Quantitative Analysis of Finance I

ECON90033

WEEK 6

AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTIC (ARCH) PROCESSES

ARCH AND GARCH MODELS OF CONDITIONAL VARIANCE

Reference:

HMPY: § 13.1-13.2

Dr László Kónya

January 2023

AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTIC (*ARCH*) PROCESSES

- In time series econometrics the basic question is whether the data has been generated by a weakly stationary stochastic processes, so it has a finite and constant *unconditional* mean and autocovariances that are not affected by a change of time origin.

  In financial econometrics, however, the *conditional* moments are occasionally more interesting than the unconditional ones.

  ⟵ For example, an asset is risky if its (log-)return r_t is volatile, i.e., changes a lot over time.

  In statistics volatility is measured by the variance, and investors wish to predict the variance from historical data, i.e., they are concerned with the conditional variance of the (log-) return,

$$\boxed{Var(r_t \mid r_{t-1}, r_{t-2}, \ldots)}$$

  Changing volatility is quite common to financial time series, especially to high frequency (weekly, daily, hourly etc.) data.

- A stylized fact about financial market is volatility clustering, meaning that a volatile period tends to be followed by another volatile period.

  Intuitively, volatility clustering occurs when some unexpected big news makes the market nervous, and it takes several periods for the market to fully digest the news.

  Statistically, volatility clustering implies time-varying conditional variance: big volatility (variance) today may lead to big volatility tomorrow.

Suppose, for example, that

$$y_t = z_{t-1}\varepsilon_t$$

where z_{t-1} is an observable independent variable and ε_t is a white-noise error term with variance σ^2.

The conditional variance of y_t given the $\Omega_{t-1} = \{z_{t-1}\}$ information set is

$$Var(y_t \mid z_{t-1}) = z_{t-1}^2 Var(\varepsilon_t) = z_{t-1}^2 \sigma^2$$

$\longrightarrow$ If the $\{z_t\}$ series is constant, $\{y_t\}$ is just some multiple of a white-noise process and its unconditional and conditional variances are also constant.

If, however, the $\{z_t\}$ series is positively autocorrelated, so is the conditional variance of $\{y_t\}$.

In this case the conditional variance is persistent and the $\{y_t\}$ series is expected to be characterised by periods of high or low volatility.

Ex 1:

Consider daily closing US dollar to Australian dollar exchange rate (*EXR*) from 16 May 2006 to 2 June 2023 (downloaded from https://finance.yahoo.com).

a) Plot *EXR.*

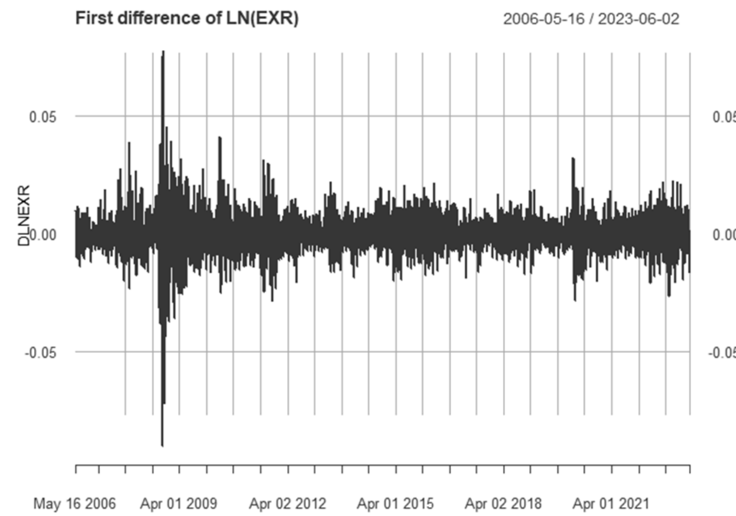
The data frequency is daily but there are many gaps and some of them are irregular. In cases like this, it is better to use *xts* objects instead of *ts* objects.

library(xts)
EXR = xts(Close, order.by = as.Date(Date))

plot.xts(EXR, xlab = "Date", ylab = "EXR", col = "darkgreen" , main = "USD to AUD Exchange Rate“)

EXR looks non-stationary, so let’s consider the first difference of its logarithm, which is the approximate rate of change of *EXR*.

```
plot.xts(diff(log(EXR), 1), xlab = "Date", ylab = "DLNEXR",
         main = "First difference of LN(EXR)", col = "red")
```



There is clearly no trend in this series, but there are periods of low and periods of high volatility.

This kind of behavior is best modelled as some basic or generalized *ARCH* process.

- What is (are) the specific factor(s) that might cause the conditional variance of y_t to change?

  In practice it is often difficult to find this (these) variables, but we may try to model the conditional mean and variance of y_t simultaneously by allowing the variance of ε_t to depend on its own history.

  Assume, for example, that

$$\boxed{y_t = \mu_t + \varepsilon_t} \qquad \boxed{\varepsilon_t : idN(0,\sigma^2)} \boxed{\;;\; \varepsilon_t \mid \Omega_t : idN(0,h_t)} \qquad \boxed{h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$$

Unconditional and conditional distributions, where *idN* stands for ‘independently normal’

Conditional variance

Given $\Omega_{t-1} = y_{t-1}$, the first equation provides the *conditional mean* of y_t,

$$\boxed{E(y_t \mid \Omega_{t-1}) = E_{t-1}(y_t) = \mu_t}$$

Some info set

The second equation specifies that the *conditional distribution* of ε_t is independently normal with time-varying conditional variance h_t.

Finally, the third equation defines the *conditional variance* of y_t,

$$Var(y_t \mid \Omega_{t-1}) = Var_{t-1}(y_t) = E_{t-1}\left[(y_t - \mu_t)^2\right] = E_{t-1}\left[\varepsilon_t^2\right] = Var_{t-1}(\varepsilon_t)$$
$$= h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

$\longrightarrow$ The common conditional variance of y_t and ε_t depends on ε^2_{t-1}.

Namely, given that $\alpha_1 > 0$, if ε_{t-1} is large (small) in absolute value, h_t is also large (small).

$\longrightarrow$ ε_t is conditionally heteroskedastic.

This error process is known as an autoregressive conditional heteroskedasticity process of order one, denoted as *ARCH*(1).

- In general, an *ARCH*(*q*) process is defined as (Engle, 1982),

$$y_t = \mu_t + \varepsilon_t \qquad \varepsilon_t \mid \Omega_t : idN(0, h_t) \qquad h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \ , \ \alpha_i \underset{i=0}{\overset{q}{\geq}} 0 \ , \ \sum_{i=1}^{q} \alpha_i < 1$$

Given these restrictions, the conditional variance is always non-negative.

Let $\eta_t \equiv \varepsilon_t^2 - h_t$

$$\longrightarrow \quad \varepsilon_t^2 = h_t + \eta_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \eta_t$$

This is an *AR*(*q*) process in ε_t^2.

Note: The second equation of *ARCH*(*q*)

$\varepsilon_t : idN(0, h_t)$ is equivalent to $\varepsilon_t = \nu_t \sqrt{h_t} \ , \ \nu_t : iidN(0,1)$

where *iidN* stands for 'identically and independently normal', and ν_t is independent of ε_{t-1}, and thus of h_t.

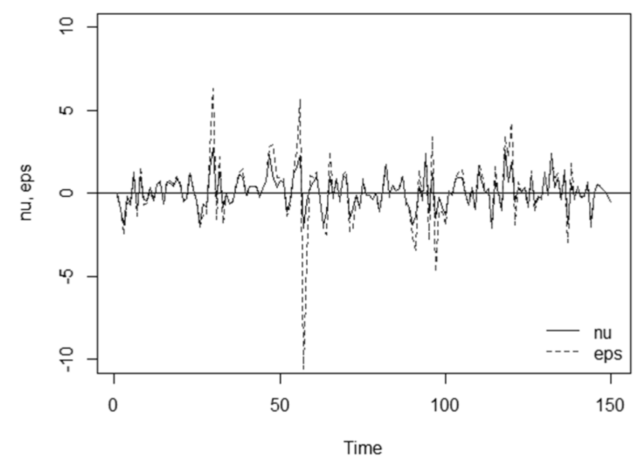
This alternative specification facilitates the simulation of *ARCH*(*q*) processes.

Ex 2:

a) Draw a set of $N(0,1)$ independent random numbers $\{\nu_t\}$ for $t = 1, ...,150$ and starting with $\varepsilon_1 = 0$ simulate and plot

$$\varepsilon_t = \nu_t \sqrt{1 + 0.8\varepsilon_{t-1}^2}$$

```
nu = ts(rnorm(150, mean = 0, sd = 1), start = 1)
eps = ts(start = 1, end = 150)
eps[1] = 0
for (t in 2:150)
  {eps[t] = ts(nu[t] * sqrt(1 + 0.8*eps[t-1]^2))}
plot.ts(nu, ylab = "nu, eps", ylim = c(-10, 10),
         col = "blue", lty = 1)
abline(h=0)
lines(eps, col = "red“, llty = 2)
legend("bottomright", bty = "n", legend = c("nu", "eps"),
          col = c("blue", "red"), lty = 1:2)
```



Both $\{\nu_t\}$ and $\{\varepsilon_t\}$ seem to fluctuate around zero and each unusually large (in absolute value) shock in ν_t is associated with relatively large volatility in $\{\varepsilon_t\}$.

b) Using $\{\varepsilon_t\}$ from part (a) and zero initial values, simulate and plot the following two stationary *AR*(1)-*ARCH*(1) processes,

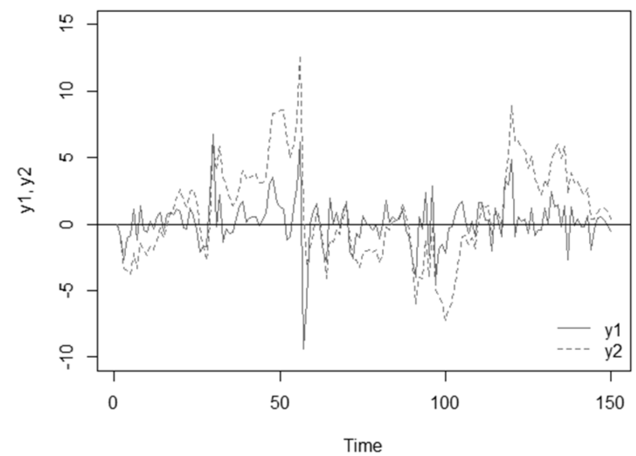
$$y_{1t} = 0.2 y_{1,t-1} + \varepsilon_t$$

```
y1 = ts(start = 1, end = 150)
y1[1] = 0
for (t in 2:150)
  {y1[t] = 0.2*y1[t-1] + eps[t]}
```

$$y_{2t} = 0.9 y_{2,t-1} + \varepsilon_t$$

```
y2 = ts(start = 1, end = 150)
y2[1] = 0
for (t in 2:150)
  {y2[t] = 0.9*y2[t-1] + eps[t]}
```

```
plot.ts(y1, ylab = "y1, y2", ylim = c(-10, 15),
        col = "pink4", lty = 1)
abline(h=0)
lines(y2, col = "seagreen", lty = 2)
legend("bottomright", bty = "n", legend = c("y1", "y2"),
   col = c("pink4", "seagreen"), lty = 1:2)
```

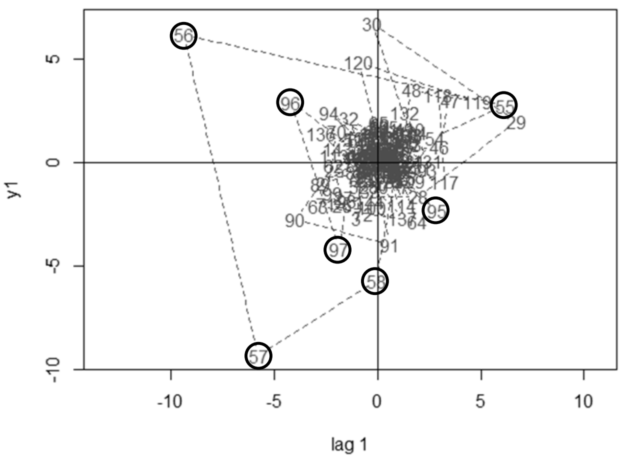


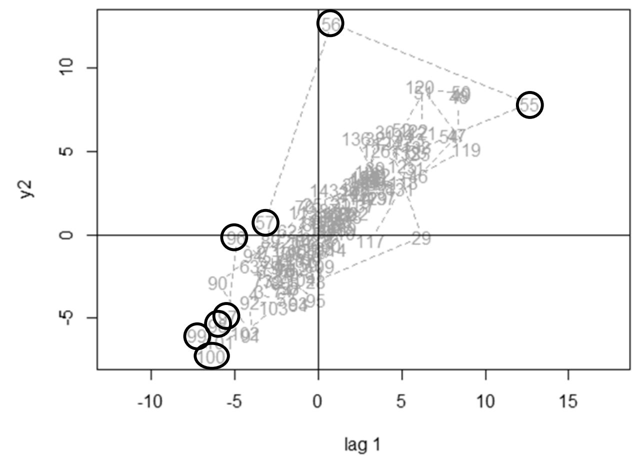
The *ARCH* error term results in changing volatility, and the bigger the *AR*(1) parameter, the more prominent any given change in y_t is.

c) In order to visualize the presence of sequences of outliers, plot y_t against y_{t-1} for both simulated series.

lag.plot(y1, lags = 1, diag = FALSE, col = "violetred", lty = 2)

lag.plot(y2, lags = 1, diag = FALSE, col = "orange2", lty = 2)





$\{y_{1t}\}$ and $\{y_{2t}\}$ have zero unconditional means, so sequences of outliers show up in series of points relatively far from the origin.

The cloud of data points around the origins demonstrate the tendency to revert to the unconditional means, zero.

The second plot also shows that $\{y_{2t}\}$ has positive first order autocorrelation.

So does $\{y_{1t}\}$ since both $AR(1)$ parameters are positive, but its autocorrelation is much weaker and thus hardly visible on the first scatter plot.

- Are *ARCH*(q) processes weakly stationary? (Yes.)

  To answer this question, we need to consider the unconditional first and second moments. To keep the manipulations simple, let's focus on *AR*1-*ARCH*(1) processes,

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \ , \ \varepsilon_t = \nu_t \sqrt{h_t} \ , \ h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

Granted that the process started sufficiently long time ago,

$$y_t = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \longrightarrow E(y_t) = \frac{a_0}{1-a_1}$$

$$Var(y_t) = \sum_{i=0}^{\infty} a_1^{2i} \underbrace{Var(\varepsilon_{t-i})}_{} = \frac{\alpha_0}{1-\alpha_1} \sum_{i=0}^{\infty} a_1^{2i} = \frac{\alpha_0}{1-\alpha_1} \frac{1}{1-a_1^2}$$

They are constant and so are the autocorrelations (not shown here).

$$Var(\varepsilon_t) = E\left(\varepsilon_t^2\right) = E\left(\nu_t^2\right) \times E\left(h_t\right)$$
$$= \alpha_0 + \alpha_1 E\left(\varepsilon_{t-1}^2\right) = \alpha_0 + \alpha_1 Var(\varepsilon_{t-1})$$

and the stationary solution is

$$Var(\varepsilon_t) = \frac{\alpha_0}{1-\alpha_1}$$

ARCH AND *GARCH* MODELS OF CONDITIONAL VARIANCE

- These models, called conditional volatility models, have three equations. The first and the third approximate the *conditional mean* dynamics and the *conditional variance* dynamics, while the second specifies the *conditional distribution* of the error variable.

  Depending on the specifications of the equations, these models are quite general, and they cover a wide range of possibilities.

For example, in the case of *ARCH*(q)

$$y_t = \mu_t + \varepsilon_t \qquad \varepsilon_t : idN(0, h_t) \qquad h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

If $\mu_t = a_0 + a_1 x_t$ and x_t is uncorrelated with ε_t, this model is a simple linear regression model with *ARCH*(q) errors. If $\mu_t = y_{t-1}$, this is an *AR*(1) model with *ARCH*(q) errors. If μ_t is the linear combination of lagged y_t's and ε_t's, this model is an *ARMA*(p,q) model with *ARCH*(q) errors.

- An extension of the *ARCH*(*q*) model is the generalized *ARCH* model, denoted as *GARCH*(*p*,*q*), that allows the conditional variance to be generated by an *ARMA* process (Bollersev, 1986).

$$\leftarrow \quad \boxed{y_t = \mu_t + \varepsilon_t} \qquad \boxed{\varepsilon_t : idN(0, h_t)}$$

$$\boxed{h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \;\; , \;\; \alpha_i \underset{i=0}{\overset{q}{\geq}} 0 \;\; , \;\; \beta_j \underset{j=1}{\overset{p}{\geq}} 0}$$

and $\boxed{\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1}$ to ensure positive but finite conditional variance and stationary volatility.

Assuming again that

$$\boxed{\eta_t = \varepsilon_t^2 - h_t} \longrightarrow \boxed{\varepsilon_t^2 = \alpha_0 + \underbrace{\sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2}_{ARCH} + \underbrace{\sum_{j=1}^{p} \beta_j h_{t-j}}_{GARCH} + \eta_t}$$

i.e., *q* is the order of the *ARCH* terms and *p* is the order of the *GARCH* terms.

The simplest version is a *GARCH*(1,1) error process:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad , \quad \alpha_1 > 0 \, , \beta_1 > 0 \, , \alpha_1 + \beta_1 < 1$$

$\longrightarrow$ ε_{t-1} has a delayed effect on h_t, and the larger α_1 is, the more pronounced this effect is. Moreover, the larger β_1 is, the more persistent h_t is.

<u>Note</u>:

a) Just like in the case of *ARCH*(*q*), the second equation of *GARCH*(*p,q*) can be rewritten as

$$\varepsilon_t = \nu_t \sqrt{h_t} \quad , \quad \nu_t : idN(0,1)$$

b) In general, under heteroskedasticity the usual estimated standard errors based on (among others) the homoskedasticity assumption are incorrect and a possible solution is to use heteroskedasticity and autocorrelation consistent (*HAC*) standard errors.

The objective of (*G*)*ARCH* modelling, however, is to capture volatility, so do not use *HAC* standard errors – model volatility instead!

Ex 3:

Using the same set of random numbers $\{\nu_t\}$ as in part (a) of Ex 2, simulate two stationary *GARCH*(1,1) error processes with $\varepsilon_1 = h_1 = 0$ and

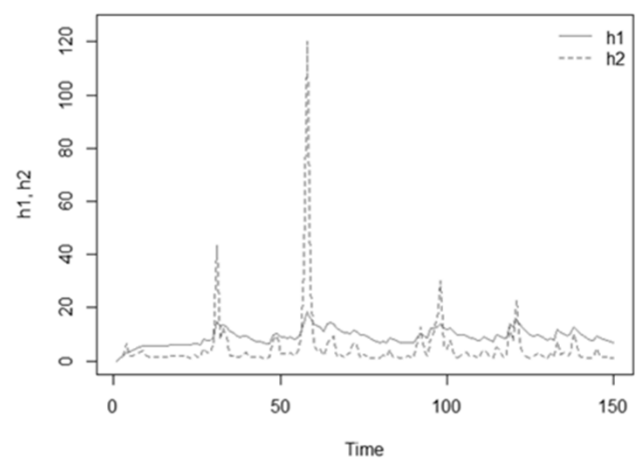
$$\varepsilon_{it} = \nu_t \sqrt{h_{it}} \quad (i = 1, 2)$$

$$h_{1t} = 1 + 0.1\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}$$

$$h_{2t} = 1 + 0.8\varepsilon_{2,t-1}^2 + 0.1h_{2,t-1}$$

```
eps1 = ts(start = 1, end = 150)
eps1[1] = 0
eps2 = ts(start = 1, end = 150)
eps2[1] = 0
h1 = ts(start = 1, end = 150)
h1[1] = 0
h2 = ts(start = 1, end = 150)
h2[1] = 0

for (t in 2:150)
  {h1[t] = ts(1 + 0.1*eps1[t-1]^2 + 0.8*h1[t-1])
   h2[t] = ts(1 + 0.8*eps2[t-1]^2 + 0.1*h2[t-1])
   eps1[t] = nu[t]*sqrt(h1[t])
   eps2[t] = nu[t]*sqrt(h2[t])}
```

```
plot.ts(h1, col = "steelblue", ylab = "h1, h2",
        ylim = c(0, 125), lty = 1)
lines(h2, col = "springgreen4", lty = 2)
legend("topright", bty = "n", legend = c("h1", "h2"),
       col = c("steelblue", "springgreen4"), lty = 1:2)
```



β_1 is 0.8 for h_{1t} but only 0.1 for h_{2t}. Consequently, as expected, $\{h_{1t}\}$ has a more persistent (less volatile) conditional variance than $\{h_{2t}\}$.

- The significance of *GARCH* models is that a high order *ARCH* model may have a more parsimonious *GARCH* representation.

For example, backward iteration of a *GARCH*(1,1) error process yields:

$$
\begin{aligned}
h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \\
&= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 (\alpha_0 + \alpha_1 \varepsilon_{t-2}^2 + \beta_1 h_{t-2}) \\
&= \alpha_0 (1 + \beta_1) + \alpha_1 (\varepsilon_{t-1}^2 + \beta_1 \varepsilon_{t-2}^2) + \beta_1^2 h_{t-2} = \ldots \\
&= \alpha_0 \sum_{i=1}^{m} \beta_1^{i-1} + \alpha_1 \sum_{i=1}^{m} \beta_1^{i-1} \varepsilon_{t-i}^2 + \beta_1^m h_{t-m}
\end{aligned}
$$

If $\beta_1 < 1$,

$$h_t \underset{m \to \infty}{\to} \frac{a_0}{1 - \beta_1} + \alpha_1 \sum_{i=1}^{\infty} \beta_1^{i-1} \varepsilon_{t-i}^2$$

ARCH(∞) error process

The effect of any shock on future volatility decreases over time.

- Models with *ARCH* or *GARCH* errors can be estimated in two steps:

    i. Estimate the mean equation (multiple regression or *ARIMA*) for y_t and save the residuals, e_t.

    ii. Take the squared residuals, e_t^2, and estimate an *ARCH* or *GARCH* variance equation.

Alternatively, it is possible to combine the two steps and to estimate the mean equation and the variance equation simultaneously with the Maximum Likelihood (ML) method.

We are going to use the ML method, keeping in mind that it heavily relies on the distribution of the ε_t error term in the mean equation.

After having estimated the mean equation for y_t, as usual, it is important the study the residuals.

← The key feature of *ARCH* and *GARCH* models is that the conditional variance of ε_t is supposed to be generated by an *AR* or *ARMA* process and this should show up in the residuals.

Once we managed to find and estimate the correct mean equation, conditional heteroskedasticity can be tested for with a Lagrange Multiplier (*LM*) test (McLeod and Li, 1983).

The *ARCH LM* test for H_0: no *ARCH* effects of orders 1,.., q versus H_A: some *ARCH* effects of orders 1,.., q consists of two steps:

i. Assuming an *ARCH*(q) error process, estimate the following auxiliary regression of e_t^2 with OLS

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \ldots + \alpha_q e_{t-q}^2 + \xi_t$$

⟵ If there are no *ARCH* effects of orders 1,.., q, this regression is insignificant.

ii. Taking R^2 from this auxiliary regression, compute the Lagrange Multiplier statistic as

$$LM = TR^2$$ where T is the usable sample size.

Under the null hypothesis, *LM* converges to a chi-square distribution with $df = q$.

⟶ Reject H_0 if *LM* is sufficiently large.

Note: If the sample size is relatively small, it is better to rely on the *F*-test of overall significance performed on the auxiliary regression.

(Ex 1)

Returning to the daily closing US dollar to Australian dollar exchange rate (*EXRF*), recall that the level series looks non-stationary but the first difference of its logarithm, which is the approximate rate of change does not.

b) Perform the *ADF* test on the level and on the first difference of the logarithm of *EXR* to confirm that it is an *I*(1) variable.

```
LNEXR = log(EXR)
library(urca)
summary(ur.df(LNEXR), type = "trend",
            selectlags = "BIC"))
```

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend

Value of test-statistic is: -2.4415

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.96 -3.41 -3.12
```

→ H_0

```
DLNEXR = na.omit(diff(LNEXR, 1))
summary(ur.df(DLNEXR, type = "drift",
            selectlags = "BIC"))
```

```
Value of test-statistic is: -48.1891

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.43 -2.86 -2.57
```

→ H_A

Hence, *EXR* is indeed *I*(1).

c) Model *DLNEXR* by running a simple regression on a constant only.

eq.mean_v1 = lm(DLNEXR ~ 1)
summary(eq.mean_v1)

```
Call:
lm(formula = DLNEXR ~ 1)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.090040 -0.004299 -0.000232  0.003991  0.077571 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.331e-05  1.209e-04   0.275    0.783
```

k = 1 + length(eq.mean_v1$coefficients)
Box.test(eq.mean_v1$residuals,
type = "Ljung-Box", lag = 10, fitdf = k)

```
	Box-Ljung test

data:  eq.mean_v1$residuals
X-squared = 38.899, df = 8, p-value = 5.133e-06
```

This regression is insignificant, and the residuals do not behave as a white noise.

d) Try to improve the specification by modelling *DLNEXR* with *auto.arima()*.

library(forecast)
best.arima = auto.arima(DLNEXR, ic = "aicc",
seasonal = FALSE, approximation = FALSE,
stepwise= FALSE)
summary(best.arima)

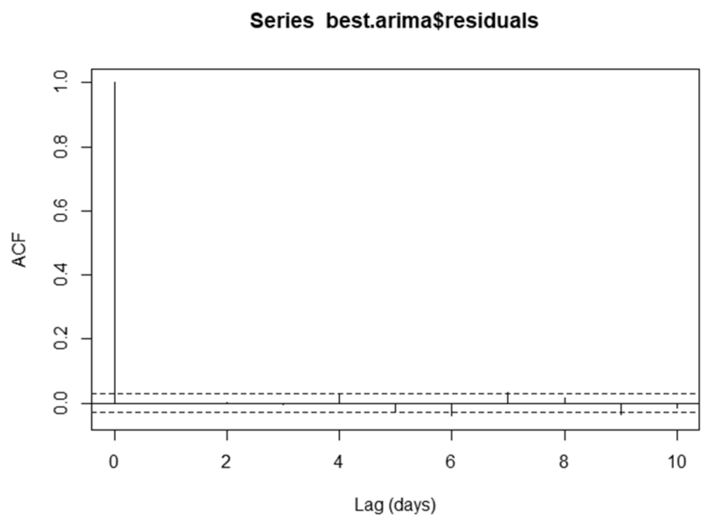
```
Series: DLNEXR 
ARIMA(1,0,0) with zero mean 

Coefficients:
          ar1
      -0.0521
s.e.   0.0150

sigma^2 = 6.471e-05:  log likelihood = 15100.06
AIC=-30196.13   AICc=-30196.13   BIC=-30183.33
```

k = 1 + length(best.arima$coef)
Box.test(best.arima$residuals,
type = "Ljung-Box", lag = 10, fitdf = k)

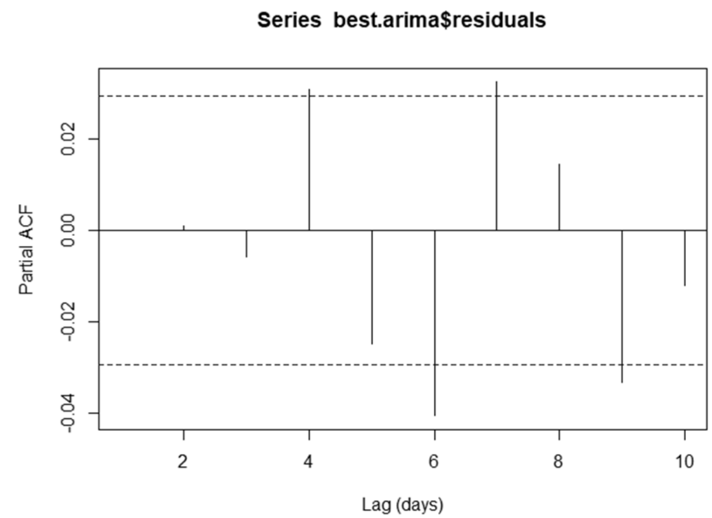
```
	Box-Ljung test

data:  best.arima$residuals
X-squared = 26.091, df = 8, p-value = 0.001013
```

This *AR*(1) model looks better (at least, the absolute value of the standard error is less than 1/3 of the slope estimate), but the residuals are still autocorrelated.

acf(best.arima$residuals, lag.max = 10, xlab = "Lag (days)", plot = TRUE)



pacf(best.arima$residuals, lag.max = 10, xlab = "Lag (days)", plot = TRUE)



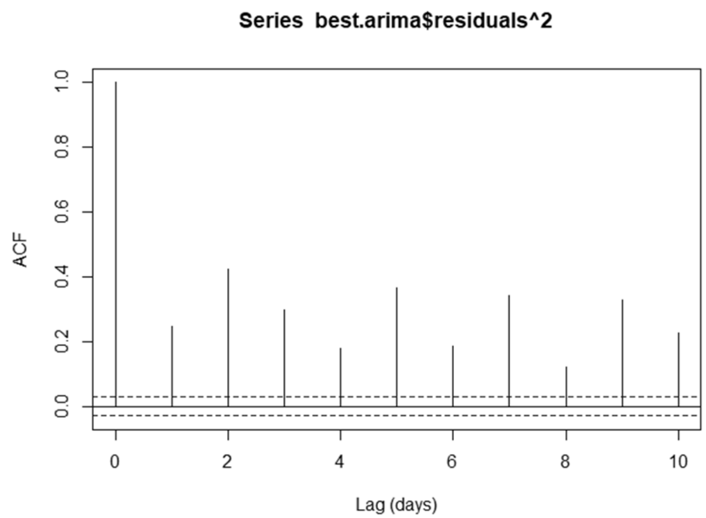
The correlograms also indicate that the residuals from the *AR*(1) model do not behave as a white noise.

In this case further *AR* and *MA* terms fail to eliminate residual autocorrelation.

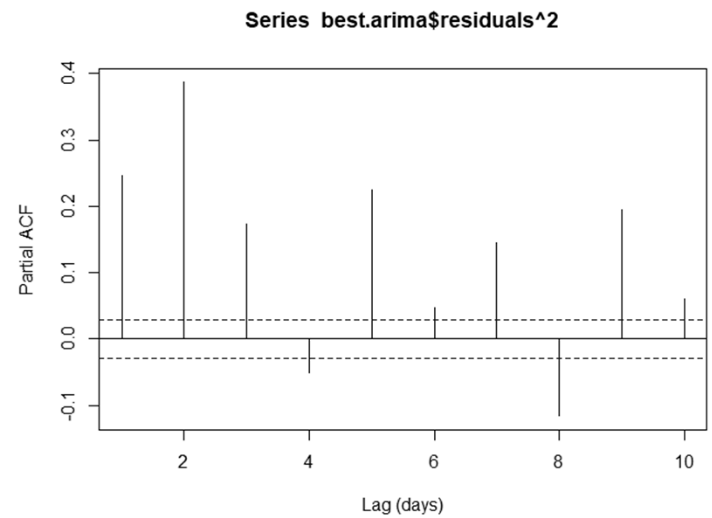
The sample autocorrelation and partial autocorrelation coefficients (for $k > 0$), however, are all very small in absolute value (≤ 0.041), so we stick to the simple and parsimonious *AR*(1) model and accept it as a reasonable mean equation.

e) Check the possibility of (*G*)*ARCH* errors by developing the sample correlogram for the squared residuals from the mean equation and performing the *LM* test for *ARCH* errors.

acf(best.arima$residuals^2, lag.max = 10, xlab="Lag (days)", plot = TRUE)



pacf(best.arima$residuals^2, lag.max = 10, xlab="Lag (days)", plot = TRUE)



Box.test(best.arima$residuals^2, type = "Ljung-Box", lag = 10, fitdf = k)

```
        Box-Ljung test

data:  best.arima$residuals^2
X-squared = 3654.9, df = 8, p-value < 2.2e-16
```

All sample autocorrelation and partial autocorrelation coefficients are significant, and so is the *LB* test statistic for H_0: no autocorrelation of orders 1-10 in the squared errors.

Hence, the *AR*(1) errors are likely generated by some (*G*)*ARCH* process.

The *ARCH LM* test can be performed with the *ArchTest()* function of the *FinTS R* package.

library(FinTS)
ArchTest(best.arima$residuals, lags = 10

```
        ARCH LM-test; Null hypothesis: no ARCH effects

data:  best.arima$residuals
Chi-squared = 1424.3, df = 10, p-value < 2.2e-16
```

Homoskedasticity can be safely rejected in favour of (*G*)*ARCH* errors, confirming our previous conclusion.

f) Estimate an *AR*(1)-*ARCH*(1) model for *DLNEXR*.

⟵ The correlograms of the squared residuals (see the previous slide) do not have cut-off points and thus provide no hint about the specification of the conditional variance equation, other than that it might be (*G*)*ARCH*. Likewise, the *ARCH LM* test does not help specify the conditional variance equation.

In cases like this, the best is to keep the specification simple and start with *ARCH*(1) or/and *GARCH*(1,1) as they are often sufficient to capture volatility clustering,

and then test the standardized residuals $\boxed{ste_t = e_t / \sqrt{\hat{h}_t}}$

for no autocorrelation and for no remaining (*G*)*ARCH* effect.

GARCH models can be set up and estimated in two steps with the *ugarchspec()* and *ugarchfit()* functions of the *rugarch R* package.

```
library(rugarch)
spec_v1 = ugarchspec(mean.model = list(armaOrder = c(1,0),
                                        include.mean = TRUE),
                     variance.model = list(model = "sGARCH",
                                           garchOrder = c(1,0)),
                     distribution.model = "norm")

fit_v1 = ugarchfit(spec = spec_v1, data = DLNEXR)
fit_v1
```

The *ugarchfit()* printout is quite long. It starts with the estimated model:

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,0)
Mean Model      : ARFIMA(1,0,0)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000022    0.000110  0.20245 0.839562
ar1     0.046190    0.015737  2.93517 0.003334
omega   0.000045    0.000001 32.71251 0.000000
alpha1  0.323179    0.031129 10.38184 0.000000
```

sGARCH: *simple GARCH*

ARFIMA: *fractionally* integrated *ARIMA*, i.e., *d* can take fractional values. (We do not discuss the details.)

Mean eq.: the intercept (*mu*) is insignificant but *ar1* is significant.

Variance eq.: the intercept (α_0 ~ *omega*) and *alpha1* are both significant.

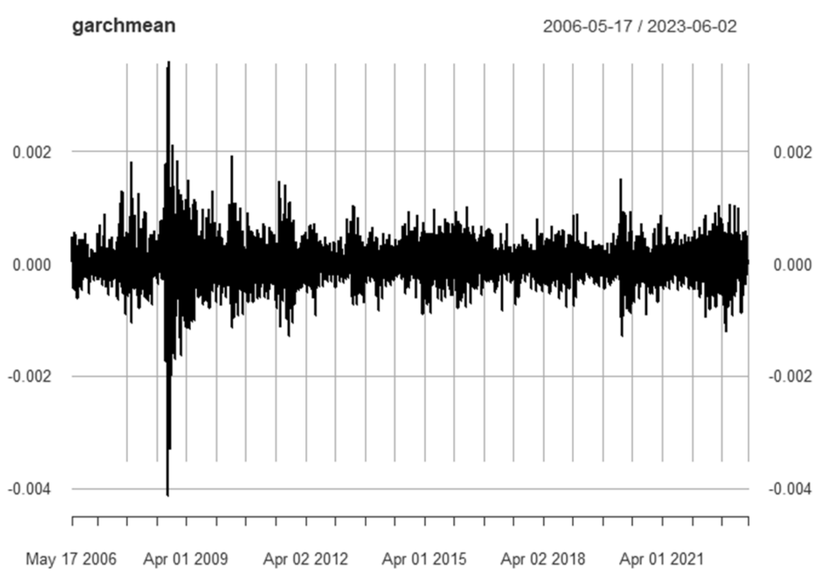
From this printout the sample *AR*(1)-*ARCH*(1) model for *DLNEXR*:

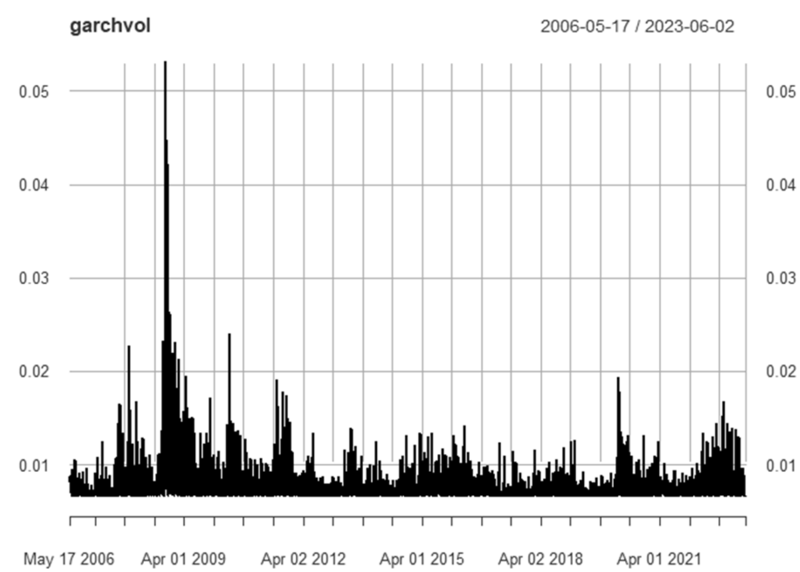
$$\widehat{DLNEXR_t} = 0.000022 + 0.046190 DLNEXR_{t-1} + e_t \quad , \quad e_t \sim N(0, \hat{h}_t)$$
$$\hat{h}_t = 0.000045 + 0.323179 e_{t-1}^2$$

The estimated mean and volatility of *DLNEXR*:

garchmean = fitted(fit_v1)
plot(garchmean)

garchvol = sigma(fit_v1)
plot(garchvol)





The conditional distribution of the ε_t mean equation error is supposed to be normal (see *Distribution: norm* on the printout).

R also reports robust standard errors based on the Quasi ML method as opposed to the regular standard errors based on ML.

Robust Standard Errors:

	Estimate	Std. Error	t value	Pr(>|t|)
mu	0.000022	0.000133	0.16726	0.867167
ar1	0.046190	0.064045	0.72122	0.470776
omega	0.000045	0.000003	14.92935	0.000000
alpha1	0.323179	0.142150	2.27350	0.022996

These standard errors are robust against violations of the distributional assumption, e.g., when the distribution is assumed to be normal, but the true distribution is Student-*t*.

Given these robust standard errors, *mu* and *ar1* become insignificant, so the distributional assumption is important.

Next, there are four information criteria on the printout:

Information Criteria

Akaike	-6.9153
Bayes	-6.9095
Shibata	-6.9153
Hannan-Quinn	-6.9132

Akaike: *AIC*, Bayes: *BIC* (week 3, slide #30)

Shibata and Hannan-Quinn are two further criteria. The rule is the same for all: the smaller the better.

They are used to rank alternative specifications so at this stage they are not informative.

ugarchfit() tests both the standardized residuals and the standardized squared residuals for autocorrelation with a weighted version of the *LB* test, which is more powerful in detecting autocorrelation in residuals from *ARMA* models.

$$Q_{LB} = T(T+2)\sum_{k=1}^{s}\frac{r_k^2}{T-k}$$

$$Q_{WLB} = T(T+2)\sum_{k=1}^{s}\frac{s-k+1}{s}\frac{r_k^2}{T-k}$$

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic   p-value
Lag[1]                      15.12 1.007e-04
Lag[2*(p+q)+(p+q)-1][2]     17.14 0.000e+00
Lag[4*(p+q)+(p+q)-1][5]     19.57 3.961e-08
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.905  0.1675
Lag[2*(p+q)+(p+q)-1][2]   142.809  0.0000
Lag[4*(p+q)+(p+q)-1][5]   287.490  0.0000
d.o.f=1
```

Not surprisingly, the standardized residuals are autocorrelated, but this is not a concern this time (see slide #21).

However, the standardized squared residuals are also autocorrelated, suggesting that the variance equation is not properly specified.

ugarchfit() also performs *weighted ARCH LM* tests for *ARCH* effects remaining in the standardized residuals (Fisher and Gallagher, 2012), which also indicate that the variance equation is inadequate.

```
Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[2]     281.6 0.500 2.000       0
ARCH Lag[4]     325.6 1.397 1.611       0
ARCH Lag[6]     399.9 2.222 1.500       0
```

The rest of the printout shows the results of three groups of tests. We do not discuss the details of those tests, but it is important to understand their purpose and the conclusions they imply.

The first group consists of joint and individual *Nyblom stability tests* for parameter stability, i.e., for structural change in the data generating process (Nyblom, 1989).

H_0: stable value (=0)

```
Nyblom stability test
------------------------------------
Joint Statistic:  7.1191
Individual Statistics:
mu      0.1200
ar1     0.1113
omega   4.7185
alpha1  3.6817

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:           1.07 1.24 1.6
Individual Statistic:      0.35 0.47 0.75
```

The joint test statistic is well above the critical values, so the joint null hypothesis that each parameter is constant is rejected.

As for the individual tests, the mean equation parameters (*mu, ar1*) pass them, unlike the variance equation parameters (*omega, alpha1*).

The second group consists of *sign bias tests* for leverage effects, i.e., that negative and positive returns have different influence on future volatility (Engle and Ng, 1993).

H_0: no leverage

```
Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           0.7795 0.4357
Negative Sign Bias  0.9238 0.3556
Positive Sign Bias  0.8809 0.3784
Joint Effect        1.6515 0.6478
```

Each test maintains the null hypothesis of no leverage effect.

Finally, the third group consists of *adjusted Pearson tests for goodness of fit* (Palm, 1996). They serve to compare the empirical distribution of the standardized residuals with the chosen conditional distribution of ε_t, which is normal in this case.

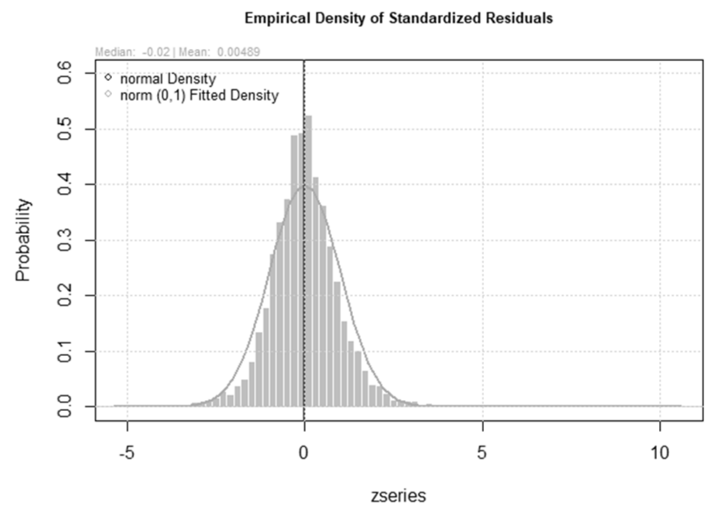
H_0: normally distributed errors

```
Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     188.7    5.966e-30
2    30     204.0    3.349e-28
3    40     231.7    3.217e-29
4    50     241.4    3.157e-27
```

No matter how many groups the observations are classified in, the null hypothesis of normally distributed errors is rejected.

This decision is supported by the empirical density and kurtosis of the standardized residuals:

plot(fit_v1,which = 8)



The empirical density is narrower than the standard normal distribution.

library(moments)
kurtosis(residuals(fit_v1, standardize = T))

```
10.84073
```

The standardized residuals have larger (excess) kurtosis than the normal distribution (0), so their distribution is leptokurtic.

WHAT SHOULD YOU KNOW?

- Volatility clustering
- The nature of *ARCH* and *GARCH* processes
- Lagrange Multiplier test for *ARCH* effect
- The specification and estimation of *ARCH* and *GARCH* models with the *rugarch* library of R

BOARD OF FAME

Robert Fry Engle III (1942-):

American economist and statistician

Nobel laureate in Economic Sciences in 2003

Michael Armellino Professor at Stern School of Business, New York University

Financial econometrics, time-varying volatility, *ARCH* modelling, cointegration, error correction

Tim Peter Bollerslev (1958-):

Danish economist

Juanita and Clifton Kreps Professor of Economics at Duke University

Financial econometrics, time-varying volatility, *GARCH* modelling