ECON90033 – QUANTITATIVE ANALYSIS OF FINANCE I

TUTORIAL 9

Download the t9e1 Excel data file from the subject website and save it to your computer or USB flash drive. Read this handout and complete the tutorial exercises before your tutorial class so that you can ask for help during the tutorial if necessary.

Heatwaves and Meteor Showers[1]

An important application of *GARCH* models examines how volatility is transmitted through different regions of the world during the course of a global financial trading day. The idea is that shocks to major markets of the world economy can have delayed effects on other markets because they operate in different time zones. One possible approach to model these spillover effects of shocks is based on the partition of each 24 hour calendar day into three major trading zones, namely, Japan (12am to 7am GMT), Europe (7am to 12:30pm GMT) and the United States (12:30pm to 9pm GMT):[2]

$$\underbrace{\overbrace{\text{12am}\cdots\text{7am}}^{\text{Japan}}\quad\overbrace{\text{7am}\cdots\text{12pm}}^{\text{Europe}}\quad\overbrace{\text{12pm}\cdots\text{9pm}}^{\text{U.S.}}}_{\text{One Trading Day}}$$

This illustrates that on any particular 24 hour GMT calendar day Europe lags behind Japan and the US lags behind Japan and Europe.

Given this calendar structure of the global trading day, the conditional variances in the three trading zones can be modelled with a trivariate *GARCH* model. Let r_{1t}, r_{2t}, and r_{3t} denote the daily log returns to the Japanese, the European, and the United States markets, respectively, and assume that their conditional distributions are normal with constant mean but potentially time varying variance, i.e.,

$$r_{it} : N(\mu_i, h_{it}) \quad , \quad i = 1,2,3$$

The shocks to these market are captured by the deviations of log returns from their respective means, i.e.,

---

[1] This part of the tutorial heavily relies on Section 13.6 of the prescribed textbook.
[2] GMT: Greenwich Mean Time is the mean solar time at the Royal Observatory in Greenwich, London, counted from midnight.

$$r_{it} - \mu_i = \varepsilon_{it}, \quad \varepsilon_{it} \mid \Omega_t : idN(0, h_{it}) \quad , \quad i = 1,2,3$$

and the contemporaneous shocks are supposed to be independent, i.e.,

$$E(\varepsilon_{1t}, \varepsilon_{2t}) = E(\varepsilon_{1t}, \varepsilon_{3t}) = E(\varepsilon_{2t}, \varepsilon_{3t}) = 0$$

As regards the conditional variances, h_{it} (i = 1,2,3), let's assume that each depends on its own lag, but not on the lags of the other two. In addition, based on the calendar structure of the global trading day, on any particular 24 hour GMT calendar day

(a) the conditional variance of the Japanese market can be affected by the previous day's shocks to all three markets,
(b) the conditional variance of the European market can be affected by the same day shock to the Japanese market and the previous day's shocks to the European and US markets,
(c) the conditional variance of the US market can be affected by the same day shocks to the Japanese and European markets and the previous day's shock to the US market.

Consequently, the conditional variance equation for Japanese market is given by

$$h_{1t} = \alpha_{10} + \alpha_{11}\varepsilon_{1,t-1}^2 + \alpha_{12}\varepsilon_{2,t-1}^2 + \alpha_{13}\varepsilon_{3,t-1}^2 + \beta_{11}h_{1,t-1}$$

for the European market it is

$$h_{2t} = \alpha_{20} + \gamma_{21}\varepsilon_{1t}^2 + \alpha_{22}\varepsilon_{2,t-1}^2 + \alpha_{23}\varepsilon_{3,t-1}^2 + \beta_{21}h_{2,t-1}$$

and for the US market it is

$$h_{3t} = \alpha_{30} + \gamma_{31}\varepsilon_{1t}^2 + \gamma_{32}\varepsilon_{2t}^2 + \alpha_{33}\varepsilon_{3,t-1}^2 + \beta_{33}h_{3,t-1}$$

In summary, each market is described with a constant mean equation and a *GARCH*(1,1) variance equation augmented with shocks to the other two markets.

Using vectors and matrices the three models can be combined and written as follows[3]:

$$\begin{bmatrix} r_{1t} - \mu_1 \\ r_{2t} - \mu_2 \\ r_{3t} - \mu_3 \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \quad , \quad \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} : N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} h_{1t} & 0 & 0 \\ 0 & h_{2t} & 0 \\ 0 & 0 & h_{3t} \end{bmatrix} \right)$$

and

$$\begin{bmatrix} h_{1t} \\ h_{2t} \\ h_{3t} \end{bmatrix} = \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \\ \alpha_{30} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \gamma_{21} & 0 & 0 \\ \gamma_{31} & \gamma_{32} & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t}^2 \\ \varepsilon_{2t}^2 \\ \varepsilon_{3t}^2 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ 0 & \alpha_{22} & \alpha_{23} \\ 0 & 0 & \alpha_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{2,t-1}^2 \\ \varepsilon_{3,t-1}^2 \end{bmatrix} + \begin{bmatrix} \beta_{11} & 0 & 0 \\ 0 & \beta_{22} & 0 \\ 0 & 0 & \beta_{33} \end{bmatrix} \begin{bmatrix} h_{1,t-1} \\ h_{2,t-1} \\ h_{3,t-1} \end{bmatrix}$$

---

[3] Do not worry about these formulas if you are unfamiliar with vectors and matrices.

In this framework there are two possible sources of volatility, termed as (a) *heatwave* and (b) *meteor shower*.

(a) Heatwave refers to volatility in any one region due to the previous day's volatility in the same region. It is captured by the β_{ii} parameters, and there is not heatwave on market i (i = 1, 2, 3) if $\beta_{ii} = 0$ (H_0) and there is if $\beta_{ii} \neq 0$ (H_A).

(b) Meteor shower refers to volatility in any one region due to the same day's volatility in the region preceding it in terms of calendar time. It is captured by the $\gamma_{i\neq j}$ parameters. Namely, there is not meteor shower from Japan to Europe if $\gamma_{21} = 0$ (H_0) and there is if $\gamma_{21} \neq 0$ (H_A). Similarly, there is not meteor shower from Japan to the US if $\gamma_{31} = 0$ (H_0) and there is if $\gamma_{31} \neq 0$ (H_A). Finally, there is not meteor shower from Europe to the US if $\gamma_{32} = 0$ (H_0) and there is if $\gamma_{32} \neq 0$ (H_A).

Exercise 1

The *t9e1.xlsx* file contains daily equity prices from 4 January 1999 to 2 April 2014, on the Dow Jones Industrial Average (*DJX*), the S&P 500 (*SPX*), the Hang Seng Index (*HSX*), the Nikkei (*NKX*), the Deutscher Aktien Index (*DAX*), and the Financial Times Stock Exchange (*UKX*).

Launch *RStudio*, create a new *RStudio* project and script, and name both *t9e1*. Import the data from the *t9e1.xlsx Excel* file, save it as *t9e1.RData*, and attach it to the project.

a) Compute the percentage log returns on the *DAX*, *NKX* and *SPX* indices. Plot the three return series, calculate the usual descriptive statistics for each of them and briefly discuss the results regarding skewness, kurtosis and normality.

The percentage log returns for *NKX* can be calculated and plotted by executing the following commands:[4]

```
library(xts)
NKX = xts(NKX, order.by = as.Date(Date))
r.NKX = (log(NKX) - lag(log(NKX), 1)) * 100
plot.xts(r.NKX, xlab = "Date", ylab = "NKX",
         main = "Percentage log returns", col = "darkgreen")
```

Similarly for *DAX* and *SPX*:

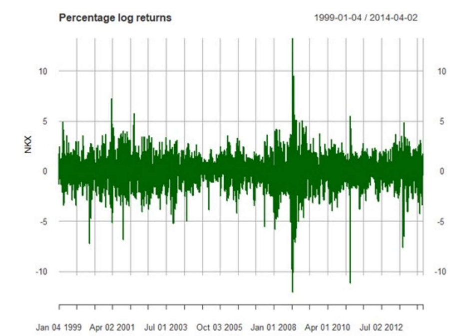
```
DAX = xts(DAX, order.by = as.Date(Date))
r.DAX = (log(DAX) - lag(log(DAX), 1)) * 100
plot.xts(r.DAX, xlab = "Date", ylab = "DAX",
         main = "Percentage log returns", col = "red")
```

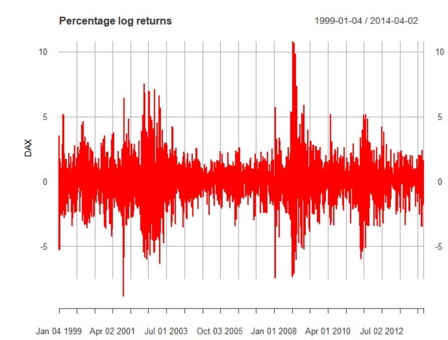
[4] Since the data is daily with gaps due to weekends and public holidays, we use *xts* objects.

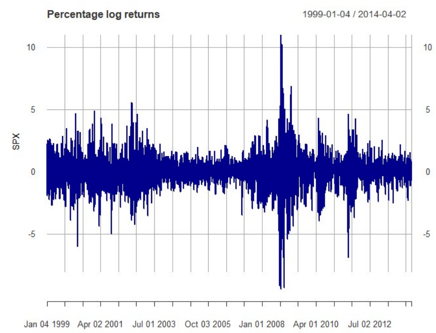
and

```
SPX = xts(SPX, order.by = as.Date(Date))
r.SPX = (log(SPX) - lag(log(SPX), 1)) * 100
plot.xts(r.SPX, xlab = "Date", ylab = "SPX",
         main = "Percentage log returns", col = "darkblue")
```

You should get these time series plots:







All three plots exhibit changing volatility and more or less coinciding periods of high and low volatilities.

Execute

library(pastecs)
options(scipen = 1000)
round(stat.desc(cbind(r.NKX, r.DAX, r.SPX), basic = FALSE, norm = TRUE), 4)

to obtain descriptive statistics for the three log return series in standard form with four decimals,

```
                r.NKX    r.DAX    r.SPX
median         0.0000   0.0482   0.0195
mean           0.0027   0.0152   0.0109
SE.mean        0.0240   0.0244   0.0202
CI.mean.0.95   0.0470   0.0479   0.0396
var            2.2875   2.3773   1.6215
std.dev        1.5125   1.5418   1.2734
coef.var     556.8179 101.2670 117.3403
skewness      -0.4184  -0.0194  -0.1713
skew.2SE      -5.3874  -0.2498  -2.2059
kurtosis       6.6227   4.4092   7.8871
kurt.2SE      42.6530  28.3975  50.7964
normtest.W     0.9421   0.9462   0.9161
normtest.p     0.0000   0.0000   0.0000
```

The reported skewness statistics are negative, implying that all three log return series are skewed to the left. However, for *r,DAX*, |*skew.2SE*| = 0.2498 < 1, so its skewness statistic is insignificant at the 5% level.

The reported kurtosis statistics are excess kurtosis statistics. They are all above 3 and each |*kurt.2SE*| > 1, implying that the three log return series have fat-tail distributions relative to the normal distribution, i.e., they are leptokurtic.

In the light of skewness and kurtosis, it is not surprising that the *SW* tests have zero *p*-values, so they reject normality at any reasonable significance level.

Note that the *options(scipen = 1000)* command forced *R* to report the results in standard form. This made the *stat.desc* results look nicer, but it might backfire later on if some values have a large number of decimals. For this reason, it is better to reset this option by executing

options(scipen = 0)

b) Compute the deviations of returns from their respective sample means,

$$e_{it} = r_{it} - \overline{r}_i \quad , \quad i = 1,2,3$$

These are the estimates of $\{\varepsilon_{it}\}$. Find the best fitting *ARMA* model for each $\{e_{it}\}$ series using *auto.arima*.

The deviations are obtained by executing the following commands[5]:

e.NKX = r.NKX - mean(r.NKX, na.rm = TRUE)
e.DAX = r.DAX - mean(r.DAX, na.rm = TRUE)
e.SPX = r.SPX - mean(r.SPX, na.rm = TRUE)

Then run

library(forecast)
auto.arima(e.NKX)

to get

```
Series: e.NKX
ARIMA(0,0,0) with zero mean

sigma^2 = 2.287:  log likelihood = -7288.04
AIC=14578.09   AICc=14578.09   BIC=14584.38
```

This printout suggests that the mean of *e.NKX* is probably best modelled as a white noise error.

[5] Note the *na.rm = TRUE* argument in the *mean()* functions. It instructs *R* to remove the missing values before calculating the mean. Without this argument the *mean()* function would return an error message.

Likewise,

auto.arima(e.DAX)
auto.arima(e.SPX)

return the following printouts.

```
Series: e.DAX
ARIMA(3,0,1) with zero mean

Coefficients:
          ar1      ar2      ar3     ma1
      -0.6014  -0.0215  -0.0461  0.5873
s.e.   0.1710   0.0187   0.0163  0.1707

sigma^2 = 2.373:  log likelihood = -7359.21
AIC=14728.43   AICc=14728.44   BIC=14759.87
```

```
Series: e.SPX
ARIMA(1,0,1) with zero mean

Coefficients:
         ar1      ma1
      0.5186  -0.6007
s.e.  0.1368   0.1284

sigma^2 = 1.607:  log likelihood = -6585.53
AIC=13177.05   AICc=13177.06   BIC=13195.92
```

The means of *e.DAX* and *e.SPX* seem to have been generated by zero mean *ARMA*(3,1) and *ARMA(1,1)* processes.

c) The *NKX*, *DAC* and *SPX* indices are from the Japanese, the European and the US market, respectively. In the light of the discussion about heatwaves and meteor showers and the suggested mean equation in part (b), estimate the following *GARCH*(1,1) model for NKX.:

$$e.NKX_t = \varepsilon_{1,t}, \quad \varepsilon_{1,t} \mid \Omega_t : idN(0, h_{1,t})$$

$$h_{1,t} = \alpha_{10} + \alpha_{11} e_{1,t-1}^2 + \alpha_{12} e_{2,t-1}^2 + \alpha_{13} e_{3,t-1}^2 + \beta_{11} h_{1,t-1}$$

To calculate the squared deviations, execute

e2.NKX = e.NKX^2
e2.DAX = e.DAX^2
e2.SPX = e.SPX^2

Call the *rugarch* library and specify the model,

```
library(rugarch)
n = length(NKX)
spec.NKX = ugarchspec(mean.model = list(armaOrder = c(0,0),
                                        include.mean = FALSE),
                      variance.model = list(model = "sGARCH",
                                            garchOrder = c(1,1),
                                            external.regressors = cbind(lag(e2.DAX)[3:n, 1],
                                                                        lag(e2.SPX)[3:n, 1])),
                      distribution.model = "norm")
```

Note the *include.mean = FALSE* argument in the mean equation and the extra regressors in the variance equation. *e2.DAX* and *e2.SPX* do not have values for the first day in the sample, and their differences do not have values for the second day either.[6] Hence, the data series that we can use in the model estimation start day 3, and *[3:n,1]* after *lag(e2.DAX)* and *lag(e2.SPX)* serves to keep only those available observations.

Fit the specified model to the data and print the results,

fit.NKX = ugarchfit(spec = spec.NKX, data = e.NKX[3:n, 1])
fit.NKX

The top part of the printout is

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

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

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
omega   0.072668    0.021482   3.3828 0.000718
alpha1  0.091223    0.010587   8.6168 0.000000
beta1   0.813929    0.029463  27.6253 0.000000
vxreg1  0.010082    0.007190   1.4023 0.160840
vxreg2  0.077399    0.017098   4.5268 0.000006
```

[6] You can easily verify this by the *print(cbind(lag(e2.NKX), lag(e2.DAX), lag(e2.SPX)))* command.

As usual, *omega* is the intercept, *alpha1* is the coefficient of the *ARCH* term, *beta1* is the coefficient of the *GARCH* term, and *vxreg1* and *vxreg2* denote the two extra regressors in the variance equation.[7]

d) Based on the suggested mean equation in part (b), estimate the following *GARCH*(1,1) model for the *DAX* index:[8]

$$e.DAX_t = \varphi_{21} e.DAX_{t-1} + \varphi_{22} e.DAX_{t-2} + \varphi_{23} e.DAX_{t-3} + \varepsilon_{2,t} + \theta_{21}\varepsilon_{2,t-1}, \quad \varepsilon_{2,t} \mid \Omega_t : idN(0, h_{2,t})$$

$$h_{2,t} = \alpha_{20} + \alpha_{21} e_{1,t}^2 + \alpha_{22} e_{2,t-1}^2 + \alpha_{23} e_{3,t-1}^2 + \beta_{21} h_{2,t-1}$$

You can do so by executing the following commands:

```
spec.DAX = ugarchspec(mean.model = list(armaOrder = c(3,1),
                                        include.mean = FALSE),
                      variance.model = list(model = "sGARCH",
                                            garchOrder = c(1,1),
                                            external.regressors = cbind(e2.NKX[3:n, 1],
                                                                        lag(e2.SPX)[3:n, 1])),
                      distribution.model = "norm")
fit.DAX = ugarchfit(spec = spec.DAX, data = e.DAX[3:n, 1])
fit.DAX
```

Again, only the top part of the printout is shown on the next page. As you can see, in the mean equation only *ar1* and *ma1* are significant[9], while in the variance equation every coefficient is strongly significant.

e) Based on the suggested mean equation in part (b), estimate the following *GARCH*(1,1) model for the *SPX* index:

$$e.SPX_t = \varphi_{31} e.SPX_{t-1} + \varepsilon_{3,t} + \theta_{31}\varepsilon_{3,t-1}, \quad \varepsilon_{3,t} \mid \Omega_t : idN(0, h_{3,t})$$

$$h_{3,t} = \alpha_{30} + \alpha_{31} e_{1,t}^2 + \alpha_{32} e_{2,t-1}^2 + \alpha_{33} e_{3,t-1}^2 + \beta_{31} h_{3,t-1}$$

[7] We do not discuss the rest of the printout this time, but check on your screen that

i. It makes a difference whether we rely on the usual or the robust standard errors. In the former case every coefficient is strongly significant, except that of *lag(e2.DAX)* errors, while in the latter case *omega* is also insignificant.
ii. The standardized residuals and the squared standardized residuals are not autocorrelated.
iii. There is no remaining *ARCH* effect.
iv. The Nyblom tests detect some structural change.
v. The sign tests detect some leverage effect.
vi. The error term in the mean equation is probably not normally distributed.

[8] This time each coefficient has two indices, the first refers to the equation and the second to the lag.
[9] One could drop the insignificant terms, re-estimate the more parsimonious model, and compare the performances of the two models on the basis of the information criteria. To save time, we do not do that.


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

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(3,0,1)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.738067    0.182682  4.04018 0.000053
ar2     0.017909    0.021324  0.83986 0.400989
ar3    -0.023993    0.017387 -1.37992 0.167610
ma1    -0.753794    0.182554 -4.12915 0.000036
omega   0.016616    0.004883  3.40279 0.000667
alpha1  0.069993    0.007815  8.95661 0.000000
beta1   0.896632    0.009928 90.31543 0.000000
vxreg1  0.007504    0.002948  2.54514 0.010923
vxreg2  0.027403    0.007325  3.74127 0.000183
```

Execute the following commands:

```
spec.SPX = ugarchspec(mean.model = list(armaOrder = c(1,1),
                                        include.mean = FALSE),
                      variance.model = list(model = "sGARCH",
                                            garchOrder = c(1,1),
                                            external.regressors = cbind(e2.NKX[3:n, 1],
                                                                        e2.DAX[3:n, 1])),
                      distribution.model = "norm")
fit.SPX = ugarchfit(spec = spec.SPX, data = e.SPX[3:n, 1])
fit.SPX
```

The top part of the printout is shown on the next page. In this case both coefficients are significant in the mean equation, but in the variance equation the two extra regressors re not.

f) Summarise the parameter estimates and their significance in a table and briefly discuss them in terms of heatwaves and meteor showers in volatility patterns between Japan, Europe, and the United States.

The table of point estimates and their significance (*: 10%, **: 5%, ***: 1%) are displayed in the table on the next page.

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

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

Optimal Parameters
------------------------------------
        Estimate  Std. Error    t value Pr(>|t|)
ar1     0.721796    0.097082   7.434914 0.000000
ma1    -0.769284    0.089283  -8.616207 0.000000
omega   0.013182    0.002793   4.719194 0.000002
alpha1  0.075023    0.007679   9.769238 0.000000
beta1   0.915485    0.008259 110.846382 0.000000
vxreg1  0.000000    0.001097   0.000009 0.999993
vxreg2  0.000000    0.001347   0.000008 0.999993
```

Variable	***NKX***	***DAX***	***SPX***
ar1		0.7381***	0.7218***
ar2		0.0179	
ar3		-0.0239	
ma1		-0.7528***	-0.7693***
omega	0.0727***	0.0166***	0.0132***
alpha1	0.0912***	0.0700***	0.0750***
beta1	0.8139***	0.8966***	0.9155***
e2.NKX		0.0075**	0.0000
lag(e2.NKX)			
e2.DAX			0.0000
lag(e2.DAX)	0.0101		
e2.SPX			
lag(e2.SPX)	0.0774***	0.0274***	

Focusing on the variance equations, two general conclusions emerge from the results.

First, *beta1* (β_{ii}), the coefficient of the lagged conditional variance, is significant in each model. This means that volatility in every region is partly determined by previous trading day's volatility in the same region, supporting the heatwave (alternative) hypothesis.

Second, there is also some evidence of the meteor shower effect since shocks to the Japanese market seem to influence the European market on the same trading day. However, there is a meteor shower neither from Japan to the US, nor from Europe to the US.

These conclusions indicate volatility of the global markets are combinations of heatwaves and meteor showers.[10]

Save your *R* code and quit *RStudio*.

[10] Caveat: In this exercise we used data from the Japanese, German and US markets. Note, however, that there is some overlap between trading on DAX and trading on the S&P 500, so they do not fully satisfy the partition discussed on page 1.