Financial Times Series

Lecture 12

Multivariate Volatility Models

- Here our aim is to generalize the previously presented univariate volatility models to their multivariate counterparts

- We assume that returns evolve according to

$$\boldsymbol{r}_t = \boldsymbol{\mu}_t + \boldsymbol{a}_t$$

where $\boldsymbol{\mu}_t = E(\boldsymbol{r}_t | F_{t-1})$ and $\boldsymbol{a}_t = (a_{1t}, \dots, a_{kt})'$ are the innovations

- Furthermore, we assume that $\boldsymbol{\mu}_t$ evolves according to

$$\boldsymbol{\mu}_t = \boldsymbol{\Upsilon} \boldsymbol{x}_t + \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{p}} \boldsymbol{\Phi}_i \boldsymbol{r}_{t-i} - \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{q}} \boldsymbol{\Theta}_i \boldsymbol{a}_{t-i}$$

where $\boldsymbol{x}_t$ is an m-dimensional vector of exogenous variables with $x_{1t} = 1$ and $\boldsymbol{\Upsilon}$ is a $k \times m$ matrix

Multivariate Volatility Models

- We also define the $k \times k$ matrix

$$\boldsymbol{\Sigma}_t = Cov(\boldsymbol{a}_t | F_{t-1})$$

- Our main interest is the evolution of $\boldsymbol{\Sigma}_t$

- It is not obvious how to generalize a univariate model, say a GARCH, to this context, since we now also have to consider (time-dependent) correlations/covariances between assets…

EWMA

- However, the generalization of an EWMA is straight forward;

- Given $F_{t-1} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{t-1}\}$ we have

$$\widehat{\boldsymbol{\Sigma}}_t = \frac{1-\lambda}{1-\lambda^{t-1}} \sum_{j=1}^{t-1} \lambda^{j-1} \boldsymbol{a}_j \boldsymbol{a}'_j$$

- For a t so large that $\lambda^{t-1} \approx 0$, we may write

$$\widehat{\boldsymbol{\Sigma}}_t = (1-\lambda)\boldsymbol{a}_{t-1}\boldsymbol{a}'_{t-1} + \lambda \widehat{\boldsymbol{\Sigma}}_{t-1}$$

Estimation

- For á priori estimates of λ and $\boldsymbol{\Sigma}_t$, $\widehat{\boldsymbol{\Sigma}}_t$ can be computed recursively using the above formula

- If we assume that $\boldsymbol{a}_t = \boldsymbol{r}_t - \boldsymbol{\mu}_t$ follows a multivariate normal distribution with mean zero and covariace matrix $\boldsymbol{\Sigma}_t$ we may use the log-likelihood function

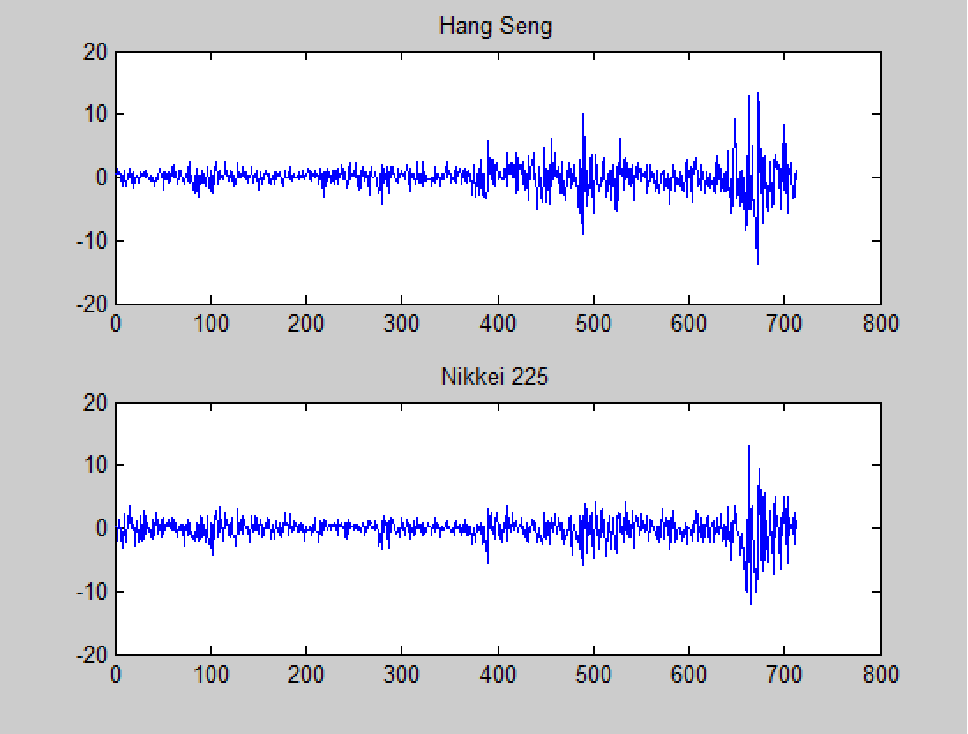
$$-\frac{1}{2}\sum_{t=1}^{T}|\boldsymbol{\Sigma}_t| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{a}'_t\boldsymbol{\Sigma}_t{}^{-\mathbf{1}}\,\boldsymbol{a}_t$$

which may be evaluated recursively, replacing $\boldsymbol{\Sigma}_t$ by $\widehat{\boldsymbol{\Sigma}}_t$

Example HS & N225

- We will use log returns in percentages for the Hang Tsen and N225 indices from 060104 to 081230 (See Tsay 3rd ed p.507)

Note the crisis in the fall of 2008!



Example HS & N225

- Univariate estimated GARCH(1,1) models are, where 1 denotes HS and 2 denotes N225;

$$r_{1t} = 0.104 + a_{1t}, a_{1t} = \sigma_{1t}\varepsilon_{1t}$$

$$\sigma^2{}_{1t} = 0.040 + 0.145r^2{}_{1t} + 0.854\sigma^2{}_{1t-1}$$

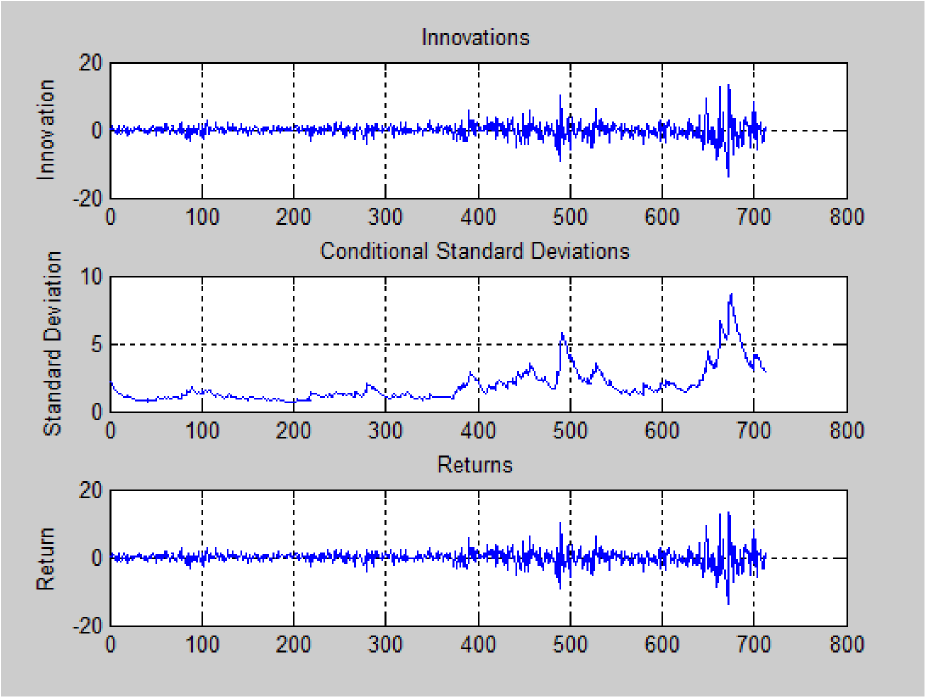
and

$$r_{1t} = 0.0017 + a_{1t}, a_{1t} = \sigma_{1t}\varepsilon_{1t}$$

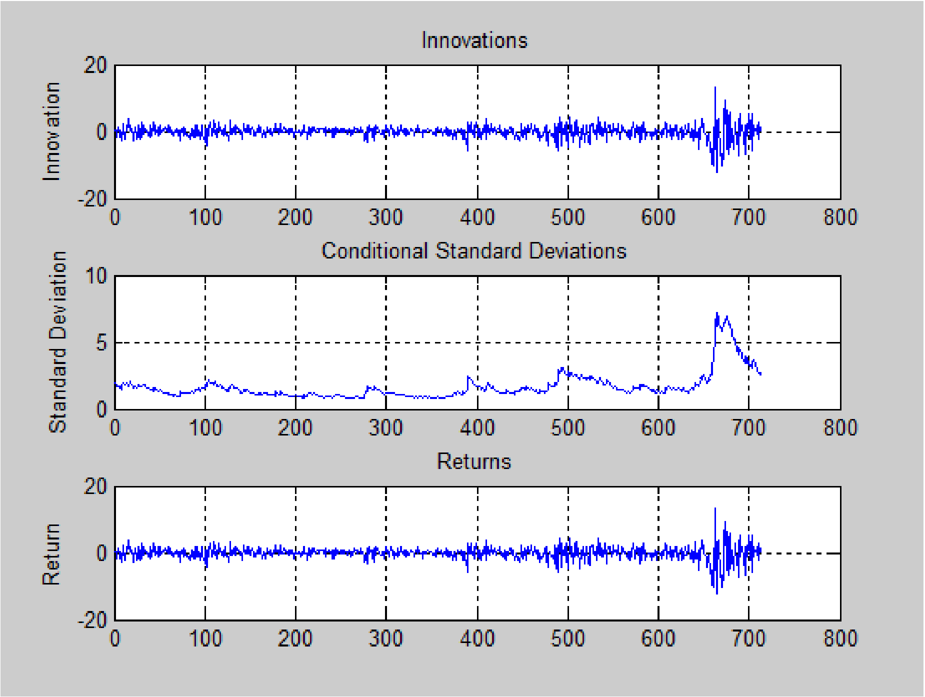
$$\sigma^2{}_{1t} = 0.045 + 0.125r^2{}_{1t} + 0.863\sigma^2{}_{1t-1}$$

Example HS & N225

HS



N225



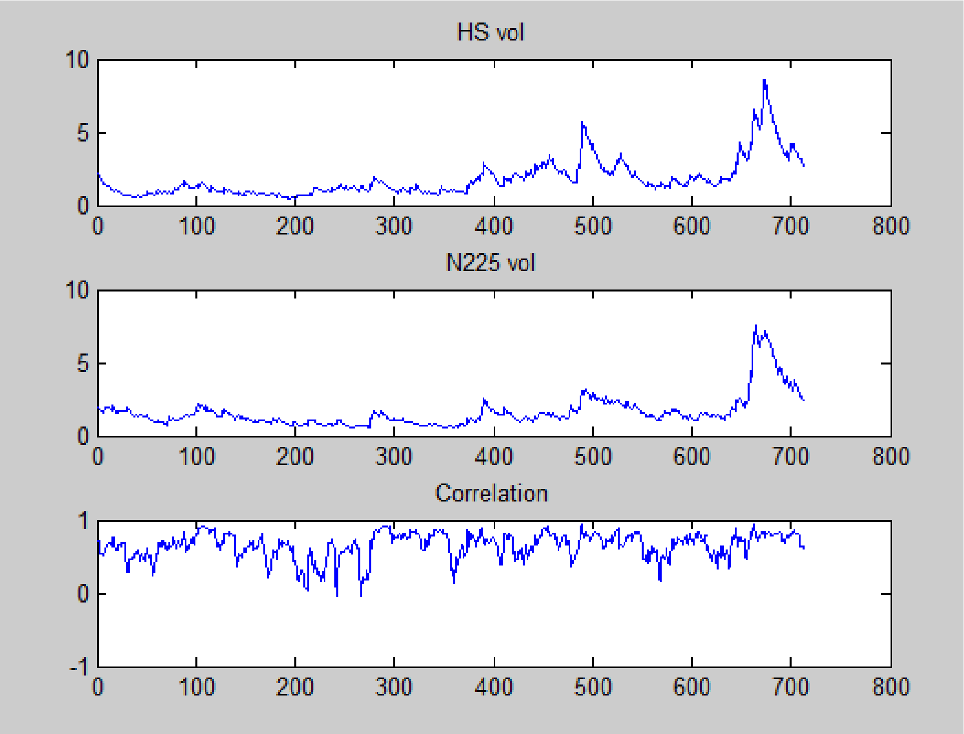
Example HS & N225

- We note that for both the models the sums of the estimated α and β are close to one, suggesting an IGARCH or EWMA model

- The large β estimates are probably due to the financial crisis

Example HS & N225

- Using $\lambda = 0.86$ as suggested by the univariate GARCH fits and the recursive formula and the unconditional covariance of the data as our á priori estimate, we get



Multivariate GARCH

- The are many different multivariate generalizations of the univariate GARCH model

- Bollerslev, Engle and Wooldridge (1988) propose the Diagonal Vectorization (DVEC) Model, defined by

$$\boldsymbol{\Sigma}_t = \boldsymbol{A}_0 + \sum_{i=1}^{m} \boldsymbol{A}_i \odot (\boldsymbol{a}_{t-i}\boldsymbol{a}'_{t-i}) + \sum_{j=1}^{s} \boldsymbol{B}_j \odot \boldsymbol{\Sigma}_{t-j}$$

where $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ are symmetric matrices and $\odot$ is element-wise multiplication (Hadamard product in French…)

DVEC(1,1)

- Since the matrices $\boldsymbol{A}_1$ and $\boldsymbol{B}_1$ are symmetric we consider only the lower triangular ”part” for which the elements are

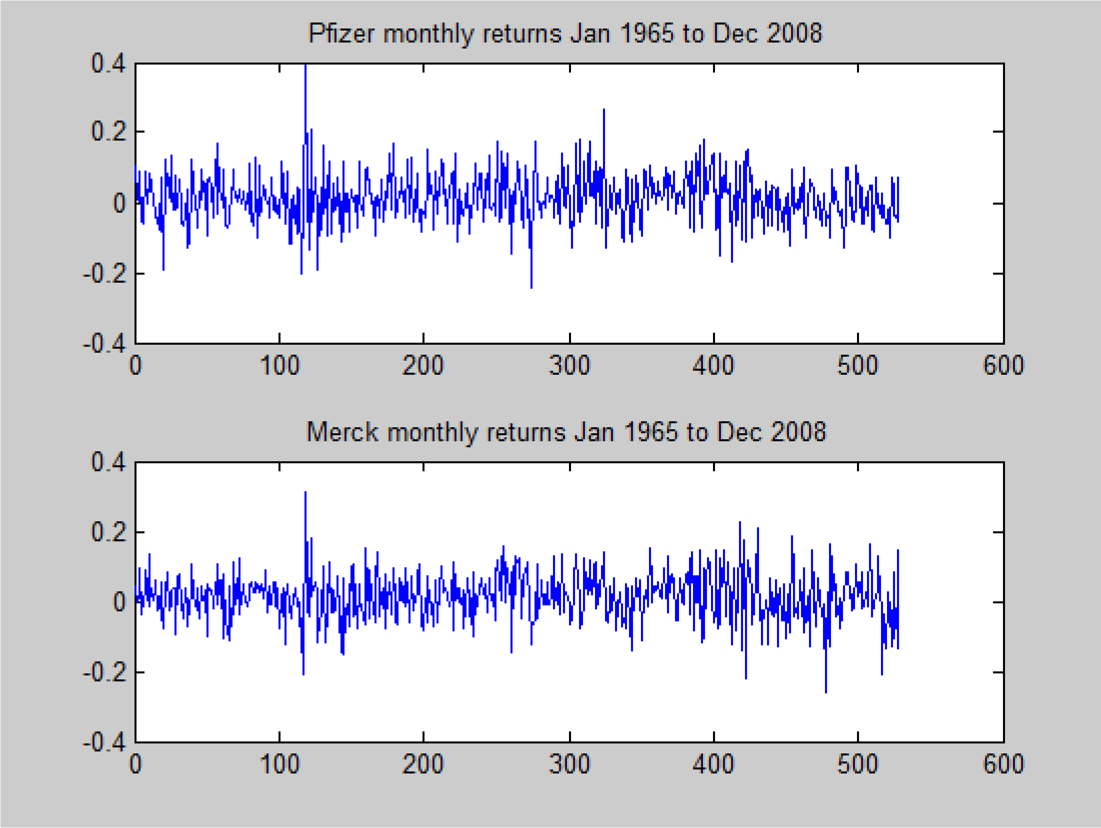
$$\sigma^2{}_{11,t} = A_{11,0} + A_{11,1}a^2{}_{1,t-1} + B_{11,1}\sigma^2{}_{11,t-1}$$

$$\sigma^2{}_{21,t} = A_{21,0} + A_{21,1}a_{1,t-1}a_{2,t-1} + B_{21,1}\sigma^2{}_{21,t-1}$$

$$\sigma^2{}_{22,t} = A_{22,0} + A_{22,1}a^2{}_{2,t-1} + B_{22,1}\sigma^2{}_{22,t-1}$$

Example

- Using the monthly returns of the Pfizer and Merck stocks, respectively



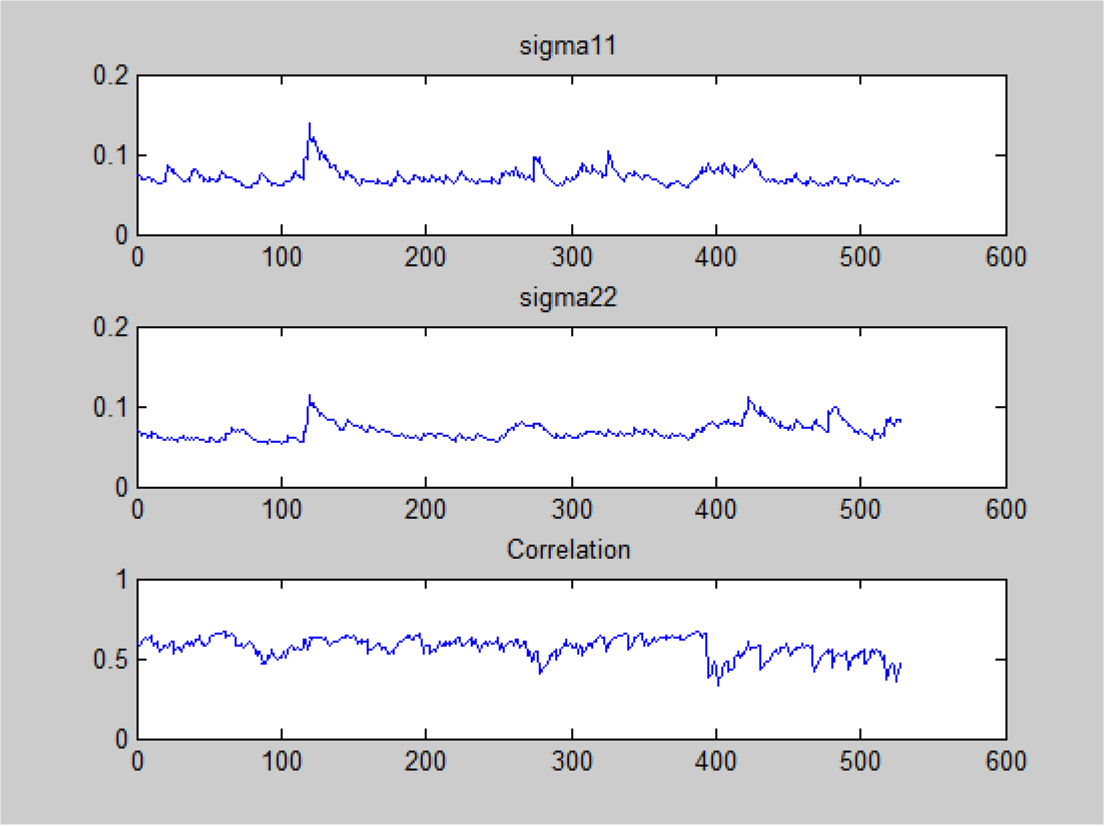
Example

- Using the ML routines of Ledoit et al (found at www.ledoit.net/ole4_abstract.htm) we get

Parameter	Estimate
$A_{11,0}$	$5.34 \cdot 10^{-4}$
$A_{12,0}$	$2.77 \cdot 10^{-4}$
$A_{22,0}$	$2.37 \cdot 10^{-4}$
$A_{11,1}$	0.0804
$A_{12,1}$	0.0493
$A_{22,1}$	0.0645
$B_{11,1}$	0.817
$B_{12,1}$	0.852
$B_{22,1}$	0.890

Example

- Here are the estimated volatilities and correlation



Cons of the DVEC

- It may produce a non-positive definite covariance matrix

- It does not allow for dynamic dependence between volatility series

BEKK

- This is a model will give a non-negative definite covariance matrix and allows for dynamic dependence between the volatility series;

$$\boldsymbol{\Sigma}_t = \boldsymbol{A}\boldsymbol{A}' + \sum_{i=1}^{m} \boldsymbol{A}_i(\boldsymbol{a}_{t-i}\boldsymbol{a}'_{t-i})\boldsymbol{A}'_i + \sum_{j=1}^{s} \boldsymbol{B}_j\boldsymbol{\Sigma}_{t-j}\,\boldsymbol{B}'_j$$

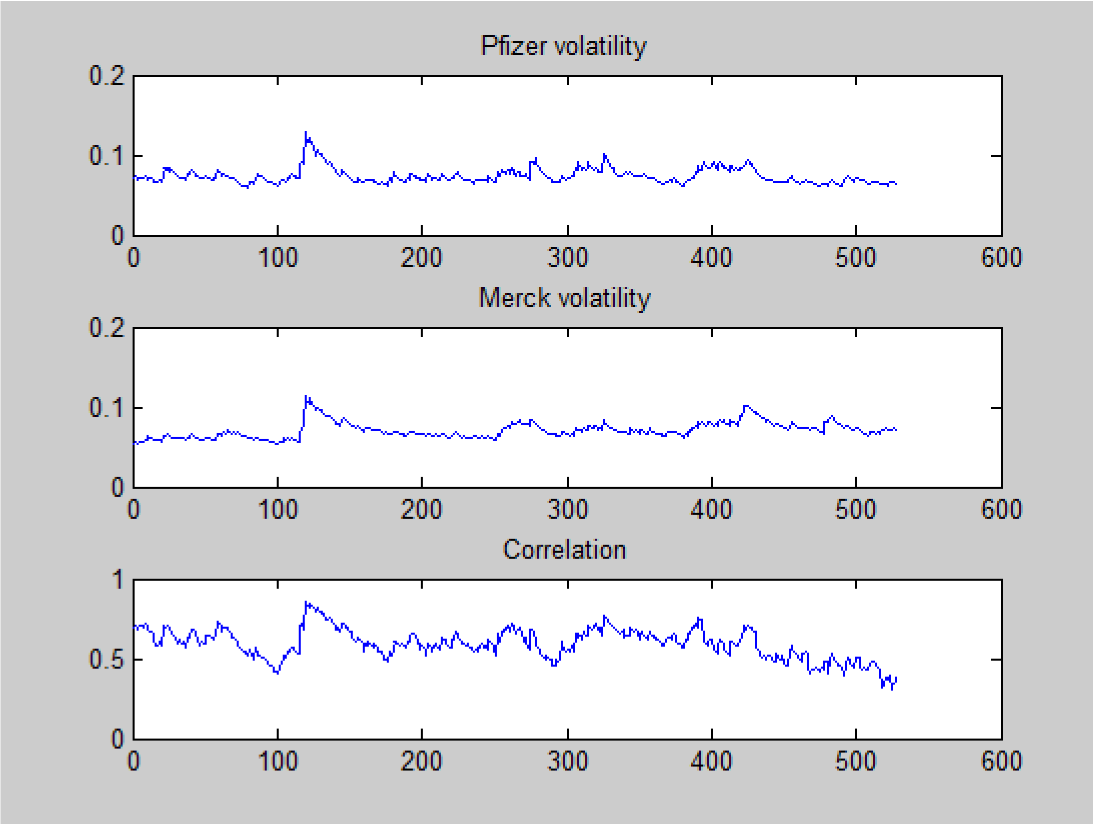
where $\boldsymbol{A}$ is a lower triangular matrix and $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ are $k \times k$ matrices

Estimation for Pfizer and Merck

- Using bekk.m of the MFE toolbox, available at www.kevinsheppard.com, we get the estimates

Parameter	Estimate	p-value
$A_{11,0}$	0.0189	0.0000
$A_{12,0}$	0.0088	0.0000
$A_{22,0}$	0.0088	0.0000
$A_{11,1}$	0.2557	0.0000
$A_{12,1}$	-0.0048	0.4491
$A_{21,1}$	0.0854	0.0002
$A_{22,1}$	0.1636	0.0000
$B_{11,1}$	0.9396	0.0000
$B_{12,1}$	-0.0093	0.0000
$B_{21,1}$	-0.0141	0.2960
$B_{22,1}$	0.9679	0.0000

BEKK for Pfizer and Merck



BEKK cons

- It should be noted that, due to the definition of the model the parameter values have no ”direct interpretation”

- We may also note that the number of parameters to be estimated grows like

$$k^2(m+s) + k(k+1)/2$$

CCC MV-GARCH

- Bollerslev (1990) proposed a multivariate GARCH with constant conditional correlation(s)

- The model may be written

$$\boldsymbol{\Sigma}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \boldsymbol{a}^2{}_{t-1} + \boldsymbol{\beta}_1 \boldsymbol{\Sigma}_{t-1}$$

- In the two-dimensional case we have

$$\begin{bmatrix} \sigma^2{}_{11,t} \\ \sigma^2{}_{22,t} \end{bmatrix} = \begin{bmatrix} \alpha_{10} \\ \alpha_{20} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} a^2{}_{11,t-1} \\ a^2{}_{22,t-1} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} \sigma^2{}_{11,t-1} \\ \sigma^2{}_{22,t-1} \end{bmatrix}$$

CCC MV-GARCH

- Letting $\boldsymbol{\eta}_t = \boldsymbol{a}^2{}_t - \boldsymbol{\Sigma}_t$ we may rewrite the model as

$$\boldsymbol{a}^2{}_t = \boldsymbol{\alpha}_0 + (\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1)\boldsymbol{a}^2{}_{t-1} - \boldsymbol{\beta}_1\boldsymbol{\eta}_t$$

- It follows (just as in the univariate case) that the model can be considered as an ARMA(1,1) for $\boldsymbol{a}^2{}_t$

Properties

- If all the eigenvalues of $\boldsymbol{\alpha}_1 + \boldsymbol{\beta}_1$ are between zero and one then the above described ARMA(1,1) is weakly stationary

- For the weakly stationary process we have that the unconditional variance of $\boldsymbol{a}_t$ is

$$(\boldsymbol{I} - \boldsymbol{\alpha}_1 - \boldsymbol{\beta}_1)^{-1}\boldsymbol{\alpha}_0$$

DCC models

- In real life correlations between time series tend to be non-constant over time
- To take this into account we may use Dymanic Conditional Correlation (DCC) models
- One such model was proposed by Engle (2002) in which

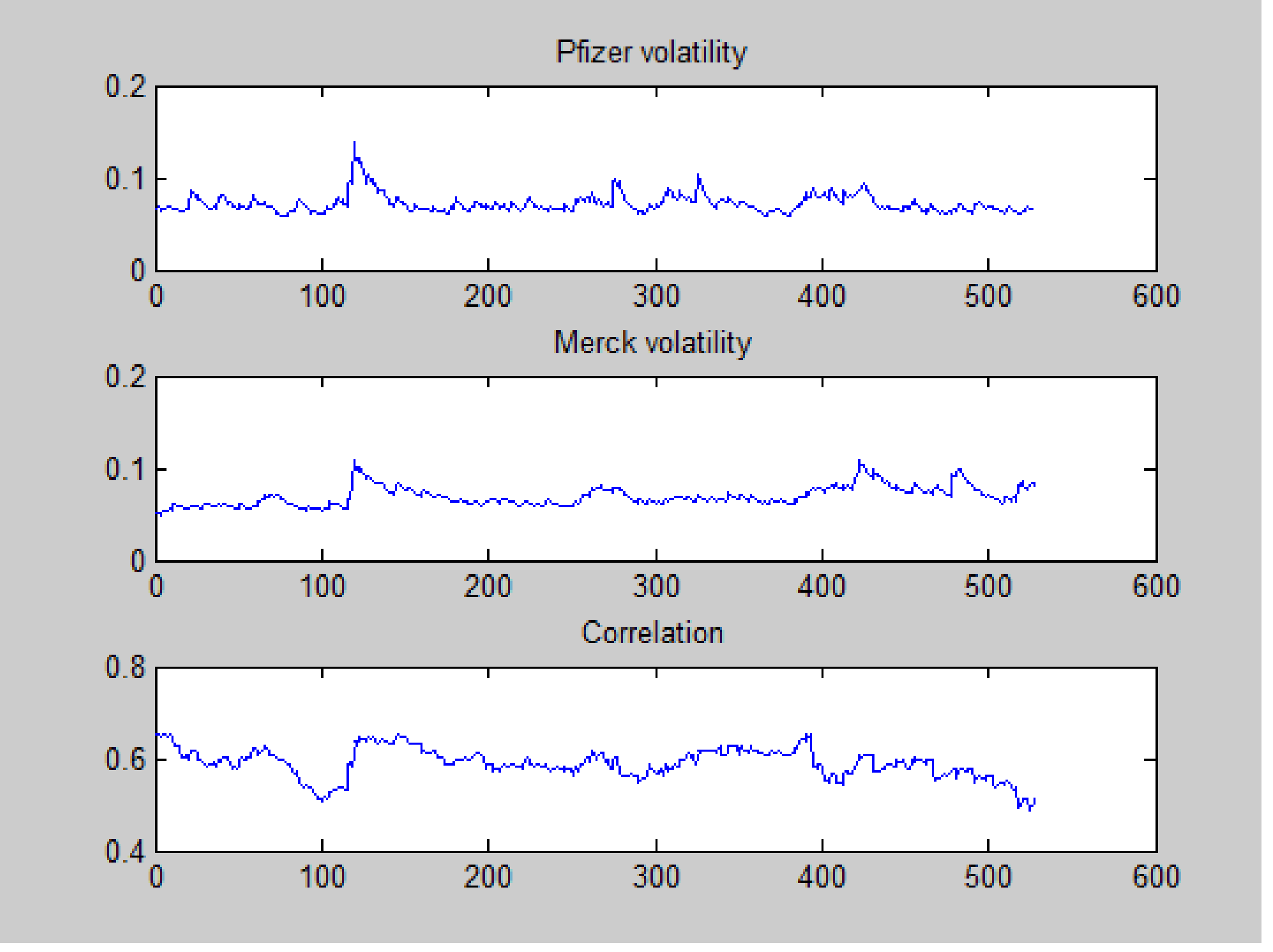
$$\boldsymbol{\rho}_t = \boldsymbol{J}_t{}^{-1/2}\boldsymbol{Q}_t\boldsymbol{J}_t{}^{-1/2}$$

where $\boldsymbol{J}_t = diag(\boldsymbol{Q}_t)$ and

$$\boldsymbol{Q}_t = (1-\theta_1-\theta_2)\overline{\boldsymbol{Q}} + \theta_1\boldsymbol{\varepsilon_t}\boldsymbol{\varepsilon}'_t + \theta_2\boldsymbol{Q}_{t-1}$$

for scalars $\theta_1 + \theta_2 < 1$, $\overline{\boldsymbol{Q}}$ the unconditional covariance of the standardized innovations $\boldsymbol{\varepsilon_t}$

Example (dcc.m from MFE)



Residuals check

- To check the different models' abilities to capture dependence and volatility of the assets we may create (observed) standardized residuals as

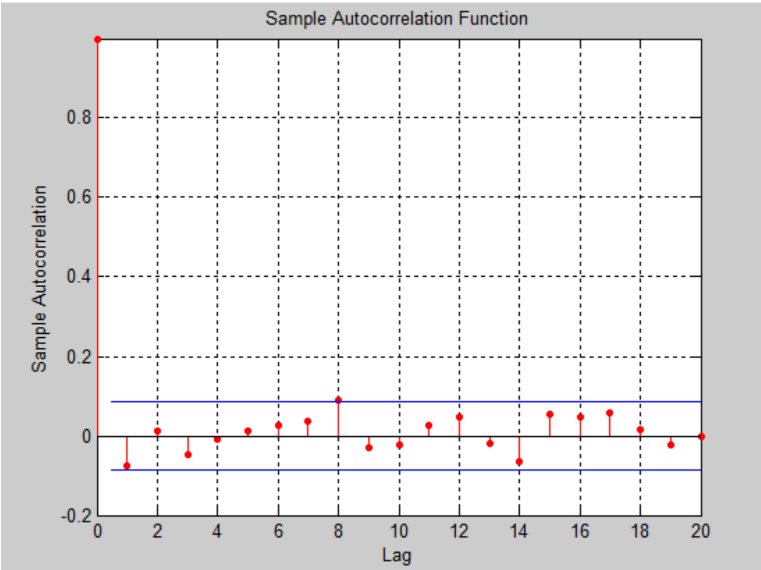
$$\hat{\boldsymbol{\varepsilon}}_t = \widehat{\boldsymbol{\Sigma}}_t^{-1/2} \boldsymbol{a}_t$$

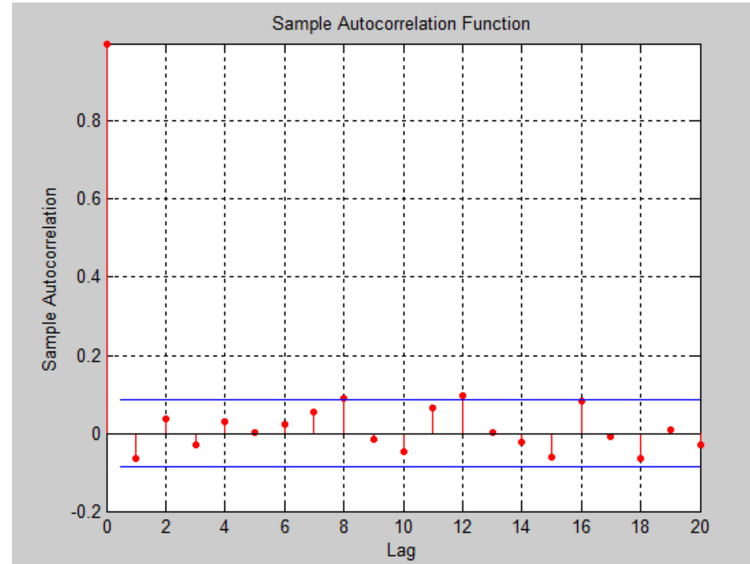
- We may apply ”autocorr” or Ljung-Box to the standard residuals and their ”squares” to check model adequacy

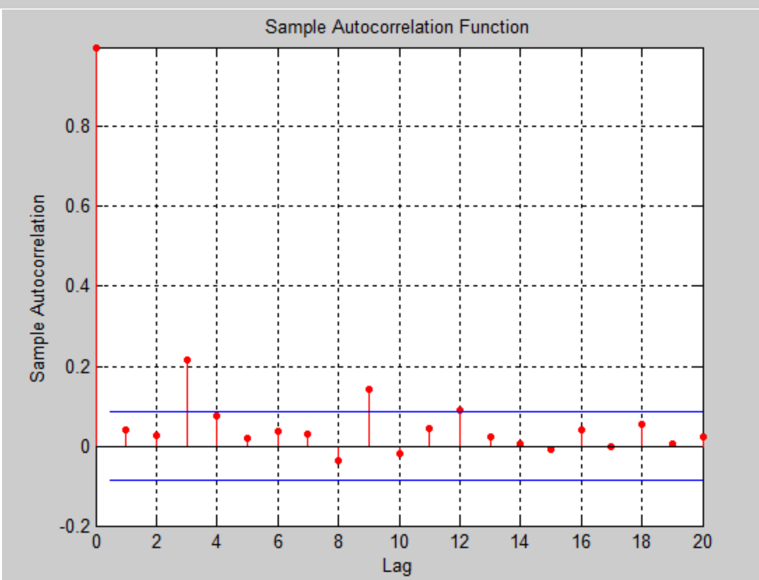
AC (squared on the bottom)

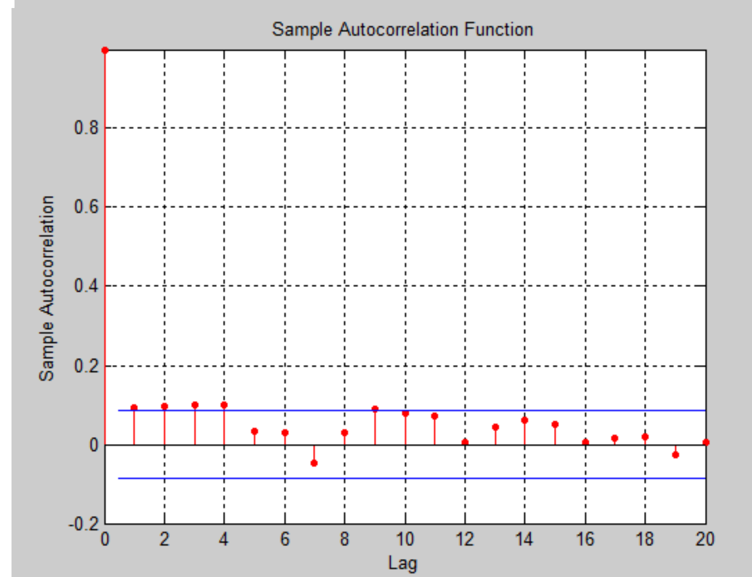
Pfizer data

Merck data



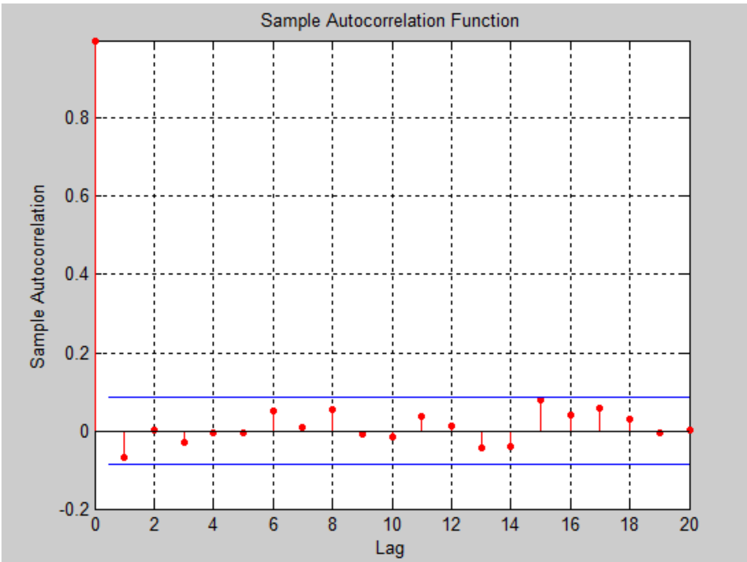


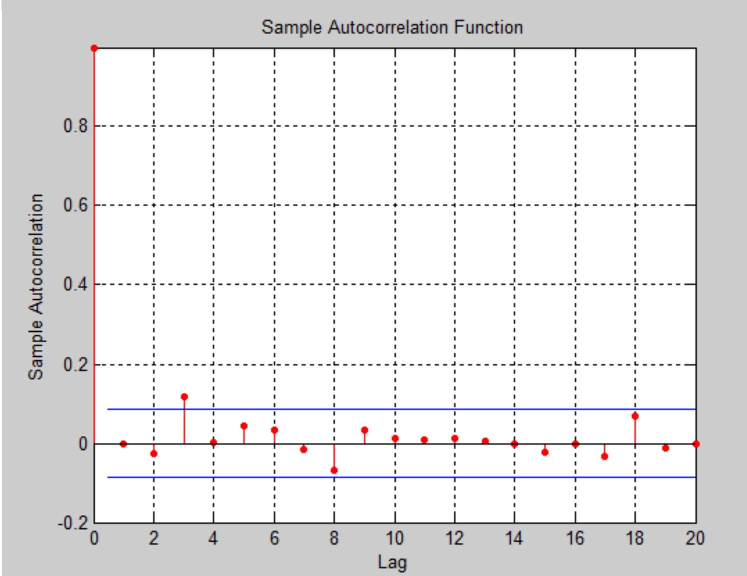




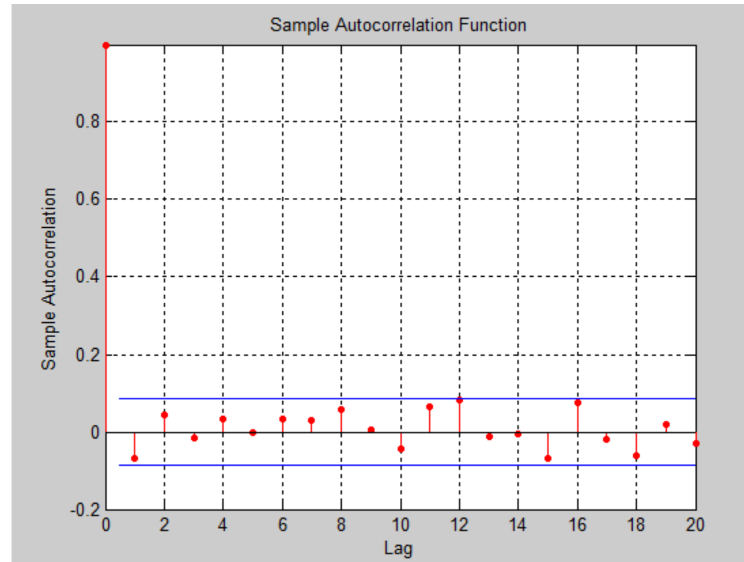
BEKK residuals

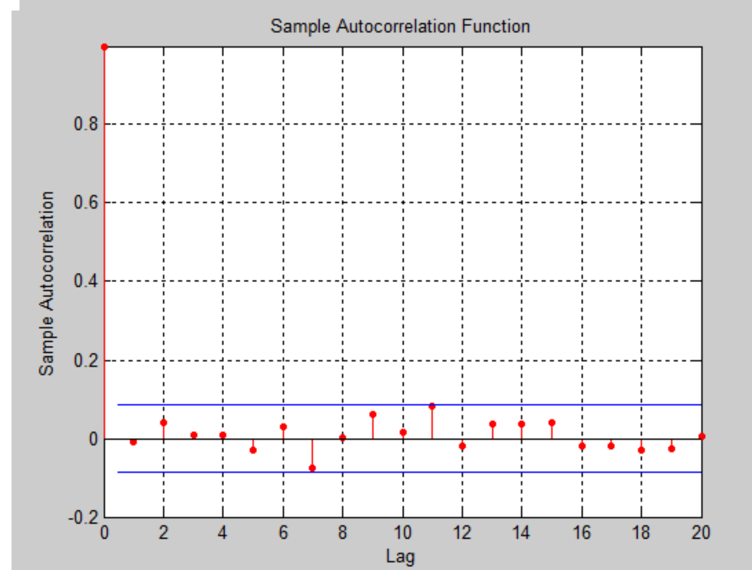
Pfizer





Merck

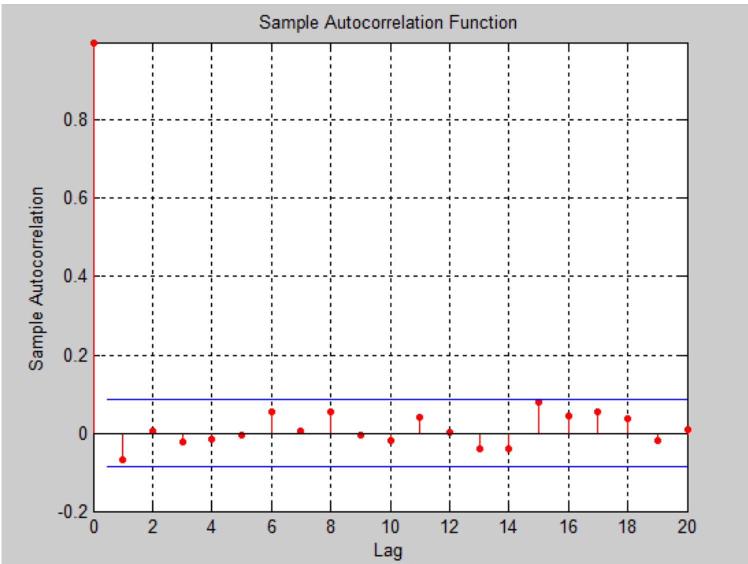


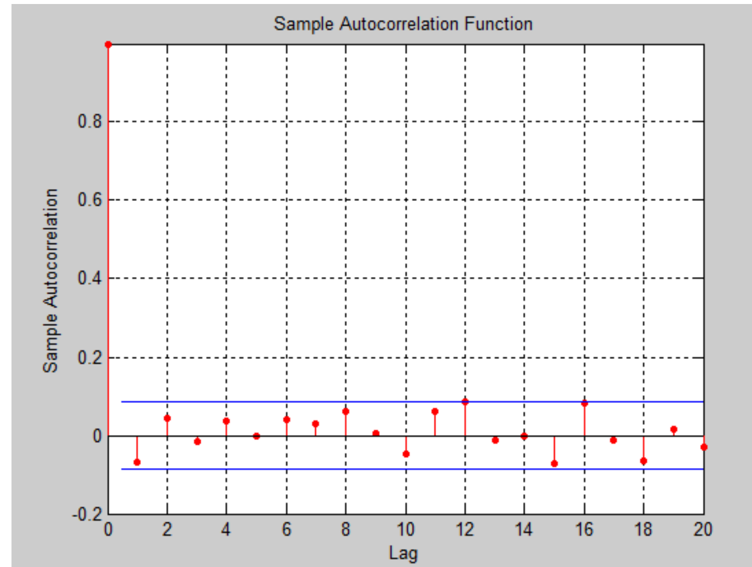


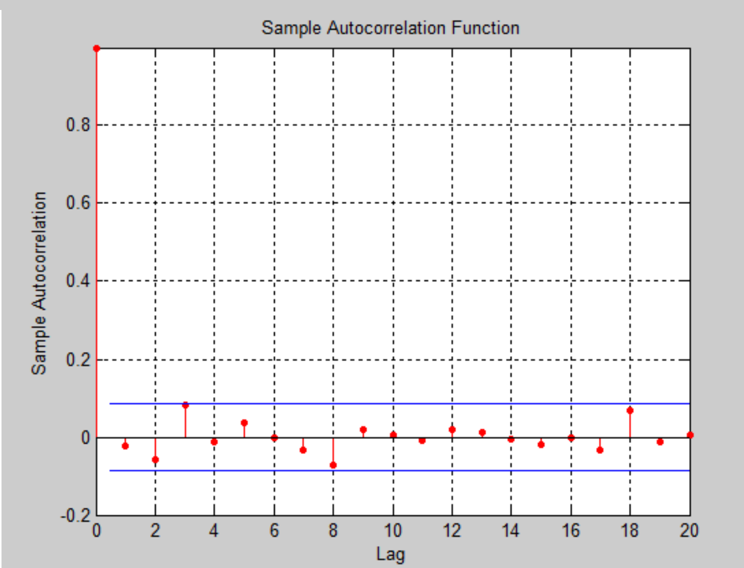
Residuals check for DCC

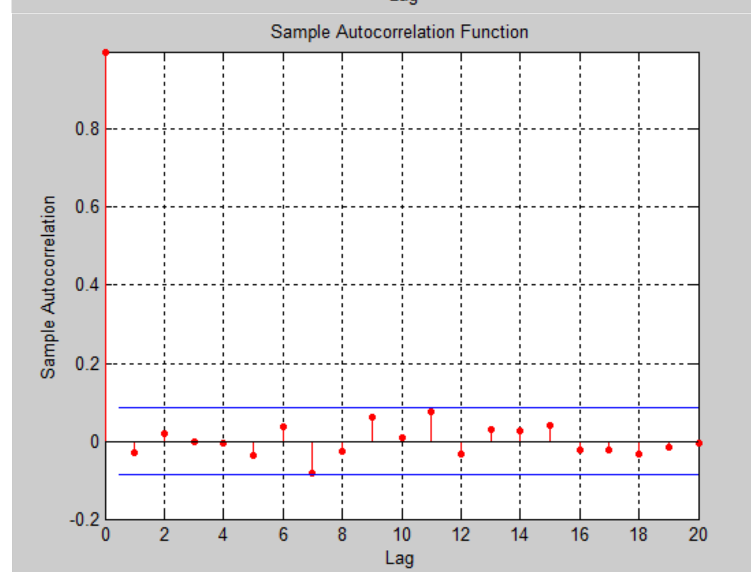
Pfizer

Merck





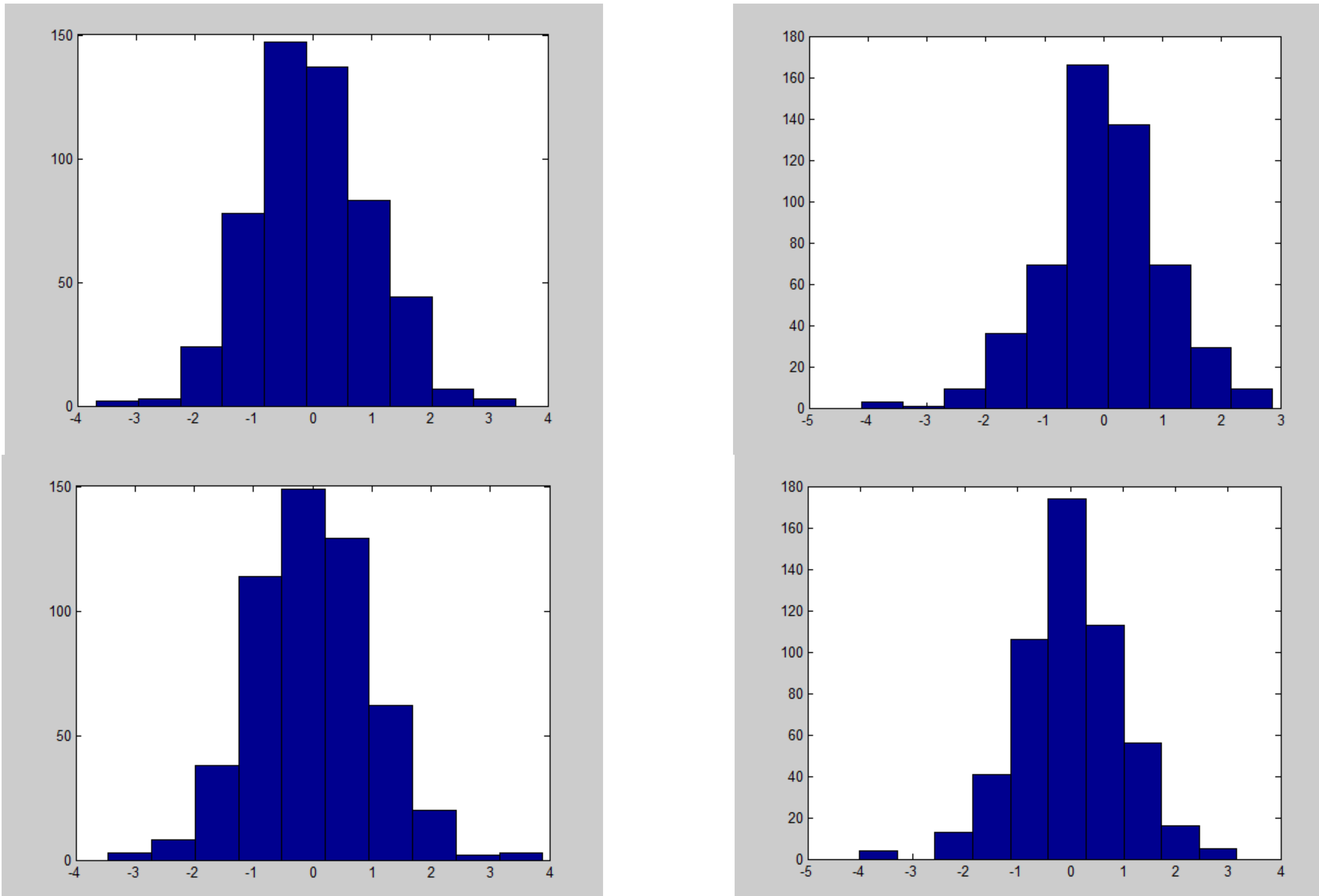




DCC vs. BEKK

- It is hard to tell from just the plots which is better
- The BIC for BEKK is 1.4692
- The BIC for DCC is 1.8595
- But these are computed under assumption of normality…

DCC vs. BEKK (DCC upper row, Pfizer left column)



DCC vs. BEKK

- Descriptives and JB-tests for residuals

Model/data	Skewness	Kurtosis	JB p-value
DCC/Pfizer	0.1094	3.4493	0.059
DCC/Merck	-0.2288	3.9802	0.001
BEKK/Pfizer	0.1899	3.6996	0.0045
BEKK/Merck	-0.2283	3.9883	0.001

- So, it is not ”safe” to assume normality…

Improvements

- Just as for the univariate GARCH the multivariate variants may be extended to include a ”skewness” part

- To be able to capture more kurtosis one may assume that the white noise has Student’s t-distribution

- There are many other multivariate volatility models, see e.g. www.kevinsheppard.com