Financial Data Analysis

Multivariate GARCH Models

Summer 2014

July 15, 2014
Multivariate Volatility Models

- Many problems in finance are inherently multivariate and require us to understand the dependence structure between assets.

- E.g.,
  - portfolio analysis,
  - volatility transmission: study of relations between the volatilities and covariances/correlations of several markets (e.g., emerging and developed markets, or different regions),
  - relation between returns, correlations and volatilities in different market regimes (e.g., bull vs. bear markets),
  - tests of asset pricing models,
  - futures hedging,
  - impact of financial volatility on other economic variables.
Multivariate GARCH Models

• Specific references:
Conditional Covariance Matrices

- Consider a return vector $r_t$ consisting of $N$ components (a column vector), i.e.,
  \[ r_t = [r_{1t}, r_{2t}, \ldots, r_{Nt}]'. \]  
  \hspace{1cm} (1)

- (Conditional) covariance matrix
  \[ H_t = \text{Cov}_{t-1}(r_t) = \begin{bmatrix} h_{1t}^2 & h_{12,t} & h_{13,t} & \ldots & h_{1N,t} \\ h_{12,t} & h_{2t}^2 & h_{23,t} & \ldots & h_{2N,t} \\ h_{13,t} & h_{23,t} & h_{3t}^2 & \ldots & h_{3N,t} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{1N,t} & h_{2N,t} & h_{3N,t} & \ldots & h_{Nt}^2 \end{bmatrix}, \hspace{1cm} (2) \]

  where
  \[ h_{it}^2 = \text{Var}_{t-1}(r_{it}), \quad h_{ij,t} = \text{Cov}_{t-1}(r_{it}, r_{jt}). \]  
  \hspace{1cm} (3)

- Covariance matrices are symmetric and positive (semi-)definite.
Conditional Covariance Matrices

- As the variance will not become negative, we know that for any linear combination of the elements of \( r_t, \sum_{i=1}^{N} w_ir_{it} \),

\[
0 \leq \text{Var}_{t-1} \left( \sum_{i=1}^{N} w_ir_{it} \right)
= \sum_{i=1}^{N} w_i^2 \text{Var}_{t-1}(r_{it}) + \sum_{i=1}^{N} \sum_{j \neq i} w_iw_j \text{Cov}_{t-1}(r_{it}, r_{jt})
= \sum_{i} w_i^2 h_{i,t}^2 + \sum_{i} \sum_{j \neq i} w_iw_j h_{ij,t}
= \sum_{i=1}^{N} \sum_{j=1}^{N} w_iw_j h_{ij,t} \quad \text{(writing } h_{ii,t} = h_{i,t}^2)\]
= \( w'H_iw \),

thus the conditional covariance matrix must be positive semidefinite for all \( t \).\(^1\)

\(^1\)This generalizes the nonnegativity constraint for \( \sigma_t^2 \) in the univariate GARCH model.
Conditional Covariance Matrices

• For example, with $N = 2$,

$$0 \leq \text{Var}_{t-1}(w_1r_{1t} + w_2r_{2t})$$

$$= w_1^2 \text{Var}_{t-1}(r_{1t}) + 2w_1w_2 \text{Cov}_{t-1}(r_{1t}, r_{2t}) + w_2^2 \text{Var}_{t-1}(r_{2t})$$

$$= w_1^2 h_{1t} + 2w_1w_2 h_{12,t} + w_2^2 h_{2t}$$

$$= \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} h_{1t}^2 & h_{12,t}^2 \\ h_{12,t}^2 & h_{2t}^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}. $$
Conditional Covariance Matrices

- The covariance matrix is positive definite (and thus nonsingular) if and only if no asset can be written as a linear combination of the other assets.

- If one asset can be written as a linear combination of the other assets, then there is a vector $\mathbf{a} \neq \mathbf{0}$ such that the linear combination

  $$\mathbf{a}' \mathbf{r}_t = \text{constant},$$

  so that the variance

  $$\text{Var}_{t-1}(\mathbf{a}' \mathbf{r}_t) = \mathbf{a}' \mathbf{H}_t \mathbf{a} = 0.$$
Conditional Covariance Matrices

• In many applications, positive definite (and thus nonsingular and invertible) covariance matrices are required (rather than just positive semidefinite matrices).

• E.g., in mean–variance (Markowitz) portfolio theory, many formulas for optimal portfolio weights require invertibility of the covariance matrix.

• Nonsingularity is also required for maximum likelihood estimation; e.g., if conditional (multivariate) normality is assumed, then the conditional density of $r_t$ is

$$
    f_{t-1}(r_t) = \frac{1}{(2\pi)^{N/2}\sqrt{\det(H_t)}} \exp\left\{ -\frac{1}{2} (r_t - \mu_t)'H_t^{-1}(r_t - \mu_t) \right\},
$$

where $\mu_t$ is the conditional mean of $r_t$. (4)
Reminder on discrete vs. continuous returns

• Recall that the linear relation between the portfolio return $r_{p,t}$ and the returns of the portfolio’s components,

$$ r_{p,t} = w^t r_t = \sum_{i=1}^{N} w_i r_{it}, $$

will only be exact if discrete returns are used.

• In the following, when continuous returns are used, we will always assume that

$$ r_{p,t} \approx w^t r_t = \sum_{i=1}^{N} w_i r_{it}, $$

which is a good approximation for daily or weekly returns.
Time–varying correlations

- The most simple and straightforward approach to detect possibly time-varying correlations is to calculate rolling averages, i.e., for demeaned returns,

\[
\hat{h}_{i,t}^2 = \frac{1}{m} \sum_{\tau=1}^{m} r_{i,t-\tau}^2, \quad i = 1, \ldots, N,
\]

\[
\hat{h}_{ij,t} = \frac{1}{m} \sum_{\tau=1}^{m} r_{i,t-\tau} r_{j,t-\tau}, \quad i, j = 1, \ldots, N,
\]

or

\[
\hat{H}_t = \frac{1}{m} \sum_{\tau=1}^{m} r_{t-\tau} r_{t-\tau}'. \quad \tag{7}
\]

- The rolling window correlations are

\[
\hat{\rho}_{ij,t} = \frac{\hat{h}_{ij,t}}{\sqrt{\hat{h}_{i,t}^2 \hat{h}_{j,t}^2}}, \quad i, j = 1, \ldots, N.
\]
Time-varying correlations

- Example 1: Daily returns of NASDAQ and Dow Jones Industrial Average (DJIA), January 1990 to September 2007, $T = 4474$ daily returns.

- $m = 250$, i.e., rolling average correlations based on approximately one year.

- Lower correlations during the boom and burst of the “New Economy” due to the NASDAQ being more vulnerable to the evolution of high-tech and internet stocks.
NASDAQ/DJIA rolling window correlations, window size $m = 250$ days
Time–varying correlations

• Example 2: Weekly returns of MSCI world stock market index and EPRA/NAREIT global real estate equity index, January 1990 to October 2011, $T = 1137$ weekly returns.

• $m = 50$, i.e., again a window size of approx. one year.
Time–varying correlations: EWMA

- As for univariate series, the rolling average (7) appears unsatisfactory due to the fact that it puts equal weight on all past returns.

- This may (again) be avoided by using an exponentially weighted moving average (EWMA), i.e.,

\[
H_t = (1 - \lambda) \sum_{\tau=1}^{\infty} \lambda^{\tau-1} r_{t-\tau} r'_{t-\tau}, \quad 0 < \lambda < 1,
\]

or, equivalently (cf. the univariate case),

\[
H_t = (1 - \lambda) r_{t-1} r'_{t-1} + \lambda H_{t-1}. \quad (8)
\]

---

This (with \( \lambda = 0.94 \)) corresponds to the 1994 RiskMetrics approach, which has been modified subsequently and may now roughly be described as a sophisticated EWMA covariance matrix, see Section 4.4.5 in K. Sheppard (2012): Forecasting High Dimensional Covariance Matrices, in Bauwens, Hafner, Laurent (eds): *Handbook of Volatility Models*, John Wiley & Sons.
Time–varying correlations: EWMA

- Note that the recursion

\[ H_t = (1 - \lambda) r_{t-1} r'_{t-1} + \lambda H_{t-1}. \]

produces a sequence of positive definite covariance matrices provided the initial covariance matrix \( H_1 \) is positive definite.

- (In the example below, the sample covariance matrix is used for the initial covariance matrix \( H_1 \).)
NASDAQ/DJIA EWMA correlations, $\lambda = 0.94$
time-varying correlations of NASDAQ/DJIA

rolling average, m = 250
EWMA, $\lambda = 0.94$
Multivariate GARCH Models

• Consider a return vector $\mathbf{r}_t$ consisting of $N$ components, i.e., $\mathbf{r}_t = [r_{1t}, r_{2t}, \ldots, r_{Nt}]'$ (a column vector),

$$
\mathbf{r}_t = \mu_t + \epsilon_t \quad (9)
$$

$$
\mu_t = \mathbb{E}(\mathbf{r}_t | I_{t-1}) = \mathbb{E}_{t-1}(\mathbf{r}_t) \quad (10)
$$

$$
\epsilon_t | I_{t-1} \sim \mathcal{N}(0, H_t) \quad (11)
$$

$$
H_t = \text{Var}(\mathbf{r}_t | I_{t-1}) = \text{Var}_{t-1}(\mathbf{r}_t) = \text{Var}_{t-1}(\epsilon_t), \quad (12)
$$

where $I_t$ is the information available at time $t$, usually $I_t = \{\mathbf{r}_t, \mathbf{r}_{t-1}, \ldots\}$.

• The error term

$$
\epsilon_t = [\epsilon_{1t}, \epsilon_{2t}, \ldots, \epsilon_{Nt}]'.
$$

• $H_t$ is the conditional covariance matrix of $\mathbf{r}_t$. 

Note I

• Similar to the univariate GARCH,

\[ r_t = \mu_t + \epsilon_t, \quad \epsilon_t = \sigma_t \eta_t, \quad \eta_t \sim \text{iid } \mathcal{N}(0, 1), \]

(11) is often written as

\[ \epsilon_t = H_t^{1/2} z_t, \quad z_t \sim \text{iid } \mathcal{N}(0, I), \]

(13)

• \( H_t^{1/2} \) is an \( N \times N \) matrix such that \( H_t^{1/2}(H_t^{1/2})' = H_t \) (matrix square root).

• As \( H_t \) is a covariance matrix, such a factorization exists, e.g., the Cholesky decomposition.
Note 1

- A symmetric positive definite matrix $A$ can be factored as $A = LL'$, where $L$ is lower triangular with positive diagonal elements (the Cholesky factorization of $A$).\(^3\)

- For example, if $N = 2$ (bivariate case), where

\[
H_t = \begin{bmatrix}
    h_{11}^2 & h_{12,t} \\
    h_{12,t} & h_{22}^2 \\
\end{bmatrix},
\]

the Cholesky factorization is

\[
L = \begin{bmatrix}
    \sqrt{h_{11}^2} & 0 \\
    h_{12,t}/\sqrt{h_{11}^2} & \sqrt{h_{22}^2 - h_{12}^2/h_{11}^2} \\
\end{bmatrix}.
\]

- $LL' = H_t$ is easily checked, and $h_{22,t}^2 - h_{12}^2/h_{11}^2 = (h_{11}^2h_{22,t}^2 - h_{12}^2)/h_{11}^2 = (\det H_t)/h_{11}^2 > 0$ since $H_t$ is positive definite.

\(^3\)Other factorizations exist.
Note 1

• It then follows from (13) that

\[ \text{Var}_{t-1}(r_t) = \text{Var}_{t-1}(\epsilon_t) \] (14)

\[ = E_{t-1}(\epsilon_t \epsilon'_t) - \underbrace{E_{t-1}(\epsilon_t)E_{t-1}(\epsilon'_t)}_{=0} \] (15)

\[ = E_{t-1}(H_{1/2}^t z_t z'_t (H_{1/2}^t))' \] (16)

\[ = H_{1/2}^t E_{t-1}(z_t z'_t) (H_{1/2}^t)' \] (17)

\[ = \text{identity matrix} \]

\[ = H_{1/2}^t \text{(}H_{1/2}^t\text{)}' = H_t, \] (18)

i.e., \( H_t \) is the conditional covariance matrix.
Note II: Alternative distributions

- It is often appropriate to replace the multivariate normal distribution in (11) with a more flexible alternative allowing for fat tails and perhaps skewness.

- E.g., we may take $z_t$ in (13) to be multivariate (unit-variance) Student’s $t$ with $\nu > 2$ degrees of freedom (common to all assets).

- In this case, the density of $z_t$ is

\[
f(z_t; \nu) = \frac{\Gamma \left( \frac{\nu+N}{2} \right)}{\Gamma(\nu/2)((\nu-2)\pi)^{N/2}} \left\{ 1 + \frac{z_t'z_t}{\nu-2} \right\}^{-(\nu+N)/2},
\]

and the conditional density of the return vector $r_t = \mu_t + H_t^{1/2}z_t$ is

\[
f_{t-1}(r_t; \nu) = \frac{\Gamma \left( \frac{\nu+N}{2} \right)}{\Gamma(\nu/2)((\nu-2)\pi)^{N/2} \sqrt{|H_t|}} \left\{ 1 + \frac{(r_t - \mu_t)'H_t^{-1}(r_t - \mu_t)}{\nu-2} \right\}^{-(\nu+N)/2}.
\]
Note II: Alternative distributions

- If the innovations display asymmetries, the Student’s $t$ distribution may not be appropriate.

- For more flexible multivariate distributions allowing for skewness and their use in multivariate GARCH models, see Section 3.1. in Bauwens et al. (2006) and the references therein.\(^4\)

---

Main Problems

• There are two main problems when it comes to the specification of multivariate GARCH models:

(i) To keep estimation feasible, we need parsimonious models (i.e., models with a moderate number of parameters) which are still flexible enough to capture the most important aspects of the volatility/covariance dynamics.

(ii) We have to make sure that the conditional covariance matrix will remain positive definite at each point of time.

• For the sake of illustration, consider a bivariate GARCH(1,1) of the general Vech–type (see below).

• The covariance matrix is given by

\[ H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{12,t} & h_{22,t} \end{bmatrix}. \]
In the most general case,
\[
\begin{align*}
    h_{1t}^2 &= c_1 + a_{11}\epsilon_{1,t-1}^2 + a_{12}\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{13}\epsilon_{2,t-1}^2 \\
        &+ b_{11}h_{1,t-1}^2 + b_{12}h_{12,t-1} + b_{13}h_{2,t-1}^2 \\
    h_{12,t} &= c_2 + a_{21}\epsilon_{1,t-1}^2 + a_{22}\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{23}\epsilon_{2,t-1}^2 \\
        &+ b_{21}h_{1,t-1}^2 + b_{22}h_{12,t-1} + b_{23}h_{2,t-1}^2 \\
    h_{2t}^2 &= c_3 + a_{31}\epsilon_{1,t-1}^2 + a_{32}\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{33}\epsilon_{2,t-1}^2 \\
        &+ b_{31}h_{1,t-1}^2 + b_{32}h_{12,t-1} + b_{33}h_{2,t-1}^2,
\end{align*}
\]
or
\[
\begin{bmatrix}
    h_{1,t}^2 \\
    h_{12,t} \\
    h_{2,t}^2
\end{bmatrix}
= h_t
= h_{1,t}^2 + a_{11}a_{12}a_{13} + a_{21}a_{22}a_{23} + a_{31}a_{32}a_{33} \epsilon_{1,t-1}^2 + a_{12}\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{23}\epsilon_{2,t-1}^2
+ b_{11}h_{1,t-1}^2 + b_{12}h_{12,t-1} + b_{13}h_{2,t-1}^2
+ b_{21}h_{1,t-1}^2 + b_{22}h_{12,t-1} + b_{23}h_{2,t-1}^2
+ b_{31}h_{1,t-1}^2 + b_{32}h_{12,t-1} + b_{33}h_{2,t-1}^2.
\]
Main Problems

- In this specification, both conditional variances, \( h_{1t}^2 \) and \( h_{2t}^2 \), and the conditional covariance, \( h_{12,t} \), may depend on all lagged squared returns and variances and all lagged cross–products \( \epsilon_{1,t-1}\epsilon_{2,t-1} \) and covariances.

- Although flexible, this model is difficult to handle in practice, since it requires estimation of 21 parameters (and this is only for the bivariate case).

- Moreover, without further restrictions, there is no guarantee that the sequence of covariance matrices implied by an estimated process will be positive definite for all \( t \).

- Such conditions are extremely tedious to work out and, in particular, to impose during estimation.
Vech Model

• The system above is a bivariate version of the Vech model, which is a straightforward generalization of univariate GARCH.

• The unrestricted case, though suffering from the problems mentioned above, is still useful, as it nests many more practicable specifications.

• The name derives from the fact that it uses the *vech operator*.

• As the $N \times N$ matrix $H_t$ is symmetric, it contains only $N(N + 1)/2$ independent elements, which may be obtained, for example, by excluding the upper triangular (redundant) part.
Vech Model

- The vech operator then stacks the remaining elements columnwise into an \( N(N + 1)/2 \) column vector, e.g., for \( N = 2 \),

\[
\text{vech}(H_t) = \text{vech}\left( \begin{bmatrix}
    h_{2t}^2 & h_{12,t} \\
    h_{12,t} & h_{2t}^2
\end{bmatrix}\right) = \begin{bmatrix}
    h_{12,t}^2 \\
    h_{12,t} \\
    h_{2t}^2
\end{bmatrix}
\]

\[
\text{vech}(\epsilon_t \epsilon_t') = \text{vech}\left( \begin{bmatrix}
    \epsilon_{1t} \\
    \epsilon_{2t}
\end{bmatrix} \begin{bmatrix}
    \epsilon_{1t} & \epsilon_{2t}
\end{bmatrix}\right)
\]

\[
= \text{vech}\left( \begin{bmatrix}
    \epsilon_{1t}^2 & \epsilon_{1t} \epsilon_{2t} \\
    \epsilon_{1t} \epsilon_{2t} & \epsilon_{2t}^2
\end{bmatrix}\right) = \begin{bmatrix}
    \epsilon_{1t}^2 \\
    \epsilon_{1t} \epsilon_{2t} \\
    \epsilon_{2t}^2
\end{bmatrix}.
\]

- (Similar to the vec operator, but the latter stacks the columns of a matrix one underneath the other \textit{without} excluding the upper triangular part.)
• The Vech(1,1) model can be written

\[ h_t = c + A\eta_{t-1} + Bh_{t-1}, \]  

(19)

where

\[ h_t = \text{vech} \ H_t \]  

(20)

\[ \eta_t = \text{vech}(\epsilon_t \epsilon_t'). \]  

(21)

• The covariance matrix has \( N(N + 1)/2 \) free parameters: \( N \) variances and \( N(N - 1)/2 \) covariances.

• Without restrictions, the are

  – \( N(N + 1)/2 \) parameters in \( c \)
  – \( N^2(N + 1)^2/4 \) parameters in \( A \)
  – \( N^2(N + 1)^2/4 \) parameters in \( B \).
  – With \( N = 2, 3, 5, 10 \) assets, we have 21, 78, 465, 6105 parameters.
Stationarity and Unconditional Covariance Matrix

- The covariance stationarity for the Vech(1,1) model (19),

\[ h_t = c + A\eta_{t-1} + Bh_{t-1}, \tag{22} \]

requires the eigenvalues of matrix

\[ Q = A + B \]

to be inside the unit circle.

- If this holds, the unconditional covariance matrix (its vech) can be obtained by taking expectations on both sides of (22),

\[
\begin{align*}
E(h_t) &= c + AE(\eta_{t-1}) + BE(h_{t-1}) \\
       &= c + A E(h_{t-1}) + B E(h_{t-1}) \\
       &= c + (A + B) E(h_t).
\end{align*}
\]
Stationarity and Unconditional Covariance Matrix

• Hence

\[ E(\text{vech } H_t) = E(h_t) = (I - A - B)^{-1} c. \]

• Covariance matrix forecasts, as required, e.g., for multi-period mean–variance portfolio optimization,

\[
\begin{align*}
    h_{t+1} &= c + A \eta_t + B h_t \\
    E_t(h_{t+2}) &= c + A E_t(\eta_{t+1}) + B h_{t+1} = c + (A + B) h_{t+1} \\
    E_t(h_{t+3}) &= c + A E_t(\eta_{t+2}) + B E_t(h_{t+2}) \\
    &= c + (A + B) E_t(h_{t+2}) \\
    &\vdots \\
    E_t(h_{t+\tau}) &= c + (A + B) E_t(h_{t+\tau-1}).
\end{align*}
\]

(23)

• Recursion (23) is convenient for calculation of multi-step covariance matrices, e.g., for mean–variance portfolio optimization.
What to do?

• As we have seen, the general Vech model
  – has too many parameters to estimate, at least for more than two or three series, and
  – does not admit conditions for positive definiteness which are straightforward to impose during estimation.

• Two possible approaches:
  1. Impose structure and simplification in the Vech model.
  2. Conditional correlation models, which separate the specification of the conditional variances from that of the conditional correlations.

• There is a trade-off between flexibility and parsimony, and the best solution depends on the particular application at hand.

• First consider a few restricted versions of the general Vech.
To reduce the number of parameters and make estimation feasible, this restricts the matrices $A$ and $B$ in (19) to be diagonal.

This means that

- each variance $h_{it}^2$ depends only on its own lagged squared error $\epsilon_{i,t-1}^2$ and its own lag (as in the univariate case)

$$h_{it}^2 = c_{ii} + a_{ii}\epsilon_{i,t-1}^2 + b_{ii}h_{i,t-1}^2, \quad i = 1, \ldots, N,$$  \hspace{1cm} (24)

- each covariance $h_{ij,t}$ depends only on the lagged cross–product of errors $i$ and $j$, $\epsilon_{i,t-1}\epsilon_{j,t-1}$, and its own lag,

$$h_{ij,t} = c_{ij} + a_{ij}\epsilon_{i,t-1}\epsilon_{j,t-1} + b_{ij}h_{ij,t-1}, \quad i, j = 1, \ldots, N.$$  \hspace{1cm} (25)
Diagonal Vech

- For many purposes, this specification may be sufficient to describe the evolution of variances and covariances.

- However, the cross-dynamics are seriously restricted, i.e., no direct interaction is allowed between the different conditional variances and covariances.

- Thus, the model will not be appropriate for studying volatility transmissions.
Even in the diagonal Vech model, necessary and sufficient conditions for positive definiteness are difficult to check and impose during estimation.

To find sufficient conditions, the diagonal Vech model (24) and (25) can be rewritten as

\[
H_t = \tilde{C} + \tilde{A} \odot (\epsilon_{t-1} \epsilon'_{t-1}) + \tilde{B} \odot H_{t-1},
\]

for suitably defined matrices \( \tilde{C}, \tilde{A}, \) and \( \tilde{B} \), and the Hadamard product \( \odot \) denotes elementwise multiplication of conformable matrices.

Note that the matrices \( \tilde{C}, \tilde{A}, \) and \( \tilde{B} \) in (26) are symmetric.

---

Diagonal Vech

- E.g., for \( N = 2 \),

\[
\begin{bmatrix}
    h_{1t}^2 & h_{12,t} \\
    h_{12,t} & h_{2t}^2
\end{bmatrix}
\]

\[
= \begin{bmatrix}
    c_{11} & c_{12} \\
    c_{12} & c_{22}
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
    a_{11} & a_{12} \\
    a_{12} & a_{22}
\end{bmatrix} \odot \begin{bmatrix}
    e_{1,t-1}^2 & e_{1,t-1} e_{2,t-1} \\
    e_{1,t-1} e_{2,t-1} & e_{2,t-1}^2
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
    b_{11} & b_{12} \\
    b_{12} & b_{22}
\end{bmatrix} \odot \begin{bmatrix}
    h_{1,t-1}^2 & h_{12,t-1} \\
    h_{12,t-1} & h_{2,t-1}^2
\end{bmatrix},
\]

that is,

\[
h_{1t}^2 = c_{11} + a_{11} e_{1,t-1}^2 + b_{11} h_{1,t-1}^2
\]

\[
h_{12,t} = c_{12} + a_{12} e_{1,t-1} e_{2,t-1} + b_{12} h_{12,t-1}
\]

\[
h_{2t}^2 = c_{22} + a_{22} e_{2,t-1}^2 + b_{22} h_{2,t-1}^2.
\]
Schur product theorem:

Consider two normally distributed zero–mean random vectors $X$ and $Y$ (of the same length), where

- $X$ has covariance matrix $\Sigma_X$, and
- $Y$ has covariance matrix $\Sigma_Y$,

and $X$ and $Y$ are independent.

Now consider the product $X \odot Y$.

The covariance between the $i$th and $j$th elements of $X \odot Y$ is

$$\text{Cov}(X_i \cdot Y_i, X_j \cdot Y_j) = \text{E}(X_i Y_i X_j Y_j) - \text{E}(X_i Y_i) \text{E}(X_j Y_j)$$

by independence

$$= \text{E}(X_i X_j) \text{E}(Y_i Y_j) - \text{E}(X_i) \text{E}(Y_i) \text{E}(X_j) \text{E}(Y_j)$$

by zero mean

$$= \text{E}(X_i X_j) \text{E}(Y_i Y_j)$$

by zero mean again

$$= \text{Cov}(X_i, X_j) \text{Cov}(Y_i, Y_j).$$
It follows that the covariance matrix of $X \odot Y$ is $\Sigma_X \odot \Sigma_Y$,

$$\text{Cov}(X \odot Y) = \Sigma_X \odot \Sigma_Y.$$ 

Thus $\Sigma_X \odot \Sigma_Y$ is a covariance matrix and therefore positive definite.

Since any positive definite matrix is the covariance matrix of a normal random vector we conclude that the Hadamard product of two positive definite matrices is likewise positive definite.

This result is often referred to as the Schur product theorem.\(^6\)

---

\(^6\)Which is slightly more general, allowing also for positive semi–definite matrices. For a “nonstatistical” derivation, see Horn and Johnson (1991): *Topics in Matrix Analysis*, Cambridge University Press.
Diagonal Vech

• The Schur product theorem can be applied to the representation

\[ H_t = \tilde{C} + \tilde{A} \circ (\epsilon_{t-1} \epsilon'_{t-1}) + \tilde{B} \circ H_{t-1} \]  

(27)

\[ H_t \]

to conclude that

– if (symmetric) matrices \( \tilde{C} \), \( \tilde{A} \), and \( \tilde{B} \) are positive definite, and
– if the initial covariance matrix \( H_0 \) is positive definite,

then \( H_t \) will remain positive definite for all \( t \).\(^7\)

\(^7\)Actually, as shown by Ledoit et al. (2003, cf. Footnote 9), it is sufficient that matrices

\[ \tilde{A}, \quad \tilde{B}, \quad \text{and} \quad \tilde{C} \circ (E - \tilde{B}) \]

are positive definite, where \( E \) is a matrix of ones, and \( \circ \) denotes elementwise division, analogous to \( \ominus \). This condition is weaker than the requirement of \( \tilde{A}, \tilde{B}, \) and \( \tilde{C} \) being positive definite. To see this is weaker, consider example

\[ B = \begin{pmatrix} 0.9 & 0.84 \\ 0.84 & 0.8 \end{pmatrix}, \quad C = \begin{pmatrix} 0.1 & 0.11 \\ 0.11 & 0.1 \end{pmatrix}. \]

Then \( C \circ (E - B) \) is positive definite but \( C \) is not.
Diagonal Vech

- The diagonal Vech model has $3 \times N(N + 1)/2$ free parameters.

- To make it applicable to high-dimensional systems, there are two ways to proceed:\footnote{Cf. K. Sheppard (2012): Forecasting High Dimensional Covariance Matrices, in Bauwens, Hafner, Laurent (eds): Handbook of Volatility Models, John Wiley & Sons.}

  (i) Introduce further simplifications, such as (greatest simplification) the scalar Vech:

  $$H_t = C + \alpha \epsilon_{t-1} \epsilon'_{t-1} + \beta H_{t-1}, \quad (28)$$

  where $\alpha$ and $\beta$ are positive scalars, and of which the EWMA is a special case with $C = 0$, $\alpha = 1 - \lambda$, $\beta = \lambda$.

BEKK

- BEKK, as suggested by Engle and Kroner (1995),\textsuperscript{10} is another restricted version of the general Vech model.

- This specifies, in its simplest (and also most frequently used) form,

\[
H_t = C^*C^{'*} + A^*\epsilon_{t-1}\epsilon_{t-1}'A^{'*} + B^*H_{t-1}B^{'*},
\]  \hspace{1cm} (29)

where $C^*$ is a triangular matrix with positive diagonal and $A^*$ and $B^*$ are $N \times N$ parameter matrices.

- This guarantees positive definiteness of $H_t$ by construction.\textsuperscript{11}

\textsuperscript{10}Multivariate Simultaneous Generalized ARCH, \textit{Econometric Theory}, 11, 122–150.

\textsuperscript{11}The decomposition of the constant term into the product of two triangular matrices also serves this purpose.
BEKK

• To figure out the restrictions imposed by (29), consider the case $N = 2$, where

$$
\begin{bmatrix}
    h_{1t}^2 & h_{12,t} \\
    h_{12,t} & h_{2t}^2
\end{bmatrix} =
\begin{bmatrix}
    c_{11}^* & 0 \\
    c_{21}^* & c_{22}^*
\end{bmatrix}
\begin{bmatrix}
    c_{11}^* & c_{21}^* \\
    0 & c_{22}^*
\end{bmatrix}
\left(\begin{array}{cc}
    \epsilon_{1,t-1}^2 & \epsilon_{1,t-1}\epsilon_{2,t-1} \\
    \epsilon_{1,t-1}\epsilon_{2,t-1} & \epsilon_{2,t-1}^2
\end{array}\right)
\begin{bmatrix}
    a_{11}^* & a_{12}^* \\
    a_{21}^* & a_{22}^*
\end{bmatrix}'

+ \begin{bmatrix}
    a_{11}^* & a_{12}^* \\
    a_{21}^* & a_{22}^*
\end{bmatrix}
\begin{bmatrix}
    \epsilon_{1,t-1}^2 & \epsilon_{1,t-1}\epsilon_{2,t-1} \\
    \epsilon_{1,t-1}\epsilon_{2,t-1} & \epsilon_{2,t-1}^2
\end{bmatrix}
\begin{bmatrix}
    a_{11}^* & a_{12}^* \\
    a_{21}^* & a_{22}^*
\end{bmatrix}'

+ \begin{bmatrix}
    b_{11}^* & b_{12}^* \\
    b_{21}^* & b_{22}^*
\end{bmatrix}
\begin{bmatrix}
    h_{1,t-1}^2 & h_{12,t-1} \\
    h_{12,t-1} & h_{2,t-1}^2
\end{bmatrix}
\begin{bmatrix}
    b_{11}^* & b_{12}^* \\
    b_{21}^* & b_{22}^*
\end{bmatrix}'

or

$$
\begin{align*}
h_{1,t}^2 &= c_1 + a_{11}^*\epsilon_{1,t-1}^2 + 2a_{11}^*a_{12}^*\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{12}^*\epsilon_{2,t-1}^2 \\
    &\quad + b_{11}^*h_{1,t-1}^2 + 2b_{11}^*b_{12}^*h_{12,t-1} + b_{12}^*h_{2,t-1}^2 \\
\end{align*}

\begin{align*}
h_{12,t} &= c_2 + a_{11}^*a_{21}^*\epsilon_{1,t-1}^2 + (a_{11}^*a_{22}^* + a_{21}^*a_{12}^*)\epsilon_{1,t-1}\epsilon_{2,t-1} + a_{22}^*a_{12}^*\epsilon_{2,t-1}^2 \\
    &\quad + b_{11}^*b_{21}^*h_{1,t-1}^2 + (b_{11}^*b_{22}^* + b_{12}^*b_{21}^*)h_{12,t-1} + b_{22}^*b_{12}^*h_{2,t-1}^2.
\end{align*}
BEKK

- E.g., the coefficient of $\epsilon_{1,t-1}\epsilon_{2,t-1}$ in the equation for $h_{1t}^2$ is given as a function of the coefficients of the lagged squared shocks $\epsilon_{1,t-1}^2$ and $\epsilon_{2,t-1}^2$.

- A drawback in some applications to testing economic theories may be that the coefficient of $h_{2,t-1}^2$ in the equation for $h_{1,t}^2$ is $b_{12}^* \epsilon_{12}$, i.e., negative volatility spillovers are excluded (similarly for the equation for $h_{2,t}^2$).\(^{12}\)

For the general Vech form of the BEKK model, recall the definition of the Kronecker product $\otimes$.

For an $m \times n$ matrix $A$ and an $p \times q$ matrix $B$, this is defined as the $mp \times nq$ matrix

$$A \otimes B = \begin{bmatrix}
a_{11}B & a_{12}B & \cdots & a_{1n}B \\
a_{21}B & a_{22}B & \cdots & a_{2n}B \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1}B & a_{m2}B & \cdots & a_{mn}B
\end{bmatrix}.$$ 

Important rule in time series analysis (see Appendix): For conformable matrices,

$$\text{vec}(ABC) = (C' \otimes A)\text{vec}(B).$$

Then (29) can be written as (without excluding the upper triangular part of $H_t$)

$$\text{vec}(H_t) = C + (A^* \otimes A^*)\text{vec}(\epsilon_{t-1}\epsilon'_{t-1}) + (B^* \otimes B^*)\text{vec}(H_{t-1}). \quad (30)$$
BEKK

• Representation (30) directly leads to stationarity conditions and covariance matrix forecasts for the BEKK model.

• E.g., covariance stationarity requires the eigenvalues of

\[ A^* \otimes A^* + B^* \otimes B^* \]

to be smaller than one in magnitude, and the unconditional covariance matrix can be recovered from

\[ E(\text{vec}(\epsilon_t \epsilon_t')) = (I - A^* \otimes A^* - B^* \otimes B^*)^{-1} \text{vec}(C^* C'^*) \].
• An implication of (30) is that each BEKK model corresponds to a unique vec model representation.

• The converse of the preceding statement is not true, however, as can be seen from the simple example

\[ A^* \otimes A^* = (-A^*) \otimes (-A^*); \quad (31) \]

thus restrictions for identification are required.

• In the practically most relevant BEKK(1,1) model of the form (29), it turns out that the model is identified if the diagonal elements of \( C^* \) as well as the top left elements of \( A^* \) and \( B^* \) are restricted to be positive.\(^{13}\)

\(^{13}\)See Proposition 2.1 of Engle and Kroner (1995).
Diagonal BEKK

• In practice, the diagonal BEKK model is sometimes used, where the parameter matrices $A^*$ and $B^*$ are diagonal.

• The latter is a restricted diagonal Vech model, where the restriction is that the coefficients in the covariance equations are directly related to those in the corresponding variance equations.
Diagonal BEKK

- E.g., for $N = 2$,

$$
\begin{bmatrix}
  h_{1t}^2 & h_{12,t} \\
  h_{12,t} & h_{2,t}^2
\end{bmatrix}
= \begin{bmatrix}
  c_{11}^* & 0 \\
  c_{21}^* & c_{22}^*
\end{bmatrix}
\begin{bmatrix}
  c_{11}^* & c_{21}^* \\
  0 & c_{22}^*
\end{bmatrix}
+ \begin{bmatrix}
  a_{11}^* & 0 \\
  0 & a_{22}^*
\end{bmatrix}
\begin{bmatrix}
  \epsilon_1^{2,t-1} & \epsilon_1^{1,t-1}\epsilon_2^{2,t-1} \\
  \epsilon_1^{1,t-1}\epsilon_2^{1,t-1} & \epsilon_2^{2,t-1}
\end{bmatrix}
\begin{bmatrix}
  a_{11}^* & 0 \\
  0 & a_{22}^*
\end{bmatrix}
+ \begin{bmatrix}
  b_{11}^* & 0 \\
  0 & b_{22}^*
\end{bmatrix}
\begin{bmatrix}
  h_{1,t-1}^2 & h_{12,t-1} \\
  h_{12,t-1} & h_{2,t-1}^2
\end{bmatrix}
\begin{bmatrix}
  b_{11}^* & 0 \\
  0 & b_{22}^*
\end{bmatrix},
$$

or

$$
\begin{align*}
  h_{1,t}^2 &= c_1 + a_{11}^{*2}\epsilon_1^{2,t-1} + b_{11}^{*2}h_{1,t-1}^2 \\
  h_{12,t} &= c_2 + a_{11}^*a_{22}^*\epsilon_1^{1,t-1}\epsilon_2^{2,t-1} + b_{11}^*b_{22}^*h_{12,t-1} \\
  h_{2,t}^2 &= c_3 + a_{22}^{*2}\epsilon_2^{2,t-1} + b_{22}^{*2}h_{2,t-1}^2
\end{align*}
$$
Diagonal Vech and BEKK and Extensions

• As argued by Bauwens et al. (2006), “diagonal vec and BEKK models are much more parsimonious but very restrictive for the cross–dynamics. They are not suitable if volatility transmission is the object of interest, but they usually do a good job in representing the dynamics of variances and covariances.”

• Asymmetries in volatility (as in the univariate case) have also been incorporated in multivariate GARCH models, see Bauwens et al. (2006) for an overview of alternative specifications and further references.
Factor Models

• This is another restricted subclass of Vech models, motivated by financial theory; see the references on the second slide for details and references.
Modeling Conditional Correlations

- Models considered so far have specified the dynamics of conditional variances and covariances, from which the conditional correlations can be recovered.

- Another approach is to model conditional variances and correlations (rather than the covariances).

- One advantage is that conditional variances (or standard deviations) and conditional correlations can be modeled separately, which often allows for consistent two-step model estimation, thus reducing the computational burden.
Modeling Conditional Correlations

- For these models, we write the conditional covariance matrix $H_t$ as

$$H_t = D_t R_t D_t$$  \hspace{1cm} (32)

$$D_t = \begin{bmatrix} \sqrt{h_{1t}^2} & 0 & \cdots & 0 \\ 0 & \sqrt{h_{2t}^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{h_{Nt}^2} \end{bmatrix},$$  \hspace{1cm} (33)

i.e., $D_t$ is a diagonal matrix with the conditional standard deviations on its main diagonal, and the conditional correlation matrix

$$R_t = \begin{bmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1N,t} \\ \rho_{12,t} & 1 & \cdots & \rho_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1N,t} & \rho_{2N,t} & \cdots & 1 \end{bmatrix}.$$  \hspace{1cm} (34)
Modeling Conditional Correlations

• In (34),

$$\rho_{ij,t} = \text{Corr}_{t-1}(\epsilon_{it}, \epsilon_{jt}), \quad i, j = 1, \ldots, N, \quad i \neq j,$$

is the conditional correlation between assets $i$ and $j$.

• The conditional covariances are

$$h_{ij,t} = \rho_{ij,t} \sqrt{h_{it}^2 h_{jt}^2}, \quad i \neq j.$$

• Positive definiteness of $H_t$ is guaranteed if the conditional correlation matrix $R_t$ is positive definite and all the conditional standard deviations in $D_t$ are positive.
Constant Conditional Correlations (CCC)\textsuperscript{14}

• Was very popular for some time, because it is very easy to estimate.

• In this model $R_t = R$ is constant in (32), i.e., the \textit{conditional correlations} are constant.

• Note that $R$ is (in general) not the unconditional correlation matrix of $\epsilon_t$.

• Thus the constant conditional correlation matrix cannot be estimated from the sample correlation matrix of the returns.

• However, a simple two–step (consistent) estimation strategy is feasible, as follows.

**Constant Conditional Correlations (CCC)**

1. In the *first step*, separately estimate $N$ univariate GARCH models for the individual returns;\(^{15}\) this gives us conditional variances

$$\hat{h}_{it}^2, \quad t = 1, \ldots, T, \quad i = 1, \ldots, N.$$  \hspace{1cm} (35)

2. Calculate the standardized residuals,

$$\hat{\xi}_{it} = \frac{r_{it}}{\sqrt{\hat{h}_{it}^2}}, \quad t = 1, \ldots, T, \quad i = 1, \ldots, N.$$  \hspace{1cm} (36)

Constant conditional correlation matrix $\mathbf{R}$ is then estimated as the sample covariance matrix of the standardized residuals in (36), i.e.,

$$\hat{\mathbf{R}} = \left[\hat{\rho}_{ij}\right]_{i,j=1,\ldots,N} = \frac{1}{T-1} \sum_{t=1}^{T} \hat{\xi}_t \hat{\xi}'_t.$$  \hspace{1cm} (37)

\(^{15}\)This allows for flexible specification of the univariate processes. For example, we may specify a standard GARCH for one component, AGARCH or EGARCH for another...
Constant Conditional Correlations (CCC)

• In finite samples, the diagonal elements of $R$ will not exactly be unity, so they need standardization.

• This is accomplished by replacing its elements $\hat{\rho}_{ij}$ by $\hat{\rho}_{ij} / \sqrt{\hat{\rho}_{ii} \hat{\rho}_{jj}}$, or, in matrix notation, replacing $\hat{R}$ with

$$\tilde{R} = (I_N \odot \hat{R})^{-1/2} \hat{R} (I_N \odot \hat{R})^{-1/2}. \quad (38)$$

• In other words, $\tilde{R}$ defined by (38) is the sample correlation matrix (rather than the sample covariance matrix) of the standardized residuals $\hat{\xi}_t$ in (36).
Dynamic Conditional Correlation (DCC) Models

- The assumption of constant correlations is very convenient but often unrealistic.

- Thus models for dynamic conditional correlations (DCC) have been proposed.

Dynamic Conditional Correlation (DCC) Models

- This can be written

\[ \epsilon_t \sim N(0, D_t R_t D_t), \]  

(39)

where \( D_t \) as before a matrix with the conditional standard deviations on the main diagonal.

- Then the standardized residuals

\[ \xi_t = D_t^{-1} \epsilon_t. \]  

(40)

- The correlation dynamics in the simplest (scalar) specification are then introduced via

\[ Q_t = (1 - \alpha - \beta) \overline{Q} + \alpha \xi_{t-1} \xi'_{t-1} + \beta Q_{t-1}, \]  

(41)

\[ \alpha, \beta \geq 0, \quad \alpha + \beta < 1, \]

\[ R_t = \{ \text{diag}(Q_t) \}^{-1/2} Q_t \{ \text{diag}(Q_t) \}^{-1/2}. \]  

(42)
Dynamic Conditional Correlation (DCC) Models

• Three–step estimation strategy for high–dimensional data (similar to two–step in the CCC):
  – First step: As in the CCC, i.e., $N$ univariate GARCH models.
  – Second step: Estimate $\mathbf{Q}$ in (41) via the correlation matrix of the standardized returns $\boldsymbol{\xi}$ in (40).
  – Third step: Estimate $\alpha$ and $\beta$, the parameters governing the correlation dynamics.

Dynamic Conditional Correlation (DCC) Models

- To spell out the correlation dynamics for the two-dimensional case, we have

\[ Q_t = \begin{bmatrix} q_{11,t} & q_{12,t} \\ q_{12,t} & q_{22,t} \end{bmatrix} = (1 - \alpha - \beta) \begin{bmatrix} 1 & \bar{q}_{12} \\ \bar{q}_{12} & 1 \end{bmatrix} + \alpha \begin{bmatrix} \bar{\xi}_{1,t-1}^2 & \bar{\xi}_{1,t-1}\bar{\xi}_{2,t-1} \\ \bar{\xi}_{1,t-1}\bar{\xi}_{2,t-1} & \bar{\xi}_{2,t-1}^2 \end{bmatrix} + \beta \begin{bmatrix} q_{11,t-1} & q_{12,t-1} \\ q_{12,t-1} & q_{22,t-1} \end{bmatrix}, \]
Dynamic Conditional Correlation (DCC) Models

and then the conditional correlation matrix

\[
\begin{align*}
R_t &= \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix} \\
&= \{\text{diag}(Q_t)\}^{-1/2} Q_t \{\text{diag}(Q_t)\}^{-1/2} \\
&= \begin{bmatrix} q_{11,t}^{-1/2} & 0 \\ 0 & q_{22,t}^{-1/2} \end{bmatrix} \begin{bmatrix} q_{11,t} & q_{12,t} \\ q_{12,t} & q_{22,t} \end{bmatrix} \begin{bmatrix} q_{11,t}^{-1/2} & 0 \\ 0 & q_{22,t}^{-1/2} \end{bmatrix} \\
&= \begin{bmatrix} 1 & q_{12,t} \sqrt{q_{11,t} q_{22,t}} \\ q_{12,t} \sqrt{q_{11,t} q_{22,t}} & 1 \end{bmatrix}. 
\end{align*}
\]
Dynamic Correlations: Simple Portfolio Illustration

- In mean–variance portfolio theory, it is assumed that investors care about the mean and the variance of the portfolio return, given by

\[
\mu_{p,t} = w_t^\prime \mu_t \quad (43)
\]

\[
\sigma^2_{p,t} = w_t^\prime H_t w_t, \quad (44)
\]

where \(\mu_t\) and \(H_t\) are the conditional mean and covariance matrix of the \(N\) assets, respectively, and \(w_t\) is the portfolio weight vector, satisfying

\[
1'_N w_t = 1. \quad (45)
\]

- Investors dislike variance (risk) and like expected return.

- Minimum variance portfolios are determined such that the variance is minimized for a prespecified portfolio mean return \(\bar{\mu}_p\), i.e.,

\[
\min_{w_t} \sigma^2_{p,t} = w_t^\prime H_t w_t \quad \text{s.t.} \quad w_t^\prime \mu_t = \bar{\mu}_p, \quad \text{and} \quad 1'_N w_t = 1. \quad (46)
\]
Dynamic Correlations: Simple Portfolio Illustration

- Due to the focus on covariance matrices (rather than expected returns), it is custom to assume that the investor wants to globally minimize the variance, i.e., choose the *global minimum variance portfolio* (GMVP), determined by

\[
\min_{\mathbf{w}_t} \sigma_{p,t}^2 = \mathbf{w}_t' \mathbf{H}_t \mathbf{w}_t \quad \text{s.t.} \quad \mathbf{1}_N \mathbf{w}_t = 1. \tag{47}
\]

- Provided there are no constraints on short sales (then in addition $\mathbf{w}_t \geq 0$), the solution of (47) can be given in closed–form as

\[
\mathbf{w}_t^{GMVP} = \frac{\mathbf{H}_t^{-1} \mathbf{1}_N}{\mathbf{1}_N' \mathbf{H}_t^{-1} \mathbf{1}_N}. \tag{48}
\]

- For example, in the two–asset case ($N = 2$), the GMVP weight of the first asset

\[
\mathbf{w}_t^{GMVP} = \frac{h_{2t}^2 - h_{12,t}}{h_{1t}^2 + h_{2t}^2 - 2h_{12,t}}. \tag{49}
\]
Dynamic Correlations: Simple Portfolio Illustration

- Note that these are *ex-ante* portfolio weights, i.e., investors determine optimal portfolio weights $w_t$ *before* the return at time $t$ is observed.

- At the end of period $t$, we know what $r_t$ (the return vector of period $t$) is.

- Thus, we can calculate the *realized* portfolio return $r_{p,t}$ as

$$r_{p,t} = w_t^{GMVP} r_t.$$  \hfill (50)

- If we repeat this for $T^*$ periods, $t = 1, \ldots, T^*$, we obtain a sample of $T^*$ realized one-step-ahead GMVP portfolio returns.
Dynamic Correlations: Simple Portfolio Illustration

- In view of the optimization criterion (i.e., minimization of the variance), we can calculate the standard deviation (or variance) of these realized returns and argue that the best model is the one which produces the smallest standard deviation of out-of-sample portfolio returns.
Dynamic Correlations: Simple Portfolio Illustration

• The same can be done for multi-step-ahead portfolio optimization.

• E.g., if we fit our model to weekly data with monthly rebalancing, then we would need four-step-ahead covariance matrix forecasts.

• With weekly data, let the $\tau$-step forecast be given by

$$\text{Cov}_t(r_{t+\tau}) = H_t(\tau). \quad (51)$$

• Then, assuming constant conditional means,$^{17}$ the monthly covariance forecast is

$$\text{Cov}_t(r_{t+1} + r_{t+2} + r_{t+3} + r_{t+4}) = \sum_{i=1}^{4} H_t(i), \quad (52)$$

and this matrix is used for determining the (estimated) optimal portfolio.

$^{17}$Otherwise covariance terms would come in.
Data

- Weekly returns of MSCI world stock market index and EPRA/NAREIT global real estate equity index, January 1990 to October 2011, $T = 1137$ weekly returns.

- The first 500 observations are used for estimation.

- Thus, we get, e.g., $T^* = 1137 - 500 = 637$ one-step-ahead realized GMVP returns, 636 two-step-ahead (two-week-ahead) realized GMVP returns, etc.\(^{18}\)

---

\(^{18}\)Using overlapping holding periods, which has its drawbacks.
scatter plot of weekly returns

weekly MSCI returns

weekly EPRA NAREIT returns
Dynamic Correlations: Simple Portfolio Illustration

- We compare CCC and DCC, with a bivariate Student’s t innovation process.
- Univariate variances are modeled via GJR (asymmetric) GARCH processes.
- (In contrast to, e.g., BEKK, multi–step covariance matrices for these models have to be evaluated by simulation, which is time–consuming.)
- Estimates of correlation dynamics in DCC model over entire sample:
  \[ \hat{\alpha} = 0.0419, \quad \hat{\beta} = 0.9504. \]  \( \text{(53)} \)
- The sample correlation of returns is 0.7946.
DCC conditional correlations (implied by estimates over the entire sample)
Percentage reduction of GMVP out-of-sample standard deviation from DCC vs. CCC
As the forecast horizon increases, the conditional correlations converge to the unconditional values, and the advantage of using dynamic correlation models disappears.
Appendix: Confirmation of $\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$

- Straightforward to verify:

$$\text{vec}(xy') = y \otimes x$$  \hspace{1cm} (54)

for vectors $x$ and $y$.

- Also straightforward to check that, for conformable matrices,

$$ (A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$  \hspace{1cm} (55)

- Let $e_j$ be the $j$th unit vector in $\mathbb{R}^N$, i.e., an $N$–dimensional (column) vector with a one in its $j$th position and zeros elsewhere, $j = 1, \ldots, N$; then we can write $N \times N$ matrix $B = [b_{ij}]_{i,j=1,\ldots,N}$ as

$$ B = \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij} e_i e'_j. $$  \hspace{1cm} (56)
Hence

$$\text{vec}(ABC') \overset{(56)}{=} \text{vec} \left( A \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij}e_ie'_jC \right) = \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij}\text{vec} \left( Ae_ie'_jC \right)$$

$$\overset{(55)}{=} \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij}(C' \otimes A)(e_j \otimes e_i)$$

$$\overset{(54)}{=} (C' \otimes A) \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij}(e_j \otimes e_i) = (C' \otimes A) \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij}\text{vec}(e_ie'_j)$$

$$\overset{(56)}{=} (C' \otimes A)\text{vec} \left( \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij}e_ie'_j \right) = (C' \otimes A)\text{vec}(B).$$