# HS-302

## End Semester Assignment

Name : P.V. Sriram
Roll. : 1801CS37

1) Usually in Finance, the volatility / risk of as asset is represented as its conditional variance. The conventional econometric models assume a constant / slightly varying conditional variance.

But more often than not we can see that time series of an asset return exhibits periods of unusually large volatility followed by periods of relative tranquility. Under such conditions, the assumption of a constant variance of the disturbances in conventional econometric models, i.e, Homoscedasticity is inappropriate.

Therefore, it is necessary to employ models capable of adjusting with variance changes over time. Time varying volatility specifications are usually reffered to as the family of ARCH (Auto regressive Conditional Heteroskedasticity) models which can allow for the variance change,

# ARCH Models

Engle suggests the heteroskedastic of conditional variance can be formulated as a linear function of past squared errors. Let $\varepsilon_t$ denote the error terms ~~~~. These $\varepsilon_t$ are split into a stochastic piece $\nu_t$ and a time-dependent standard deviation $\sigma_t$.

$$\varepsilon_t = \nu_t \times z_t$$

where $z_t = \sqrt{\alpha_0 + \sum_{i=1}^{n} \alpha_i \varepsilon_{t-i}^2}$

where $\alpha_0 \geq 0$, $\alpha_i \geq 0$

and $\nu_t$ = white noise term $\left( E(\nu_t) = 0; \; var(\nu_t) = 1 \right)$

Verification

$$E(\varepsilon_t) = E\left(\nu_t \times \sqrt{\alpha_0 + \sum_{i=1}^{n} \alpha_i \varepsilon_{t-i}^2}\right) = 0$$

$=$ ~~E~~ Mean is 0

$$E(\varepsilon_t^2) = E\left(\nu_t^2 \left(\alpha_0 + \sum_{i=1}^{n} \alpha_i \varepsilon_{t-i}\right)\right)$$

$$= 1\left(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2\right)$$

$\therefore$ Increasing with previous values $\left(\varepsilon_{t-i}^2\right)$

## Fitting ARCH Models

(i) Estimate the best fitting AR(q) model.

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_q y_{t-q} + \varepsilon_t$$

$$\Rightarrow \quad y_t = a_0 + \sum_{i=1}^{q} a_i y_{t-i} + \varepsilon_t$$

(ii) Obtain the squares of the error $\hat{\varepsilon}^2$ and regress them on a const. and q lagged values:

$$\hat{\varepsilon}_t^2 = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \varepsilon_{t-i}^2$$

If $\alpha_1 = \alpha_2 = \cdots = 0$.

Then there is no ARCH effect

Past values aren't effecting the present errors significantly.

# GARCH Models

Bollerslev proposed the generalized ARCH or GARCH models, which specified the conditional variance to be a function of lagged squared errors and past conditional variance

GARCH (p, q) model (where p is the order of the GARCH terms $\sigma^2$ and q is the order of the ARCH terms $\varepsilon^2$), following is the model

$$\varepsilon_t = \vartheta_t \times \sqrt{h_t}$$

$$\text{where } h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

$$\vartheta_t = \text{white noise terms } \left( \begin{array}{l} E(\vartheta_t) = 0 \\ Var(\vartheta_t) = 1 \end{array} \right)$$

## Fitting GARCH

(i) Estimate the best fitting AR(q) model

$$y_t = a_0 + a_1 y_{t-1} + \cdots + a_q y_{t-q} + \varepsilon_t$$

$$= a_0 + \sum_{i=1}^{q} a_i y_{t-i} + \varepsilon_t$$

(ii) Compute and plot the autocorrelations of $\varepsilon^2$ by

$$\rho = \frac{\sum_{t=i+1}^{T} (\varepsilon_t^2 - \sigma_t^2)(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)}{\sum_{t=1}^{T} (\varepsilon_t^2 - \sigma_t^2)^2}$$

(iii) Standard deviation for large samples, of $P(i) = \frac{1}{\sqrt{T}}$.

Q statistic: $T(T+2) \sum_{i=1}^{n} P_i^2 / (T-i)$

If Q-statistic is significantly different from '0' then reject the null hypothesis.

$\downarrow$

$P_i \; \forall i$ are insignificant

## Threshold ARCH

The idea of Threshold ARCH (TARCH) models is to divide the distribution of the innovations into disjoint intervals and then approximate a piecewise linear function for the conditional standard deviation and conditional variance respectively. If there are only two intervals, the division is normally at zero.

The influence of positive and negative innovations on the volatility is differented.

$$r_t = \alpha + \beta r_{t-1} + \varepsilon_t$$

$$\varepsilon_t = v_t \times \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \lambda D \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

If $\varepsilon_{t-1} < 0$ then $D = 1$

$\varepsilon_{t-1} \geq 0$ then $D = 0$

$\varepsilon_{t-1} < 0$ (Bad news)

$$h_t = \alpha_0 + (\alpha_1 + \lambda)\,\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$\varepsilon_{t-1} \geq 0$ (Good news)

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

This is based on the notion of leverage effect where,

Tendency for volatility ↓ses when return ↑ses and vice versa

## 2) Cointegration

Cointegration is the presence of long run or multiple long run, stable equilibrium relationship between variables.

Let X and Y are two stocks (Both being stationary or non-stationary) then a linear combination of X and Y can generate a stationary variable

$$\Rightarrow \underbrace{\alpha X + \beta Y}_{\text{Non-Stationary}} = \underbrace{Z}_{\text{Stationary}}$$

If this condition takes place, then X and Y are cointegrated.

In order to analyze time series with classical methods like OLS, an assumption is made -

* The variances and means of the series are constants i.e) Time independent.

But, non stationary time series dont meet this assumption. So, the results from any hypothesis test will be biased or misleading.

So, to analyze such models, we use cointegration.

Tests for cointegration

a) Engle - Granger Test (2 step method)

b) Johansen test

Steps to test Cointegration

Let X and Y be two time series

a) Test the series X and Y for unit roots

b) Regress Y over X

i.e) $Y = A + BX + \varepsilon_y$

On running the above regression,

$Y = BX + A + \varepsilon \Rightarrow Y - BX = \varepsilon'$

c) Now, test these residuals $\varepsilon'$ for unit root

If we reject the null of a unit root in the residuals; then we cannot reject that the two variables cointegrate.

If $\varepsilon'$ is stationary, then X and Y are cointegrated. Here $[1, -\beta]$ is reffered as cointegrated vector

# Error Correction model

Error correction model (ECM) describes how dependent variable y and the independent variable (X) behave in short run consistent with a long run cointegrating relationship.

First, the long term relationship between the cointegrated variable is captured by regressing the values y and x.

Then the error terms of regression, together with other short term drivers is leveraged to correct for the short term trends, in turn aligning with the long term equilibrium.

=> ECM models make no ad hoc assumptions of how the variables change over time.

There are two types of ECM:

a) Two step ECM. (Engle and Granger procedure)

b) One step ECM

a) Two step ECM

Consider a simple model, where $y_t$ and $x_t$ are both I (1) i.e) Non-Stationary.

If there is a linear combination of $y_t$ and $x_t$ i.e) stationary then $x_t$ and $y_t$ are cointegrated.

The estimated residuals are stationary

$$\hat{u}_t = y_t - \hat{\beta}_1 - \hat{\beta}_2 x_t \quad (1)$$

Long run coefficient

$$y_t = \beta_1 + \beta_2 x_t + u_t$$

Step-1

The Engle and Granger Test is done by first running the cointegrating regression.

~~Step~~ If the residuals are stationary, it indicates there is a stationary cointegrating relationship

Step-2

Expressing the relationship between $y_t$ and $x_t$ with an ECM

$$\Delta y_t = \beta_3 + \beta_4 \Delta x_t - \pi_1 \hat{u}_{t-1} + \varepsilon_t$$

white noise

$\beta_3$ = const. term , $\beta_2$ = long run coefficient

$\beta_4$ = Short run coefficient, measure the immediate impact of a change in $x_t$ will have on a change in $y_t$

$\pi_1$ = Coefficient of estimated lagged residual of eq(1), shows the feedback effect, or the adjustment effect or error correction coefficient (how much of the disequilibrium is being corrected).

Now, from eq(1)

$$\Delta y_t = \beta_3 + \beta_4 \Delta x_t - \pi_1 (y_{t-1} - \hat{\beta}_1 - \hat{\beta}_2 x_{t-1}) + \varepsilon_t$$

Here, y can wander away from its long run path in the short run but will be pulled back to it over the longer term by the ECM.

b) One Step ECM model

(i) Generate new stationary series from original (Non-stationary) series by first differencing

(ii) Estimate the first difference of dependent variable followed by lagged 1 period dependent and independent variables (long-run) and the first differented independent variable (short run)

$$\Delta y_t = \alpha_0 + \alpha_1 \underbrace{\left(y_{t-1} + \frac{\alpha_2}{\alpha_1} x_{t-1}\right)}_{\text{long-run}} + \underbrace{\alpha_3 \Delta x_t}_{\text{Short Run}} + \varepsilon_t$$