Time Series Analysis

Caston Sigauke (STA 5242)

Lecture H8

Volatility Forecasting Models

Volatility Forecasting Models

Volatility Forecasting Models

Volatility Forecasting Models

Volatility Forecasting Models

Volatility Forecasting Models

Some stylized financial facts

1. Time series of share prices X_t and other basic financial instruments are not stationary and possess a local trend at least.

Some stylized financial facts

1. Time series of share prices X_t and other basic financial instruments are not stationary and possess a local trend at least.
2. Returns r_t have a leptokurtic distribution. The empirically estimated kurtosis is mostly greater than 3.

Some stylized financial facts

1. Time series of share prices X_t and other basic financial instruments are not stationary and possess a local trend at least.
2. Returns r_t have a leptokurtic distribution. The empirically estimated kurtosis is mostly greater than 3.
3. Leverage effect. Price movements are negatively correlated with volatility. This means that the volatility of stock tends to increase when the prices drops.

Some stylized financial facts

1. Time series of share prices X_t and other basic financial instruments are not stationary and possess a local trend at least.
2. Returns r_t have a leptokurtic distribution. The empirically estimated kurtosis is mostly greater than 3.
3. Leverage effect. Price movements are negatively correlated with volatility. This means that the volatility of stock tends to increase when the prices drops.
4. Volatility clustering: Large movements in returns tend to be followed by further large movements. Likewise for small movements.

Some volatility forecasting models

1. Historical volatility models: Historical average model, simple moving average, exponential smoothing, and exponential weighted moving average.

Some volatility forecasting models

1. Historical volatility models: Historical average model, simple moving average, exponential smoothing, and exponential weighted moving average.
2. Implied volatility models: is the market's forecast of a likely movement in a security's price. It is often used to price options contracts: High implied volatility results in options with higher premiums and vice versa.

Some volatility forecasting models

1. Historical volatility models: Historical average model, simple moving average, exponential smoothing, and exponential weighted moving average.
2. Implied volatility models: is the market's forecast of a likely movement in a security's price. It is often used to price options contracts: High implied volatility results in options with higher premiums and vice versa.
3. Autoregressive and Heteroskedastic models: Autoregressive Conditional Heteroskedasticity (ARCH), Generalized ARCH (GARCH).

Some volatility forecasting models

1. Historical volatility models: Historical average model, simple moving average, exponential smoothing, and exponential weighted moving average.
2. Implied volatility models: is the market's forecast of a likely movement in a security's price. It is often used to price options contracts: High implied volatility results in options with higher premiums and vice versa.
3. Autoregressive and Heteroskedastic models: Autoregressive Conditional Heteroskedasticity (ARCH), Generalized ARCH (GARCH).
4. Models based on artificial neural networks.

Comparison of the volatility forecasting models

Historical volatility models

Advantages: Easy to use, relatively good results (except historical average).

Comparison of the volatility forecasting models

Historical volatility models

Advantages: Easy to use, relatively good results (except historical average).
Disadvantages: after sudden price movements these models tend to significantly overestimate the volatility.

Comparison of the volatility forecasting models

Historical volatility models

Advantages: Easy to use, relatively good results (except historical average).
Disadvantages: after sudden price movements these models tend to significantly overestimate the volatility.

Implied volatility

Advantages: Good results, these model are based on strong theoretical results.

Comparison of the volatility forecasting models

Historical volatility models

Advantages: Easy to use, relatively good results (except historical average).
Disadvantages: after sudden price movements these models tend to significantly overestimate the volatility.

Implied volatility

Advantages: Good results, these model are based on strong theoretical results.
Disadvantages: complex to implement; not suitable for all products (only for that, which are an underlying for an actively trading options); are dependent on sometimes irrational expectations of investors.

Comparison of the volatility forecasting models

Historical volatility models

Advantages: Easy to use, relatively good results (except historical average).
Disadvantages: after sudden price movements these models tend to significantly overestimate the volatility.

Implied volatility

Advantages: Good results, these model are based on strong theoretical results.
Disadvantages: complex to implement; not suitable for all products (only for that, which are an underlying for an actively trading options); are dependent on sometimes irrational expectations of investors.

Comparison of the volatility forecasting models

Autoregressive and Heteroskedastic models

Advantages: an existence of a big number of theoretical researches of these models.

Comparison of the volatility forecasting models

Autoregressive and Heteroskedastic models

Advantages: an existence of a big number of theoretical researches of these models.
Disadvantages: moderate forecasting accuracy; complex to implement and optimize.

Comparison of the volatility forecasting models

Autoregressive and Heteroskedastic models

Advantages: an existence of a big number of theoretical researches of these models.
Disadvantages: moderate forecasting accuracy; complex to implement and optimize.

Models based on Neural Networks

Advantages: relatively good results; an ability to build models which use not only historical returns as an input, but also other related financial time series and variables.

Comparison of the volatility forecasting models

Autoregressive and Heteroskedastic models

Advantages: an existence of a big number of theoretical researches of these models.
Disadvantages: moderate forecasting accuracy; complex to implement and optimize.

Models based on Neural Networks

Advantages: relatively good results; an ability to build models which use not only historical returns as an input, but also other related financial time series and variables.
Disadvantages: complex to implement and to find suitable architecture; can forecast only dependencies from previous observations.

ARCH and GARCH Models

Let X_t be a discrete stochastic process, then the return of the process X_t is defined as:

$$r_t = \nabla \log X_t = (1-B)\log X_t = \log \frac{X_t}{X_{t-1}}$$

ARCH and GARCH Models

Let X_t be a discrete stochastic process, then the return of the process X_t is defined as:

$$r_t = \nabla \log X_t = (1 - B) \log X_t = \log \frac{X_t}{X_{t-1}}$$

$$r_t = \log \frac{X_t}{X_{t-1}} \times 100$$

ARCH and GARCH Models

Let X_t be a discrete stochastic process, then the return of the process X_t is defined as:

$$r_t = \nabla \log X_t = (1 - B) \log X_t = \log \frac{X_t}{X_{t-1}}$$

$$r_t = \log \frac{X_t}{X_{t-1}} \times 100$$

A general model is (this is the mean equation):

ARCH and GARCH Models

Let X_t be a discrete stochastic process, then the return of the process X_t is defined as:

$$r_t = \nabla \log X_t = (1-B)\log X_t = \log \frac{X_t}{X_{t-1}}$$

$$r_t = \log \frac{X_t}{X_{t-1}} \times 100$$

A general model is (this is the mean equation):

$$r_t = \mu_t + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2), E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$$

ARCH and GARCH Models

Let X_t be a discrete stochastic process, then the return of the process X_t is defined as:

$$r_t = \nabla \log X_t = (1-B)\log X_t = \log \frac{X_t}{X_{t-1}}$$

$$r_t = \log \frac{X_t}{X_{t-1}} \times 100$$

A general model is (this is the mean equation):

$$r_t = \mu_t + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2), E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$$

$$\mu_t = \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

ARCH Models

ARCH(1): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(1), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

ARCH Models

ARCH(1): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(1), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

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

with $\omega \geq 0, 0 \leq \alpha < 1$ and

ARCH Models

ARCH(1): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(1), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

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

with $\omega \geq 0, 0 \leq \alpha < 1$ and

- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ and $Z_t = \frac{\varepsilon_t}{\sigma_t}$ is i.i.d. (strong ARCH).

ARCH Models

ARCH(1): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(1), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

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

with $\omega \geq 0, 0 \leq \alpha < 1$ and

- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ and $Z_t = \frac{\varepsilon_t}{\sigma_t}$ is i.i.d. (strong ARCH).
- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ (semi-strong ARCH).

ARCH Models

ARCH(1): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(1), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

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

with $\omega \geq 0, 0 \leq \alpha < 1$ and

- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ and $Z_t = \frac{\varepsilon_t}{\sigma_t}$ is i.i.d. (strong ARCH).
- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ (semi-strong ARCH).

Setting $Z_t = \frac{\varepsilon_t}{\sigma_t}$, it holds for the semi-strong and the strong ARCH models that $E(Z_t) = 0$ and $\text{Var}(Z_t) = 1$. It is frequently assumed that Z_t is normally distributed, which means ε_t is conditionally normally distributed:

ARCH Models

ARCH(1): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(1), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

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

with $\omega \geq 0, 0 \leq \alpha < 1$ and

- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ and $Z_t = \frac{\varepsilon_t}{\sigma_t}$ is i.i.d. (strong ARCH).
- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ (semi-strong ARCH).

Setting $Z_t = \frac{\varepsilon_t}{\sigma_t}$, it holds for the semi-strong and the strong ARCH models that $E(Z_t) = 0$ and $\text{Var}(Z_t) = 1$. It is frequently assumed that Z_t is normally distributed, which means ε_t is conditionally normally distributed:

$$\varepsilon_t|\mathcal{F}_{t-1} \sim N(0, \sigma_t^2)$$

Theorem: Unconditional variance of the ARCH(1))

Assume the process ε_t is a semi-strong ARCH(1) process with $\text{Var}(\varepsilon_t) = \sigma^2 < \infty$. Then it holds that

$$\sigma^2 = \frac{\omega}{1 - \alpha}$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$Z_t = \frac{\varepsilon_t}{\sigma_t} \implies \varepsilon_t = Z_t \sigma_t$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$Z_t = \frac{\varepsilon_t}{\sigma_t} \implies \varepsilon_t = Z_t \sigma_t$$

$$\implies E(\varepsilon_t^2) = E(Z_t^2 \sigma_t^2) \implies E(\varepsilon_t^2) = E(Z_t^2)E(\sigma_t^2) = E(\sigma_t^2)$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$Z_t = \frac{\varepsilon_t}{\sigma_t} \implies \varepsilon_t = Z_t \sigma_t$$

$$\implies E(\varepsilon_t^2) = E(Z_t^2 \sigma_t^2) \implies E(\varepsilon_t^2) = E(Z_t^2)E(\sigma_t^2) = E(\sigma_t^2)$$

$$\text{But } \mathsf{Var}(\varepsilon_t) = E(\varepsilon_t^2) - (E(\varepsilon_t))^2 = E(\varepsilon_t^2)$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$Z_t = \frac{\varepsilon_t}{\sigma_t} \implies \varepsilon_t = Z_t \sigma_t$$

$$\implies E(\varepsilon_t^2) = E(Z_t^2 \sigma_t^2) \implies E(\varepsilon_t^2) = E(Z_t^2)E(\sigma_t^2) = E(\sigma_t^2)$$

But $\text{Var}(\varepsilon_t) = E(\varepsilon_t^2) - (E(\varepsilon_t))^2 = E(\varepsilon_t^2)$

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

Proof of Theorem: Unconditional variance of the ARCH(1))

$$Z_t = \frac{\varepsilon_t}{\sigma_t} \implies \varepsilon_t = Z_t \sigma_t$$

$$\implies E(\varepsilon_t^2) = E(Z_t^2 \sigma_t^2) \implies E(\varepsilon_t^2) = E(Z_t^2)E(\sigma_t^2) = E(\sigma_t^2)$$

But $\text{Var}(\varepsilon_t) = E(\varepsilon_t^2) - (E(\varepsilon_t))^2 = E(\varepsilon_t^2)$

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

$$\sigma^2 = \omega + \alpha\sigma^2 \implies \sigma^2 = \frac{\omega}{1-\alpha}$$

Unconditional distribution of ε_t is leptokurtic

Let ε_t be a strong ARCH(1) process, $Z_t \sim N(0,1)$ and $E[\varepsilon_t^4] = c < \infty$ then

1. $$E[\varepsilon_t^4] = \frac{3\omega^2}{(1-\alpha)^2} \frac{1-\alpha^2}{1-3\alpha^2}$$
   with $3\alpha^2 < 1$.

2. the unconditional distribution of ε_t is leptokurtic

Proof of Theorem: Unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(X) = \frac{E(X-\mu)^4}{\left(E(X-\mu)^2\right)^2} = \frac{E(X-\mu)^4}{\sigma^4}$$

Proof of Theorem: Unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(X) = \frac{E(X-\mu)^4}{\left(E(X-\mu)^2\right)^2} = \frac{E(X-\mu)^4}{\sigma^4}$$

$$\varepsilon_t = Z_t \sigma_t \implies E(\varepsilon_t^2)^2 = E(Z_t^2 \sigma_t^2)^2$$

Proof of Theorem: Unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(X) = \frac{E(X-\mu)^4}{\left(E(X-\mu)^2\right)^2} = \frac{E(X-\mu)^4}{\sigma^4}$$

$$\varepsilon_t = Z_t\sigma_t \implies E(\varepsilon_t^2)^2 = E(Z_t^2\sigma_t^2)^2$$

$$\text{Now } E(\varepsilon_t^4) = E(Z_t^4)E\left[(\sigma_t^2)^2\right]$$

Proof of Theorem: Unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(X) = \frac{E(X-\mu)^4}{(E(X-\mu)^2)^2} = \frac{E(X-\mu)^4}{\sigma^4}$$

$$\varepsilon_t = Z_t\sigma_t \implies E(\varepsilon_t^2)^2 = E(Z_t^2\sigma_t^2)^2$$

$$\text{Now } E(\varepsilon_t^4) = E(Z_t^4)E\left[(\sigma_t^2)^2\right]$$

But $E(Z_t^4) = 3$ since the kurtosis of a normal distribution is 3.
Note $\text{Kurt}(Z_t) = \frac{E(Z_t-\mu)^4}{(\sigma^2)^2} = E(Z_t^4)$

Proof of Theorem: Unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(X) = \frac{E(X-\mu)^4}{\left(E(X-\mu)^2\right)^2} = \frac{E(X-\mu)^4}{\sigma^4}$$

$$\varepsilon_t = Z_t\sigma_t \implies E(\varepsilon_t^2)^2 = E(Z_t^2\sigma_t^2)^2$$

$$\text{Now } E(\varepsilon_t^4) = E(Z_t^4)E\left[(\sigma_t^2)^2\right]$$

But $E(Z_t^4) = 3$ since the kurtosis of a normal distribution is 3.
Note $\text{Kurt}(Z_t) = \frac{E(Z_t-\mu)^4}{(\sigma^2)^2} = E(Z_t^4)$

$$\implies E(\varepsilon_t^4) = 3E\left[(\omega + \alpha\varepsilon_{t-1}^2)^2\right]$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$E(\varepsilon_t^4) = 3E\left(\omega^2 + 2\omega\alpha\varepsilon_{t-1}^2 + \alpha^2\varepsilon_{t-1}^4\right)$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$E(\varepsilon_t^4) = 3E\left(\omega^2 + 2\omega\alpha\varepsilon_{t-1}^2 + \alpha^2\varepsilon_{t-1}^4\right)$$

$$E(\varepsilon_t^4) = 3\left[\omega^2 + 2\omega\alpha E(\varepsilon_{t-1}^2) + \alpha^2 E(\varepsilon_{t-1}^4)\right]$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$E(\varepsilon_t^4) = 3E\left(\omega^2 + 2\omega\alpha\varepsilon_{t-1}^2 + \alpha^2\varepsilon_{t-1}^4\right)$$

$$E(\varepsilon_t^4) = 3\left[\omega^2 + 2\omega\alpha E(\varepsilon_{t-1}^2) + \alpha^2 E(\varepsilon_{t-1}^4)\right]$$

$$E(\varepsilon_t^4) = 3\left[\omega^2 + 2\omega\alpha\frac{\omega}{1-\alpha} + \alpha^2 c\right]$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$E(\varepsilon_t^4) = 3E\left(\omega^2 + 2\omega\alpha\varepsilon_{t-1}^2 + \alpha^2\varepsilon_{t-1}^4\right)$$

$$E(\varepsilon_t^4) = 3\left[\omega^2 + 2\omega\alpha E(\varepsilon_{t-1}^2) + \alpha^2 E(\varepsilon_{t-1}^4)\right]$$

$$E(\varepsilon_t^4) = 3\left[\omega^2 + 2\omega\alpha\frac{\omega}{1-\alpha} + \alpha^2 c\right]$$

$$\implies 3\left[\omega^2 + 2\omega\alpha\frac{\omega}{1-\alpha} + \alpha^2 c\right] = c, \text{since } E[\varepsilon_t^4] = c < \infty$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$E(\varepsilon_t^4) = 3E\left(\omega^2 + 2\omega\alpha\varepsilon_{t-1}^2 + \alpha^2\varepsilon_{t-1}^4\right)$$

$$E(\varepsilon_t^4) = 3\left[\omega^2 + 2\omega\alpha E(\varepsilon_{t-1}^2) + \alpha^2 E(\varepsilon_{t-1}^4)\right]$$

$$E(\varepsilon_t^4) = 3\left[\omega^2 + 2\omega\alpha\frac{\omega}{1-\alpha} + \alpha^2 c\right]$$

$$\implies 3\left[\omega^2 + 2\omega\alpha\frac{\omega}{1-\alpha} + \alpha^2 c\right] = c, \text{since } E[\varepsilon_t^4] = c < \infty$$

$$c = \frac{3(\omega^2(1-\alpha) + 2\omega^2\alpha)}{1-\alpha} \cdot \frac{1}{1-3\alpha^2}$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$c = \frac{3(\omega^2 + \alpha\omega^2)}{1-\alpha} . \frac{1}{1-3\alpha^2}$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$c = \frac{3(\omega^2 + \alpha\omega^2)}{1-\alpha} \cdot \frac{1}{1-3\alpha^2}$$

$$c = \frac{3\omega^2(1+\alpha)(1-\alpha)}{(1-\alpha)(1-\alpha)} \cdot \frac{1}{1-3\alpha^2}$$

Proof of Theorem: Unconditional variance of the ARCH(1))

$$c = \frac{3(\omega^2 + \alpha\omega^2)}{1-\alpha} \cdot \frac{1}{1-3\alpha^2}$$

$$c = \frac{3\omega^2(1+\alpha)(1-\alpha)}{(1-\alpha)(1-\alpha)} \cdot \frac{1}{1-3\alpha^2}$$

$$\Longrightarrow \; E(\varepsilon_t^4) = c = \frac{3\omega^2}{(1-\alpha)^2} \cdot \frac{1-\alpha^2}{1-3\alpha^2}, \; \text{with } 3\alpha^2 < 1$$

Proof of Theorem: Unconditional variance of the ARCH(1))

Now considering part 2 of the theorem: the unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(\varepsilon_t) = \frac{E(\varepsilon_t - \mu)^4}{(E(\varepsilon_t^2))^2} = \frac{E(\varepsilon_t^4)}{(\sigma^2)^2}$$

Proof of Theorem: Unconditional variance of the ARCH(1))

Now considering part 2 of the theorem: the unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(\varepsilon_t) = \frac{E(\varepsilon_t - \mu)^4}{(E(\varepsilon_t^2))^2} = \frac{E(\varepsilon_t^4)}{(\sigma^2)^2}$$

$$\text{Kurt}(\varepsilon_t) = \frac{3\omega^2}{(1-\alpha)^2} \cdot \frac{1-\alpha^2}{1-3\alpha^2} \times \frac{(1-\alpha)^2}{\omega^2}$$

Proof of Theorem: Unconditional variance of the ARCH(1))

Now considering part 2 of the theorem: the unconditional distribution of ε_t is leptokurtic

$$\text{Kurt}(\varepsilon_t) = \frac{E(\varepsilon_t - \mu)^4}{(E(\varepsilon_t^2))^2} = \frac{E(\varepsilon_t^4)}{(\sigma^2)^2}$$

$$\text{Kurt}(\varepsilon_t) = \frac{3\omega^2}{(1-\alpha)^2} . \frac{1-\alpha^2}{1-3\alpha^2} \times \frac{(1-\alpha)^2}{\omega^2}$$

$$\text{Kurt}(\varepsilon_t) = 3 . \frac{1-\alpha^2}{1-3\alpha^2} \geq 3$$

ARCH(q): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(q), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

ARCH(q): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(q), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

ARCH(q): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(q), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

with $\omega > 0, \alpha_1 \geq 0, \ldots, \alpha_q \geq 0$ and

ARCH(q): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(q), if $E(\varepsilon_t | \mathcal{F}_{t-1}) = 0$,

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

with $\omega > 0, \alpha_1 \geq 0, ..., \alpha_q \geq 0$ and

- $\text{Var}(\varepsilon_t | \mathcal{F}_{t-1}) = \sigma_t^2$ and $Z_t = \frac{\varepsilon_t}{\sigma_t}$ is i.i.d. (strong ARCH)

ARCH(q): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is ARCH(q), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

with $\omega > 0, \alpha_1 \geq 0, ..., \alpha_q \geq 0$ and

- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ and $Z_t = \frac{\varepsilon_t}{\sigma_t}$ is i.i.d. (strong ARCH)
- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ (semi-strong ARCH)

ARCH(q): Definition and Properties

The conditional variance equation is a function of two terms:

ARCH(q): Definition and Properties

The conditional variance equation is a function of two terms:

- A constant term, ω.

ARCH(q): Definition and Properties

The conditional variance equation is a function of two terms:

- A constant term, ω.
- News about volatility from the previous period, measured as the lag of the squared residual from the mean equation: ε_{t-i}^2.

Theorem: Unconditional variance of the ARCH(q))

Assume the process ε_t is a semi-strong ARCH(q) process with $\text{Var}(\varepsilon_t) = \sigma^2 < \infty$. Then it holds that

$$\sigma^2 = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i}$$

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

Proof

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

Proof

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

$$\sigma^2 = \omega + \sigma^2 \sum_{i=1}^{q} \alpha_i$$

Proof

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

$$\sigma^2 = \omega + \sigma^2 \sum_{i=1}^{q} \alpha_i$$

$$\implies \sigma^2 = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i}$$

Theorem: Representation of ARCH(1) as an AR(1) process

Let ε_t be a stationary strong ARCH(1) process with $E(\varepsilon_t^4) = c < \infty$ and $Z_t \sim N(0,1)$. It holds that

$$\eta_t = \sigma_t^2(Z_t^2 - 1) \text{ is white noise} \tag{1}$$

$$\varepsilon_t^2 \text{ is an AR(1) process with } \varepsilon_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \eta_t \tag{2}$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$E(\varepsilon_t^4) = c, \; Z_t \sim N(0,1), \; \varepsilon_t = Z_t\sigma_t$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$E(\varepsilon_t^4) = c, \; Z_t \sim N(0,1), \; \varepsilon_t = Z_t\sigma_t$$

$$\text{Now } \eta_t = \sigma_t^2(Z_t^2 - 1) = \sigma_t^2 Z_t^2 - \sigma_t^2$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$E(\varepsilon_t^4) = c, \; Z_t \sim N(0,1), \; \varepsilon_t = Z_t \sigma_t$$

$$\text{Now } \eta_t = \sigma_t^2 (Z_t^2 - 1) = \sigma_t^2 Z_t^2 - \sigma_t^2$$

$$\implies E(\eta_t) = E(\sigma_t^2) E(Z_t^2) - E(\sigma_t^2) = 0$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$E(\varepsilon_t^4) = c,\ Z_t \sim N(0,1),\ \varepsilon_t = Z_t\sigma_t$$

$$\text{Now } \eta_t = \sigma_t^2(Z_t^2 - 1) = \sigma_t^2 Z_t^2 - \sigma_t^2$$

$$\Longrightarrow\ E(\eta_t) = E(\sigma_t^2)E(Z_t^2) - E(\sigma_t^2) = 0$$

$$\mathsf{Var}(\eta_t) = \mathsf{Var}(\sigma_t^2 Z_t^2 - \sigma_t^2) = \mathsf{Var}(\sigma_t^2 Z_t^2) + \mathsf{Var}(\sigma_t^2)$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$E(\varepsilon_t^4) = c, \ Z_t \sim N(0,1), \ \varepsilon_t = Z_t \sigma_t$$

$$\text{Now } \eta_t = \sigma_t^2 (Z_t^2 - 1) = \sigma_t^2 Z_t^2 - \sigma_t^2$$

$$\implies E(\eta_t) = E(\sigma_t^2) E(Z_t^2) - E(\sigma_t^2) = 0$$

$$\mathrm{Var}(\eta_t) = \mathrm{Var}(\sigma_t^2 Z_t^2 - \sigma_t^2) = \mathrm{Var}(\sigma_t^2 Z_t^2) + \mathrm{Var}(\sigma_t^2)$$

$$= \mathrm{Var}(\sigma_t^2)\mathrm{Var}(Z_t^2) + \mathrm{Var}(\sigma_t^2) = \mathrm{Var}(\sigma_t^2)\left[E(Z_t^4) - \left(E(Z_t^2)\right)^2\right] + \mathrm{Var}(\sigma_t^2)$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$= \mathsf{Var}(\sigma_t^2)\,[3-1] + \mathsf{Var}(\sigma_t^2) = 3\mathsf{Var}(\sigma_t^2)$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$= \mathsf{Var}(\sigma_t^2)\,[3-1] + \mathsf{Var}(\sigma_t^2) = 3\mathsf{Var}(\sigma_t^2)$$

$$\mathsf{Var}(\eta_t) = 3\mathsf{Var}(\omega + \alpha\varepsilon_{t-1}^2)$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$= \mathsf{Var}(\sigma_t^2)\,[3-1] + \mathsf{Var}(\sigma_t^2) = 3\mathsf{Var}(\sigma_t^2)$$

$$\mathsf{Var}(\eta_t) = 3\mathsf{Var}(\omega + \alpha\varepsilon_{t-1}^2)$$

$$\mathsf{Var}(\eta_t) = 3\alpha^2\mathsf{Var}(\varepsilon_{t-1}^2) = 3\alpha^2\left[E(\varepsilon_{t-1}^4) - \left(E(\varepsilon_{t-1}^2)\right)^2\right]$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$= \mathrm{Var}(\sigma_t^2)\,[3-1] + \mathrm{Var}(\sigma_t^2) = 3\mathrm{Var}(\sigma_t^2)$$

$$\mathrm{Var}(\eta_t) = 3\mathrm{Var}(\omega + \alpha\varepsilon_{t-1}^2)$$

$$\mathrm{Var}(\eta_t) = 3\alpha^2\mathrm{Var}(\varepsilon_{t-1}^2) = 3\alpha^2\left[E(\varepsilon_{t-1}^4) - \left(E(\varepsilon_{t-1}^2)\right)^2\right]$$

$$\mathrm{Var}(\eta_t) = 3\alpha^2(c - (\sigma^2)^2) = 3\alpha^2(c - \sigma^4)$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(1)

$$= \mathsf{Var}(\sigma_t^2)\,[3-1] + \mathsf{Var}(\sigma_t^2) = 3\mathsf{Var}(\sigma_t^2)$$

$$\mathsf{Var}(\eta_t) = 3\mathsf{Var}(\omega + \alpha\varepsilon_{t-1}^2)$$

$$\mathsf{Var}(\eta_t) = 3\alpha^2\mathsf{Var}(\varepsilon_{t-1}^2) = 3\alpha^2\left[E(\varepsilon_{t-1}^4) - \left(E(\varepsilon_{t-1}^2)\right)^2\right]$$

$$\mathsf{Var}(\eta_t) = 3\alpha^2(c - (\sigma^2)^2) = 3\alpha^2(c - \sigma^4)$$

Since both the mean and variance of η_t are constant, η_t is a white noise process.

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(2)

$$\eta_t = \sigma_t^2 Z_t^2 - \sigma_t^2 \implies \eta_t + \sigma_t^2 = \sigma_t^2 Z_t^2$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(2)

$$\eta_t = \sigma_t^2 Z_t^2 - \sigma_t^2 \implies \eta_t + \sigma_t^2 = \sigma_t^2 Z_t^2$$

$$\text{but } \varepsilon_t^2 = \sigma_t^2 Z_t^2 \implies \varepsilon_t^2 = \sigma_t^2 + \eta_t$$

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(2)

$$\eta_t = \sigma_t^2 Z_t^2 - \sigma_t^2 \implies \eta_t + \sigma_t^2 = \sigma_t^2 Z_t^2$$

$$\text{but } \varepsilon_t^2 = \sigma_t^2 Z_t^2 \implies \varepsilon_t^2 = \sigma_t^2 + \eta_t$$

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

Proof of Theorem: Representation of ARCH(1) as an AR(1) process

(2)

$$\eta_t = \sigma_t^2 Z_t^2 - \sigma_t^2 \implies \eta_t + \sigma_t^2 = \sigma_t^2 Z_t^2$$

$$\text{but } \varepsilon_t^2 = \sigma_t^2 Z_t^2 \implies \varepsilon_t^2 = \sigma_t^2 + \eta_t$$

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

$$\implies \varepsilon_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \eta_t$$

Persistence

$$\alpha_1 + ... + \alpha_q < 1$$

Persistence of a ARCH/GARCH model refers to how fast large volatilities decay after a shock. If the sum is greater than 1, then the predictions of volatility are explosive.

The Generalized ARCH (GARCH)

GARCH(p,q): Definition and Properties

The process $\varepsilon_t, t \in \mathbb{Z}$, is GARCH(p,q), if $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$,

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

and

- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ and $Z_t = \frac{\varepsilon_t}{\sigma_t}$ is i.i.d. (strong GARCH)
- $\text{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2$ (semi-strong GARCH)

The Generalized ARCH (GARCH)

GARCH(p,q): Definition and Properties

The conditional variance equation is a function of three terms:

- A constant term, ω.

The Generalized ARCH (GARCH)

GARCH(p,q): Definition and Properties

The conditional variance equation is a function of three terms:

- A constant term, ω.
- News about volatility from the previous period, measured as the lag of the squared residual from the mean equation: ε_{t-i}^2 (the ARCH term).

The Generalized ARCH (GARCH)

GARCH(p,q): Definition and Properties

The conditional variance equation is a function of three terms:

- A constant term, ω.
- News about volatility from the previous period, measured as the lag of the squared residual from the mean equation: ε_{t-i}^2 (the ARCH term).
- Last period's forecast variance, σ_{t-j}^2 (the GARCH term).

Theorem: Unconditional variance of a GARCH(p,q))

Assume the process ε_t is a semi-strong GARCH(p,q) process with $\text{Var}(\varepsilon_t) = \sigma^2 < \infty$. Then

$$\sigma^2 = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}$$

$$\text{with } \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$$

Proof

Left as an exrecise for the students.

Example

For a GARCH(1,1) process show that

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

Example

For a GARCH(1,1) process show that

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

Proof

For the GARCH(1,1) process

$$\sigma^2 = \frac{\omega}{1-\alpha_1-\beta_1} \implies \omega = \sigma^2(1-\alpha_1-\beta_1)$$

Example

For a GARCH(1,1) process show that

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

Proof

For the GARCH(1,1) process

$$\sigma^2 = \frac{\omega}{1-\alpha_1-\beta_1} \implies \omega = \sigma^2(1-\alpha_1-\beta_1)$$

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

Persistence of volatility

1. The sum of the coefficients (α_i, β_j) in the conditional variance equation of the GARCH process measures the speed of decay of shocks to volatility.

Persistence of volatility

1. The sum of the coefficients (α_i, β_j) in the conditional variance equation of the GARCH process measures the speed of decay of shocks to volatility.
2. This is referred to as the persistence of the GARCH model, and it indicates how fast large volatilities declines after a shock.

Persistence of volatility

1. The sum of the coefficients (α_i, β_j) in the conditional variance equation of the GARCH process measures the speed of decay of shocks to volatility.
2. This is referred to as the persistence of the GARCH model, and it indicates how fast large volatilities declines after a shock.
3. Shocks to the conditional variability are highly persistent when the sum is greater than one, suggesting that the forecasts of volatility are explosive.

Persistence of volatility

1. The sum of the coefficients (α_i, β_j) in the conditional variance equation of the GARCH process measures the speed of decay of shocks to volatility.
2. This is referred to as the persistence of the GARCH model, and it indicates how fast large volatilities declines after a shock.
3. Shocks to the conditional variability are highly persistent when the sum is greater than one, suggesting that the forecasts of volatility are explosive.
4. If the sum of the coefficients equals one, then the persistence of shocks to volatility is felt forever, and the model will be unable to determine the unconditional variance of the process.

Half life

Is the number of days (hours,weeks, etc) it takes for half of the expected reversion back towards $E(\sigma^2)$ after a shock (or crisis). Let K denote the half life, then

$$\begin{aligned}(\alpha_1+\beta_1)^K &= 0.5\\ \Rightarrow K &= \frac{\log(0.5)}{\log(\alpha_1+\beta_1)}\end{aligned}$$

Half life

Is the number of days (hours,weeks, etc) it takes for half of the expected reversion back towards $E(\sigma^2)$ after a shock (or crisis). Let K denote the half life, then

$$\begin{aligned} (\alpha_1 + \beta_1)^K &= 0.5 \\ \Rightarrow K &= \frac{\log(0.5)}{\log(\alpha_1 + \beta_1)} \end{aligned}$$

Example: For a GARCH(1,1) if $\alpha_1 + \beta_1 = 0.97$, then $K \approx 23$ days.

Conditional distributions

The Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Conditional distributions

The Student distribution

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma(v/2)\sqrt{\beta v \pi}} \left(1 + \frac{(x-\alpha)^2}{\beta v}\right)^{-\left(\frac{v+1}{2}\right)}$$

where α, β and v are the location, scale and shape parameters respectively.

Conditional distributions

The Generalized Error distribution

$$f(x) = \frac{\kappa e^{-0.5\left|\frac{x-\alpha}{\beta}\right|^{\kappa}}}{2^{1+\kappa^{-1}}\beta\Gamma(\kappa^{-1})}$$

where α, β and κ are the location, scale and shape parameters respectively.

Estimation of GARCH Models

Maximum likelihood method

The log-likelihood function under the assumption that the error term follows a standard Gaussian distribution is given by

$$\log L_{\text{norm}} = -\frac{1}{2}\sum_{t=1}^{n}\left[\log(2\pi) + \log(\sigma_t^2) + \nu_t\right]$$

Estimation of GARCH Models

Maximum likelihood method

Under the assumption of a Student-t distribution the log-likelihood is

$$\log L_{\text{Stud}} = n\Big[\Big]\log\Gamma\left(\frac{u+1}{2}\right) - \log\Gamma\left(\frac{u}{2}\right)$$

$$-\frac{1}{2}\log\left[\pi(u-2)\right]\Big] - \frac{1}{2}\sum_{t=1}^{n}\Big[\log(\sigma_t^2) + (1+u)$$

$$\log\left(1+\frac{\nu_t^2}{u-2}\right)\Big]$$

Estimation of GARCH Models

Maximum likelihood method

where u is the degrees of freedom. The GED log-likelihood function is given by

$$\log L_{\text{GED}} = \sum_{t=1}^{n}\left[\log\left(\frac{u}{\lambda_u}\right) - 0.5\left|\frac{\nu_t}{\lambda_u}\right|^{u} - (1+u^{-1})\log(2) - \log\Gamma\left(\frac{1}{u}\right) - 0.5\log(\sigma_t^2)\right]$$

where $u > 0$ and $\lambda_u = \sqrt{\frac{\Gamma\left(\frac{1}{u}\right)2^{-2/u}}{\Gamma\left(\frac{3}{u}\right)}}$.

Model selection

Information criteria

Akaike
Bayes
Hannan-Quinn
Shibata

Model selection

Akaike information criterion

- AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.

Model selection

Akaike information criterion

- AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.
- AIC deals with both the risk of overfitting and the risk of underfitting.

Model selection

Akaike information criterion

- AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.
- AIC deals with both the risk of overfitting and the risk of underfitting.
- AIC $= 2p - 2\ell$

Model selection

Akaike information criterion

- AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.
- AIC deals with both the risk of overfitting and the risk of underfitting.
- AIC $= 2p - 2\ell$
- AIC however is criticized for not being asymptotically consistent because the computation of the AIC does not directly involve sample size.

Model selection

Bayesian information criterion

- The BIC discourages overparameterization by imposing more heavy penalties on model complexity than does the AIC.

Model selection

Bayesian information criterion

- The BIC discourages overparameterization by imposing more heavy penalties on model complexity than does the AIC.
- The BIC is consistent such that, as the sample size increases, the criterion will select a true model if it is included among the candidate models.

Model selection

Bayesian information criterion

- The BIC discourages overparameterization by imposing more heavy penalties on model complexity than does the AIC.
- The BIC is consistent such that, as the sample size increases, the criterion will select a true model if it is included among the candidate models.
- BIC $= p\log(n) - 2\ell$

Model selection

Bayesian information criterion

- The BIC discourages overparameterization by imposing more heavy penalties on model complexity than does the AIC.
- The BIC is consistent such that, as the sample size increases, the criterion will select a true model if it is included among the candidate models.
- BIC $= p\log(n) - 2\ell$
- The BIC can be used to compare estimated models only when the numerical values of the dependent variable are identical for all models being compared.

Model selection

Hannan-Quinn information criterion

- It satisfies the consistency property.

Model selection

Hannan-Quinn information criterion

- It satisfies the consistency property.
- Is not asymptotically efficient. It is rarely used in practice.

Model selection

Hannan-Quinn information criterion

- It satisfies the consistency property.
- Is not asymptotically efficient. It is rarely used in practice.
- HQIC $= -2\ell + 2p\log(\log(n))$

Model diagnostics

Ljung-Box and ARCH LM tests

1. Ljung-Box test on standardized residuals
2. Ljung-Box test on standardized squared residuals
3. ARCH LM test

Model diagnostics

Ljung-Box test on standardized squared residuals

- The Ljung-Box test on standardized squared residuals is used to test for auto-correlation within the squared residuals.

Model diagnostics

Ljung-Box test on standardized squared residuals

- The Ljung-Box test on standardized squared residuals is used to test for auto-correlation within the squared residuals.
- The null hypothesis is that the auto-correlation between the residuals for a set of lags $k = 0$.

Model diagnostics

ARCH Lagrange Multiplier test

- ARCH LM test is a standard test to detect ARCH effects.

Model diagnostics

ARCH Lagrange Multiplier test

- ARCH LM test is a standard test to detect ARCH effects.
- ARCH LM test provides a means of testing for serial dependence (auto-correlation) due to a conditional variance process by testing for auto-correlation within the squared residuals.

Model diagnostics

ARCH Lagrange Multiplier test

- ARCH LM test is a standard test to detect ARCH effects.
- ARCH LM test provides a means of testing for serial dependence (auto-correlation) due to a conditional variance process by testing for auto-correlation within the squared residuals.
- The null hypothesis is that the auto-correlation between the residuals for a set of lags $k = 0$.

Forecasting: $l-$step ahead forecasts

1. Consider a GARCH(1,1) process, derive
   1. $\hat{\sigma}^2_{t+1}$,
   2. $\hat{\sigma}^2_{t+2}$,
   3. $\hat{\sigma}^2_{t+l}$.
2. Find

$$\lim_{l\to\infty} \hat{\sigma}^2_{t+l}$$

Solution

1.1

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

Solution

1.1

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

$$\hat{\sigma}_{t+1}^2 = E(\sigma_{t+1}^2 | I_{t-1})$$

Solution

1.1

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

$$\hat{\sigma}_{t+1}^2 = E(\sigma_{t+1}^2 | I_{t-1})$$

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

Solution

1.1

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

$$\hat{\sigma}_{t+1}^2 = E(\sigma_{t+1}^2 | I_{t-1})$$

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

$$\hat{\sigma}_{t+1}^2 = \omega + \alpha_1 \sigma_t^2 + \beta_1 \sigma_t^2$$

Solution

1.1

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

$$\hat{\sigma}_{t+1}^2 = E(\sigma_{t+1}^2|I_{t-1})$$

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

$$\hat{\sigma}_{t+1}^2 = \omega + \alpha_1\sigma_t^2 + \beta_1\sigma_t^2$$

Solution

1.1

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

Solution

1.1

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

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

Solution

1.1

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

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

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

Solution

1.2

$$\hat{\sigma}^2_{t+2} = \omega + \alpha_1 E(\varepsilon^2_{t+1}|I_{t-1}) + \beta_1 E(\sigma^2_{t+1}|I_{t-1})$$

Solution

1.2

$$\hat{\sigma}^2_{t+2} = \omega + \alpha_1 E(\varepsilon^2_{t+1}|I_{t-1}) + \beta_1 E(\sigma^2_{t+1}|I_{t-1})$$

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

Solution

1.2

$$\hat{\sigma}^2_{t+2} = \omega + \alpha_1 E(\varepsilon^2_{t+1}|I_{t-1}) + \beta_1 E(\sigma^2_{t+1}|I_{t-1})$$

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

$$\hat{\sigma}^2_{t+2} = \sigma^2 - (\alpha_1 + \beta_1)\sigma^2 + (\alpha_1 + \beta_1)(\sigma^2 + (\alpha_1 + \beta_1)(\sigma^2_t - \sigma^2))$$

Solution

1.2

$$\hat{\sigma}^2_{t+2} = \omega + \alpha_1 E(\varepsilon^2_{t+1}|I_{t-1}) + \beta_1 E(\sigma^2_{t+1}|I_{t-1})$$

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

$$\hat{\sigma}^2_{t+2} = \sigma^2 - (\alpha_1 + \beta_1)\sigma^2 + (\alpha_1 + \beta_1)(\sigma^2 + (\alpha_1 + \beta_1)(\sigma^2_t - \sigma^2))$$

$$\hat{\sigma}^2_{t+2} = \sigma^2 - (\alpha_1 + \beta_1)\sigma^2 + (\alpha_1 + \beta_1)\sigma^2 + (\alpha_1 + \beta_1)^2(\sigma^2_t - \sigma^2)$$

Solution

1.2

$$\hat{\sigma}^2_{t+2} = \omega + \alpha_1 E(\varepsilon^2_{t+1}|I_{t-1}) + \beta_1 E(\sigma^2_{t+1}|I_{t-1})$$

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

$$\hat{\sigma}^2_{t+2} = \sigma^2 - (\alpha_1 + \beta_1)\sigma^2 + (\alpha_1 + \beta_1)(\sigma^2 + (\alpha_1 + \beta_1)(\sigma^2_t - \sigma^2))$$

$$\hat{\sigma}^2_{t+2} = \sigma^2 - (\alpha_1 + \beta_1)\sigma^2 + (\alpha_1 + \beta_1)\sigma^2 + (\alpha_1 + \beta_1)^2(\sigma^2_t - \sigma^2)$$

$$\hat{\sigma}^2_{t+2} = \sigma^2 + (\alpha_1 + \beta_1)^2(\sigma^2_t - \sigma^2)$$

Solution

1.3

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

Solution

1.3

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

2.0

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

Solution

1.3

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

2.0

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

For stability (stationarity)

Solution

1.3

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

2.0

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

For stability (stationarity) $\alpha_1 + \beta_1 < 1$, then

Solution

1.3

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

2.0

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

For stability (stationarity) $\alpha_1 + \beta_1 < 1$, then

$$\lim_{l\to\infty} \hat{\sigma}^2_{t+l} = \lim_{l\to\infty} \left[\sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)\right]$$

Solution

1.3

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

2.0

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

For stability (stationarity) $\alpha_1 + \beta_1 < 1$, then

$$\lim_{l \to \infty} \hat{\sigma}^2_{t+l} = \lim_{l \to \infty} \left[\sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)\right]$$

$$\lim_{l \to \infty} \hat{\sigma}^2_{t+l} = \sigma^2$$

Solution

1.3

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

2.0

$$\hat{\sigma}^2_{t+l} = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)$$

For stability (stationarity) $\alpha_1 + \beta_1 < 1$, then

$$\lim_{l\to\infty} \hat{\sigma}^2_{t+l} = \lim_{l\to\infty} \left[\sigma^2 + (\alpha_1 + \beta_1)^l(\sigma^2_t - \sigma^2)\right]$$

$$\lim_{l\to\infty} \hat{\sigma}^2_{t+l} = \sigma^2$$

since $(\alpha_1 + \beta_1)^l \to 0,$ as $l \to \infty$.

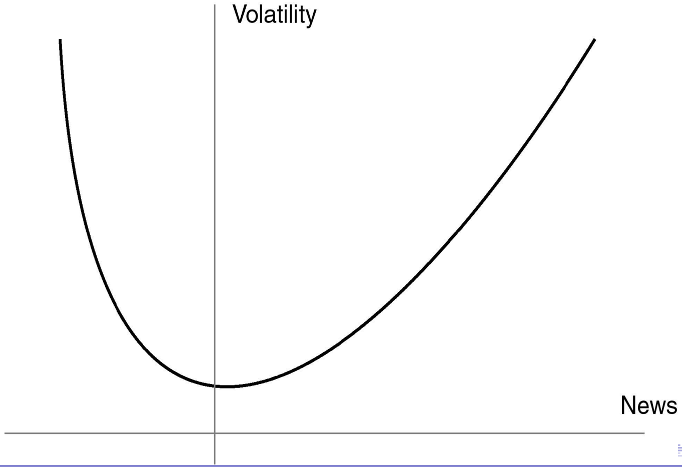
Threshold GARCH (TGARCH)

- Standard GARCH models (Bollerslev, 1986) do not distinguish the differential impacts of good news from bad news on volatility

Threshold GARCH (TGARCH)

- Standard GARCH models (Bollerslev, 1986) do not distinguish the differential impacts of good news from bad news on volatility
- To overcome this the Threshold GARCH (TGARCH) (introduced independently by Zakoian (1990) and Glosten *et al.* (1993)) can be used to model the asymmetric responses of volatility to positive and negative shocks.

News impact curve with asymmetric response to good and bad news



Threshold GARCH (TGARCH)

- $$\text{Mean equation } r_t = \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Threshold GARCH (TGARCH)

- $$\text{Mean equation } r_t = \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$
- The variance equation is a TGARCH(1,1) model given as :

Threshold GARCH (TGARCH)

- $$\text{Mean equation } r_t = \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$
- The variance equation is a TGARCH(1,1) model given as :

$$\text{Variance equation} \quad \sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 d_{t-1} + b_1 \sigma_{t-1}^2$$

Threshold GARCH (TGARCH)

- $$\text{Mean equation } r_t = \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$$

- The variance equation is a TGARCH(1,1) model given as :

$$\text{Variance equation} \quad \sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 d_{t-1} + b_1 \sigma_{t-1}^2$$

$$d_t = \begin{cases} 1 & \text{if } \varepsilon_t < 0 \\ 0 & \text{otherwise,} \end{cases}$$

- where γ represents the threshold. When $\gamma = 0$ the TGARCH reduces to the GARCH model.

Threshold GARCH (TGARCH)

- In financial time series the effects of good and bad news are measured by ε_t in the variance equation.

Threshold GARCH (TGARCH)

- In financial time series the effects of good and bad news are measured by ε_t in the variance equation.
- If $\varepsilon_t > 0$ we have good news meaning that the volatility does not increase significantly.

Threshold GARCH (TGARCH)

- In financial time series the effects of good and bad news are measured by ε_t in the variance equation.
- If $\varepsilon_t > 0$ we have good news meaning that the volatility does not increase significantly.
- We have bad news if $\varepsilon_t < 0$ resulting in an increase in volatility.

Threshold GARCH (TGARCH)

- In financial time series the effects of good and bad news are measured by ε_t in the variance equation.
- If $\varepsilon_t > 0$ we have good news meaning that the volatility does not increase significantly.
- We have bad news if $\varepsilon_t < 0$ resulting in an increase in volatility.
- The impacts of good and bad news are measured by a and $a + \gamma$ respectively.

Threshold GARCH (TGARCH)

- In financial time series the effects of good and bad news are measured by ε_t in the variance equation.
- If $\varepsilon_t > 0$ we have good news meaning that the volatility does not increase significantly.
- We have bad news if $\varepsilon_t < 0$ resulting in an increase in volatility.
- The impacts of good and bad news are measured by a and $a + \gamma$ respectively.
- For $\gamma \neq 0$ the news impact is asymmetric and for $\gamma > 0$ the leverage effect exists meaning that bad news increases volatility.

Component GARCH (CGARCH)

- The Component GARCH (CGARCH) developed by Engle and Lee (1999) decomposes volatility into a permanent and transitory comonent.

Component GARCH (CGARCH)

- The Component GARCH (CGARCH) developed by Engle and Lee (1999) decomposes volatility into a permanent and transitory comonent.
- The objective is investigate the long- and short-run movements of volatility.

Component GARCH (CGARCH)

- The Component GARCH (CGARCH) developed by Engle and Lee (1999) decomposes volatility into a permanent and transitory comonent.
- The objective is investigate the long- and short-run movements of volatility.
- The CGARCH model allows for the reversion of the conditional volatility to a time-varying trend component q.

Component GARCH (CGARCH)

- The Component GARCH (CGARCH) developed by Engle and Lee (1999) decomposes volatility into a permanent and transitory comonent.
- The objective is investigate the long- and short-run movements of volatility.
- The CGARCH model allows for the reversion of the conditional volatility to a time-varying trend component q.
- Let q_t represent the permanent component of the conditional variance, then the CGARCH(1,1) model can be written as:

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

$$q_t = \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)$$

Component GARCH (CGARCH)

- where σ_t^2 is volatility, q_t is the time-varying long-run volatility, $\sigma_t^2 - q_t$ is the transitory component of volatility.

Component GARCH (CGARCH)

- where σ_t^2 is volatility, q_t is the time-varying long-run volatility, $\sigma_t^2 - q_t$ is the transitory component of volatility.
- For the permanent component, the speed of mean reversion is determined by q, value of which typically lies between $(\alpha + \beta) < q < 1$.

Component GARCH (CGARCH)

- where σ_t^2 is volatility, q_t is the time-varying long-run volatility, $\sigma_t^2 - q_t$ is the transitory component of volatility.
- For the permanent component, the speed of mean reversion is determined by q, value of which typically lies between $(\alpha + \beta) < q < 1$.
- A concern in using standard CGARCH model is that it assumes symmetry between positive and negative shocks.

Asymmetric CGARCH (ACGARCH)

- $$\sigma_t^2 - q_t = \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \gamma(\varepsilon_{t-1}^2 - q_{t-1})d_{t-1} + \beta(\sigma_{t-1}^2 - q_{t-1})$$
  $$q_t = \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)$$

Asymmetric CGARCH (ACGARCH)

- $$\sigma_t^2 - q_t = \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \gamma(\varepsilon_{t-1}^2 - q_{t-1})d_{t-1} + \beta(\sigma_{t-1}^2 - q_{t-1})$$
$$q_t = \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)$$

- where d is a dummy variable indicating a negative shock and takes a value of 1 if $\varepsilon_{t-1} < 0$, and 0 otherwise.

Asymmetric CGARCH (ACGARCH)

- $$\sigma_t^2 - q_t = \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \gamma(\varepsilon_{t-1}^2 - q_{t-1})d_{t-1} + \beta(\sigma_{t-1}^2 - q_{t-1})$$
  $$q_t = \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)$$
- where d is a dummy variable indicating a negative shock and takes a value of 1 if $\varepsilon_{t-1} < 0$, and 0 otherwise.
- The effect of a positive news on volatility is measured by the coefficient α_1.

Asymmetric CGARCH (ACGARCH)

- $$\sigma_t^2 - q_t = \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \gamma(\varepsilon_{t-1}^2 - q_{t-1})d_{t-1} + \beta(\sigma_{t-1}^2 - q_{t-1})$$
  $$q_t = \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2)$$
- where d is a dummy variable indicating a negative shock and takes a value of 1 if $\varepsilon_{t-1} < 0$, and 0 otherwise.
- The effect of a positive news on volatility is measured by the coefficient α_1.
- Effect of negative news is measured by the coefficients $(\alpha + \gamma)$.

Computer Parctical II