Introduction to Financial Econometrics

Chapter 6: Introduction to GARCH Models

Christophe Hurlin

Univ Orléans

May, 2019

1. Introduction

Definition (volatility)

The **volatility** is defined as the degree of variation of a trading price series over time.

Notes

1. Volatility is an important factor in **options trading** (for instance, see the Black–Scholes option pricing formula).
2. Volatility is a key factor of **risk management** (risk measures, risk-adjusted performance measures, etc.).

Overview

- A special feature of stock volatility is that it is unobservable.
- The daily volatility is not directly observable from the returns R_t because there is only one observation in a trading day t, denoted r_t.
- So, we need to estimate it.

1. Introduction

Several approaches have been proposed to measure and to forecast the volatility:

1. The **implied volatility** which is defined as the volatility of an underlying instrument which, when input in an option pricing model (such as Black–Scholes) will return a theoretical value equal to the current market price of the option. Example: the CBOE Volatility Index (ticker: VIX).

2. The estimators based on high frequency data. For instance, the **realized volatility** is computed as the sum of squared intraday returns for a particular day (Andersen and Bollerslev, 1998).

3. The **conditional variance** issued from dynamic models such as the ARCH and GARCH type models.

1. Introduction

The outline of this chapter is the following:

- **Section 2:** ARCH models
  - **Sub-Section 2.1:** Properties of ARCH models
  - **Sub-Section 2.2:** Building an ARCH model
- **Section 3:** GARCH models
- **Section 4:** Extensions of GARCH models

1. Introduction

References

Brooks, C., Introductory Econometrics for Finance, Cambridge University Press, 3rd edition, 2014.

Campbell, J.,Y., A.W. Lo and A.C. MacKinlay, The Econometrics of Financial Markets, Princeton University Press, 1997.

Francq, C. and J.M. Zakoian, GARCH Models: Structure, Statistical Inference and Financial Applications, Wiley, 2010. (**main reference**)

Tsay, R., 2002, Analysis of Financial Time Series, Wiley Series.

Section 2

ARCH Models

2. ARCH models

The ARCH model has been introduced by **Engle (1982)**

ARCH = **A**uto**R**egressive **C**onditional **H**eteroskedasticity

Robert F. Engle Nobel Prize 2003

Engle, R.F. (1982), AutoRegressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, Econometrica, 50, 987-1008.

2. ARCH models

ARCH = **A**uto**R**egressive **C**onditional **H**eteroskedasticity

- The term **heteroscedasticity** refers to a time-varying variance (cf. Chapter 2).
- In an ARCH model, it is the **conditional variance** (and not the variance itself) which changes with time, in a specific way, depending on the available data.
- The ARCH model assumes that the conditional variance is a linear function of the **past squared return**.
- Thus, the squared return follows an **autoregressive model**.

Definition (ARCH(1))

The process $\{X_t, t \in \mathbb{Z}\}$ is said to be an **ARCH**(1) process, if

$$X_t = Z_t \sigma_t$$

where Z_t is a sequence of independent and identically distributed (i.i.d.) random variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and σ_t is a non-negative process such that

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

with $\alpha_0 > 0$ and $0 \leq \alpha_1 < 1$.

2. ARCH models

Definition (conditional variance)

The process σ_t^2 corresponds to the **conditional variance** of X_t.

$$\mathbb{V}\left(X_t | \mathcal{F}_{t-1}\right) \equiv \mathbb{V}\left(X_t | \underline{X}_{t-1}\right) = \sigma_t^2$$

where $\mathcal{F}_{t-1} \equiv \underline{X}_{t-1} = \{X_{t-1}, X_{t-2}, \ldots\}$ is the information set available at time $t-1$.

Interpretation: Some authors denote the conditional variance by h_t, with

$$X_t = Z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2$$

Interpretation

Consider an ARCH(1) process

$$X_t = Z_t \sigma_t$$

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

Then, we have

$$\mathbb{V}\left(X_t | \underline{X}_{t-1}\right) = \mathbb{V}\left(Z_t \sigma_t | \underline{X}_{t-1}\right) = \sigma_t^2 \mathbb{V}\left(Z_t | \underline{X}_{t-1}\right) = \sigma_t^2 \mathbb{V}\left(Z_t\right) = \sigma_t^2$$

1. Given the past information $\underline{X}_{t-1}$, the conditional variance $\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2$ is **deterministic**, since x_{t-1} is a constant.

2. The process $\{Z_t, t \in \mathbb{Z}\}$ is an IID noise (cf. Chapter 5), so $\mathbb{V}\left(Z_t | \underline{X}_{t-1}\right) = \mathbb{V}\left(Z_t\right)$, i.e. there is no "memory" in Z_t.

3. The normalization $\mathbb{V}\left(Z_t\right) = 1$ is not a restriction: the scaling implied by any other variance would be absorbed by the parameters α_0 and α_1.

Example (ARCH(1) process)

Consider the following ARCH(1) process with a Gaussian innovation

$$X_t = Z_t \sigma_t$$

$$\sigma_t^2 = 1 + 0.5 X_{t-1}^2$$

$$Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$$

with an initial condition $X_0 = 0$. **Question:** simulate 500 values of the conditional variance $\sigma_t^2 = \mathbb{V}\left(X_t | \underline{X}_{t-1}\right)$ for $t = 1, \ldots, 500$.

Solution

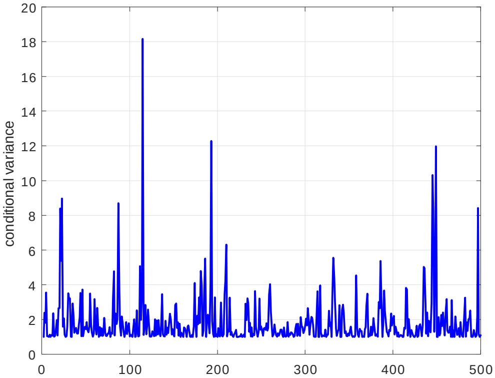
Consider the following realizations for the innovation process Z_t.

$$X_t = Z_t \sigma_t \qquad \sigma_t^2 = 1 + 0.5 X_{t-1}^2$$

Realizations	z_t	x_{t-1}	$\sigma_t^2 = 1 + 0.5 x_{t-1}^2$	x_t
$t = 0$	—	—	—	0
$t = 1$	-1.6718	0	$1 + 0.5 \times (0)^2 = 1.0000$	-1.6718
$t = 2$	0.8267	-1.6718	$1 + 0.5 \times (-1.6718)^2 = 2.3975$	1.2801
$t = 3$	-1.6764	1.2801	$1 + 0.5 \times (1.2801)^2 = 1.8193$	-2.2612
$t = 4$	-0.1554	-2.2612	$1 + 0.5 \times (-2.2612)^2 = 3.5564$	-0.2931

2. ARCH models

Figure: Simulated conditional variance for an ARCH(1) process with Gaussian innovations



Sub-Section 2.1

Properties of ARCH Models

Objectives

1. To understand the **main properties** of an ARCH(1) model
2. To make the distinction between the **conditional** and **unconditional** variances
3. To establish a link between an **ARCH model** on X_t and an **AR representation** on X_t^2
4. To understand the source of the **ARCH effect**
5. To show that a ARCH process is a **martingale difference**
6. To make the distinction between the **conditional** and **unconditional** distributions

Definition

Property 1: *If* $\{X_t, t \in \mathbb{Z}\}$ *has a* $ARCH(1)$ *representation, with*

$$X_t = Z_t \sigma_t$$

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

then $\left\{X_t^2, t \in \mathbb{Z}\right\}$ *has an* **AR**(1) **representation**, *with*

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + v_t$$

where v_t *is an innovation process*

$$\mathbb{E}\left(v_t \middle| \underline{X}_{t-1}\right) = 0$$

Proof: Consider an ARCH(1) model such that

$$X_t = z_t \sigma_t$$

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

Add X_t^2 on both sides of the equation of σ_t^2, then we get

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

and rewrite X_t^2 as

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + v_t$$

with

$$v_t = X_t^2 - \sigma_t^2$$

Proof (cont'd): Show that $v_t = X_t^2 - \sigma_t^2$ is an innovation, i.e. $\mathbb{E}\left(v_t | \underline{X}_{t-1}\right) = 0$.

$$
\begin{aligned}
\mathbb{E}\left(v_t | \underline{X}_{t-1}\right) &= \mathbb{E}\left(X_t^2 | \underline{X}_{t-1}\right) - \mathbb{E}\left(\sigma_t^2 | \underline{X}_{t-1}\right) \\
&= \mathbb{E}\left(z_t^2 \sigma_t^2 | \underline{X}_{t-1}\right) - \sigma_t^2 \\
&= \sigma_t^2 \mathbb{E}\left(z_t^2 | \underline{X}_{t-1}\right) - \sigma_t^2 \\
&= \sigma_t^2 \mathbb{V}\left(z_t | \underline{X}_{t-1}\right) - \sigma_t^2 \\
&= \sigma_t^2 \mathbb{V}\left(z_t\right) - \sigma_t^2 \\
&= \sigma_t^2 - \sigma_t^2 \\
&= 0
\end{aligned}
$$

So, X_t^2 has an AR(1) representation. $\square$

Consequences

1. If $\{X_t, t \in \mathbb{Z}\}$ has a **ARCH(1)** representation

$$X_t = z_t \sigma_t$$

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

2. Then $\{X_t^2, t \in \mathbb{Z}\}$ has an **AR(1)** representation

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + v_t$$

3. Thus X_t^2 and X_{t-k}^2 are correlated: => **ARCH effect**

$$\rho_k = \mathbb{C}orr\left(X_t^2, X_{t-k}^2\right) \neq 0$$

especially for small values of k.

Definition (Yule-Walker equations)

Consider an AR(p) process $\{Y_t, \ t \in \mathbb{Z}\}$ such that

$$\Phi(L) \, Y_t = c + \varepsilon_t \iff Y_t = c + \phi_1 Y_{t-1} ... + \phi_p Y_{t-p} + \varepsilon_t$$

where $\varepsilon_t \sim$WN$\left(0, \sigma^2\right)$ is a white noise process, then the **autocorrelation function** $\rho_k = \mathbb{C}orr\left(Y_t, Y_{t-k}\right)$ satisfies a recurrence relation of the form

$$\Phi(L) \, \rho_k = 0 \iff \rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \ldots + \phi_p \rho_{k-p} \quad \forall k \in \mathbb{Z}^*$$

with $\rho_0 = 1$. These equations are called the **Yule-Walker equations**.

2.1. Properties of ARCH models

Yule Walker equation and ARCH effect

If the process $\left\{X_t^2, t \in \mathbb{Z}\right\}$ has an AR(1) representation with

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + v_t$$

$$\mathbb{E}\left(v_t | \underline{X}_{t-1}\right) = 0$$

then its **autocorrelation function** $\rho_k = \mathbb{C}orr\left(X_t^2, X_{t-k}^2\right)$ satisfies the following recurrence relation

$$\rho_k = \alpha_1 \rho_{k-1}$$

$$\rho_0 = 1$$

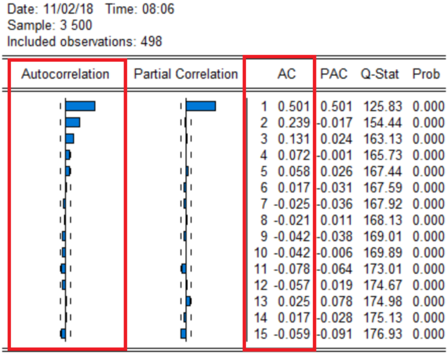
Reminder Chapter 1. Stylized Fact 7: (ARCH effect)

Fact (ARCH effect)

The daily squared returns often exhibit significant correlations. These autocorrelations are often referred as an ARCH effect.

2.1. Properties of ARCH models

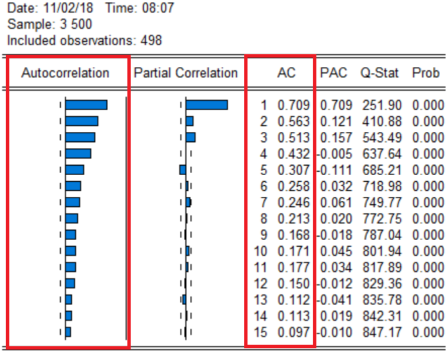
Figure: ACF of a simulated AR(1) process


Date: 11/02/18 Time: 08:06
Sample: 3 500
Included observations: 498

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.501	0.501	125.83	0.000
		2	0.239	-0.017	154.44	0.000
		3	0.131	0.024	163.13	0.000
		4	0.072	-0.001	165.73	0.000
		5	0.058	0.026	167.44	0.000
		6	0.017	-0.031	167.59	0.000
		7	-0.025	-0.036	167.92	0.000
		8	-0.021	0.011	168.13	0.000
		9	-0.042	-0.038	169.01	0.000
		10	-0.042	-0.006	169.89	0.000
		11	-0.078	-0.064	173.01	0.000
		12	-0.057	0.019	174.67	0.000
		13	0.025	0.078	174.98	0.000
		14	0.017	-0.028	175.13	0.000
		15	-0.059	-0.091	176.93	0.000

2.1. Properties of ARCH models

Figure: ACF of a simulated AR(3) process



Date: 11/02/18 Time: 08:07
Sample: 3 500
Included observations: 498

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.709	0.709	251.90	0.000
		2	0.563	0.121	410.88	0.000
		3	0.513	0.157	543.49	0.000
		4	0.432	-0.005	637.64	0.000
		5	0.307	-0.111	685.21	0.000
		6	0.258	0.032	718.98	0.000
		7	0.246	0.061	749.77	0.000
		8	0.213	0.020	772.75	0.000
		9	0.168	-0.018	787.04	0.000
		10	0.171	0.045	801.94	0.000
		11	0.177	0.034	817.89	0.000
		12	0.150	-0.012	829.36	0.000
		13	0.112	-0.041	835.78	0.000
		14	0.113	0.019	842.31	0.000
		15	0.097	-0.010	847.17	0.000

Figure: ACF for the S&P500 squared returns (August 19, 2013 to August 17, 2018)

Date: 09/09/18 Time: 11:25
Sample: 8/19/2013 8/17/2018
Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.302	0.302	114.94	0.000
		2	0.251	0.176	194.81	0.000
		3	0.288	0.196	299.82	0.000
		4	0.251	0.116	379.50	0.000
		5	0.115	-0.052	396.30	0.000
		6	0.135	0.018	419.45	0.000
		7	0.095	-0.020	430.83	0.000
		8	0.108	0.044	445.63	0.000
		9	0.102	0.043	458.93	0.000
		10	0.123	0.060	478.14	0.000
		11	0.084	0.003	487.21	0.000
		12	0.050	-0.041	490.35	0.000
		13	0.048	-0.020	493.24	0.000
		14	0.040	-0.012	495.29	0.000
		15	0.031	0.010	496.51	0.000

Definition

Property 2: *if $\{X_t, t \in \mathbb{Z}\}$ is an ARCH(1) process, then it is a martingale difference*

$$\mathbb{E}\left(X_t | \mathcal{F}_{t-1}\right) \equiv \mathbb{E}\left(X_t | \underline{X}_{t-1}\right) = 0$$

Consequences

- The very best (linear or nonlinear) predictor of X_t based on the available information at time $t-1$ is simply the trivial predictor, namely the series mean, 0.
- In terms of point forecasting of the series itself, then, the ARCH models offer no advantages over the linear ARMA models.
- This property implies that $\mathbb{C}ov\left(X_t, X_{t-k}\right) = 0$ for $k \neq 0$, i.e. that the process X_t has no "memory".

2.1. Properties of ARCH models

Innovation		**ARCH model**		**Output:** X_t
Z_t i.i.d. noise	$\Longrightarrow$	$X_t = Z_t \sigma_t$	$\Longrightarrow$	$\mathbb{E}\left(X_t \mid \mathcal{F}_{t-1}\right) = 0$
No **dependence**		$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$		X_t is a martingale difference
between Z_t and Z_{t-k}				**No correlation**
				between X_t and X_{t-k}
				but $\mathbb{C}orr\left(X_t^2, X_{t-k}^2\right) \neq 0$

Reminder Chapter 1. Stylized Fact 2 (Absence of autocorrelations)

Fact (absence of autocorrelations)

The **autocorrelations** *of asset returns R_t are often insignificant, except for very small intraday time scales ($\approx$ 20 minutes) for which microstructure effects come into play.*

2.1. Properties of ARCH models

Proof: Consider an ARCH(1) model such that

$$X_t = Z_t \sigma_t$$

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

It is possible to show that X_t is a martingale difference, since

$$\begin{aligned} \mathbb{E}\left(X_t | \underline{X}_{t-1}\right) &= \mathbb{E}\left(Z_t \sigma_t | \underline{X}_{t-1}\right) \\ &= \sigma_t \mathbb{E}\left(Z_t | \underline{X}_{t-1}\right) \\ &= \sigma_t \mathbb{E}\left(Z_t\right) \\ &= 0 \quad \square \end{aligned}$$

since Z_t is an i.i.d. process $\mathbb{E}\left(Z_t\right) = 1$.

2.1. Properties of ARCH models

Definition

Property 3: *if* $\{X_t, t \in \mathbb{Z}\}$ *is an ARCH(1) process with*

$$X_t = Z_t \sigma_t$$

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

then, its two first **unconditional moments** *are equal to*

$$\mathbb{E}(X_t) = 0 \qquad \mathbb{V}(X_t) = \frac{\alpha_0}{1 - \alpha_1}$$

with $\alpha_0 > 0$ *and* $0 \leq \alpha_1 < 1$.

2.1. Properties of ARCH models

Proof: Consider an ARCH(1) model such that

$$X_t = Z_t \sigma_t$$

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

The (unconditional) mean of X_t is defined as

$$\begin{aligned}
\mathbb{E}(X_t) &= \mathbb{E}(Z_t \sigma_t) \\
&= \mathbb{E}\left(\mathbb{E}\left(Z_t \sigma_t | \underline{X}_{t-1}\right)\right) \\
&= \mathbb{E}\left(\sigma_t \mathbb{E}\left(Z_t | \underline{X}_{t-1}\right)\right) \\
&= \mathbb{E}(\sigma_t \times 0) \\
&= 0 \quad \square
\end{aligned}$$

Proof (cont'd): Consider an ARCH(1) model such that

$$X_t = Z_t \sigma_t$$

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

Compute the unconditional variance of X_t. Since $\mathbb{E}(X_t) = 0$, we have

$$\mathbb{V}(X_t) = \mathbb{E}\left(X_t^2\right)$$

We know that X_t^2 has an AR(1) representation with

$$X_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + v_t \Longleftrightarrow \Phi(L) X_t^2 = \alpha_0 + v_t$$

with $\Phi(L) = 1 - \alpha_1$. Then, we have

$$\mathbb{V}(X_t) = \mathbb{E}\left(X_t^2\right) = \alpha_0 \Phi(1)^{-1} = \frac{\alpha_0}{1-\alpha_1} \quad \square$$

Consequences

1. An ARCH(1) process is unconditionally **homoscedastic**

   Unconditional variance $\mathbb{V}\left(X_t\right) = \frac{\alpha_0}{1-\alpha_1} = \text{constant } \forall t$

2. An ARCH(1) process is conditionally **heteroscedastic**

   Conditional variance $\mathbb{V}\left(X_t \middle| \mathcal{F}_{t-1}\right) = \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$ varies with $\mathcal{F}_{t-1}$

Consequences (cont'd)

1. If $\alpha_1 < 1$, the two first moments of an ARCH(1) process are finite and constant

$$\mathbb{E}\left(X_t\right) = 0 \qquad \gamma\left(0\right) = \mathbb{V}\left(X_t\right) = \frac{\alpha_0}{1-\alpha_1}$$

$$\gamma\left(k\right) = \mathbb{C}ov\left(X_t, X_{t-k}\right) = 0 \quad \text{for } k \neq 0$$

   and, the ARCH(1) process is (weakly) **stationary**.

2. For more details on the stationarity conditions of an ARCH/GARCH process, see Francq and Zakoian (2010).

Reminder Chapter 1. Stylized Fact 1 (Stationarity)

Fact (stationarity)

In general, the prices are non-stationary whereas the returns are ***stationary****.*

Definition

Property 4: *If* $\{X_t, t \in \mathbb{Z}\}$ *is an ARCH(1) process with Gaussian innovations* $Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, *then, its conditional and unconditional* **fourth moments** *are equal to*

$$\mathbb{E}\left(X_t^4 | \underline{X}_{t-1}\right) = 3\left(\alpha_0 + \alpha_1 X_{t-1}^2\right)^2 = 3\sigma_t^4$$

$$\mathbb{E}\left(X_t^4\right) = \frac{3\alpha_0^2\left(1+\alpha_1\right)}{\left(1-3\alpha_1^2\right)\left(1-\alpha_1\right)}$$

Corollary

If $\{X_t, t \in \mathbb{Z}\}$ *is an ARCH(1) process with Gaussian innovations* $Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, *then, its unconditional and conditional* **Kurtosis coefficients** *are equal to*

$$\mathbb{K}\left(X_t^4 | \underline{X}_{t-1}\right) = \frac{\mathbb{E}\left(X_t^4 | \underline{X}_{t-1}\right)}{\left(\mathbb{V}\left(X_t | \underline{X}_{t-1}\right)\right)^2} = 3$$

$$\mathbb{K}(X_t) = \frac{\mathbb{E}\left(X_t^4\right)}{\left(\mathbb{V}(X_t)\right)^2} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3 \quad \textit{if } \alpha_1^2 < 1/3$$

2.1. Properties of ARCH models

Innovation		**ARCH model**		**Output:** X_t
$Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$	$\Longrightarrow$	$X_t = Z_t \sigma_t$	$\Longrightarrow$	$\mathbb{K}(X_t) > 3$ if $\alpha_1^2 < 1/3$
$\mathbb{K}(Z_t) = 3$		$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$		**leptokurtic** distribution
mesokurtic distribution				

Reminder Chapter 1 Stylized Fact 3 (Heavy tails)

Fact (heavy tails)

The return distribution often exhibits **heavier tails** *than those of a normal distribution.*

Consequences

1. Even if the innovation Z_t has a normal distribution, the **marginal distribution** of X_t is not Gaussian since

$$X_t \sim \text{unknown distribution with}$$

$$\mathbb{E}\left(X_t\right) = 0 \quad \mathbb{V}\left(X_t\right) = \frac{\alpha_0}{1-\alpha_1} \quad \mathbb{S}\left(X_t\right) = 0 \quad \mathbb{K}\left(X_t\right) = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3 \text{ if } \alpha_1^2 < 1/3$$

2. If the innovation Z_t has a normal distribution, the **conditional distribution** of X_t is Gaussian

$$X_t | \underline{X}_{t-1} \sim \mathcal{N}\left(0, \sigma_t^2\right)$$

$$\mathbb{E}\left(X_t | \underline{X}_{t-1}\right) = 0 \quad \mathbb{V}\left(X_t | \underline{X}_{t-1}\right) = \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

$$\mathbb{S}\left(X_t | \underline{X}_{t-1}\right) = 0 \quad \mathbb{K}\left(X_t | \underline{X}_{t-1}\right) = 3$$

Proof: Consider an ARCH(1) model such that

$$X_t = Z_t \sigma_t$$

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

So, we have

$$\begin{aligned} \mathbb{E}\left(X_t^4 | \underline{X}_{t-1}\right) &= \mathbb{E}\left(Z_t^4 \sigma_t^4 | \underline{X}_{t-1}\right) \\ &= \mathbb{E}\left(Z_t^4 | \underline{X}_{t-1}\right) \sigma_t^4 \\ &= \mathbb{E}\left(Z_t^4\right) \left(\sigma_t^2\right)^2 \end{aligned}$$

If $Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, $\mathbb{K}(Z_t) = \mathbb{E}\left(Z_t^4\right)/1^2 = 3$, or equivalently $\mathbb{E}\left(Z_t^4\right) = 3$. So, we get

$$\mathbb{E}\left(X_t^4 | \underline{X}_{t-1}\right) = 3\sigma_t^4 \quad \square$$

2.1. Properties of ARCH models

Proof (cont'd): Then, we have

$$\begin{aligned}
\mathbb{E}\left(X_t^4\right) &= \mathbb{E}\left(\mathbb{E}\left(X_t^4 | \underline{X}_{t-1}\right)\right) \\
&= 3\,\mathbb{E}\left(\left(\alpha_0 + \alpha_1 X_{t-1}^2\right)^2\right) \\
&= 3\left(\alpha_0^2 + 2\alpha_0\alpha_1 \mathbb{E}\left(X_{t-1}^2\right) + \alpha_1^2 \mathbb{E}\left(X_{t-1}^4\right)\right) \\
&= 3\left(\alpha_0^2 + \frac{2\alpha_0^2\alpha_1}{1-\alpha_1} + \alpha_1^2 \mathbb{E}\left(X_{t-1}^4\right)\right) \\
&= 3\alpha_0^2\left(\frac{1+\alpha_1}{1-\alpha_1}\right) + 3\alpha_1^2 \mathbb{E}\left(X_{t-1}^4\right)
\end{aligned}$$

If X_t is fourth-order stationary, then $\mathbb{E}\left(X_t^4\right) = \mathbb{E}\left(X_{t-1}^4\right)$ and we get

$$\mathbb{E}\left(X_t^4\right) = \frac{3\alpha_0^2\left(1+\alpha_1\right)}{\left(1-3\alpha_1^2\right)\left(1-\alpha_1\right)} \quad \square$$

Proof (cont'd): We have

$$\mathbb{E}\left(X_t^4 | \underline{X}_{t-1}\right) = 3\sigma_t^4 \qquad \mathbb{V}\left(X_t | \underline{X}_{t-1}\right) = \sigma_t^2$$

The conditional Kurtosis coefficient is equal to

$$\mathbb{K}\left(X_t^4 | \underline{X}_{t-1}\right) = \frac{\mathbb{E}\left(X_t^4 | \underline{X}_{t-1}\right)}{\left(\mathbb{V}\left(X_t | \underline{X}_{t-1}\right)\right)^2} = \frac{3\sigma_t^4}{\sigma_t^4} = 3 \quad \square$$

The conditional distribution is **mesokurtic**.

Proof (cont'd): We have

$$\mathbb{E}\left(X_t^4\right) = \frac{3\alpha_0^2\left(1+\alpha_1\right)}{\left(1-3\alpha_1^2\right)\left(1-\alpha_1\right)} \qquad \mathbb{V}\left(X_t\right) = \frac{\alpha_0}{1-\alpha_1}$$

The unconditional Kurtosis coefficient is equal to

$$\mathbb{K}\left(X_t^4\right) = \frac{\mathbb{E}\left(X_t^4\right)}{\left(\mathbb{V}\left(X_t\right)\right)^2} = \frac{3\alpha_0^2\left(1+\alpha_1\right)}{\left(1-3\alpha_1^2\right)\left(1-\alpha_1\right)}\frac{\left(1-\alpha_1\right)^2}{\alpha_0^2} = 3\left(\frac{1-\alpha_1^2}{1-3\alpha_1^2}\right) > 3 \quad \square$$

The Kurtosis is finite and positive as soon as $\alpha_1^2 < 1/3$. Moreover, the conditional distribution is **leptokurtic**.

2.1. Properties of ARCH models

Summary *If* $\{X_t, t \in \mathbb{Z}\}$ *is an ARCH(1) process with* **Gaussian innovations**, *then*

	Property	Consequences / Interpretation
P1	X_t^2 is an AR(1) process	ARCH effect: $\mathbb{C}ov\left(X_t^2, X_{t-k}^2\right) \neq 0$ for "small" k
P2	X_t is a martingale difference	$\mathbb{E}\left(X_t \mid \underline{X}_{t-1}\right) = 0$ and $\mathbb{C}ov\left(X_t, X_{t-k}\right) = 0 \ \forall k \neq 0$
P3	$\mathbb{E}\left(X_t\right) = 0, \ \mathbb{V}\left(X_t\right) = \frac{\alpha_0}{1-\alpha_1}$ $\mathbb{V}\left(X_t \mid \underline{X}_{t-1}\right) = \sigma_t^2$	$\{X_t\}$ is stationary, unconditionally homoscedastic, and conditionally heteroscedastic
P4	$\mathbb{K}\left(X_t\right) > 3$	The ARCH model generates leptokurtosis
	$\mathbb{K}\left(X_t \mid \underline{X}_{t-1}\right) = 3$	The **marginal** distribution of X_t is not Gaussian
		The **conditional** distribution of X_t is Gaussian

2.1. Properties of ARCH models

The properties of the ARCH(1) allows to capture most of the stylized facts of financial data (cf. Chapter 1)

1. **The returns are stationary**
2. **Absence of autocorrelations**
3. **Heavy tails**
4. **Asymmetry**
5. **Volatility clustering**
6. **Aggregational Gaussianity**
7. **ARCH effect**
8. **Leverage effect**

2.1. Properties of ARCH models

The properties of the ARCH(1) allows to capture most of the stylized facts of financial data (cf. Chapter 1)

1. **The returns are stationary** $=>$ X_t is stationary
2. **Absence of autocorrelations** $=>$ X_t is a martingale difference
3. **Heavy tails** $=>$ $\mathbb{K}(X_t)$ may be larger than 3 given the value of α_1
4. **Asymmetry**
5. **Volatility clustering** $=>$ $\mathbb{C}ov\left(X_t^2, X_{t-k}^2\right) \neq 0$
6. **Aggregational Gaussianity** $=>$ The marginal distribution of X_t is not normal
7. **ARCH effect** $=>$ X_t^2 has an AR(1) representation and $\mathbb{C}ov\left(X_t^2, X_{t-k}^2\right) \neq 0$
8. **Leverage effect**

Reminder: Chapter 1

Example (Intel Corp)

In order to illustrate some of these stylized facts, we consider a sample of 1,259 daily prices and (log-) returns for the S&P500 index (ticker: SPY) from August 19, 2013 to August 17, 2018 (5 years). The data are available in Data_SP500.xlsx.

2.1. Properties of ARCH models

Figure: Daily returns for the S&P500 index are stationary

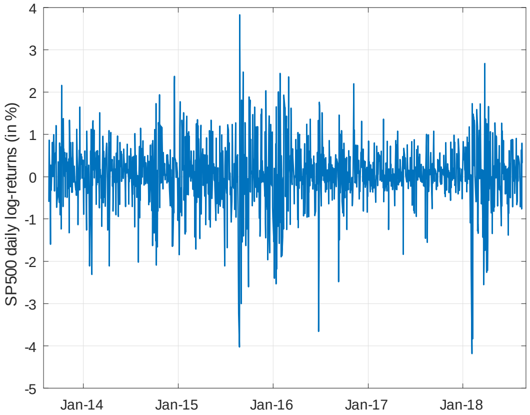


Figure: ACF for the S&P500 returns

Date: 09/09/18 Time: 15:19
Sample: 8/19/2013 8/17/2018
Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.019	-0.019	0.4455	0.504
		2	-0.054	-0.054	4.1067	0.128
		3	0.024	0.022	4.8464	0.183
		4	-0.043	-0.045	7.1488	0.128
		5	-0.028	-0.027	8.1122	0.150
		6	-0.008	-0.014	8.1869	0.225
		7	0.013	0.011	8.3965	0.299
		8	-0.037	-0.039	10.165	0.254
		9	-0.054	-0.057	13.887	0.126
		10	-0.019	-0.028	14.324	0.159
		11	0.002	-0.003	14.330	0.215
		12	-0.001	-0.004	14.330	0.280
		13	-0.015	-0.022	14.617	0.332
		14	-0.020	-0.027	15.102	0.371
		15	-0.073	-0.079	21.926	0.110

2.1. Properties of ARCH models

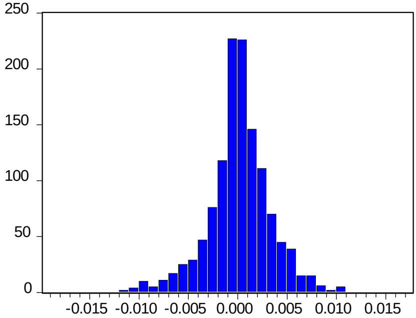
Figure: ACF for the S&P500 squared returns (August 19, 2013 to August 17, 2018)

Date: 09/09/18 Time: 11:25
Sample: 8/19/2013 8/17/2018
Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.302	0.302	114.94	0.000
		2	0.251	0.176	194.81	0.000
		3	0.288	0.196	299.82	0.000
		4	0.251	0.116	379.50	0.000
		5	0.115	-0.052	396.30	0.000
		6	0.135	0.018	419.45	0.000
		7	0.095	-0.020	430.83	0.000
		8	0.108	0.044	445.63	0.000
		9	0.102	0.043	458.93	0.000
		10	0.123	0.060	478.14	0.000
		11	0.084	0.003	487.21	0.000
		12	0.050	-0.041	490.35	0.000
		13	0.048	-0.020	493.24	0.000
		14	0.040	-0.012	495.29	0.000
		15	0.031	0.010	496.51	0.000

2.1. Properties of ARCH models

Figure: Descriptive statistics for the daily returns of the S&P500 index



Series: SP500_RETURNS
Sample 8/20/2013 6/15/2018
Observations 1259

Mean	0.000189
Median	0.000231
Maximum	0.016630
Minimum	-0.018172
Std. Dev.	0.003390
Skewness	-0.566934
Kurtosis	6.405986
Jarque-Bera	675.9989
Probability	0.000000

Weaknesses of ARCH Models

Tsay (2002) identifies three main limits of the ARCH models.

1. The model assumes that positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks. In practice, the return of a financial asset responds differently to positive and negative shocks.
2. The ARCH model is rather restrictive. For instance, the fourth moment $\mathbb{E}\left(X_t^4\right)$ exists only if $\alpha_1^2 < 1/3$.
3. The ARCH model does not provide any insight for understanding the source of volatility. It only provides a *mechanical way* to describe the behavior of the conditional variance. It gives no indication about what causes such behavior to occur.

2.1. Properties of ARCH models

Key Concepts

If $\{X_t, t \in \mathbb{Z}\}$ *has an ARCH(1) representation with Gaussian innovations, then*

1. X_t^2 has an AR(1) representation
2. X_t is a martingale difference
3. X_t is a stationary process under some conditions on the parameters
4. X_t is (unconditionally) homoscedastic
5. X_t is conditionally heteroscedastic
6. The (marginal) distribution of X_t is leptokurtic
7. The conditional distribution of X_t is normal

Sub-Section 2.2

Building an ARCH Model

Objectives

1. To introduce the ARCH model **of order p** or ARCH(p)
2. To introduce a conditional mean model with **ARCH errors**
3. To **estimate** the ARCH model parameters
4. To **check the validity** of an ARCH model
5. To compute **volatility forecasts** with ARCH model

2.2. Building an ARCH model

Definition (ARCH(q))

The process $\{X_t, t \in \mathbb{Z}\}$ is said to be an **ARCH**(p) process, if

$$X_t = Z_t \sigma_t$$

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and σ_t is a non-negative process such that

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

with $\alpha_0 > 0$, $\alpha_i \in \mathbb{R}, \forall i < p$, $\alpha_p \in \mathbb{R}^*$, and $\sum_{i=1}^{p} \alpha_i < 1$.

2.2. Building an ARCH model

Example (ARCH(2))

The process $\{X_t, t \in \mathbb{Z}\}$ is an **ARCH(3)** if

$$X_t = Z_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \alpha_2 X_{t-2}^2$$

where Z_t is an IID noise with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$.

Example (ARCH(3))

The process $\{Y_t, t \in \mathbb{Z}\}$ is an **ARCH(3)** if

$$Y_t = Z_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2 + \alpha_2 Y_{t-2}^2 + \alpha_3 Y_{t-3}^2$$

where Z_t is an IID noise with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$.

ARMA model with ARCH errors

In general, the structure of a **(conditional) volatility model** can be described as:

$$R_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = \sigma_t Z_t$$

where Z_t is a sequence of i.i.d. random variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

$$\mu_t \equiv \mathbb{E}(R_t | \mathcal{F}_{t-1}) = \mu_t(\underline{R}_{t-1}; \theta)$$

$$\sigma_t^2 \equiv \mathbb{V}(R_t | \mathcal{F}_{t-1}) = \sigma_t^2(\underline{R}_{t-1}; \theta)$$

where θ denotes the **set of parameters** for the conditional mean and variance and μ_t is typically an **ARMA**-type model .

2.2. Building an ARCH model

Denote by R_t the **daily return** of an asset or a portfolio at time t.

$$R_t = \underbrace{\mu_t}_{\text{cond mean model}} + \underbrace{\varepsilon_t}_{\text{innovation (martingale diff)}}$$

$$\varepsilon_t = \sigma_t Z_t$$

$$\sigma_t^2 = \underbrace{\alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2}_{\text{cond variance model}}$$

2.2. Building an ARCH model

Example (ARCH(1))

The process $\{R_t, t \in \mathbb{Z}\}$ has an **ARCH(1)** representation if

$$R_t = \phi_0 + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

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

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$. We have

$$\mu_t = \mathbb{E}(R_t | \mathcal{F}_{t-1}) = \phi_0$$

$$\sigma_t^2 = \mathbb{V}(R_t | \mathcal{F}_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

and $\theta = (\phi_0, \alpha_0, \alpha_1)'$ is the vector of parameters to estimate.

Example (AR(1)-ARCH(2))

The process $\{R_t, t \in \mathbb{Z}\}$ has an **AR(1)-ARCH(2)** representation if

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2$$

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$. We have

$$\mu_t = \mathbb{E}(R_t | \mathcal{F}_{t-1}) = \phi_0 + \phi_1 R_{t-1}$$

$$\sigma_t^2 = \mathbb{V}(R_t | \mathcal{F}_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2$$

and $\theta = (\phi_0, \phi_1, \alpha_0, \alpha_1, \alpha_2)'$ is the vector of parameters to estimate.

2.2. Building an ARCH model

Example (ARMA(1, 1)-ARCH(1))

The process $\{R_t, t \in \mathbb{Z}\}$ has an **ARMA(1,1)-ARCH(1)** representation if

$$R_t = \phi_0 + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

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

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$. We have

$$\mu_t = \mathbb{E}(R_t | \mathcal{F}_{t-1}) = \phi_0 + \phi_1 R_{t-1}$$

$$\sigma_t^2 = \mathbb{V}(R_t | \mathcal{F}_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

and $\theta = (\phi_0, \phi_1, \theta_1, \alpha_0, \alpha_1)'$ is the vector of parameters to estimate.

2.2. Building an ARCH model

Remark: do not make a confusion between the constant term of the conditional mean equation and the constant term of the conditional variance equation

$$R_t = \underbrace{\phi_0}_{\text{constant term}} + \phi_1 R_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \underbrace{\alpha_0}_{\text{constant term}} + \alpha_1 \varepsilon_{t-1}^2$$

2.2. Building an ARCH model

Estimation: The parameters θ are estimated by

1. **Maximum Likelihood (ML)** when one puts a distributional assumption on the innovations term Z_t.

2. **Quasi Maximum Likelihood (QML)** when the distribution of Z_t is unknown. The QML only assumes that the true (unknown) distribution of Z_t belongs to a given family (typically the exponential family).

3. In most of the statistical software, the model parameters are estimated by ML and the normality assumption is considered by default.

$$Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$$

Gouriéroux, C., Monfort, A., and A. Trognon (1984). Pseudo Maximum Likelihood Methods: Theory. *Econometrica* 52:681–700

2.2. Building an ARCH model

Remarks

1. For more details about the ML estimation method
   http://www.univ-orleans.fr/deg/masters/ESA/CH/Chapter2_MLE.pdf
   http://www.univ-orleans.fr/deg/masters/ESA/CH/Chapter2_Exercises.pdf

2. For more details about the ML estimation of AR model parameters

Francq C. and J.M Zakoian (2004), Maximum Likelihood Estimation of Pure GARCH and ARMA-GARCH Processes. *Bernoulli*, 10, 605-637.

Francq, C. and J.M. Zakoian (2010), *GARCH Models: Structure, Statistical Inference and Financial Applications*, Wiley.

2.2. Building an ARCH model

Figure: Estimation results, AR(1)-ARCH(2) model, US GDP annual growth rate (1961-2017)

Dependent Variable: Y_US
Method: ML - ARCH
Date: 11/02/18 Time: 22:44
Sample(adjusted): 1962 2017
Included observations: 56 after adjusting endpoints
Convergence achieved after 15 iterations

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	2.941298	0.410801	7.159914	0.0000
AR(1)	0.367916	0.152417	2.413879	0.0158
Variance Equation				
C	1.330471	0.622486	2.137351	0.0326
ARCH(1)	0.315348	0.251548	1.253627	0.2100
ARCH(2)	0.370306	0.257393	1.438682	0.1502

R-squared	0.117311	Mean dependent var	3.110150
Adjusted R-squared	0.048080	S.D. dependent var	2.071902
S.E. of regression	2.021480	Akaike info criterion	4.173308
Sum squared resid	208.4054	Schwarz criterion	4.354143
Log likelihood	-111.8526	F-statistic	1.694492
Durbin-Watson stat	1.909427	Prob(F-statistic)	0.165664
Inverted AR Roots	.37		

2.2. Building an ARCH model

Figure: Estimation results for an AR(1)-ARCH(1) model, Intel Corp. (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML ARCH - Normal distribution (Marquardt / EViews legacy)
Date: 11/05/18 Time: 11:15
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 15 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(3) + C(4)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000273	0.000250	1.092443	0.2746
AR(1)	-0.058922	0.050497	-1.166840	0.2433
		Variance Equation		
C	3.56E-05	1.16E-06	30.72748	0.0000
RESID(-1)^2	0.211148	0.056371	3.745674	0.0002

R-squared	0.005142	Mean dependent var	0.000250
Adjusted R-squared	0.003164	S.D. dependent var	0.006591
S.E. of regression	0.006580	Akaike info criterion	-7.229207
Sum squared resid	0.021779	Schwarz criterion	-7.195745
Log likelihood	1829.375	Hannan-Quinn criter.	-7.216082
Durbin-Watson stat	2.032375		
Inverted AR Roots	-.06		

Model Checking

- For an ARCH model, the **standardized innovations**

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

are i.i.d. random variates (following either a standard normal or Student-t distribution).

- Therefore, one can check the adequacy of a fitted ARCH model by examining the series of **standardized residuals**

$$\widehat{Z}_t = \frac{\widehat{\varepsilon}_t}{\widehat{\sigma}_t}$$

Model Checking (cont'd)

Tsay (2002) recommend three types of tests on the series $\{\widehat{z}_t\}_{t=1}^{T}$.

1. The Ljung–Box Q-statistics of $\widehat{z}_t$ can be used to check the adequacy of the mean equation.
2. The Ljung–Box Q-statistics of $\widehat{z}_t^2$ can be used to check the adequacy of the volatility equation.
3. The skewness, kurtosis, and QQ-plot of $\widehat{z}_t$ can be used to check the validity of the distribution assumption on Z_t.

2.2. Building an ARCH model

Forecasting

We have to distinguish:

1. The forecasts on the series R_t itself (typically the returns).
2. The forecasts on the volatility (or the variance) of R_t.

Example (forecasting model)

Consider the following AR(1)-ARCH(1) process

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

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

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$.

Forecasting

Fact (Forecasting of the series R_t)

The **best linear forecast** *of R_t given the information set $\mathcal{F}_{t-1}$ will be no different with or without an ARCH error because the process ε_t is a martingale difference.*

$$\widehat{R}_{t|t-1} = \mathbb{E}\left(R_t | \mathcal{F}_{t-1} \right)$$

Example (forecasting)

If the process $\{R_t, t \in \mathbb{Z}\}$ has an AR(1)-ARCH(1) representation

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t \qquad Z_t \text{ i.i.d.}\,(0,1)$$

then, the **forecast** of R_{t+1} given $\mathcal{F}_t$ is given by

$$\widehat{R}_{t+1|t} = \mathbb{E}\left(R_{t+1} \middle| \mathcal{F}_t \right) = \phi_0 + \phi_1 R_t$$

since the process ε_t is a martingale difference.

$$\mathbb{E}\left(\varepsilon_{t+1} \middle| \mathcal{F}_t \right) = 0$$

2.2. Building an ARCH model

Volatility forecasting

Definition (conditional variance forecast)

The **conditional variance forecast** at the horizon h is defined by

$$\widehat{\sigma}^2_{t+h|t} = \mathbb{V}\left(\left. R_{t+h} \right| \mathcal{F}_t \right)$$

For the forecast horizon $h = 1$, we have

$$\widehat{\sigma}^2_{t+1|t} = \mathbb{V}\left(\left. R_{t+1} \right| \mathcal{F}_t \right) = \sigma^2_{t+1}$$

Note: The volatility forecast is defined as

$$\widehat{\sigma}_{t+h|t} = \sqrt{\mathbb{V}\left(\left. R_{t+h} \right| \mathcal{F}_t \right)}$$

Definition (conditional variance forecast)

If the process $\{R_t, t \in \mathbb{Z}\}$ has an AR(1)-ARCH(1) representation

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t \qquad Z_t \text{ i.i.d.}\,(0,1)$$

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

then, the **conditional variance forecast** at horizon $h = 1$ is given by

$$\widehat{\sigma}^2_{t+1|t} = \mathbb{V}\left(R_{t+1} \middle| \mathcal{F}_t \right) = \alpha_0 + \alpha_1 \varepsilon_t^2 = \alpha_0 + \alpha_1 \left(R_t - \phi_0 - \phi_1 R_{t-1} \right)^2$$

Example

Using the following estimation results for an AR(1)-ARCH(1) model and the Intel Corp. daily returns, compute the **conditional variance forecast** for the 7/24/2018 given that

$$R_{7/23/2018} = -0.002661 \qquad R_{7/20/2018} = -0.005998$$

Dependent Variable: RETURNS
Method: ML ARCH - Normal distribution (Marquardt / EViews legacy)
Date: 11/05/18 Time: 11:15
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 15 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(3) + C(4)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000273	0.000250	1.092443	0.2746
AR(1)	-0.058922	0.050497	-1.166840	0.2433
		Variance Equation		
C	3.56E-05	1.16E-06	30.72748	0.0000
RESID(-1)^2	0.211148	0.056371	3.745674	0.0002

R-squared	0.005142	Mean dependent var	0.000250
Adjusted R-squared	0.003164	S.D. dependent var	0.006591
S.E. of regression	0.006580	Akaike info criterion	-7.229207
Sum squared resid	0.021779	Schwarz criterion	-7.195745
Log likelihood	1829.375	Hannan-Quinn criter.	-7.216082
Durbin-Watson stat	2.032375		

Inverted AR Roots	-.06

Solution

	Day	R_t
$t = 1$	7/20/2018	-0.005998
$t = 2$	7/23/2018	-0.002661
$t = 3$	7/24/2018	—

Dependent Variable: RETURNS
Method: ML ARCH - Normal distribution (Marquardt / EViews legacy)
Date: 11/05/18 Time: 11:15
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 15 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(3) + C(4)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000273	0.000250	1.092443	0.2746
AR(1)	-0.058922	0.050497	-1.166840	0.2433
Variance Equation				
C	3.56E-05	1.16E-06	30.72748	0.0000
RESID(-1)^2	0.211148	0.056371	3.745674	0.0002

R-squared	0.005142	Mean dependent var	0.000250
Adjusted R-squared	0.003164	S.D. dependent var	0.006591
S.E. of regression	0.006580	Akaike info criterion	-7.229207
Sum squared resid	0.021779	Schwarz criterion	-7.195745
Log likelihood	1829.375	Hannan-Quinn criter.	-7.216082
Durbin-Watson stat	2.032375		

Inverted AR Roots	-.06

$$
\begin{aligned}
\widehat{\sigma}^2_{3|2} &= \mathbb{V}\left(R_3 | \mathcal{F}_2 \right) \\
&= 3.56e^{-05} + 0.211148 \times \left(R_2 - 0.000273 - 0.058922 R_1 \right)^2 \\
&= 3.56e^{-05} + 0.211148 \times \left(-0.002661 - 0.000273 - 0.058922 \times 0.005998 \right)^2 \\
&= 3.7882e^{-05}
\end{aligned}
$$

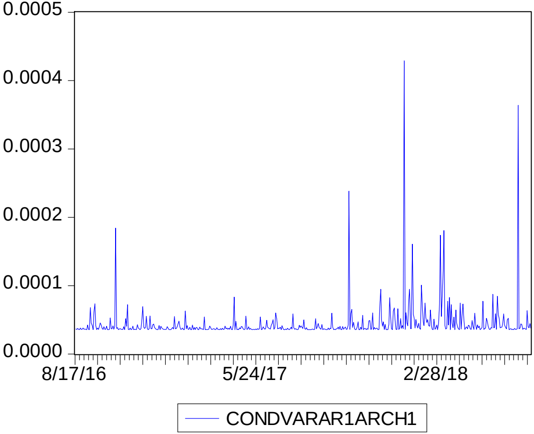
2.2. Building an ARCH model

Figure: Conditional variance forecasts, Intel Corp.

obs	RETURNS	CONDVARAR
6/13/18	0.004123	4.23E-05
6/14/18	0.002505	3.92E-05
6/15/18	-0.008068	3.68E-05
6/18/18	0.009647	4.98E-05
6/19/18	-0.001080	5.22E-05
6/20/18	-0.001750	3.57E-05
6/21/18	-0.002176	3.65E-05
6/22/18	-0.000252	3.70E-05
6/25/18	0.002178	3.57E-05
6/26/18	-0.000585	3.63E-05
6/27/18	0.003334	3.57E-05
6/28/18	-0.001081	3.75E-05
6/29/18	0.002076	3.58E-05
7/02/18	-0.002242	3.62E-05
7/03/18	-0.039001	3.68E-05
7/04/18	9.11E-05	0.000364
7/05/18	0.003718	3.69E-05
7/06/18	0.006364	3.80E-05
7/09/18	0.005921	4.39E-05
7/10/18	0.001315	4.32E-05
7/11/18	-0.002897	3.60E-05
7/12/18	0.003509	3.76E-05
7/13/18	0.002266	3.75E-05
7/16/18	0.001562	3.66E-05
7/17/18	-0.011320	3.60E-05
7/18/18	-0.003571	6.36E-05
7/19/18	-0.002968	3.99E-05
7/20/18	-0.005998	3.81E-05
7/23/18	-0.002662	4.44E-05
7/24/18	-0.000645	3.79E-05

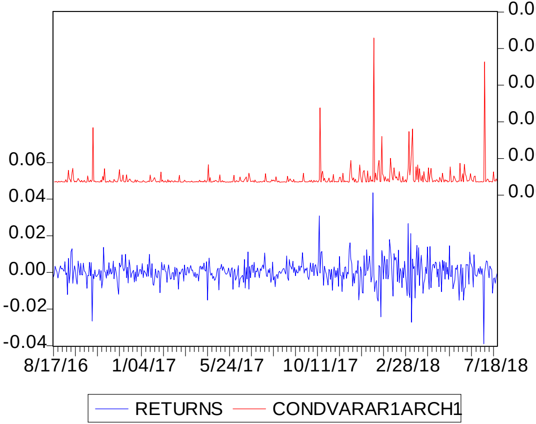
2.2. Building an ARCH model

Figure: Estimated conditional variance AR(1)-GARCH(1) model, Intel Corp. daily returns (8/17/2016 to 7/24/2018)



2.2. Building an ARCH model

Figure: Estimated conditional variance and observed returns, AR(1)-GARCH(1) model, Intel Corp. daily returns (8/17/2016 to 7/24/2018)



Definition (conditional variance forecasts)

If the process $\{R_t, t \in \mathbb{Z}\}$ has an AR(1)-ARCH(1) representation

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t \qquad Z_t \text{ i.i.d. } (0,1)$$

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

then, the **conditional variance forecast** at any horizon $h \geq 1$ is given by

$$\widehat{\sigma}^2_{t+h|t} = \left(\frac{\phi_0}{1-\alpha_1} \right) \left(\left(\frac{1-\phi_1^{2h}}{1-\phi_1^2} \right) - \alpha_1 \left(\frac{\alpha_1^h - \phi_1^{2h}}{\alpha_1 - \phi_1^2} \right) \right) + \alpha_1 \left(\frac{\alpha_1^h - \phi_1^{2h}}{\alpha_1 - \phi_1^2} \right) \varepsilon_t^2$$

2.2. Building an ARCH model

Key Concepts

1. ARCH(q) model
2. Conditional mean and conditional variance models
3. Estimation methods for ARCH model parameters
4. Model checking
5. (Conditional) volatility forecasting

Section 3

GARCH Models

3. GARCH models

Objectives

1. To introduce the **GARCH** model
2. To present the main **properties** of GARCH processes
3. To discuss the choice of the **conditional distributions**
4. To present the parameter **estimation**
5. To compute a volatility **forecast**

3. GARCH models

Due to the large persistence in volatility, ARCH models often require a large p to fit the data. A more **parsimonious** specification is provided by **GARCH** models.

GARCH = **G**eneralized **A**uto**R**egressive **C**onditional **H**eteroskedasticity

Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307–327

Notations

- Denote R_t the **daily return** of an asset or a portfolio at time t.

- Consider a (conditional mean) model with an **ARCH/GARCH error** for the process $\{R_t, t \in \mathbb{Z}\}$ given by:

$$R_t = \mathbb{E}\left(R_t | \mathcal{F}_{t-1}\right) + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}\left(Z_t\right) = 0$ and $\mathbb{V}\left(Z_t\right) = 1$, and σ_t^2 is the **conditional variance** of ε_t and R_t.

$$\sigma_t^2 = \mathbb{V}\left(\varepsilon_t | \mathcal{F}_{t-1}\right) = \mathbb{V}\left(R_t | \mathcal{F}_{t-1}\right)$$

- Denote by $\mu_t = \mathbb{E}\left(R_t | \mathcal{F}_{t-1}\right)$ the **conditional mean** of R_t.

Definition (GARCH model)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is said to be a **GARCH**(p, q) process, if

$$\varepsilon_t = Z_t \sigma_t$$

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

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

with $\omega > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$ and $\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i < 1$.

3. GARCH models

The conditional variance of a GARCH(p, q) depends on:

- The first p lag of the ε_t^2 (e.g., the squared error terms).
- The first q lag of the conditional variance σ_t^2.

$$\sigma_t^2 = \omega + \underbrace{\sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2}_{\text{ARCH component}} + \underbrace{\sum_{i=1}^{q} \beta_i \sigma_{t-i}^2}_{\text{GARCH component}}$$

- The parameters α_i are often called the **ARCH parameters**.
- The parameters β_i are often called the **GARCH parameters**.

3. GARCH models

Example (GARCH model)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ has a **GARCH(2,1)** representation if

$$\varepsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \beta_1 \sigma_{t-1}^2$$

where Z_t is i.i.d. with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$.

Example (GARCH model)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ has a **GARCH(1,2)** representation if

$$\varepsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2$$

where Z_t is i.i.d. with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$.

Fact (GARCH lag orders)

From a practical point, **GARCH(1,1)** *specifications are generally sufficient to capture the dynamics of the conditional variance and higher-order lags are not required.*

Figure: Estimation results for a GARCH(2,1) model, Intel Corp. (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 22:14
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 20 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000147	0.000252	0.583453	0.5596
Variance Equation				
C	7.09E-07	1.45E-07	4.887164	0.0000
ARCH(1)	0.176223	0.055091	3.198780	0.0014
ARCH(2)	-0.141629	0.052286	-2.708747	0.0068
GARCH(1)	0.951795	0.006419	148.2774	0.0000

R-squared	-0.000246	Mean dependent var	0.000250
Adjusted R-squared	-0.008248	S.D. dependent var	0.006591
S.E. of regression	0.006618	Akaike info criterion	-7.314493
Sum squared resid	0.021897	Schwarz criterion	-7.272666
Log likelihood	1851.910	Durbin-Watson stat	2.145515

Note: the ARCH parameters α_1 and α_2 cannot be negative.

Figure: Estimation results for a GARCH(1,2) model, Intel Corp. (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 22:15
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 17 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000179	0.000267	0.669737	0.5030
Variance Equation				
C	1.36E-06	6.19E-07	2.194689	0.0282
ARCH(1)	0.047376	0.026505	1.787474	0.0739
GARCH(1)	0.496865	0.757230	0.656161	0.5117
GARCH(2)	0.423764	0.720050	0.588520	0.5562

R-squared	-0.000118	Mean dependent var	0.000250
Adjusted R-squared	-0.008119	S.D. dependent var	0.006591
S.E. of regression	0.006617	Akaike info criterion	-7.308236
Sum squared resid	0.021894	Schwarz criterion	-7.266409
Log likelihood	1850.330	Durbin-Watson stat	2.145789

Figure: Estimation results for a GARCH(1,1) model, Intel Corp. (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 22:08
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 25 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000188	0.000268	0.701286	0.4831
Variance Equation				
C	6.20E-07	9.43E-08	6.570948	0.0000
ARCH(1)	0.028547	0.004738	6.025243	0.0000
GARCH(1)	0.960202	0.004786	200.6078	0.0000

R-squared	-0.000089	Mean dependent var	0.000250
Adjusted R-squared	-0.006078	S.D. dependent var	0.006591
S.E. of regression	0.006611	Akaike info criterion	-7.315541
Sum squared resid	0.021893	Schwarz criterion	-7.282079
Log likelihood	1851.174	Durbin-Watson stat	2.145852

Definition (GARCH(1,1) model)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is said to be a **GARCH**$(1,1)$ process, if

$$\varepsilon_t = Z_t \sigma_t$$

where Z_t is a sequence of i.i.d. variables with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

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

with $\omega > 0$, $\alpha \geq 0, \beta \geq 0$ and $\alpha + \beta < 1$.

Remarks

- The conditional variance $\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$ depends on two effects:
  1. An **intrinsic persistence** effect through the first lag of the conditional variance.
  2. An **extrinsic persistence** effect.
- Following a positive (or negative) shock at time $t-1$, the conditional variance at time t increases (impact effect) and thus it has an impact on $\varepsilon_t = Z_t \sigma_t$.

$$\text{shock } z_{t-1} > 0 \Longrightarrow \varepsilon_{t-1} \uparrow \Longrightarrow \sigma_t \uparrow ....$$

- Starting from the next period (i.e., at time t), the effect of the shock (at time $t-1$) on the conditional variance at $t+1$ (and thus on ε_{t+1}) passes through the conditional variance at time t (intrinsic persistence effect).

$$.... \Longrightarrow \sigma_t \uparrow \Longrightarrow \sigma_{t+1}^2 \uparrow$$

- The overall impact of a shock can be decomposed into a "**contemporaneous effect**", which depends on α, and a "**persistence effect**", which depends on β.

Remarks

One often observes that:

1. The sum of the estimates of α and β are generally close (but below 1).
2. The estimate of β is generally greater than the one of α.
3. The estimate of β is generally larger than 0.90 for daily returns and the estimate of α is below 0.1.

Be careful: it is not a general rule, just an observation.

3. GARCH models

Figure: GARCH(1,1) model, Intel Corp. daily returns (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 22:08
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 25 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000188	0.000268	0.701286	0.4831
		Variance Equation		
C	6.20E-07	9.43E-08	6.570948	0.0000
ARCH(1)	0.028547	0.004738	6.025243	0.0000
GARCH(1)	0.960202	0.004786	200.6078	0.0000

R-squared	-0.000089	Mean dependent var	0.000250
Adjusted R-squared	-0.006078	S.D. dependent var	0.006591
S.E. of regression	0.006611	Akaike info criterion	-7.315541
Sum squared resid	0.021893	Schwarz criterion	-7.282079
Log likelihood	1851.174	Durbin-Watson stat	2.145852

3. GARCH models

The main **properties** of a GARCH process are similar to those of an ARCH process.

1. ε_t^2 has an ARMA representation
2. ε_t is a martingale difference
3. ε_t is a stationary process under some conditions on the parameters α and β
4. ε_t is (unconditionally) homoscedastic
5. ε_t is conditionally heteroscedastic
6. The (marginal) distributions of ε_t and R_t are leptokurtic
7. If Z_t has a normal distribution, the conditional distributions of ε_t and R_t are normal

Theorem (ARMA representation)

If $\{\varepsilon_t, t \in \mathbb{Z}\}$ has a GARCH(p, q) representation, with

$$\varepsilon_t = Z_t \sigma_t$$

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

then $\{\varepsilon_t^2, t \in \mathbb{Z}\}$ has an **ARMA**$(\max(p, q), q)$ **representation**, *with*

$$\varepsilon_t^2 = \omega + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + v_t - \sum_{i=1}^{q} \beta_i v_{t-i}$$

where $v_t = \varepsilon_t^2 - h_t$ is an innovation process, i.e. $\mathbb{E}(v_t | \mathcal{F}_{t-1}) = 0$.

Definition

if $\{\varepsilon_t, t \in \mathbb{Z}\}$ *is a GARCH(1) process with*

$$\varepsilon_t = Z_t \sigma_t$$

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

then, its two first **unconditional moments** *are equal to*

$$\mathbb{E}(\varepsilon_t) = 0 \qquad \mathbb{V}(\varepsilon_t) = \frac{\omega}{1 - \alpha - \beta}$$

with $\omega > 0$, $\alpha \geq 0, \beta \geq 0$ and $\alpha + \beta < 1$.

Other properties

1. Since $\mathbb{V}(\varepsilon_t) = \omega / (1 - \alpha - \beta)$, the GARCH(1,1) is sometimes written as

$$\varepsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \mathbb{V}(\varepsilon_t)(1 - \alpha - \beta) + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

2. Bollerslev (1986) shows that the **Kurtosis coefficient** of a GARCH(1,1) is equal to

$$\mathbb{K}(\varepsilon_t) = \frac{\mathbb{E}\left(\varepsilon_t^4\right)}{\mathbb{E}\left(\varepsilon_t^2\right)^2} = \frac{3\left(1 - (\alpha + \beta)^2\right)}{1 - (\alpha + \beta)^2 - 2\alpha^2}$$

as soon as $(\alpha + \beta)^2 + 2\alpha^2 < 1$.

Estimation:

1. The set of parameters θ of an ARMA-GARCH model is estimated by **Maximum Likelihood (ML)** or Quasi maximum Likelihood (QML).

2. When the model is estimated by ML, the most often used distributions for Z_t are:

   1. The **normal distribution**, $Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. **IMPORTANT**: the normality assumption on Z_t does not imply that the return R_t has a normal (marginal) distribution.

   2. The **Student t-distribution**, $Z_t \overset{i.i.d.}{\sim} t(v)$, which is symmetric and leptokurtic (if v is "small").

   3. The **skewed Student t-distribution**, $Z_t \overset{i.i.d.}{\sim} Skewed\ t(\delta, v)$, which is asymmetric (if $\delta \neq 1$) and leptokurtic (if v is "small").

   4. The **Generalized Error Distribution (GED)**, $Z_t \overset{i.i.d.}{\sim} GED(v)$, which is symmetric and leptokurtic (if $v < 2$).

3. GARCH models

Example (GARCH model with Gaussian innovations)

The process $\{R_t, t \in \mathbb{Z}\}$ has a GARCH(1,1) representation with **Gaussian innovations** if

$$R_t = c + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

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

$$Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$$

The set of parameters to be estimated is $\theta = (c, \omega, \alpha, \beta)$. Notice that R_t has a conditional normal distribution

$$R_t | \mathcal{F}_{t-1} \sim \mathcal{N}\left(c, \sigma_t^2\right)$$

but a marginal distribution which is not Gaussian

$$R_t \sim \text{leptokurtic distribution}$$

Example (GARCH model with Student innovations)

The process $\{R_t, t \in \mathbb{Z}\}$ has a GARCH(1,1) representation with **Student innovations** if

$$R_t = c + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

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

$$Z_t \overset{i.i.d.}{\sim} t(v)$$

where $t(v)$ is a standardized Student t-distribution with v degrees of freedom. The set of parameters to be estimated is now $\theta = (c, \omega, \alpha, \beta, v)$.

Why considering non-Gaussian distributions for the innovation Z_t?

1. The use of a **leptokurtic distribution** for Z_t allows to increase the kurtosis of R_t.

$$\begin{aligned} \text{kurtosis of a GARCH process} \quad &= \quad \text{kurtosis generated by the model (dynamics)} \\ &\quad + \text{kurtosis of the innovation } Z_t \end{aligned}$$

   In order to reproduce the level of kurtosis of the financial returns, the kurtosis generated by the model is not sufficient. That is why, we generally consider a leptokurtic distribution for Z_t: Student, GED, etc.

2. The use of an **a skewed distribution** for Z_t allows to reproduce the skewness observed in the distribution of the financial returns.

$$\text{skewed distribution for } Z_t \Longrightarrow \text{skewed distribution for } R_t$$

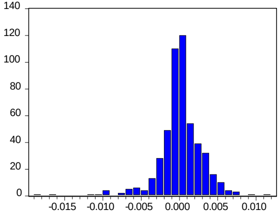
3. GARCH models

Reminder Chapter 1. Stylized Fact 4 (Asymmetry)

Fact (asymmetry)

The distribution of return is **asymmetric** *and often* **negatively skewed**, *reflecting the fact that the downturns of financial markets are often much steeper than the recoveries. Investors tend to react more strongly to negative news than to positive news*

Figure: Daily returns S&P500 (8/17/2016 to 7/24/2018)



Series: RETURNS
Sample 8/17/2016 7/24/2018
Observations 505

Mean	0.000231
Median	0.000253
Maximum	0.011637
Minimum	-0.018172
Std. Dev.	0.002895
Skewness	-1.166176
Kurtosis	9.903015
Jarque-Bera	1117.133
Probability	0.000000

Figure: GARCH(1,1) model with Gaussian innovations, Intel Corp. (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML ARCH - Normal distribution (Marquardt / EViews legacy)
Date: 11/05/18 Time: 08:56
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 14 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000192	0.000270	0.712549	0.4761
Variance Equation				
C	6.38E-07	9.65E-08	6.606469	0.0000
RESID(-1)^2	0.028572	0.004822	5.925374	0.0000
GARCH(-1)	0.959825	0.004906	195.6437	0.0000

R-squared	-0.000078	Mean dependent var	0.000250
Adjusted R-squared	-0.000078	S.D. dependent var	0.006591
S.E. of regression	0.006591	Akaike info criterion	-7.315561
Sum squared resid	0.021893	Schwarz criterion	-7.282099
Log likelihood	1851.179	Hannan-Quinn criter.	-7.302436
Durbin-Watson stat	2.145875		

3. GARCH models

Figure: GARCH(1,1) model with Student innovations, Intel Corp. (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML ARCH - Student's t distribution (Marquardt / EViews legacy)
Date: 11/05/18 Time: 08:50
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 17 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000387	0.000190	2.036120	0.0417
		Variance Equation		
C	5.62E-07	3.38E-07	1.664750	0.0960
RESID(-1)^2	0.050711	0.021913	2.314185	0.0207
GARCH(-1)	0.944133	0.019231	49.09416	0.0000
T-DIST. DOF	3.328842	0.497571	6.690185	0.0000

R-squared	-0.000430	Mean dependent var	0.000250
Adjusted R-squared	-0.000430	S.D. dependent var	0.006591
S.E. of regression	0.006592	Akaike info criterion	-7.579864
Sum squared resid	0.021901	Schwarz criterion	-7.538037
Log likelihood	1918.916	Hannan-Quinn criter.	-7.563458
Durbin-Watson stat	2.145121		

3. GARCH models

Figure: GARCH(1,1) model with GED innovations, Intel Corp. (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML ARCH - Generalized error distribution (GED) (Marquardt / EViews legacy)
Date: 11/05/18 Time: 08:54
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 18 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000355	0.000173	2.056731	0.0397
		Variance Equation		
C	5.14E-07	2.48E-07	2.071835	0.0383
RESID(-1)^2	0.036673	0.014939	2.454783	0.0141
GARCH(-1)	0.952122	0.016155	58.93793	0.0000
GED PARAMETER	1.011089	0.062464	16.18668	0.0000

R-squared	-0.000256	Mean dependent var	0.000250
Adjusted R-squared	-0.000256	S.D. dependent var	0.006591
S.E. of regression	0.006591	Akaike info criterion	-7.541885
Sum squared resid	0.021897	Schwarz criterion	-7.500057
Log likelihood	1909.326	Hannan-Quinn criter.	-7.525479
Durbin-Watson stat	2.145494		

Forecasting

Definition (conditional variance forecast)

If the process $\{R_t, t \in \mathbb{Z}\}$ has a GARCH(1,1) representation

$$R_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

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

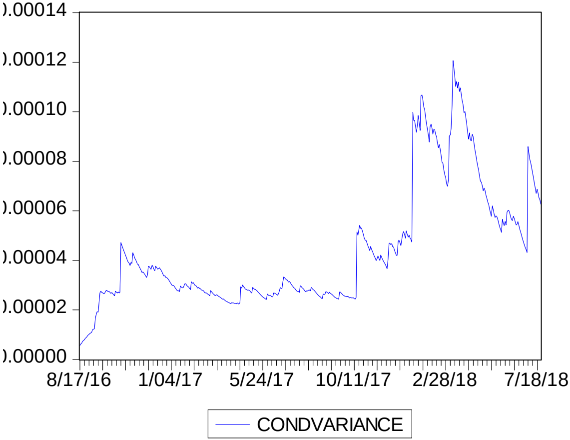
the **conditional variance forecast** at time t for an horizon $h = 1$ is defined by the recurrence relation

$$\widehat{\sigma}_{t+1|t}^2 = \mathbb{V}\left(R_{t+1} \middle| \mathcal{F}_t \right) = \sigma_{t+1}^2 = \omega + \alpha \left(R_t - \mu_t \right)^2 + \beta \sigma_t^2$$

where σ_1^2 is fixed.

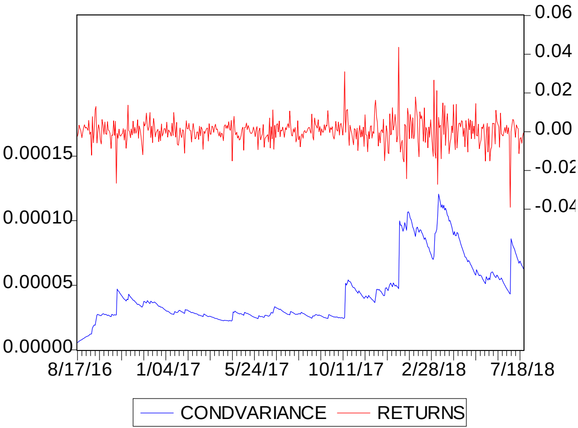
3. GARCH models

Figure: Estimated (in sample) conditional variance, GARCH(1,1) model with Gaussian innovations, Intel Corp.daily returns (8/17/2016 to 7/24/2018)



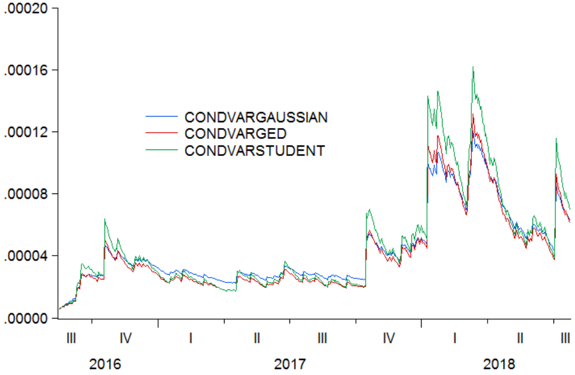
3. GARCH models

Figure: Estimated conditional variance and observed returns, GARCH(1,1) model with Gaussian innovations, Intel Corp. dialy returns (8/17/2016 to 7/24/2018)



3. GARCH models

Figure: Estimated conditional variance for various conditional distributions, Intel Corp. daily returns (8/17/2016 to 7/24/2018)



Key Concepts

1. GARCH(p, q) model
2. GARCH$(1, 1)$ model
3. ARCH and GARCH parameters
4. Properties of a GARCH process
5. ARMA representation associated to a GARCH process
6. Conditional distributions (normal, Student, skewed Student, GED, etc.)
7. Parameter estimation
8. Volatility forecasting

Section 4

Extensions of GARCH Models

Objectives

1. To introduce the **IGARCH** model
2. To introduce the **GARCH-M** model
3. To introduce the **asymmetric** GARCH models
4. To establish a link between these models and the **leverage effect**
5. To introduce the **GJR-GARCH** model
6. To introduce the **TGARCH** model
7. To introduce the **EGARCH** model

Overview

Some relevant extensions of the GARCH model have been proposed in order to accommodate particular features of financial series (asymmetry, leverage effect, etc.).

Among others, GARCH models have been refined by introducing:

1. **Asymmetric** responses to negative and positive innovations to handle the observed asymmetry in the reaction of conditional volatility to the arrivals of news.
2. **Persistence**.
3. **Long-memory** (the dependency for a large number of lags).

Overview

The following GARCH models are often encountered in the empirical financial literature as well as in the industry:

1. Asymmetric GARCH models: Exponential GARCH model (EGARCH), Threshold GARCH model (TGARCH), GJR model;

2. Integrated GARCH model (IGARCH).

3. Long-memory GARCH model (LMGARCH) or Fractionally integrated GARCH model (FIGARCH).

4. Extensions of GARCH models

We will focus on the following models:

1. **IGARCH** model
2. **GARCH-M** model
3. **GJR-GARCH** model
4. **TGARCH** model
5. **EGARCH** model

4. Extensions of GARCH models

We will focus on the following models:

1. **IGARCH** model
2. **GARCH-M** model
3. **GJR-GARCH** model
4. **TGARCH** model
5. **EGARCH** model

Definition (IGARCH)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is an **Integrated**-GARCH(1,1) process or **IGARCH**$(1,1)$, if

$$\varepsilon_t = Z_t \sigma_t$$

where Z_t is i.i.d. with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

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

with $\omega > 0$ and $\alpha \in [0, 1[$.

4. Extensions of GARCH models

Remarks

1. The impact of past squared shocks $\eta_{\tau-k} = \varepsilon_{t-k}^2 - \sigma_{t-k}^2$ on X_t^2 in the ARMA representation is **persistent**.

2. The unconditional variance of ε_t, hence that of R_t, is **not defined** under a IGARCH(1, 1) model.

3. A special case of the IGARCH(1,1) is the **RiskMetrics** volatility model defined as

$$\sigma_t^2 = \lambda \varepsilon_{t-1}^2 + (1-\lambda)\,\sigma_{t-1}^2$$

which is a model used to compute the Value-at-Risk.

For more details

Nelson, D. B. (1990), Stationarity and persistence in the GARCH(1, 1) model, *Econometric Theory*, 6, 318–334.

4. Extensions of GARCH models

Figure: IGARCH(1,1) model, S&P500 daily returns (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML ARCH - Normal distribution (Marquardt / EViews legacy)
Date: 11/05/18 Time: 09:07
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 11 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(2)*RESID(-1)^2 + (1 - C(2))*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000496	4.94E-05	10.03791	0.0000
Variance Equation				
RESID(-1)^2	0.073433	0.004180	17.56655	0.0000
GARCH(-1)	0.926567	0.004180	221.6533	0.0000

R-squared	-0.008401	Mean dependent var	0.000231
Adjusted R-squared	-0.008401	S.D. dependent var	0.002895
S.E. of regression	0.002907	Akaike info criterion	-8.954779
Sum squared resid	0.004260	Schwarz criterion	-8.938048
Log likelihood	2263.082	Hannan-Quinn criter.	-8.948216
Durbin-Watson stat	2.148913		

4. Extensions of GARCH models

We will focus on the following models:

1. **IGARCH** model
2. **GARCH-M** model
3. **GJR-GARCH** model
4. **TGARCH** model
5. **EGARCH** model

GARCH-M model

- The return of a security may depend on its volatility.
- To model such a phenomenon, one may consider the GARCH-M model, where "M" stands for GARCH **in mean**.
- The GARCH-M has been introduced by Engle, Lilien and Robbins (1987).

Engle, R., Lilien, D., and R. Robins (1987). Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model. *Econometrica*, 55(2), 391-407.

Definition (GARCH-M)

An example of **GARCH-M(1,1)** is given by

$$R_t = c + \delta\sigma_t^2 + \varepsilon_t$$

$$\varepsilon_t = \sigma_t Z_t$$

where Z_t is i.i.d. with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

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

with $\omega > 0$, $\alpha \geq 0, \beta \geq 0$ and $\delta \in \mathbb{R}$.

Interpretation

- The parameter δ is called the risk premium parameter.
- A positive δ indicates that the return is positively related to its past volatility.

Alternative GARCH-M specifications

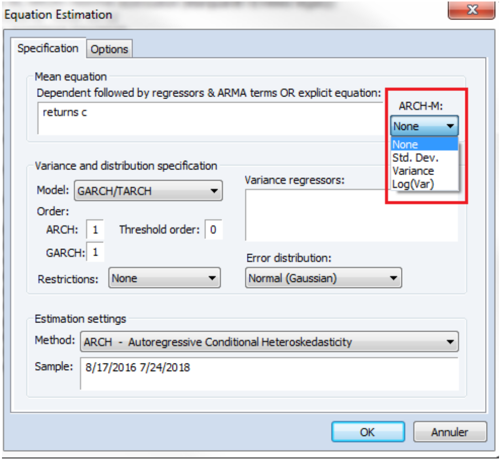
$$R_t = c + \delta \sigma_t^2 + \varepsilon_t \qquad \text{Conditional variance effect}$$

$$R_t = c + \delta \sigma_t + \varepsilon_t \qquad \text{Conditional volatility effect}$$

$$R_t = c + \delta \ln \left(\sigma_t^2 \right) + \varepsilon_t \qquad \text{Log-linear specification}$$

4. Extensions of GARCH models

Figure: Examples of GARCH-M specifications available in Eviews 9



4. Extensions of GARCH models

Figure: GARCH-M (volatility specification), S&P500 daily returns (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 20:50
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 15 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
SQR(GARCH)	0.228092	0.205836	1.108123	0.2678
C	-0.000183	0.000489	-0.374348	0.7081
Variance Equation				
C	8.34E-07	1.70E-07	4.909816	0.0000
ARCH(1)	0.168322	0.029165	5.771471	0.0000
GARCH(1)	0.722121	0.044882	16.08929	0.0000

R-squared	-0.003518	Mean dependent var	0.000231
Adjusted R-squared	-0.011546	S.D. dependent var	0.002895
S.E. of regression	0.002912	Akaike info criterion	-9.115285
Sum squared resid	0.004239	Schwarz criterion	-9.073458
Log likelihood	2306.610	Durbin-Watson stat	2.126942

Figure: GARCH-M (variance specification), S&P500 daily returns (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 20:54
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 17 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	25.21145	29.78123	0.846555	0.3972
C	0.000206	0.000201	1.026419	0.3047
Variance Equation				
C	8.16E-07	1.67E-07	4.879708	0.0000
ARCH(1)	0.162550	0.028669	5.669930	0.0000
GARCH(1)	0.729257	0.044119	16.52919	0.0000

R-squared	0.000041	Mean dependent var	0.000231
Adjusted R-squared	-0.007959	S.D. dependent var	0.002895
S.E. of regression	0.002907	Akaike info criterion	-9.114192
Sum squared resid	0.004224	Schwarz criterion	-9.072364
Log likelihood	2306.333	F-statistic	0.005141
Durbin-Watson stat	2.126117	Prob(F-statistic)	0.999947

4. Extensions of GARCH models

We will focus on the following models:

1. **IGARCH** model
2. **GARCH-M** model
3. **GJR-GARCH** model
4. **TGARCH** model
5. **EGARCH** model

Asymmetric GARCH models

- The GARCH model assumes that positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks.
- In practice, the return of a financial asset responds differently to positive and negative shocks.
- The GARCH model does not allow to capture the **leverage effect**.

Stylized Fact 8: Leverage effect (reminder Chapter 1)

Fact (leverage effect)

Asset returns are negatively correlated with the changes of their volatilities: this negative correlation is called **leverage effect***.*

- As asset prices decline, companies become more leveraged (debt to equity ratios increase) and riskier, and hence their stock prices become more volatile.
- On the other hand, when stock prices become more volatile, investors demand high returns and hence stock prices go down.

Asymmetric GARCH models

- The asymmetric GARCH models are designed to capture the **non linearities** of the conditional variance dynamics, including the leverage effect.
- Many asymmetric GARCH models have been proposed: GJR-GARCH, TGARCH, EGARCH, APARCH, VSGARCH, QGARCH, LSTGARCH, ANSTGARCH, etc.
- One of the most often used asymmetric models is the **GJR-GARCH** model, where "GJR" stands for Glosten, Jagannathan and Runkle (1993).

Glosten, L., Jagannathan, R., and D. Runkle, D. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*, 48(5), 1779-1801.

Definition (GJR-GARCH)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is to be a **GJR-GARCH**(1,1) process, if

$$\varepsilon_t = Z_t \sigma_t$$

where Z_t is i.i.d. with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \gamma \mathbb{I}_{(\varepsilon_{t-1}<0)}\, \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

with $\omega > 0$, $\alpha \geq 0, \beta \geq 0$, $\gamma \in \mathbb{R}$, and where $\mathbb{I}_{(.)}$ is the indicator function that takes a value 1 if the condition is true and 0 otherwise.

Interpretation

- The term ε_t can be interpreted as a shock (surprise) on the return, since

$$\varepsilon_t = R_t - \mu_t = R_t - \mathbb{E}\left(R_t | \mathcal{F}_{t-1}\right)$$

- In a GJR-GARCH model, the influence of the past return shock ε_{t-1} on the current conditional variance σ_t^2 depends on its sign

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \gamma \mathbb{I}_{\varepsilon_{t-1}<0}\, \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

$$\frac{\partial \sigma_t^2}{\partial \varepsilon_{t-1}^2} = \begin{cases} \alpha + \gamma & \text{if } \varepsilon_{t-1} < 0 \\ \alpha & \text{otherwise} \end{cases}$$

- A **leverage effect** implies that $\gamma > 0$, i.e. the increase in volatility caused by a negative return is larger than the appreciation due a positive return of the same magnitude.

Figure: GJR-GARCH, S&P500 daily returns (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 20:55
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 5 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000261	0.000101	2.585792	0.0097
Variance Equation				
C	8.35E-07	1.45E-07	5.775906	0.0000
ARCH(1)	-0.007291	0.028332	-0.257332	0.7969
(RESID<0)*ARCH(1)	0.220719	0.041923	5.264862	0.0000
GARCH(1)	0.750475	0.040774	18.40561	0.0000

R-squared	-0.000107	Mean dependent var	0.000231
Adjusted R-squared	-0.008108	S.D. dependent var	0.002895
S.E. of regression	0.002907	Akaike info criterion	-9.167772
Sum squared resid	0.004225	Schwarz criterion	-9.125945
Log likelihood	2319.862	Durbin-Watson stat	2.166735

4. Extensions of GARCH models

We will focus on the following models:

1. **IGARCH** model
2. **GARCH-M** model
3. **GJR-GARCH** model
4. **TGARCH** model
5. **EGARCH** model

TGARCH model

- The TGARCH, where "T" stands for **Threshold**, is an asymmetric GARCH model designed to capture the leverage effect.
- The TGARCH is similar to the GJR model, different only because of the use of the **conditional volatility**, instead of the variance, in the specification.
- The TGARCH has been introduced by Zakoian (1994).

Rabemananjara R. and J.M. Zakoian (1993), Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics*, 8, 31-49.

Zakoian J.M. (1994), Threshold Heteroskedastic Models. *Journal of Economic Dynamic and Control*, 18, 931-955, 1994.

Definition (TGARCH)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is to be a **TGARCH**(1,1) process, if

$$\varepsilon_t = Z_t \sigma_t$$

where Z_t is i.i.d. with $\mathbb{E}(Z_t) = 0$ and $\mathbb{V}(Z_t) = 1$, and

$$\sqrt{\sigma_t^2} = \omega + \alpha_+ \varepsilon_{t-1} \mathbb{I}_{(\varepsilon_{t-1} \geq 0)} + \alpha_- \varepsilon_{t-1} \mathbb{I}_{(\varepsilon_{t-1} < 0)} + \beta \sqrt{\sigma_{t-1}^2}$$

with $(\omega, \alpha_+, \alpha_-, \beta) \in \mathbb{R}^4$ and $\mathbb{I}_{(.)}$ the indicator function.

TGARCH model

- One advantage of the TGARCH is that it does not require any **positivity constraints** on the parameters, since we have $\forall\left(\omega,\alpha_{+},\alpha_{-},\beta\right)\in\mathbb{R}^4$

$$\sigma_t^2=\left(\omega+\alpha_{+}\varepsilon_{t-1}\mathbb{I}_{(\varepsilon_{t-1}\geq 0)}+\alpha_{-}\varepsilon_{t-1}\mathbb{I}_{(\varepsilon_{t-1}<0)}+\beta\sqrt{\sigma_{t-1}^2}\right)\geq 0$$

- The TGARCH allows to capture an **asymmetry** between positive and negative shocks, as

$$\frac{\partial\sigma_t}{\partial\varepsilon_{t-1}}=\begin{cases}\alpha_{-} & \text{if } \varepsilon_{t-1}<0\\ \alpha_{+} & \text{otherwise}\end{cases}$$

- The **leverage effect** implies that $|\alpha_{-}|>|\alpha_{+}|$, i.e. the increase in volatility caused by a negative return is larger than the appreciation due a positive return of the same magnitude.

4. Extensions of GARCH models

We will focus on the following models:

1. **IGARCH** model
2. **GARCH-M** model
3. **GJR-GARCH** model
4. **TGARCH** model
5. **EGARCH** model

EGARCH model

- The EGARCH, where "E" stands for **Exponential**, is an asymmetric GARCH model.
- The EGARCH is designed to capture both (1) the asymmetric effects between positive and negative shocks on the returns and (2) the effects of "big" shocks.
- The TGARCH has been introduced by Nelson (1991).

Nelson, D. B. (1991), "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica*, 59, 347–370.

Definition (EGARCH)

The process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is to be a **EGARCH**(1,1) process, if

$$\varepsilon_t = Z_t \sigma_t$$

$$\ln\left(\sigma_t^2\right) = \omega + \alpha Z_{t-1} + \gamma\left(|Z_{t-1}| - \mathbb{E}\left(|Z_{t-1}|\right)\right) + \beta \ln\left(\sigma_{t-1}^2\right)$$

with $(\omega, \alpha, \gamma, \beta) \in \mathbb{R}^4$ and where Z_t is i.i.d. with $\mathbb{E}\left(Z_t\right) = 0$ and $\mathbb{V}\left(Z_t\right) = 1$.

4. Extensions of GARCH models

GARCH model

$$\varepsilon_t = Z_t \sigma_t$$

$$\sigma_t^2 = \omega + \alpha \underbrace{\varepsilon_{t-1}^2}_{\text{depends on } \varepsilon_{t-1}} + \beta \sigma_{t-1}^2$$

EGARCH model

$$\varepsilon_t = Z_t \sigma_t$$

$$\ln\left(\sigma_t^2\right) = \omega + \underbrace{\alpha Z_{t-1} + \gamma\left(|Z_{t-1}| - \mathbb{E}\left(|Z_{t-1}|\right)\right)}_{\text{depends on the standardized error } Z_{t-1}} + \beta \ln\left(\sigma_{t-1}^2\right)$$

EGARCH model

- The EGARCH model does not require any **restriction** on the parameters because, since the equation is on log variance instead of variance itself, the positivity of the variance is automatically satisfied $\forall (\omega, \alpha, \gamma, \beta) \in \mathbb{R}^4$

$$\sigma_t^2 = \exp\left(\omega + \alpha Z_{t-1} + \gamma \left(|Z_{t-1}| - \mathbb{E}\left(|Z_{t-1}|\right)\right) + \beta \ln\left(\sigma_{t-1}^2\right)\right) > 0$$

- The EGARCH model captures the asymmetric effects between positive and negative shocks on the returns, since

$$\frac{\partial \ln\left(\sigma_t^2\right)}{\partial |Z_{t-1}|} = \begin{cases} \gamma - \alpha & \text{if } z_{t-1} < 0 \\ \gamma + \alpha & \text{otherwise} \end{cases}$$

- The **leverage effect**, i.e. the fact that negative shocks at time $t-1$ have a stronger impact on the variance at time t than positive shocks, implies that $\alpha < 0$.

EGARCH model

- The term $\left(|Z_{t-1}| - \mathbb{E}\left(|Z_{t-1}|\right)\right)$ measures the **magnitude** of the (positive or negative) shocks.
- If the parameter γ is positive, then the "**big**" (compared to their expected value) shocks have a stronger impact on the variance than the "**small**" shocks.
- The mean $\mathbb{E}\left(|Z_{t-1}|\right)$ is a constant that depends on the distribution of Z_t.

$$\mathbb{E}\left(|Z_t|\right) = \sqrt{\frac{2}{\pi}} \quad \text{Gaussian distribution}$$

$$\mathbb{E}\left(|Z_t|\right) = 2\frac{\Gamma\left(\frac{v}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\left(v-1\right)\Gamma\left(\frac{v}{2}\right)} \quad \text{Student } t\left(v\right) \text{ distribution}$$

where $\Gamma\left(x\right) = \int_0^\infty u^{x-1}e^{-u}du$ is the **gamma function**.

Example (EGARCH)

Consider a AR(1)-**EGARCH**(1,1) with Gaussian innovation for the returns $\{R_t, t \in \mathbb{Z}\}$

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t$$

$$\varepsilon_t = Z_t \sigma_t$$

$$\ln\left(\sigma_t^2\right) = \omega + \alpha Z_{t-1} + \gamma\left(|Z_{t-1}| - \sqrt{\frac{2}{\pi}}\right) + \beta \ln\left(\sigma_{t-1}^2\right)$$

or equivalently

$$\ln\left(\sigma_t^2\right) = \omega + \alpha\left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right) + \gamma\left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| - \sqrt{\frac{2}{\pi}}\right) + \beta \ln\left(\sigma_{t-1}^2\right)$$

with Z_t i.i.d. $\mathcal{N}(0,1)$. The vector of parameters to be estimated is

$$\theta = (\phi_0, \phi_1, \omega, \alpha, \gamma, \beta)'$$

4. Extensions of GARCH models

Figure: EGARCH, S&P500 daily returns (8/17/2016 to 7/24/2018)

Dependent Variable: RETURNS
Method: ML - ARCH
Date: 11/04/18 Time: 20:57
Sample: 8/17/2016 7/24/2018
Included observations: 505
Convergence achieved after 38 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000238	0.000108	2.214508	0.0268
Variance Equation				
C	-1.477426	0.251807	-5.867296	0.0000
|RES|/SQR[GARCH](1	0.214937	0.041854	5.135363	0.0000
RES/SQR[GARCH](1)	-0.238090	0.029121	-8.175833	0.0000
EGARCH(1)	0.890261	0.019755	45.06581	0.0000

R-squared	-0.000006	Mean dependent var	0.000231
Adjusted R-squared	-0.008006	S.D. dependent var	0.002895
S.E. of regression	0.002907	Akaike info criterion	-9.179873
Sum squared resid	0.004224	Schwarz criterion	-9.138046
Log likelihood	2322.918	Durbin-Watson stat	2.166954

4. Extensions of GARCH models

Key Concepts

1. IGARCH model
2. GARCH-M model
3. GJR-GARCH model
4. TGARCH model
5. EGARCH model

End of Chapter 6

Christophe Hurlin