Modeling Volatility

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
  - volatility clustering or volatility pooling
    - the tendency for volatility in financial markets to appear in bunches. Thus large returns are expected to follow large returns, and small returns to follow small returns.
  - leverage effects
    - the tendency for volatility to rise more following a large price fall than following a price rise of the same magnitude.
- Our "traditional" structural model could be something like:

$$y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + \varepsilon_t, \; \varepsilon_t \sim \mathrm{N}(0,\sigma^2).$$

Types of non-linear models

- There are many types of non-linear models, e.g.
  - ARCH / GARCH
  - switching models
- One particular non-linear model that has proved very useful in finance is the ARCH model due to Engle (1982, Econometrica).

Heteroscedasticity

- An example of a linear model is
  $y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \varepsilon_t$, $\varepsilon_t \sim \mathrm{N}(0,\sigma_\varepsilon^2)$.

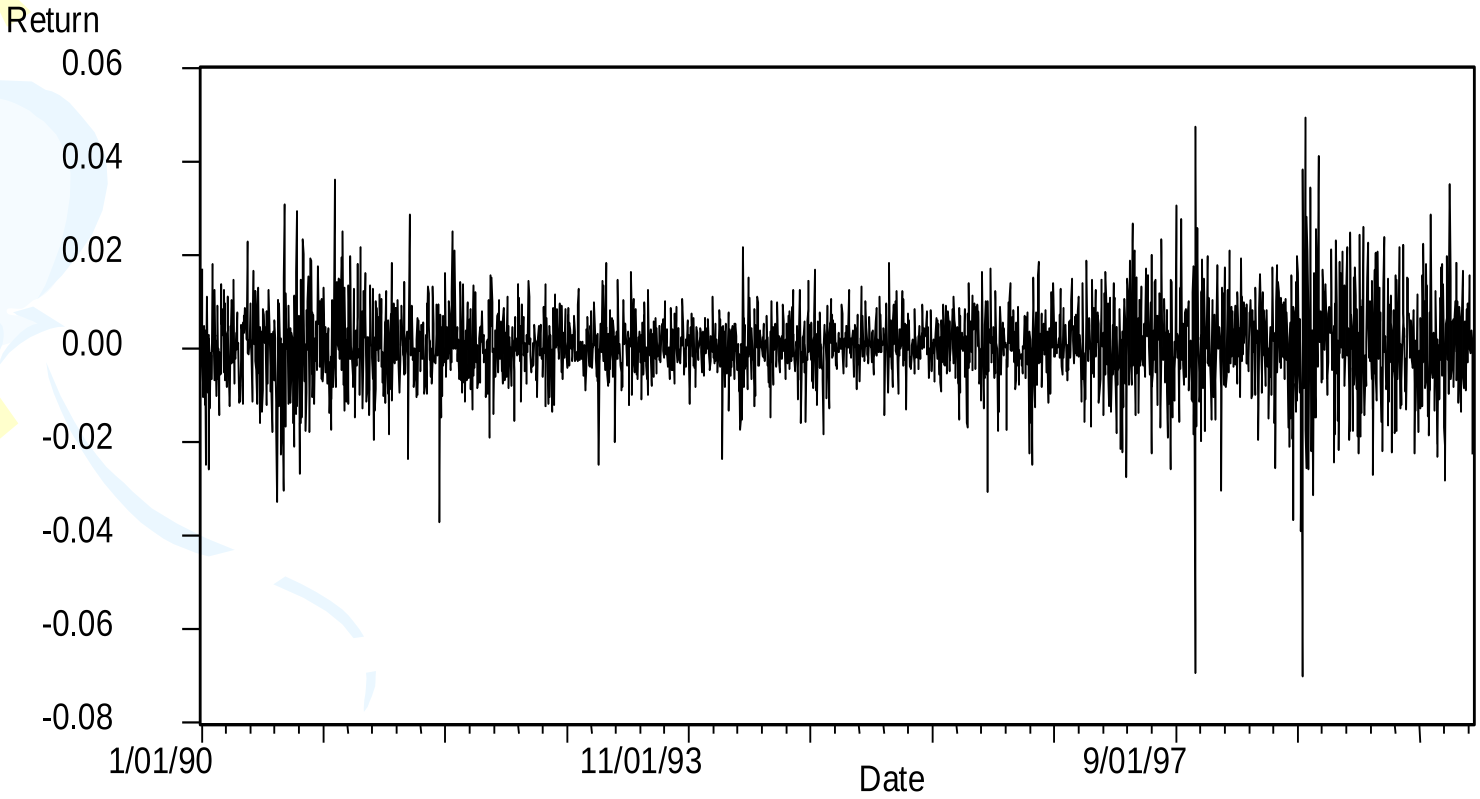
- The assumption that the variance of the errors is constant is known as <u>homoscedasticity</u>, i.e. $\mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2$.

- What if the variance of the errors is not constant?
  - heteroscedasticity
  - 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.

A Sample Financial Asset Returns Time Series

Daily S&P 500 Returns for January 1990 – December 1999



Autoregressive Conditional Heteroscedastic (ARCH) Models

- So, use a model which does not assume that the variance is constant.
- Definition of the variance of ε_t:

  $$\sigma_t^2 = \mathrm{Var}(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2},...) = \mathrm{E}[(\varepsilon_t\text{-}\mathrm{E}(\varepsilon_t))^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2},...]$$

  We usually assume that $\mathrm{E}(\varepsilon_t) = 0$

  so $\sigma_t^2 = \mathrm{Var}(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2},...) = \mathrm{E}[\varepsilon_t^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2},...]$.
- What could the current value of the variance of the errors plausibly depend upon?
  - Previous squared error terms.
- This leads to the ARCH model for the variance of the errors:

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

  - This is known as an ARCH(1) model.

Engle (1982, Econometrica) called this form of heteroscedasticity, where σ_t^2 depends on $\varepsilon_{t-1}{}^2$, $\varepsilon_{t-2}{}^2$, $\varepsilon_{t-3}{}^2$, etc. "autoregressive conditional heteroscedasticity (ARCH)". More formally, the model is

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + \varepsilon_t; \;\; \varepsilon_t \,|\psi_{t-1} \sim N\left(0,\; \sigma_t^2\right)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

where $\psi_{t-1} = \{y_{t-1}, x_{t-1}, x_{t-2}, \ldots\}$ represents the past realized values of the series.

This is an ARCH(q) model.

This equation is called an ARCH(q) model. We require that $\alpha_0 > 0$ and $\alpha_\mathrm{i} \geq 0$ to ensure that the conditional variance is positive. Stationarity of the series requires that

$$\sum_{i=1}^{q} \alpha_i < 1.$$

Weaknesses of ARCH Models

- Assumes that positive and negative shocks have the same effects on volatility.
- Gives no indication about what causes the variations of a financial time series.
- Rather restrictive.

Building an ARCH Model

- Model the mean effect and test for ARCH effects
- Specify the ARCH order and perform estimation
- Check the fitted ARCH model

Testing for ARCH effect

- Two tests are available
  - Ljung-Box test
    - The null hypothesis is that the first m lags of ACF of the e_t^2 series are zero.
  - Lagrange multiplier test

Lagrange multiplier test

H_0: $\alpha_1 = \alpha_2 = \dots \alpha_q = 0$

H_1: at least one $\alpha_i \neq 0$, $i = 1, \dots, q$

To conduct the test,

1. First, run any postulated linear regression $y_t = \beta_1 + \beta_2 x_{2t} + \dots + \beta_k x_{kt} + \varepsilon_t$ saving the residuals, e_t.

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

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_q e_{t-q}^2 + v_t, \qquad t = q+1, \dots, n,$$

where v_t is iid.
Obtain R^2 from this regression.

3. The LM statistic is defined as $LM = (n-q)R^2 \overset{a}{\sim} \chi_q^2$ under H_0.

Order Determination

- Use the PACF of ε_t^2 to determine the ARCH order
  - Define $u_t = \varepsilon_t^2 - \sigma_t^2$.
  - The ARCH model then becomes

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 + u_t,$$

  which is in the form of an AR(q) model for ε_t^2, except that $\{u_t\}$ is not an iid series.

In general, the ℓ-step ahead forecast is

$$\hat{\sigma}^2_{t+\ell} = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i \hat{\sigma}^2_{t+\ell-i}$$

$$\hat{\sigma}^2_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 e_t^2 + \hat{\alpha}_2 e_{t-1}^2 + \ldots + \hat{\alpha}_q e_{t+1-q}^2$$

$$\hat{\sigma}^2_{t+2} = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\sigma}^2_{t+1} + \hat{\alpha}_2 e_t^2 + \ldots + \hat{\alpha}_q e_{t+2-q}^2$$

Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

- ARCH model often requires many parameters to adequately describe the volatility process of an asset return.
- Using the concept of an ARMA process, Bollerslev (1986, Journal of Econometrics) proposed a useful extension known as the generalized ARCH (GARCH) model.

Specifically, a GARCH(p, q) model is defined as

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \varepsilon_t ; \quad \varepsilon_t \big| \psi_{t-1} \sim N\left(0, \sigma_t^2\right)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \gamma_j \sigma_{t-j}^2$$

with $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \ldots q$, $\gamma_j \geq 0$, $j = 1, \ldots p$ imposed to ensure that the conditional variances are positive. It is also required that $\sum_{i=1}^{\max(q,p)} (\alpha_i + \gamma_i) < 1$.

Usually, we only consider lower order GARCH processes such as GARCH (1, 1), GARCH (1, 2), GARCH (2, 1) and GARCH (2, 2) processes

For a GARCH (1, 1) process, for example, the forecasts are

$$\hat{\sigma}_{t+1}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 e_t^2 + \hat{\gamma}_1 \hat{\sigma}_t^2$$

$$\hat{\sigma}_{t+\ell}^2 = \hat{\alpha}_0 + \left(\hat{\alpha}_1 + \hat{\gamma}_1\right)\hat{\sigma}_{t+\ell-1}^2 \quad \text{for } \ell > 1$$

Unconditional Variance

- The unconditional variance of ε_t is given by

$$Var(\varepsilon_t) = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \gamma_j}$$

when $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \gamma_j < 1$

- $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \gamma_j \geq 1$ is termed "non-stationarity" in variance
- $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \gamma_j = 1$ is termed intergrated GARCH

- For stationarity in variance, the conditional variance forecasts will converge on their unconditional value as the horizon increases to infinity.

Model Checking

- For a properly specified (G)ARCH model, the standardized residuals

$$\hat{\eta}_t = e_t / \hat{\sigma}_t$$

form a sequence of iid random variables.

  – Check the adequacy of a fitted (G)ARCH model by conducting Ljung-Box test for $\hat{\eta}_t$ and $\hat{\eta}_t^2$.

- AIC, BIC