ARCH Models

Series de Tiempo

UTDT

2020

- Different ARCH models are proposed to capture the "Stylized Facts" of Asset returns.

Uses of Univariate ARCH models:

- Forecasting Volatility: Need to find a model to characterize the series which then is used to produce a forecast.
- Forecasting volatility might be useful when pricing Derivatives.

STRUCTURE OF THE LECTURE

- Definition and Properties.
- Estimation and testing
- Multivariate extensions

Empirical Regularities of Asset Returns

- Thick tails : returns tend to be leptokurtic
- Volatility clustering
- Leverage effects
- Non Trading Periods
- Forecastable Events

ARCH (1)

Consider the following model:

$$y_t = \mu + \varepsilon_t$$

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

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

ARCH (1)

- $E(\varepsilon_t|I_{t-1}) = E(v_t|I_{t-1})(\omega + \alpha\varepsilon_{t-1}^2)^{1/2} = 0$
  - Since $E(v_t|I_{t-1}) = E(v_t) = 0$
- $Var(\varepsilon_t|I_{t-1}) = E(\varepsilon_t^2|I_{t-1}) = E(v_t|I_{t-1})(\omega + \alpha\varepsilon_{t-1}^2) = (\omega + \alpha\varepsilon_{t-1}^2)$
  - Since $E(v_t|I_{t-1}) = E(v_t) = 1$
- Then we define the first order ARCH model as

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

- Notice that whenever the process is stationary applying iterative expectations we find that :

$$\begin{aligned} V(\varepsilon_t) &= E(\varepsilon_t^2) = E(\sigma_t^2) = E(\omega + \alpha\varepsilon_{t-1}^2) \\ E(\varepsilon_t^2) &= \omega/(1-\alpha) \end{aligned}$$

- An alternative representation
  We can always write the following expression

$$
\begin{aligned}
\varepsilon_t^2 &= E(\varepsilon_t^2 | I_{t-1}) + \eta_t, \text{ where } \eta_t \text{ is an innovation.} \\
& \quad \textit{or} \\
\varepsilon_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \eta_t
\end{aligned}
$$

then if $0 < \alpha < 1$, we find that

$$E(\varepsilon_t^2) = \omega / (1 - \alpha)$$

ARCH (1)

Consider Now the properties of $y_t = \mu + \varepsilon_t$ when there are ARCH(1) effects:

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

- Consider:

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

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

- $E(\varepsilon_t | I_{t-1}) = E(v_t | I_{t-1})(\omega + \alpha\varepsilon_{t-1}^2)^{\frac{1}{2}} = 0 \;\; (E(v_t | I_{t-1}) = E(v_t) = 0)$

First order autoregressive process with ARCH effects

- Conditional variance:

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

- Conditional mean and variance of y_t :

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

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

First order autoregressive process with ARCH effects

- Unconditional variance of y_t:

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

- Thus:

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

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

then

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

ARCH Models

Definition

An ARCH (q) model for the time varying conditional variance is

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

Remark A sufficient condition for the conditional variance to be positive is that the parameters of the model satisfy: $\omega > 0, (\forall i)\alpha_i \geq 0$

ARCH Models

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 $\sigma_t^2 = E(\varepsilon_t^2 | I_{t-1})$)

- AR(q) model for the squared innovations, ε_t^2 .
- Covariance stationary $\Leftrightarrow$ the sum of the positive autoregressive parameters is less than one, which gives:

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

- ε_t 's are clearly not independent through time.

GARCH Models

- ARCH(q): too many parameters,
  $\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \sum_{i=2}^{q} \alpha_i \varepsilon_{t-i}^2$
- Generalized ARCH or GARCH (1,1):

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

- Generalized ARCH or GARCH (p,q):

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

- A sufficient condition for the conditional variance in the model to be well defined is that all the coefficients in the infinite order linear ARCH are positive

GARCH Models

- Rearranging the GARCH(p,q) model by defining

$$\nu_t = \varepsilon_t - \sigma_t^2$$

it follows that

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

- ARMA(Max(p, q),p) model for ε_t^2

GARCH Models

- Covariance stationary $\Leftrightarrow$ all the coefficients of $(1 - \alpha(L) - \beta(L))$ lie outside the unit circle
- $\Leftrightarrow$ sum of the autoregressive coefficients is less than one
- Standard time series techniques in the identification of the orders of p and q

Persistence and stationarity

- GARCH (1,1)
- Assuming $\alpha + \beta < 1$, the unconditional variance of ε_{t+1} is:

$$\frac{\omega}{1 - (\alpha + \beta)}$$

- We can rewrite the GARCH (1,1) as:

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

Persistence and stationarity

- The conditional expectation of volatility j periods ahead is:

$$E_{t-1}[\sigma^2_{t+j}] = \frac{w}{(1-\alpha-\beta)} + (\sigma^2_t - \frac{w}{(1-\alpha-\beta)})(\alpha+\beta)^j$$

Since, using iterative expectations,
$E_{t-1}[\varepsilon^2_{t+j}] = E_{t-1}[E_{t+j-1}\varepsilon^2_{t+j}] = E_{t-1}[\sigma^2_{t+j}]$

IGARCH Models

Definition

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

$$E_t[\sigma^2_{t+j}] = \sigma^2_t + j\omega.$$

- Looks very much like a random walk with drift ω.
- The unconditional variance does not exist

EGARCH Models

- GARCH models not well suited to capture the "leverage effect"
- In the exponential GARCH (EGARCH) σ_t^2 depends on both the size and the sign of lagged residuals.

EGARCH Models

- EGARCH(1,1):

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

- Obviously the EGARCH model always produces a positive conditional variance σ_t^2 for any choice of α_0, β_1, γ_0
- Because of the use of both $|\varepsilon_t/\sigma_t|$ and (ε_t/σ_t), σ_t^2, it 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 $\varepsilon_{t-1} < 0$.

Additional Explanatory Variables:

It is straightforward to add other explanatory variables to a GARCH specification:

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

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 .
- *GARCH in Mean Models* allow for the conditional variance to have mean effects.
- Time varying risk premium.

GARCH-in-Mean Models

- Consider the following model:

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

and

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

ARCH-in-Mean Models

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

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

ARCH-in-Mean Models

Example

ARCH(1)-M

Consider a simple version of the above model:

$$y_t = \mu + \lambda\sigma_t^2 + \varepsilon_t$$

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

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

ARCH-in-Mean Models

Example

Then y_t may be expressed as

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

Then the expected value of y_t is

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

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

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

Lagrange Multiplier test

- Check there are ARCH effects in the residuals
- Lagrange Multiplier Test for ARCH :

  Under the null, the AR(p) process for y_t

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

  where ε_t is a Gaussian white noise process, $\varepsilon_t | I_{t-1}$ ~N(0,σ^2)

Lagrange Multiplier Test

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

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

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

Lagrange Multiplier Test

Caution! While this is the most widely used test we should be careful in interpreting the results. If the model is misspecified it is quite likely to reject the null

- Breaks in the variance will look as ARCH effects when using the whole sample
- In these cases, it is better to divide the sample and test for ARCH effects in the subperiods

Problem: How to divide the sample ex ante

Garch Effects and Sampling Frequency

- GARCH models do not temporaly aggregate
- Difficult to determine at which sampling frequency the data presents GARCH effects
- Empirical regularity: the higher is the sampling frequency, the higher the GARCH effects

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 ε_t ~N(0,σ_t^2), and 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

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

Maximum Likelihood Estimation with Gaussian Errors

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_0^2 and σ_0^2:

$$\sigma_0^2 = \omega / (1 - \alpha - \beta)$$

and

$$\varepsilon_0^2 = (\omega / (1 - \alpha - \beta))$$

Maximum Likelihood Estimation with Non-Gaussian Errors

- Some unconditional distributions seem to have fatter tails than the normal:
  - another distribution for ε_t : t-distribution.

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

  n represents the degrees of freedom (estimate)

How to compare between Different GARCH Specifications

- Most GARCH models are non nested
  - this makes comparison not straight forward

Misspecification tests on the standardized residuals

- ARCH(1):

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

- Test for the existence ARCH effects in the standardized residuals

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

Comparison between alternative models based on auxiliary regressions

- Auxiliary regression as a mean of choosing between different ARCH models:

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

- Regress the squared residuals on the fitted variance of the alternative GARCH models.
- Pagan and Scwhert (1989) propose to test the joint hypothesis:

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

Comparison between alternative models based on auxiliary regressions

- As a second step, they propose to compare the models that were not rejected on the basis of goodness of fit.
- The better the fit the better it mimics the conditional variance.
- Scale effects: logs

Measuring the accuracy of Forecasts of Different ARCH Models

- Forecasting ability of the different ARCH models as a way of comparing them
- Various *loss functions* have been proposed :
  - *Mean Squared Error*

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

  - *Mean Absolute Error.*

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

- *Mean Squared Error of the log of the squared residuals.*

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

- *Mean Absolute Error of the log of the squared residuals.*

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

Forecasting performance at different horizons

- *Mean Squared Error*

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

- *Mean Absolute Error.*

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

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

- *Mean Squared Error of the log of the squared residuals.*

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

- *Mean Absolute Error of the log of the squared residuals.*

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

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

Multivariate GARCH Models

- Multivariate extension of the GARCH(p,q) as follows:

$$y_t = \mu + u_t$$

$$vech(H_t) = C + \sum_{i=1}^{p} A_i vech(u_{t-i} u'_{t-i}) + \sum_{i=1}^{q} B_i vech(H_{t-i})$$

where $u_t | I_{t-1} \tilde{} N(0, H_t)$

Multivariate GARCH Models

- $H_t = E(u_t u_t' | I_{t-1})$:$n \times n$ conditional variance matrix associated with u_t'
- $vech(H_t)$: $(n(n+1))/2 \times 1$ vector of all the unique elements of H_t
    - μ is $n \times 1$
    - C is $(n(n+1))/2 \times 1$
    - $A_1, A_2, ...A_p, B_1, ...B_q$ are $(n(n+1))/2 \times (n(n+1))/2$

Multivariate GARCH Models

Example

A bivariate GARCH (1,1)

$$\begin{bmatrix} x_t \\ w_t \end{bmatrix} = \begin{bmatrix} \mu_x \\ \mu_w \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \nu_t \end{bmatrix}$$

Multivariate GARCH Models

Example

where

$$vech(H_t) = C + Avech(u_{t-1}u'_{t-1}) + Bvech(H_{t-1})$$

or

$$\begin{bmatrix} V_{t-1}(x_t) \\ COV_{t-1}(x_t w_t) \\ V_{t-1}(w_t) \end{bmatrix} = \begin{bmatrix} c_x \\ c_{xw} \\ c_w \end{bmatrix} + \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} (\varepsilon_{t-1})^2 \\ (\varepsilon_{t-1}\nu_{t-1}) \\ (\nu_{t-1})^2 \end{bmatrix}$$
$$+ \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} V_{t-2}(x_{t-1}) \\ COV_{t-2}(x_{t-1} w_{t-1}) \\ V_{t-2}(w_{t-1}) \end{bmatrix}$$

A Diagonal Vech Parametrization

- Assume that each covariance depends only on its own past values and innovations
- Each element of H_t follows an univariate GARCH model driven by the corresponding cross product $u_t u_t'$.

$$\begin{bmatrix} c_x & c_{xw} \\ c_{xw} & c_w \end{bmatrix}, \begin{bmatrix} a_1 & a_5 \\ a_5 & a_9 \end{bmatrix}, \begin{bmatrix} b_1 & b_5 \\ b_5 & b_9 \end{bmatrix}$$

are all positive definite.

- This can be ensured by doing simple Cholesky transformations to each of these matrices.

A Quadratic Specification

$$H_t = C'C + A\prime u_{t-1}u'_{t-1}A + B'H_{t-1}B.$$

C is a lower triangular with $n(n+1)/2$ parameters A and B are square matrices with n^2 parameters each

A Constant Correlation Specification

- Correlation between the assets is constant

$$COV_{t-1}(x_t w_t) = \rho\sqrt{V_{t-1}(x_t)V_{t-1}(w_t)}$$

- ρ is also estimated with the rest of the parameter set

Stationarity

- Conditions are similar to those discussed in the univariate case.
- The minimum square error forecast for $vech(H_t)$ ~$GARCH(1,1)$

$$E_s(vech(H_t)) = [\sum_{k=0}^{t-s-1}(A_1+B_1)^k]C + (A_1+B_1)^{t-s}vech(H_s)$$

- This can be derived in the usual way noting that

$$\begin{aligned} vech(H_t) &= C + Avech(u_{t-1}u'_{t-1}) + Bvech(H_{t-1}), \\ vech(u_t u'_t) &= vech(H_t) + vech(\eta_t) \\ vech(u_{t-1}u'_{t-1}) &= vech(H_{t-1}) + vech(\eta_{t-1}), \end{aligned}$$

writing the stochastic representation

$$vech(u_t u'_t) = C + (A+B)vech(u_{t-1}u'_{t-1}) - Bvech(\eta_{t-1}) + vech(\eta_t),$$

and doing the same derivation as in the univariate case.

Stationarity

- Let $V\Lambda V^{-1}$ denote the Jordan decomposition of the matrix $A_1 + B_1$, so that

$$(A_1 + B_1)^{t-s} = V\Lambda^{t-s}V^{-1}$$

- $E_s(vech(H_t))$ converges to the vech of the unconditional covariance matrix, $vech(\Sigma) = (I - A_1 - B_1)^{-1}C \Leftrightarrow$ the absolute value of the largest eigen value of $A_1 + B_1$ is strictly less than one.
- If the absolute value of the largest eigen value of $A_1 + B_1$ is strictly less than one, then the forecast can be written

$$E_s(vech(H_t)) = vech(\Sigma) + (A_1 + B_1)^{t-s}(vech(H_s) - vech(\Sigma))$$

Co-Persistence in Variance

- The empirical estimates for univariate and multivariate ARCH models often indicate a high degree of persistence in the forecast moments of the conditional variances,

$$E_s(H_t)_{ii}, i = 1, 2, ...N, \textit{for } t \rightarrow \infty.$$

- This persistence may be common across different series.

Definition

The multivariate ARCH process is said to be *co-persistent* in variance if at least one of the elements in $E_s(H_t)$ is non-convergent for increasing forecasts, $t - s$, but there exists a linear combination $\gamma' \varepsilon_t$, such that for every forecast origin s, the forecasts of the corresponding future conditional variances $E_s(\gamma' H_t \gamma)$ converge to a finite limit, independent of time s information.

Multivariate GARCH-M Models

Consider a system of n regression equations,

$$y_t = BX_t + Dvech(H_t) + u_t$$

$$vech(H_t) = C + \sum_{i=1}^{p} A_i vech(u_{t-i}u'_{t-i}) + \sum_{i=1}^{q} B_i vech(H_{t-i})$$

where

$$u_t | I_{t-1} \tilde{} N(0, H_t) B_{n\times k}, D_{n\times(n(n+1))/2}, C_{(n(n+1))/2\times 1},$$
$$A_1, A_2, ...A_p, B_1, ...B_q are (n(n+1))/2 \times (n(n+1))/2$$

A Bivariate GARCH-M (1,1) Model

$$\begin{bmatrix} x_t \\ w_t \end{bmatrix} = \begin{bmatrix} \mu_x \\ \mu_w \end{bmatrix} + \begin{bmatrix} d_1 & d_2 & d_3 \\ d_4 & d_5 & d_6 \end{bmatrix} \begin{bmatrix} V_{t-1}(x_t) \\ COV_{t-1}(x_t w_t) \\ V_{t-1}(w_t) \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \nu_t \end{bmatrix}$$

A Bivariate GARCH-M (1,1) Model

$$
\begin{aligned}
\begin{bmatrix} V_{t-1}(x_t) \\ COV_{t-1}(x_t w_t) \\ V_{t-1}(w_t) \end{bmatrix} &= \begin{bmatrix} c_x \\ c_{xw} \\ c_w \end{bmatrix} + \begin{bmatrix} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \\ c_4 & c_5 & c_6 \end{bmatrix} \begin{bmatrix} (\varepsilon_{t-1})^2 \\ (\varepsilon_{t-1} \nu_{t-1}) \\ (\nu_{t-1})^2 \end{bmatrix} \\
&+ \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_4 & b_5 & b_6 \end{bmatrix} \begin{bmatrix} V_{t-2}(x_{t-1}) \\ COV_{t-2}(x_{t-1} w_{t-1}) \\ V_{t-2}(w_{t-1}) \end{bmatrix}
\end{aligned}
$$

Estimation

- Conditional maximum likelihood estimation.

$$L_t(\theta) = -(n/2)\log(2\pi) - (1/2)\log(|H_t(\theta)|) - (1/2)u_t(\theta)' H_t^{-1}(\theta) u_t(\theta)$$

where θ represents a vector of parameters, n represents number of equations and t represents time.

Estimation

- Conditional on initial values for u_0 and H_0 , the likelihood function for the sample:

$$L(\theta) = \sum_{t=1}^{T} L_t(\theta)$$

- Model is highly non-linear and very unstable.

Testing the CAPM

- Problems:
  - The CAPM is a statement about the relationships between ex ante risk premiums and betas, both of which there are not directly observable.
  - The Roll- critique
  - The CAPM is a single-period model

The Unconditional CAPM

- In the simple version of the CAPM the expected return for asset j is

$$E(r_j) = r_f + \beta_{jm}(E(r_m) - r_f)$$

where r_j, r_f and r_m are the asset j, risk free and market rate of return respectively.

The Unconditional CAPM

Defining excess returns $\tilde{r}$, we can rewrite the above expressions as

$$E(\tilde{r}_j) = \beta_{jm}(E\tilde{r}_m)$$

where

$$\tilde{r}_j = r_j - r_f$$

$$\tilde{r}_m = r_m - r_f$$

The Conditional CAPM

- Models for r_{jt} in which the conditional density rather than the unconditional density returns is used.
- Conditional asset pricing model has

$$E(\tilde{r}_{jt}|I_{t-1}) = \beta_{jt}E(\tilde{r}_{mt}|I_{t-1})$$

where $\beta_{jt} = cov(\tilde{r}_{jt}\tilde{r}_{mt}|I_{t-1})/var(\tilde{r}_{mt}|I_{t-1})$.

Multivariate GARCH-M Models: A CAPM with time varying covariances

- The conditional CAPM can be written as

$$E(r_{jt}|I_{t-1}) - r_{ft-1} = \beta_{jt}[E(r_{mt}|I_{t-1}) - r_{ft-1}]$$

where

$$\beta_{jt} = cov(r_{jt}r_{mt}|I_{t-1})/var(r_{mt}|I_{t-1}).$$

and

$$H = \left[\begin{array}{cc} var(r_{jt}|I_{t-1}) & cov(r_{jt}r_{mt}|I_{t-1}) \\ cov(r_{jt}r_{mt}|I_{t-1}) & var(r_{mt}|I_{t-1}) \end{array} \right]$$

Multivariate GARCH-M Models: A CAPM with time varying covariances

- If we assume that the "market price of risk", λ is constant, where

$$\lambda = (E(r_{mt}|I_{t-1}) - r_{ft-1})/var(r_{mt}|I_{t-1})$$

- Then we can write

$$\begin{aligned} E(r_{jt}|I_{t-1}) - r_{ft-1} &= \beta_{jt}[E(r_{mt}|I_{t-1}) - r_{ft-1}] \\ &= \frac{cov(r_{jt}r_{mt}|I_{t-1})}{var(r_{mt}|I_{t-1})}[E(r_{mt}|I_{t-1}) - r_{ft-1}]. \\ &= \lambda cov(r_{jt}, r_{mt}|I_{t-1}) \end{aligned}$$

- And finally

$$r_{jt} = r_{ft-1} + \lambda cov(r_{jt}, r_{mt}|I_{t-1}) + u_{jt}$$

- Notice that

$$E(r_{mt}|I_{t-1}) - r_{ft-1} = \lambda var(r_{mt}|I_{t-1})$$

(since $\lambda = (E(r_{mt}|I_{t-1}) - r_{ft-1})/var(r_{mt}|I_{t-1})$)

$$r_{mt} = r_{ft-1} + \lambda var(r_{mt}|I_{t-1}) + u_{mt}$$

where u_{jt} and u_{mt} are the innovations.

- Time varying CAPM can be put into multivariate GARCH-M form as

$$y_t = b + d\,vech(H_t) + u_t$$

where $y_t = (r_{jt} - r_{ft-1}, r_{mt} - r_{ft-1})'$, $vech(H_t) = (var(r_{jt}|I_{t-1}),$ $cov(r_{jt}r_{mt}|I_{t-1}),\ var(r_{mt}|I_{t-1}))'$, $u_t = (u_{jt}, u_{mt})'$ and

$$d = \lambda \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- The zero restrictions implied by the theory may be tested by a likelihood ratio test which is asymptotically distributed $\chi^2(5)$.