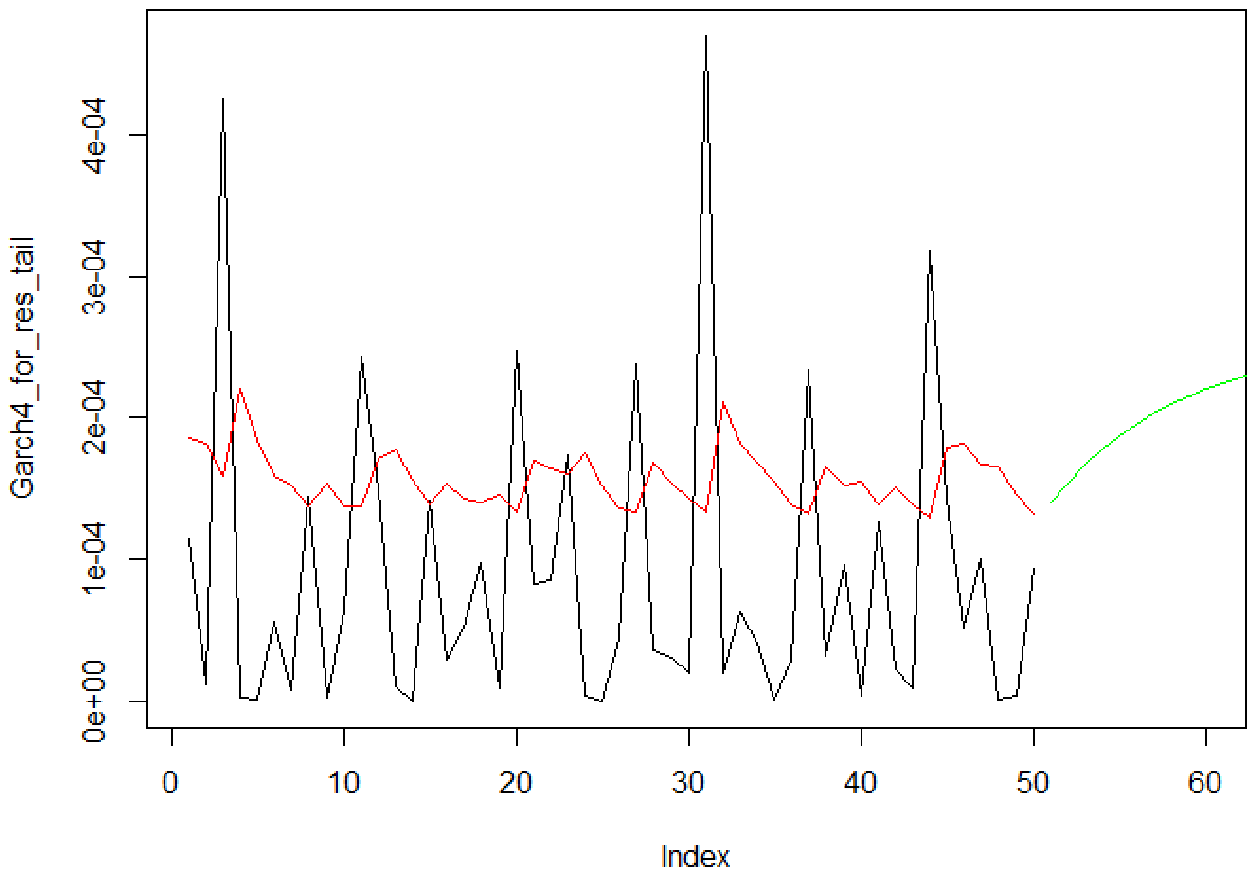


Figure 46 displays the forecasted path of the variance over a 20 period horizon (green line) taking into account the conditional estimated variance (red line). As spotted in Figure 45, the sigma displays an upward trend, that is reflected in the forecasted green line in Figure 46. From an economic perspective, this implies that the volatility of the Microsoft returns in the near future should increase.

iii. The Multivariate analysis on the Google and Microsoft stock returns

C. ARCH in Mean analysis

In order to do an ARCH in Mean analysis on our data, we have to adapt our GARCH model by removing the q component as seen in equation **(4)** which leaves us with a standard ARCH model (only with the p component). Next, we have to calibrate our new GARCH(1,0) model (which is an ARCH(1) model) with a mean analysis. To do so, we just have to include an "*archm = TRUE*" option specification in the model.

As we encounter some convergence issues with the S&P500 and the Google time series, we chose to conduct the analysis with the Microsoft returns.

Our ARCH-M model specification is the following :

```
*---------------------------------*
*       GARCH Model Spec          *
*---------------------------------*

Conditional Variance Dynamics
------------------------------------
GARCH Model             : sGARCH(1,0)
Variance Targeting      : FALSE

Conditional Mean Dynamics
------------------------------------
Mean Model              : ARFIMA(4,0,4)
Include Mean            : TRUE
GARCH-in-Mean           : TRUE

Conditional Distribution
------------------------------------
Distribution    :  norm
Includes Skew   :  FALSE
Includes Shape  :  FALSE
Includes Lambda :  FALSE
```

As stated above, we have a standard GARCH(1,0) model that constitutes our ARCH(1) model. The ARIMA specification corresponds to the ARMA(4,4) result we found for the Microsoft series through the "auto.arima" function and we have enabled the GARCH-in-Mean option. With those specifications, our model should be an ARCH-M(1) model.

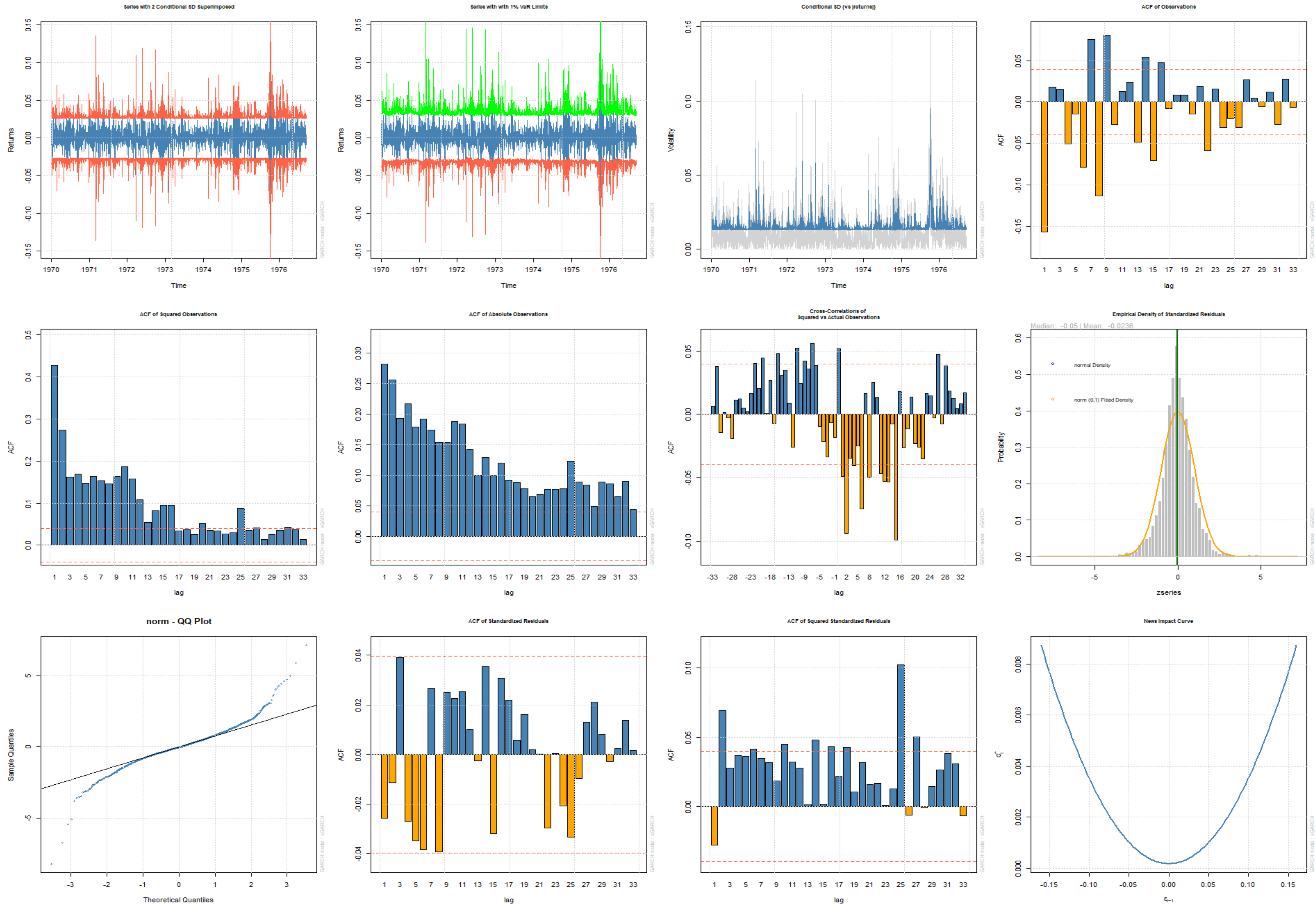
The estimations of our model are displayed in the following R-Code 59 box.

R-Code 59

ARCH-in-mean model for the Microsoft returns time series

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,0)
Mean Model      : ARFIMA(4,0,4)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error     t value Pr(>|t|)
mu     -0.002486    0.000181 -1.3766e+01        0
ar1    -1.112032    0.003887 -2.8607e+02        0
ar2     0.238184    0.001076  2.2145e+02        0
ar3     1.160101    0.004188  2.7700e+02        0
ar4     0.397013    0.004797  8.2755e+01        0
ma1     1.076767    0.000887  1.2142e+03        0
ma2    -0.290652    0.000001 -2.0226e+05        0
ma3    -1.192363    0.000000 -4.2046e+07        0
ma4    -0.394178    0.000001 -4.8022e+05        0
archm   0.256085    0.001650  1.5524e+02        0
omega   0.000171    0.000005  3.2357e+01        0
alpha1  0.334265    0.006785  4.9262e+01        0

Robust Standard Errors:
        Estimate  Std. Error     t value Pr(>|t|)
mu     -0.002486    0.000274 -9.0775e+00        0
ar1    -1.112032    0.003562 -3.1220e+02        0
ar2     0.238184    0.001313  1.8147e+02        0
ar3     1.160101    0.007428  1.5618e+02        0
ar4     0.397013    0.006811  5.8291e+01        0
ma1     1.076767    0.001962  5.4874e+02        0
ma2    -0.290652    0.000004 -7.4057e+04        0
ma3    -1.192363    0.000000 -1.4566e+07        0
ma4    -0.394178    0.000002 -1.7275e+05        0
archm   0.256085    0.007630  3.3562e+01        0
omega   0.000171    0.000014  1.2365e+01        0
alpha1  0.334265    0.015234  2.1941e+01        0

LogLikelihood : 6822.992
```

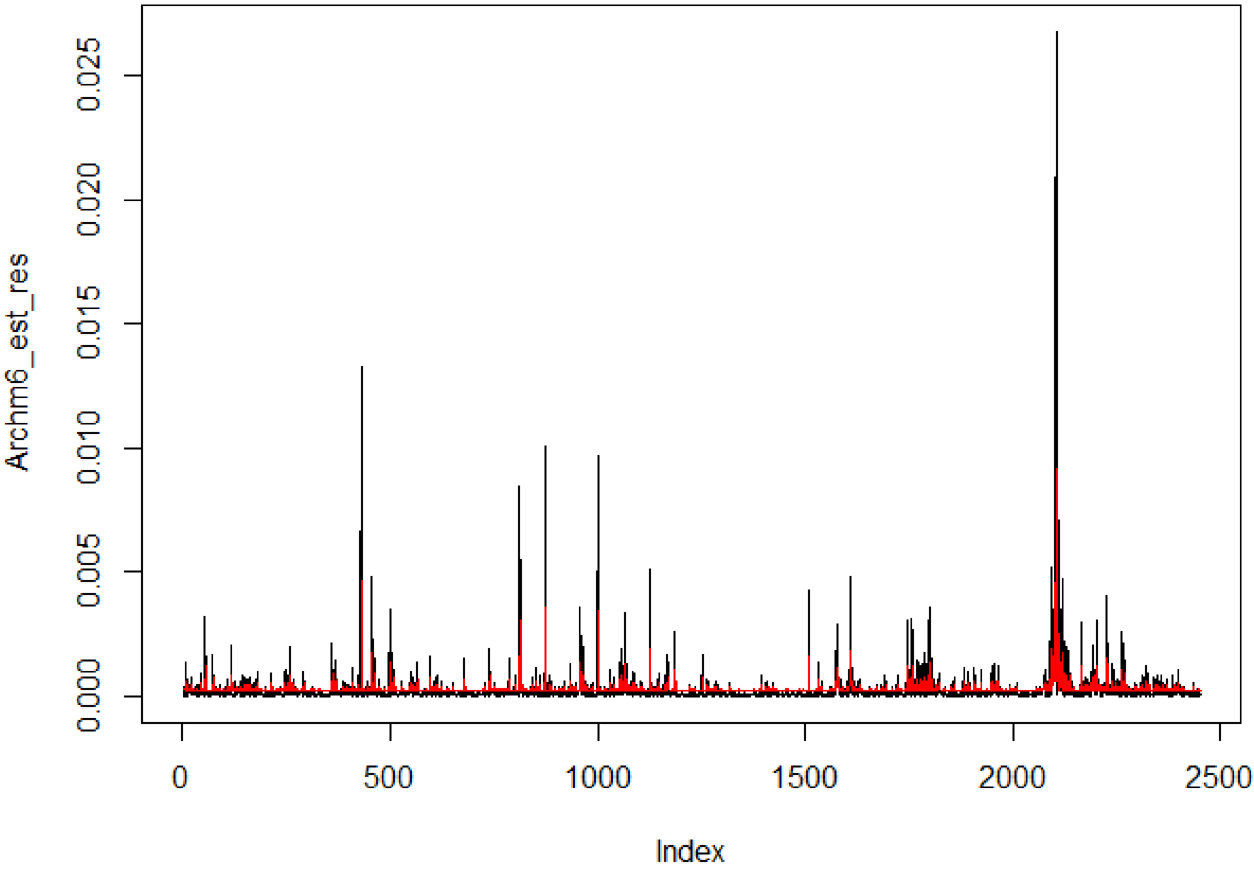
In the displayed results, we see that we have the α_1 estimate that corresponds to the ARCH effects we try to identify. Also, note that we have the $archm$ estimate for the mean analysis. All the estimates are statistically significant at the 1% significance level. The complementary tests are displayed in the annex section. In the next figure, we present all the plotted results of our estimation.


Series with 2 Conditional SD Superimposed
Returns
Time
Series with with 1% VaR Limits
Returns
Time
Conditional SD (vs |returns|)
Volatility
Time
ACF of Observations
ACF
lag
ACF of Squared Observations
ACF
lag
ACF of Absolute Observations
ACF
lag
Cross-Correlations of
Squared vs Actual Observations
ACF
lag
Empirical Density of Standardized Residuals
Median: -0.05 | Mean: -0.0236
normal Density
norm (0,1) Fitted Density
Probability
zseries
norm - QQ Plot
Sample Quantiles
Theoretical Quantiles
ACF of Standardized Residuals
ACF
lag
ACF of Squared Standardized Residuals
ACF
lag
News Impact Curve

In addition, we present the estimation of the variance for the Microsoft in Figure 47.

Figure 47

Plotted estimation of the variance for the Microsoft returns time series (ARCH-M analysis)



Now that we have estimated the disturbance of the variance, we can turn to the forecast of our model in order to build over this estimation a prediction for the future variance of the Microsoft time series. In order to do so, we forecast our ARCH-M model to extract the Sigma values for a 20 period horizon.

R-Code 60

Forecasted Sigma values

```
*------------------------------------*
*       GARCH Model Forecast         *
*------------------------------------*
Model: sGARCH
Horizon: 20
Roll Steps: 0
Out of Sample: 0

0-roll forecast [T0=1976-09-17 02:00:00]:
         Series   Sigma
T+1  -0.0002791 0.01355
T+2   0.0015469 0.01525
T+3   0.0005557 0.01578
T+4   0.0011144 0.01595
T+5   0.0015616 0.01601
T+6   0.0004526 0.01603
T+7   0.0019454 0.01603
T+8   0.0007285 0.01604
T+9   0.0013172 0.01604
T+10  0.0016604 0.01604
T+11  0.0005987 0.01604
T+12  0.0020606 0.01604
T+13  0.0008138 0.01604
T+14  0.0014530 0.01604
T+15  0.0017196 0.01604
T+16  0.0007093 0.01604
T+17  0.0021428 0.01604
T+18  0.0008711 0.01604
T+19  0.0015605 0.01604
T+20  0.0017529 0.01604
```

Figure 48

Plot of forecasted Sigma values

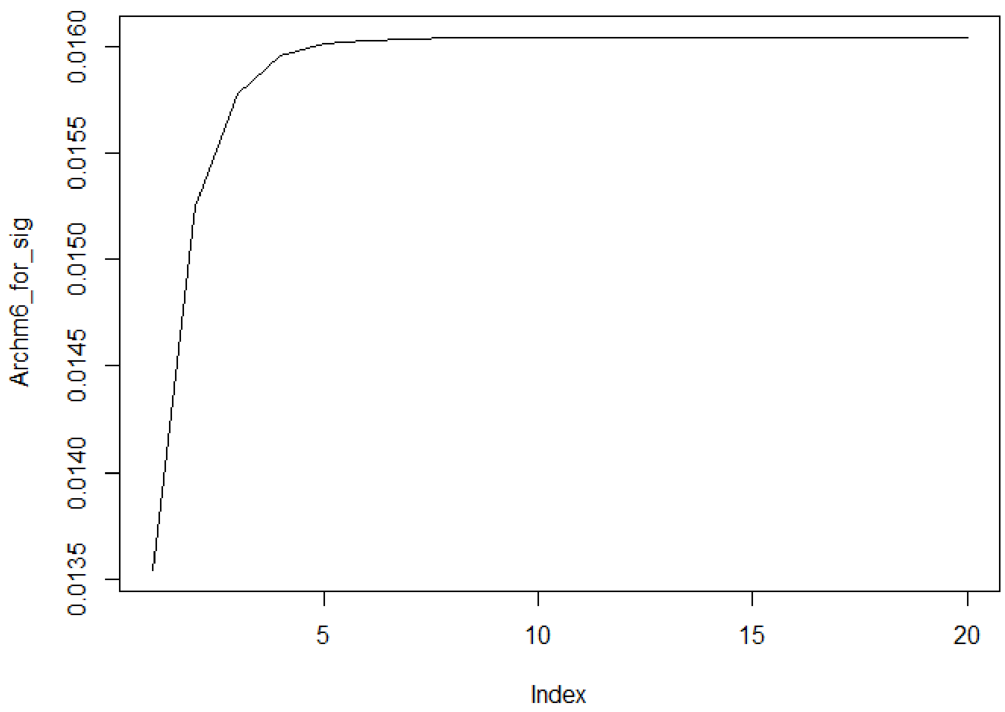
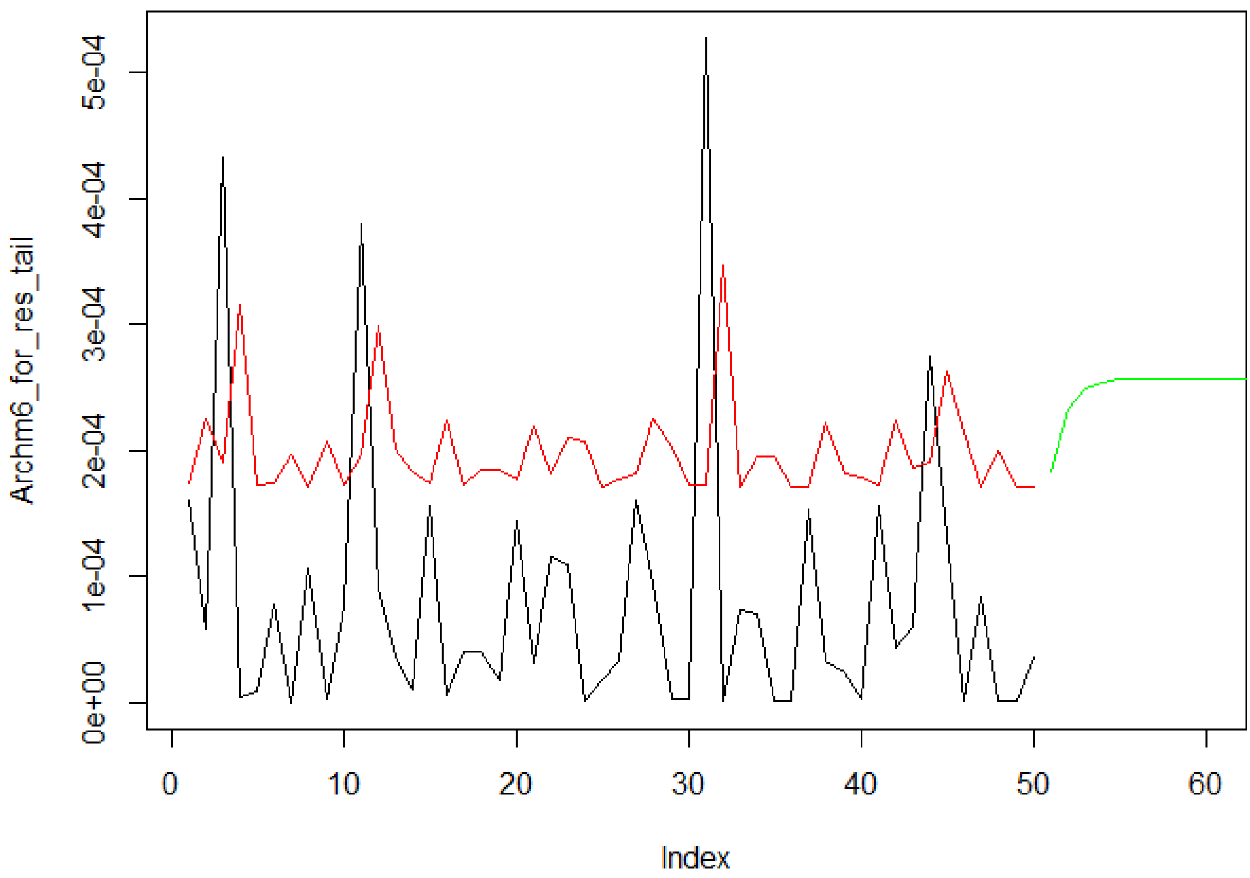


Figure 49
Forecasted variance for the Microsoft time series



As spotted in Figure 48, the sigma displays a sharp upward trend, that is reflected in the forecasted green line in Figure 49. From an economic perspective, this implies that the volatility of the Microsoft returns in the near future should sharply increase.

ANNEXES – EMPIRICAL APPLICATION 4

R-Code 4.1

Standard GARCH model for the S&P500 time series (complementary tests)

```
Information Criteria
------------------------------------

Akaike       -6.8269
Bayes        -6.7961
Shibata      -6.8269
Hannan-Quinn -6.8157

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                         statistic p-value
Lag[1]                      0.8943  0.3443
Lag[2*(p+q)+(p+q)-1][26]   14.1890  0.1218
Lag[4*(p+q)+(p+q)-1][44]   23.2535  0.3959
d.o.f=9
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                   0.009143  0.9238
Lag[2*(p+q)+(p+q)-1][5]  0.915397  0.8785
Lag[4*(p+q)+(p+q)-1][9]  2.782471  0.7943
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]   0.05275 0.500 2.000  0.8183
ARCH Lag[5]   1.98908 1.440 1.667  0.4737
ARCH Lag[7]   3.43216 2.315 1.543  0.4359

Nyblom stability test
------------------------------------
Joint Statistic:  6.6478
Individual Statistics:
mu     0.42510
ar1    0.07064
ar2    0.05756
ar3    0.09557
ar4    0.04370
ma1    0.08275
ma2    0.04127
ma3    0.10857
ma4    0.04492
ma5    0.10849
omega  1.12950
alpha1 0.29565
beta1  0.17792

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         2.89 3.15 3.69
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value      prob sig
Sign Bias           3.8557 1.184e-04 ***
Negative Sign Bias  0.5901 5.551e-01
Positive Sign Bias  0.4214 6.735e-01
Joint Effect       25.6461 1.131e-05 ***


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     114.7    1.076e-15
2    30     122.4    1.961e-13
3    40     136.3    1.031e-12
4    50     166.7    9.676e-15
```

R-Code 4.2

Standard GARCH model for the Google returns time series (complementary tests)

```
Information Criteria
------------------------------------

Akaike       -5.5673
Bayes        -5.5555
Shibata      -5.5673
Hannan-Quinn -5.5630

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.2622  0.6086
Lag[2*(p+q)+(p+q)-1][2]    0.4152  0.9811
Lag[4*(p+q)+(p+q)-1][5]    1.7388  0.7848
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.2457  0.6201
Lag[2*(p+q)+(p+q)-1][5]    0.9890  0.8622
Lag[4*(p+q)+(p+q)-1][9]    1.3173  0.9691
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.2779 0.500 2.000  0.5981
ARCH Lag[5]    0.3071 1.440 1.667  0.9384
ARCH Lag[7]    0.5177 2.315 1.543  0.9768

Nyblom stability test
------------------------------------
Joint Statistic:  1.3125
Individual Statistics:
mu     0.1618
ma1    0.2759
omega  0.1346
alpha1 0.1421
beta1  0.2141

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.28 1.47 1.88
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           0.4199 0.6746
Negative Sign Bias  0.1186 0.9056
Positive Sign Bias  0.4438 0.6572
Joint Effect        0.3448 0.9514


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     159.0    3.917e-24
2    30     177.6    2.783e-23
3    40     184.2    1.004e-20
4    50     203.7    9.134e-21
```

R-Code 4.3

Standard GARCH model for the Microsoft returns time series

(complementary tests)

```
Information Criteria
------------------------------------

Akaike       -5.6316
Bayes        -5.6032
Shibata      -5.6317
Hannan-Quinn -5.6213

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                         statistic p-value
Lag[1]                        2.47 0.11603
Lag[2*(p+q)+(p+q)-1][23]     13.18 0.02522
Lag[4*(p+q)+(p+q)-1][39]     18.85 0.61665
d.o.f=8
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.6077  0.4357
Lag[2*(p+q)+(p+q)-1][5]    1.2789  0.7940
Lag[4*(p+q)+(p+q)-1][9]    2.0332  0.9010
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.6306 0.500 2.000  0.4271
ARCH Lag[5]    1.3319 1.440 1.667  0.6376
ARCH Lag[7]    1.5341 2.315 1.543  0.8145

Nyblom stability test
------------------------------------
Joint Statistic:  2.7993
Individual Statistics:
mu     0.17064
ar1    0.17821
ar2    0.10875
ar3    0.06294
ar4    0.35585
ma1    0.04800
ma2    0.04646
ma3    0.04969
ma4    0.07331
omega  0.19959
alpha1 0.08863
beta1  0.08760

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         2.69 2.96 3.51
Individual Statistic:    0.35 0.47 0.75

 Sign Bias Test
 ------------------------------------
                    t-value   prob sig
 Sign Bias           0.1154 0.9081
 Negative Sign Bias  0.6188 0.5361
 Positive Sign Bias  0.3856 0.6999
 Joint Effect        1.1810 0.7576


 Adjusted Pearson Goodness-of-Fit Test:
 ------------------------------------
   group statistic p-value(g-1)
 1    20     124.5    1.570e-17
 2    30     136.7    6.473e-16
 3    40     152.2    2.737e-15
 4    50     169.9    3.007e-15
```

R-Code 4.4

ARCH-in-mean model for the Microsoft returns time series (complementary tests)

```
Information Criteria
------------------------------------

Akaike       -5.5577
Bayes        -5.5293
Shibata      -5.5578
Hannan-Quinn -5.5474

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                         statistic  p-value
Lag[1]                       1.606 0.205064
Lag[2*(p+q)+(p+q)-1][23]    22.656 0.000000
Lag[4*(p+q)+(p+q)-1][39]    30.564 0.002525
d.o.f=8
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic   p-value
Lag[1]                      1.898 0.1683541
Lag[2*(p+q)+(p+q)-1][2]     7.824 0.0072009
Lag[4*(p+q)+(p+q)-1][5]    14.537 0.0005915
d.o.f=1

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale   P-Value
ARCH Lag[2]     11.83 0.500 2.000 5.814e-04
ARCH Lag[4]     14.97 1.397 1.611 2.622e-04
ARCH Lag[6]     18.70 2.222 1.500 8.272e-05

Nyblom stability test
------------------------------------
Joint Statistic:  2.6005
Individual Statistics:
mu     0.29884
ar1    0.16245
ar2    0.03576
ar3    0.08013
ar4    0.23263
ma1    0.13006
ma2    0.02988
ma3    0.04582
ma4    0.11484
archm  0.34353
omega  0.29016
alpha1 0.94532
Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         2.69 2.96 3.51
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           0.2680 0.7887
Negative Sign Bias  0.5811 0.5613
Positive Sign Bias  0.3469 0.7287
Joint Effect        1.3878 0.7084


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     154.4    3.103e-23
2    30     175.1    7.948e-23
3    40     200.7    1.249e-23
4    50     214.5    1.407e-22
```