Multivariate GARCH models: Inference and Evaluation

Alessandro Palandri*

October, 2004

* Dept. of Economics, Duke University, Durham, NC 27708. E-mail: palandri@econ.duke.edu .

Outline

- MGARCH models: current feasible alternatives
- The Proposed Model
- Issues regarding the Realized Correlations
- Evaluation
- Results
- Conclusions
- Future Work

DCC: scalar - BEKK

- This model for conditional correlations, originally proposed by Engle and Sheppard (2001), has been changed to improve comparability across models:

$$R_t = (\bar{R} - \delta\bar{R} - \theta\bar{\Psi} - \beta\bar{\Upsilon}) + \delta R_{t-1} + \theta\Psi_{t-1} + \beta\Upsilon_{t-1}$$

  where:
  - Ψ_{t-1} is the matrix of realized correlations,
  - Υ_{t-1} is a realized correlation matrix introduced to model asymmetries,
  - the coefficients are such that $\delta, \theta, \beta \geq 0$, and $(\bar{R} - \delta\bar{R} - \theta\bar{\Psi} - \beta\bar{\Upsilon})$ is p.d.
- Advantages: estimation is always feasible, regardless of the dimensionality M of the problem.
- Disadvantages: it is an over simplistic model as it constraints the $\frac{M(M-1)}{2}$ correlations to share the same dynamic.

DCC: diagonal - BEKK

- Introduced by Sheppard (2002), and here presented in a different representation, this model allows for asset specific news impact parameters:

$$R_t = (\bar{R} - \Delta\bar{R}\Delta - \Theta\bar{\Psi}\Theta - B\bar{\Upsilon}B) + \Delta R_{t-1}\Delta + \Theta\Psi_{t-1}\Theta + B\Upsilon_{t-1}B$$

  where:
  - Δ, Θ, and B are diagonal matrices of coefficients,
  - the coefficients are such that $(\bar{R} - \Delta\bar{R}\Delta - \Theta\bar{\Psi}\Theta - B\bar{\Upsilon}B)$ is p.d.
- Advantages: there are M parameters governing the dynamics of the $\frac{M(M-1)}{2}$ correlations.
- Disadvantage: since the total number of parameters is $3M$, estimation is infeasible for large M.

Other parsimonious models for R_t and H_t

- CCC: coincides with the DCC - scalar - BEKK with $\delta = \theta = \beta = 0$.
- O-GARCH: fits univariate GARCHes to $N < M$ principal components.
  - Advantages: estimation is always feasible.
  - Disadvantages: covariances are modeled as a linear combination of the GARCH processes with weights determined by the orthogonalization.
- Flexible GARCH
- Copulas

Desirable Properties of an MGARCH model

- Separate estimation of the conditional variances and correlations:
  - allows to choose from a rich set of parameterizations of conditional variances (univariate GARCHes)
  - parsimonious parametrization of the conditional correlations without penalizing that of the variances
- Its correlation structure should allow for asset specific news impact parameters.
- Estimation should be feasible for any M:
  - it must be possible to estimate each conditional correlation separately
  - it must be possible to re-construct a positive definite correlation matrix

Proposed Model

1. Univariate GARCH models are estimated for each of the M assets
2. The standardized returns are used to estimate the conditional correlation:
   (a) estimate the conditional correlation of two returns
   (b) use the resulting $\rho_{ij,t}$ to rotate the returns
   (c) proceed until all pairwise correlations have been estimated
   (d) use all previous results to re-construct R_t.

Estimation of the Correlation Matrix

- Given M standardized returns, fit a correlation model to series 1 and 2:
  - this will produce the following conditional correlations $\{\rho_t^{1,2}\}_{t=1}^{T}$
  - rotate series 2 by pre-multiplying the data by the inverse of a Cholesky:

$$\varepsilon_t^{1,2} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ a & b & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \cdot \varepsilon_t \qquad \forall t$$

where $a = -\rho_t(1-\rho_t^2)^{-1/2}$ and $b = (1-\rho_t)^{-1/2}$.

Estimation of the Correlation Matrix

- Proceed estimating the remaining correlations:
  - $(1, 3), (1, 4), ...(1, M)$
  - $(2, 3), (2, 4), ...(2, M)$
  - $...$
  - $(M - 1, M)$
- In general, once $\rho_t^{p,q}$ has been estimated, $\varepsilon_t^{p,q}$ is computed in the following way:

$$\varepsilon_t^{p,q} = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots \\ 0 & 1 & \dots & 0 & \dots \\ \vdots & \vdots & & & \\ 0 & a & \dots & b & \dots \\ \vdots & \vdots & & \vdots & \ddots \end{pmatrix} \cdot \varepsilon_t^{p,q-1} \qquad \forall t$$

where $a = -\rho_t(1 - \rho_t^2)^{-1/2}$ is the (q, p) element of the matrix and $b = (1 - \rho_t)^{-1/2}$ is the (q, q) element.

Estimation of the Correlation Matrix

- Once all the conditional correlations have been estimated and the series rotated we are going to have $\varepsilon_t^{M-1,M}$ s.t.

$$E_{t-1}\left(\varepsilon_t^{M-1,M}\right) = I$$

- Hence, to reconstruct the correlation matrix of ε_t the procedure must be repeated backward: just that this time we have the estimates of all the conditional correlations.

$$\mathbf{R_t^{M-1,M}} = K_t^{M-1,M} \cdot I \cdot \left(K_t^{M-1,M}\right)'$$

where $K_t^{M-1,M}$ is an identity matrix with element $(M, M-1)$ equal to ρ_t and element (M, M) equal to $(1-\rho_t^2)^{-1/2}$.

Estimation of the Correlation Matrix

- The conditional correlation matrix R_t is positive definite by construction.
- It is a correlation matrix:
  - Assume that every element on the main diagonal of the matrix $R_t^{p,q+1}$ is equal to 1,
  - $R_t^{p,q}$ is obtained by pre-multiplying $R_t^{p,q+1}$ by $K_t^{p,q}$ and post-multiplying by its transpose,
  - such multiplication affects only one of the terms on the main diagonal of $R_t^{p,q+1}$ and that is element (q, q),
  - for notational purposes let: $Z = R_t^{p,q}$ and $Q = R_t^{p,q+1}$:

$$Z_{[q,q]} = \rho^2 Q_{[p,p]} + 2\rho(1-\rho^2)^{1/2} Q_{[p,q]} + (1-\rho^2) Q_{[q,q]}$$

  - since the main diagonal of $R_t^{p,q+1}$ is made of 1s we have $Q_{[i,i]} = 1$:

$$Z_{[q,q]} = 1 + 2\rho(1-\rho^2)^{1/2} Q_{[p,q]}$$

– the element (p, q) of $R_t^{p,q+1}$ is zero (it gets filled in only in $R_t^{p,q}$):

$$Z_{[q,q]} = 1$$

A simple Example

- Suppose that the 'real' conditional correlation matrix at time t is equal to:

$$\mathbf{R_t} = \begin{pmatrix} 1 & 0.6 & -0.4 \\ 0.6 & 1 & 0.2 \\ -0.4 & 0.2 & 1 \end{pmatrix}$$

- Assume that model produces an exact estimate of the conditional correlation between assets 1 and 2: $\rho_t^{1,2} = 0.6$. Through a Cholesky decomposition we obtain $R_t^{1,2}$:

$$\begin{aligned} \mathbf{R_t^{1,2}} &= \begin{pmatrix} 1 & 0 & 0 \\ -0.75 & 1.25 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{R_t} \begin{pmatrix} 1 & -0.75 & 0 \\ 0 & 1.25 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & -0.4 \\ 0 & 1 & 0.55 \\ -0.4 & 0.55 & 1 \end{pmatrix} \end{aligned}$$

A simple Example

- On the rotated data, the estimated conditional correlation between assets 1 and 3 is $\rho_t^{1,3} = -0.4$:

$$\begin{aligned} \mathbf{R_t^{1,3}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.44 & 0 & 1.09 \end{pmatrix} \mathbf{R_t^{1,2}} \begin{pmatrix} 1 & 0 & 0.44 \\ 0 & 1 & 0 \\ 0 & 0 & 1.09 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.6 \\ 0 & 0.6 & 1 \end{pmatrix} \end{aligned}$$

- Finally, the estimated conditional correlation between assets 2 and 3 is $\rho_t^{2,3} = 0.6$.

A simple Example

- The conditional correlation matrix of the rotated series is now equal to the identity matrix.
- The conditional correlation matrix of the original series can be obtained by repeating the whole procedure backward (just that this time we know $\rho_t^{i,j}$.

$$
\begin{aligned}
\mathbf{R_t^{1,2}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -0.4 & 0 & 0.92 \end{pmatrix} \mathbf{R_t^{1,3}} \begin{pmatrix} 1 & 0 & -0.4 \\ 0 & 1 & 0 \\ 0 & 0 & 0.92 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & -0.4 \\ 0 & 1 & 0.55 \\ -0.4 & 0.55 & 1 \end{pmatrix} \\
\mathbf{R_t} &= \begin{pmatrix} 1 & 0 & 0 \\ 0.6 & 0.8 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{R_t^{1,3}} \begin{pmatrix} 1 & 0.6 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

A Model for Correlations

- Let's start with a bivariate scalar BEKK:

$$Q_t = (\bar{Q} - \delta\bar{Q} - \theta\bar{\Psi}) + \delta Q_{t-1} + \theta\Psi_{t-1}$$

- Since the elements on the main diagonal are always equal to 1 we can simply write the equation for the correlation:

$$\rho_t = (\bar{\rho} - \delta\bar{\rho} - \theta\bar{\psi}) + \delta\rho_{t-1} + \theta\psi_{t-1}$$

- ρ_t is a correlation IFF $|\bar{\rho} - \delta\bar{\rho} - \theta\bar{\psi}| < 1 - \delta - \theta$
- A preliminary analysis showed that the previous constraint binds **too** often !
  - from the constraint we can solve for one of the model's parameters,
  - the model now has only one parameter.

A Model for Correlations

- Instead of using a model with 2 parameters where 1 of them ends up being determined by a binding constraint rather than the data we can use a simple exponential smoothing:

$$\rho_t = \delta \rho_{t-1} + (1 - \delta)\psi_{t-1}$$

- A possible generalization that accounts for asymmetries could be:

$$\rho_t = \begin{cases} \theta \rho_{t-1} + (1 - \theta)\psi_{t-1} & \text{if } \varepsilon_{1,t} < 0 \text{ and } \varepsilon_{2,t} < 0 \\ \delta \rho_{t-1} + (1 - \delta)\psi_{t-1} & \text{otherwise} \end{cases}$$

The Realized Correlations: ψ_{t-1}

- The product of the GARCH-standardized residuals $\eta_{i,t-1}\eta_{j,t-1}$ is not exactly a measure of correlation as it is not bounded to the interval $[-1, 1]$.
- Realized correlations ψ_{t-1} can be constructed from the product of the residuals:

a) Proxy 1:

$$\begin{aligned} x_{i,t-1} &= \max\left(\min\left(\eta_{i,t-1}, 1\right), -1\right) \\ x_{j,t-1} &= \max\left(\min\left(\eta_{j,t-1}, 1\right), -1\right) \\ \psi_{t-1} &= x_{i,t-1} \cdot x_{j,t-1} \end{aligned}$$

b) Proxy 2:

$$\begin{aligned} x_{i,t-1} &= \eta_{i,t-1}/\eta_{j,t-1} \\ x_{j,t-1} &= \eta_{j,t-1}/\eta_{i,t-1} \\ \psi_{t-1} &= \mathbf{1}_{[x_{i,t-1}\in[-1,1]]} \cdot x_{i,t-1} + \mathbf{1}_{[x_{j,t-1}\in[-1,1]]} \cdot x_{j,t-1} \end{aligned}$$

The Realized Correlations: ψ_t in a Monte Carlo Study

- Let y_{1s} and y_{2s} be two standardized random variables with constant correlation ρ.
- Compute the naive-proxy for the correlation: $\hat{\rho}_s = y_{1s} \cdot y_{2s}$
- Compute the ψ_s proxies: $\hat{\rho}_s = \psi_s^{(1)}$ and $\hat{\rho}_s = \psi_s^{(2)}$
- Also, let $\psi_s^{(3)} \sim U[-1, 1]$ and $\psi_s^{(4)} \sim U[0, 1]$ or $U[-1, 0]$ depending on the sign of ρ
- Repeat for $S = 10^6$.
- Report the percentage reduction in MAD and MSE of ψ^i w.r.t the naive-proxy $\hat{\rho}_s$

The Realized Correlations: ψ_t in a Monte Carlo Study

	$\rho = \pm 0.9$	$\rho = \pm 0.7$	$\rho = \pm 0.5$	$\rho = \pm 0.3$	$\rho = \pm 0.1$	$\rho = 0$
			Gaussian			
ψ^1	-46% -79%	-41% -78%	-38% -76%	-37% -74%	-37% -74%	-37% -73%
ψ^2	-54% -80%	-44% -73%	-39% -72%	-35% -72%	-32% -73%	-31% -73%
ψ^3	-1% -36%	-11% -45%	-18% -53%	-21% -61%	-22% -66%	-21% -66%
ψ^4	-55% -86%	-65% -92%	-67% -93%	-58% -89%	-37% -76%	-21% -66%
			Student-t_4			
ψ^1	-48% -96%	-44% -95%	-42% -92%	-40% -89%	-39% -85%	-39% -82%
ψ^2	-60% -97%	-49% -94%	-41% -90%	-31% -85%	-21% -79%	-17% -74%
ψ^3	-16% -87%	-19% -87%	-18% -83%	-13% -78%	-4% -71%	0% -65%
ψ^4	-62% -98%	-68% -98%	-67% -98%	-54% -94%	-21% -78%	0% -68%

MGARCH Evaluation

- Comparison between predictions and realizations of the conditional variance-covariance matrix is infeasible:
  - the realization of the variance-covariance matrix is not observable
  - $\varepsilon_t \varepsilon_t' = H_t^{1/2} \eta_t \eta_t' H_t^{1/2}$ is a noisy measurement of the realized variances because of the idiosyncratic term $\eta_t \eta_t'$; Andersen and Bollerslev (1998)
- Ledoit, Santa-Clara, and Wolf (2002) proxy for *integrated volatility*:
  - conditioning at time $t = \tau$ compute the cumulative realizations and in-sample predictions of the time intervals $(\tau, \tau + 20)$
  - compare the two measures of the 1-month variances and covariances using MSE and MAD
- Comparison between predicted and realized correlations:
  - conditioning at time $t = \tau$ compute the average realizations and in-sample predictions of the time intervals $(\tau, \tau + 20)$
  - compare the two measures of the 1-month correlations using MSE and MAD

- Engle and Sheppard (2001):
  - construct a portfolio x_t using weights ω_t
  - standardize the returns using the model's variance $\omega_t' H_t \omega_t$
  - test $VAR\left(\frac{x_t}{\sqrt{\omega_t' H_t \omega t}}\right) = 1$
  - $VAR(\bullet) < 1$: over-estimation of the correlations
  - $VAR(\bullet) > 1$: under-estimation of the correlations

The Data Set

- 69 time series from the NASDAQ-100.
- Sample is 10 years of daily observations (from 9/1/1994 to 8/31/2004) for a total of 2517 returns.
- 31 series were not included in the study because there were not enough observations available: from Apollo Group (2454 obs) to Kmart Holding (340 obs).

Symbols of Included Stocks

ADBE	ALTR	APCC	AMGN	AAPL	AMAT	BBBY	BIIB	BMET	CDWC	CEPH	CHIR
CTAS	CSCO	CMCSA	CPWR	CMVT	COST	DELL	XRAY	ERTS	EXPD	ESRX	FAST
FHCC	FISV	FLEX	GNTX	GENZ	GILD	IACI	INTC	INTU	JDSU	KLAC	LRCX
LNCR	LLTC	MXIM	MEDI	MERQ	MCHP	MSFT	MOLX	NXTL	NVLS	ORCL	PCAR
PDCO	PTEN	PAYX	PSFT	PETM	QLGC	QCOM	ROST	SANM	SIAL	SSCC	SPLS
SBUX	SUNW	SYMC	SNPS	TLAB	TEVA	VRTS	WFMI	XLNX			

Univariate GARCHes selection

- Both a ZARCH and an EGARCH were fit to each series in the sample.
- BIC and MAD on 20 days in-sample predictions were used as a measure fit:

% of wins	BIC	MAD
ZARCH	52.2%	82.6%
EGARCH	47.8%	17.4%

- Selected model was ZARCH for each stock:
  - it tied with EGARCH in terms of BIC but clearly prevailed in terms of MAD
  - in an automated procedure it is more reliable: easier to implement and stable.

The Models

- Constant Correlations:
  - conditional variances have been modeled using univariate ZARCHes and conditional correlations have been been set to their unconditional mean.
- Symmetric Scalar DCC:
  - conditional variances have been modeled using univariate ZARCHes and conditional correlations have been modeled with a scalar-BEKK.
- Proposed Model:
  - ZARCHes for conditional variances and *asymmetric exponential smoothing* for conditional correlations estimated one at the time.

Results on R_t

	CCC	Scalar DCC	Proposed Model
MAD 20-days	1.21e+06	1.27e+06	1.16e+06
MSE 20-days	438,242	497,409	410,001

% improvement of the Proposed Model	CCC	Scalar DCC
MAD 20-days	-4.5%	-8.7%
MSE 20-days	-6.4%	-17.6%

Results on H_t

	Constant Var-Cov	CCC	Scalar DCC	Proposed Model
MAD 20-days	4.94e+08	3.51e+08	3.50e+08	3.39e+08
MSE 20-days	1.28e+11	6.23e+10	6.23e+10	5.89e+10

% improvement of the Proposed Model	Constant Var-Cov	CCC	Scalar DCC
MAD 20-days	-31.4%	-3.5%	-3.1%
MSE 20-days	-54.1%	-5.5%	-5.5%

Minimum Variance Portfolios

	Scalar DCC	Proposed Model
mean(z^2)	1.114606	1.094261
mean(σ^2)	0.780815	0.772310
min(σ)	0.055412	0.055631
max(σ)	1.413843	1.430626
var(z/σ)	1.357249	1.341710

Descriptive Statistics

Minimum Variance Portfolios

	Scalar DCC	Proposed Model
var$(x) = \sigma_x^2$, mean$(x) = \bar{x}$	t=7.33	t=6.95
var$(x) = \sigma_x^2$, mean$(x) = 1$	t=7.26	t=6.88
Newey-West (500-lags), mean$(x) = \bar{x}$	t=3.79	t=4.06
Newey-West (500-lags), mean$(x) = 1$	t=7.37	t=7.68
bootstrap* 95%	(1.26, 1.46)	(1.25, 1.44)
bootstrap* 99%	(1.24, 1.49)	(1.22, 1.47)

Testing the null hypothesis that mean$(x) = 1$ where $x = (z^2/\sigma^2)$.
* S=100,000 replications of independent draws.

Conclusions

- The Proposed Model for correlation:
  - directly models the conditional correlations, one at the time
  - uses an improved proxy for the realized correlations
  - retains positive definiteness of R_t
  - does not suffer from any type of course of dimensionality
  - it is based on a *large* number of very *simple* estimations
  - it performs at least as well as the benchmark Scalar-DCC
  - does not seem to suffer from any problem that may derive from the over-parametrization of the conditional correlations (4692 parameters)
- The *Climber*:
  - feasibly estimated estimates the conditional correlation model even at the boundary of the parameters' space
  - in the univariate cases it provided such good starting values that in general the Newton-Raphson algorithm based on analytical derivatives converges in no more than 3 iterations.

Conclusions

- From the tracking results on R_t and H_t:
  - returns of the 69 stocks of the NASDAQ-100 at the daily frequency appear to have constant correlations.
- From the Monte Carlo Experiment on the proxies for realized correlations:
  - it might be the case that dynamic conditional correlations models cannot outperform constant conditional correlations because of the bad quality of the regressors
  - the latter could be such a noisy measurement that the model cannot do anything better than capture the level of the unconditional correlations
- For this data-set the ZARCH for the conditional variances:
  - is at least as good as an EGARCH according to the BIC
  - performs better according to the MAD in terms of variance tracking
  - it is the optimal choice for an automated procedure

Future Work

- Complete and polish the study of this data-set.
- Include more and possibly more accurate Specification Tests.
- Test different specifications of the conditional correlations.
- Further investigation of the *'realized correlation problem'*.
- Evaluation of the Proposed Model in a setting where conditional correlations are strongly time-varying.