NASDAQ Composite

Trabajo Final - Modelos de Volatilidad - Financial Statistics

Alexander J Ohrt

13 mars, 2022

Contents

1 Abstract

The mean and conditional variance of the NASDAQ Composite index are modelized. The work shows that ...

MAKE REFERENCES BETWEEN PLOTS (NUMBERS AND CITATIONS IN TEXT, bookdown is necessary maybe?) LATER.

2 Introduction

Describe

- Scenario and objective of the work. What will be analyzed.
- Precise description of variable (NASDAG Composite) used in the analysis and description of where the data is gathered from (Yahoo Finance)
- Summary of structure of the work (description of what is done in each part)

https://finance.yahoo.com/quote/%5EIXIC?p=%5EIXIC

3 Empirical Application

3.1 Load Data

First, we load the NASDAQ Composite data from Yahoo Finance.

```
getSymbols("^IXIC",from="2012-01-01", to="2022-03-01", warnings = F)
```

```
#> [1] "^IXIC"
```

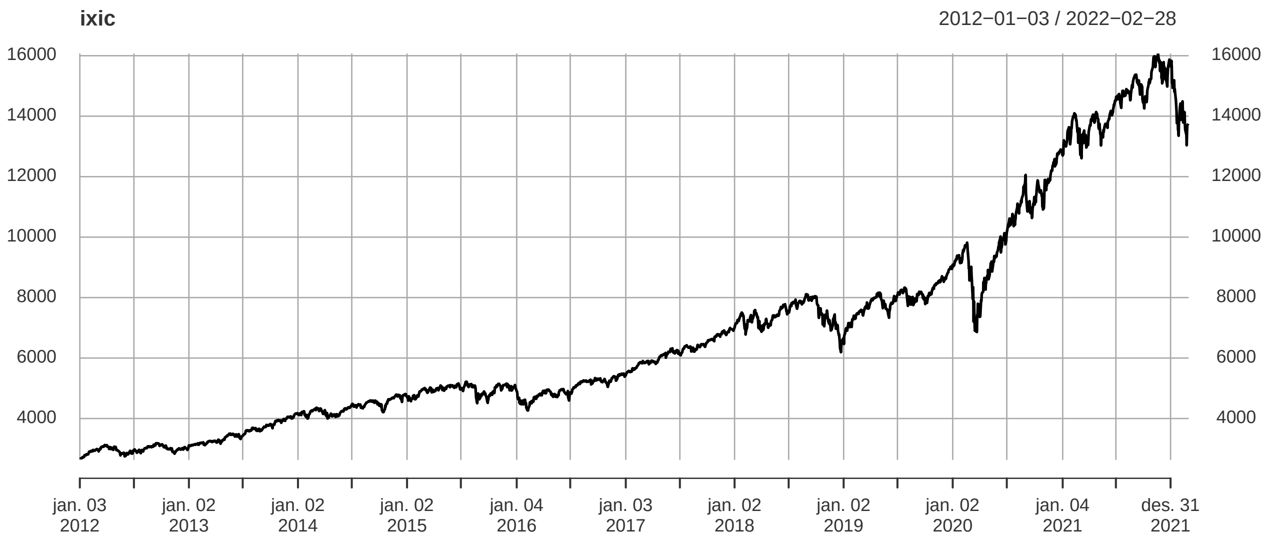
```
dim(IXIC)          # <=== find the size of the data downloaded
```

```
#> [1] 2556    6
```

```
#write.csv(IXIC, file = "IXIC.csv", row.names = F)
#data <- read.csv2("IXIC.csv")

# Want the adjusted closed data.
ixic <- IXIC[,6]
```

The data does not have any NA values (Weekends and holidays have been removed already), we can start working with the data directly.

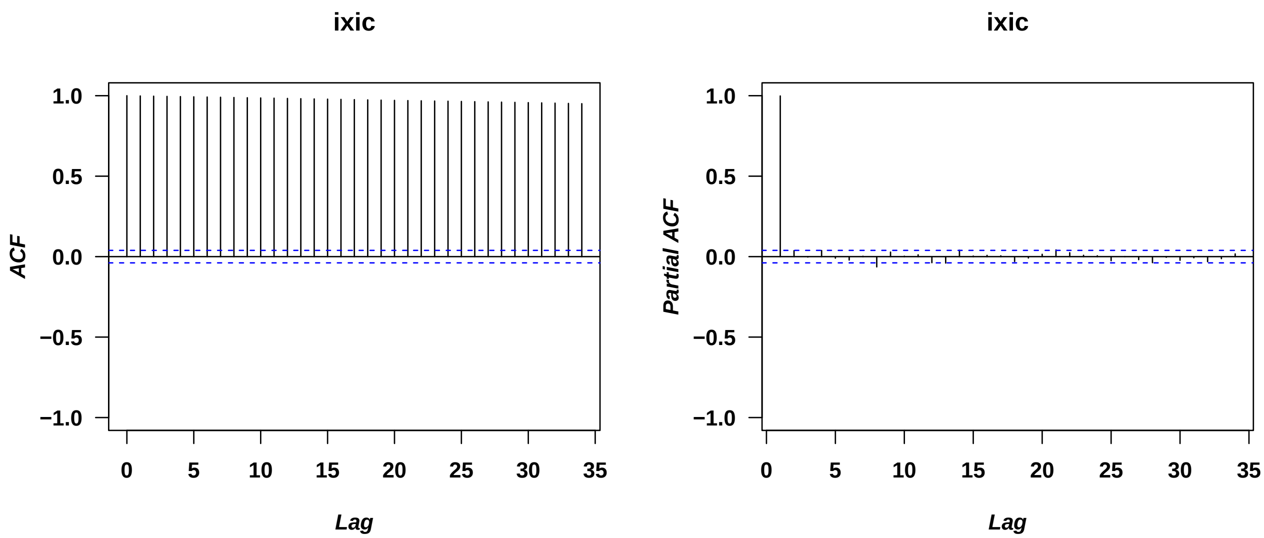


COMMENT: DRAW SOME HAPPENINGS IN THE SERIES (Covid March 2020 and Russia-Ukraine in Feb 2022 + leading up to Feb in the beginning of 2022). Did something happen in Jan 2019 in the US (with tech-companies?) Did something happen in Jan 2016 (small regression).

3.2 Analysis of Stationarity

In order to see if the series is stationary, we will employ both informal and formal tests. Immediately, by looking at the plot above (reference later), the series does not look stationary, since the mean of the process looks to change quite dramatically with time. Some more informal tests are done. The function of autocorrelation and partial autocorrelation (empirical) for the series are plotted below.

```
par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
acf(ixic,ylim=c(-1,1),main="ixic")
pacf(ixic,ylim=c(-1,1),main="ixic")
```



As is seen from the function of autocorrelation (ACF), the coefficients decrease slowly. This suggests that the time series is non-stationary, since a stationary series would show exponentially decreasing coefficients in the ACF.

SJEKK AT ALT DETTE GIR MENING (OG BRUKER KORREKTE BEGREPER) SENERE! (TIL SLUTT)

Next, some Ljung-Box tests are done. Here we are testing the joint hypothesis that all m of the correlation coefficients are simultaneously equal to zero. Below we are testing for $m \in \{1, 5, 10, 15, 20\}$.

MAKE A SUMMARY-TABLE HERE INSTEAD LATER! JUST SHOW THE LAGS AND THE P-VALUES.

```
Box.test(ixic, lag = 1, type = c("Ljung-Box"))
```

```
#>
#>  Box-Ljung test
#>
#> data:  ixic
#> X-squared = 2551.4, df = 1, p-value < 2.2e-16
```

```
Box.test(ixic, lag = 5, type = c("Ljung-Box"))
```

```
#>
#>  Box-Ljung test
#>
#> data:  ixic
#> X-squared = 12698, df = 5, p-value < 2.2e-16
```

```
Box.test(ixic, lag = 10, type = c("Ljung-Box"))
```

```
#>
#>  Box-Ljung test
#>
#> data:  ixic
#> X-squared = 25246, df = 10, p-value < 2.2e-16
```

```
Box.test(ixic, lag = 15, type = c("Ljung-Box"))
```

```
#>
#>  Box-Ljung test
#>
#> data:  ixic
```

```
#> X-squared = 37633, df = 15, p-value < 2.2e-16
```

```
Box.test(ixic, lag = 20, type = c("Ljung-Box"))
```

```
#>
#>  Box-Ljung test
#>
#> data:  ixic
#> X-squared = 49855, df = 20, p-value < 2.2e-16
```

All the p-values from the Ljung-Box tests are low, which means that we would reject the null hypothesis that all m correlation coefficients are simultaneously equal to zero. This further suggests that the series is non-stationary.

Next, some formal tests are done to check stationarity of the series. First, the Augmented-Dickey-Fuller (ADF) unit root test is done. The null hypothesis for this case states that the series is integrated of order 1, i.e. that it is non-stationary. Below, the ADF test is done assuming both a stochastic and deterministic trend in the data. The maximum number of lags considered are 20 and the number of lags used are chosen by BIC.

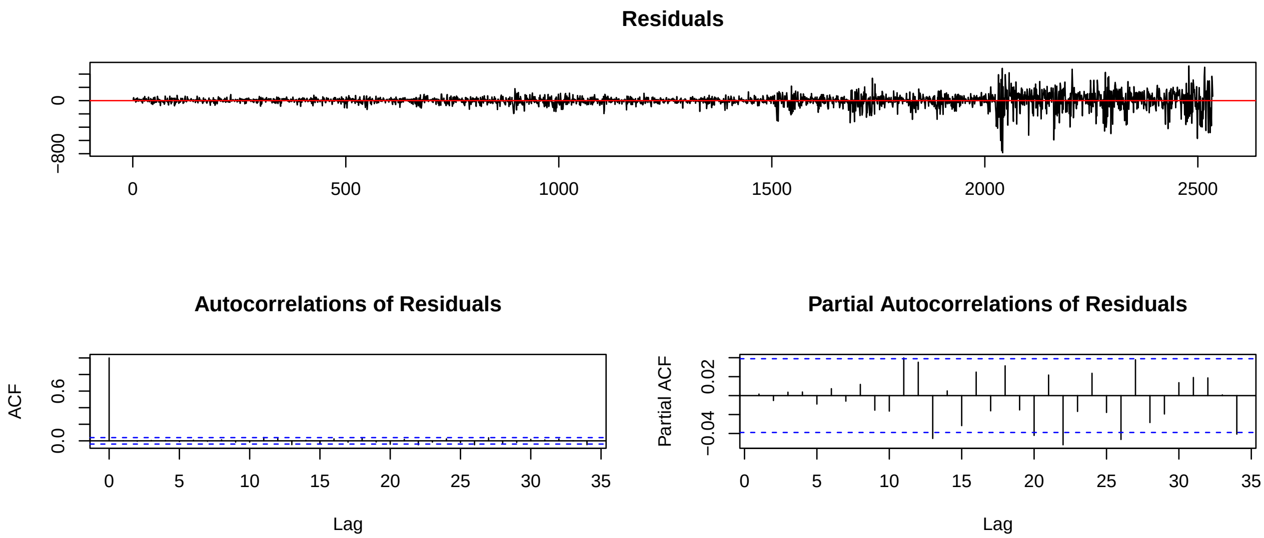
```
ixic.df<-ur.df(ixic, type = c("trend"), lags=20, selectlags = c("BIC"))
summary(ixic.df)
```

```
#>
#> ###############################################
#> # Augmented Dickey-Fuller Test Unit Root Test #
#> ###############################################
#>
#> Test regression trend
#>
#>
#> Call:
#> lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
#>
#> Residuals:
#>     Min      1Q  Median      3Q     Max
#> -785.57  -28.42    3.81   35.46  523.61
#>
#> Coefficients:
#>              Estimate Std. Error t value Pr(>|t|)
#> (Intercept)  4.013754   4.319321   0.929  0.35285
#> z.lag.1     -0.002490   0.001457  -1.709  0.08751 .
#> tt           0.013739   0.006845   2.007  0.04482 *
#> z.diff.lag1 -0.090268   0.019879  -4.541 5.87e-06 ***
#> z.diff.lag2  0.051188   0.019870   2.576  0.01005 *
#> z.diff.lag3 -0.004620   0.019903  -0.232  0.81646
#> z.diff.lag4 -0.062694   0.019939  -3.144  0.00168 **
#> z.diff.lag5  0.007115   0.019994   0.356  0.72197
#> z.diff.lag6 -0.026554   0.019982  -1.329  0.18401
#> z.diff.lag7  0.084723   0.020069   4.222 2.51e-05 ***
#> z.diff.lag8 -0.104673   0.020111  -5.205 2.10e-07 ***
#> z.diff.lag9  0.062169   0.020173   3.082  0.00208 **
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 97.42 on 2523 degrees of freedom
#> Multiple R-squared:  0.05128,    Adjusted R-squared:  0.04715
#> F-statistic:  12.4 on 11 and 2523 DF,  p-value: < 2.2e-16
```

```
#>
#>
#> Value of test-statistic is: -1.7093 3.3031 2.0816
#>
#> Critical values for test statistics:
#>       1pct  5pct 10pct
#> tau3 -3.96 -3.41 -3.12
#> phi2  6.09  4.68  4.03
#> phi3  8.27  6.25  5.34
```

From the output it is apparent that BIC chooses 9 lags in the DF test. Moreover, the value of the test-statistic clearly suggests that we cannot reject the null-hypothesis, since the value is much larger than the critical values for this left-sided test. Thus, we would conclude that the series is non-stationary. Note that the test leads to the same conclusion when assuming no trends and when assuming only a drift. Moreover, the same amount of lags are chosen for all three variants. Below, the residuals and the autocorrelation functions of the residuals are plotted, in order to check if the number of lags chosen via BIC is satisfactory.

```
plot(ixic.df)
```

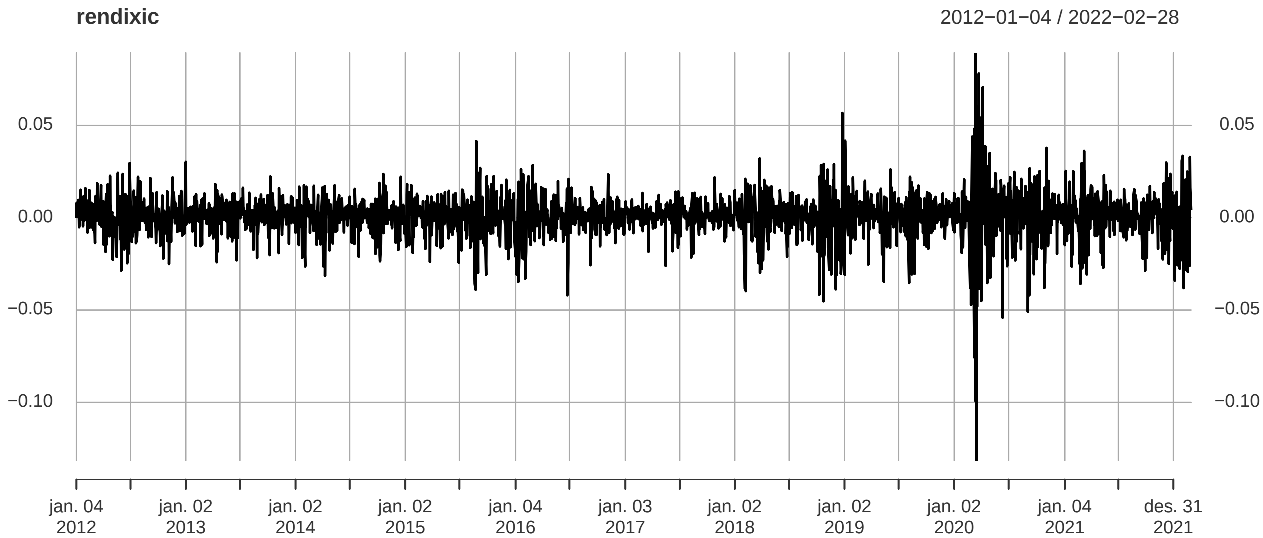


The autocorrelation function of the residuals has no significant coefficients, which leads us to conclude that the amount of lags chosen via BIC is satisfactory.

Next, we check if the returns (rendimientos) are stationary. Below we calculate the returns and remove the first difference, since it is not a numerical value.

```
rendixic <- diff(log(ixic))
rendixic <- rendixic[-1] # The first difference is NA, needs to be removed.
```

```
plot(rendixic)
```



Then, the ADF test is calculated without trends, since there does not look to be any trends in the plot of the returns. Note that, as earlier, the conclusion of the test and the amount of lags that are chosen via BIC are the same when assuming a drift or both types of trends.

```
rendixic.df<-ur.df(rendixic, type = c("none"), lags=20, selectlags = c("BIC"))
summary(rendixic.df)
```

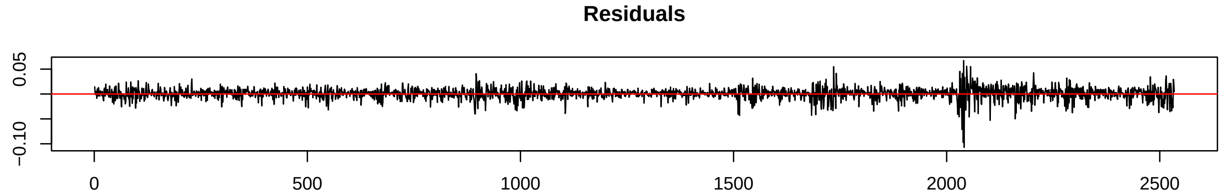
```
#>
#> ###############################################
#> # Augmented Dickey-Fuller Test Unit Root Test #
#> ###############################################
#>
#> Test regression none
#>
#>
#> Call:
#> lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
#>
#> Residuals:
#>       Min        1Q    Median        3Q       Max
#> -0.107159 -0.004275  0.001424  0.006875  0.067084
#>
#> Coefficients:
#>             Estimate Std. Error t value Pr(>|t|)
#> z.lag.1     -1.07192    0.06437 -16.651  < 2e-16 ***
#> z.diff.lag1 -0.02421    0.06026  -0.402 0.687881
#> z.diff.lag2  0.02301    0.05659   0.407 0.684281
#> z.diff.lag3  0.02650    0.05193   0.510 0.609830
#> z.diff.lag4 -0.02609    0.04723  -0.552 0.580728
#> z.diff.lag5 -0.01914    0.04188  -0.457 0.647707
#> z.diff.lag6 -0.06599    0.03630  -1.818 0.069208 .
#> z.diff.lag7  0.02827    0.02966   0.953 0.340668
#> z.diff.lag8 -0.06861    0.01996  -3.437 0.000597 ***
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.01167 on 2525 degrees of freedom
#> Multiple R-squared:  0.5826, Adjusted R-squared:  0.5811
```

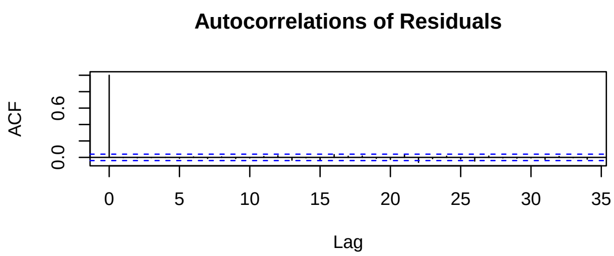
```
#> F-statistic: 391.6 on 9 and 2525 DF,  p-value: < 2.2e-16
#>
#>
#> Value of test-statistic is: -16.6512
#>
#> Critical values for test statistics:
#>       1pct  5pct 10pct
#> tau1 -2.58 -1.95 -1.62
```

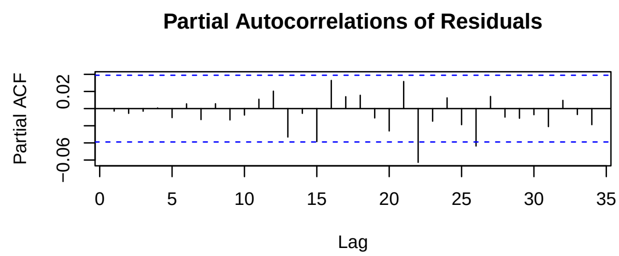
It is apparent that 8 lags are chosen. Moreover, from the test-statistic above we would reject the null-hypothesis, which means that we have found evidence against the hypothesis that the returns are $I(1)$, i.e. evidence against the hypothesis that the original series is $I(0)$. Thus, we conclude that the returns are $I(0)$ or the original series is $I(1)$. This means that, through the results from this test, the original series is not stationary (which we have seen earlier), but the returns are stationary and can be used further in the analysis. SJEKK AT DET JEG SKRIVER HER STEMMER, TROR DET GJØR DET! HAR NOEN NOTATER FRA ET EKSEMPEL PÅ TAVLA PÅ DETTE!

As earlier, the plot below shows that the amount of lags for the ADF test chosen via BIC is satisfactory.

```r
plot(rendixic.df)
```







For completeness, we also use the Philips-Perron (PP) test to check stationarity of the series. This test defines the same null-hypothesis as the ADF test, which means that this is a left-tailed test as well. COULD MAKE A TABLE FROM THESE TWO LAST TESTS (WITH THE MOST IMPORTANT STATISTICS), TO SAVE SOME ROOM IF NEEDED, SINCE THE CONCLUSIONS ARE THE SAME AS EARLIER (AS EXPECTED).

```r
ixic.pp<-ur.pp(ixic, type = c("Z-tau"), model = c("trend"), lags = c("long"))
summary(ixic.pp)
```

```
#>
#> ###################################
#> # Phillips-Perron Unit Root Test #
#> ###################################
#>
#> Test regression with intercept and trend
#>
#>
#> Call:
```

```
#> lm(formula = y ~ y.l1 + trend)
#>
#> Residuals:
#>     Min      1Q  Median      3Q     Max
#> -983.02  -26.63    2.68   34.73  658.52
#>
#> Coefficients:
#>              Estimate Std. Error t value Pr(>|t|)
#> (Intercept) 23.065091  10.220234   2.257   0.0241 *
#> y.l1         0.997252   0.001472 677.449   <2e-16 ***
#> trend        0.014404   0.006891   2.090   0.0367 *
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 99.37 on 2552 degrees of freedom
#> Multiple R-squared:  0.9992, Adjusted R-squared:  0.9992
#> F-statistic: 1.543e+06 on 2 and 2552 DF,  p-value: < 2.2e-16
#>
#>
#> Value of test-statistic, type: Z-tau  is: -1.669
#>
#>            aux. Z statistics
#> Z-tau-mu              2.9272
#> Z-tau-beta            1.9503
#>
#> Critical values for Z statistics:
#>                      1pct      5pct     10pct
#> critical values -3.967077 -3.414184 -3.128848
```

All combinations of trend assumptions and/or long or short lags yield the same conclusions as from the output above; namely that we have not found sufficient evidence to reject the null-hypothesis of non-stationarity of the series. Below the PP-test is done with the returns.

```r
rendixic.pp<-ur.pp(rendixic, type = c("Z-tau"), model = c("constant"), lags = c("short"))
summary(rendixic.pp)
```

```
#>
#> ##################################
#> # Phillips-Perron Unit Root Test #
#> ##################################
#>
#> Test regression with intercept
#>
#>
#> Call:
#> lm(formula = y ~ y.l1)
#>
#> Residuals:
#>       Min        1Q    Median        3Q       Max
#> -0.120202 -0.004652  0.000702  0.006097  0.076985
#>
#> Coefficients:
#>               Estimate Std. Error t value Pr(>|t|)
#> (Intercept)  0.0007315  0.0002344   3.121  0.00182 **
#> y.l1        -0.1345383  0.0196155  -6.859 8.68e-12 ***
```

```
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 0.01183 on 2552 degrees of freedom
#> Multiple R-squared:  0.0181, Adjusted R-squared:  0.01772
#> F-statistic: 47.04 on 1 and 2552 DF,  p-value: 8.683e-12
#>
#>
#> Value of test-statistic, type: Z-tau  is: -57.7836
#>
#>          aux. Z statistics
#> Z-tau-mu            3.1176
#>
#> Critical values for Z statistics:
#>                       1pct      5pct     10pct
#> critical values -3.435853 -2.863173 -2.567664
```

When referring to the returns, the conclusion is the same as for the ADF test; the returns are stationary while the original series is not.

Finally, we use the KPSS test to check stationarity of the series. The null hypothesis for this test states that the series is stationary. In the test below we have chosen to assume the deterministic component as a constant with a linear trend, and we have used short lags. Note that the conclusion is the same with all different variations of assumptions for the test.

```
ixic.kpss<-ur.kpss(ixic, type = c("tau"), lags = c("short"))
summary(ixic.kpss)
```

```
#>
#> #######################
#> # KPSS Unit Root Test #
#> #######################
#>
#> Test is of type: tau with 8 lags.
#>
#> Value of test-statistic is: 4.8701
#>
#> Critical value for a significance level of:
#>                  10pct  5pct 2.5pct  1pct
#> critical values 0.119 0.146  0.176 0.216
```

Since this is a right-tailed test, the test-statistic is clearly sufficiently large to reject the null-hypothesis to the lowest significance level shown (0.01). Thus, we conclude that the series is non-stationary, as expected. The test below shows that the returns are stationary, in line with what we have concluded earlier, since we cannot find strong evidence against the null-hypothesis.

```
rendixic.kpss <- ur.kpss(rendixic, type = c("mu"), lags = c("short"))
summary(rendixic.kpss)
```

```
#>
#> #######################
#> # KPSS Unit Root Test #
#> #######################
#>
#> Test is of type: mu with 8 lags.
#>
#> Value of test-statistic is: 0.0313
```

```
#>
#> Critical value for a significance level of:
#>                 10pct  5pct 2.5pct  1pct
#> critical values 0.347 0.463  0.574 0.739
```

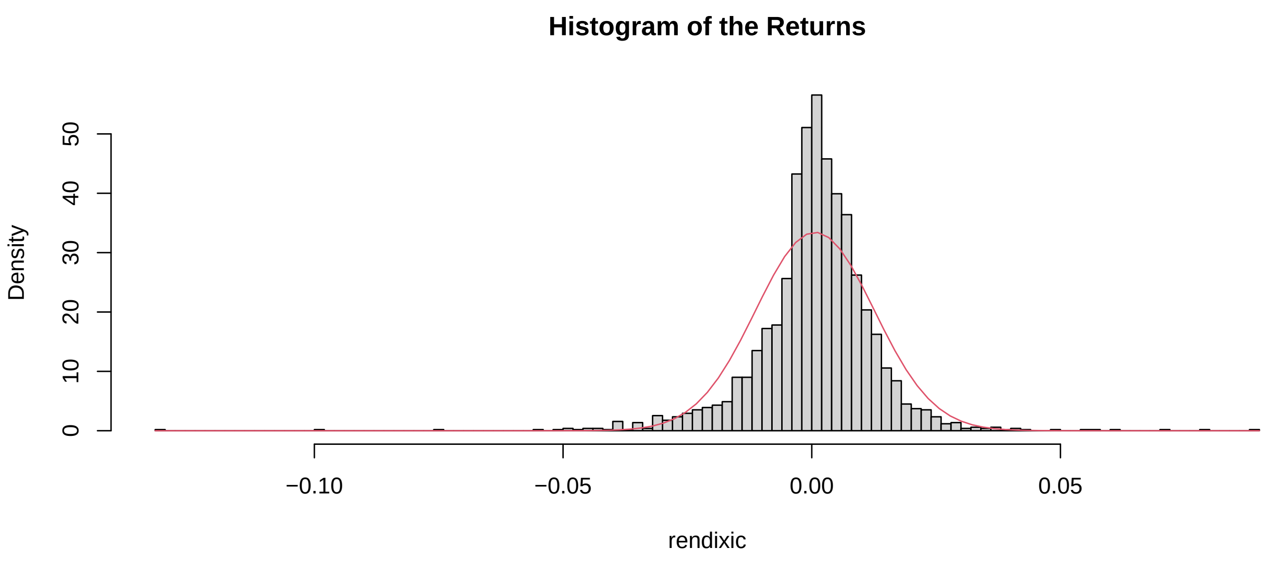
Conclusively, the original time series is not stationary, but the returns are stationary, which means that the returns will be used in the following analysis. We can be relatively certain that this is the case, since all three formal tests, as well as the informal tests, point to this conclusion.

3.3 Basic Statistical Properties of the Stationary Series

Some basic statistical properties of the stationary series, the returns, are shown below.

```
basicStats(rendixic)
```

```
#>              IXIC.Adjusted
#> nobs           2555.000000
#> NAs               0.000000
#> Minimum          -0.131492
#> Maximum           0.089347
#> 1. Quartile      -0.003980
#> 3. Quartile       0.006759
#> Mean              0.000645
#> Median            0.001093
#> Sum               1.647064
#> SE Mean           0.000236
#> LCL Mean          0.000182
#> UCL Mean          0.001108
#> Variance          0.000142
#> Stdev             0.011933
#> Skewness         -0.841975
#> Kurtosis         12.309021
```

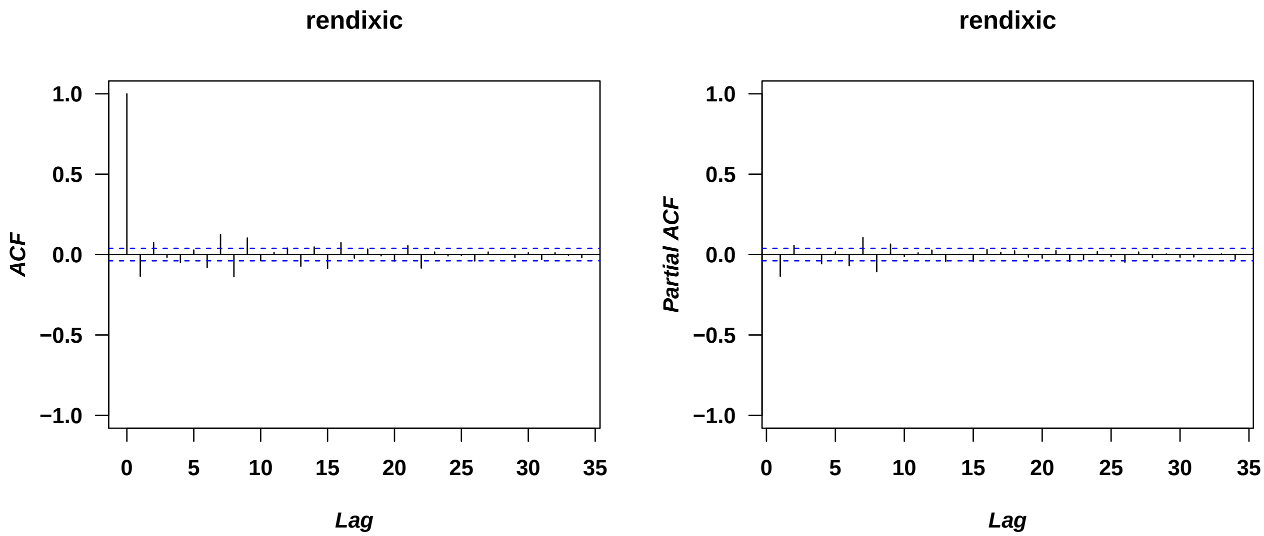


We can see that the series is Leptokurtic, both by the kurtosis value and from the histogram above. The superposed red curve is a Gaussian distribution with empirical mean and standard error according to the returns of the NASDAQ Composite series. Moreover, the skewness is negative, which means that the distribution of the returns are heavy-tailed in the left tail. This is also apparent from the histogram above.

3.4 Identification, Estimation and Diagnostics of a Model for the Mean

Plotting the autocorrelation functions of the returns.

```
par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
acf(rendixic,ylim=c(-1,1),main="rendixic")
pacf(rendixic,ylim=c(-1,1),main="rendixic")
```



It looks like both the ACF and the partial ACF (PACF) have 9 significant coefficients. This seems like a large order of ARMA-model to estimate, so I will try with smaller models instead. Note that the third coefficient of both ACF and PACF seems to be non-significant, which might be a hint to what order of model would be fitting. The table below shows the BIC and the AIC for different orders of ARMA-models.

Table 1: AIC and BIC of different estimated models for the returns of NASDAQ Composite

Model	BIC	AIC
AR(1)	-15420.249041659	-15402.7116191511
MA(1)	-15414.9737924899	-15397.436369982
ARMA(1,1)	-15424.2940143262	-15400.9107843157
AR(2)	-15426.5374072231	-15403.1541772126
MA(2)	-15426.312921927	-15402.9296919165
ARMA(2,2)	-15505.4106732282	-15470.3358282124
AR(3)	-15424.5375105372	-15395.308473024
MA(3)	-15426.644706053	-15397.4156685398

HVA BETYR P-VERDIER PÅ 0.000e+00? DETTE FÅR JEG FOR ARMA(2,2)!! SE NEDENFOR!

The table above clearly shows that ARMA(2,2) yields the lowest AIC and BIC. Moreover, the estimated coefficients of the model are significant to a level of 0.05. SIKKER!?!? The estimated model is shown below

```
mean.model <- arima(rendixic, order = c(2,0,2),include.mean = TRUE)
mean.model
```

```
#>
#> Call:
#> arima(x = rendixic, order = c(2, 0, 2), include.mean = TRUE)
#>
#> Coefficients:
```

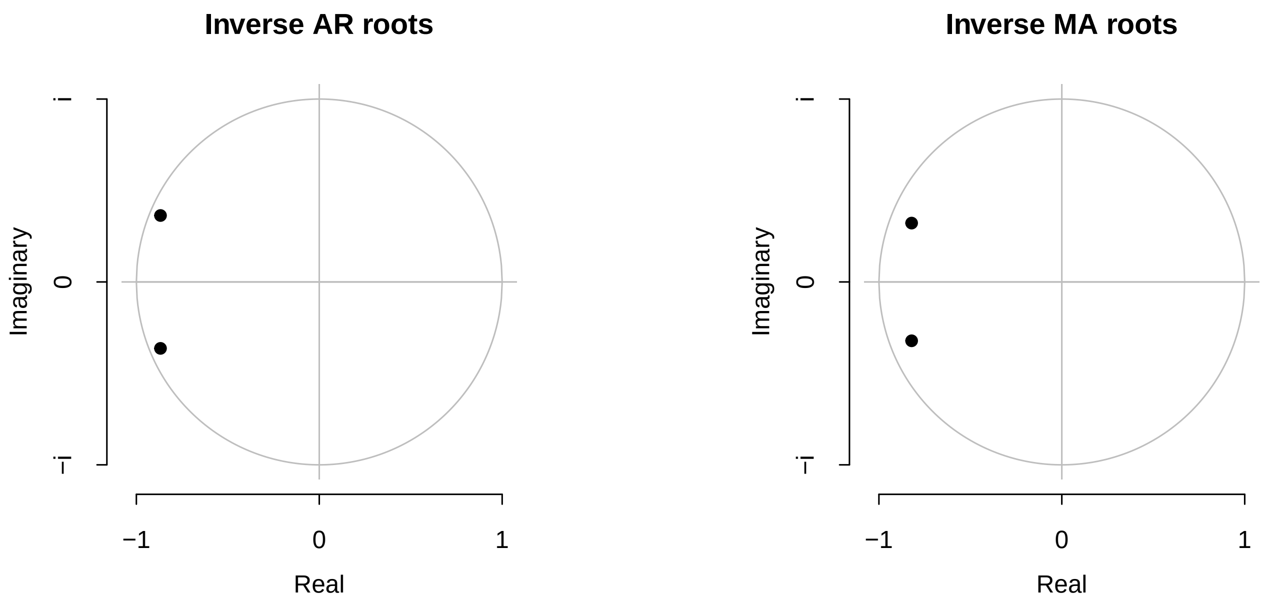
```
#>            ar1      ar2     ma1    ma2  intercept
#>        -1.7362  -0.8856  1.6425  0.778      6e-04
#> s.e.    0.0241   0.0226  0.0326  0.030      2e-04
#>
#> sigma^2 estimated as 0.0001349:  log likelihood = 7758.71,  aic = -15505.41
```

```
pnorm(c(abs(mean.model$coef)/sqrt(diag(mean.model$var.coef))), mean=0, sd=1, lower.tail=FALSE)
```

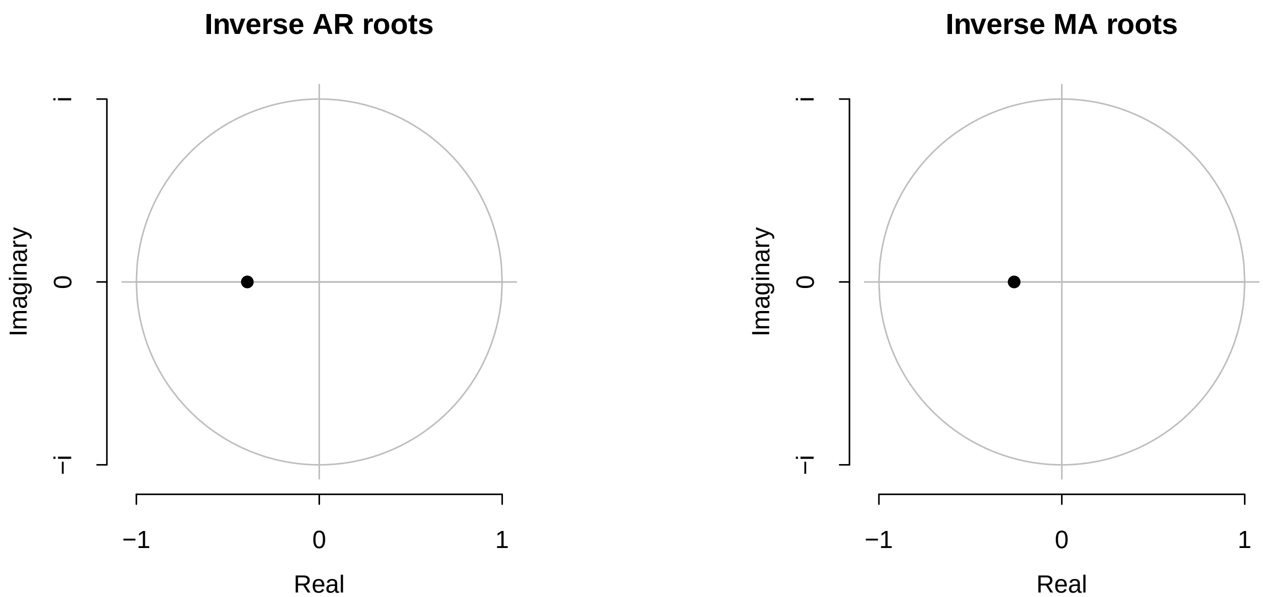
```
#>           ar1           ar2           ma1           ma2     intercept
#>  0.000000e+00  0.000000e+00  0.000000e+00 7.280516e-149  1.595550e-03
```

Some model diagnostics follow. We must check if the model is stationary.

```
plot(mean.model)
```



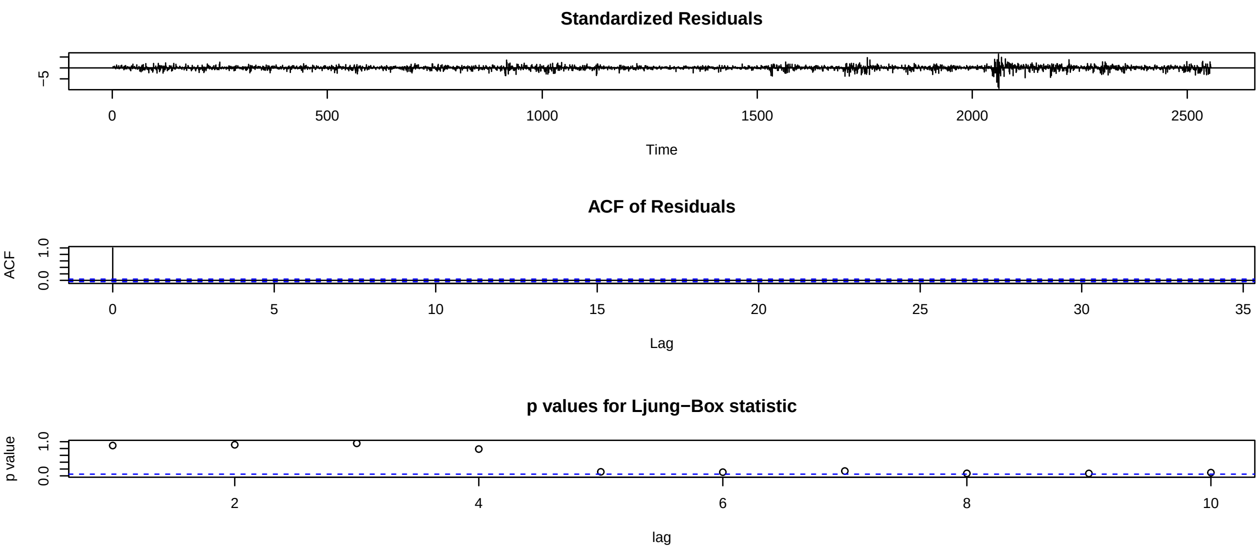

```
plot(model3) # Or just use model3, because of parsimony? The AIC is much larger though.
```



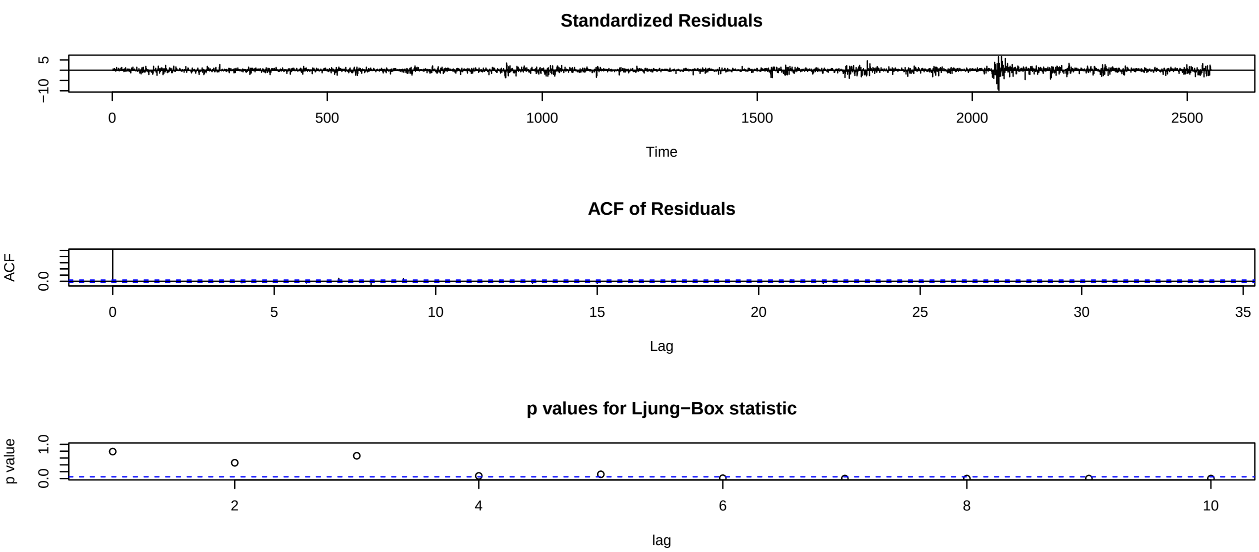
The inverse roots of the characteristic polynomial of AR and MA are shows above. DO WE WANT ALL OF THE INVERSE ROOTS TO FALL INSIDE, FOR BOTH THE AR AND THE MA PROCESS? The stationarity condition for the AR-process is satisfied, since the roots have absolute values greater than one. Moreover, the invertibility condition holds for the MA process, since the roots of this process also have absolute values greater than one. Thus, the model is stationary.

The residuals of the model are analyzed below.

```
# Denne kodeblokken trengs kanskje ikke?
tsdiag(mean.model)
```

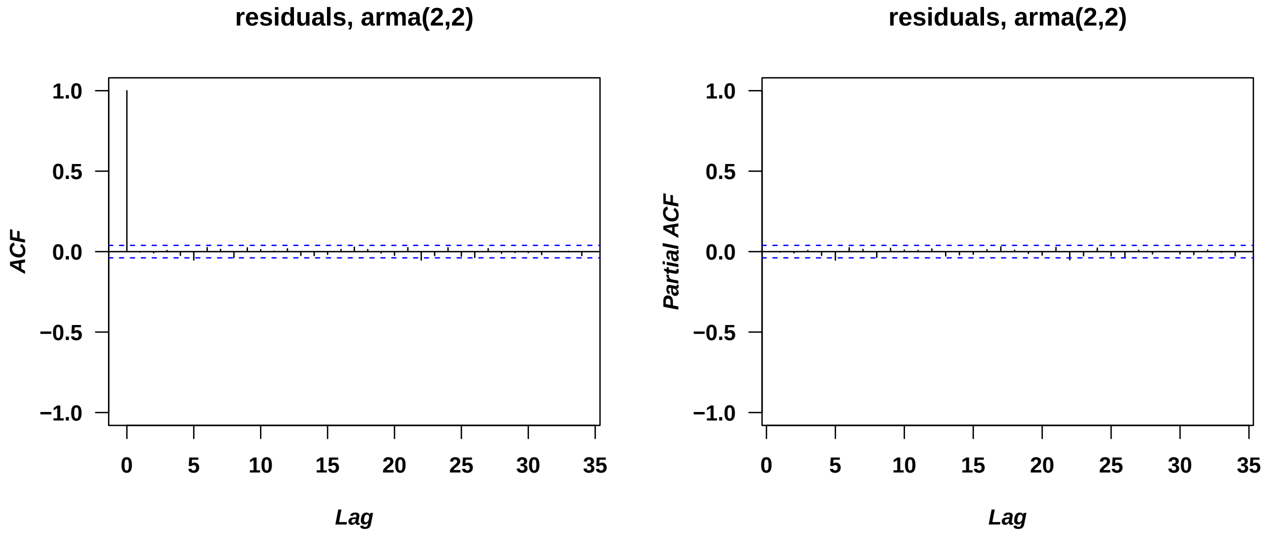



```
tsdiag(model3) # Or just use model3, because of parsimony?
```



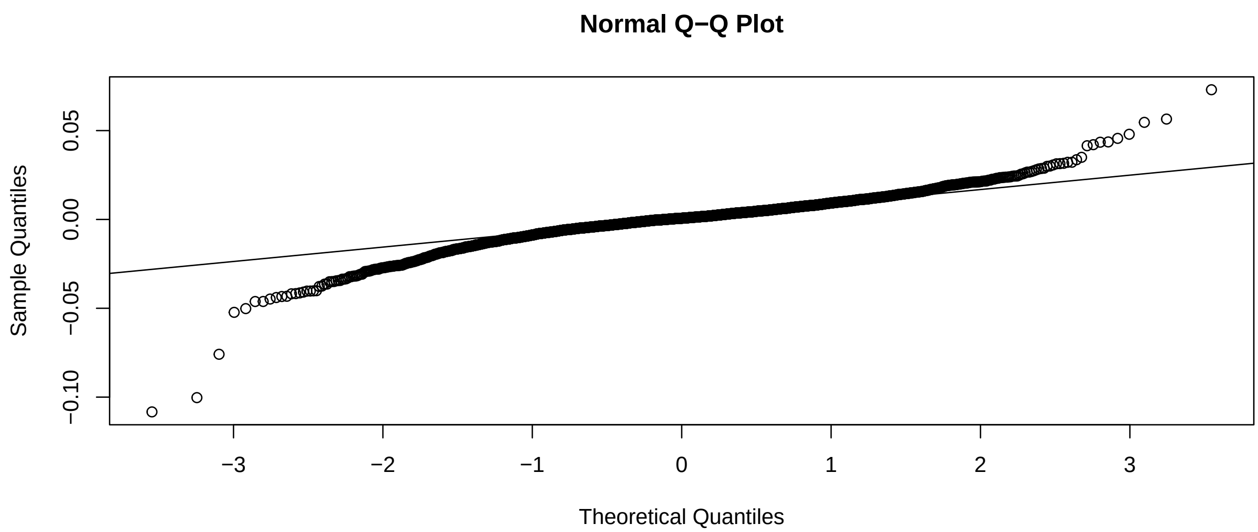
There are no significant coefficients in the autocorrelation function above, which suggests that the model has adequately captured the information in the data. Moreover, the Ljung-Box statistic p-values are all relatively large, which means that we will not reject the Ljung-Box null hypothesis that all m of the correlation coefficients are simultaneously equal to zero. Thus, this further suggests that the residuals are not correlated and we have found a model that seems reasonable in this regard.

```
par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
acf(mean.model$residuals,ylim=c(-1,1),main="residuals, arma(2,2)")
pacf(mean.model$residuals,ylim=c(-1,1),main="residuals, arma(2,2)")
```

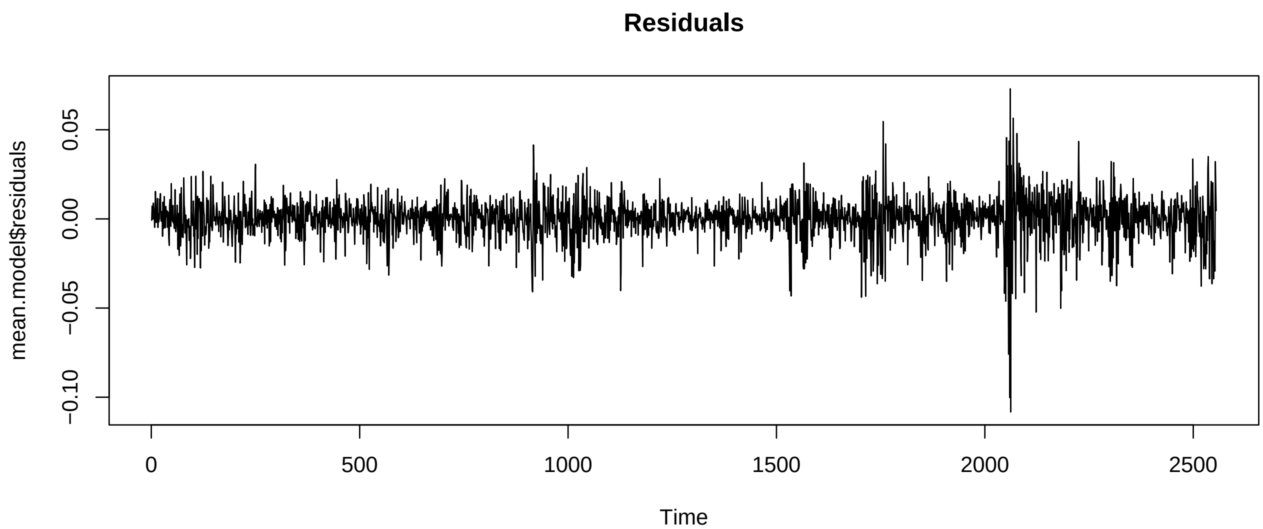


The same conclusion is made from the ACF and PACF plotted above.

```
qqnorm(mean.model$residuals)
qqline(mean.model$residuals, datax = FALSE)
```




```
plot(mean.model$residuals)
title (main="Residuals")
```




```
normalTest(mean.model$residuals,method="jb")
```

```
#>
#> Title:
#>  Jarque - Bera Normalality Test
#>
#> Test Results:
#>   STATISTIC:
#>     X-squared: 7472.6779
#>   P VALUE:
#>     Asymptotic p Value: < 2.2e-16
#>
#> Description:
#>  Sun Mar 13 22:31:46 2022 by user: ajo
```

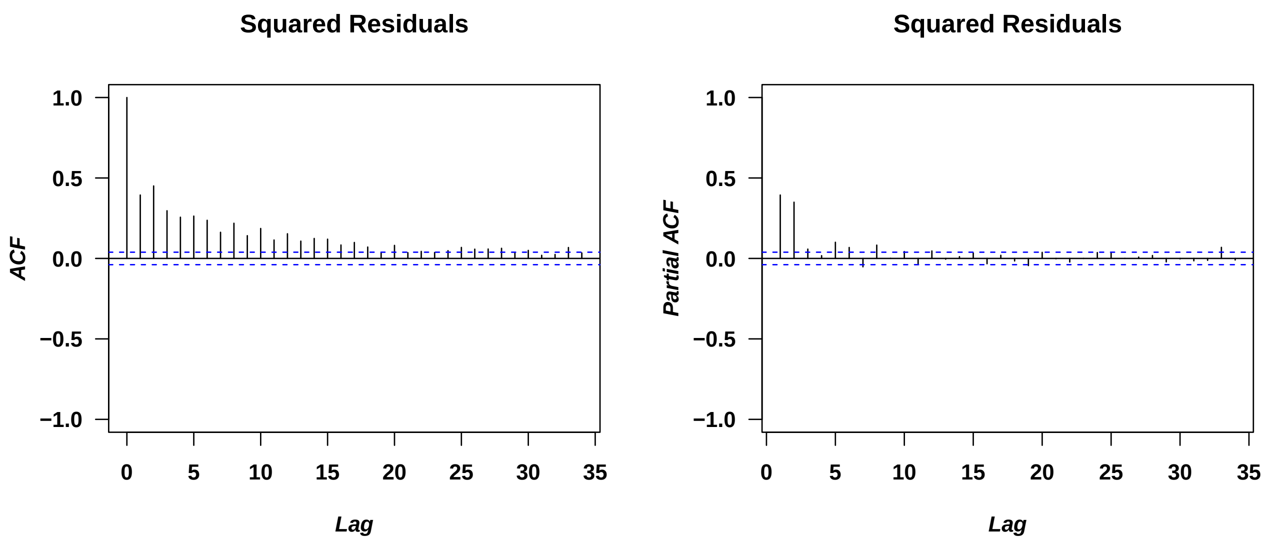
A QQ-plot of the residuals is plotted below, followed by the residuals (not standardized) are plotted above. It is apparent that the residuals have heavy tails. It is not reasonable to assume normality of the residuals, an argument that the Jarque-Bera Normality test further substantiates because its null hypothesis of normality is rejected following the very small p-value.

3.5 Identification, Estimation and Diagnostics of a Model for the Variance

HVA ER KODEN / METODEN NEDENFOR GODT FOR? VIRKER SOM OM DETTE BLIR BRUKT I "IDENTIFICACION" AV GARCH MODELL, MEN HVORDAN?

```
residuals <- mean.model$residuals
residuals2 <- residuals^2

par(mfrow=c(1,2),font=2,font.lab=4,font.axis=2,las=1)
acf(residuals2,ylim=c(-1,1),main="Squared Residuals")
pacf(residuals2,ylim=c(-1,1),main="Squared Residuals")
```




```r
# Har acf ovenfor noe sammenheng med plottet nedenfor (se slides 33++ i Volatility Models)?
# For de ser veldig like ut. Men vet ikke helt om residuals^2 fra modellen har noe med dette å gjøre?
# Mulig det er fordi vi estimerer mean vha ARMA-modellen og trekker den fra + kvadrat, noe vi egt gjør o
acf((rendixic-mean(rendixic))^2, ylim = c(-1,1))

Box.test(residuals2,lag=1,type='Ljung')
```

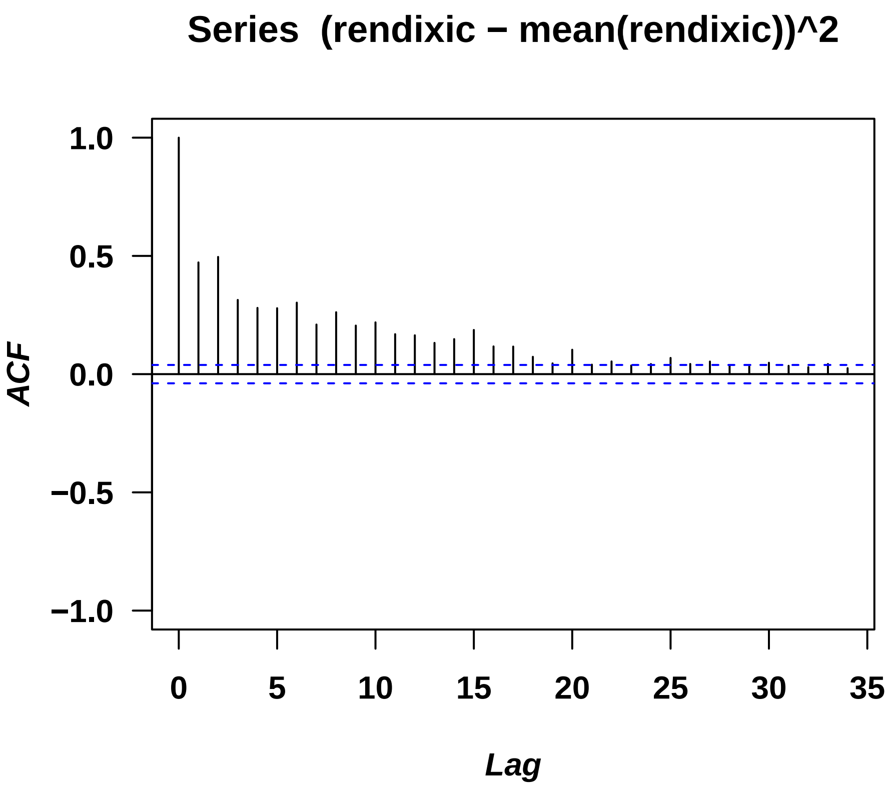
```
#>
#>  Box-Ljung test
#>
#> data:  residuals2
#> X-squared = 397.36, df = 1, p-value < 2.2e-16
```

```r
Box.test(residuals2,lag=5,type='Ljung')
```

```
#>
#>  Box-Ljung test
#>
#> data:  residuals2
#> X-squared = 1487.7, df = 5, p-value < 2.2e-16
```

```r
Box.test(residuals2,lag=15,type='Ljung')
```

```
#>
#>  Box-Ljung test
#>
#> data:  residuals2
#> X-squared = 2163.8, df = 15, p-value < 2.2e-16
```



The Ljung-Box tests above lead to a conclusion that the squared residuals are correlated.

Now over to estimation of GARCH models for the variance of the returns. First we estimate a ARMA(2,2)-GARCH(1,1) with a t-student distribution. WHY CHOSE A T-STUDENT OVER A NORMAL (OR VICE VERSA)?

```
spec1 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1,1)), mean.model=list(armaOr
(m <- ugarchfit(spec = spec1, data = rendixic))
```

```
#>
#> *---------------------------------*
#> *          GARCH Model Fit        *
#> *---------------------------------*
#>
#> Conditional Variance Dynamics
#> -----------------------------------
#> GARCH Model  : sGARCH(1,1)
#> Mean Model   : ARFIMA(2,0,2)
#> Distribution : std
#>
#> Optimal Parameters
#> ------------------------------------
#>         Estimate  Std. Error  t value Pr(>|t|)
#> mu      0.001230    0.000085  14.4538 0.000000
#> ar1     0.411215    0.010896  37.7392 0.000000
#> ar2     0.468396    0.002407 194.5656 0.000000
#> ma1    -0.467394    0.010241 -45.6404 0.000000
#> ma2    -0.455194    0.009398 -48.4364 0.000000
#> omega   0.000005    0.000002   2.2329 0.025555
#> alpha1  0.163393    0.020408   8.0062 0.000000
```

```
#> beta1   0.813486    0.023304  34.9081 0.000000
#> shape   5.240364    0.590405   8.8759 0.000000
#>
#> Robust Standard Errors:
#>         Estimate  Std. Error   t value Pr(>|t|)
#> mu      0.001230    0.000101  12.15400  0.00000
#> ar1     0.411215    0.026259  15.65987  0.00000
#> ar2     0.468396    0.026987  17.35607  0.00000
#> ma1    -0.467394    0.016152 -28.93808  0.00000
#> ma2    -0.455194    0.015014 -30.31821  0.00000
#> omega   0.000005    0.000005   0.95667  0.33873
#> alpha1  0.163393    0.020752   7.87359  0.00000
#> beta1   0.813486    0.039750  20.46509  0.00000
#> shape   5.240364    0.726690   7.21127  0.00000
#>
#> LogLikelihood : 8262.277
#>
#> Information Criteria
#> ------------------------------------
#>
#> Akaike       -6.4605
#> Bayes        -6.4399
#> Shibata      -6.4605
#> Hannan-Quinn -6.4530
#>
#> Weighted Ljung-Box Test on Standardized Residuals
#> ------------------------------------
#>                          statistic p-value
#> Lag[1]                       1.450  0.2286
#> Lag[2*(p+q)+(p+q)-1][11]     3.676  1.0000
#> Lag[4*(p+q)+(p+q)-1][19]     8.810  0.6701
#> d.o.f=4
#> H0 : No serial correlation
#>
#> Weighted Ljung-Box Test on Standardized Squared Residuals
#> ------------------------------------
#>                         statistic p-value
#> Lag[1]                    0.02101  0.8847
#> Lag[2*(p+q)+(p+q)-1][5]   0.92104  0.8773
#> Lag[4*(p+q)+(p+q)-1][9]   2.69363  0.8083
#> d.o.f=2
#>
#> Weighted ARCH LM Tests
#> ------------------------------------
#>             Statistic Shape Scale P-Value
#> ARCH Lag[3]    0.1391 0.500 2.000  0.7092
#> ARCH Lag[5]    2.1473 1.440 1.667  0.4396
#> ARCH Lag[7]    3.0338 2.315 1.543  0.5072
#>
#> Nyblom stability test
#> ------------------------------------
#> Joint Statistic:  1.8626
#> Individual Statistics:
#> mu     0.37609
```

```
#> ar1    0.06996
#> ar2    0.12329
#> ma1    0.07412
#> ma2    0.11176
#> omega  0.24242
#> alpha1 0.64422
#> beta1  0.45667
#> shape  0.47015
#>
#> Asymptotic Critical Values (10% 5% 1%)
#> Joint Statistic:          2.1 2.32 2.82
#> Individual Statistic:     0.35 0.47 0.75
#>
#> Sign Bias Test
#> ------------------------------------
#>                    t-value     prob sig
#> Sign Bias           2.3681 0.017953  **
#> Negative Sign Bias  0.1143 0.909045
#> Positive Sign Bias  0.9489 0.342756
#> Joint Effect       15.5615 0.001395 ***
#>
#>
#> Adjusted Pearson Goodness-of-Fit Test:
#> ------------------------------------
#>   group statistic p-value(g-1)
#> 1    20     66.58    3.371e-07
#> 2    30     89.80    3.904e-08
#> 3    40     96.76    8.200e-07
#> 4    50    118.17    1.208e-07
#>
#>
#> Elapsed time : 0.6370857
```

We observe that all the parameter estimates are significant to a 5% significance level. Moreover, we note that the condition of positivity holds, because $\hat{\alpha}_1 > 0$ and $\hat{\beta}_1$, where we follow the standard statistical notation of a hat indicating an estimate. Also, we note that the condition of stationarity holds, because $\hat{\alpha}_1 + \hat{\beta}_1 < 1$.

```
#plot(m) # Vet ikke om denne er noe interessant?
```

Estimated volatilities.

```
head(v <- sigma(m)) # Estimated volatility.
```

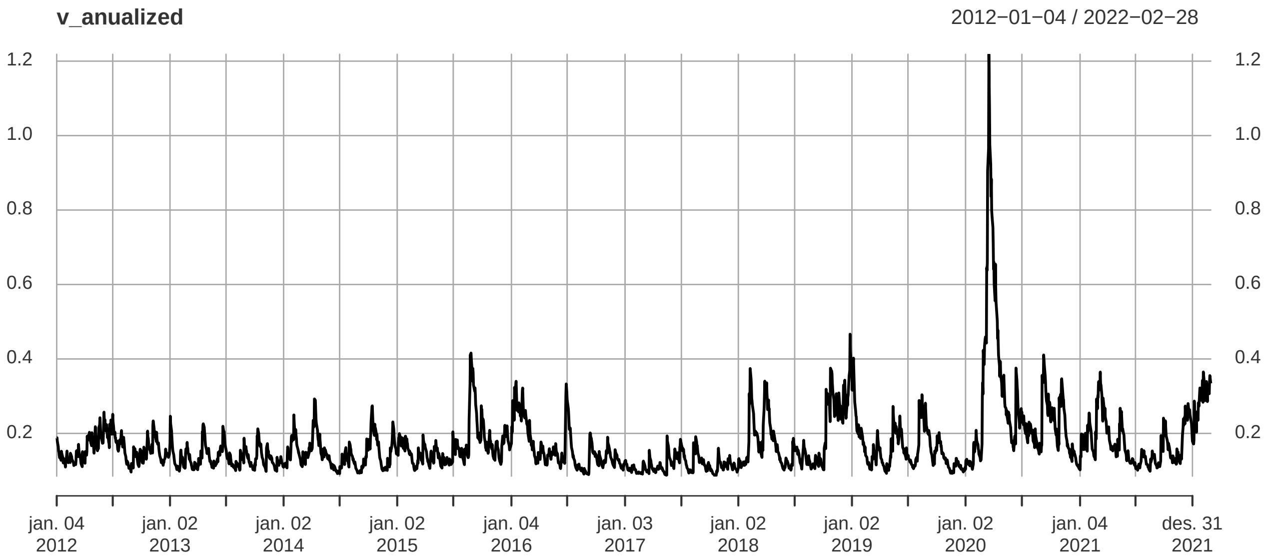
```
#>                   [,1]
#> 2012-01-04 0.011905236
#> 2012-01-05 0.011905236
#> 2012-01-06 0.011291248
#> 2012-01-09 0.010408316
#> 2012-01-10 0.009625509
#> 2012-01-11 0.009594342
```

```
head(v_anualized <- (250)^0.5*v) # Anualized estimated volatility. .
```

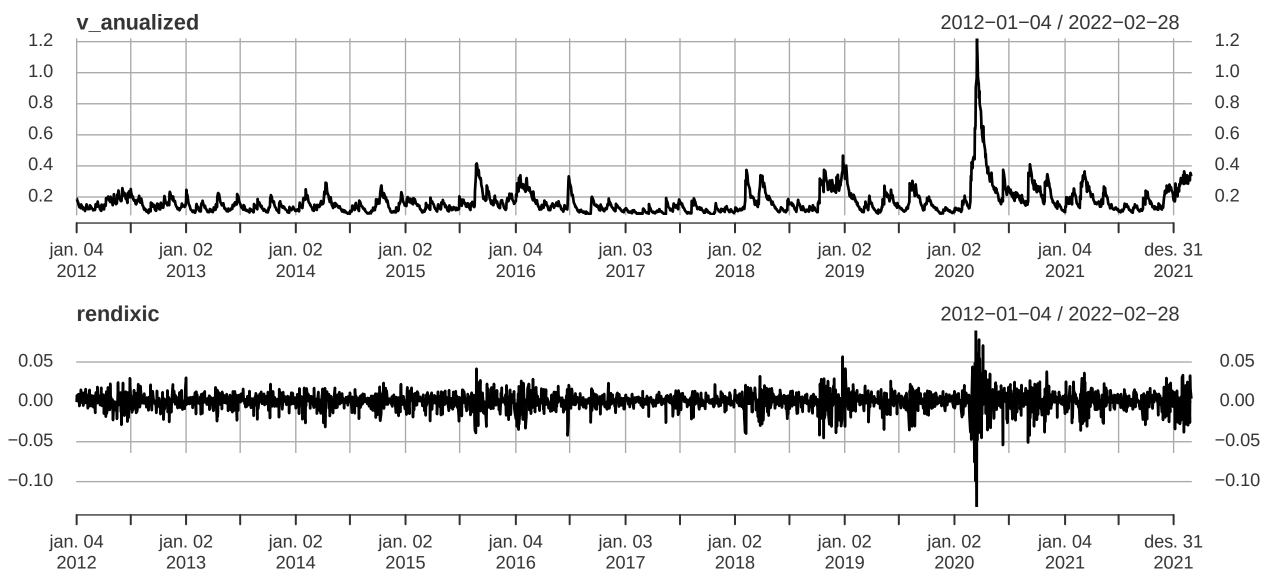
```
#>                 [,1]
#> 2012-01-04 0.1882383
#> 2012-01-05 0.1882383
#> 2012-01-06 0.1785303
```

```
#> 2012-01-09 0.1645699
#> 2012-01-10 0.1521927
#> 2012-01-11 0.1516999
```

```
plot(v_anualized)
```

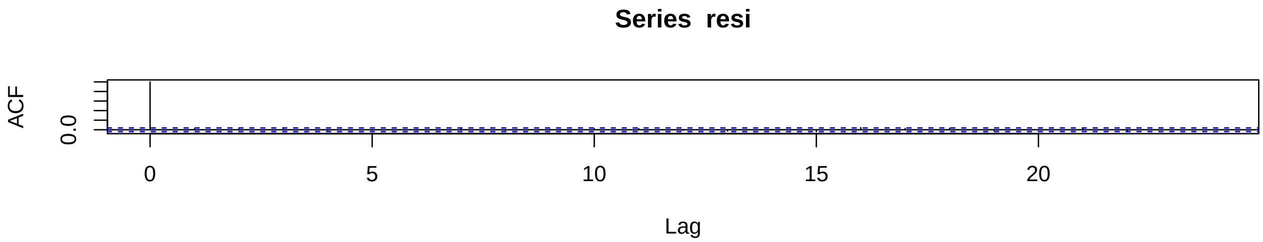


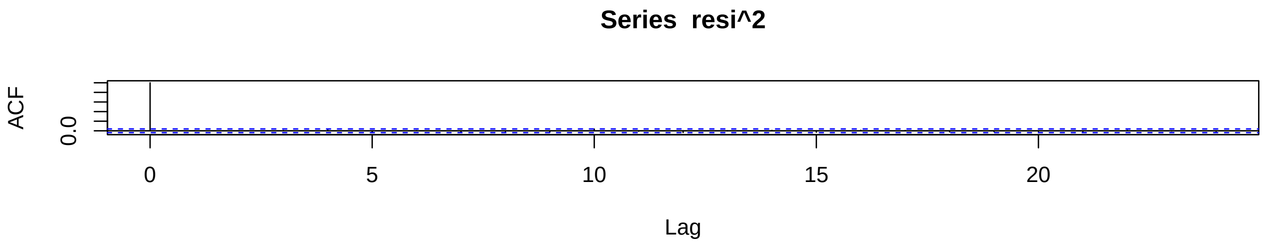

```
par(mfcol=c(2,1))  # Show volatility and returns simultaneously.
plot(v_anualized)
plot(rendixic)
```



The residuals of the model are found SURELY SOME RESIDUAL ANALYSIS IS NECESSARY FOR THESE MODELS AS WELL?

```
resi <- residuals(m,standardize=T) # Standardized residuals
par(mfcol=c(2,1)) # Obtain ACF & PACF
acf(resi,lag=24)
acf(resi^2,lag=24)
```





There is nothing significant in the plots above.

Next we will fit a ARMA(2,2)-GJR-GARCH(1,1) model, assuming a t-distribution. WHY DO THIS?

```
spec2 <- ugarchspec(variance.model=list(model="gjrGARCH", garchOrder = c(1,1)), mean.model=list(armaOrde
(mgjr <- ugarchfit(spec = spec2, data = rendixic))
```

```
#>
#> *---------------------------------*
#> *          GARCH Model Fit        *
#> *---------------------------------*
#>
#> Conditional Variance Dynamics
#> -----------------------------------
#> GARCH Model  : gjrGARCH(1,1)
#> Mean Model   : ARFIMA(2,0,2)
#> Distribution : std
#>
#> Optimal Parameters
#> ------------------------------------
#>         Estimate  Std. Error   t value Pr(>|t|)
#> mu      0.001047    0.000138  7.606420 0.000000
#> ar1     0.275227    0.319606  0.861146 0.389157
#> ar2     0.470193    0.205718  2.285615 0.022277
#> ma1    -0.320370    0.321325 -0.997029 0.318750
#> ma2    -0.455726    0.214102 -2.128549 0.033292
#> omega   0.000005    0.000000 15.301373 0.000000
#> alpha1  0.000000    0.009195  0.000002 0.999999
#> beta1   0.823974    0.014252 57.815933 0.000000
#> gamma1  0.260477    0.031890  8.168122 0.000000
#> shape   5.594916    0.591457  9.459543 0.000000
#>
#> Robust Standard Errors:
#>         Estimate  Std. Error   t value Pr(>|t|)
#> mu      0.001047    0.000130  8.037487 0.000000
#> ar1     0.275227    0.196628  1.399736 0.161593
#> ar2     0.470193    0.119749  3.926475 0.000086
#> ma1    -0.320370    0.198050 -1.617626 0.105743
```

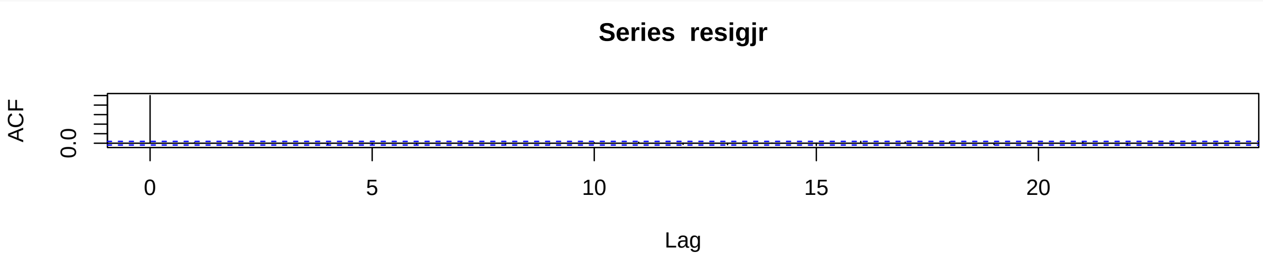
```
#> ma2    -0.455726    0.122730 -3.713240 0.000205
#> omega   0.000005    0.000001  9.596872 0.000000
#> alpha1  0.000000    0.012423  0.000001 0.999999
#> beta1   0.823974    0.013038 63.195840 0.000000
#> gamma1  0.260477    0.035268  7.385755 0.000000
#> shape   5.594916    0.550880 10.156318 0.000000
#>
#> LogLikelihood : 8303.976
#>
#> Information Criteria
#> ------------------------------------
#>
#> Akaike       -6.4923
#> Bayes        -6.4695
#> Shibata      -6.4924
#> Hannan-Quinn -6.4841
#>
#> Weighted Ljung-Box Test on Standardized Residuals
#> ------------------------------------
#>                          statistic p-value
#> Lag[1]                      0.2822  0.5953
#> Lag[2*(p+q)+(p+q)-1][11]    3.0139  1.0000
#> Lag[4*(p+q)+(p+q)-1][19]    8.8996  0.6552
#> d.o.f=4
#> H0 : No serial correlation
#>
#> Weighted Ljung-Box Test on Standardized Squared Residuals
#> ------------------------------------
#>                         statistic p-value
#> Lag[1]                     0.2543  0.6141
#> Lag[2*(p+q)+(p+q)-1][5]    0.5938  0.9423
#> Lag[4*(p+q)+(p+q)-1][9]    1.6828  0.9393
#> d.o.f=2
#>
#> Weighted ARCH LM Tests
#> ------------------------------------
#>             Statistic Shape Scale P-Value
#> ARCH Lag[3]   0.06952 0.500 2.000  0.7920
#> ARCH Lag[5]   0.85359 1.440 1.667  0.7768
#> ARCH Lag[7]   1.59629 2.315 1.543  0.8019
#>
#> Nyblom stability test
#> ------------------------------------
#> Joint Statistic:  17.2754
#> Individual Statistics:
#> mu     0.7414
#> ar1    0.1534
#> ar2    0.3645
#> ma1    0.1807
#> ma2    0.3876
#> omega  1.3411
#> alpha1 2.3979
#> beta1  0.9204
#> gamma1 0.8213
```

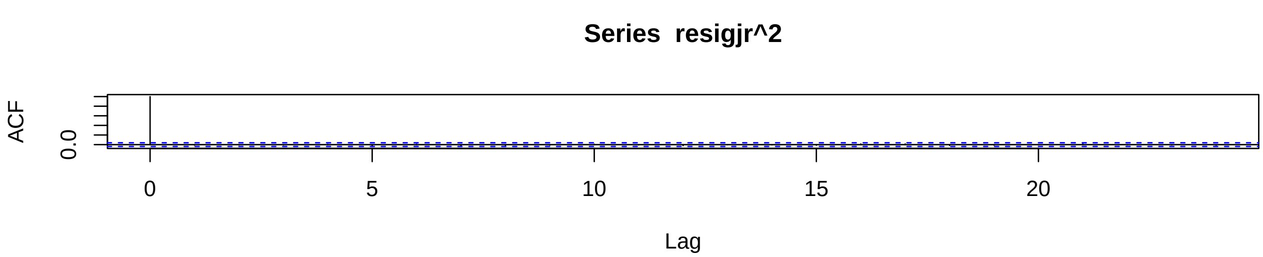
```
#> shape  0.3773
#>
#> Asymptotic Critical Values (10% 5% 1%)
#> Joint Statistic:          2.29 2.54 3.05
#> Individual Statistic:     0.35 0.47 0.75
#>
#> Sign Bias Test
#> ------------------------------------
#>                     t-value   prob sig
#> Sign Bias            1.2277 0.2197
#> Negative Sign Bias  1.2552 0.2095
#> Positive Sign Bias  0.2208 0.8253
#> Joint Effect         2.7117 0.4383
#>
#>
#> Adjusted Pearson Goodness-of-Fit Test:
#> ------------------------------------
#>   group statistic p-value(g-1)
#> 1    20     70.26    8.326e-08
#> 2    30     85.76    1.611e-07
#> 3    40    109.44    1.324e-08
#> 4    50    129.36    3.569e-09
#>
#>
#> Elapsed time : 1.022597
#plot(mgjr)
```

Still, the parameter estimates are significant (except for `omega`). The positivity and the stationarity conditions also hold here, since all the parameter estimates are positive THIS SEEMS TO BE FALSE SINCE alpha is 0! and $\hat{\alpha}_1 + \hat{\beta}_1 + \frac{1}{2}\hat{\gamma} \approx 0.954 < 1$.

Residuals.

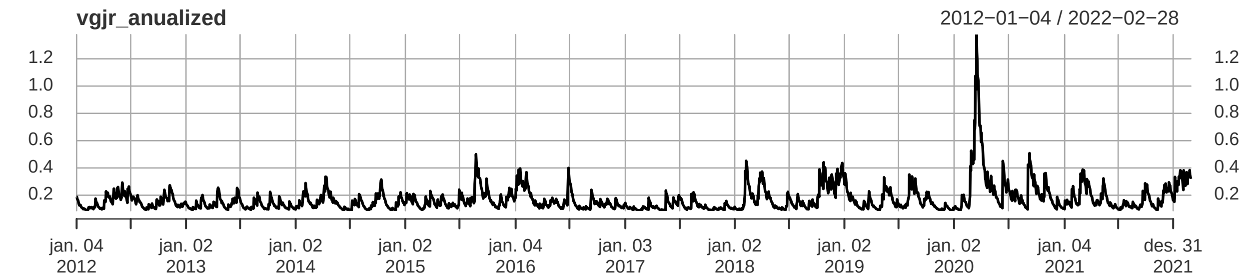
```
resigjr <- residuals(mgjr,standardize=T) # Standardized residuals
par(mfcol=c(2,1)) # Obtain ACF & PACF
acf(resigjr,lag=24)
acf(resigjr^2,lag=24)
```

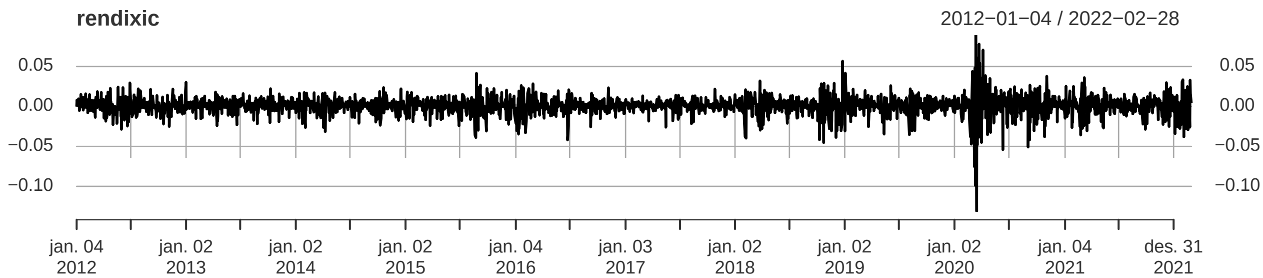




There is nothing significant in the plots above.

```
vgjr <- sigma(mgjr) # Estimated volatility.
vgjr_anualized <- (250)^0.5*vgjr
par(mfcol=c(2,1))
plot(vgjr_anualized)
plot(rendixic)
```



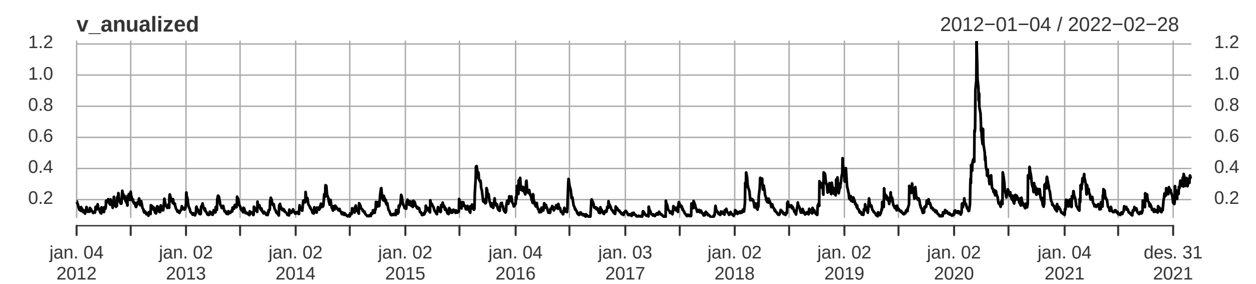


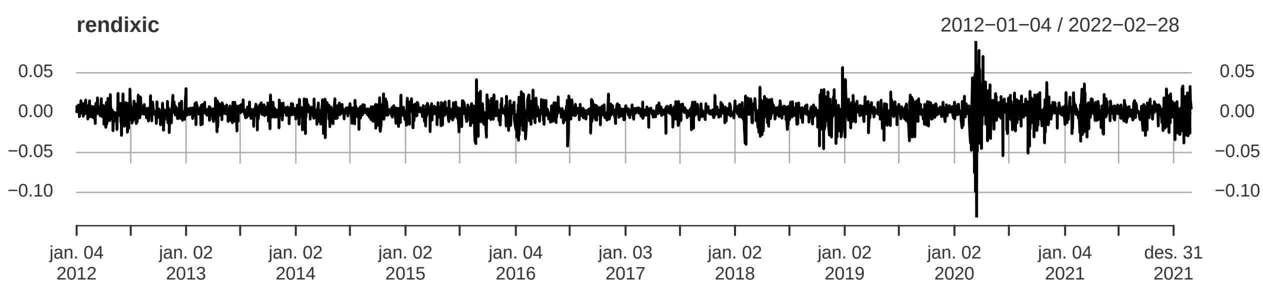
Next, we will fit an ARMA(2,2)-L-GARCH NOT SURE HOW TO DO THIS YET.

4 Grafic and Interpretation of the Estimated Series of Volatility

TAKE THIS FROM THE MODEL CHOSEN ABOVE. FOR NOW HAVE TAKEN THE ONE FROM THE FIRST MODEL (GARCH(1,1))

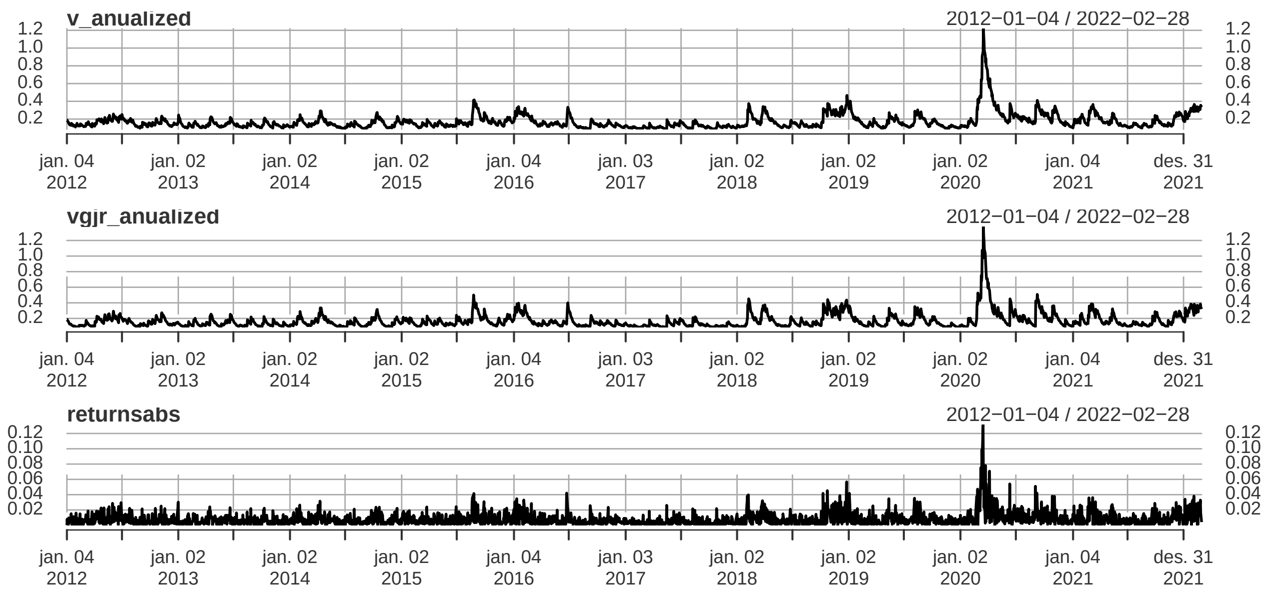
```
par(mfcol=c(2,1))  # Show volatility and returns simultaneously.
plot(v_anualized)
plot(rendixic)
```





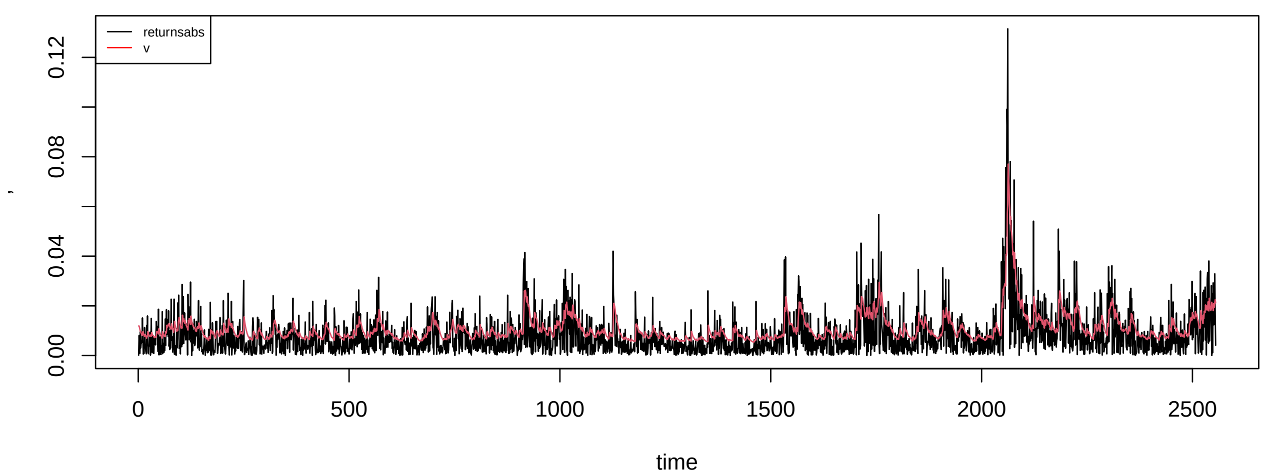
Comparison of the estimated volatilities and the absolute value of the returns.

```
returnsabs <- abs(rendixic)
par(mfcol=c(3,1))  # Show volatility and returns
plot(v_anualized)
plot(vgjr_anualized)
plot(returnsabs)
```



They all seem to follow a similar pattern, which is a good sign, even though we cannot compare the absolute values of the data shown in the plots.

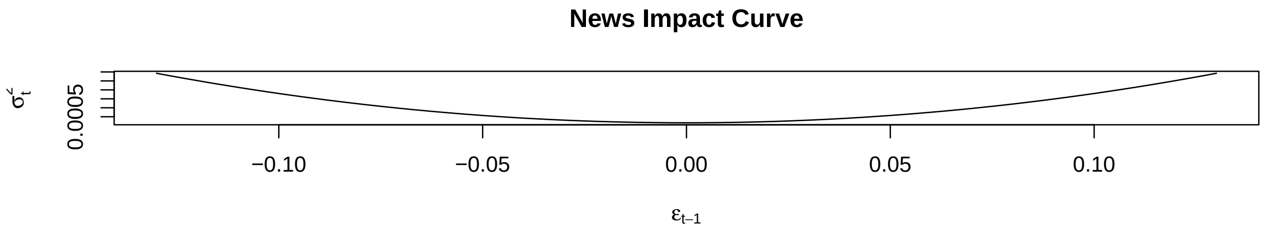
```
time <-  data.frame(returnsabs, v)
ts.plot(time,gpars= list(xlab="time", ylab=",", col = 1:ncol(time)))
legend("topleft", c("returnsabs","v"), lty=c(1,1), col=c("black","red"), cex=0.6)
```

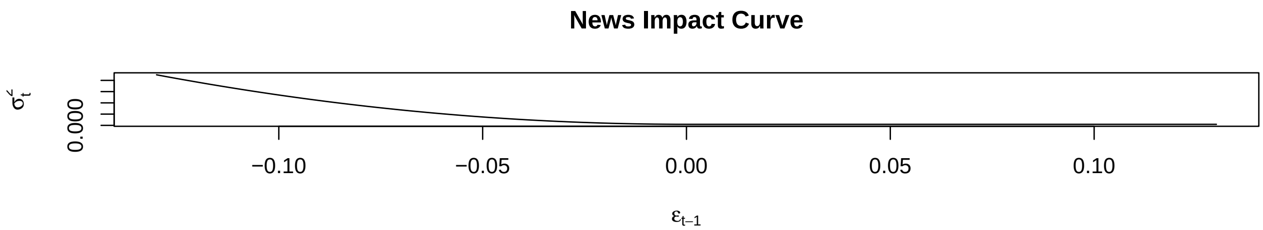


5 Grafic and Interpretation of the News Impact Curve

```
par(mfrow=c(2,1))
ni <- newsimpact(z = NULL, m)
plot(ni$zx, ni$zy, ylab=ni$yexpr, xlab=ni$xexpr, type="l", main = "News Impact Curve")
```

```r
ni2 <- newsimpact(z = NULL, mgjr)
plot(ni2$zx, ni2$zy, ylab=ni2$yexpr, xlab=ni2$xexpr, type="l", main = "News Impact Curve")
```





INTERPRETATIONS?!

6 Volatility Predictions and Interpretations

7 Conclusions