Module 9 - Time Series Models of Heteroscedasticity

MATH1318 Time Series Analysis

Prepared by: Dr. Haydar Demirhan based on the textbook by Cryer and Chan, Time Series Analysis with R, Springer, 2008.

Introduction

In the previous modules, our models are concerned with the conditional mean structure of time series data. So, the previous ARIMA models focus on how to predict the conditional mean of future values based on current and past data. However, mostly in financial time series data, series consists of a high variation between the current and past values of the process. For example, daily returns of stocks are often observed to have larger conditional variance following a period of violent price movement than a relatively stable period. This violates the constancy of variance of further steps ahead for ARIMA processes. We can handle this situation by modelling the conditional variance structure of time series data.

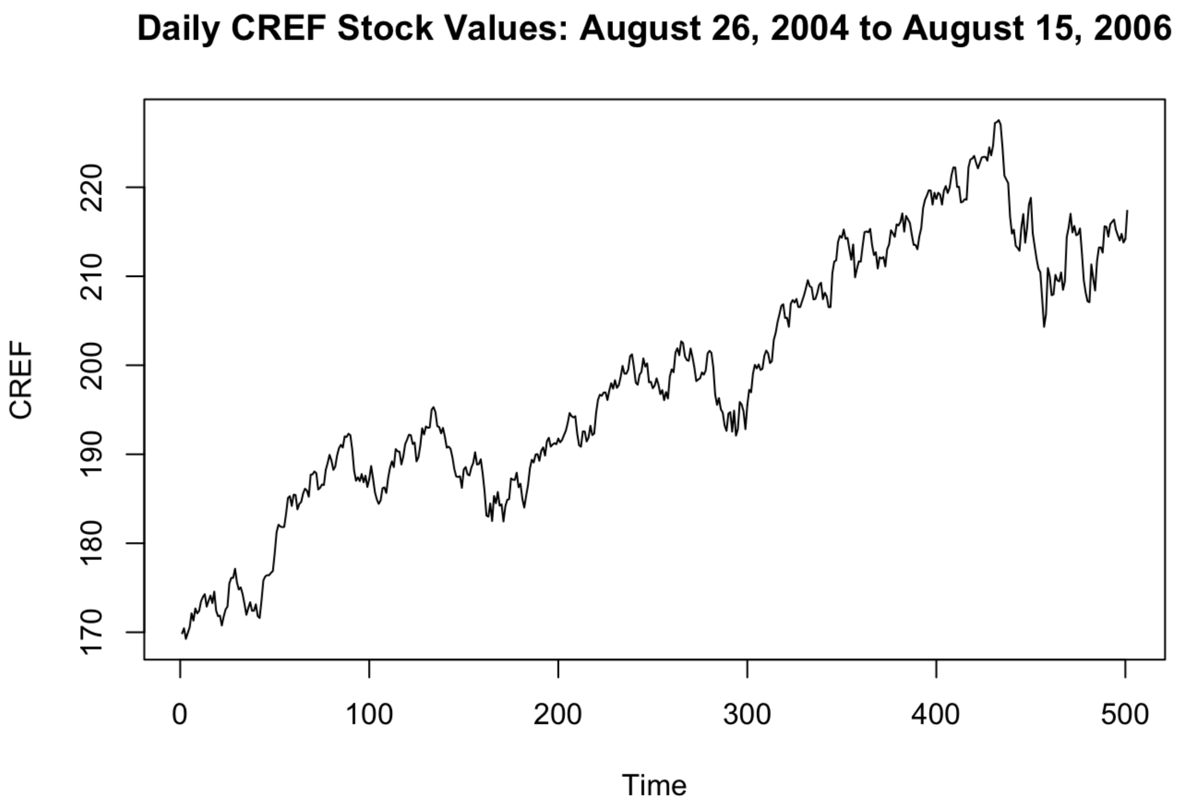
In this module, we will focus on models for the conditional variance process with which we can predict the variability of future values based on current and past data.

For ease of exposition, we will assume in the first few sections that the data are returns of some financial asset and are white noise; that is, serially uncorrelated data. Later, we will discuss some simple schemes for simultaneously modeling the conditional mean and conditional variance structure by combining an ARIMA model with a model of conditional heteroscedasticity.

Some Common Features of Financial Time Series

Let's consider the daily values of a unit of the CREF stock fund, which is a fund of several thousand stocks and is not openly traded in the stock market, over the period from August 26, 2004, to August 15, 2006. Since stocks are not traded over weekends or on holidays, only on so-called trading days, the CREF data do not change over weekends and holidays. For simplicity, we will analyze the data as if they were equally spaced. The following is the time series plot of CREF data:

```
data(CREF)
plot(CREF, main="Daily CREF Stock Values: August 26, 2004 to August 15,
   2006")
```



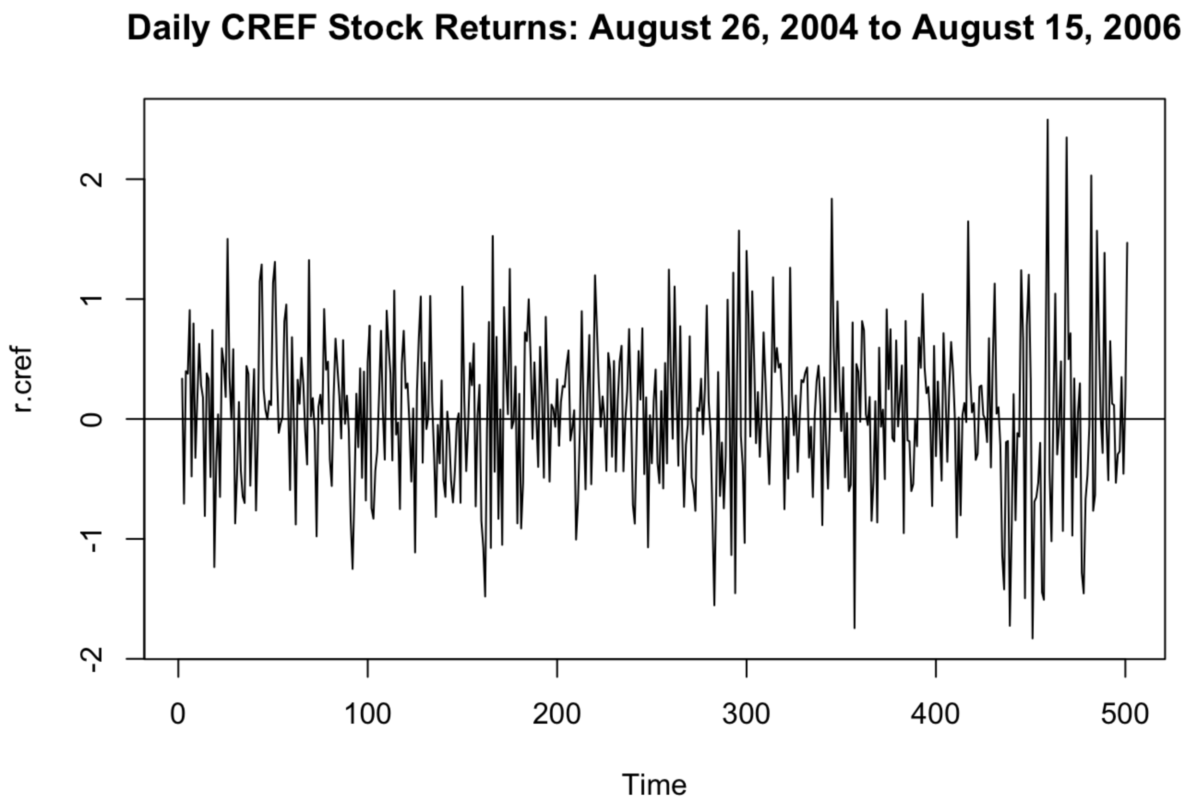
There is an increasing trend with a variation that gets higher as the time passes. So, this indicates a higher variability with a higher level of the stock value. Let $\{p_t\}$ be the time series of, say, the daily price of some financial asset. The (continuously compounded) return on the t-th day is defined as

$$r_t = \log(p_t) - \log(p_{t-1}).$$

Sometimes the returns are then multiplied by 100 so that they can be interpreted as percentage changes in the price. The multiplication may also reduce numerical errors as the raw returns could be very small numbers and render large rounding errors in some calculations.

The following plot shows the time series plot of the CREF return series with a sample size of 500.

```
r.cref=diff(log(CREF))*100
plot(r.cref,main="Daily CREF Stock Returns: August 26, 2004 to August 1
  5, 2006")
abline(h=0)
```

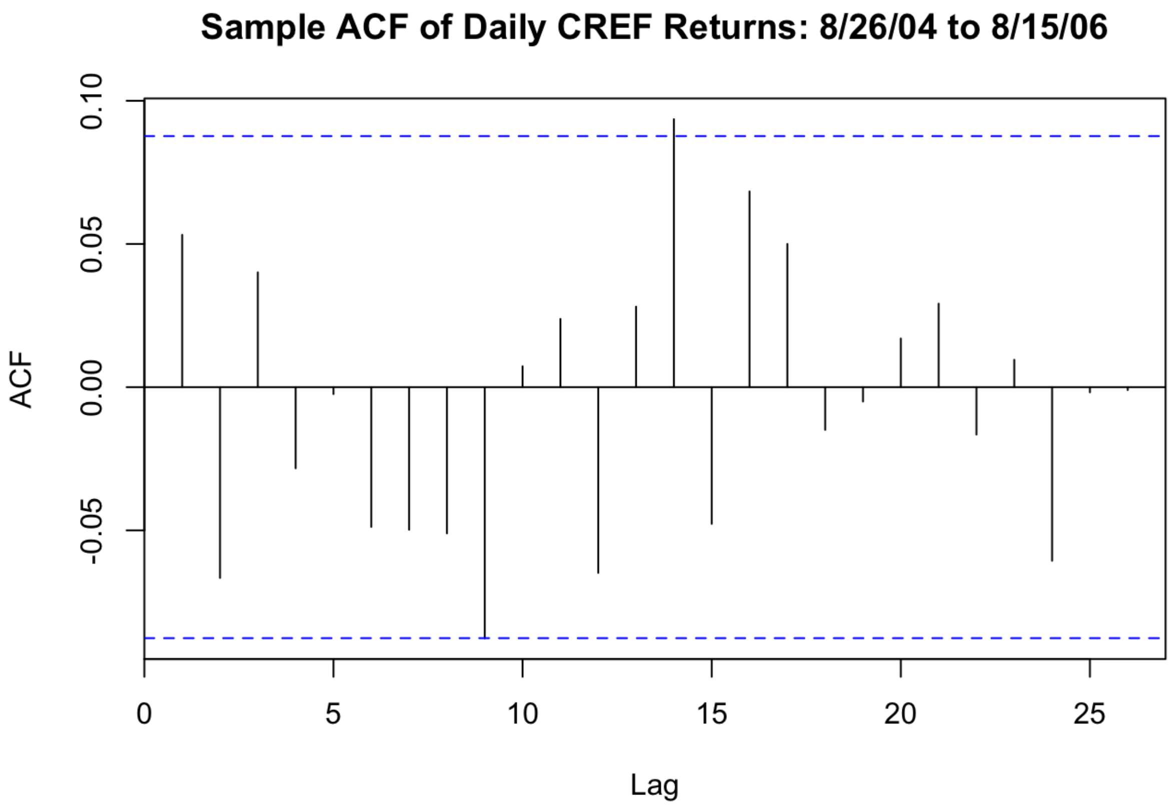


It is seen in this plot that the returns were more volatile over some time periods and became very volatile toward the end of the study period. This might be triggered by the instability in the Middle East due to a war in southern Lebanon from July 12 to August 14, 2006. This pattern of alternating quiet and volatile periods of substantial duration is called **volatility clustering** in the literature.

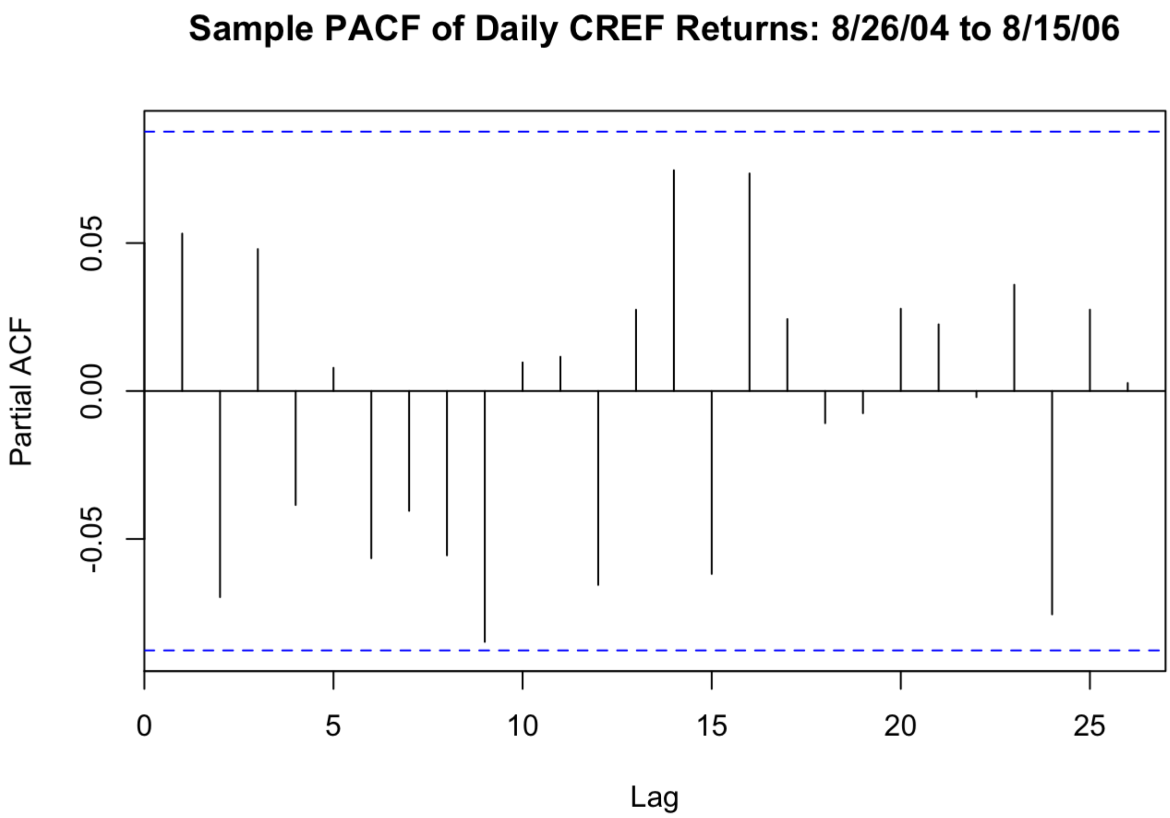
Volatility in a time series refers to the phenomenon where the conditional variance of the time series varies over time.

The sample ACF and PACF of the daily CREF returns are given below.

```
acf(r.cref,main="Sample ACF of Daily CREF Returns: 8/26/04 to 8/15/06")
```




```
pacf(r.cref,main="Sample PACF of Daily CREF Returns: 8/26/04 to 8/15/0
  6")
```



There is a slight serial correlation in the return series. The sample EACF is as follows:

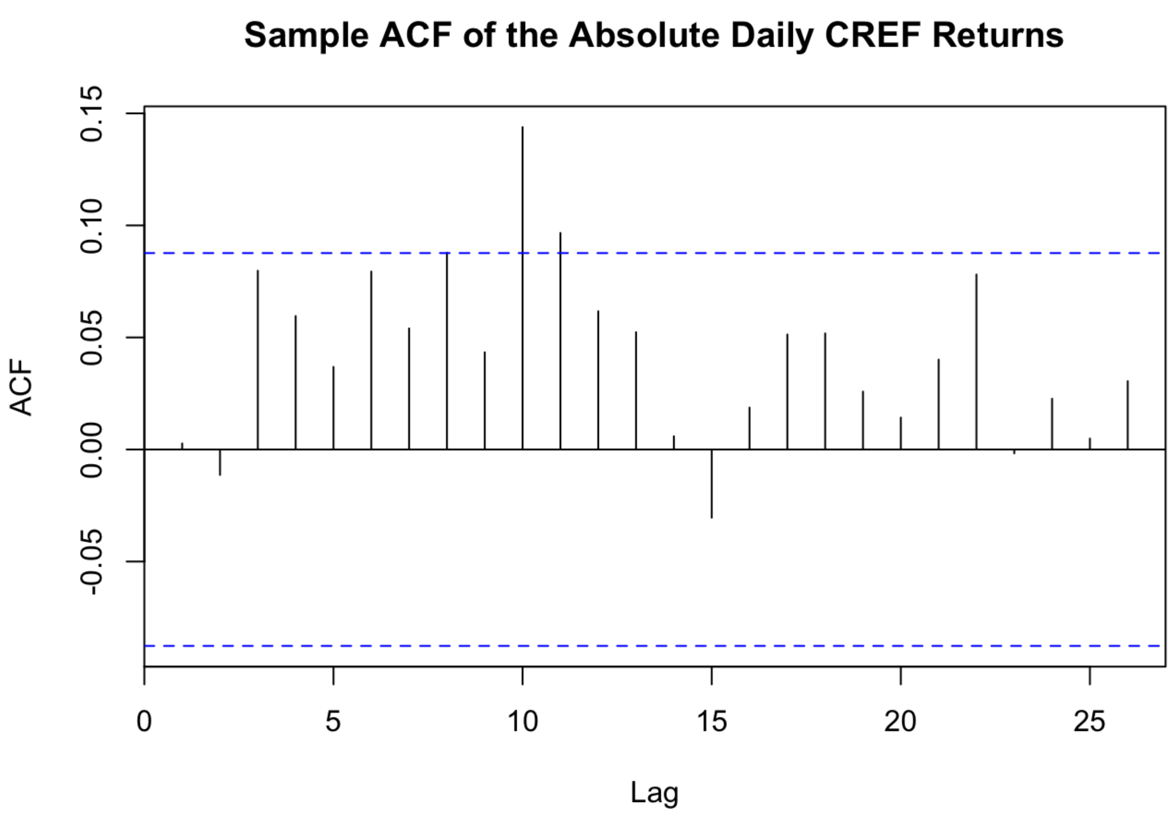
```
eacf(r.cref)
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o o o o o o o o o o o  o  o  x
## 1 x o o o o o o o x o o  o  o  o
## 2 x x o o o o o o x o o  o  o  o
## 3 x x o o o o o o o o o  o  o  o
## 4 x x x o o o o o o o o  o  o  o
## 5 x x x o x o o o o o o  o  o  o
## 6 x x x x x o o o o o o  o  o  o
## 7 x o o x o o x o o o o  o  o  o
```

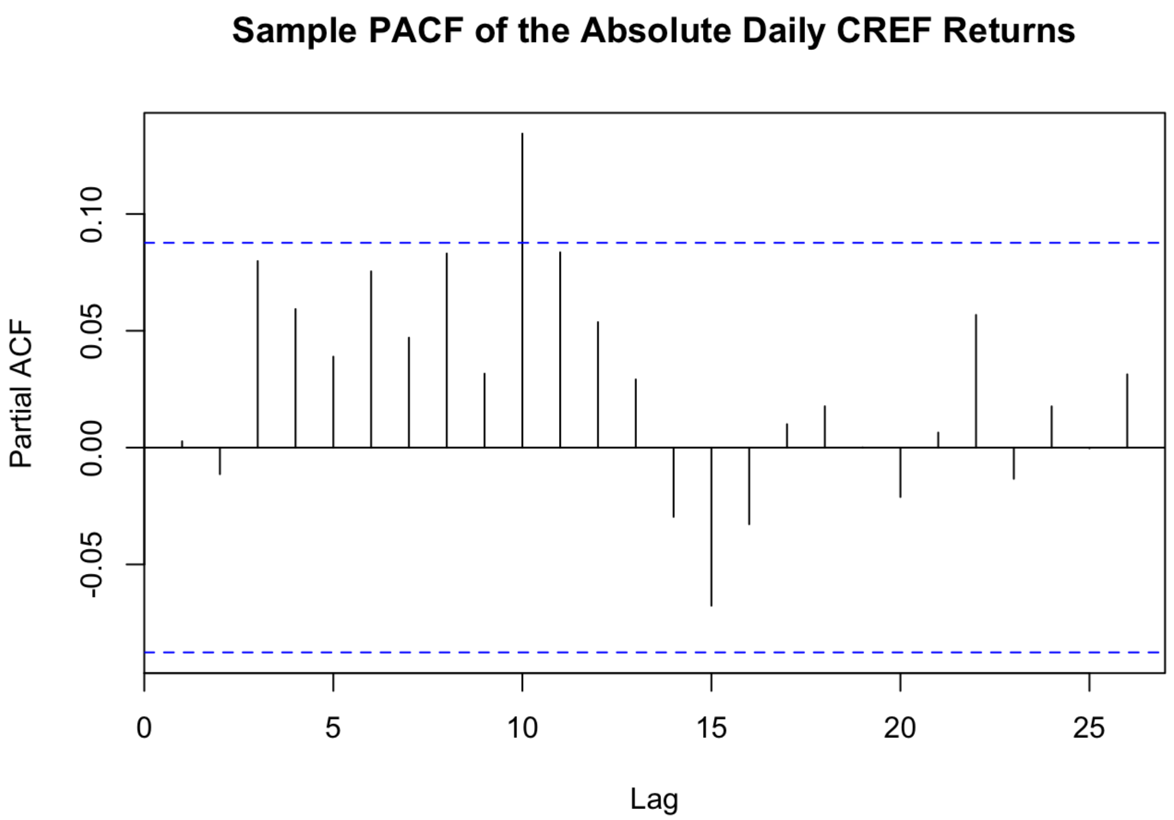
The sample EACF also confirms the existence of little serial correlation by suggesting a white noise series. The average CREF return equals 0.0493 with a standard error of 0.02885. Thus the mean of the return process is not statistically significantly different from zero. this is in accordance with the efficient-market hypothesis.

The problem here is that the data is not independently and identically distributed due to the significant volatility. If series values are truly independent, then nonlinear transformations such as log, absolute value, or square root preserve independence. However, the same is not true of correlation, since correlation is only a measure of linear dependence. Thus, we can detect violations of iid assumption over the ACF or PACF of transformed series. The following plots show the sample ACF and the sample PACF of absolute and squared returns.

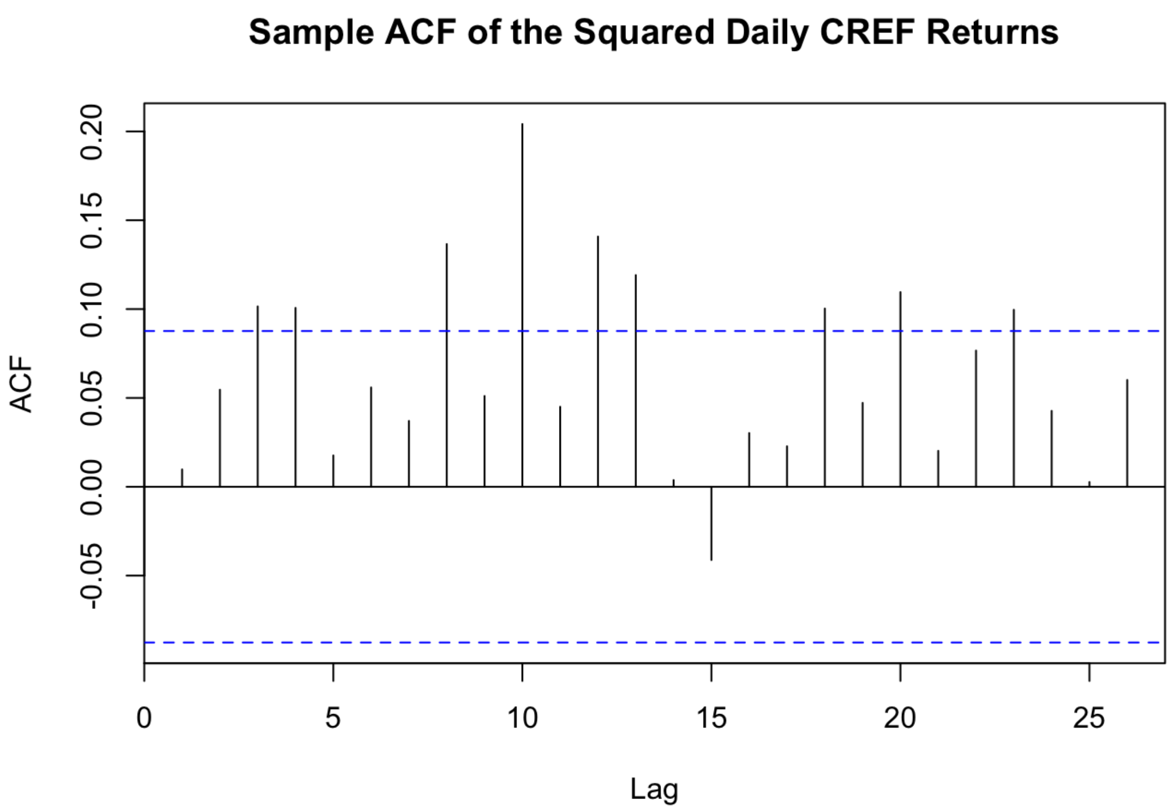
```
acf(abs(r.cref),main="Sample ACF of the Absolute Daily CREF Returns")
```



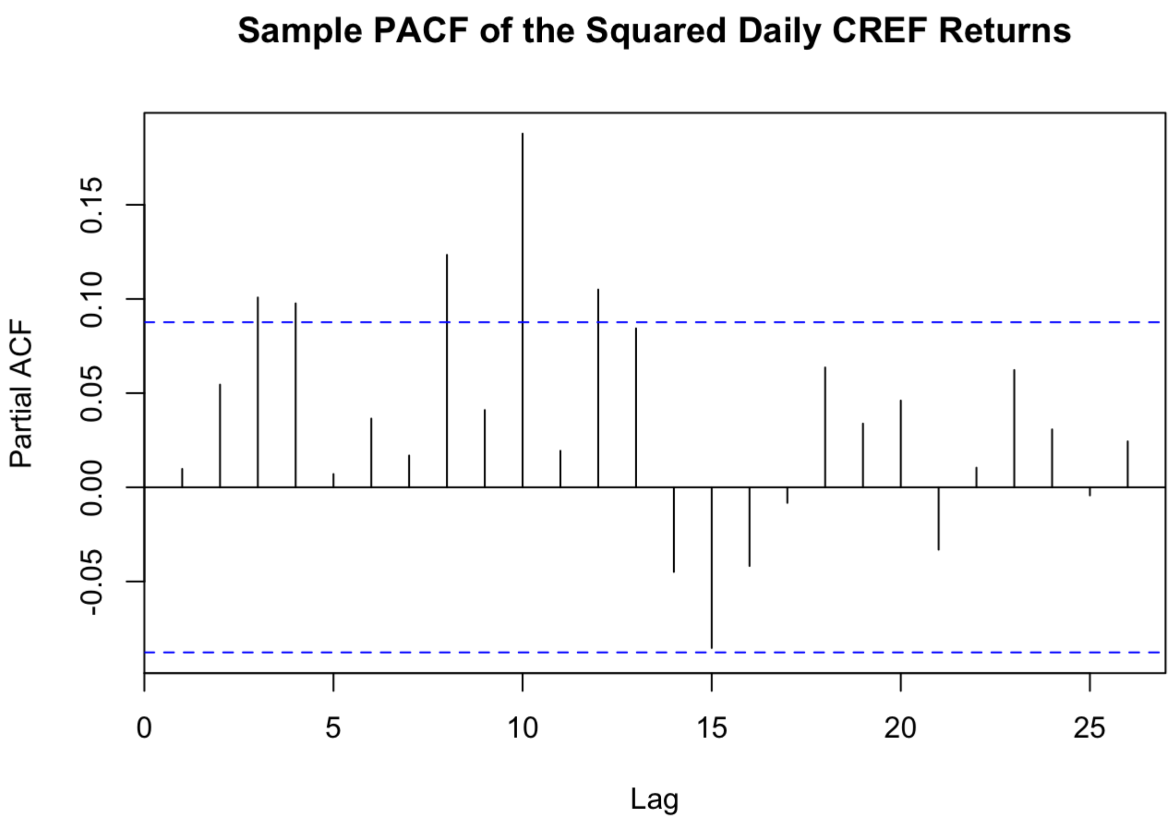
```
pacf(abs(r.cref),main="Sample PACF of the Absolute Daily CREF Returns")
```




```
acf(r.cref^2,main="Sample ACF of the Squared Daily CREF Returns")
```



```
pacf(r.cref^2,main="Sample PACF of the Squared Daily CREF Returns")
```

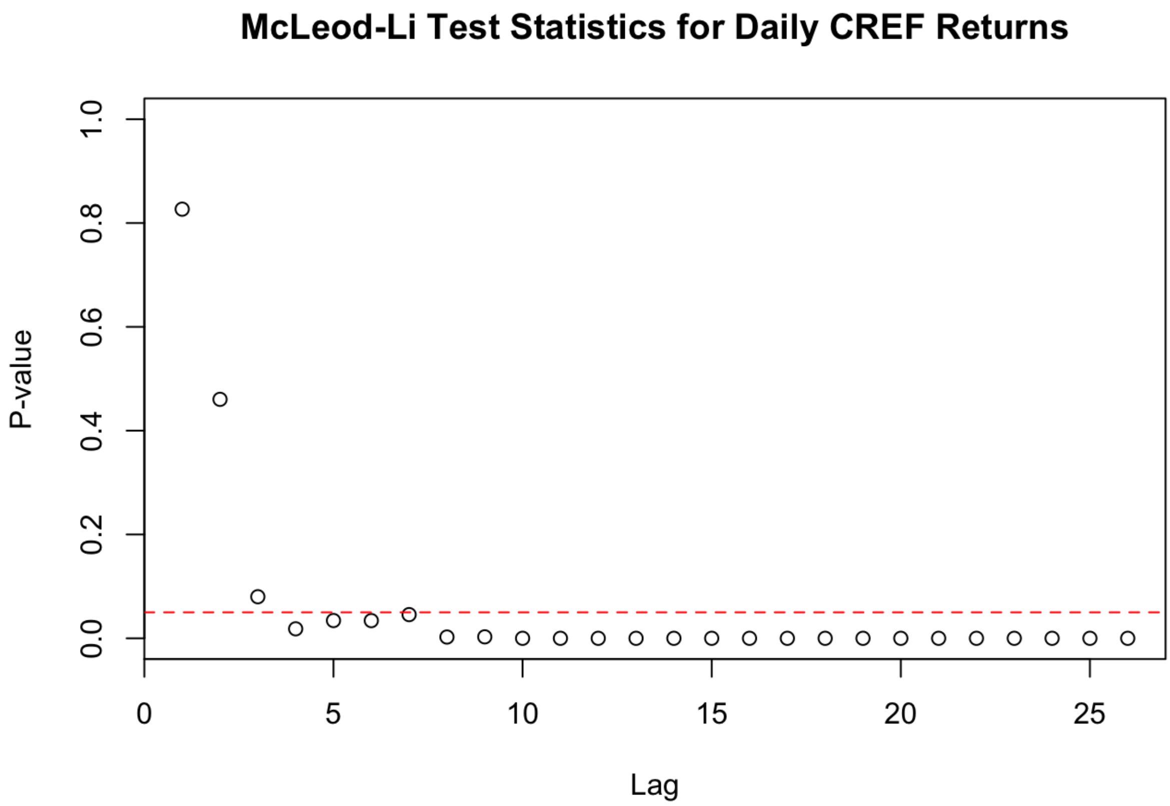


Because we observe some significant correlations in these plots, we have some evidence that the daily CREF returns are not independently and identically distributed.

In addition to the visual tools, we can apply the Ljung-Box test to the transformed data. The test statistic of this test will be approximately chi-square distributed if there is no **Autoregressive Conditional Heteroskedasticity (ARCH)**. This approach can be extended to the case when the conditional mean of the process is non-zero and if an ARMA model is adequate in describing the autocorrelation structure of the data. In which case, the first m autocorrelations of the squared residuals from this model can be used to test for the presence of ARCH. We will refer to the test for ARCH effects using the Box-Ljung statistic with the squared residuals or data as the **McLeod- Li test**.

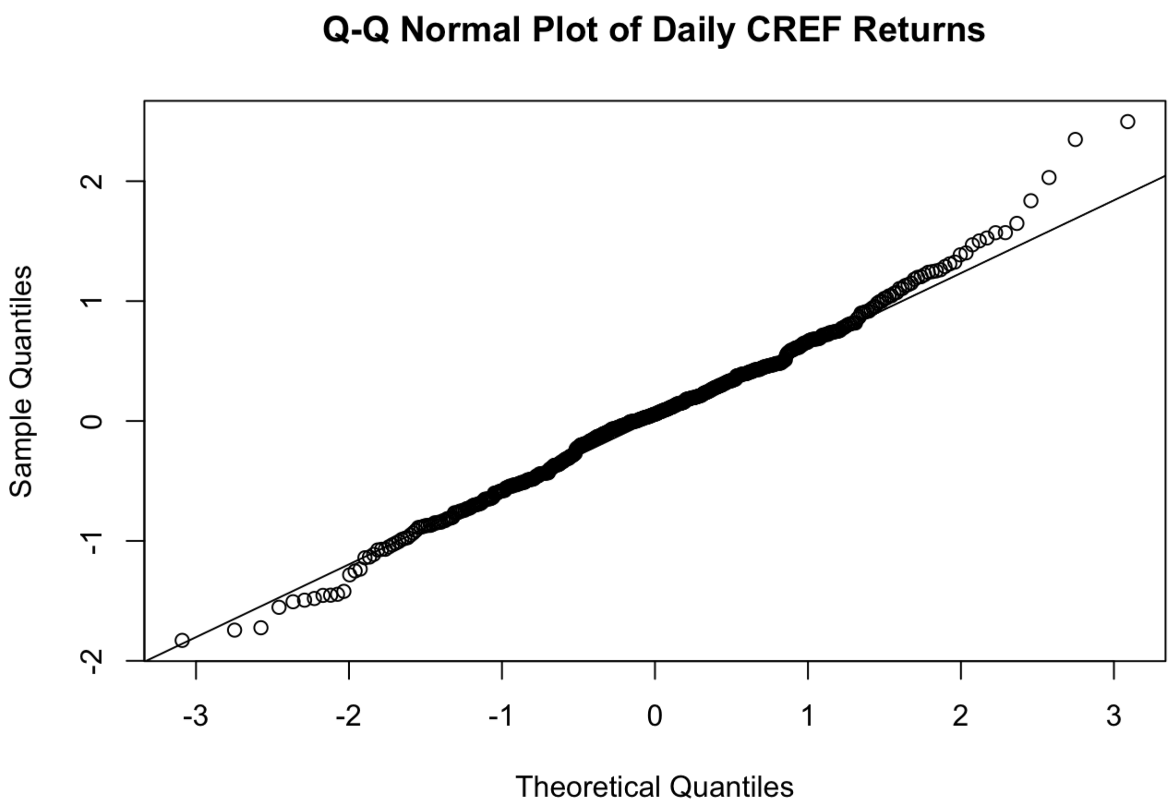
In practice, it is useful to apply the McLeod-Li test for ARCH using a number of lags and plot the p-values of the test. The following plot shows that the McLeod-Li tests are all significant at the 5% significance level when more than 3 lags are included in the test. This is strong evidence for ARCH in this data.

```
McLeod.Li.test(y=r.cref,main="McLeod-Li Test Statistics for Daily CREF
   Returns")
```



We can check the normality of CREF returns by using the Q-Q plot.

```
qqnorm(r.cref,main="Q-Q Normal Plot of Daily CREF Returns")
qqline(r.cref)
```



The QQ plot suggests that the distribution of returns may have a tail thicker than that of a normal distribution and may be somewhat skewed to the right. Then, we apply the Shapiro-Wilk test to get a more formal idea about the normality of the CREF returns.

```
shapiro.test(r.cref)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  r.cref
## W = 0.99324, p-value = 0.02412
```

The p-value of 0.024 suggests a significant divergence from the normal distribution at 5% level of significance. Please note that you can apply other normality tests if you get a p-value close to 0.05. But you should keep in mind that each test has its own advantages, disadvantages, and characteristics in terms of the detection of non-normality.

In summary, the CREF return data are found to be serially uncorrelated but admit a higher-order dependence structure, namely volatility clustering, and a heavy-tailed distribution. It is commonly observed that such characteristics are rather prevalent among financial time series data.

The ARCH(1) Model

As mentioned, the return series of a financial asset, $\{r_t\}$, is often a serially uncorrelated sequence with zero mean, even as it exhibits volatility clustering. This suggests that the conditional variance of r_t given past returns is not constant. The conditional variance, also referred to as the conditional volatility, of r_t will be denoted by $\sigma^2_{t|t-1}$, with the subscript $t-1$ denoting that the conditioning is upon returns through time $t-1$. When r_t is available, the squared return r_t^2 provides an unbiased estimator of $\sigma^2_{t|t-1}$. A series of large squared returns may indicate a relatively volatile period while a series of small squared returns may indicate a relatively quiet period. The ARCH model is formally a regression model with the conditional volatility as the response variable and the past lags of the squared return as the covariates. The ARCH(1) model assumes that the return series $\{r_t\}$ is generated as follows:

$$\begin{aligned} r_t &= \sigma_{t|t-1}\epsilon_t \\ \sigma^2_{t|t-1} &= \omega + \alpha r^2_{t-1}, \end{aligned}$$

or

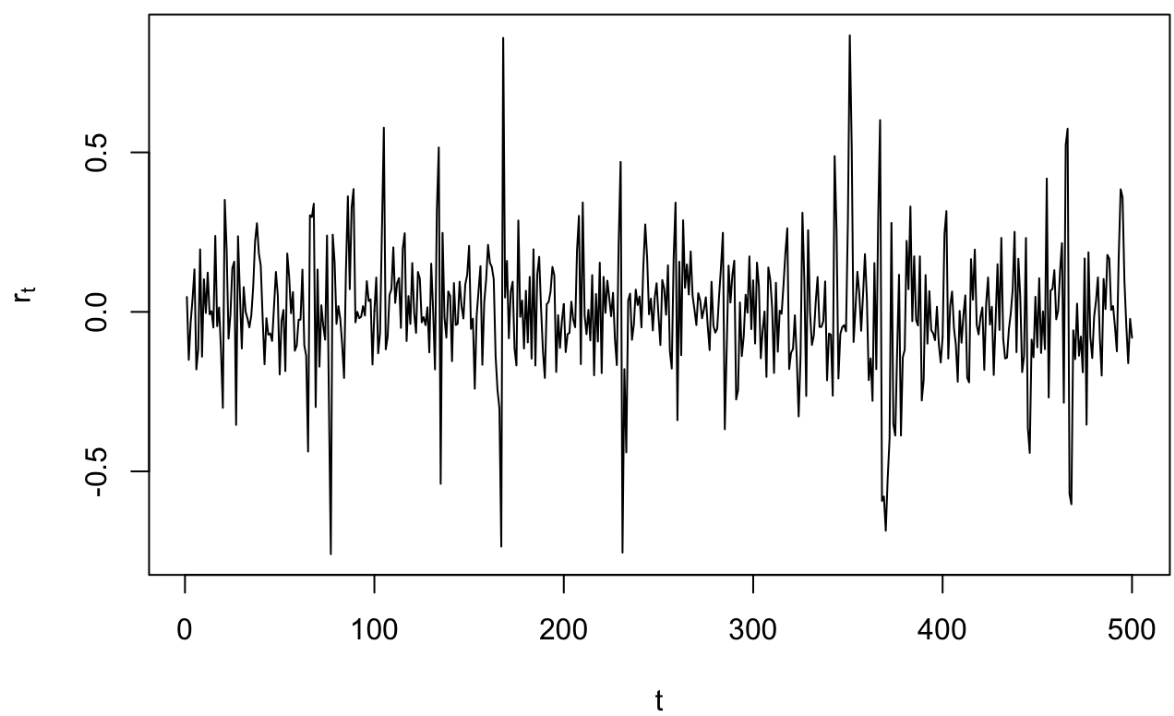
$$r_t = (\omega + \alpha r^2_{t-1})\epsilon_t,$$

where α and ω are unknown parameters, $\{\epsilon_t\}$ is a sequence of independently and identically distributed random variables each with zero mean and unit variance (also known as the innovations), and $\{\epsilon_t\}$ is independent of $r_{t-j}, j = 1, 2, \ldots$. Because the innovation ϵ_t is presumed to have unit variance, the conditional variance of r_t equals $\sigma^2_{t|t-1}$. Because $\sigma^2_{t|t-1}$ is known given the past observations, ϵ_t is independent of past returns, and the variance of errors is 1, we obtain

$$
\begin{aligned}
E(r_t^2|r_{t-j}, j = 1, 2, \ldots) &= E(\sigma_{t|t-1}^2 \epsilon_t^2 | r_{t-j}, j = 1, 2, \ldots) \\
&= \sigma_{t|t-1}^2 E(\epsilon_t^2 | r_{t-j}, j = 1, 2, \ldots) \\
&= \sigma_{t|t-1}^2 E(\epsilon_t^2) \\
&= \sigma_{t|t-1}^2.
\end{aligned}
$$

The following plot shows the time series plot of an ARCH(1) process with $\omega = 0.01$ and $\alpha_1 = 0.9$.

```
set.seed(1235678)
library(tseries)
# Randomly generate ARCH(1) series using GARCH.sim() function
garch01.sim=garch.sim(alpha=c(.01,.9),n=500)
plot(garch01.sim,type='l',ylab=expression(r[t]), xlab='t')
```



We observe in this series that larger fluctuations cluster together and series recovers from these larger fluctuations. This implies that there is volatility clustering in the series and it has a short memory due to the conditional variance properties.

Because the conditional variance is not observable in ARCH models, we replace it with some observable quantities to explore the regression relationship graphically.

Let's introduce a serially uncorrelated series with zero mean such that

$$\eta_t = r_t^2 - \sigma_{t|t-1}^2.$$

The series η_t is also uncorrelated with past returns. When we write $\sigma_{t|t-1}^2 = r_t^2 - \eta_t$ in ARCH(1) model, we obtain the following

$$r_t^2 - \omega + \alpha r_{t-1}^2 + \eta_t.$$

Thus, the squared return series satisfies an AR(1) model under the assumption of an ARCH(1) model for the return series. Based on this useful observation, an ARCH(1) model may be specified if an AR(1) specification for the squared returns is warranted.

Besides its value in terms of data analysis, the deduced AR(1) model for the squared returns can be exploited to gain theoretical insights on the parameterization of the ARCH model. For example, because the squared returns must be nonnegative, it makes sense to always restrict the parameters ω and α to be nonnegative. Also, if the return series is stationary with a variance σ^2, then taking expectation on both sides of the above equation gives

$$\sigma^2 = \omega + \alpha\sigma^2.$$

Here, $0 \leq \alpha < 1$ is the necessary and sufficient condition for the (weak) stationarity of the ARCH(1) model. It can be checked that the ARCH(1) process is white noise. Hence, it is an example of a white noise that admits a nonconstant conditional variance process that varies with the lag one of the squared process. An ARCH(1) process has fat tails as a result of the volatility clustering even with normal innovations.

Main use of the ARCH model is to predict the future conditional variances. For example, one might be interested in forecasting the h-step-ahead conditional variance

$$\sigma_{h+t|t}^2 = E(r_{t+h}^2 | r_t, r_{t-1}, \ldots).$$

For $h = 1$, the ARCH(1) model implies that

$$\sigma_{t+1|t}^2 = \omega + \alpha r_t^2 = (1-\alpha)\sigma^2 + \alpha r_t^2.$$

which is a weighted average of the long-run variance and the current squared return. Similarly, using the iterated expectation formula, we can compute h-step ahead variance recursively:

$$\begin{aligned}
\sigma_{t+h|t}^2 &= E(r_{t+h}^2 | r_t, r_{t-1}, \ldots) \\
&= E[E(\sigma_{t+h|t+h-1}^2 \varepsilon_{t+h}^2 | r_{t+h-1}, r_{t+h-2}, \ldots) | r_t, r_{t-1}, \ldots] \\
&= E[\sigma_{t+h|t+h-1}^2 E(\varepsilon_{t+h}^2) | r_t, r_{t-1}, \ldots] \\
&= E(\sigma_{t+h|t+h-1}^2 | r_t, r_{t-1}, \ldots) \\
&= \omega + \alpha E(r_{t+h-1}^2 | r_t, r_{t-1}, \ldots) \\
&= \omega + \alpha \sigma_{t+h-1|t}^2
\end{aligned}$$

where we adopt the convention that $\sigma_{t+h|t}^2 = r_{t+h}^2$ for $h < 0$.

GARCH Models

In ARCH models, we predict the future conditional variances only involves the most recent squared return. However, we might obtain more efficient estimates by including all past squared returns with lesser weight for more distant volatilities. One approach is to include further lagged squared returns in the model. The ARCH(q) model generalizes the conditional variance equation by specifying that

$$\sigma^2_{t|t-1} = \omega + \alpha_1 r^2_{t-1} + \alpha_2 r^2_{t-2} + \cdots + \alpha_q r^2_{t-q}$$

where q is ARCH order. Another approach introduces p lags of the conditional variance in the model, where p is referred to as the GARCH order. The combined model is called the generalized autoregressive conditional heteroscedasticity, GARCH(p,q), model.

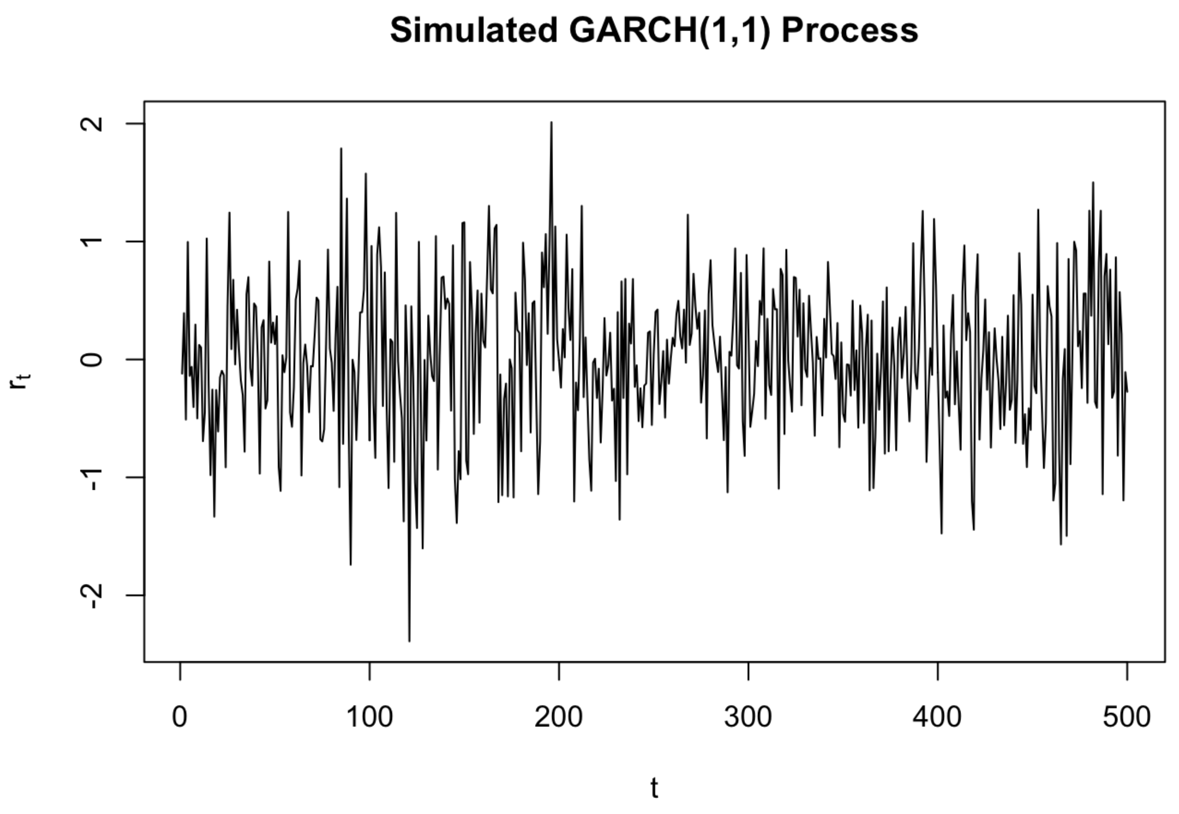
$$\sigma^2_{t|t-1} = \omega + \beta_1 \sigma^2_{t-1|t-2} + \cdots + \beta_p \sigma^2_{t-p|t-p-1} + \alpha_1 r^2_{t-1} + \alpha_2 r^2_{t-2} + \cdots + \alpha_q r^2_{t-q}$$

We note that in some of the literature, the notation GARCH(p,q) is written as GARCH(q,p); that is, the orders are switched. It can be rather confusing but true that the two different sets of conventions are used in different software! You must find out which convention is used by the software on hand before fitting or interpreting a GARCH model.

Because conditional variances must be nonnegative, the coefficients in a GARCH model are often constrained to be nonnegative. However, the nonnegative parameter constraints are not necessary for a GARCH model to have nonnegative conditional variances with probability 1.

The following plot is the time series plot of a simulated GARCH(1,1) model with 500 observations, standard normal innovations, and parameters $\omega = 0.02, \alpha = 0.05$ and $\beta = 0.9$.

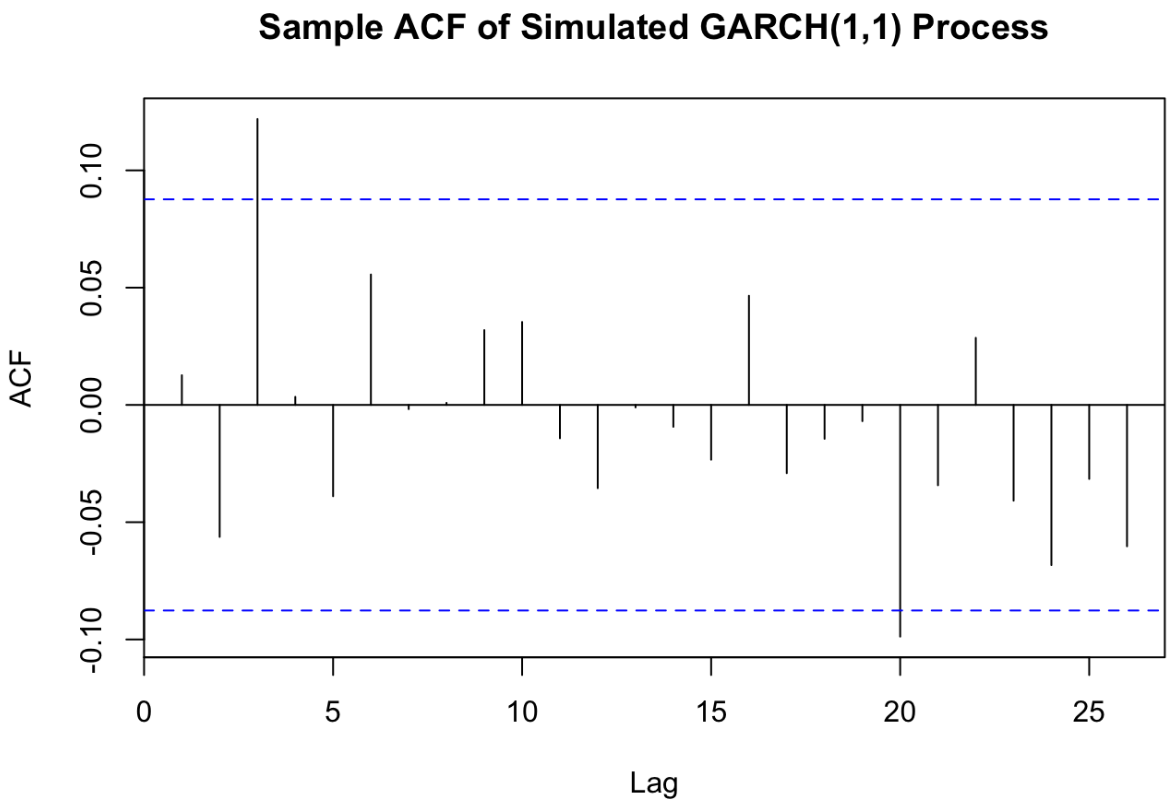
```
set.seed(1234567)
garch11.sim=garch.sim(alpha=c(0.02,0.05),beta=.9,n=500)
plot(garch11.sim,type='l',ylab=expression(r[t]), xlab='t',main="Simulat
  ed GARCH(1,1) Process")
```



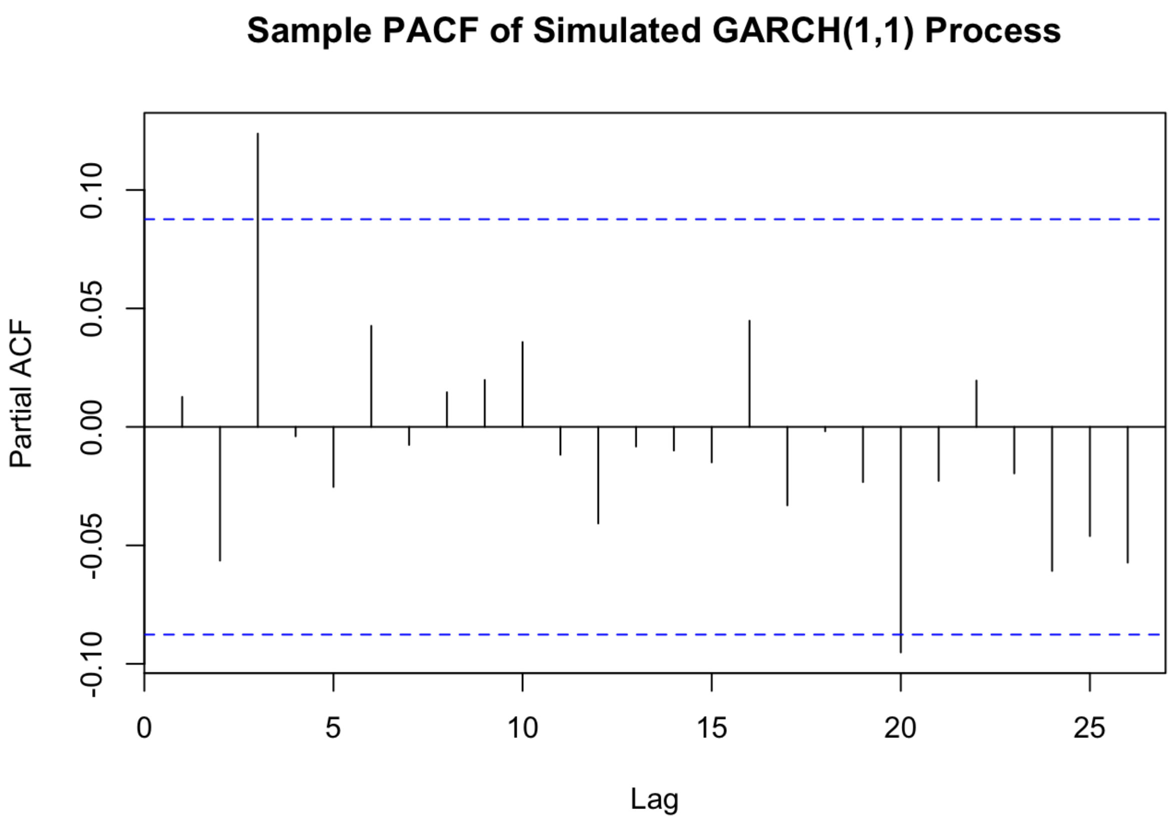
Volatility clustering is evident in this plot as large (small) fluctuations are usually succeeded by large (small) fluctuations. Moreover, the inclusion of the lag 1 of the conditional variance in the model successfully enhances the memory in the volatility.

The sample ACF and PACF plots of the simulated GARCH(1,1) model is given below:

```
acf(garch11.sim,main="Sample ACF of Simulated GARCH(1,1) Process")
```



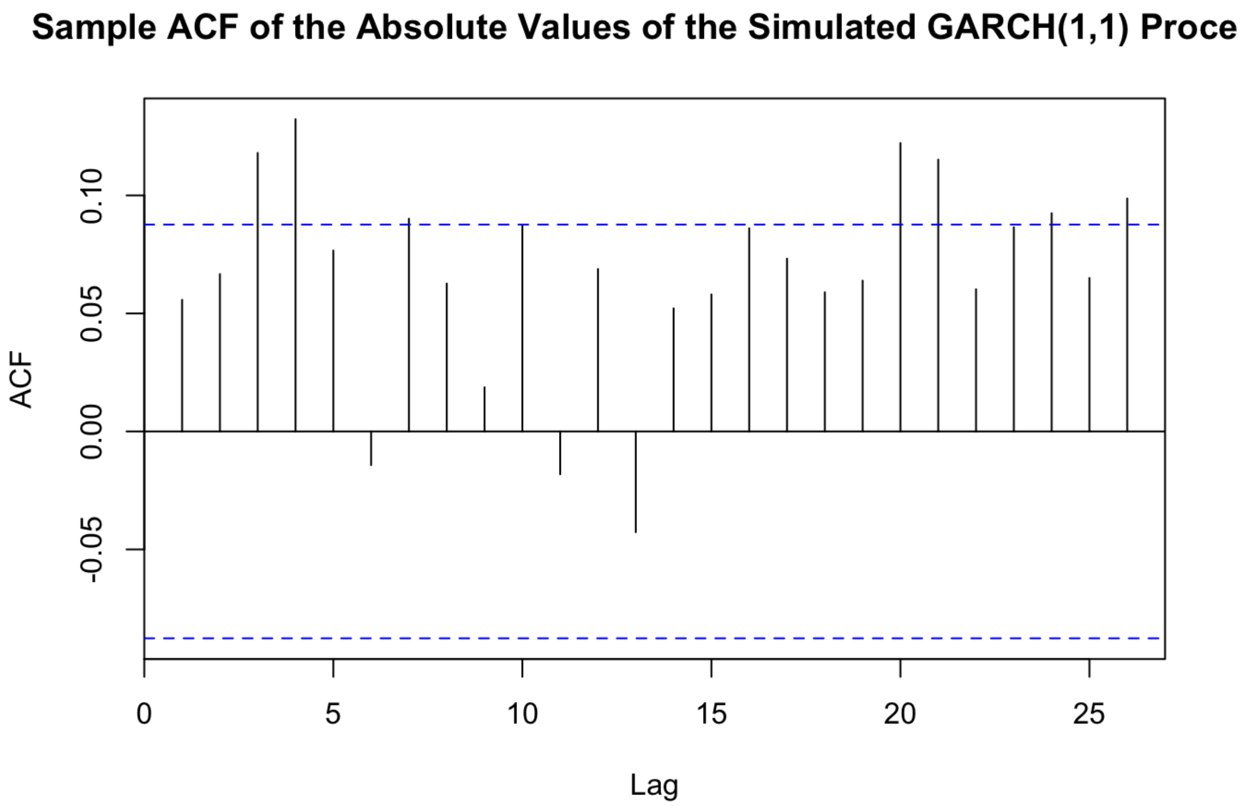

```
pacf(garch11.sim,main="Sample PACF of Simulated GARCH(1,1) Process")
```



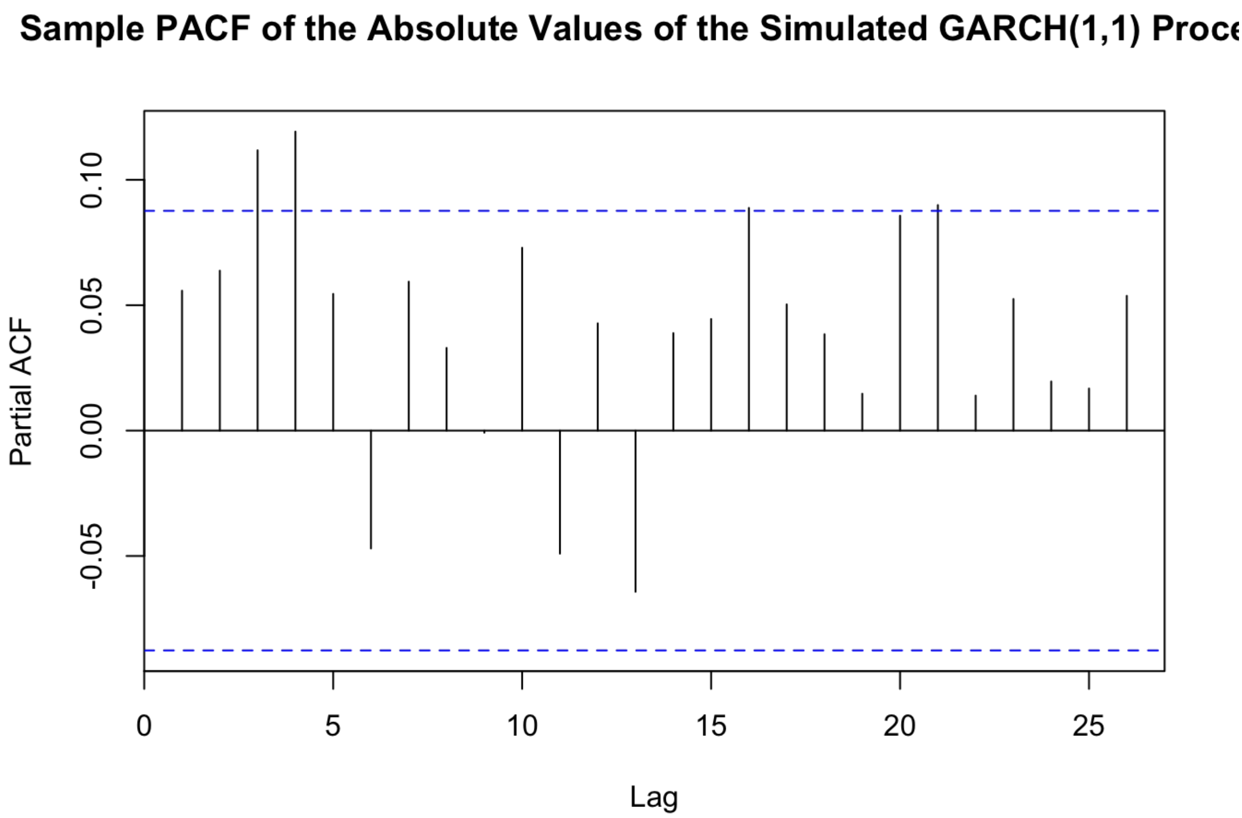
There is no significant correlation in ACF and PACF plots but the mildly significant ones at lags 3 and 20. The series seems to be serially uncorrelated as it should be.

Then we apply absolute value and square transformations and obtain the following ACF and PACF plots:

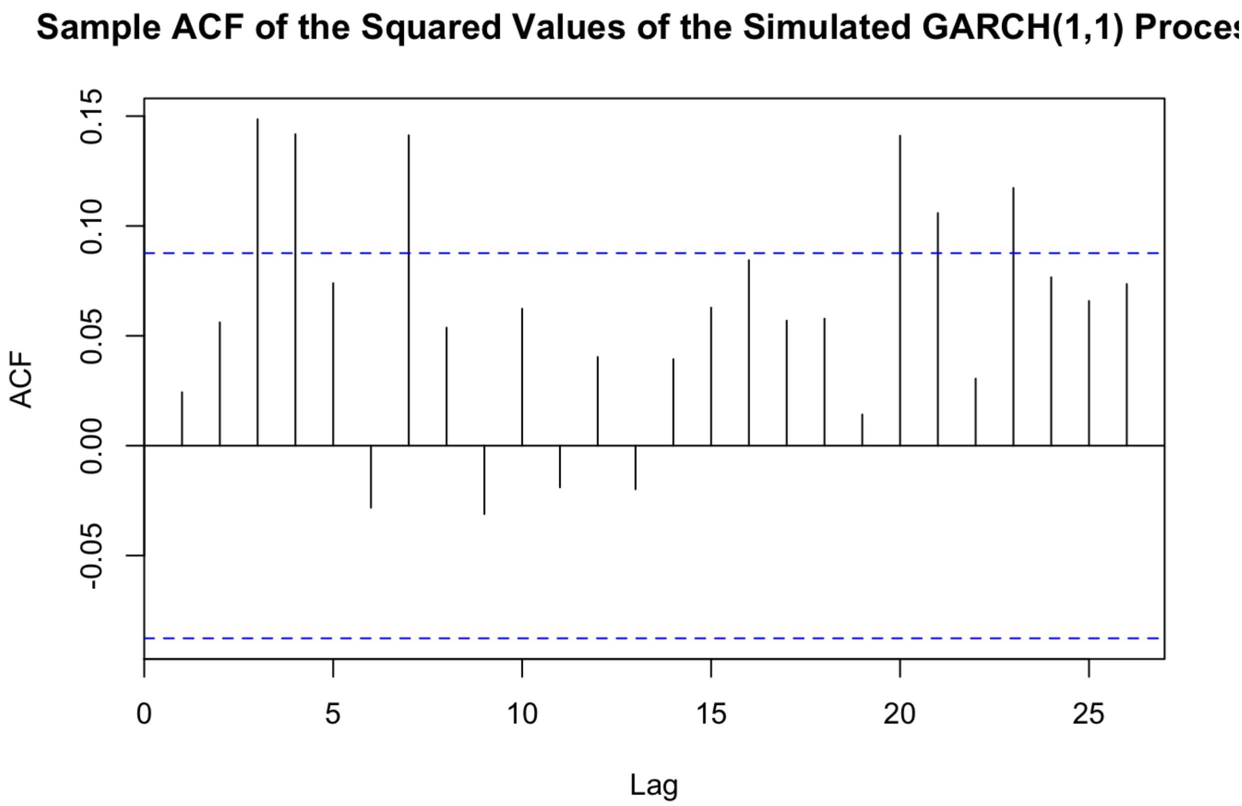
```
acf(abs(garch11.sim),main="Sample ACF of the Absolute Values of the Sim
  ulated GARCH(1,1) Process")
```



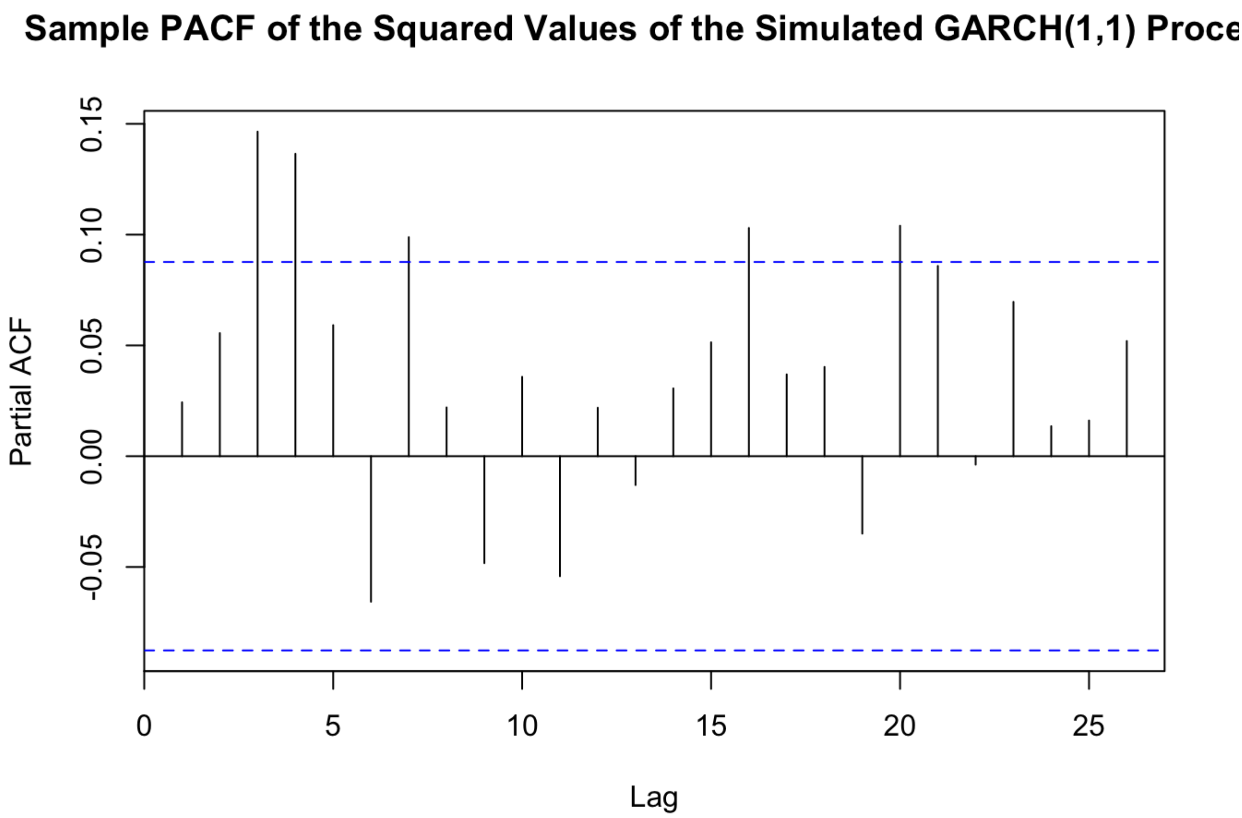

```
pacf(abs(garch11.sim),main="Sample PACF of the Absolute Values of the S
  imulated GARCH(1,1) Process")
```




```
acf(garch11.sim^2,main="Sample ACF of the Squared Values of the Simulat
  ed GARCH(1,1) Process")
```




```
pacf(garch11.sim^2,main="Sample PACF of the Squared Values of the Simul
  ated GARCH(1,1) Process")
```



Because we observe significant serial correlations in the transformed series, we infer that the simulated process is in fact serially dependent. Interestingly, the lag 1 autocorrelations are not significant in any of these last four plots.

For model the specification of the GARCH models, we express the model for the conditional variances in terms of the squared returns. We defined $\eta_t = r_t^2 - \sigma_{t|t-1}^2$. Similar to the ARCH(1) model, we can show that $\{\eta_t\}$ is a serially uncorrelated sequence. Moreover, $\{\eta_t\}$ is uncorrelated with past squared returns. Substituting the expression $\sigma_{t|t-1}^2 = r_t^2 - \eta_t$ into GARCH(p,q) model equation yields

$$r_t^2 = \omega + (\beta_1 + \alpha_1) r_{t-1}^2 + \cdots + (\beta_{max(p,q)} + \alpha_{max(p,q)}) r_{t-max(p,q)}^2 \\ + \eta_t - \beta_1 \eta_{t-1} - \cdots - \beta_p \eta_{t-p}$$

where $\beta_k = 0$ for all integers $k > p$ and $\alpha_k = 0$ for $k > q$. This implies that the GARCH(p,q) model for the return series corresponds to an ARMA(max(p,q),p) model.

Thus, we can apply the model identification techniques for ARMA models to the squared return series to identify p and $max(p, q)$. Notice that if q is smaller than p, it will be masked in the model identification. In such cases, we can first fit a GARCH(p,p) model and then estimate q by examining the significance of the resulting ARCH coefficient estimates.

To illustrate the specification of orders of a GARCH model, we obtain the EACF of the squared values from the simulated GARCH(1,1) series.

```
eacf((garch11.sim)^2)
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o o x x o o x o o o o  o  o  o
## 1 x o o o x o x x o o o  o  o  o
## 2 x o o o o o x o o o o  o  o  o
## 3 x x x o o x o o o o o  o  o  o
## 4 x x o x x o o o o o o  o  o  o
## 5 x o x x o o o o o o o  o  o  o
## 6 x o x x o x o o o o o  o  o  o
## 7 x x x x x x o o o o o  o  o  o
```

There is not a very clear pattern in the EACF. The fuzziness of the signal in the EACF table is likely caused by the larger sampling variability when we deal with higher moments. Shin and Kang (2001) argued that, to a first-order approximation, a power transformation preserves the theoretical autocorrelation function and hence the order of a stationary ARMA process. Their result suggests that the GARCH order may also be identified by studying the absolute returns. The sample EACF for absolute returns is shown below:

```
eacf(abs(garch11.sim))
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o o x x o o x o o o o  o  o  o
## 1 x o o o x o o o o o o  o  o  o
## 2 x x o o o o o o o o o  o  o  o
## 3 x x o o o x o o o o o  o  o  o
## 4 x x o x o x o o o o o  o  o  o
## 5 x o x x x o o o o o o  o  o  o
## 6 x o x x x x o o o o o  o  o  o
## 7 x x x x x o x o o o o  o  o  o
```

Now, the EACF of absolute returns suggests an ARMA(1,1) model, and therefore a GARCH(1,1) model for the original data, although there is also a hint of a GARCH(2,2) model.

As for the CREF returns, we construct the EACF table after the absolute value and squared value transformations.

```
eacf(abs(r.cref))
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o o o o o o o o o x x  o  o  o
## 1 x o o o o o o o o o o  o  o  o
## 2 x o o o o o o o o o o  o  o  o
## 3 x o x o o o o o o o o  o  o  o
## 4 x o x o o o o o o o o  o  o  o
## 5 x x x x o o o o o o o  o  o  o
## 6 x x x x o o o o o o o  o  o  o
## 7 x x x x o o o o o o o  o  o  o
```

```
eacf(r.cref^2)
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o o x x o o o x o x o  x  x  o
## 1 x o o o o o o x o x o  o  x  o
## 2 x x o o o o o o o o o  o  o  o
## 3 x x o x o o o o o o o  x  o  o
## 4 o x x x o o o o o o o  x  x  o
## 5 x x o x o o o o o o o  x  o  o
## 6 x x o x x o o o o o o  o  x  o
## 7 x x o x o o x o o o o  o  x  o
```

This EACF table for absolute returns suggests a GARCH(1,1) model while the one for squared returns suggests a GARCH(2,2) model but it is less clear.

Furthermore, the parameter estimates of the fitted ARMA model for the absolute data may yield initial estimates for maximum likelihood estimation of the GARCH model. When we fit an ARMA(1,1) model for the absolute simulated GARCH(1,1) process, we obtain the following parameter estimates.

```
arima(abs(garch11.sim),order=c(1,0,1))
```

```
##
## Call:
## arima(x = abs(garch11.sim), order = c(1, 0, 1))
##
## Coefficients:
##          ar1      ma1  intercept
##       0.9821  -0.9445     0.5077
## s.e.  0.0134   0.0220     0.0499
##
## sigma^2 estimated as 0.1486:  log likelihood = -232.97,  aic = 471.9
4
```

The estimates of β and α are 0.9445 and 0.03763, respectively. And ω can be estimated as the variance of the original data times the estimate of $1 - \alpha - \beta$, which equals 0.0073.

Estimation of Parameters

We use maximum likelihood estimators of parameters for GARCH models. The maximum likelihood estimators can be shown to be approximately normally distributed with the true parameter values as their means.

For the simulated GARCH(1,1) process, we specified a GARCH(1,1) model and considered the GARCH(2,2) model as a potentially suitable model. Parameter estimates of these models are obtained with the following code chunk.

```
g1=garch(garch11.sim,order=c(2,2),trace=FALSE)
summary(g1)
```

```
##
## Call:
## garch(x = garch11.sim, order = c(2, 2), trace = FALSE)
##
## Model:
## GARCH(2,2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -3.346827 -0.631881  0.008473  0.736112  3.202344
##
## Coefficient(s):
##     Estimate  Std. Error  t value Pr(>|t|)
## a0 1.835e-02   1.515e-02    1.211   0.2257
## a1 9.976e-16   4.723e-02    0.000   1.0000
## a2 1.136e-01   5.855e-02    1.940   0.0524 .
## b1 3.369e-01   3.696e-01    0.911   0.3621
## b2 5.100e-01   3.575e-01    1.426   0.1538
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 0.41859, df = 2, p-value = 0.8112
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 0.005298, df = 1, p-value = 0.942
```

```
g2=garch(garch11.sim,order=c(1,1),trace=FALSE)
summary(g2)
```

```
##
## Call:
## garch(x = garch11.sim, order = c(1, 1), trace = FALSE)
##
## Model:
## GARCH(1,1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -3.307030 -0.637977  0.009156  0.741977  3.019441
##
## Coefficient(s):
##     Estimate  Std. Error  t value Pr(>|t|)
## a0  0.007575    0.007590    0.998   0.3183
## a1  0.047184    0.022308    2.115   0.0344 *
## b1  0.935377    0.035839   26.100   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 0.82911, df = 2, p-value = 0.6606
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 0.53659, df = 1, p-value = 0.4638
```

For the GARCH(2,2) model, all of the parameters are insignificant at 5% level while all parameters but a0 of GARCH(1,1) are significant at 5% level. AIC of GARCH(1,1) model is less than that of GARCH(2,2) model. Thus, the GARCH(1,1) model gives a better fit for the series. Notice that all of the reported confidence intervals for ARCH(1,1) model contains their true values.

For the CREF return data, we detected either a GARCH(1,1) or GARCH(2,2) model. Their parameter estimates are given below.

```
m1=garch(x=r.cref,order=c(1,1),trace=FALSE)
summary(m1)
```

```
##
## Call:
## garch(x = r.cref, order = c(1, 1), trace = FALSE)
##
## Model:
## GARCH(1,1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.78577 -0.61949  0.08695  0.67933  3.30810
##
## Coefficient(s):
##     Estimate  Std. Error  t value Pr(>|t|)
## a0   0.01633     0.01237    1.320   0.1869
## a1   0.04414     0.02097    2.105   0.0353 *
## b1   0.91704     0.04570   20.066   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 1.0875, df = 2, p-value = 0.5806
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 0.77654, df = 1, p-value = 0.3782
```

```
m2=garch(x=r.cref,order=c(2,2),trace=FALSE)
summary(m2)
```

```
##
## Call:
## garch(x = r.cref, order = c(2, 2), trace = FALSE)
##
## Model:
## GARCH(2,2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.82899 -0.60897  0.08606  0.69719  3.27461
##
## Coefficient(s):
##     Estimate  Std. Error  t value Pr(>|t|)
## a0 1.876e-02   2.826e-02    0.664    0.507
## a1 7.658e-03   4.982e-02    0.154    0.878
## a2 4.079e-02   8.168e-02    0.499    0.618
## b1 9.067e-01   1.564e+00    0.580    0.562
## b2 8.107e-15   1.441e+00    0.000    1.000
##
## Diagnostic Tests:
##  Jarque Bera Test
##
## data:  Residuals
## X-squared = 1.0338, df = 2, p-value = 0.5964
##
##
##  Box-Ljung test
##
## data:  Squared.Residuals
## X-squared = 0.085809, df = 1, p-value = 0.7696
```

We get a better fit with GARCH(1,1) for the CREF data in terms of significant tests and AIC.

Model Diagnostics

We need to confirm that the model is correctly specified with all model assumptions are supported by the data. The standardized residuals are defined as $\hat{\epsilon}_t = r_t/\hat{\Sigma}_{t|t-1}$ which are approximately iid if the model is correctly specified. We use the same diagnostic tools like those used to check ARIMA models for GACH models.

For the simulated GARCH(1,1) series, both the Shapiro-Wilk test and the Jarque-Bera test suggest the normality of standardised residuals.

```
jarque.bera.test(garch11.sim)
```

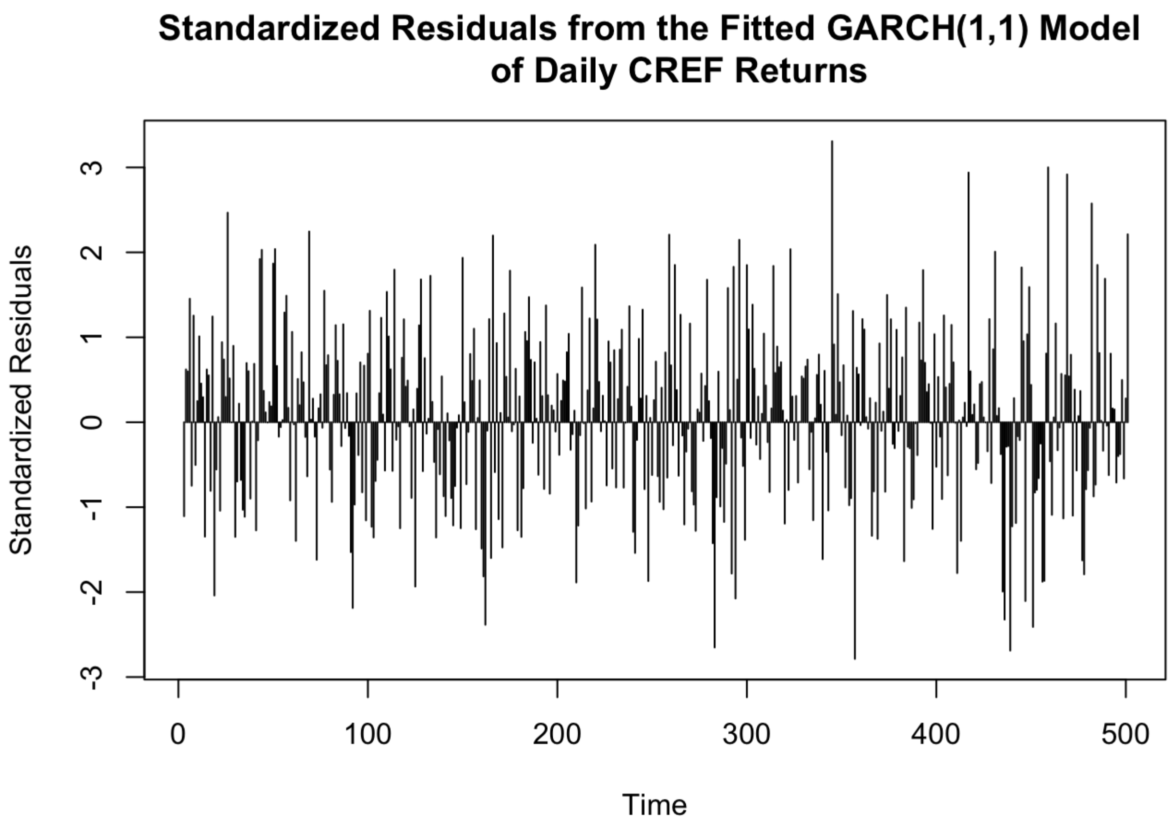
```
##
##  Jarque Bera Test
##
## data:  garch11.sim
## X-squared = 1.6269, df = 2, p-value = 0.4433
```

```
shapiro.test(garch11.sim)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  garch11.sim
## W = 0.99765, p-value = 0.714
```

Time series plot of standardised residuals is given below.

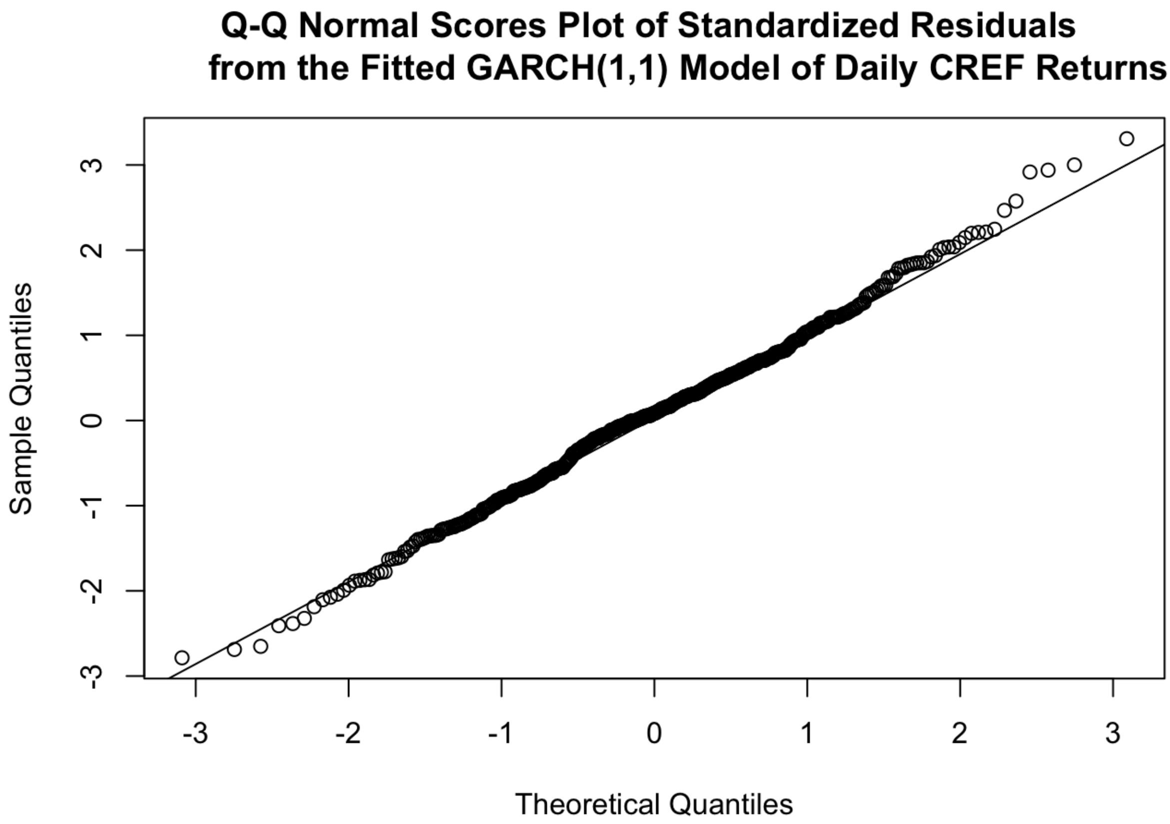
```
plot(residuals(m1),type='h',ylab='Standardized Residuals',main="Standar
  dized Residuals from the Fitted GARCH(1,1) Model
      of Daily CREF Returns")
```



There is some tendency for the residuals to be larger in magnitude towards the end of the study period, perhaps suggesting that there is some residual pattern in the volatility.

The Q-Q plot of standardised residuals is shown below.

```
qqnorm(residuals(m1),main="Q-Q Normal Scores Plot of Standardized Resid
  uals
       from the Fitted GARCH(1,1) Model of Daily CREF Returns")
qqline(residuals(m1))
```

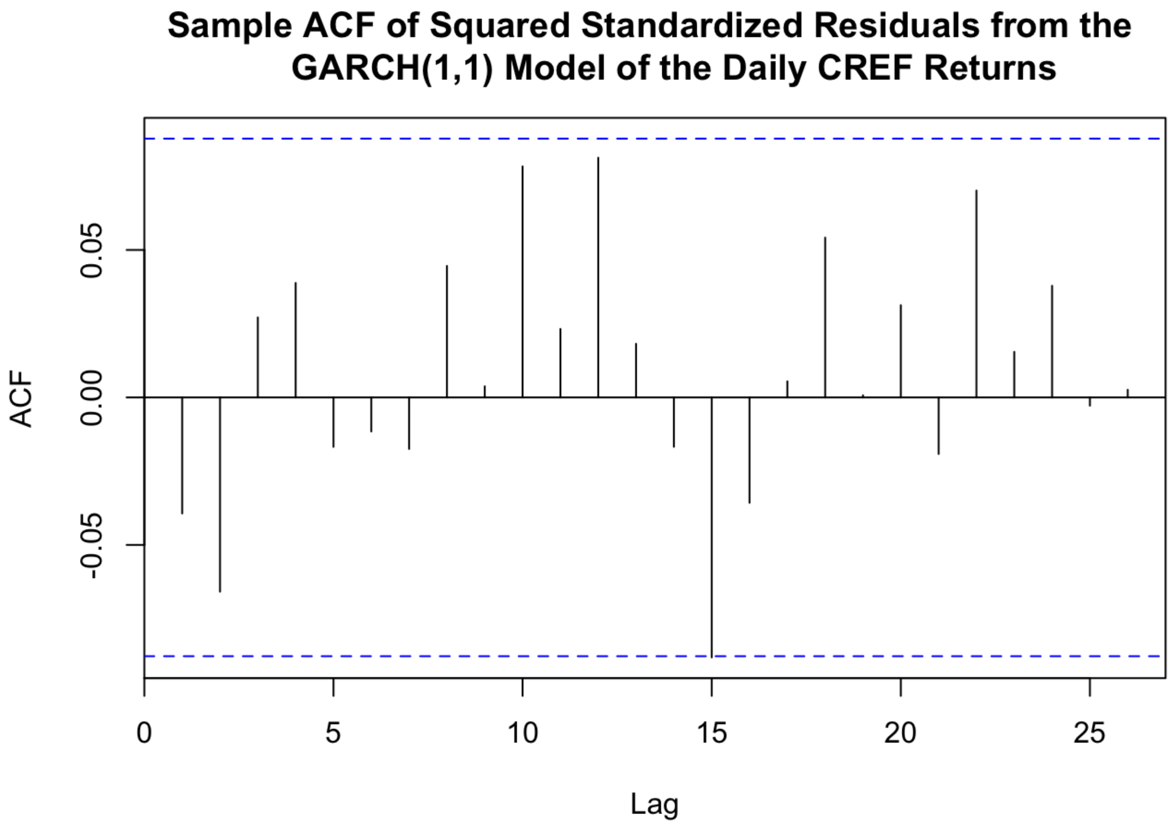


The Q-Q plot supports the inference of not rejecting the normality of residuals.

To check whether the residuals are iid or not, we use the sample ACF of residuals. As a formal test, the Ljung-Box test can be used. However, the portmanteau test does not have strong power against uncorrelated and yet serially dependent innovations. In fact, we start out with the assumption that the return data are uncorrelated, so the preceding test is of little interest.

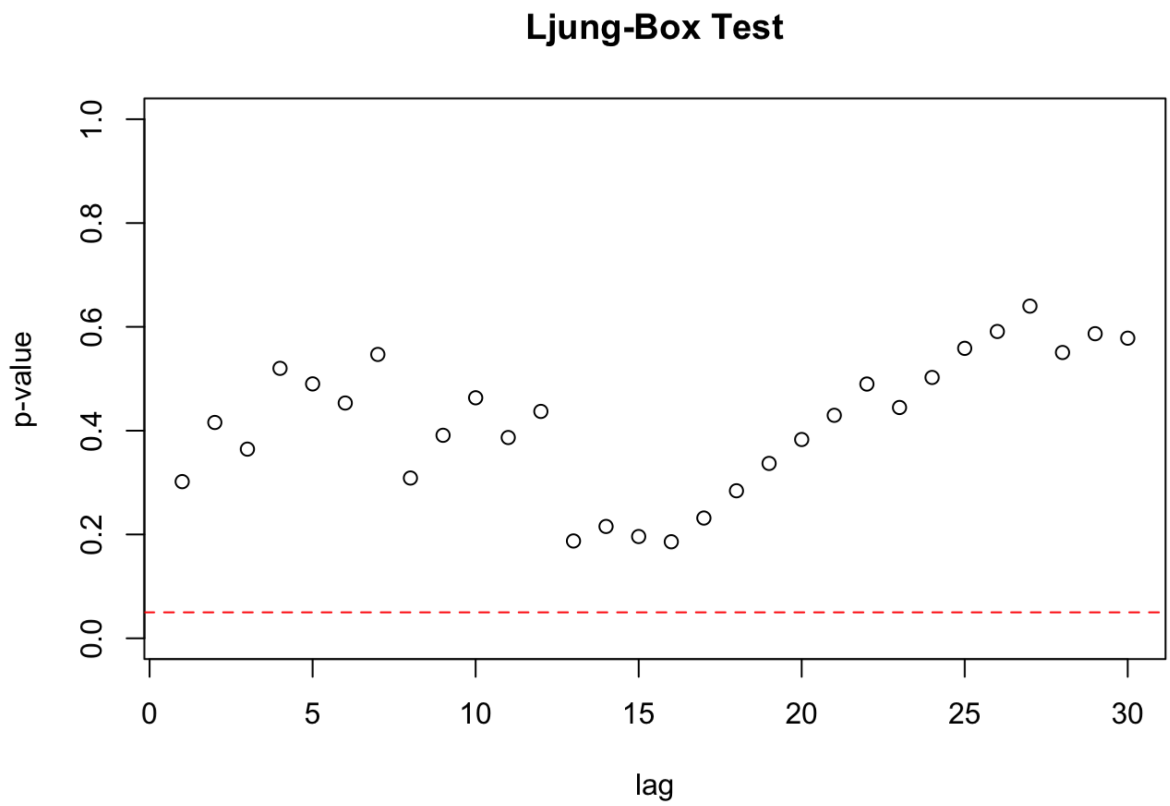
To reach more powerful conclusions, we can use *generalised portmanteau test*. For the CREF data, sample ACF of the squared standardized residuals from the fitted GARCH(1,1) model and the p-values of the generalized portmanteau tests with the squared standardized residuals are obtained with the following code chunk.

```
acf(residuals(m1)^2,na.action=na.omit,main="Sample ACF of Squared Stand
  ardized Residuals from the
    GARCH(1,1) Model of the Daily CREF Returns")
```




```
res.model = m1$residuals[2:500]
LBQPlot(res.model, lag.max = 30, SquaredQ = TRUE)
```

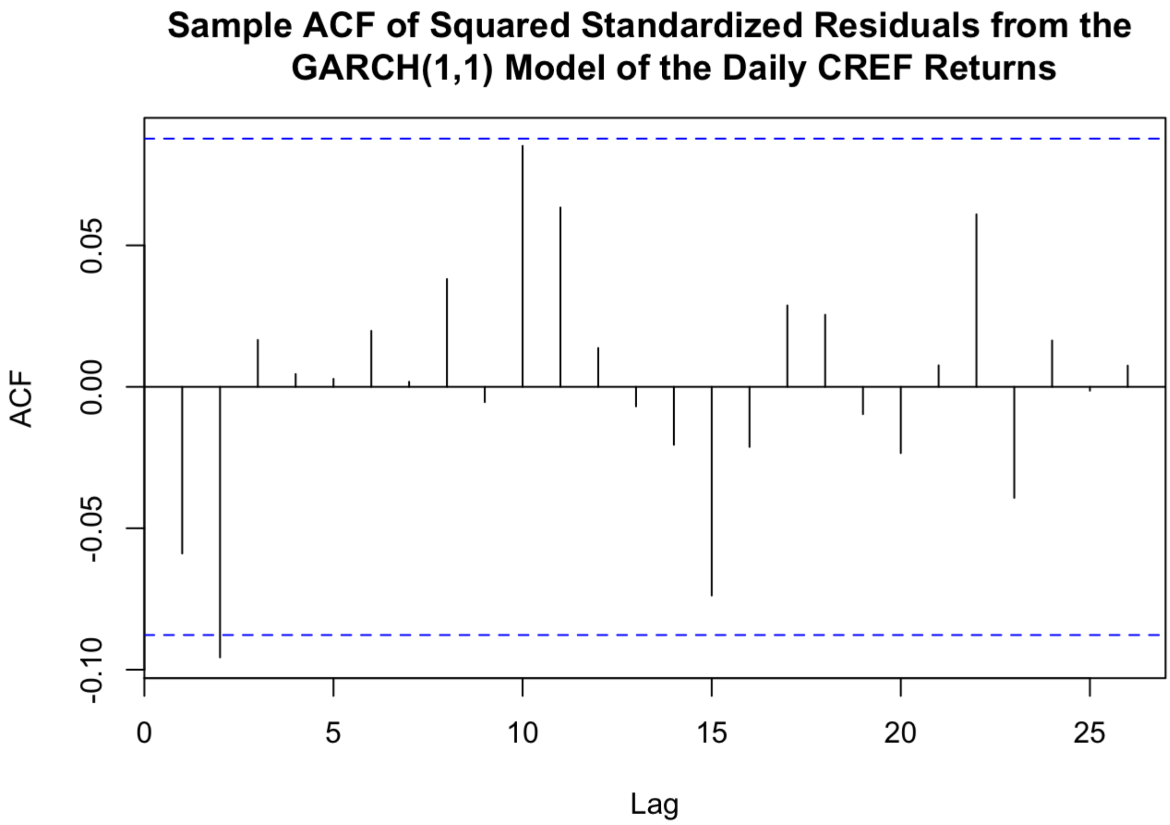
```
## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a m
ultiple
## of shorter object length
```



There is no lag for which ACF values exceed the limits and all p-values are higher than 5%, suggesting that the squared residuals are uncorrelated over time, and hence the standardized residuals may be independent.

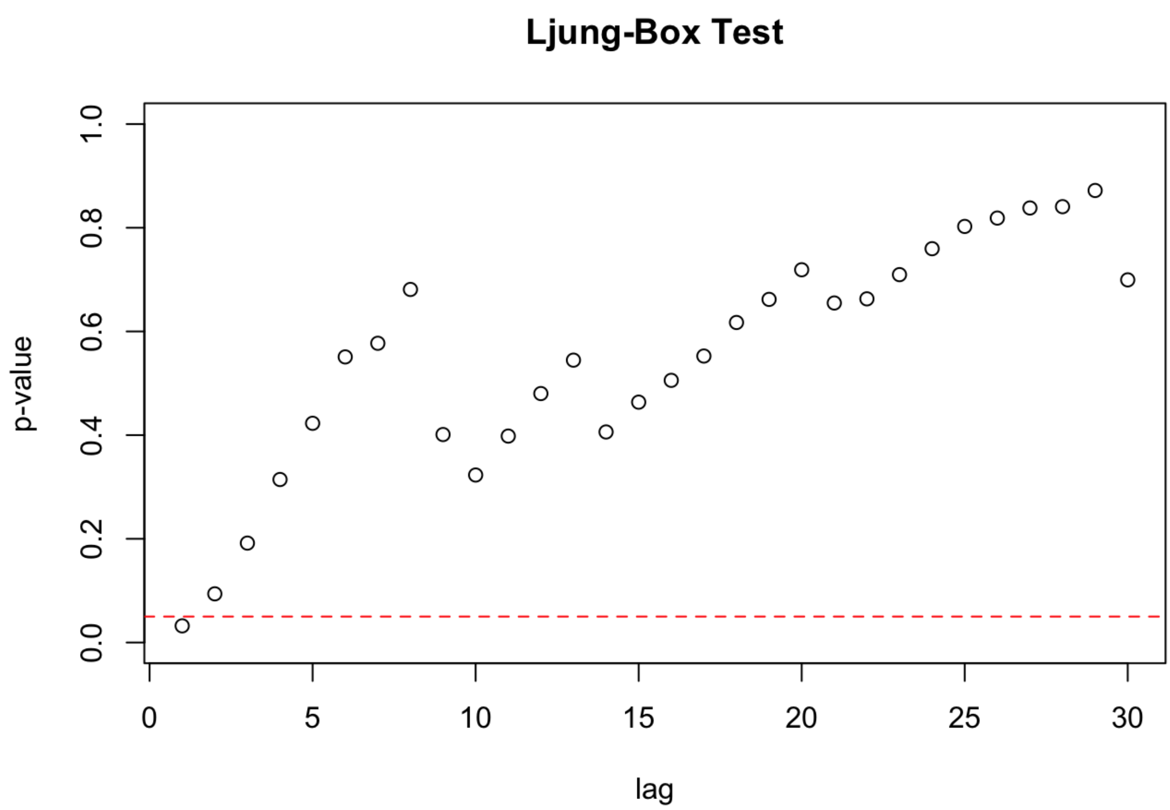
For the absolute standardised residuals we get the following plots:

```
acf(abs(residuals(m1)),na.action=na.omit,main="Sample ACF of Squared St
  andardized Residuals from the
     GARCH(1,1) Model of the Daily CREF Returns")
```




```
res.model = m1$residuals[2:500]
LBQPlot(abs(res.model), lag.max = 30)
```

```
## Warning in (ra^2)/(n - (1:lag.max)): longer object length is not a m
ultiple
## of shorter object length
```



The lag 2 autocorrelation of the absolute residuals is significant according to the nominal critical limits shown. Furthermore, the generalized portmanteau test is significant for lag 2. So, we further have a look at the EACF table.

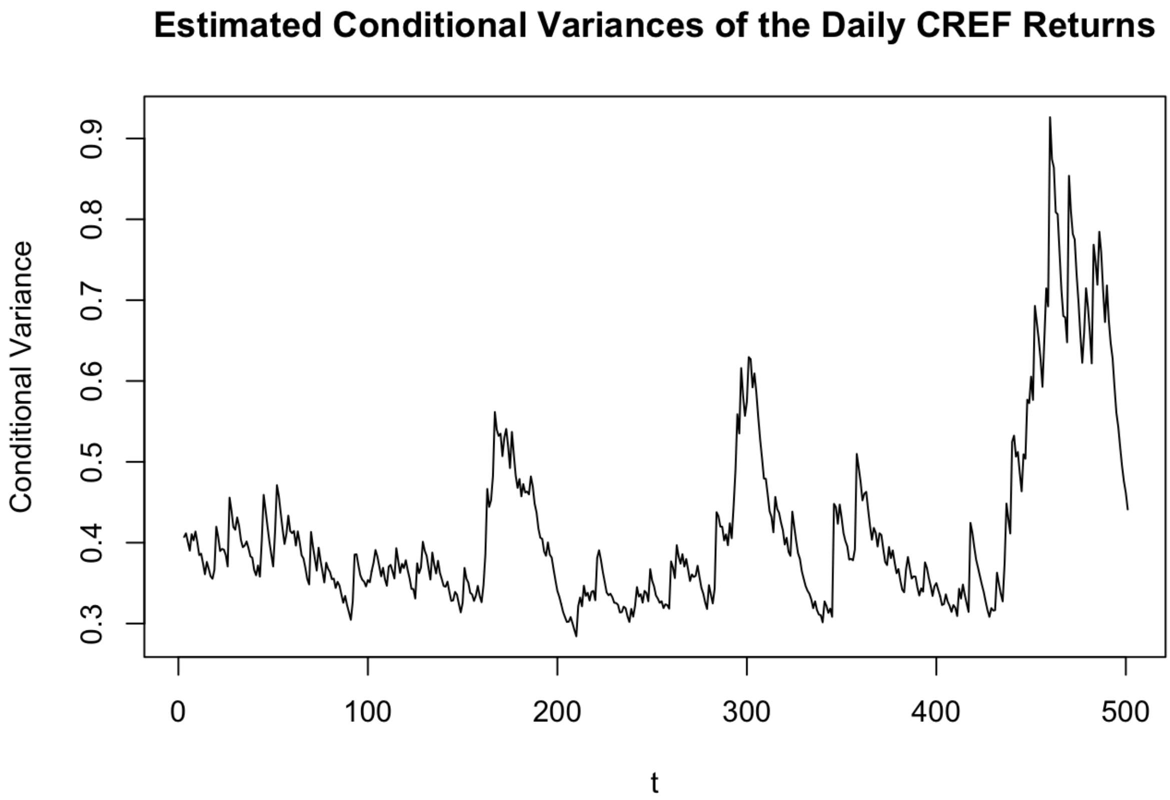
```
eacf(abs(res.model))
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 o x o o o o o o o o o  o  o  o
## 1 x x o o o o o o o o o  o  o  o
## 2 o o o o o o o o o o o  o  o  o
## 3 x o o o o o o o o o o  o  o  o
## 4 x o o o o o o o o o o  o  o  o
## 5 x x x o o o o o o o o  o  o  o
## 6 x x o o o o o o o o o  o  o  o
## 7 x x x o x o o o o o o  o  o  o
```

EACF table (not shown) of the absolute standardized residuals suggests an MA(2) model for the absolute residuals and hence points to the possibility that the CREF returns may be identified as a GARCH(1,2) process. However, the fitted GARCH(1,2) model to the CREF data did not improve the fit, as its AIC was 978.2-much higher than 969.6, that of the GARCH(1,1) model. Therefore, we conclude that the fitted GARCH(1,1) model provides a good fit to the CREF data.

We use GACRH(1,1) model to obtain future predictions of conditional variances. The following plot shows the within-sample estimates of the conditional variances, which capture several periods of high volatility, especially the one at the end of the study period.

```
plot((fitted(m1)[,1])^2,type='l',ylab='Conditional Variance',xlab='t',m
  ain="Estimated Conditional Variances of the Daily CREF Returns")
```



At the final time point, the squared return equals 2.159, and the conditional variance is estimated to be 0.4411. For example, the one-step-ahead forecast of the conditional variance equals $0.01633 + 0.04414 * 2.159 + 0.91704 * 0.4411 = 0.5161$ The two-step forecast of the conditional variance equals $0.01633 + 0.04414 * 0.5161 + 0.91704 * 0.5161 = 0.5124$ and so forth, with the longer lead forecasts eventually approaching 0.42066, the long-run variance of the model.

Some Extensions of the GARCH Model

The GARCH model assumes that the conditional mean of the time series is zero. Even for financial time series, this strong assumption need not always hold. In the more general case, the conditional mean structure may be modeled by some ARMA(u,v) model, with the white noise term of the ARMA model modeled by some GARCH(p, q) model. Specifically, let $\{Y_t\}$ be a time series given by

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \cdots + \phi_u Y_{t-u} \theta_0 + e_t + \theta_1 e_{t-1} + \cdots + \theta_v e_{t-v} \\ e_t &= \sigma_{t|t-1} \varepsilon_t \\ \sigma^2_{t|t-1} &= \omega + \alpha_1 e^2_{t-1} + \cdots + \alpha_q e^2_{t-q} + \beta_1 \sigma^2_{t-1|t-2} + \cdots + \beta_p \sigma^2_{t-p|t-p-1} \end{aligned}$$

In this setting, the ARMA orders can be identified based on the time series $\{Y_t\}$, whereas the GARCH orders may be identified based on the squared residuals from the fitted ARMA model.

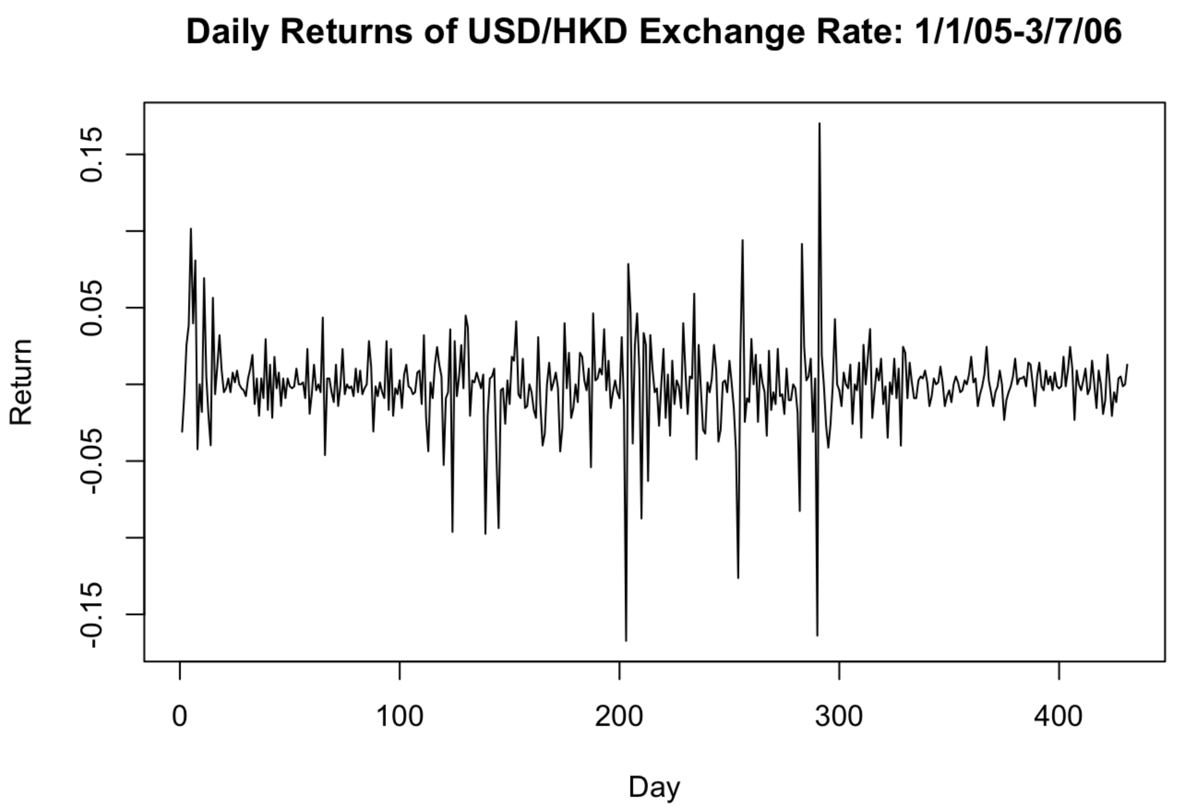
Another direction of generalization concerns nonlinearity in the volatility process. For financial data, this is motivated by a possible asymmetric market response that may, for example, react more strongly to a negative return than a positive return of the same magnitude. The idea can be simply illustrated in the setting of an ARCH(1) model, where the asymmetry can be modeled by specifying that

$$\sigma_{t|t-1}^2 = \omega + \alpha e_{t-1}^2 + \gamma \min(e_{t-1}, 0)^2.$$

Such a model is known as a GJR model-a variant of which allows the threshold to be unknown and other than 0.

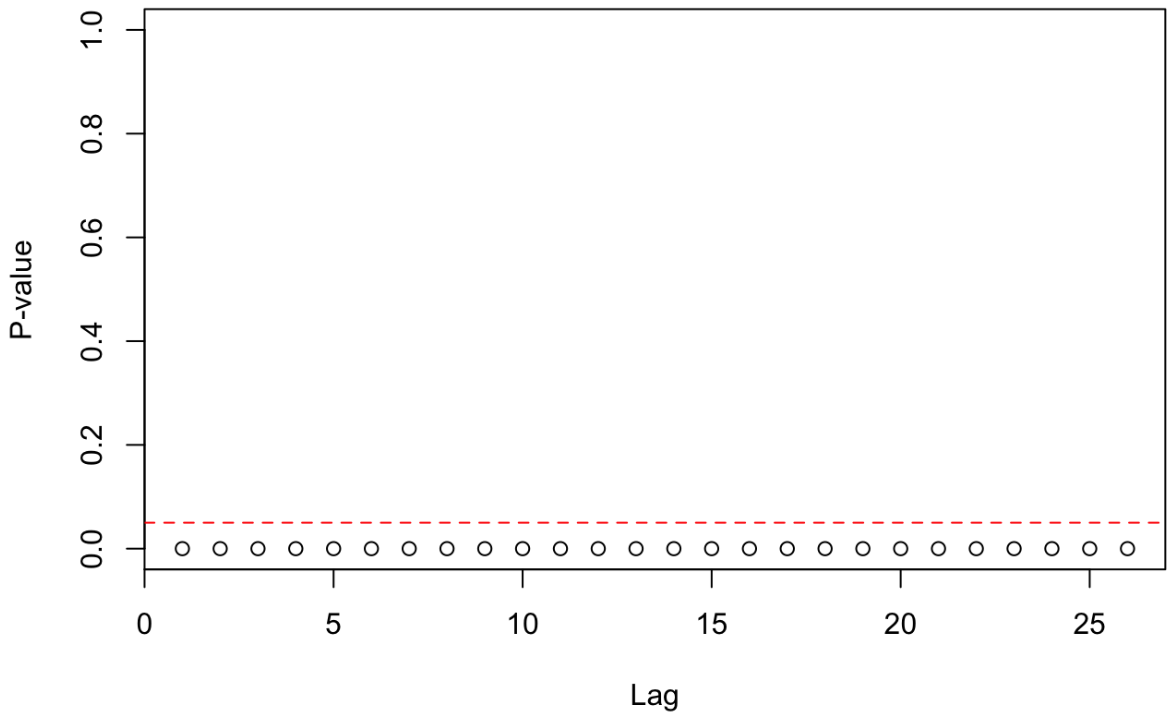
As an illustration for the ARIMA + GARCH model, we consider the daily USD/HKD (U.S. dollar to Hong Kong dollar) exchange rate from January 1, 2005, to March 7, 2006, altogether 431 days of data. The returns of the daily exchange rates are shown in the following time series plot. The series appears to be stationary, although volatility clustering is evident in the plot.

```
data(usd.hkd)
plot(ts(usd.hkd$hkrate,freq=1),type='l',xlab='Day', ylab='Return',main=
  "Daily Returns of USD/HKD Exchange Rate: 1/1/05-3/7/06")
```



It is interesting to note that the need for incorporating ARCH in the data is also supported by the McLeod-Li test applied to the residuals of the AR(1) + outlier model. The tests are all significant when the number of lags of the autocorrelations of the squared residuals ranges from 1 to 26, displaying strong evidence of conditional heteroscedasticity.

```
attach(usd.hkd)
McLeod.Li.test(arima(hkrate,order=c(1,0,0), xreg=data.frame(outlier1)))
```



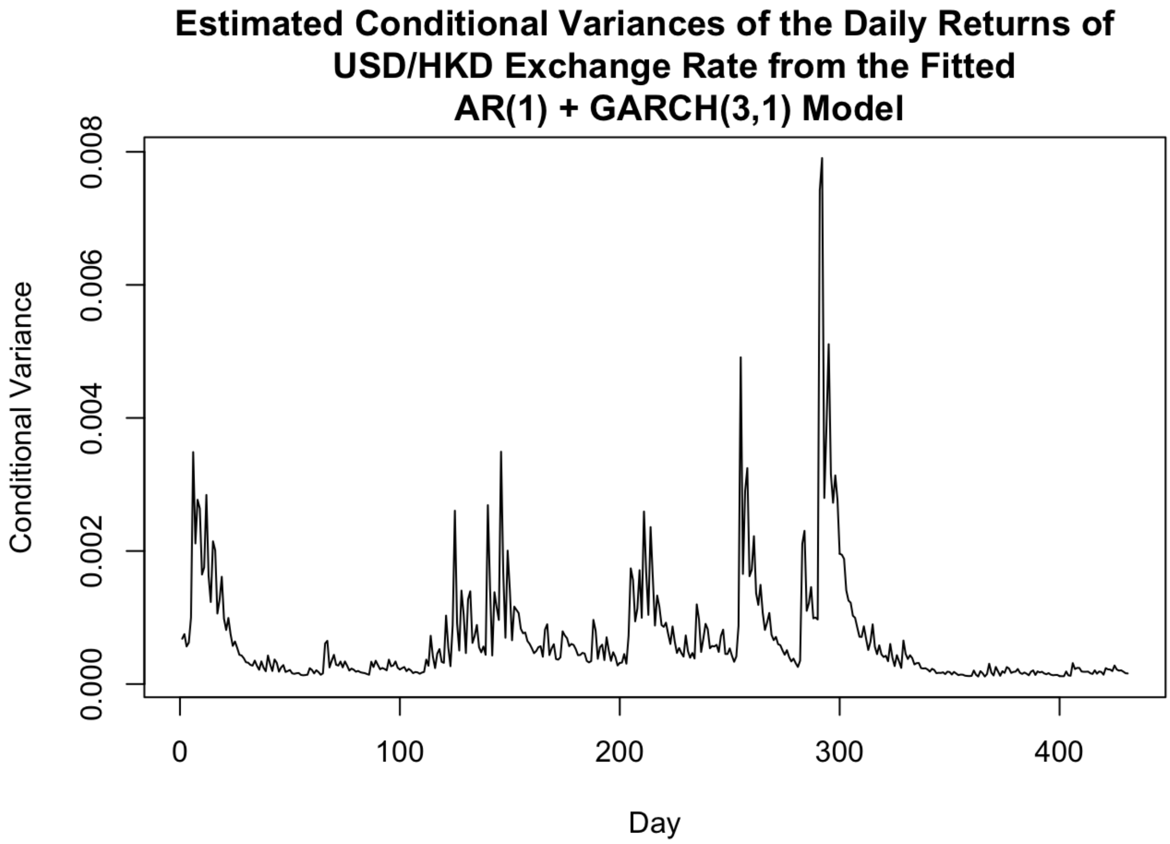
An AR(1) + GARCH(3,1) model was fitted to the (raw) return data with an additive outlier one day after July 22, 2005, the date when China revalued the yuan by 2.1% and adopted a floating-rate system for it. The intercept term in the conditional mean function was found to be insignificantly different from zero and hence is omitted from the model. Thus we take the returns to have zero mean unconditionally. The fitted model has an AIC = 2070.9, being smallest among various competing (weakly) stationary models. Interestingly, for lower GARCH orders $(p \leq 2)$, the fitted models are nonstationary, but the fitted models are largely stationary when the GARCH order is higher than 2. As the data appear to be stationary, we choose the AR(1) + GARCH(3,1) model as the final model.

AIC Values for Various Fitted Models for the Daily Returns of the USD/HKD Exchange Rate

AR order	GARCH order (p)	ARCH order (q)	AIC	Stationarity
0	3	1	−1915.3	nonstationary
1	1	1	−2054.3	nonstationary
1	1	2	−2072.5	nonstationary
1	1	3	−2051.0	nonstationary
1	2	1	−2062.2	nonstationary
1	2	2	−2070.5	nonstationary
1	2	3	−2059.2	nonstationary
1	3	1	−2070.9	stationary
1	3	2	−2064.8	stationary
1	3	3	−2062.8	stationary
1	4	1	−2061.7	nonstationary
1	4	2	−2054.8	stationary
1	4	3	−2062.4	stationary
2	3	1	−2066.6	stationary

Estimated conditional variances from the model fitted to the Hong Kong exchange rate data are shown below:

```
plot(ts(usd.hkd$v,freq=1),type='l',xlab='Day', ylab='Conditional Varian
  ce',main = "Estimated Conditional Variances of the Daily Returns of
      USD/HKD Exchange Rate from the Fitted
      AR(1) + GARCH(3,1) Model")
```



And parameter estimates are shown below:

Fitted AR(1) + ARCH(3,1) Model for Daily Returns of USD/HKD Exchange Rate

Coefficient	Estimate	Std. error	*t*-ratio	*p*-value
AR1	0.1635	0.005892	21.29	0.0022
ARCH0 (ω)	2.374×10^{-5}	6.93×10^{-6}	3.42	0.0006
ARCH1 (α_1)	0.2521	0.0277	9.09	< 0.0001
GARCH1 (β_1)	0.3066	0.0637	4.81	< 0.0001
GARCH2 (β_2)	−0.09400	0.0391	−2.41	0.0161
GARCH3 (β_3)	0.5023	0.0305	16.50	< 0.0001
Outlier	−0.1255	0.00589	−21.29	< 0.0001

Summary

In this module, we dealt with autoregressive conditional heteroscedasticity (ARCH) and generalized autoregressive conditional heteroscedasticity (GARCH) models. In this sense

- ARHC models were then introduced in an attempt to model the changing variance of a time series.
- The ARCH model of order 1 was thoroughly explored from identification through parameter estimation and prediction.
- The GARCH models were also thoroughly explored with respect to identification, maximum likelihood estimation, prediction, and model diagnostics. - Examples with both simulated and real time series data were used to illustrate the ideas.