Introductory Financial Econometrics

Modelling Volatility

January 17, 2009

Table of contents

Fundamental Ideas

The standard deviation of the returns of financial assets, commonly referred to as the volatility, is a crucial aspect of much of modern finance theory, because it is a key input to areas such as portfolio construction, risk management and option pricing. In this session the particular nonlinear specifications in the variance investigated are motivated by the Autoregressive Conditional Heteroskedastity (ARCH) class of models introduced by Engle, which has become an important class of models for explaining the time series characteristics of asset returns. The models discussed here are univariate as they focus just on modelling the variance. A natural extension is to consider a multivariate model and specify both conditional variances as well as conditional covariances.

Volatility Clustering

- Plots of the returns of a financial asset show the random nature of returns. This is highlighted by the fact that the autocorrelation function of returns is typically flat.
- Closer inspection, however, shows that there are periods of time when returns hardly change (market tranquility) and others where changes in returns are followed by further large changes (market turbulence). This is a common characteristic of financial returns which is referred to as volatility clustering.
- Another way to highlight this volatility clustering property, is to look at the time series plot of the squares of returns $\left(y_t^2\right)$ where volatility clustering manifests itself in the squares of returns exhibiting autocorrelation.
- The effect of volatility clustering is also evident in the estimated unconditional probability distribution of returns. Typically the estimated empirical distribution has a sharper peak and fatter tails than the normal distribution.

A first model

Intuitively, the autocorrelation properties of the levels and squares of returns shows that whilst it is not possible to predict the direction of asset returns, it is possible to predict their volatility. These two properties suggest the following model

$$y_t = \rho y_{t-1} + u_t \quad (1)$$
$$y_t^2 = \gamma + \delta y_{t-1}^2 + v_t, \quad (2)$$

where $u_t \sim N\left(0, \sigma_u^2\right)$ and $v_t \sim N\left(0, \sigma_v^2\right)$ are two independent disturbances and

$$\begin{aligned} \rho &= 0 \\ \gamma, \delta &> 0. \end{aligned}$$

The restriction $\rho = 0$ ensures that y_t exhibits no autocorrelation, whereas the restriction $\delta > 0$ ensures that the squares of y_t exhibits (positive) autocorrelation.

Unconditional variance

With the restriction $\rho = 0$, the unconditional mean and variance of y_t are respectively

$$\begin{aligned} E\left[y_t\right] &= E\left[u_t\right] = 0 \\ E\left[y_t^2\right] &= E\left[u_t^2\right] = \sigma_u^2 \end{aligned}$$

An alternative expression of the unconditional variance of y_t is obtained by taking unconditional expectations

$$E\left[y_t^2\right] = E\left[\gamma + \delta y_{t-1}^2 + v_t\right] = \gamma + \delta E\left[y_{t-1}^2\right] + E\left[v_t\right],$$

and hence

$$E\left[y_t^2\right] = \gamma + \delta \sigma_u^2 + 0$$

or

$$\sigma_u^2 = \frac{\gamma}{1-\delta}.$$

Unconditional variance

Let the conditional expectation of y_t^2 based on information at time $t-1$ be defined as

$$h_t = E_{t-1}\left[y_t^2\right].$$

It follows that the conditional expectation is

$$h_t = E_{t-1}\left[\gamma + \delta y_{t-1}^2 + v_t\right] = \gamma + \delta E_{t-1}\left[y_{t-1}^2\right] + E_{t-1}\left[v_t\right] = \gamma + \delta y_{t-1}^2.$$

in which the variance of y_t is a nonlinear function of y_{t-1}.

Implications

- For small values of y_{t-1} the conditional variance is drawn from a relatively compact distribution with mean of zero and approximate variance γ. This means that there is a high probability of drawing another small value of y in the next period.
- For large values of y_{t-1} the conditional variance is drawn from a more dispersed distribution with mean zero, but approximate variance $\gamma + \delta y_{t-1}^2$. There is therefore a high probability of drawing another large value of y in the next period.
- The autocorrelation present in the variance results in a relatively small (large) variance tending to produce values of y_t relatively close (far) from the mean of zero, which, in turn, produces yet another small (large) variance thereby reinforcing the sequence.

ARCH(q)

The model is specified as

$$\begin{aligned} y_t &= u_t \\ u_t &\sim N(0, h_t) \\ h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_j u_{t-i}^2. \end{aligned} \tag{3}$$

This model is referred to as ARCH(q), where q refers to the order of the lagged squared returns included in the model. The conditional variance is given by

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \ldots + \alpha_q u_{t-q}^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \ldots + \alpha_q y_{t-q}^2,$$

where $\theta = (\alpha_0, \alpha_1, \cdots, \alpha_q)$ is a vector of parameters to be estimated.

ARCH(1)

A special case of this model is the ARCH(1) model given by

$$\begin{aligned} y_t &= u_t \\ u_t &\sim N(0, h_t) \\ h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 = \alpha_0 + \alpha_1 y_{t-1}^2 . \end{aligned} \tag{4}$$

Using the normality assumption, the ARCH(1) model is conveniently summarized by the conditional distribution of y_t

$$\begin{aligned} f(y_t | y_{t-1}) &= \frac{1}{\sqrt{2\pi h_t}} \exp\left[-\frac{y_t^2}{2h_t}\right] \\ &= \frac{1}{\sqrt{2\pi\left(\alpha_0 + \alpha_1 y_{t-1}^2\right)}} \exp\left[-\frac{y_t^2}{2\left(\alpha_0 + \alpha_1 y_{t-1}^2\right)}\right] . \end{aligned} \tag{5}$$

Estimation

This model can be estimated by maximum likelihood using a standard gradient algorithm. For a sample of $t = 1, 2, \cdots, T$ observations the log-likelihood is

$$\begin{aligned} \ln L &= \sum_{t=1}^{T} \ln L_t \\ &= \sum_{t=1}^{T} \ln f\left(y_t | y_{t-1}\right) + \ln f\left(y_0\right), \end{aligned}$$

The log-likelihood function at observation t is therefore

$$\begin{aligned} \ln L_t &= \ln f\left(y_t | y_{t-1}\right) \\ &= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(h_t\right) - \frac{1}{2}\frac{y_t^2}{h_t} \\ &= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\alpha_0 + \alpha_1 y_{t-1}^2\right) - \frac{1}{2}\frac{y_t^2}{\alpha_0 + \alpha_1 y_{t-1}^2}. \qquad (6) \end{aligned}$$

Two details

1. **Starting values**:
   The log likelihood at $t = 1$ needs a starting value for h_1 given by

   $$h_1 = \alpha_0 + \alpha_1 y_0^2,$$

   which in turn, requires a starting value for y_0. The simplest solution is to choose h_1 as the unconditional variance of y_t. Another approach is to compute h_1 as immediately above, by setting $y_0 = 0$ which is the unconditional mean of y_t for the model.
2. **Parameter restrictions**:
   - 2.1 To ensure that the variance is positive both α_0 and α_1 must be restricted to be positive.
   - 2.2 It is also usual to confine attention to cases in which the process generating the disturbances is variance stationary, that this, the unconditional variance of u_t is unchanging over time. This requires that $\alpha_1 < 1$.

Testing

The null and alternative hypotheses are thus

$$H_0: \quad \alpha_1 = 0 \quad \text{[No ARCH]}$$

$$H_1: \quad \alpha_1 \neq 0 \quad \text{[ARCH]}$$

Two tests

1. **Autocorrelation of Squared Returns**
   Autocorrelation in squared returns reflects the volatility clustering characteristically observed in returns and is evidence of ARCH. This can be tested by computing both r_j and the Ljung-Box statistic using squared returns. It is common to denote the Ljung-Box statistic when based on squared returns as $Q_{xx}(j)$.
2. **Formal ARCH test**
   A test of ARCH is given by testing that $\alpha_1 = 0$ in which case the model under the null hypothesis reduces a normal distribution with zero mean and constant variance

$$y_t \sim N\left(0, \sigma^2\right).$$

ARCH test

The test of ARCH can be performed by using a Lagange Multiplier test that involves estimating an auxiliary OLS regression.

Step 1 Regress y_t^2 on a constant and y_{t-1}^2

Step 2 Compute TR^2 from this regression and compare the computed value of the test statistic to the critical values obtained from the χ_1^2 distribution. A value of TR^2 in excess of the critical value is evidence of ARCH in y_t.

This test is available in EViews.

The GARCH Model

The ARCH(q) model has the property that the memory in the variance stops at the lag q. This means that for processes that exhibit long memory in the variance, it would be necessary to specify and estimate a high dimensional model. A natural way to circumvent this problem is to specify the conditional variance as a function of its own lags. The equation for the conditional variance then becomes

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}, \tag{7}$$

which is known as GARCH(p,q) where the p and the q identify the lags of the model and the "G" stands for Generalised ARCH.

GARCH(1,1)

$$\begin{aligned} y_t &= u_t \\ u_t &\sim N(0, h_t) \\ h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}. \end{aligned}$$

To highlight the long memory properties of this model, rewrite the expression for the conditional variance, h_t, using the lag operator $L^k y_t = y_{t-k}$ to yield

$$(1 - \beta_1 L) h_t = \alpha_0 + \alpha_1 y_{t-1}^2.$$

Assuming that $|\beta_1| < 1$, and using the properties of the lag operator, the conditional variance can be expressed as

$$\begin{aligned} h_t &= (1 - \beta_1 L)^{-1} \alpha_0 + \alpha_1 (1 - \beta_1 L)^{-1} y_{t-1}^2 \\ &= \frac{\alpha_0}{1 - \beta_1} + \alpha_1 y_{t-1}^2 + \alpha_1 \sum_{i=1}^{\infty} \beta_1^i y_{t-1-i}^2, \end{aligned} \tag{8}$$

which is instantly recognisable as an ARCH(∞) model.

Forecast error

Another way to highlight the memory characteristics of the GARCH conditional variance is to define the (forecast) error

$$v_t = y_t^2 - h_t,$$

which has the property $E_{t-1}[v_t] = 0$. Rearranging and substituting for h gives

$$\begin{aligned}
y_t^2 &= h_t + v_t \\
y_t^2 &= \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1} + v_t \\
y_t^2 &= \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \left(y_{t-1}^2 - v_{t-1}\right) + v_t \\
y_t^2 &= \alpha_0 + (\alpha_1 + \beta_1) y_{t-1}^2 - \beta_1 v_{t-1} + v_t,
\end{aligned}$$

which is an ARMA(1,1) model in terms of y_t^2. The memory of this process is determined by the autoregressive parameter $\alpha_1 + \beta_1$. The closer is $\alpha_1 + \beta_1$ to unity, the longer is the effect of a shock on volatility.

Long-run effects

The effect of a shock in the long-run on volatility is obtained from the unconditional variance of y_t, defined as $h = E\left[y_t^2\right]$. Taking unconditional expectations of (9) and using the result $E\left[v_t\right] = E\left[v_{t-1}\right] = 0$

$$\begin{aligned} E\left[y_t^2\right] &= E\left[\alpha_0 + (\alpha_1+\beta_1)\, y_{t-1}^2 - \beta_1 v_{t-1} + v_t,\right] \\ h &= \alpha_0 + (\alpha_1+\beta_1)\, y_{t-1}^2 h, \end{aligned}$$

upon rearranging gives an expression of the unconditional, or long-run, variance

$$h = \frac{\alpha_0}{1-\alpha_1-\beta_1}. \tag{9}$$

Estimation

As with the ARCH model, the GARCH model can be estimated by maximizing the log likelihood

$$
\begin{aligned}
\ln L_t &= \ln f\left(y_t | y_{t-1}\right) \\
&= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(h_t\right) - \frac{1}{2}\frac{y_t^2}{h_t} \\
&= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}\right) \\
&\quad -\frac{1}{2}\frac{y_t^2}{\alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}}.
\end{aligned}
$$

In estimating this model it may be necessary to restrict the parameters $\{\alpha_0, \alpha_1, \beta_1\}$ to be positive to ensure that the conditional variance is positive for all t.

Additional variables

A further extension of the ARCH class of model is to include additional variables in both the mean and the variance. An example, using the GARCH(1,1) conditional variance specification, is given by

$$
\begin{aligned}
y_t &= \gamma_0 + \gamma_1 x_t + u_t \\
u_t &\sim N(0, h_t) \\
h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + \lambda w_t,
\end{aligned}
$$

where x_t and w_t are additional explanatory variables which are important in explaining the mean and the variance respectively. In modelling the volatility of financial data, some examples of w_t are the volume of trades, dummy variables to allow for day-of-the-week effects and data on policy announcements.

Additional variables

An LM test of ARCH can be derived as before. The steps to test for ARCH of order q are as follows:

Step 1 Regress y_t on a constant and x_t and compute the least squares residuals $\widehat{u}_t$.

Step 2 Regress $\widehat{u}_t^2$ on a constant and the lagged squared residuals

$$\left\{\widehat{u}_{t-1}^2, \widehat{u}_{t-2}^2, \cdots, \widehat{u}_{t-q}^2\right\}.$$

Step 3 Compute TR^2 from the regression in Step 2 and compare this value to the critical value obtained from the χ_q^2 distribution. A value of TR^2 in excess of the critical value is evidence of ARCH in y_t.

Asymmetries

The GARCH model assumes that negative and positive shocks have the same impact on volatility, an assumption which ensures that the so-called News Impact Curve is symmetric. A natural extension of the GARCH model is to allow the effects of negative and positive shocks on the conditional variance to differ. This is especially important in modelling equity markets where negative shocks are expected to have a relatively bigger effect on volatility than positive shocks as a result of leverage effects.

The TARCH Model

In the Threshold GARCH model the variance is specified as

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} + \phi u_{t-1}^2 d_{t-1},$$

where d_{t-1} is a dummy variable given by

$$d_{t-1} = \begin{cases} 1: & u_{t-1} < 0 \\ 0: & u_{t-1} \geq 0 \end{cases} .$$

If $\phi > 0$, then good "news", as given by $u_{t-1} \geq 0$, has an effect on volatility equal to α_1 whereas bad "news", as given by $u_{t-1} < 0$, has an effect on volatility equal to $\alpha_1 + \phi$. If downward movements in the equity market are followed by relatively higher volatility, this implies that $\phi > 0$.

Specifically ...

These features of the model can be conveniently summarized for the case of $p = q = 1$ lags, as

$$\begin{array}{rcccccccc} h_t\,(good) & = & \alpha_0 & + & \alpha_1 & u_{t-1}^2 & + & \beta_1 & h_{t-1} \\ h_t\,(bad) & = & \alpha_0 & + & (\alpha_1 + \phi) & u_{t-1}^2 & + & \beta_1 & h_{t-1} \end{array} .$$

The effect of $\phi \neq 0$, is to make the News Impact Curve asymmetric. The TARCH model can be estimated using maximum likelihood methods, while a test of this model can be conducted using a Wald test of the hypothesis $\phi = 0$.

The E-GARCH Model

An alternative asymmetric model of the variance is the exponential GARCH model (EGARCH) model

$$\ln h_t = \alpha_0 + \alpha_1 \left| \frac{u_{t-1}}{h_{t-1}} \right| + \phi \frac{u_{t-1}}{h_{t-1}} + \beta_1 \ln h_{t-1}.$$

The asymmetry in the conditional variance is once again governed by the parameter ϕ. If $\phi < 0$ bad news increases volatility.
Note that the left-hand side is the log of the conditional variance. This implies

1. Forecasts of the conditional variance are guaranteed to be nonnegative. This makes estimation of the EGARCH model marginally more simple than the GARCH model, as the latter model may require imposing positivity restrictions on the parameters.
2. The leverage effect is exponential, rather than quadratic.

Model Specification

1. *ARCH test on standardized residuals*
   If the model is correctly specified then there should be no remaining evidence of ARCH in the residuals of the model. Apply the ARCH TR^2 test for various lags to the standardized residuals.
2. *Overfitting*
   Having estimated a GARCH model, overfit the model by extending the q and p lags. For each model compute the AIC, SIC and HIC and choose the model with the smallest value.

The problem

An important feature of the volatility models discussed so far is the role of conditional normality. Whilst the combination of conditional normality and GARCH conditional variances yields unconditional financial returns distributions which are leptokurtotic, fat-tails and a sharp peak relative to the normal distribution, in practice this class of models is not able to capture all of the leptokurtosis in the data. A problem with specifying the incorrect distribution as the basis of maximum likelihood estimation is that incorrect inference may result. Two possible solutions to this problem are discussed here.

Parametric solution

Consider the GARCH(p,q) model

$$\begin{aligned} y_t &= \sqrt{h_t} z_t \\ z_t &\sim St(0,1,\nu) \\ h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \end{aligned}$$

where $St(0,1,\nu)$ is the standardized Student t distribution

$$f(z_t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{z_t^2}{\nu-2}\right)^{-\left(\frac{\nu+1}{2}\right)}.$$

The log of the likelihood $\ln L = \sum_t \ln L_t$ is maximized with respect to the parameters in the conditional mean and conditional variance equations, plus the degrees of freedom parameter ν.
EViews provides the t-distribution as an alternative assumption for construction of the likelihood.

Correcting the Standard Errors

A simpler approach is simply to compute quasi-maximum likelihood (QMLE) or Bollerslev-Wooldridge standard errors, computed as the square root of the following matrix

$$H^{-1}\left(\widehat{\theta}\right) J\left(\widehat{\theta}\right) H^{-1}\left(\widehat{\theta}\right), \tag{10}$$

where

$$\begin{aligned} H &= \frac{\partial^2 \ln L}{\partial\theta\partial\theta'} \\ J &= G'G + \sum_{l=1}^{L} w_i \left(G'G_{-i} + (G'G_{-i})'\right) \\ G\left(\widehat{\theta}\right) &= \left[\frac{\partial \ln L_1}{\partial\theta}, \cdots \frac{\partial \ln L_T}{\partial\theta}\right]'_{\theta=\widehat{\theta}} \end{aligned}$$

and $w_i = 1 - \frac{i}{L+1}$ are the Newey-West weights. Notice that the point estimates of θ will be unchanged.

EViews provides QMLE standard errors as an option.

Attitudes to Risk

An important application of the ARCH class of volatility models is in modelling the trade-off between the expected return (μ) and risk (σ) of an asset. In general, the higher is the risk, the higher is the expected return that is needed to compensate the asset holder. However, the increase in the expected return required for an increase in risk varies across asset holders depending on their attitudes towards risk. Three categories of risk behaviour are usually identified:

1. Risk averter : μ increases at an increasing rate as σ increases.
2. Risk neutral : μ increases at a constant rate as σ increases.
3. Risk lover : μ increases at a decreasing rate as σ increases.

GARCH-M(1,1)

The ARCH model provides a natural and convenient way to model the tradeoff between expected return and risk by simply including the variance equation into the mean equation in terms of the GARCH-M model

$$\begin{aligned} y_t &= \mu_t + u_t \\ u_t &\sim N(0, h_t) \\ \mu_t &= \gamma_0 + \varphi h_t^{\rho} \\ h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}. \end{aligned}$$

The third equation gives the expected return (μ_t), and the fourth equation gives the conditional variance which is related to risk according to the convention in empirical finance $\sigma_t = \sqrt{h_t}$.

Testing

Thus the relationship between μ_t and σ_t is given by

$$\mu_t = \gamma_0 + \varphi \sigma_t^{2\rho},$$

so that the alternative types of risk aversion may be classified as

$$\begin{aligned} \text{Risk aversion} &: \quad \rho > 0.5 \\ \text{Risk neutral} &: \quad \rho = 0.5 \\ \text{Risk lover} &: \quad \rho < 0.5 \end{aligned} \tag{11}$$

The parameters

$$\theta = \{\gamma_0, \varphi, \rho, \alpha_0, \alpha_1 \beta_1\},$$

can be estimated by maximum likelihood methods as usual and hypothesis tests can be performed.