Arch Models

Most investors dislike risk taking and require a premium for holding assets with risky payoffs. The variance of an asset has been used to measure risk, and split the risk into a company specific component, which is diversifiable, and a market component which cannot be diversified. This measure of the unconditional volatility does not recognize that there may be predictable patterns in stock market volatility. We will analyze models of conditional (on information at time t-1) volatility. These type of models have the implication for finance that investors can predict the risk. This type of models successfully characterize the fact that stock prices seem to go through long periods of high and long periods of low volatility.

The fact that market participants may predict volatility has important implications The most important is that for periods where the investor has forecasted prices to be very volatile, she should either exit the market or require a large premium as a compensation for bearing an unusual high risk.

Empirical Regularities of Asset Returns.

i Thick Tails

Asset returns tend to be leptokurtotic . The documentation of this empirical regularity is presented in Mandelbrot (1965).

ii Volatility Clustering

" ... large changes tend to be followed by large changes, of either sign and small changes tend to be followed by small changes "

iii Leverage Effects

The so-called "leverage effect" first noted by Black(1976) refers to the tendency for stock prices to be negatively correlated with changes in stock volatility. A firm with debt and equity outstanding typically becomes more highly leveraged when the value of the firm falls. This raises the equity return volatility.

iv) Non-Trading Periods

Information that accumulates when financial markets are closed is reflected in prices after the markets reopen. If for example, information accumulates at a constant rate over calendar time, then the variance of the returns over the period from Friday close to the Monday close should be three times the variance from the Monday close to the Tuesday close.

v) Forecastable Events

Patell and Wolfson (1979,1981) show that individual firm's stock returns volatility is high around earning announcements.

Introduction: Conditional and Unconditional moments

Before presenting the alternative Arch type models, we will briefly review the difference between conditional and unconditional moments.

Let us assume that y_t follows a random walk, i.e.

$$y_t = y_{t-1} + \varepsilon_t$$

Then

$$y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i$$

Unconditional Moments

The unconditional mean and variance are;

$$E(y_t) = y_0$$
$$V(y_t) = t\sigma^2$$

A RW has a constant unconditional mean but a time varying unconditional variance.

Conditional Moments

The conditional mean and variance are;

$$E(y_t|y_{t-1}) = y_{t-1}$$
$$V(y_t|y_{t-1}) = E(y_t - E(y_t|y_{t-1}))^2 = E(y_{t-1} + \varepsilon_t - E(y_t|y_{t-1}))^2 = \sigma^2$$

A RW has a constant unconditional mean but a time varying unconditional variance.

So while the unconditional variance of a random walk model tends to infinite as t increase, the conditional variance is constant.

Univariate Parametric Models

Arch Models

In the linear Arch(q) model originally introduced by Engle(1982), the time varying conditional variance is postulated to be a linear function of the past q squared innovations.

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 = w + \alpha(L)\varepsilon_{t-1}^2$$

A sufficient condition for the conditional variance to be positive is that the parameters of the model satisfy the following constraints; $\omega > 0$ and $\alpha_1 > 0, \alpha_2 > 0, , ...\alpha_q > 0$

Defining $\nu_t \equiv \varepsilon_t^2 - \sigma_t^2$, the ARCH(q) model can be re-written as

$$\varepsilon_t^2 = \omega + \alpha(L)\varepsilon_{t-1}^2 + \nu_t$$

(Notice that $\sigma_t^2 = E(\varepsilon_t^2 | \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, ...)$) Since $E_{t-1}(\nu_t) = 0$, the model corresponds to an AR(q) model for the squared innovations, ε_t^2 . Then, the process is covariance stationary if and only if the sum of the positive autoregressive parameters is less than one, in which case the unconditional variance equals

$$Var(\varepsilon_t^2) = \omega/(1 - \alpha_1 - \alpha_2 ... - \alpha_q).$$

Even though ε_t 's are serially uncorrelated they are clearly not independent through time. In accordance with the stylized facts for assets returns discussed above, there is a tendency for large (small) absolute values of the process to be followed by other large (small) values of unpredictable sign.

The ARCH(1) Model

Constant unconditional Variance but non-constant conditional Variance.

Some useful statistical results are given below for the simplest ARCH(1) model, which is identical to the one used by Engle (1982). The main result is that this simple model exhibits constant unconditional variance but non-constant conditional variance.

Consider the following model

$$y_t = \mu + \varepsilon_t$$
$$\varepsilon_t = u_t(\omega + \alpha\varepsilon_{t-1}^2)^{1/2}, u_t \tilde{\ } IIN(0,1), \omega > 0, \alpha > 0$$

(NOTICE that $(\omega + \alpha\varepsilon_{t-1}^2)^{1/2}$ is the conditional standard deviation, σ_t defined as $(E(\varepsilon_t^2 | \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, ...))^{1/2}$.

i) The conditional expectation of ε_t is equal to zero

$$E(\varepsilon_t | \varepsilon_{t-1}) = E(u_t | \varepsilon_{t-1})(\omega + \alpha\varepsilon_{t-1}^2)^{1/2} = 0$$

Notice that $E(u_t|\varepsilon_{t-1}) = E(u_t) = 0$, since u_t ~IIN(0,1)

$ii)$ The conditional variance is given by the following formula

$$Var(\varepsilon_t|\varepsilon_{t-1}) = E(u_t^2|\varepsilon_{t-1})(\omega + \alpha\varepsilon_{t-1}^2) = (\omega + \alpha\varepsilon_{t-1}^2)$$

Notice that $E(u_t^2|\varepsilon_{t-1}) = E(u_t^2) = 1$, since u_t ~IIN(0,1)

Then the conditional mean and variance of y_t are given by the following formulae;

$$E(y_t|y_{t-1}) = \mu$$

$$Var(y_t|y_{t-1}) = (\omega + \alpha\varepsilon_{t-1}^2)$$

Then, the conditional variance of y_t is time varying. On the other hand it can be easily seen that the unconditional variance is time invariant whenever ε_t^2 is stationary, i.e.

$$V(y_t) = V(\varepsilon_t) = \omega/(1-\alpha)$$

whenever the process is stationary.

(since $V(\varepsilon_t) = E(\varepsilon_t^2) = E(\omega + \alpha\varepsilon_{t-1}^2) = \omega + \alpha E(\varepsilon_{t-1}^2)$)

First Order Autoregressive Process with ARCH effects.

For more complicated models such as AR(1)-ARCH(1), we obtain similar results provided that the process for y is stationary, i.e. that the autoregressive parameter is smaller than one in absolute value.

Assume the following first order autoregressive process

$$y_t = \theta y_{t-1} + \varepsilon_t$$

where $\varepsilon_t = u_t(\omega + \alpha\varepsilon_{t-1}^2)$ and u_t ~IIN(0,1) , ω >0 , α> 0

then

$i)$ The conditional expectation of ε_t is equal to zero

$$E(\varepsilon_t|\varepsilon_{t-1}) = E(u_t^2|\varepsilon_{t-1})(\omega + \alpha\varepsilon_{t-1}^2) = 0 \ since \ E(u_t|\varepsilon_{t-1}) = E(u_t) = 0$$

$ii)$ The conditional variance is given by the following formula

$$Var(\varepsilon_t|\varepsilon_{t-1}) = E(u_t^2|\varepsilon_{t-1})(\omega + \alpha\varepsilon_{t-1}^2) = (\omega + \alpha\varepsilon_{t-1}^2)$$

since $E(u_t^2|\varepsilon_{t-1}) = E(u_t) = 1$

Then the conditional mean and variance of y_t are given by the following formulae;

$$E(y_t|y_{t-1}) = \theta y_{t-1}$$

$$Var(y_t|y_{t-1}) = (\omega + \alpha\varepsilon_{t-1}^2)$$

To find the unconditional variance of y_t we recall the following property for the variance;

$$Var(y_t) = E(Var(y_t|y_{t-1})) + Var(E(y_t|y_{t-1}))$$

then

i) $E(Var(y_t|y_{t-1})) = E(\omega + \alpha\varepsilon_{t-1}^2) = \omega + \alpha E(\varepsilon_{t-1}^2) = \omega + \alpha Var(\varepsilon_{t-1})$

ii) $Var(E(y_t|y_{t-1})) = \theta^2 Var(y_{t-1})$

Then if the process is covariance stationary we have

$$\begin{aligned} Var(y_t) &= \frac{\omega + \alpha Var(\varepsilon_{t-1})}{(1-\theta^2)} \\ &= \frac{\omega}{(1-\alpha)(1-\theta^2)} \end{aligned}$$

(Since $Var(\varepsilon_{t-1}) = \omega/((1-\alpha))$

GARCH Models

In empirical applications it is often difficult to estimate models with large number of parameters, say ARCH(q). To circumvent this problem Bollerslev (1986) proposed the Generalized ARCH or GARCH(p, q) model,

$$\begin{aligned} \sigma_t^2 &= \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \\ &= \omega + \alpha(L)\varepsilon_{t-1}^2 + \beta(L)\sigma_{t-1}^2 \end{aligned}$$

A sufficient condition for the conditional variance in the GARCH(p, q) model to be well defined is that all the coefficients in the infinite order linear ARCH model must be positive. Provided that $\alpha(L)$ and $\beta(L)$ have no common roots and that the roots of the polynomial in L, $(1 - \beta(L)) = 0$ lie outside the unit circle, this positive constraint is satisfied, if and only if, the coefficients of the infinite power series expansion for $\alpha(L)/(1 - \beta(L))$ are non-negative.

Rearranging the GARCH(p, q) model by defining $\nu_t \equiv \varepsilon_t^2 - \sigma_t^2$, it follows that

$$\varepsilon_t^2 = \omega + (\alpha(L) + \beta(L))\varepsilon_{t-1}^2 - \beta(L)\nu_{t-1} + \nu_t$$

which defines an ARMA(Max(p,q),p) model for ε_t^2

By standard arguments, the model is covariance stationary if and only if all the roots of $(1 - \alpha(L) - \beta(L))$ lie outside the unit circle.

If all the coefficients are positive, this is equivalent to the sum or the autoregressive coefficients being smaller that 1.

The analogy to ARMA class of models also allows for the use of standard time series techniques in the identification of the orders of p and q.

In most empirical applications with finitely sampled data, the simple GARCH$(1,1)$ is found to provide a fair description of the data.

Persistence and Stationarity

Using the GARCH(1,1) model it is easy to construct multi period forecasts of volatility. When $\alpha+\beta < 1$, the unconditional variance of ε_{t+1} is $\omega/(1-\alpha-\beta)$.

If we re-write the following GARCH(1,1) as

$$\begin{aligned}\sigma_t^2 &= \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \\ &= \omega + \alpha(\varepsilon_{t-1}^2 - \sigma_{t-1}^2) + (\alpha+\beta)\sigma_{t-1}^2\end{aligned}$$

The coefficient measures the extent to which a volatility shock today feeds through into next periods volatility, while $(\alpha+\beta)$ measures the rate at which this effect dies over time. Recursively substituting and using the law of iterated expectations, the conditional expectation of volatility j periods ahead is,

$$E_t[\sigma_{t+j}^2] = (\alpha+\beta)^j(\sigma_t^2 - \omega/(1-\alpha-\beta)) + w/(1-\alpha-\beta)$$

The multi period volatility forecast reverts to its unconditional mean at rate $(\alpha+\beta)$.

IGARCH Models

Integrated GARCH models are processes were the autorregresive part of the square residuals has a unit root, i.e., $(\alpha+\beta) = 1$. For this case the conditional expectation of the volatility j periods ahead is

$$E_t[\sigma_{t+j}^2] = \sigma_t^2 + j\omega.$$

This process looks very much as a random walk with drift ω. Then, if ε_t follows an IGARCH process the unconditional variance does not exist and therefore it is not covariance stationary. Nelson(1990) shows that the analogy with the random walk process should be treated with caution since the IGARCH process is not covariance stationary but it may be proved to be strictly stationary.

For example when $\omega = 0$, $E_t[\sigma_{t+j}^2] = \sigma_t^2$, so volatility is a martingale. But the volatility remains bounded , since it cannot be negative, and then using the fact that a bounded martingale must converge, we can show that it converges to zero, a degenerate distribution.

Despite the fact that it seems to be an empirical regularity that volatility is IGARCH (many estimated models have coefficients that sum near 1) we regard this type of process as unlikely. (see section on structural breaks and GARCH models)

EGARCH Models

Even if GARCH models successfully capture thick tailed returns, and volatility clustering, are not well suited to capture the "leverage effect" since the conditional variance is a function only of the magnitudes of the lagged residuals and not their signs

In the exponential GARCH (EGARCH) model of Nelson (1991) σ_t^2 depends on both the size and the sign of lagged residuals.

EGARCH(1,1) Models

$$\ln \sigma_t^2 = \alpha_0 + \beta_1 \ln \sigma_{t-1}^2 + \gamma_0([|\varepsilon_{t-1}/\sigma_{t-1}| - (2/\pi)^{1/2}] + \delta(\varepsilon_{t-1}/\sigma_{t-1}))$$

Obviously the EGARCH model always produces a positive conditional variance σ_t^2 for any choice of α_0, β_1, γ_0 and so that no restrictions need to be placed on these coefficients (except $|\beta_1| < 1$). Because of the use of both $|\varepsilon_t/\sigma_t|$ and (ε_t/σ_t), σ_t^2 will also be non-symmetric in ε_t and, for negative δ, will exhibit higher volatility for large negative ε_t .

Other ARCH Specifications

Glosten, Jagannathan and Runkle (1989) proposed the following specification:

$\varepsilon_t = \sigma_t \nu_t$, where v_t is iid.

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2 I_{t-1},$$

where, $I_{t-1} = 1$ if $\varepsilon_{t-1} \geq 0$ and $I_{t-1} = 0$ if $\varepsilon_{t-1} < 0$.

The non-negativity condition is satisfied provided that all the parameters are positive. If leverage effects do exist, $\alpha_2 < 0$.

Additional Explanatory Variables.

It is straightforward to add other explanatory variables to a GARCH specification. Glosten, Jagannathan and Runkle(1993) add a short-term nominal interest rate to various GARCH models and show that it has a significant positive effect on stock market volatility.

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + \gamma X_{t-1},$$

where X is any positive variable.

GARCH in Mean Models

Many theories in finance assume some kind of relationship between the mean of a return and its variance . A way to take this into account is to explicitly write the returns as a function of the conditional variance or, in other words, to include the conditional variance as another regressor. *GARCH in Mean Models* allow for the conditional variance to have mean effects. Most of the time this conditional variance term will have the interpretation of a time varying risk premium.

Consider the following model.

$$y_t = \theta x_t + \psi\sigma_t^2 + \varepsilon_t$$

and

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_{t-1}^2 + \beta(L)\sigma_{t-1}^2$$

Consistent estimation of θ and ψ is dependent on the correct specification of the entire model. The estimation of GARCH in Mean type of models is numerically unstable and many empirical applications have used ARCH-M type of models which are easier to estimate.

An ARCH in Mean model simply models the conditional variance as an ARCH model instead of modeling as GARCH, i.e.

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_{t-1}^2$$

Example of an ARCH(1)-M

Consider a simple version of the above model.

$$y_t = \psi\sigma_t^2 + \varepsilon_t$$

where $\varepsilon_t = v_t\sigma_t \quad v_t$ ~N(0,1)

$$\sigma_t^2 = w + \omega + \alpha\varepsilon_{t-1}^2$$

Then y_t may be expressed as

$$y_t = \psi(\omega + \alpha\varepsilon_{t-1}^2) + \varepsilon_t$$

Then the expected value of y_t is

$$E(y_t) = \psi\omega + \psi\alpha E(\varepsilon_{t-1}^2)$$

and using that $E(\varepsilon_{t-1}^2) = \omega/(1-\alpha)$then

$$E(y_t) = \psi\omega + \psi\alpha\omega/(1-\alpha)$$

Which can be viewed in finance models as the unconditional expected return for holding a risky asset.

Testing for Arch

Before attempting to estimate a GARCH model you should first check if there are ARCH effects in the residuals of the model. Clearly we should not explicitly model (and estimate) the conditional volatility of series as GARCH when there are not signs of Arch effects.

The original Lagrange Multiplier test for ARCH proposed by Engle (1982) is very simple to compute, and relatively easy to derive. Under the null hypothesis it is assumed that the model is say an AR(p) model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t$$

where ε_t is a Gaussian white noise process, $\varepsilon_t|I_{t-1}$ ~N(0,σ^2)where I_t is the information set. The Alternative hypothesis is that the errors are ARCH(q).

The test for ARCH(q) effect simply consists on regressing

$$\hat{\varepsilon}_t^2 = \alpha_1\hat{\varepsilon}_{t-1}^2 + \alpha_2\hat{\varepsilon}_{t-2}^2 + ... + \alpha_q\hat{\varepsilon}_{t-q}^2 + \psi_t$$

Under the null hypothesis that $\alpha_1 = \alpha_2 = ... = \alpha_q = 0$, and TR^2 is asymptotically distributed $\chi(q)$, where T is the number of observations.

While this is the most widely used test we should be cautious in interpreting the results. If the model is misspecified it is quite likely to reject the null hypothesis simply because most of the time serial correlation in the residuals will induce serial correlation in the squared residuals.

Structural Breaks and ARCH effects

It has been shown (Diebold (1986)) that breaks in the variance, which are not taken into account by the econometrician, will look as ARCH effects when the whole sample is used. In other words, it might be that for a sub sample the unconditional variance changes from say to and then back to the previous level. In this case to model the conditional variance as an ARCH model will be the wrong thing to do. In this case, it is recommended to divide the sample and test for ARCH for the sub periods, if no ARCH effects are found for any of the sub periods but are found for the whole sample that is a clear indication of a break in the unconditional variance and not of ARCH effects. Many researchers wrongly estimated GARCH Models in many situations where there was only a change in regime. For example many papers use GARCH models to fit interest rate series for USA when the change in the Volatility was simply a result of the different operative procedures of the Federal Reserve (a different distribution).

GARCH Effects and Sampling Frequency

It can be proved that GARCH models do not temporarily aggregate, or in other words if a model is GARCH using daily data cannot be GARCH with weakly data and so on. Given that we don't observe the data generating process in practice is very difficult to determine at which sampling frequency the data presents GARCH effects (if it has at all). Nevertheless there are some well established empirical regularities that show that the higher is the sampling frequency (say daily) the higher the GARCH effects. Weakly and every forth night data seem to also present GARCH effect. Monthly data usually does not have GARCH effects and whenever these are detected, are usually due to a structural break of the unconditional variance.

Estimating GARCH Models

Maximum likelihood Estimation with Gaussian Errors

The estimation of GARCH type models is easily done by conditional maximum likelihood.

If the model to be estimated is

$$y_t = x_t\theta + \varepsilon_t$$

Where x_t is a (row) vector of predetermined variables, which could include lagged variables, θ is a parameter vector and ε_t ~N(0,σ_t^2), where the conditional variance is assumed to be GARCH(1,1), i.e. ;

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

Then the conditional distribution of y_t is

$$f(y_t|x_t, I_{t-1}) = (2\pi\sigma_t^2)^{-.5}\exp(-.5(y_t - x_t\theta)^2/\sigma_t^2)$$

Then the conditional log likelihood is

$$\log L(\theta, \omega, \alpha, \beta|I_{t-1}) = \sum_{t=1}^{T}(-.5\log(2\pi) - .5\log(\sigma_t^2) - .5\sigma_t^{-2}(y_t - x_t\theta)^2)$$

Notice that at time 1 we need initial values for ε_0 and σ_0^2. These values are usually assumed to be the equilibrium values, that is $\sigma_0^2 = \omega/(1 - \alpha - \beta)$ and $\varepsilon_0 = (\omega/(1 - \alpha - \beta))^{.5}$

Maximum likelihood Estimation with non Gaussian Errors

The unconditional distribution of many financial time series seems to have fatter tails than the normal. GARCH effects may not account for this and we need to use another distribution for ε_t . A tractable distribution is the t-distribution. We proceed as before but replace the Normal density function by

$$f(\varepsilon_t) = (\Gamma[(\nu+1)/2]/\Gamma(\nu/2))((\nu-2)\pi\sigma_t^2)^{-.5}[1+(\varepsilon_t^2/(\sigma_t^2(\nu-2))]^{-(\nu+1)/2}$$

Where n is a parameter to be estimated which represents the degrees of freedom. We estimate as before numerically subject to the constraint that n is greater than 2.

Stochastic Volatility Models

A possible response to non-normality of returns conditional upon past returns is to assume that there is a random variable conditional upon which returns are normal, but this variable-which we may call stochastic volatility-is not directly observed.

A simple example of a stochastic-volatility model is the following:

$$\eta_t = \varepsilon_t e^{\alpha_t/2}, \alpha_t = \phi\alpha_{t-1} + \xi_t$$

where ε_t ~N(0,σ_ε^2), ξ_t ~N(0,σ_ξ^2), and we assume that ε_t and ξ_t are serially uncorrelated and independent of each other.

If we squared η_t, the returns equation, and take logs we can write this expression as

$$\log(\eta_t^2) = \alpha_t + \log(\varepsilon_t^2), \quad \alpha_t = \phi\alpha_{t-1} + \xi_t$$

This is in linear state-space (to be covered in the course) form except that has an error with a log χ^2 distribution instead of a normal distribution.

How to compare Between Different GARCH - Specifications

Most of the GARCH models are non-nested (they cannot be written as a restricted version of a more general process). Therefore the comparison between different GARCH Models is no straight forward.

Misspecification Tests on the standardized residuals.

We have seen above that the residuals may be written as the product of a WN and the conditional standard deviation. For example for an ARCH(1) this can be written as

$$\varepsilon_t = v_t(\omega + \alpha\varepsilon_{t-1}^2)^{1/2}$$

Therefore we can test for the existence ARCH effects in the standardized residuals

$$\hat{v}_t = \hat{\varepsilon}_t/(\hat{\omega} + \hat{\alpha}\hat{\varepsilon}_{t-1}^2)^{1/2}$$

The model that "cleans" the standardized residuals is a candidate to be the "true" model.

Some Other Ways Of Discriminating between alternative ARCH models.

We are going to present two alternative ways of discriminating between ARCH models; (i) based on the use of auxiliary regressions of the squared residuals, (ii) based in their forecasting ability.

(i) Comparison between alternative models based on the use of auxiliary regressions

Pagan and Scwhert (1989) suggest to use the following auxiliary regression as a mean of choosing between different Arch models.

$$\hat{\varepsilon}_t^2 = \alpha + \beta\hat{\sigma}_t^2 + \xi_t$$

This regress the squared residuals on the fitted variance of the alternative GARCH models. If the chosen GARCH model is appropriate to explain the conditional volatility of the series under scrutiny, you should expect α to be zero, β to be one and the fit (R^2) to be good.

Pagan and Scwhert (1989) propose to test the joint hypothesis

$H_0) \alpha = 0$, $\beta = 1$
$H_1) \alpha \neq 0$, $\beta \neq 1$

As a second step, they propose to compare the models that were not rejected on the basis of goodness of fit. The argument being, the one with better fit the better that mimics the conditional variance.

They also propose to express the previous regression in logarithms to account for scale effects and then compare the goodness of fit of this alternative auxiliary regression.

(ii) Measuring the Accuracy of Forecasts of Different Arch Models.

Hamilton (1994) propose to use the forecasting ability of the different ARCH models as a way of comparing these models. As we said before, the ARCH type of models have the property that they allow to forecast the conditional variance of a series, therefore a criteria which may enable us to choose between different models is to choose that one that forecast better.

Various measures (*loss functions*) have been proposed for assessing the predictive accuracy of the forecasting ARCH models.

The Mean Squared Error

$$MSE = (1/T)(\sum_{t=1}^{T}(\hat{\varepsilon}_t^2 - \hat{\sigma}_t^2)^2)$$

The Mean Absolute Error.

$$MAE = (1/T)(\sum_{t=1}^{T}|\hat{\varepsilon}_t^2 - \hat{\sigma}_t^2|)$$

The Mean Squared Error of the log of the squared residuals.

$$[LE]^2 = (1/T)(\sum_{t=1}^{T}(\ln(\hat{\varepsilon}_t^2) - \ln(\hat{\sigma}_t^2))^2)$$

The Mean Absolute Error of the log of the squared residuals.

$$[MAE]^2 = (1/T)(\sum_{t=1}^{T}|ln(\hat{\varepsilon}_t^2) - ln(\hat{\sigma}_t^2)|)$$

For all the models we calculate the proportional improvement over a model which assumes constant variance, i.e., $\hat{\sigma}_t^2 = \hat{\sigma}^2$ (to account for scale effects). The model that provides the largest proportional improvement is the one to be preferred.

Hamilton also propose to compare the forecasting performance at different horizons (4 and 8 periods). That will slightly modify the above formulae in the following way;

The Mean Squared Error

$$MSE = (1/T)(\sum_{t=1}^{T}(\hat{\varepsilon}_{t+\tau}^2 -_{\tau} \hat{\sigma}_t^2)^2)$$

The Mean Absolute Error.

$$MAE = (1/T)(\sum_{t=1}^{T}|\hat{\varepsilon}_{t+\tau}^2 -_{\tau} \hat{\sigma}_t^2|)$$

The Mean Squared Error of the log of the squared residuals.

$$[LE]^2 = (1/T)(\sum_{t=1}^{T}(\ln(\hat{\varepsilon}_{t+\tau}^2) - \ln(\hat{\sigma}_t^2))^2)$$

The Mean Absolute Error of the log of the squared residuals.

$$[MAE]^2 = (1/T)(\sum_{t=1}^{T} | \ln(\hat{\varepsilon}^2_{t+\tau}) - \ln(\hat{\sigma}^2_t)|)$$

where ${}_\tau\hat{\sigma}^2_t$ is the forecast of the variance τ periods ahead given information at time t.