Time Series Analysis

GARCH

Binzhi Chen

Department of Economics
University of Birmingham

MSc in 2020 Fall

An Excursion into Non-linearity Land

Motivation:

the linear structural (and time series) models cannot explain a number of important features common to much financial data:

1. leptokurtosis
2. volatility clustering or volatility pooling
3. leverage effects

Our "traditional" structural model could be something like:

$$y = \beta_1 + \beta_2 x_2 + \ldots + \beta_k x_{kt} + u_t$$

or more compactly $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$.

We also assumed that $u_t \sim \mathrm{N}(0, \sigma^2)$.

A Sample Financial Asset Returns Time Series

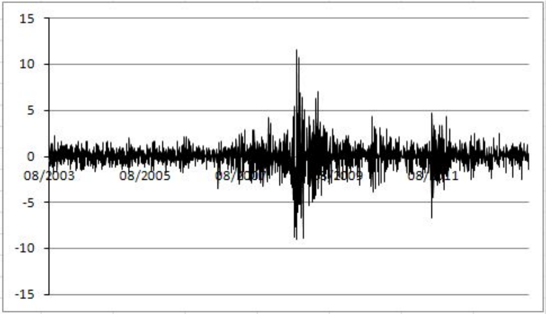


Figure: Daily S&P 500 Returns for August 2003 – August 2013

Non-linear Models: A Definition

Campbell, Lo and MacKinlay (1997) define a non-linear data generating process as one that can be written:

$$y_t = f(u_t, u_{t-1}, u_{t-2}, \ldots)$$

where u_t is an iid error term and $f()$ is a non-linear function.

They also give a slightly more specific definition as:

$$y_t = g(u_{t-1}, u_{t-2}, \ldots) + u_t\ \sigma^2(u_{t-1}, u_{t-2}, \ldots)$$

where $g()$ is a function of past error terms only and σ^2 is a variance term.

Models with nonlinear $g(\bullet)$ are "non-linear in mean", while those with non-linear $\sigma(\bullet)^2$ are "non-linear in variance".

Types of non-linear models

The linear paradigm is a useful one.

Many apparently non-linear relationships can be made linear by a suitable transformation.

On the other hand, it is likely that many relationships in finance are intrinsically non-linear.

There are many types of non-linear models, e.g.

1. ARCH / GARCH
2. switching models
3. bilinear models

Testing for Non-linearity – The RESET Test

The "traditional" tools of time series analysis (acf's, spectral analysis) may find no evidence that we could use a linear model, but the data may still not be independent.

Portmanteau tests for non-linear dependence have been developed.

The simplest is Ramsey's RESET test, which took the form:

$$\hat{u}_t = \beta_0 + \beta_1 \hat{y}_t^2 + \beta_2 \hat{y}_t^3 + \ldots + \beta_{p-1} \hat{y}_t^p + v_t$$

Here the dependent variable is the residual series and the independent variables are the squares, cubes, ... , of the fitted values.

Testing for Non-linearity – The BDS Test

Many other non-linearity tests are available - e.g., the BDS and bispectrum test

BDS is a pure hypothesis test. That is, it has as its null hypothesis that the data are pure noise (completely random)

It has been argued to have power to detect a variety of departures from randomness – linear or non-linear stochastic processes, deterministic chaos, etc)

The BDS test follows a standard normal distribution under the null

The test can also be used as a model diagnostic on the residuals to 'see what is left'

If the proposed model is adequate, the standardised residuals should be white noise.

Models for Volatility

Modelling and forecasting stock market volatility has been the subject of vast empirical and theoretical investigation

There are a number of motivations for this line of inquiry:

1. Volatility is one of the most important concepts in finance
2. Volatility, as measured by the standard deviation or variance of returns, is often used as a crude measure of the total risk of financial assets
3. Many value-at-risk models for measuring market risk require the estimation or forecast of a volatility parameter
4. The volatility of stock market prices also enters directly into the Black–Scholes formula for deriving the prices of traded options

We will now examine several volatility models.

Historical Volatility

The simplest model for volatility is the historical estimate

Historical volatility simply involves calculating the variance (or standard deviation) of returns in the usual way over some historical period

This then becomes the volatility forecast for all future periods

Evidence suggests that the use of volatility predicted from more sophisticated time series models will lead to more accurate forecasts and option valuations

Historical volatility is still useful as a benchmark for comparing the forecasting ability of more complex time models

Heteroscedasticity Revisited

An example of a structural model is

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + u_t$$

with $u_t \sim \mathrm{N}(0, \sigma_u^2)$.

The assumption that the variance of the errors is constant is known as homoscedasticity, i.e. $\text{Var}(u_t) = \sigma_u^2$.

What if the variance of the errors is not constant?

1. heteroscedasticity
2. would imply that standard error estimates could be wrong.

Is the variance of the errors likely to be constant over time? Not for financial data.

Autoregressive Conditionally Heteroscedastic (ARCH)

So use a model which does not assume that the variance is constant.

Recall the definition of the variance of u_t:

$$\sigma_t^2 = \mathrm{var}(u_t \mid u_{t-1}, u_{t-2}, \ldots) = \mathrm{E}[(u_t - \mathrm{E}(u_t))^2 \mid u_{t-1}, u_{t-2}, \ldots]$$

We usually assume that $E(u_t) = 0$

$$\text{so } \sigma_t^2 = \mathrm{var}(u_t \mid u_{t-1}, u_{t-2}, \ldots) = \mathrm{E}[u_t^2 | u_{t-1}, u_{t-2}, \ldots]$$

What could the current value of the variance of the errors plausibly depend upon?

* Previous squared error terms.

Autoregressive Conditionally Heteroscedastic (ARCH)

This leads to the autoregressive conditionally heteroscedastic model for the variance of the errors:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$

This is known as an ARCH(1) model

The ARCH model due to Engle (1982) has proved very useful in finance.

The full model would be

$$y_t = \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t, \qquad u_t \sim \mathrm{N}(0, \sigma_t^2)$$

where $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$

Autoregressive Conditionally Heteroscedastic (ARCH)

We can easily extend this to the general case where the error variance depends on q lags of squared errors:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_q u_{t-q}^2$$

This is an ARCH(q) model.

Instead of calling the variance σ_t^2, in the literature it is usually called h_t, so the model is

$$y_t \quad = \quad \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t \qquad u_t \sim \mathrm{N}(0, h_t)$$

where $h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_q u_{t-q}^2$

Another Way of Writing ARCH

For illustration, consider an ARCH(1). Instead of the above, we can write

$$
\begin{aligned}
y_t &= \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t, \qquad u_t = v_t \sigma_t \\
\sigma_t &= \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2} \qquad v_t \sim \mathrm{N}(0,1)
\end{aligned}
$$

The two are different ways of expressing exactly the same model. The first form is easier to understand while the second form is required for simulating from an ARCH model, for example.

Testing for "ARCH Effects"

1. First, run any postulated linear regression of the form given in the equation above,

$$y_t = \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t$$

saving the residuals, $\hat{u}_t$.

2. Then square the residuals, and regress them on <u>q</u> own lags to test for ARCH of order <u>q</u>, i.e. run the regression

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \gamma_2 \hat{u}_{t-2}^2 + \cdots + \gamma_q \hat{u}_{t-q}^2 + v_t$$

where v_t is iid.

Obtain R^2 from this regression

Testing for "ARCH Effects"

3. The test statistic is defined as TR^2 (the number of observations multiplied by the coefficient of multiple correlation) from the last regression, and is distributed as a $\chi^2(q)$.

The null and alternative hypotheses are

$$
\begin{aligned}
\mathrm{H}_0 &: \gamma_1 = 0 \text{ and } \gamma_2 = 0 \text{ and } \gamma_3 = 0 \text{ and} \ldots \text{and } \gamma_q = 0 \\
\mathrm{H}_1 &: \gamma_1 \neq 0 \text{ or } \gamma_2 \neq 0 \text{ or } \gamma_3 \neq 0 \text{ or} \ldots \text{or } \gamma_q \neq 0
\end{aligned}
$$

If the value of the test statistic is greater than the critical value from the χ^2 distribution, then reject the null hypothesis.

Note that the ARCH test is also sometimes applied directly to returns instead of the residuals from Stage 1 above.

Problems with ARCH (q) Models

How do we decide on q?

The required value of q might be very large

Non-negativity constraints might be violated.

* When we estimate an ARCH model, we require $\alpha_i \geq 0 \, \forall \, i = 0, 1, 2, \ldots, q$ (since variance cannot be negative)

A natural extension of an ARCH(q) model which gets around some of these problems is a GARCH model.

Generalised ARCH (GARCH) Models

Due to Bollerslev (1986). Allow the conditional variance to be dependent upon previous own lags

The variance equation is now

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{1}$$

This is a GARCH(1,1) model, which is like an ARMA(1,1) model for the variance equation.

We could also write

$$\begin{aligned}
\sigma_{t-1}^2 &= \alpha_0 + \alpha_1 u_{t-2}^2 + \beta\sigma_{t-2}^2 \\
\sigma_{t-2}^2 &= \alpha_0 + \alpha_1 u_{t-3}^2 + \beta\sigma_{t-3}^2
\end{aligned}$$

Generalised ARCH (GARCH) Models

Substituting into (1) for σ^2_{t-1}:

$$\begin{aligned}
\sigma^2_t &= \alpha_0 + \alpha_1 u^2_{t-1} + \beta(\alpha_0 + \alpha_1 u^2_{t-2} + \beta\sigma^2_{t-2}) \\
&= \alpha_0 + \alpha_1 u^2_{t-1} + \alpha_0\beta + \alpha_1\beta u^2_{t-2} + \beta^2\sigma^2_{t-2}
\end{aligned}$$

Generalised ARCH (GARCH) Models

Now substituting into (2) for σ^2_{t-2}

$$\begin{aligned}
\sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_0\beta + \alpha_1\beta u_{t-2}^2 + \beta^2(\alpha_0 + \alpha_1 u_{t-3}^2 \\
&\quad + \beta\sigma_{t-3}^2) \\
\sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_0\beta + \alpha_1\beta u_{t-2}^2 + \alpha_0\beta^2 + \alpha_1\beta^2 u_{t-3}^2 \\
&\quad + \beta^3\sigma_{t-3}^2 \\
\sigma_t^2 &= \alpha_0(1 + \beta + \beta^2) + \alpha_1 u_{t-1}^2(1 + \beta L + \beta^2 L^2) + \beta^3\sigma_{t-3}^2
\end{aligned}$$

Generalised ARCH (GARCH) Models

An infinite number of successive substitutions would yield

$$
\begin{aligned}
\sigma_t^2 &= \alpha_0(1 + \beta + \beta^2 + \cdots) + \alpha_1 u_{t-1}^2 (1 + \beta L + \beta^2 L^2 + \cdots) \\
&\quad + \beta^\infty \sigma_0^2
\end{aligned}
$$

So the GARCH(1,1) model can be written as an infinite order ARCH model.

We can again extend the GARCH(1,1) model to a GARCH(p, q):

$$
\begin{aligned}
\sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 \\
&\quad + \beta_2 \sigma_{t-2}^2 + \cdots + \beta_p \sigma_{t-p}^2 \\
\sigma_t^2 &= \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2
\end{aligned}
$$

Generalised ARCH (GARCH) Models

But in general a GARCH(1,1) model will be sufficient to capture the volatility clustering in the data.

Why is GARCH Better than ARCH?

1. more parsimonious - avoids overfitting
2. less likely to breech non-negativity constraints

The Unconditional Variance under GARCH

The unconditional variance of u_t is given by

$$\mathrm{var}(u_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta)}$$

when $\alpha_1 + \beta < 1$

$\alpha_1 + \beta \geq 1$ is termed "non-stationarity" in variance

$\alpha_1 + \beta = 1$ is termed intergrated GARCH

For non-stationarity in variance, the conditional variance forecasts will not converge on their unconditional value as the horizon increases.

Estimation of ARCH / GARCH Models

Since the model is no longer of the usual linear form, we cannot use OLS.

We use another technique known as maximum likelihood.

The method works by finding the most likely values of the parameters given the actual data.

More specifically, we form a log-likelihood function and maximise it.

The steps involved in actually estimating an ARCH or GARCH model are as follows

Estimation of ARCH / GARCH Models

1. Specify the appropriate equations for the mean and the variance - e.g. an AR(1)- GARCH(1,1) model:

$$\begin{aligned} y_t &= \mu + \phi y_{t-1} + u_t, \;\; u_t \sim \mathrm{N}(0, \sigma_t^2) \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned}$$

2. Specify the log-likelihood function to maximise:

$$L = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log\left(\sigma_t^2\right) - \frac{1}{2}\sum_{t=1}^{T}(y_t - \mu - \phi y_{t-1})^2/\sigma_t^2$$

3. The computer will maximise the function and give parameter values and their standard errors

Estimation of GARCH Models Using Maximum Likelihood

Now we have y

$$y_t = \mu + \phi y_{t-1} + u_t, \qquad u_t \sim \mathrm{N}(0, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2$$

$$L = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log\left(\sigma_t^2\right) - \frac{1}{2}\sum_{t=1}^{T}(y_t - \mu - \phi y_{t-1})^2/\sigma_t^2$$

Unfortunately, the LLF for a model with time-varying variances cannot be maximised analytically, except in the simplest of cases. So a numerical procedure is used to maximise the log-likelihood function. A potential problem: local optima or multimodalities in the likelihood surface.

Estimation of GARCH Models Using Maximum Likelihood

The way we do the optimisation is:

1. Set up LLF.
2. Use regression to get initial guesses for the mean parameters.
3. Choose some initial guesses for the conditional variance parameters.
4. Specify a convergence criterion - either by criterion or by value.

Non-Normality and Maximum Likelihood

Recall that the conditional normality assumption for u_t is essential.

We can test for normality using the following representation

$$
\begin{aligned}
u_t &= v_t \sigma_{t,} & v_t &\sim \mathrm{N}(0,1) \\
\sigma_t &= \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2} & v_t &= \frac{u_t}{\sigma_t}
\end{aligned}
$$

The sample counterpart is

$$\hat{v}_t = \frac{\hat{u}_t}{\hat{\sigma}_t}$$

Are the $\hat{v}_t$ normal? Typically $\hat{v}_t$ are still leptokurtic, although less so than the $\hat{u}_t$. Is this a problem? Not really, as we can use the ML with a robust variance/covariance estimator. ML with robust standard errors is called Quasi- Maximum Likelihood or QML.

Extensions to the Basic GARCH Model

Since the GARCH model was developed, a huge number of extensions and variants have been proposed. Three of the most important examples are EGARCH, GJR, and GARCH-M models.

Problems with GARCH(p, q) Models:

1. Non-negativity constraints may still be violated
2. GARCH models cannot account for leverage effects

Possible solutions: the exponential GARCH (EGARCH) model or the GJR model, which are asymmetric GARCH models.

The EGARCH Model

Suggested by Nelson (1991). The variance equation is given by

$$\ln\left(\sigma_t^2\right) = \omega + \beta \ln\left(\sigma_{t-1}^2\right) + \gamma \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right]$$

Advantages of the model

1. Since we model the $\log(\sigma_t^2)$, then even if the parameters are negative, σ_t^2 will be positive.
2. We can account for the leverage effect: if the relationship between volatility and returns is negative, γ, will be negative.

The GJR Model

Due to Glosten, Jaganathan and Runkle

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma u_{t-1}^2 I_{t-1}$$

where $I_{t-1} = 1$ if $u_{t-1} < 0$
$\qquad\qquad = 0$ otherwise

For a leverage effect, we would see $\gamma > 0$.

We require $\alpha_1 + \gamma \geq 0$ and $\alpha_1 \geq 0$ for non-negativity.

An Example of the use of a GJR Model

Using monthly S&P 500 returns, December 1979–June 1998

Estimating a GJR model, we obtain the following results.

$$\underset{(3.198)}{y_t = 0.172}$$

$$\sigma_t^2 = \underset{(16.372)}{1.243} + \underset{(0.437)}{0.015 u_{t-1}^2} + \underset{(14.999)}{0.498 \sigma_{t-1}^2} + \underset{(5.772)}{0.604 u_{t-1}^2 I_{t-1}}$$

News Impact Curves

The news impact curve plots the next period volatility (h_t) that would arise from various positive and negative values of u_{t-1}, given an estimated model.

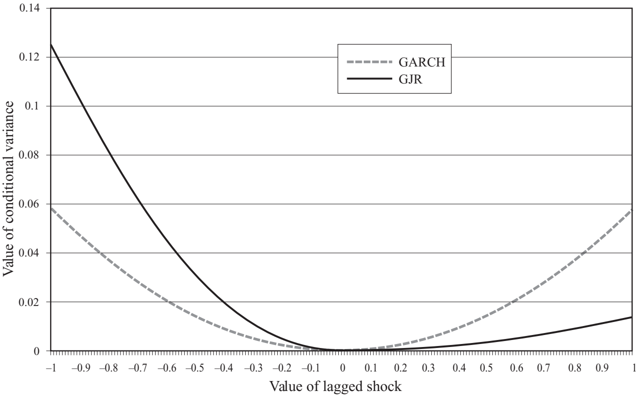


Figure: News Impact Curves for S&P 500 Returns using Coefficients from GARCH and GJR Model Estimates

GARCH-in Mean

We expect a risk to be compensated by a higher return. So why not let the return of a security be partly determined by its risk?

Engle, Lilien and Robins (1987) suggested the ARCH-M specification. A GARCH-M model would be

$$\begin{aligned} y_t &= \mu + \delta\sigma_{t-1} + u_t, \qquad u_t \sim \mathrm{N}(0, \sigma_t^2) \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2 \end{aligned}$$

δ can be interpreted as a sort of risk premium.

It is possible to combine all or some of these models together to get more complex "hybrid" models - e.g. an ARMA-EGARCH(1,1)-M model.

What Use Are GARCH-type Models?

GARCH can model the volatility clustering effect since the conditional variance is autoregressive. Such models can be used to forecast volatility.

We could show that

$$\text{var}\left(y_t \,|\, y_{t-1}, y_{t-2}, \ldots\right) = \text{var}\left(u_t \,|\, u_{t-1}, u_{t-2}, \ldots\right)$$

So modelling σ_t^2 will give us models and forecasts for y_t as well.

Variance forecasts are additive over time.

Forecasting Variances using GARCH Models

Producing conditional variance forecasts from GARCH models uses a very similar approach to producing forecasts from ARMA models.

It is again an exercise in iterating with the conditional expectations operator.

Consider the following GARCH(1,1) model:

$$y_t = \mu + u_t, \quad u_t \sim \mathrm{N}(0, \sigma_t^2), \quad \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2$$

What is needed is to generate are forecasts of ${\sigma_{T+1}}^2|\Omega_T, {\sigma_{T+2}}^2|\Omega_T, \ldots, {\sigma_{T+s}}^2|\Omega_T$ where Ω_T denotes all information available up to and including observation T.

Forecasting Variances using GARCH Models

Adding one to each of the time subscripts of the above conditional variance equation, and then two, and then three would yield the following equations

$$
\begin{aligned}
\sigma^2_{T+1} &= \alpha_0 + \alpha_1 u^2_T + \beta\sigma^2_T \\
\sigma^2_{T+2} &= \alpha_0 + \alpha_1 u^2_{T+1} + \beta\sigma^2_{T+1} \\
\sigma^2_{T+3} &= \alpha_0 + \alpha_1 u^2_{T+2} + \beta\sigma^2_{T+2}
\end{aligned}
$$

Let $\sigma^{f^2}_{1,T}$ be the one step ahead forecast for σ^2 made at time T. This is easy to calculate since, at time T, the values of all the terms on the RHS are known.

Forecasting Variances using GARCH Models

$\sigma^{f2}_{1,T}$ would be obtained by taking the conditional expectation of the first equation presented on the previous slide:

$$\sigma_{1,T}{}^2 \quad = \quad \alpha_0 + \alpha_1 u_T^2 + \beta\sigma_T^2$$

Given, $\sigma^{f2}_{1,T}$ how is $\sigma^{f2}_{2,T}$, the 2-step ahead forecast for σ^2 made at time T, calculated? Taking the conditional expectation of the second equation of the previous slide:

$$\sigma^{f2}_{2,T} = \alpha_0 + \alpha_1 \mathrm{E}(u^2_{T+1} \mid \Omega_T) + \beta\sigma^{f2}_{1,T}$$

where $\mathrm{E}(u^2_{T+1} \mid \Omega_T)$ is the expectation, made at time T, of u^2_{T+1}, which is the squared disturbance term.

Forecasting Variances using GARCH Models

We can write

$$\mathrm{E}(u_{T+1} \mid \Omega_t)^2 = \sigma^2_{T+1}$$

but σ^2_{T+1} is not known at time T, so it is replaced with the forecast for it, $\sigma^{f2}_{1,T}$, so that the 2-step ahead forecast becomes

$$\begin{aligned}
\sigma^{f2}_{2,T} &= \alpha_0 + \alpha_1 \sigma^{f2}_{1,T} + \beta \sigma^{f2}_{1,T} \\
\sigma^{f2}_{2,T} &= \alpha_0 + (\alpha_1 + \beta) \sigma^{f2}_{1,T}
\end{aligned}$$

Forecasting Variances using GARCH Models

By similar arguments, the 3-step ahead forecast will be given by

$$\begin{aligned}
\sigma^{f2}_{3,T} &= \mathrm{E}_T(\alpha_0 + \alpha_1 u^2_{T+2} + \beta\sigma^2_{T+2}) \\
&= \alpha_0 + (\alpha_1 + \beta)\sigma^{f2}_{2,T} \\
&= \alpha_0 + (\alpha_1 + \beta)[\alpha_0 + (\alpha_1 + \beta)\sigma^{f2}_{1,T}] \\
&= \alpha_0 + \alpha_0(\alpha_1 + \beta) + (\alpha_1 + \beta)^2\sigma^{f2}_{1,T}
\end{aligned}$$

Any s-step ahead forecast ($s \geq 2$) would be produced by

$$\sigma^{f2}_{s,T} = \alpha_0 \sum_{i=1}^{s-1} (\alpha_1 + \beta)^{i-1} + (\alpha_1 + \beta)^{s-1}\sigma^{f2}_{1,T}$$

What Use Are Volatility Forecasts?

1. Option pricing

$$C=f(S,\ X,\ \sigma^2,\ T,\ r_f)$$

2. Conditional betas

$$\beta_{i,t} = \frac{\sigma_{im,t}}{\sigma_{m,t}^2}$$

3. Dynamic hedge ratios

The Hedge Ratio - the size of the futures position to the size of the underlying exposure, i.e. the number of futures contracts to buy or sell per unit of the spot good.

What is the optimal value of the hedge ratio?

What Use Are Volatility Forecasts?

Assuming that the objective of hedging is to minimise the variance of the hedged portfolio, the optimal hedge ratio will be given by

$$h = p\frac{\sigma_s}{\sigma_F}$$

where	h =	hedge ratio
	p =	correlation coefficient between change in spot price (S) and change in futures price (F)
	$\sigma_S =$	standard deviation of S
	$\sigma_F =$	standard deviation of F

What if the standard deviations and correlation are changing over time? We use

$$h_t = p_t\frac{\sigma_{s,t}}{\sigma_{F,t}}$$

Testing Non-linear Restrictions or Testing Hypotheses about Non-linear Models

Usual t- and F-tests are still valid in non-linear models, but they are not flexible enough.

There are three hypothesis testing procedures based on maximum likelihood principles: Wald, Likelihood Ratio, Lagrange Multiplier.

Consider a single parameter, θ to be estimated, Denote the MLE as $\hat{\theta}$ and a restricted estimate as $\tilde{\theta}$.

Likelihood Ratio Tests

Estimate under the null hypothesis and under the alternative.

Then compare the maximised values of the LLF.

So we estimate the unconstrained model and achieve a given maximised value of the LLF, denoted L_u

Then estimate the model imposing the constraint(s) and get a new value of the LLF denoted L_r.

Which will be bigger?

$L_r \leq L_u$ comparable to $RRSS \geq URSS$

Likelihood Ratio Tests

The LR test statistic is given by

$$LR = -2(L_r - L_u) \sim \chi^2(m)$$

where m = number of restrictions

Example: We estimate a GARCH model and obtain a maximised LLF of 66.85. We are interested in testing whether $\beta = 0$ in the following equation.

$$y_t = \mu + \phi y_{t-1} + u_t,\ u_t \sim N(0, \sigma_t^2)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2$$

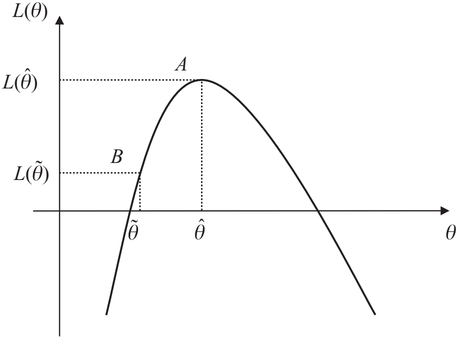
We estimate the model imposing the restriction and observe the maximised LLF falls to 64.54. Can we accept the restriction?

Likelihood Ratio Tests

$LR = -2(64.54 - 66.85) = 4.62$

The test follows a $\chi^2(1) == 3.84$ at 5%, so reject the null.

Denoting the maximised value of the LLF by unconstrained ML as $L(\hat{\theta})$ and the constrained optimum as $L(\tilde{\theta})$. Then we can illustrate the 3 testing procedures in the following diagram:



Hypothesis Testing under Maximum Likelihood

The vertical distance forms the basis of the LR test.

The Wald test is based on a comparison of the horizontal distance.

The LM test compares the slopes of the curve at A and B.

We know at the unrestricted MLE, $L(\hat{\theta})$, the slope of the curve is zero.

But is it "significantly steep" at $L(\tilde{\theta})$?

This formulation of the test is usually easiest to estimate.

An Example of the Application of GARCH Models

- Day & Lewis (1992)

Purpose

To consider the out of sample forecasting performance of GARCH and EGARCH Models for predicting stock index volatility.

Implied volatility is the markets expectation of the “average” level of volatility of an option:

Which is better, GARCH or implied volatility?

An Example of the Application of GARCH Models

- Day & Lewis (1992) Data

Weekly closing prices (Wednesday to Wednesday, and Friday to Friday) for the S&P100 Index option and the underlying 11 March 83–31 Dec. 89

Implied volatility is calculated using a non-linear iterative procedure.

The Models

The “Base” Models

For the conditional mean

$$R_{Mt} - R_{Ft} = \lambda_0 + \lambda_1 \sqrt{h_t} + u_t \tag{2}$$

and for the variance

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} \tag{3}$$

or

$$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 \left(\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left[\left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| - \left(\frac{2}{\pi} \right)^{1/2} \right] \right) \tag{4}$$

The Models

where R_{Mt} denotes the return on the market portfolio
R_{Ft} denotes the risk-free rate
h_t denotes the conditional variance from the GARCH-type models while
σ_t^2 denotes the implied variance from option prices.

Add in a lagged value of the implied volatility parameter to equations 3 and 4.

The Models

(3) becomes

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + \delta \sigma_{t-1}^2 \tag{5}$$

and (4) becomes

$$\begin{aligned} \ln(h_t) &= \alpha_0 + \beta_1 \ln(h_{t-1}) \\ &+ \alpha_1 \left(\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left[\left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| - \left(\frac{2}{\pi} \right)^{1/2} \right] \right) \\ &+ \delta \ln(\sigma_{t-1}^2) \end{aligned} \tag{6}$$

We are interested in testing H_0: $\delta = 0$ in (5) or (6).

The Models

Also, we want to test $H_0 : \alpha_1 = 0$ and $\beta_1 = 0$ in (5),

and $H_0 : \alpha_1 = 0$ and $\beta_1 = 0$ and $\theta = 0$ and $\gamma = 0$ in (6).

If this second set of restrictions holds, then (5) & (6) collapse to

$$h_t = \alpha_0 + \delta \sigma_{t-1}^2 \tag{7}$$

and (6) becomes

$$\ln(h_t) = \alpha_0 + \delta \ln(\sigma_{t-1}^2) \tag{8}$$

We can test all of these restrictions using a likelihood ratio test.

In-sample Likelihood Ratio Test Results: GARCH Versus Implied Volatility

$$R_{Mt} - R_{Ft} = \lambda_0 + \lambda_1\sqrt{h_t} + u_t \qquad (9.79)$$
$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} \qquad (9.80)$$
$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + \delta\sigma_{t-1}^2 \qquad (9.82)$$
$$h_t = \alpha_0 + \delta\sigma_{t-1}^2 \qquad (9.82')$$

Equation for variance	λ_0	λ_1	$\alpha_0 \times 10^{-4}$	α_1	β_1	δ	Log-L	χ^2
(9.80)	0.0072	0.071	5.428	0.093	0.854	–	767.321	17.77
	(0.005)	(0.01)	(1.65)	(0.84)	(8.17)			
(9.82)	0.0015	0.043	2.065	0.266	−0.068	0.318	776.204	–
	(0.028)	(0.02)	(2.98)	(1.17)	(−0.59)	(3.00)		
(9.82′)	0.0056	−0.184	0.993	–	–	0.581	764.394	23.62
	(0.001)	(−0.001)	(1.50)			(2.94)		

Notes: t-ratios in parentheses, Log-L denotes the maximised value of the log-likelihood function in each case. χ^2 denotes the value of the test statistic, which follows a $\chi^2(1)$ in the case of (9.82) restricted to (9.80), and a $\chi^2(2)$ in the case of (9.82) restricted to (9.82′).

Source: Day and Lewis (1992). Reprinted with the permission of Elsevier.

In-sample Likelihood Ratio Test Results: EGARCH Versus Implied Volatility

$$R_{Mt} - R_{Ft} = \lambda_0 + \lambda_1 \sqrt{h_t} + u_t \qquad (9.79)$$

$$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 \left(\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left[\left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| - \left(\frac{2}{\pi} \right)^{1/2} \right] \right) \qquad (9.81)$$

$$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 \left(\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left[\left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| - \left(\frac{2}{\pi} \right)^{1/2} \right] \right) + \delta \ln(\sigma^2_{t-1} g) \qquad (9.83)$$

$$\ln(h_t) = \alpha_0 + \delta \ln(\sigma^2_{t-1} g) \qquad (9.83')$$

Equation for variance	λ_0	λ_1	$\alpha_0 \times 10^{-4}$	β_1	θ	γ	δ	*Log-L*	χ^2
(9.81)	-0.0026	0.094	-3.62	0.529	0.273	0.357	–	776.436	8.09
	(-0.03)	(0.25)	(-2.90)	(3.26)	(-4.13)	(3.17)			
(9.83)	0.0035	-0.076	-2.28	0.373	-0.282	0.210	0.351	780.480	–
	(0.56)	(-0.24)	(-1.82)	(1.48)	(-4.34)	(1.89)	(1.82)		
(9.83′)	0.0047	-0.139	-2.76	–	–	–	0.667	765.034	30.89
	(0.71)	(-0.43)	(-2.30)				(4.01)		

Notes: t-ratios in parentheses, Log-L denotes the maximised value of the log-likelihood function in each case. χ^2 denotes the value of the test statistic, which follows a $\chi^2(1)$ in the case of (9.83) restricted to (9.81), and a $\chi^2(3)$ in the case of (9.83) restricted to (9.83′).

Source: Day and Lewis (1992). Reprinted with the permission of Elsevier.

Conclusions for In-sample Model Comparisons & Out-of-Sample Procedure

IV has extra incremental power for modelling stock volatility beyond GARCH.

But the models do not represent a true test of the predictive ability of IV.

So the authors conduct an out of sample forecasting test.

There are 729 data points. They use the first 410 to estimate the models, and then make a 1-step ahead forecast of the following week's volatility.

Then they roll the sample forward one observation at a time, constructing a new one step ahead forecast at each step.

Out-of-Sample Forecast Evaluation

They evaluate the forecasts in two ways:

The first is by regressing the realised volatility series on the forecasts plus a constant:

$$\sigma^2_{t+1} = b_0 + b_1 \sigma^2_{ft} + \xi_{t+1} \tag{9}$$

where σ^2_{t+1} is the 'actual' value of volatility at time $t+1$, and σ^2_{ft} is the value forecasted for it during period t.

Perfectly accurate forecasts imply $b_0 = 0$ and $b_1 = 1$.

But what is the "true" value of volatility at time t?

Day & Lewis use 2 measures

1. The square of the weekly return on the index, which is SR.
2. The variance of the week's daily returns multiplied by the number of trading days in that week.

Out-of Sample Model Comparisons

$\sigma^2_{t+1} = b_0 + b_1\sigma^2_{ft} + \xi_{t+1}$ (9.84)

Forecasting model	Proxy for *ex post* volatility	b_0	b_1	R^2
Historic	SR	0.0004	0.129	0.094
		(5.60)	(21.18)	
Historic	WV	0.0005	0.154	0.024
		(2.90)	(7.58)	
GARCH	SR	0.0002	0.671	0.039
		(1.02)	(2.10)	
GARCH	WV	0.0002	1.074	0.018
		(1.07)	(3.34)	
EGARCH	SR	0.0000	1.075	0.022
		(0.05)	(2.06)	
EGARCH	WV	−0.0001	1.529	0.008
		(−0.48)	(2.58)	
Implied volatility	SR	0.0022	0.357	0.037
		(2.22)	(1.82)	
Implied volatility	WV	0.0005	0.718	0.026
		(0.389)	(1.95)	

Notes: 'Historic' refers to the use of a simple historical average of the squared returns to forecast volatility; *t*-ratios in parentheses; SR and WV refer to the square of the weekly return on the S&P100, and the variance of the week's daily returns multiplied by the number of trading days in that week, respectively.

Source: Day and Lewis (1992). Reprinted with the permission of Elsevier.

Encompassing Test Results: Do the IV Forecasts Encompass those of the GARCH Models?

$\sigma^2_{t+1} = b_0 + b_1\sigma^2_{It} + b_2\sigma^2_{Gt} + b_3\sigma^2_{Et} + b_4\sigma^2_{Ht} + \xi_{t+1}$

Forecast comparisons	b_0	b_1	b_2	b_3	b_4	R^2
Implied versus GARCH	−0.00010	0.601	0.298	−	−	0.027
	(−0.09)	(1.03)	(0.42)			
Implied versus GARCH	0.00018	0.632	−0.243	−	0.123	0.038
versus Historical	(1.15)	(1.02)	(−0.28)		(7.01)	
Implied versus EGARCH	−0.00001	0.695	−	0.176	−	0.026
	(−0.07)	(1.62)		(0.27)		
Implied versus EGARCH	0.00026	0.590	−0.374	−	0.118	0.038
versus Historical	(1.37)	(1.45)	(−0.57)		(7.74)	
GARCH versus EGARCH	0.00005	−	1.070	−0.001	−	0.018
	(0.370)		(2.78)	(−0.00)		

Notes: t-ratios in parentheses; the *ex post* measure used in this table is the variance of the week's daily returns multiplied by the number of trading days in that week.

Source: Day and Lewis (1992). Reprinted with the permission of Elsevier.

Conclusions of Paper

Within sample results suggest that IV contains extra information not contained in the GARCH / EGARCH specifications.

Out of sample results suggest that nothing can accurately predict volatility!

Stochastic Volatility Models

It is a common misconception that GARCH-type specifications are stochastic volatility models

However, as the name suggests, stochastic volatility models differ from GARCH principally in that the conditional variance equation of a GARCH specification is completely deterministic given all information available up to that of the previous period

There is no error term in the variance equation of a GARCH model, only in the mean equation

Stochastic volatility models contain a second error term, which enters into the conditional variance equation.

Autoregressive Volatility Models

A simple example of a stochastic volatility model is the autoregressive volatility specification

This model is simple to understand and simple to estimate, because it requires that we have an observable measure of volatility which is then simply used as any other variable in an autoregressive model

The standard Box-Jenkins-type procedures for estimating autoregressive (or ARMA) models can then be applied to this series

Autoregressive Volatility Models

For example, if the quantity of interest is a daily volatility estimate, we could use squared daily returns, which trivially involves taking a column of observed returns and squaring each observation

The model estimated for volatility, σ_t^2, is then

$$\sigma_t^2 = \beta_0 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \varepsilon_t$$

A Stochastic Volatility Model Specification

The term 'stochastic volatility' is usually associated with a different formulation to the autoregressive volatility model, a possible example of which would be

$$y_t = \mu + u_t\sigma_t, u_t \sim N(0,1)$$

$$\log(\sigma_t^2) = \alpha_0 + \beta_1 \log(\sigma_{t-1}^2) + \sigma_\eta \eta_t$$

where η_t is another N(0,1) random variable that is independent of u_t.

The volatility is latent rather than observed, and so is modelled indirectly

Stochastic volatility models are superior in theory compared with GARCH-type models, but the former are much more complex to estimate.

Covariance Modelling: Motivation

A limitation of univariate volatility models is that the fitted conditional variance of each series is entirely independent of all others

This is potentially an important limitation for two reasons:

1. If there are 'volatility spillovers' between markets or assets, the univariate model will be mis-specified
2. It is often the case that the covariances between series are of interest too
3. The calculation of hedge ratios, portfolio value at risk estimates, CAPM betas, and so on, all require covariances as inputs

Multivariate GARCH models can be used for estimation of:

1. Conditional CAPM betas
2. Dynamic hedge ratios
3. Portfolio variances

Simple Covariance Models: Historical and Implied

In exactly the same fashion as for volatility, the historical covariance or correlation between two series can be calculated from a set of historical data

Implied covariances can be calculated using options whose payoffs are dependent on more than one underlying asset

The relatively small number of such options that exist limits the circumstances in which implied covariances can be calculated

Examples include rainbow options, 'crack spread' options for different grades of oil, and currency options.

Implied Covariance Models

To give an illustration for currency options, the implied variance of the cross-currency returns is given by

$$\tilde{\sigma}^2(xy) = \tilde{\sigma}^2(x) + \tilde{\sigma}^2(y) - 2\tilde{\sigma}(x, y)$$

where $\tilde{\sigma}^2(x)$ and $\tilde{\sigma}^2(y)$ are the implied variances of the x and y returns, respectively, and $\tilde{\sigma}(x, y)$ is the implied covariance between x and y.

So if the implied covariance between USD/DEM and USD/JPY is of interest, then the implied variances of the returns of USD/DEM and USD/JPY and the returns of the cross-currency DEM/JPY are required.

EWMA Covariance Models

A EWMA specification gives more weight in the covariance to recent observations than an estimate based on the simple average

The EWMA model estimates for variances and covariances at time t in the bivariate setup with two returns series x and y may be written as

$$h_{ij,t} = \lambda h_{ij,t-1} + (1 - \lambda) x_{t-1} y_{t-1}$$

where $i \neq j$ for the covariances and $i = j; x = y$ for the variance specifications.

The fitted values for h also become the forecasts for subsequent periods

EWMA Covariance Models

$\lambda (0 < \lambda < 1)$ denotes the decay factor determining the relative weights attached to recent versus less recent observations

This parameter could be estimated but is often set arbitrarily (e.g., Risk-metrics use a decay factor of 0.97 for monthly data but 0.94 for daily).

EWMA Covariance Models - Limitations

This equation can be rewritten as an infinite order function of only the returns by successively substituting out the covariances:

$$h_{ij,t} = (1-\lambda)\sum_{i=0}^{\infty}\lambda^i x_{t-i} y_{t-i}$$

The EWMA model is a restricted version of an integrated GARCH (IGARCH) specification, and it does not guarantee the fitted variance-covariance matrix to be positive definite

EWMA models also cannot allow for the observed mean reversion in the volatilities or covariances of asset returns that is particularly prevalent at lower frequencies.

Multivariate GARCH Models

Multivariate GARCH models are used to estimate and to forecast covariances and correlations.

The basic formulation is similar to that of the GARCH model, but where the covariances as well as the variances are permitted to be time-varying.

There are 3 main classes of multivariate GARCH formulation that are widely used: VECH, diagonal VECH and BEKK.

VECH and Diagonal VECH

e.g. suppose that there are two variables used in the model. The conditional covariance matrix is denoted H t, and would be 2×2. H_t and VECH(H_t) are

$$H_t = \begin{bmatrix} h_{11t} & h_{12t} \\ h_{21t} & h_{22t} \end{bmatrix}, \qquad VEC(H_t) = \begin{bmatrix} h_{11t} \\ h_{22t} \\ h_{12t} \end{bmatrix}$$

VECH and Diagonal VECH

In the case of the VECH, the conditional variances and covariances would each depend upon lagged values of all of the variances and covariances and on lags of the squares of both error terms and their cross products.

In matrix form, it would be written

$$VECH(H_t) = C + AVECH(\Xi_{t-1}\Xi'_{t-1}) + BVECH(H_{t-1})$$
$$\Xi_t|\psi_{t-1} \sim N(0, H_t)$$

VECH and Diagonal VECH

Writing out all of the elements gives the 3 equations as

$$
\begin{aligned}
h_{11t} &= c_{11} + a_{11}u_{1t-1}^2 + a_{12}u_{2t-1}^2 + a_{13}u_{1t-1}u_{2t-1} + b_{11}h_{11t-1} \\
&\quad + b_{12}h_{22t-1} + b_{13}h_{12t-1} \\
h_{22t} &= c_{21} + a_{21}u_{1t-1}^2 + a_{22}u_{2t-1}^2 + a_{23}u_{1t-1}u_{2t-1} + b_{21}h_{11t-1} \\
&\quad + b_{22}h_{22t-1} + b_{23}h_{12t-1} \\
h_{12t} &= c_{31} + a_{31}u_{1t-1}^2 + a_{32}u_{2t-1}^2 + a_{33}u_{1t-1}u_{2t-1} + b_{31}h_{11t-1} \\
&\quad + b_{32}h_{22t-1} + b_{33}h_{12t-1}
\end{aligned}
$$

VECH and Diagonal VECH

Such a model would be hard to estimate. The diagonal VECH is much simpler and is specified, in the 2 variable case, as follows:

$$
\begin{aligned}
h_{11t} &= \alpha_0 + \alpha_1 u_{1t-1}^2 + \alpha_2 h_{11t-1} \\
h_{22t} &= \beta_0 + \beta_1 u_{2t-1}^2 + \beta_2 h_{22t-1} \\
h_{12t} &= \gamma_0 + \gamma_1 u_{1t-1} u_{2t-1} + \gamma_2 h_{12t-1}
\end{aligned}
$$

BEKK and Model Estimation for M-GARCH

Neither the VECH nor the diagonal VECH ensure a positive definite variance-covariance matrix.

An alternative approach is the BEKK model (Engle & Kroner, 1995).

The BEKK Model uses a Quadratic form for the parameter matrices to ensure a positive definite variance / covariance matrix H_t .

In matrix form, the BEKK model is

$$H_t = W'W + A'H_{t-1}A + B'\Xi_{t-1}\Xi'_{t-1}B$$

BEKK and Model Estimation for M-GARCH

Model estimation for all classes of multivariate GARCH model is again performed using maximum likelihood with the following LLF:

$$\ell(\theta) = -\frac{TN}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T} g(\log |H_t| + \Xi_t' H_t^{-1} \Xi_t g)$$

where N is the number of variables in the system (assumed 2 above), θ is a vector containing all of the parameters, and T is the number of obs.

Correlation Models and the CCC

The correlations between a pair of series at each point in time can be constructed by dividing the conditional covariances by the product of the conditional standard deviations from a VECH or BEKK model

A subtly different approach would be to model the dynamics for the correlations directly

In the constant conditional correlation (CCC) model, the correlations between the disturbances to be fixed through time

Correlation Models and the CCC

Thus, although the conditional covariances are not fixed, they are tied to the variances

The conditional variances in the fixed correlation model are identical to those of a set of univariate GARCH specifications (although they are estimated jointly):

$$h_{ii,t} = c_i + a_i \epsilon^2_{i,t-i} + b_i h_{ii,t-1}, \qquad i = 1, \ldots, N$$

More on the CCC

The off-diagonal elements of H_t, $h_{ij,t}(i \neq j)$, are defined indirectly via the correlations, denoted ρ_{ij}:

$$h_{ij,t} = \rho_{ij} h_{ii,t}^{1/2} h_{jj,t}^{1/2}, \qquad i, j = 1, \ldots, N, i < j$$

Is it empirically plausible to assume that the correlations are constant through time?

Several tests of this assumption have been developed, including a test based on the information matrix due and a Lagrange Multiplier test

There is evidence against constant correlations, particularly in the context of stock returns.

The Dynamic Conditional Correlation Model

Several different formulations of the dynamic conditional correlation (DCC) model are available, but a popular specification is due to Engle (2002)

The model is related to the CCC formulation but where the correlations are allowed to vary over time.

Define the variance-covariance matrix, H_t, as $H_t = D_t R_t D_t$

The Dynamic Conditional Correlation Model

D_t is a diagonal matrix containing the conditional standard deviations (i.e. the square roots of the conditional variances from univariate GARCH model estimations on each of the N individual series) on the leading diagonal

R_t is the conditional correlation matrix

Numerous parameterisations of R_t are possible, including an exponential smoothing approach

The DCC Model – A Possible Specification

A possible specification is of the MGARCH form:

$$Q_t = S \circ (\iota\iota' - A - B) + A \circ u_{t-1}u'_{t-1} + B \circ Q_{t-1}$$

where:
S is the unconditional correlation matrix of the vector of standardised residuals (from the first stage estimation), $u_t = D_t^{-1}\epsilon_t$.

ι is a vector of ones

Q_t is an $N \times N$ symmetric positive definite variance-covariance matrix.

$\circ$ denotes the *Hadamard* or element-by-element matrix multiplication procedure.

This specification for the intercept term simplifies estimation and reduces the number of parameters.

The DCC Model – A Possible Specification

Engle (2002) proposes a GARCH-esque formulation for dynamically modelling Q_t with the conditional correlation matrix, R_t then constructed as

$$R_t = diag\{Q_t^*\}^{-1} Q_t diag\{Q_t^*\}^{-1}$$

where $diag(\cdot)$ denotes a matrix comprising the main diagonal elements of $(\cdot)$ and Q^* is a matrix that takes the square roots of each element in Q.

This operation is effectively taking the covariances in Q_t and dividing them by the product of the appropriate standard deviations in Q_t^* to create a matrix of correlations.

DCC Model Estimation

The model may be estimated in a single stage using ML although this will be difficult. So Engle advocates a two-stage procedure where each variable in the system is first modelled separately as a univariate GARCH

A joint log-likelihood function for this stage could be constructed, which would simply be the sum (over N) of all of the log-likelihoods for the individual GARCH models

In the second stage, the conditional likelihood is maximised with respect to any unknown parameters in the correlation matrix

DCC Model Estimation

The log-likelihood function for the second stage estimation will be of the form

$$\ell(\theta_2|\theta_1) = \sum_{t=1}^{T} g(\log|R_t| + u_t' R_t^{-1} u_t g)$$

where θ_1 and θ_2 denote the parameters to be estimated in the 1^{st} and 2^{nd} stages respectively.

Asymmetric Multivariate GARCH

Asymmetric models have become very popular in empirical applications, where the conditional variances and / or covariances are permitted to react differently to positive and negative innovations of the same magnitude

In the multivariate context, this is usually achieved in the Glosten et al. (1993) framework

Kroner and Ng (1998), for example, suggest the following extension to the BEKK formulation (with obvious related modifications for the VECH or diagonal VECH models)

$$H_t = W'W + A'H_{t-1}A + B'\Xi_{t-1}\Xi'_{t-1}B + D'z_{t-1}z'_{t-1}D$$

where z_{t-1} is an N-dimensional column vector with elements taking the value $-\epsilon_{t-1}$ if the corresponding element of ϵ_{t-1} is negative and zero otherwise.

An Example: Estimating a Time-Varying Hedge Ratio for FTSE Stock Index Returns

Data comprises 3580 daily observations on the FTSE 100 stock index and stock index futures contract spanning the period 1 January 1985–9 April 1999.

Several competing models for determining the optimal hedge ratio are constructed. Define the hedge ratio as β.

1. No hedge (β=0)
2. Naïve hedge (β=1)
3. Multivariate GARCH hedges:
   * Symmetric BEKK
   * Asymmetric BEKK
   * In both cases, estimating the OHR involves forming a 1-step ahead forecast and computing

$$OHR_{t+1} = -\frac{h_{CF,t+1}}{h_{F,t+1}} | \Omega_t$$

OHR Results

		In-sample		
	Unhedged	Naive hedge	Symmetric time-varying hedge	Asymmetric time-varying hedge
	$\beta = 0$	$\beta = -1$	$\beta_t = \frac{h_{FS,t}}{h_{F,t}}$	$\beta_t = \frac{h_{FS,t}}{h_{F,t}}$
(1)	(2)	(3)	(4)	(5)
Return	0.0389	-0.0003	0.0061	0.0060
	{2.3713}	{-0.0351}	{0.9562}	{0.9580}
Variance	0.8286	0.1718	0.1240	0.1211
		Out-of-sample		
	Unhedged	Naive hedge	Symmetric time-varying hedge	Asymmetric time-varying hedge
	$\beta = 0$	$\beta = -1$	$\beta_t = \frac{h_{FS,t}}{h_{F,t}}$	$\beta_t = \frac{h_{FS,t}}{h_{F,t}}$
Return	0.0819	-0.0004	0.0120	0.0140
	{1.4958}	{0.0216}	{0.7761}	{0.9083}
Variance	1.4972	0.1696	0.1186	0.1188

Plot of the OHR from Multivariate GARCH

1. OHR is time-varying and less than 1
2. M-GARCH OHR provides a better hedge, both in-sample and out-of-sample.
3. No role in calculating OHR for asymmetries