**Time Series and Prediction (Part II)**

*Assignment 1*

Student 1: Niall O’Neill

Student 2: Karolina Sidlauskaite

**Introduction**

In this report we will be analysing the financial return time series data of the Financial Times Stock Exchange 100 (FTSE 100), which is a share index of the 100 companies listed on the London Stock Exchange with the highest market capitalisation. The series will be analysed in terms of its sample properties, and also a univariate GARCH model will be fitted to the series.

**Data analysis**

The data analysed contains daily series of quotes during the period of January 5, 2000 and March 11, 2019. The analysis was performed using the Matlab software. This is the snapshot of the code run to generate the plots for the sample properties of the financial time series:

|  |
| --- |
| function [s\_mean, s\_sd, s\_k, s\_sk] = sample\_properties(namefile)    data = readtable(namefile);    % pre-process the data, show returns  n\_rows = height(data);  ret = [0];  for i = 2:n\_rows  temp = 100 \* (log(data.price(i)) - log(data.price(i-1)));  ret = cat(1, ret, temp);  end  data.ret = ret;    data = table2timetable(data);    subplot(3,2,1); stackedplot(data, 'Title', 'Plot of Prices and FTSE 100 Returns', 'Fontsize', 16, 'DisplayLabels', {'Price', 'Returns in %'});    ret = data.ret;  s\_mean = mean(ret);  s\_sd = sqrt(var(ret));  s\_k = kurtosis(ret);  s\_sk = skewness(ret);    % next we obtain the squared observations  ret2 = (ret).^2;  subplot(3,2,2); plot(ret2);  title('Plot of Squared Returns');  xlim([1 4840]);  legend('hide');  grid('off');    % now compute autocorrelation function of returns  subplot(3,2,3); autocorr(ret, 50);  title('ACF of Returns');    % and squared returns  subplot(3,2,4); autocorr(ret2,50);  title('ACF of Squared Returns');    % finally, calculate the cross-correlation between squared returns and past returns  subplot(3,2,5); crosscorr(ret2, ret, 50);  title('CCF Between Squared Returns and Past Returns'); |

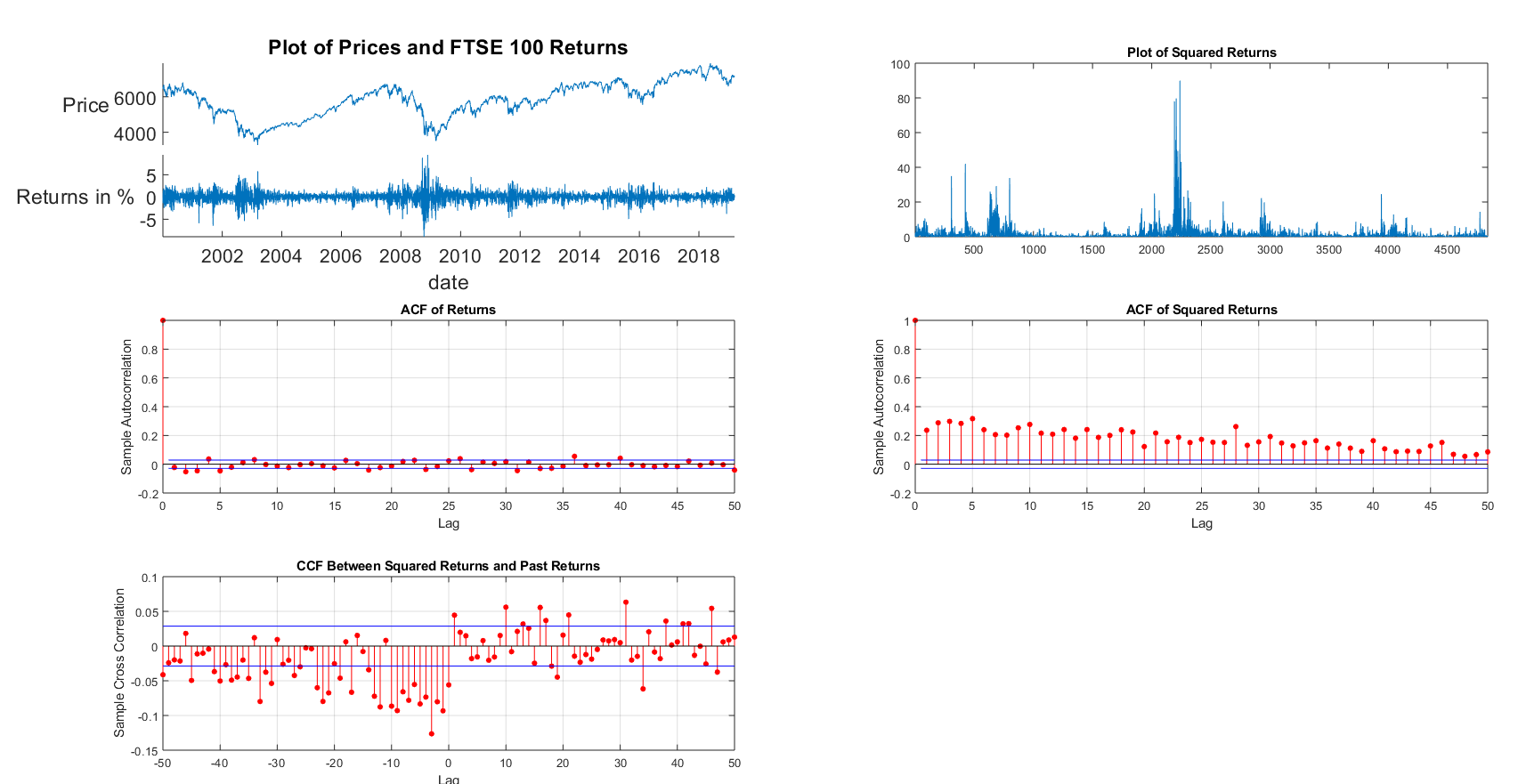
We can then run the above function by:

|  |
| --- |
| [s\_mean, s\_sd, s\_k, s\_sk] = sample\_properties(‘ftse.xlsx’); |

This returns the following:

|  |  |
| --- | --- |
| **Mean** | 0.0019 |
| **Variance** | 1.1456 |
| **Skewness** | 9.4472 |
| **Kurtosis** | -0.1661 |

From which we can see that the series is positively skewed with a negative kurtosis.

And generates the plots as below:

Plot of of prices and FTSE 100 returns: from this plot we can see that there were quite a few smaller shocks in between 2001 and 2003 (most during 2003), but there were also quite a few big shocks around the year of 2009, when the price of the stock has decreased quite a bit. It appears that this drop in the price has been caused by the HSBC which dragged the market downwards by confirming a 12.5 billion pounds rights issue. Historically, this meant that the stock price of FTSE 100 has fallen to a six-year low, which was last seen in 2003 *[The Telegraph, 2009]*.

Plot of squared returns: as previously, we can see some shocks appear at the beginning, most likely referring to the same 2003, and large shocks around the middle, which almost likely refer to the events mentioned above in 2009.

ACF of returns: this displays the ACF of returns along with a 95% confidence interval, from where we can see that most autocorrelations lie within the interval, however, the first autocorrelation appears to be widely outside of the bounds, while the rest are just around zero, which means there is amoving average term in the data. Since there is only one significant correlation, it indicates that the order of the moving average term is 1.

ACF of squared returns: here we can see that the autocorrelation is significant even at long lags, hence providing evidence for the predictability of volatily (or, in other words, measure for market uncertainty), given the persistence of the autocorrelations. Note also that the ACF of squared returns is always significantly positive decaying very slowly.

CCF between squared and past returns: the cross correlation function is used to determine whether there is a relationship between the squared returns and past returns. The most dominant cross correlation (a negative one) occurs at lag -4.

**Univariate GARCH models**

In this section we will be fitting a GARCH model to the FTSE 100 time series returns. First, we will be fitting a GARCH(1,1) model with errors that follow a Gaussian distribution to the daily return. As previously, we use Matlab to perform this analysis (note we are using the same pre-processed *data* object).

First, we test for the existence of conditional heteroscedasticity using the ARCH test:

|  |
| --- |
| ret = data.ret;  ret1 = ret - ones(size(ret,1), 1) \* mean(ret);  [h, pValue, stat, cValue] = archtest(ret1); |

Which returns the following:

|  |  |
| --- | --- |
| **Logical value** | 1 |
| **p-value** | 0 |
| **Test statistic** | 270.1192 |
| **Critical value** | 3.8415 |

And so the null hypothesis can be rejected, and then we fit the GARCH(1,1) with normal errors:

|  |
| --- |
| Mdl = garch('GARCHLags', 1, 'ARCHLags', 1, 'distribution', 'Gaussian');  [EstMdl, EstParamCov, logL] = estimate(Mdl, ret1); |

Where we get:

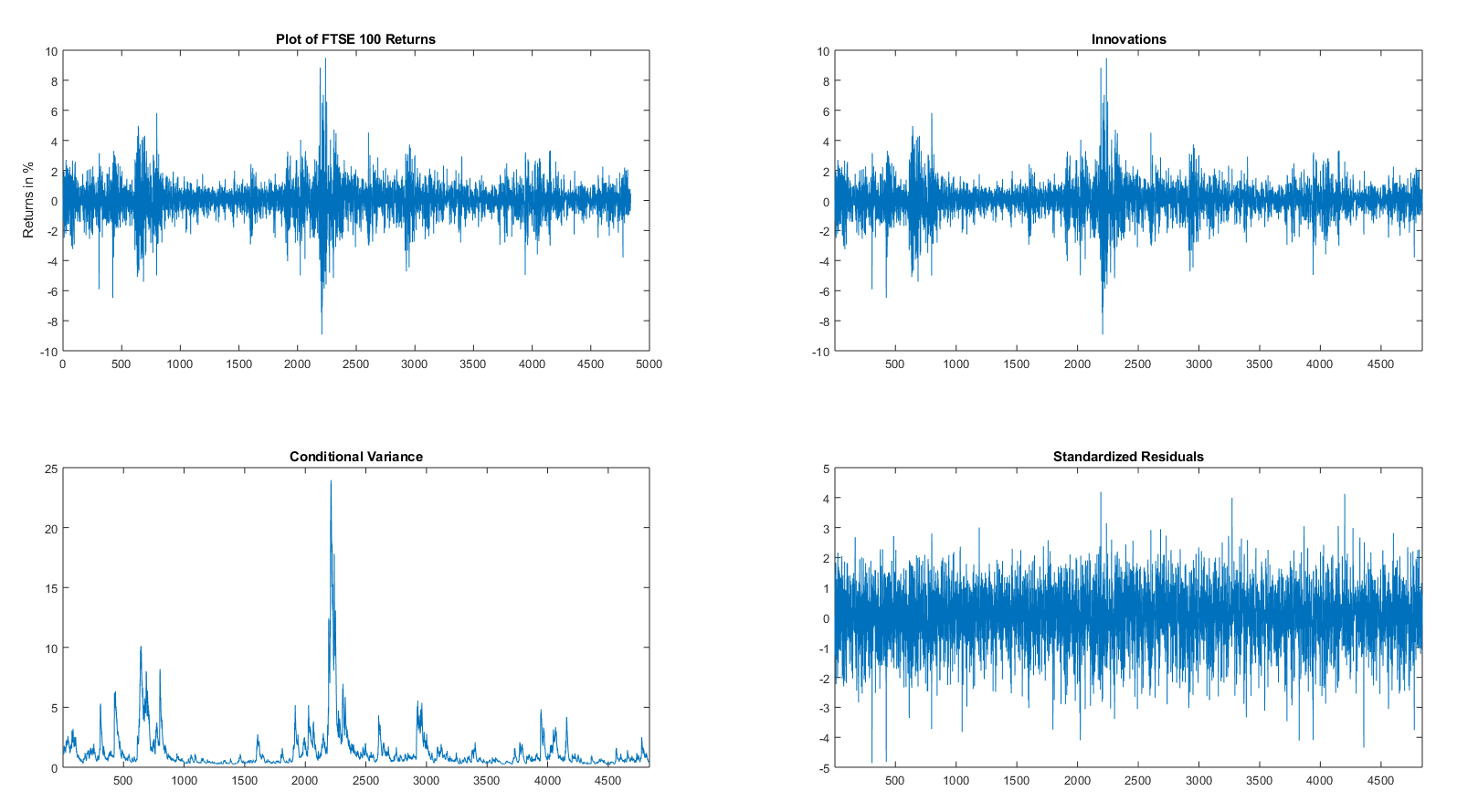
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Value** | **Standard Error** | **t-statistic** | **p-value** |
| **Constant** | 0.016998 | 0.0024843 | 6.8419 | 7.8138e-12 |
| **GARCH(1)** | 0.87764 | 0.0077763 | 112.86 | 0 |
| **ARCH(1)** | 0.1093 | 0.0068781 | 15.891 | 7.3451e-57 |

Now we check the goodness of fit for this GARCH(1,1) using the AIC and BIC values:

|  |
| --- |
| numParams = sum(any(EstParamCov));  T=length(ret1);  [aic,bic] = aicbic(logL,numParams,T); |

Which returns AIC = 1.3098e+04 and BIC = 1.3118e+04, and both are quite large, so we may consider fitting a different model to minimize those quantities. Now, to further evaluate the model, let’s plot the innovations, estimated volatility and the standardised residuals:

|  |
| --- |
| [cond\_variance\_garch] = infer(EstMdl,ret1);  subplot(2,2,1); plot(ret);    title('Plot of FTSE 100 Returns');  ylabel('Returns in %');  legend('hide');  grid('off');    Innovations = ret1;  subplot(2,2,2); plot(Innovations);  title('Innovations');  legend('hide');  grid('off');  xlim([1 4840]);    subplot(2,2,3); plot(cond\_variance\_garch);  title('Conditional Variance');  legend('hide');  grid('off');  xlim([1 4840]);    sd\_residuals = ret1./cond\_variance\_garch.^0.5;  subplot(2,2,4); plot(sd\_residuals);  title('Standardized Residuals');  legend('hide');  grid('off');  xlim([1 4840]); |

And so we get the following:

From above we can see very big shocks around days 400 – 800 and 2100 – 2400, which, as mentioned previously, refer to the negative shocks in 2003 and 2009 respectively. It is known that larger negative shocks create more volatility than positive shocks of the same magnitude, hence GARCH model will under-estimate the amount of volatility generated by those negative shocks (especially the one that appeared in 2009) and will over-estimate the amount of volatitlity following any positive shocks, however little. And since large shocks cause more volatility than small ones, this means that the volatility will be under-estimated after a large shock (for example the large negative shock in 2009) and over-estimated after a smaller one.

To test whether the model is able to fit the asymmetry characteristics of the data, we use Engle and NG test:

|  |
| --- |
| v2=sd\_residuals.^2;  v2=v2(2:size(sd\_residuals,1),1);  %Here we lag the Innovations  innovationsL = lagmatrix(Innovations,1);  innovationsL=innovationsL(2:size(Innovations,1),1);    %Here we build the dummy  d=zeros(size(innovationsL,1),1);  for i=1:size(innovationsL,1)  if (innovationsL(i,1)<0)  d(i,1)=1;  end  end    %Now we do the regression of the test  X=[ones(size(innovationsL,1),1) d d.\*innovationsL innovationsL.\*(1-d)];    [b,bint,r,rint,stats]=regress(v2,X)    test=size(v2,1).\*stats(1,1);  pvalue = 1-chi2cdf(test,3);    C='We reject the null hypothesis';  if pvalue<0.05  disp(C);  end |

This returns the following statistics:

|  |  |
| --- | --- |
| **R2** | 0.0087 |
| **F-statistic** | 14.2096 |
| **p-value** | 0 |
| **Estimate of error variance** | 2.7942 |

Hence, we can reject the null hypothesis, which means that the model is not able to fit the asymmetry of the data. And so additional modelling structure incorporating the possibility of asymmetry into the variance equation is required.

Now to capture the asymmetry, let’s consider using a Threshold GARCH model, which is denoted by GJR(1,1):

|  |
| --- |
| Mdl = gjr('GARCHLags', 1, 'ARCHLags', 1, 'LeverageLags', 1, 'distribution', 't');  [EstMdl, EstParamCov, logL] = estimate(Mdl,ret1); |

Where we get:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Value** | **Standard Error** | **t-statistic** | **p-value** |
| **Constant** | 0.017104 | 0.0023857 | 7.1693 | 7.5374e-13 |
| **GARCH(1)** | 0.89863 | 0.0090591 | 99.197 | 0 |
| **Leverage(1)** | 0.16856 | 0.014316 | 11.774 | 5.2947e-32 |
| **DoF** | 12.921 | 2.2835 | 5.6585 | 1.5273e-08 |

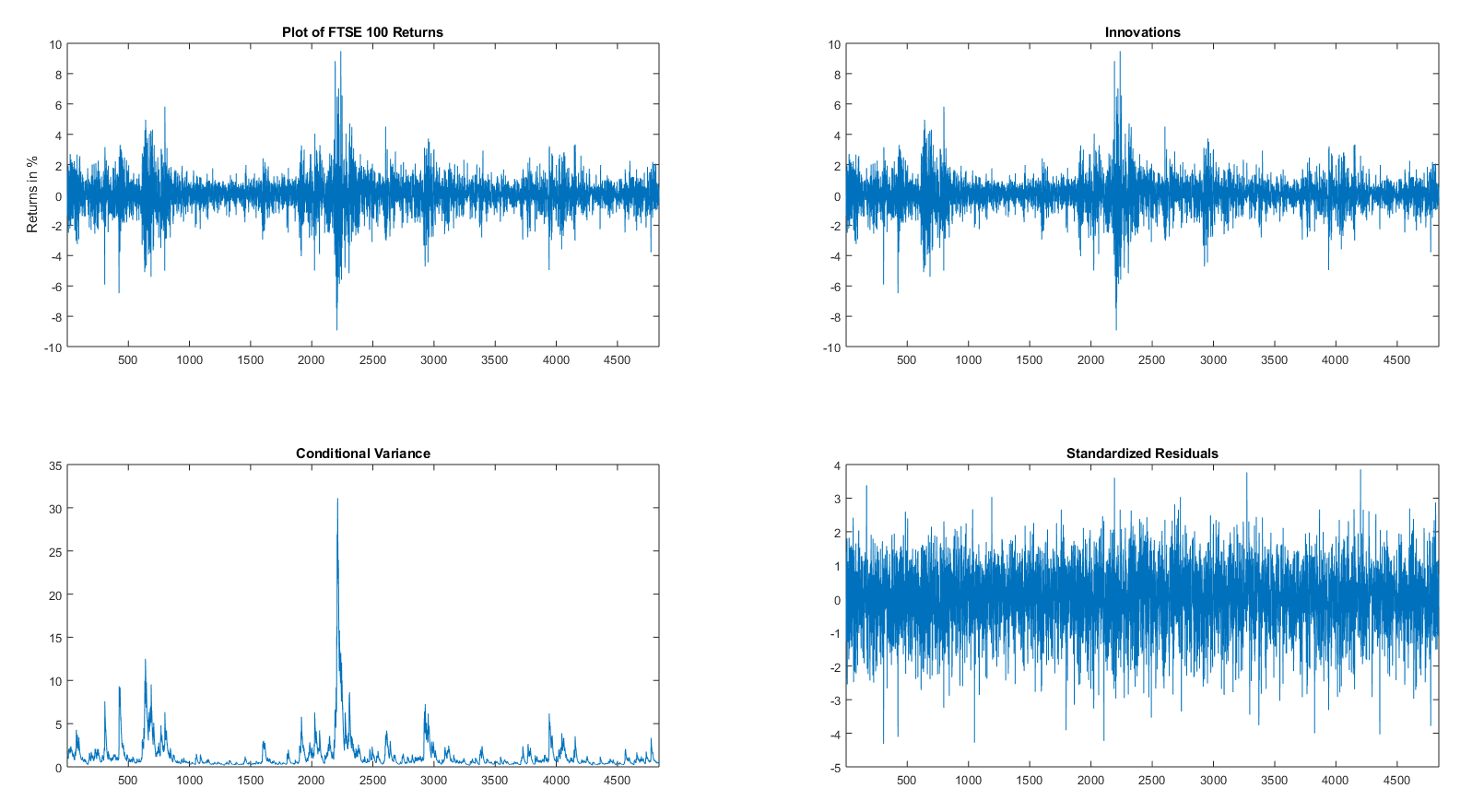
Same as for GARCH(1,1) we check the goodness of fit of GJR(1,1) using AIC and BIC values:

|  |
| --- |
| numParams = sum(any(EstParamCov));  T=length(ret1);  [aic,bic] = aicbic(logL,numParams,T); |

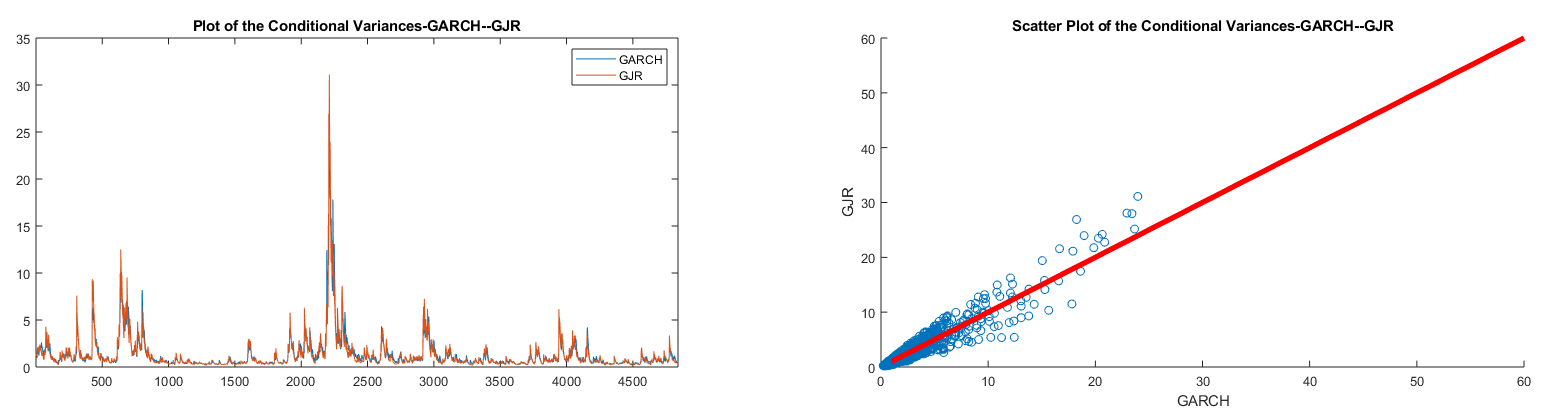
Which gives AIC = 1.2848e+04 and BIC = 1.2874e+04, which is slightly lower than what we got for GARCH(1,1), and so means fits our data better. As for GARCH(1,1), let’s plot the innovations, estimated volatility and the standardised residuals:

|  |
| --- |
| [cond\_variance\_gjr] = infer(EstMdl,ret1);  subplot(2,2,1); plot(ret1);    title('Plot of FTSE 100 Returns');  ylabel('Returns in %');  legend('hide');  grid('off');  xlim([1 4840]);    Innovations=ret1;  subplot(2,2,2); plot(Innovations);  title('Innovations');  legend('hide');  grid('off');  xlim([1 4840]);    subplot(2,2,3); plot(cond\_variance\_gjr);  title('Conditional Variance');  legend('hide');  grid('off');  xlim([1 4840]);    %construct the series os standardized residuals  sd\_residuals=ret1./cond\_variance\_gjr.^0.5;    subplot(2,2,4); plot(sd\_residuals);  title('Standardized Residuals');  legend('hide');  grid('off');  xlim([1 4840]); |

Which generates:



Now we can compare the two using their conditional variances:



References

The Telegraph (2009) - <https://www.telegraph.co.uk/finance/markets/4926462/FTSE-100-falls-to-six-year-low.html>