Time Series

Course Lecturer: Loukia Meligkotsidou
Department of Mathematics, University of Athens

MSc in Statistics and Operations Research

Time series models of heteroscedasticity

- Introduction
- Characteristics of financial/economic data
- Time series models of heteroscedasticity and basic properties
  - ARCH - Autoregressive conditional heteroscedastic models
  - GARCH - Generalised Autoregressive conditional heteroscedastic models
  - EGARCH - Exponential Generalised Autoregressive conditional heteroscedastic models
- Estimation of time-varying volatility models
- Forecasting time-varying volatility models

Introduction

- Introduce time series models of time-varying variance
- Uncertainty i.e. volatility is very crucial (theoretical and practical aspects)
- model building (due to the presence of heteroscedasticity and non-normality of the data)
- empirical financial/economic applications (portfolio allocation decisions, risk management, option pricing, asset pricing)

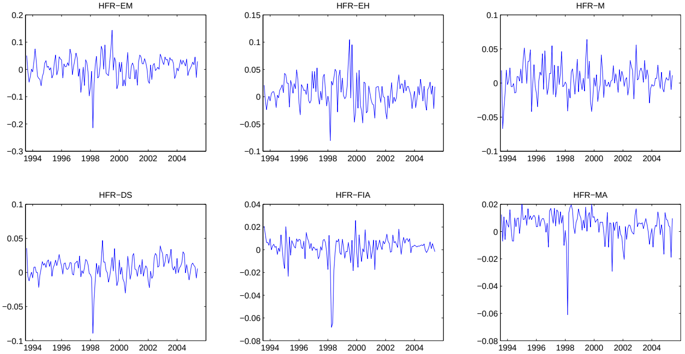
Characteristics of financial data

- volatility clustering (sub periods of high/low variability)
- non-normality, fat tails, excess kurtosis
- leverage effect
- co-movement in volatility changes across assets

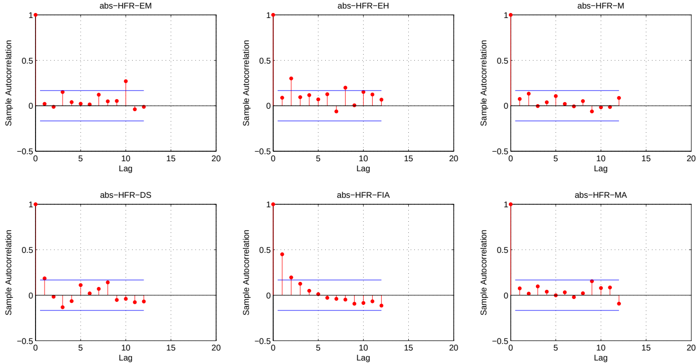
Descriptive statistics of hedge fund returns

Assets	*Mean*	*St.D.*	*Kurt*	$LB_{12}^{y_t}$	$LB_{12}^{\|y_t\|}$	$LB_{12}^{y_t^2}$	*JB*
EM	0.58%	4.2%	7.67	20.6	18.5	13.5	150.74*
EH	0.83%	2.6%	4.80	15.2	34.4*	36.2*	20.78*
M	0.50%	2.1%	3.85	9.0	7.7	6.5	4.33
DS	0.65%	1.6%	11.31	31.1*	16.3	9.1	466.34*
FIA	0.17%	1.1%	19.41	30.7*	43.9*	41.2*	1836.2*
MA	0.45%	1.0%	14.71	13.3	9.9	2.8	958.1*

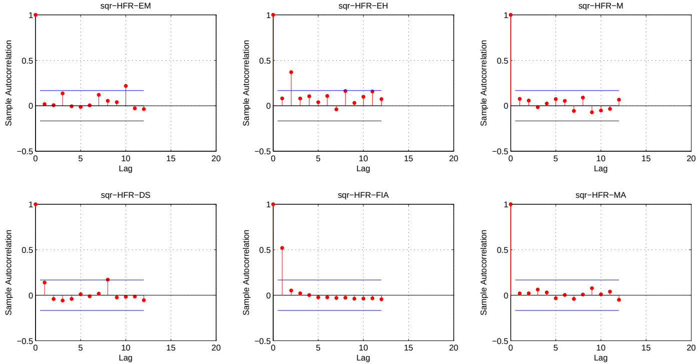
Time series plots - volatility clustering



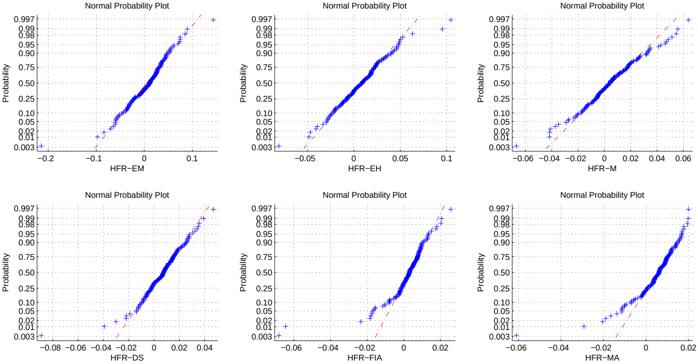
Autocorrelation plots of absolute returns



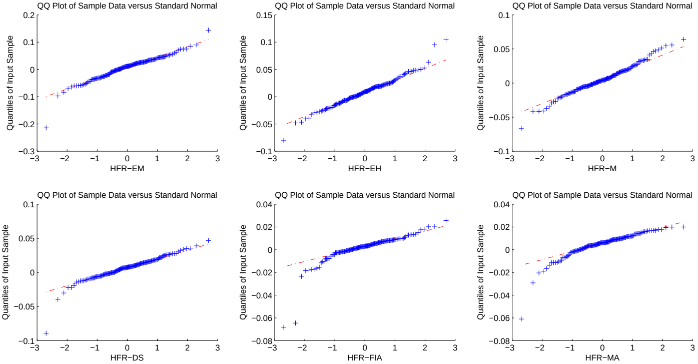
Autocorrelation plots of squared returns



Normal probability plots



Normal quantile plots



Time series models of heteroscedasticity

- Unconditional and conditional mean and variance
- ARCH models: Autoregressive conditional heteroscedastic models
- GARCH models: Generalized Autoregressive conditional heteroscedastic models
- EGARCH models: Exponential Generalized Autoregressive conditional heteroscedastic models
- Model Properties and characteristics

Unconditional and Conditional mean

Consider the AR(1) model: $y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d.(0, \sigma^2)$

- Unconditional mean: constant across time

  $E(y_t) = E(\delta + \phi_1 y_{t-1} + \varepsilon_t) = E(\delta) + E(\phi_1 y_{t-1}) + E(\varepsilon_t)$

  $\Rightarrow \mu = \delta + \phi_1 \mu \Rightarrow \mu(1 - \phi_1) = \delta \Rightarrow \mu = \frac{\delta}{1-\phi_1} = E(y_t)$

- Conditional mean: time-varying

  $E(y_t|\Phi_t) = E(\delta + \phi_1 y_{t-1} + \varepsilon_t|\Phi_t) =$

  $= E(\delta|\Phi_t) + E(\phi_1 y_{t-1}|\Phi_t) + E(\varepsilon_t|\Phi_t)$

  $\Rightarrow E(y_t|\Phi_t) = \delta + \phi_1 y_{t-1}$

Unconditional and Conditional Variance

Consider the AR(1) model: $y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d.(0, \sigma^2)$

- Unconditional variance: constant across time

$$V(y_t) = V(\delta + \phi_1 y_{t-1} + \varepsilon_t) = V(\delta) + V(\phi_1 y_{t-1}) + V(\varepsilon_t)$$

$$\Rightarrow v = \phi_1^2 v + \sigma^2 \Rightarrow v(1-\phi_1^2) = \sigma^2 \Rightarrow v = \frac{\sigma^2}{1-\phi_1^2} = V(y_t)$$

- Conditional variance: constant over time - to be modeled i.e. to be time-varying

$$V(y_t|\Phi_t) = V(\delta + \phi_1 y_{t-1} + \varepsilon_t|\Phi_t) =$$

$$= V(\delta|\Phi_t) + V(\phi_1 y_{t-1}|\Phi_t) + V(\varepsilon_t|\Phi_t)$$

$$\Rightarrow V(y_t|\Phi_t) = \sigma^2$$

Modeling conditional variance

At the conditional heteroscedasticity models presented below, we model the conditional variance at time t, σ_t^2

Study and model the conditional variance for different reasons:

- to understand the risk of a time series
- to achieve efficient estimates of a time series model
- to construct accurate confidence intervals for a forecast (i.e. time-varying)
- to capture the stylized facts i.e. the characteristics of a time series in empirical financial applications

Autoregressive Conditional Heteroscedasticity models [ARCH(p)]

The ARCH(p) model (Engle, 1982) can be written in the form:

Mean equation: $y_t = \gamma_0 + \gamma_1 x_{1,t} + \gamma_2 x_{2,t} + \ldots + \gamma_k x_{k,t} + \varepsilon_t$

Conditional distribution: $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$

Variance equation: $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_p \varepsilon_{t-p}^2$

where, $\alpha_0 > 0, \alpha_1, \ldots, \alpha_p \geq 0$ in order to be well defined the variance σ_t^2

The conditional variance depends on lagged squared errors

ARCH(1) model

The simple ARCH(1) model can be written:

Mean equation: $y_t = \varepsilon_t$

Conditional distribution: $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$

Variance equation: $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$, $\alpha_0 > 0, \alpha_1 \geq 0$

- the conditional variance depends only on the lagged one squared error, ε_{t-1}^2
- the ARCH(1) model captures the volatility clustering phenomenon
- the ARCH(1) model does not capture the leverage effect

ARCH(1) - AR(1) representation

The ARCH(1) model can be written as a non-Gaussian AR(1) model for the squared errors:

$\varepsilon_t^2 = \sigma_t^2 + (\varepsilon_t^2 - \sigma_t^2) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + v_t$, where $v_t = \varepsilon_t^2 - \sigma_t^2$

The conditional mean of v_t is

$$E(v_t|\Phi_{t-1}) = E(\varepsilon_t^2 - \sigma_t^2|\Phi_{t-1}) = E(\varepsilon_t^2|\Phi_{t-1}) - E(\sigma_t^2|\Phi_{t-1}) = \sigma_t^2 - \sigma_t^2 = 0$$

- the ARCH(1) model has significant partial autocorrelation of squared errors at lag 1
- the ARCH(p) model can be written as an AR(p) model for the squared errors
- the ARCH(p) model has significant the first p partial autocorrelations of the squared errors

ARCH(1) - kurtosis

Engle (1982) proved that the unconditional moments of an ARCH(1) process can be given by:

$E(\varepsilon_t^2) = \frac{\alpha_0}{1-\alpha_1}$ and $E(\varepsilon_t^4) = \frac{3\alpha_0^2}{(1-\alpha_1)^2}\frac{1-\alpha_1^2}{1-3\alpha_1^2}$, $\alpha_1 < 1$ and $3\alpha_1^2 < 1$

Then, the kurtosis is given by:

$$k = \frac{E(\varepsilon_t^4)}{[E(\varepsilon_t^2)]^2} = \frac{3\alpha_0^2}{(1-\alpha_1)^2}\frac{1-\alpha_1^2}{1-3\alpha_1^2} / \frac{\alpha_0^2}{(1-\alpha_1)^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2}$$

- the kurtosis is always larger than 3, i.e. larger than the kurtosis of a normal random variable
- the ARCH(1) model captures the fat tail characteristic of financial data
- similar arguments hold for the ARCH(p) model, which also produces kurtosis larger than 3

Generalised Autoregressive Conditional Heteroscedasticity models [GARCH(p,q)]

The GARCH(p,q) model (Bollerslev, 1986) can be written in the form:

Mean equation: $y_t = \gamma_0 + \gamma_1 x_{1,t} + \gamma_2 x_{2,t} + \ldots + \gamma_k x_{k,t} + \varepsilon_t$

Conditional distribution: $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$

Variance equation:
$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_q \sigma_{t-q}^2$

where, $\alpha_0 > 0, \alpha_1, \ldots, \alpha_p \geq 0, \beta_1, \ldots, \beta_q \geq 0$ in order to be well defined the variance σ_t^2

The conditional variance depends on lagged squared errors and on lagged variances

GARCH(1,1) model

The simple GARCH(1,1) model can be written:

Mean equation: $y_t = \varepsilon_t$

Conditional distribution: $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$

Variance equation: $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$, $\alpha_0 > 0, \alpha_1, \beta_1 \geq 0$

- the conditional variance depends only on the lagged one squared error, ε_{t-1}^2 and on the lagged one variance, σ_{t-1}^2
- the GARCH(1,1) model captures the volatility clustering phenomenon
- the GARCH(1,1) model does not capture the leverage effect

GARCH(1,1) - ARMA(1,1) representation

The GARCH(1,1) model can be written as a non-Gaussian ARMA(1,1) model for the squared errors:

$$\varepsilon_t^2 = \sigma_t^2 + (\varepsilon_t^2 - \sigma_t^2) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + v_t, \text{ where } v_t = \varepsilon_t^2 - \sigma_t^2$$

$$= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 (\varepsilon_{t-1}^2 - v_{t-1}) + v_t$$

$$= \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 - \beta_1 v_{t-1} + v_t$$

The conditional mean of v_t is

$$E(v_t|\Phi_{t-1}) = E(\varepsilon_t^2 - \sigma_t^2|\Phi_{t-1}) = E(\varepsilon_t^2|\Phi_{t-1}) - E(\sigma_t^2|\Phi_{t-1}) = \sigma_t^2 - \sigma_t^2 = 0$$

- the GARCH(1,1) model has significant autocorrelation and partial autocorrelation of squared errors at lag 1
- the GARCH(p,q) model can be identified through the autocorrelation and partial autocorrelation plot of squared residuals

GARCH(1,1) - kurtosis

Bollerslev (1986) proved that the unconditional moments of a GARCH(1,1) process are given by:

$E(\varepsilon_t^2) = \frac{\alpha_0}{1-\alpha_1-\beta_1}$ and $E(\varepsilon_t^4) = \frac{3\alpha_0^2(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)(1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2)}$

Then, the kurtosis is given by:

$$k = \frac{E(\varepsilon_t^4)}{[E(\varepsilon_t^2)]^2} = 3 + \frac{6\alpha_1^2}{1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2}$$

- the kurtosis is always larger than 3, i.e. larger than the kurtosis of a normal random variable
- the GARCH(1,1) model captures the fat tail characteristic of financial data
- similar arguments hold for the GARCH(p,q) model, which also produces kurtosis larger than 3

Exponential Generalised Autoregressive Conditional Heteroscedasticity models [EGARCH(p,q)]

The EGARCH(p,q) model (Nelson, 1991) can be written in the form:

Mean equation: $y_t = \gamma_0 + \gamma_1 x_{1,t} + \gamma_2 x_{2,t} + \ldots + \gamma_k x_{k,t} + \varepsilon_t$, $\varepsilon_t = z_t \sigma_t$

Conditional distribution: $z_t|\Phi_{t-1} \sim N(0,1)$ or $z_t|\Phi_{t-1} \sim GED(0,1)$

Variance equation:
$ln(\sigma_t^2) = \alpha_0 + \sum_{j=1}^{q} \beta_j ln(\sigma_{t-j}^2) + \sum_{i=1}^{p} [\theta_i z_{t-i} + \alpha_i(|z_{t-i}| - E\,|z_{t-i}|)]$

The logarithm of the conditional variance depends on lagged standardized errors, lagged absolute standardized errors and on lagged variances

EGARCH(1,1) model

The simple EGARCH(1,1) model can be written:

Mean equation: $y_t = \varepsilon_t$

Conditional distribution: $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$

Variance equation:
$ln(\sigma_t^2) = \alpha_0 + \beta_1 ln(\sigma_{t-1}^2) + \theta_1 z_{t-1} + \alpha_1(|z_{t-1}| - E\,|z_{t-1}|)$

- the logarithm of conditional variance depends only on the lagged one standardized error, $z_{t-1} = \frac{\varepsilon_{t-1}}{\sigma_{t-1}}$, lagged one absolute standardized error, and on the lagged one variance, σ_{t-1}^2
- the EGARCH(1,1) model captures the volatility clustering phenomenon
- the EGARCH(1,1) model captures the leverage effect

Maximum Likelihood Estimation: Regression-GARCH(1,1)

Consider a GARCH(1,1) model of the form:

Mean equation: $y_t = \gamma_0 + \gamma_1 x_{1,t} + \gamma_2 x_{2,t} + \ldots + \gamma_k x_{k,t} + \varepsilon_t$

Conditional distribution: $\varepsilon_t | \Phi_{t-1} \sim N(0, \sigma_t^2)$

Variance equation: $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$, $\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0$

Aim: Estimate the parameter vector $\theta = (\gamma_0, \gamma_1, \ldots, \gamma_k, \alpha_0, \alpha_1, \beta_1)$

Maximum Likelihood Estimation: Computing densities

To compute the conditional likelihood for the regression-GARCH(1,1) model, we condition on the initial values of errors and variances

$$y_1 = \gamma_0 + \gamma_1 x_{1,1} + \gamma_2 x_{2,1} + \ldots + \gamma_k x_{k,1} + \varepsilon_1$$

$$\Rightarrow \varepsilon_1 = y_1 - \gamma_0 - \gamma_1 x_{1,1} - \gamma_2 x_{2,1} - \ldots - \gamma_k x_{k,1}$$

$$y_1|\theta \sim N(\gamma_0 + \gamma_1 x_{1,1} + \gamma_2 x_{2,1} + \ldots + \gamma_k x_{k,1}, \sigma_1^2)$$

$$\sigma_1^2 = \alpha_0 + \alpha_1 \varepsilon_0^2 + \beta_1 \sigma_0^2$$

different alternatives for ε_0^2 and σ_0^2

The conditional density of the first observation is given by:

$$f(y_1|\theta) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[\frac{-(\varepsilon_1)^2}{2\sigma_1^2}\right]$$

Maximum Likelihood Estimation: Computing densities

At time t the density $f(y_t|\Phi_{t-1},\theta)$ is computed as follows:

$$y_t = \gamma_0 + \gamma_1 x_{1,t} + \gamma_2 x_{2,t} + \ldots + \gamma_k x_{k,t} + \varepsilon_t$$

$$\Rightarrow \varepsilon_t = y_t - \gamma_0 - \gamma_1 x_{1,t} - \gamma_2 x_{2,t} - \ldots - \gamma_k x_{k,t}$$

$$y_t|\Phi_{t-1},\theta \sim N(\gamma_0 + \gamma_1 x_{1,t} + \gamma_2 x_{2,t} + \ldots + \gamma_k x_{k,t}, \sigma_t^2)$$

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

The conditional density of $f(y_t|\Phi_{t-1},\theta)$ is given by:

$$f(y_t|\Phi_{t-1},\theta) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left[\frac{-(\varepsilon_t)^2}{2\sigma_t^2}\right]$$

Maximum Likelihood Estimation: likelihood

Therefore, the likelihood is computed by:

$$\text{Conditional Likelihood} = L(\theta|y,x) = f(y_1, y_2, \ldots, y_T|\theta) =$$

$$= f(y_T|\Phi_{T-1},\theta)\cdot f(y_{T-1}|\Phi_{T-2},\theta)\cdot\ldots f(y_2|\Phi_1,\theta)\cdot f(y_1|\theta) =$$

$$= \prod_{t=2}^{T} f(y_t|\Phi_{t-1},\theta)\cdot f(y_1|\theta) =$$

$$= \prod_{t=1}^{T}[\frac{1}{\sqrt{2\pi\sigma_t^2}}]\exp\left(-\frac{1}{2}\sum_{t=1}^{T}[\frac{\varepsilon_t^2}{\sigma_t^2}]\right)$$

$$= (2\pi)^{-T/2}\cdot\prod_{t=1}^{T}[(\alpha_0+\alpha_1\varepsilon_{t-1}^2+\beta_1\sigma_{t-1}^2)^{-1/2}]\cdot$$

$$\cdot\exp\left(-\frac{1}{2}\sum_{t=1}^{T}[\frac{(y_t-\gamma_0-\gamma_1 x_{1,t}-\ldots-\gamma_k x_{k,t})^2}{\alpha_0+\alpha_1\varepsilon_{t-1}^2+\beta_1\sigma_{t-1}^2}]\right)$$

Maximum Likelihood Estimation: log-likelihood

The log-likelihood for the regression GARCH(1,1) model is given by:

$$log[L(\theta|y,x)] = log[f(y_1|\theta)] + \sum_{t=2}^{T} log[f(y_t|\Phi_{t-1},\theta)] =$$

$$= -\frac{T}{2} log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} log(\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2)$$

$$-\frac{1}{2}\sum_{t=1}^{T}[\frac{(y_t-\gamma_0-\gamma_1 x_{1,t}-\ldots-\gamma_k x_{k,t})^2}{\alpha_0+\alpha_1\varepsilon_{t-1}^2+\beta_1\sigma_{t-1}^2}]$$

Diagnostic checking

After estimating an identified model, the residuals must be (resemble) a white noise process, i.e. must be:

- Uncorrelated
- Homoskedastic
- Normal distributed

Conduct diagnostic tests as in the case of regression-type and ARMA-type models

Forecasting ARCH(1) process

Suppose we are interested in forecasting the values of σ^2_{t+i}, $i = 1, \ldots, s$

Let $\hat{\sigma}^2_{t+i|t}$ denote the forecasts of σ^2_{t+i}

Consider an ARCH(1) model: $\sigma^2_t = \alpha_0 + \alpha_1 \varepsilon^2_{t-1}$

$$\hat{\sigma}^2_{t+1|t} = E(\alpha_0 + \alpha_1 \varepsilon^2_t | \Phi_t) = E(\alpha_0 | \Phi_t) + E(\alpha_1 \varepsilon^2_t | \Phi_t) = \alpha_0 + \alpha_1 \varepsilon^2_t$$

$$\hat{\sigma}^2_{t+2|t} = E(\alpha_0 + \alpha_1 \varepsilon^2_{t+1} | \Phi_t) = E(\alpha_0 | \Phi_t) + E(\alpha_1 \varepsilon^2_{t+1} | \Phi_t) = \alpha_0 + \alpha_1 \hat{\sigma}^2_{t+1|t}$$

$\ldots$

$$\hat{\sigma}^2_{t+s|t} = E(\alpha_0 + \alpha_1 \varepsilon^2_{t+s-1} | \Phi_t) = E(\alpha_0 | \Phi_t) + E(\alpha_1 \varepsilon^2_{t+s-1} | \Phi_t) =$$

$$= \alpha_0 + \alpha_1 \hat{\sigma}^2_{t+s-1|t}$$

Forecasting GARCH(1,1) process

Consider a GARCH(1,1) model: $\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2$

$$\hat{\sigma}_{t+1|t}^2 = E(\alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\sigma_t^2|\Phi_t) =$$

$$= E(\alpha_0|\Phi_t) + E(\alpha_1\varepsilon_t^2|\Phi_t) + E(\beta_1\sigma_t^2|\Phi_t) = \alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1\sigma_t^2$$

$$\hat{\sigma}_{t+2|t}^2 = E(\alpha_0 + \alpha_1\varepsilon_{t+1}^2 + \beta_1\sigma_{t+1}^2|\Phi_t) =$$

$$= E(\alpha_0|\Phi_t) + E(\alpha_1\varepsilon_{t+1}^2|\Phi_t) + E(\beta_1\sigma_{t+1}^2|\Phi_t) =$$

$$= \alpha_0 + \alpha_1\hat{\sigma}_{t+1|t}^2 + \beta_1\hat{\sigma}_{t+1|t}^2 = \alpha_0 + (\alpha_1 + \beta_1)\hat{\sigma}_{t+1|t}^2$$

$$\hat{\sigma}_{t+s|t}^2 = E(\alpha_0 + \alpha_1\varepsilon_{t+s-1}^2 + \beta_1\sigma_{t+s-1}^2|\Phi_t) =$$

$$= E(\alpha_0|\Phi_t) + E(\alpha_1\varepsilon_{t+s-1}^2|\Phi_t) + E(\beta_1\sigma_{t+s-1}^2|\Phi_t) =$$

$$= \alpha_0 + \alpha_1\hat{\sigma}_{t+s-1|t}^2 + \beta_1\hat{\sigma}_{t+s-1|t}^2 = \alpha_0 + (\alpha_1 + \beta_1)\hat{\sigma}_{t+s-1|t}^2$$

Application to financial and economic series

- Example 1: GARCH modeling of the Intel stock returns
- Example 2: GARCH modeling of the S&P500 index
- Example 3: Regression - ARMA - GARCH modeling of hedge fund returns
- Discussion on financial empirical applications i.e. performance evaluation, predictability, value at risk