UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Financial Econometrics

Topic 5:
Modelling Volatility

Lecturer: Nico Katzke
nicokatzke@sun.ac.za

Department of Economics

DEPARTMENT OF ECONOMICS

Texts used

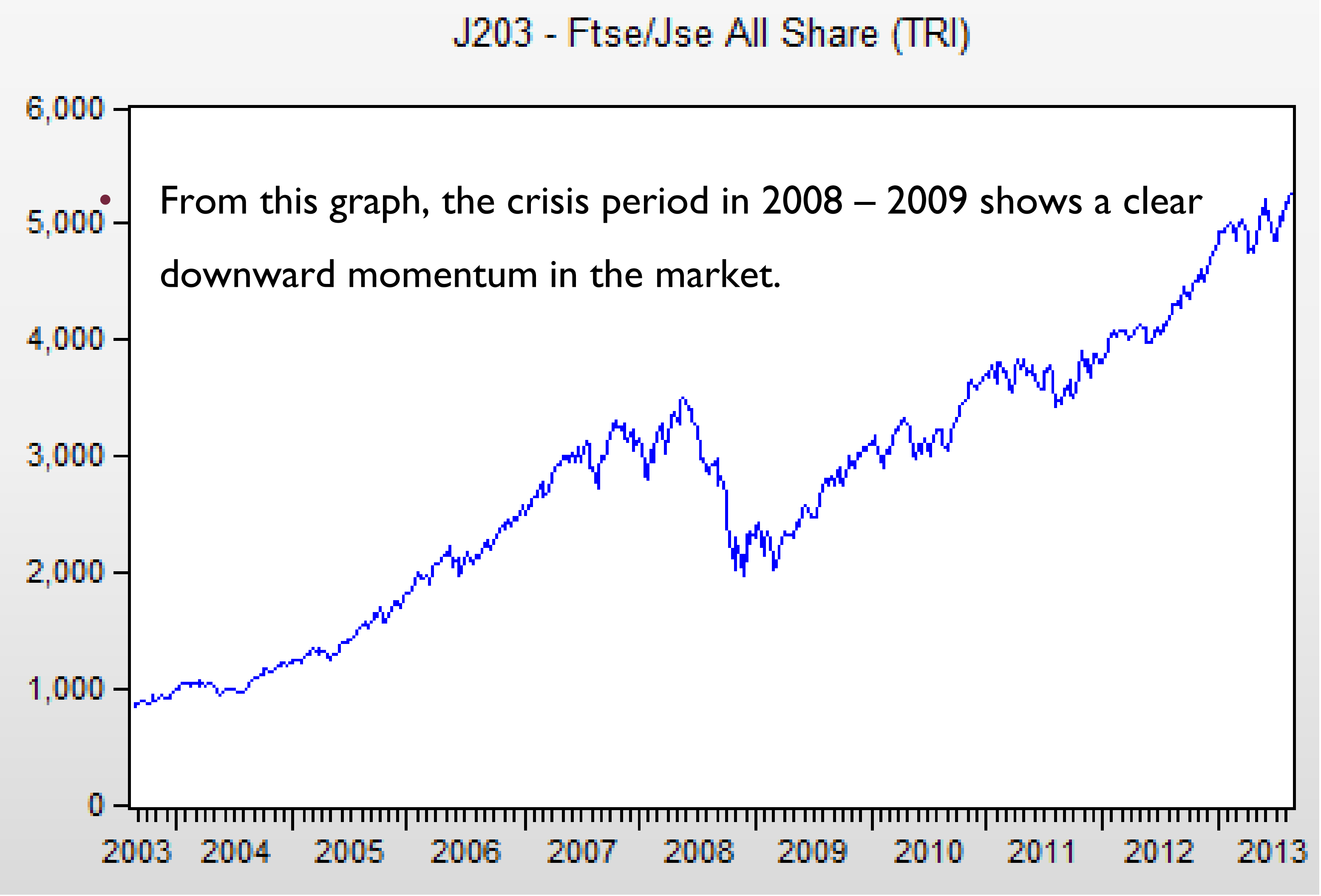
- The notes and code in R were created using many references.
- Intuitively, Brookes explains volatility modelling well, as does Tsay (2012,2014) and Ruppert (2011)

- Also, read the following paper which explains the different GARCH models and distributions very well: Hentschel (1995). *All in the family…*

Modelling Volatility - importance

- In financial markets we take as a stylized fact that risk requires a premium in terms of its expected pay-off. This implies that investors **should be compensated for taking on higher risk.**
- In this section we look at uncovering predictable patterns in the volatility of financial time-series data.
- In particular we will see whether we could fit an **autoregressive process on the second moments (variance component)** of our series in order to see whether it shows persistence during certain periods.
- Think back on the Global Financial Crisis – remember how volatile global markets were… Let's look at the J203 (JSE All Share Index) TRI

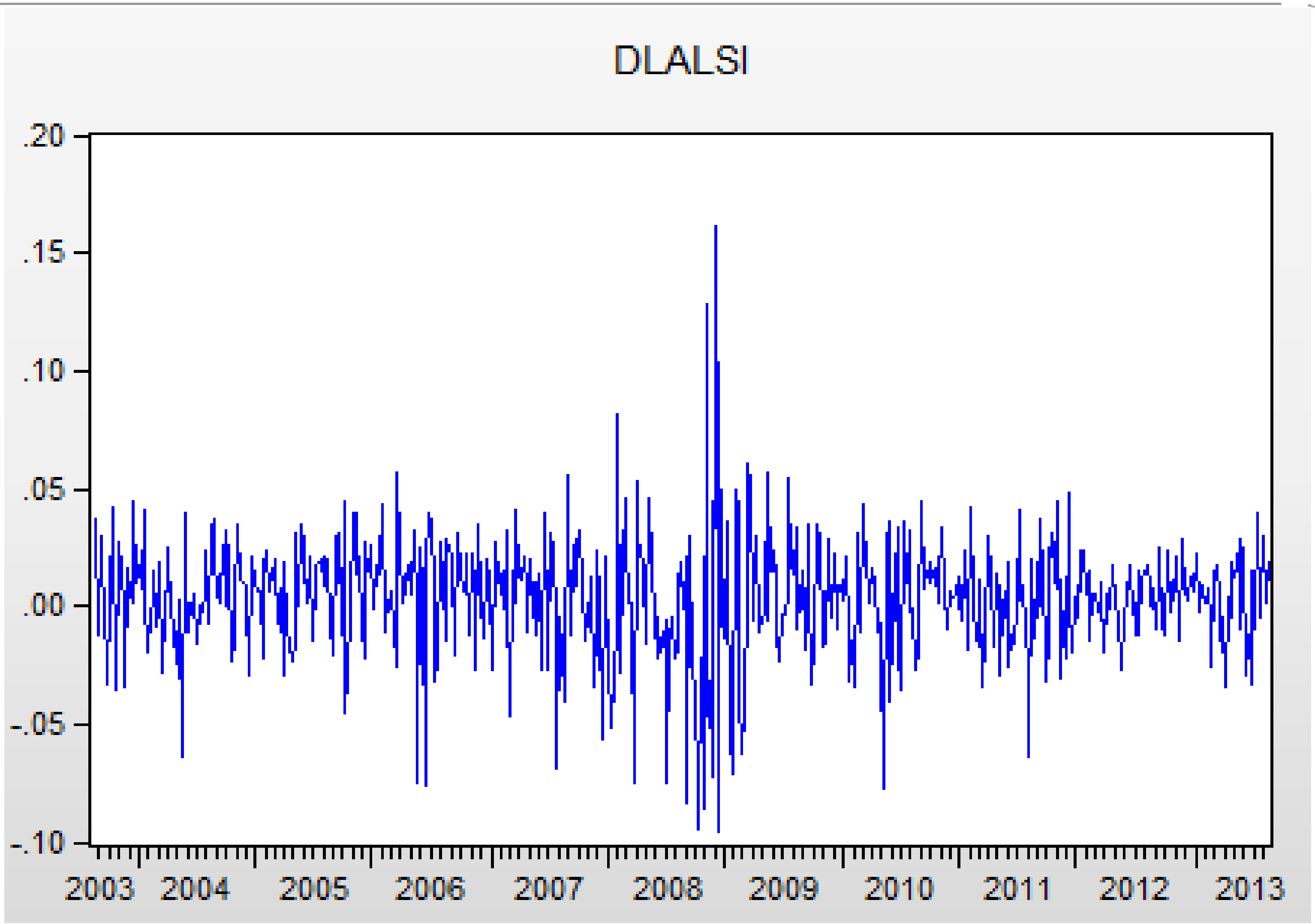
Weekly J203 Total Return Index



- From this graph, the crisis period in 2008 – 2009 shows a clear downward momentum in the market.

Standard procedure follows…

- If you look closely, you will see the level increase in volatility over time – requiring the Logarithmic form.
- You will also notice strong mean persistence and also time dependency on the **mean**.
- After taking the Log, we see that the ADF test rejects stationarity, requiring us to take the **difference** to remove the unit root.
- That means we take the **dlog** of the J203 series, which yields the following graph:


DLALSI
.20
.15
.10
.05
.00
-.05
-.10
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013

Fitting ARIMA model on J203

- From the graph of the Dlog(J203), we can clearly see there still remains significant persistence in the series… Although it now resembles something closer to stationarity…
- In order to make the residuals of the J203 White Noise – we now include autoregressive components to the model {The correlogram suggests including the AR(1) MA(1) terms} – thus an ARIMA(111) model is fitted.

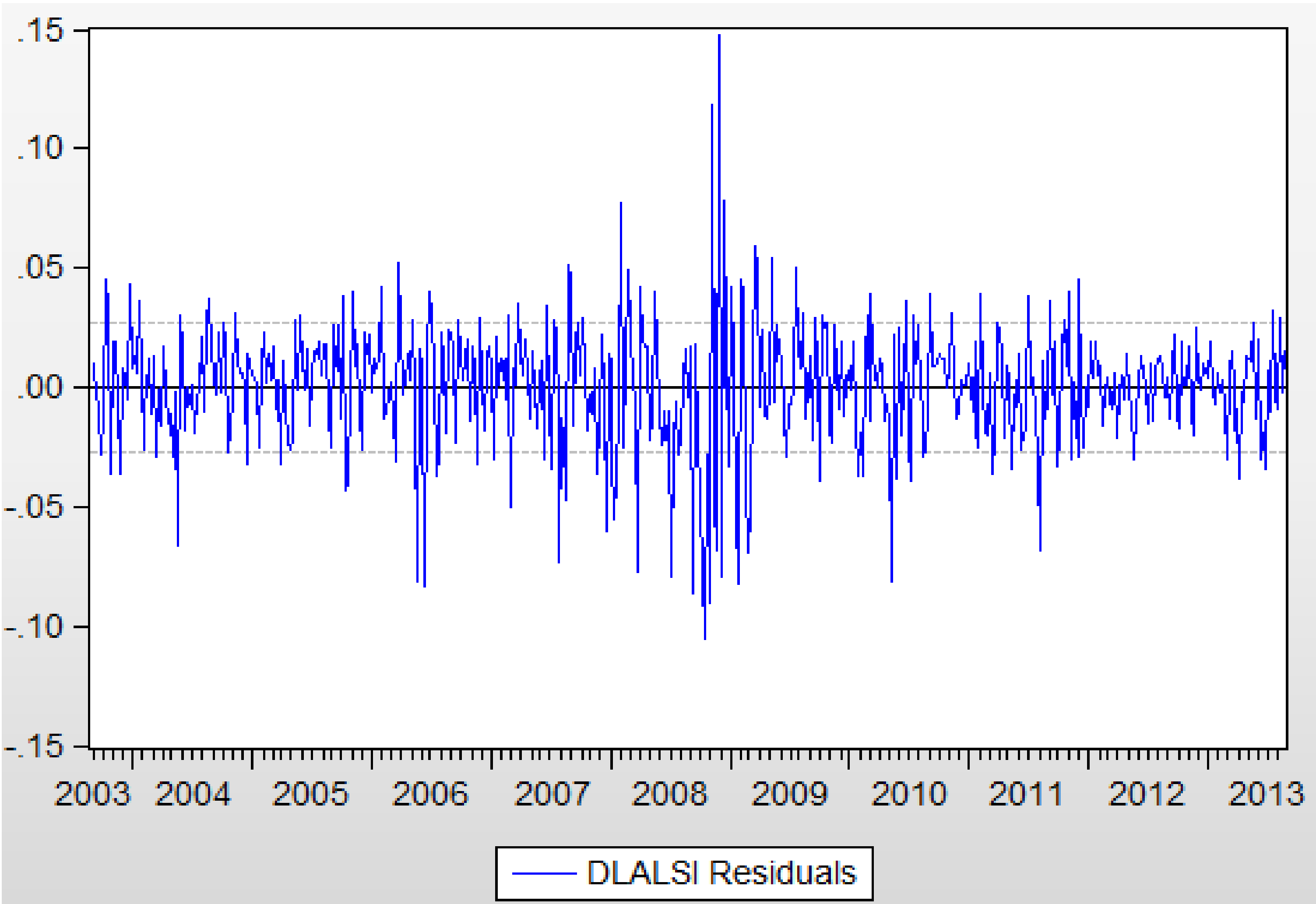
- Looking at the Ljung-Box Q-stats and the correlogram of the ARIMA(111), the residuals seem now to be well-behaved (White Noise).
- This allows us then to use the model in making forecasts / including it in other regressions.

Correlogram of ARIMA(1,1,1) for J203

Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.016	-0.016	0.1400	
		2	-0.049	-0.049	1.3950	
		3	-0.013	-0.014	1.4789	0.224
		4	0.014	0.011	1.5780	0.454
		5	-0.035	-0.036	2.2294	0.526
		6	0.049	0.049	3.5047	0.477
		7	-0.001	-0.002	3.5050	0.623
		8	-0.054	-0.051	5.0624	0.536
		9	0.062	0.063	7.1240	0.416
		10	-0.005	-0.011	7.1358	0.522
		11	-0.026	-0.019	7.5071	0.584
		12	-0.092	-0.094	12.054	0.281

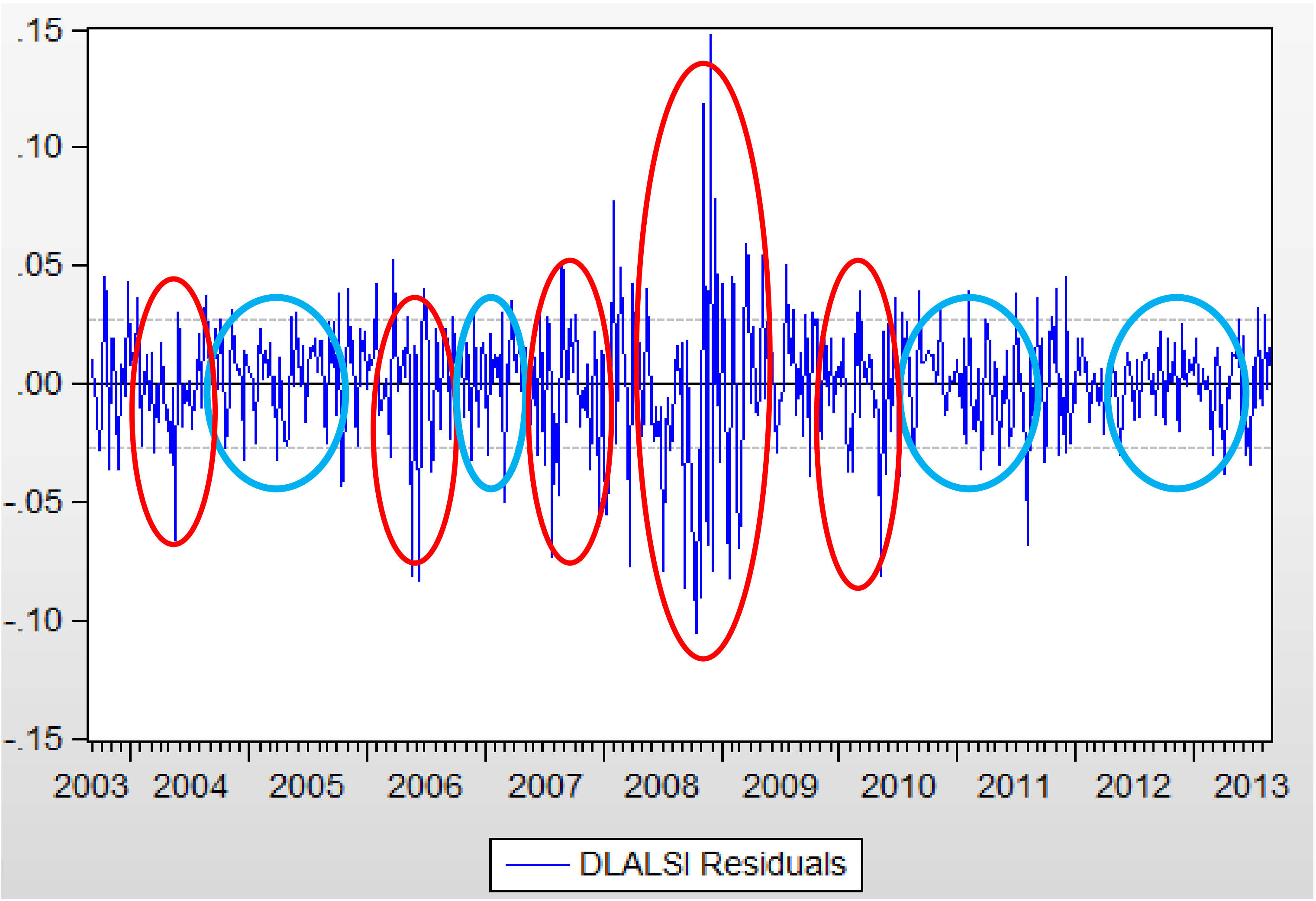
Residuals graph of the ARIMA(111) J203:



Are the residuals ***really*** well behaved?

- Considering the correlogram and the graph of the residuals, it seems as though the residuals of the univariate autoregressive model is well-behaved.
- It seems now appropriate to conduct time-series analysis using this model, as it shows stationarity with White Noise residuals.
- But let's look a bit closer… Notice that there are periods where the volatility seems to cluster, i.e. periods where there seems to be market momentum (check how the periods in blue differ from those in red in terms of their variance…)

Residuals graph of the ARIMA(1 1 1) J203:

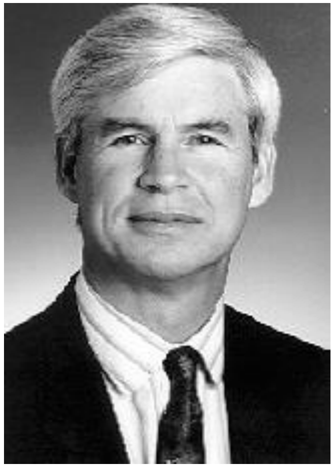


Periods of persistence

- As seen in the above graph of the J203 series, there are periods where the series displays strong persistence and periods of higher than normal volatility (red vs blue periods).
- Often such period's volatility is conditional upon the previous period's volatility being above normal, creating a sort of **volatility momentum** carried into the next period.
- If left uncontrolled for, this momentum in the residuals will negatively impact the fit of our model.
- Thus controlling for this *volatility* ***conditional*** *on the previous period's volatility* (or conditional heteroskedastic periods) is the aim of this session.

Conditional heteroskedasticity

- Heteroskedasticity we normally associate with cross sectional studies, whereas time series data sets we assume them to be made homoskedastic.
- Engle (1982, 1983), however, presented evidence that most financial time-series data sets display <u>periods of persistence in volatility</u> in terms of error variances.
- Initially, work was done on models of inflation where varying sizes of forecast errors seemed to **cluster**, suggesting a form of heteroskedasticity in which the variance of a forecast error depends on the size of the previous period’s disturbance(s).

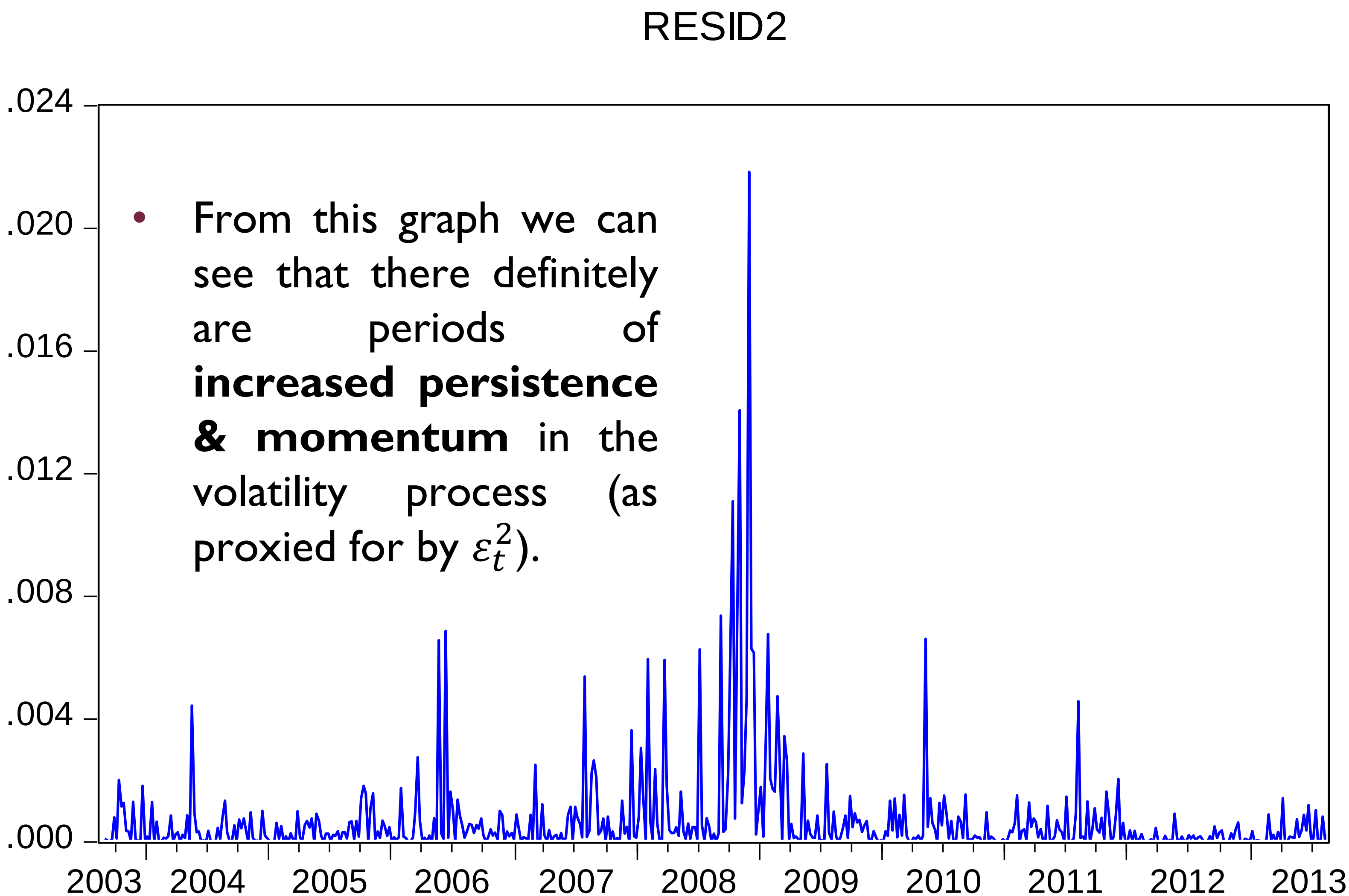


Importance of homoskedastic assumption

- Remember from our definition of stationarity – we assumed the residuals to be homoskedastic (have a constant, time-invariant variance).
- As we assume the residuals to have a zero mean, that would imply that $Var(\varepsilon_t) = E(\varepsilon_t^2) - E(\varepsilon_t)^2 \approx E(\varepsilon_t^2) = \text{c}$ under homoskedasticity.

- But this is where Engle showed that there can be persistence still remaining – in the squared residual term – even though the residual term may display White Noise.

Resid squared graph (ε_t^2)



- From this graph we can see that there definitely are periods of **increased persistence & momentum** in the volatility process (as proxied for by ε_t^2).

ARCH background

- The Conditional Heteroskedastic (ARCH) model was first proposed to take into account this **autocorrelation** (or serially dependent) **of the** (or serially dependent) **error variances** displayed at certain periods.
- The ARCH process similar to that of the ARIMA family, as it controls for periods of volatility persistence in the residual process specifically.

- Soon after the initial specification on inflation data, other studies pointed to the application of such techniques to fields ranging from term-structures of interest rates, stock market return volatility, foreign exchange behaviour, inflation modelling, etc.

No “con” in conditional…

- Although we assume our time series (after stationarity transformations) is made **time invariant**, the earlier models only consider the **<u>unconditional</u>** (or long run) constant -means and –variance processes over time.
  - Thus unconditional forecasts would imply forecasting using only the long run mean.
- Following Engle’s approach: although the residuals may be homoskedastic in the long run (Unconditionally homoskedastic), the short run behaviour of the variance structure might be time-dependent – i.e. showing the presence of **conditional heteroskedasticity** (which is persistence in the variance structure **conditional upon** a past period higher-than-normal variance).
- This gives us, by definition (See Enders : 126), better forecasts than unconditional forecasts.

Up to now…

- We made: $E(\varepsilon_t) = 0$ (using an ARMA model / a regression fit of y_t)
- And we assumed: $\mathrm{Var}(\varepsilon_t) = E(\varepsilon_t)^2 - 0 = \mathrm{h}^2 = time\ invariant$

- One way to now model a non-constant error variance that displays periods of persistent volatility, is by fitting an AR(p) process on the **squared residuals**:

$$\widehat{\varepsilon_t}^2 = \alpha_0 + \alpha_1 \widehat{\varepsilon_{t-1}}^2 + \alpha_2 \widehat{\varepsilon_{t-2}}^2 + \cdots + \alpha_p \widehat{\varepsilon_{t-p}}^2 + \eta_t$$

$$With:\ \eta_t \equiv White\ noise$$

- From this, if the $\alpha's$ are all zero, it implies the error variance $[\mathrm{E}(\widehat{\varepsilon_t}^2)]$ is constant, and thus homoskedastic (and no conditional heteroskedasticity is present)
- This process of fitting an AR to the squared residuals is known as **ARCH** modelling.

ARCH specification

- We normally specify multiplicative disturbance terms, using ML techniques to simultaneously estimate the parameters. Thus the simplest form of ARCH specification is the ARCH(1) model:

$$y_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = \eta_t . h_t = \eta_t . \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}{}^2}$$

With:

$\boldsymbol{y_t}$ made stationary by the model $\boldsymbol{\mu_t}$ (which can be an ARIMA process, or contain explanatory variables, etc.).

$\boldsymbol{\varepsilon_t}$ displays stationarity, but the squared residuals ($\boldsymbol{\varepsilon_t}^2$) show **autoregressive** [AR(1)] **behaviour**

$$\eta_t \equiv White\ Noise \sim N(0,1)$$

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}{}^2 \rightarrow ARCH(1)\ \text{variance process}$$

Restriction : $\alpha_0 > 0; \quad 0 < \alpha_1 < 1$, cause we can't have negative variance…

Unconditional moments: Mean

From the ARCH(1) model: $y_t = \mu_t + \varepsilon_t$

$$\varepsilon_t = \eta_t . h_t = \eta_t . \sqrt{\alpha_0 + \alpha_1 {\varepsilon_{t-1}}^2}$$

it follows (as before) that:

- $E(\varepsilon_t | \mu_t) = E\left[\eta_t \sqrt{\alpha_0 + \alpha_1 {\varepsilon_{t-1}}^2}\right] = E(\eta_t). \sqrt{\alpha_0 + \alpha_1 E({\varepsilon_{t-1}}^2)} = 0$

  (Due to independence of η_t & h_t & the fact that: $\eta_t \equiv WN$)

- Thus : $E(y_t) = \mu_t$, {which is $\beta_0 + \beta_1 y_{t-1}$ if μ_t is an AR(1) process}

  (which is the classic regression specification)

- **But : the variance of ε_t differs from before…**
- In particular, we will show that it is **Unconditionally** constant, but **Conditionally** dependent on past variances…

Unconditional moments: Variance

- The **unconditional variance**:

$$Var(\varepsilon_t|\mu_t) = E(\varepsilon_t{}^2|\mu_t) = E[\eta_t{}^2.(\alpha_0 + \alpha_1\varepsilon_{t-1}{}^2)|\,\mu_t]$$

$$= E(\eta_t{}^2).E\,(\alpha_0 + \alpha_1\varepsilon_{t-1}{}^2) = 1.\left(\alpha_0 + \alpha_1 E(\varepsilon_{t-1}{}^2)\right)$$

- So that : $Var(\varepsilon_t) = \alpha_0 + \alpha_1 Var(\varepsilon_{t-1})$

- If the residuals can be assumed stationary, the unconditional variances are equal over time, so that we have: $\{Var(\varepsilon_t) = Var(\varepsilon_{t-1})\}$:

$$Var(\varepsilon_t|\mu_t) = \alpha_0 + \alpha_1 Var(\varepsilon_{t-1}) = \alpha_0/1 - \alpha_1$$

→ **Which is constant**

- And we then have the **unconditional** (LR) distribution of the residuals as:

$$\varepsilon_t \sim \mathrm{N}(0, \alpha_0/1 - \alpha_1)$$

- Also, **unconditional autocovariances** are zero : $\mathrm{E}(\varepsilon_t, \varepsilon_{t-1}|\mu_t) = 0$

Conditional moments

- Notice that the unconditional mean, variance & autocovariances of the residuals was unaffected by the presence of the ARCH error process.
- The **conditional mean** (conditional on **past** residuals) is found as:

$$E(\varepsilon_t|\varepsilon_{t-1}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i\, E_{t-1}(\varepsilon_t) = E_{t-1}(\eta_t\sqrt{\alpha_0 + \alpha_1\varepsilon_{t-1}{}^2}) = 0$$

- And the **conditional variance** is found as:

$$Var(\varepsilon_t|\varepsilon_{t-1}) = Var_{t-1}\left(\eta_t\sqrt{\alpha_0 + \alpha_1\varepsilon_{t-1}{}^2}\right) =$$

$$(1).\, E_{t-1}(\alpha_0 + \alpha_1\varepsilon_{t-1}{}^2) = \alpha_0 + \alpha_1\varepsilon_{t-1}{}^2 \quad \neq constant$$

- Thus the variance of ε_t conditional upon its past residuals (and not conditional on μ_t as studied before) **is not constant**, and thus displays ***conditional heteroskedasticity***

Martingale System

- Remember, by definition – a martingale series has the following attributes wrt itself and any other stochastic process (X_t) – making it essentially unpredictable:

$$E(Y_{t+1}|X_t, \dots X_{t-p}) = 0, \qquad \forall t$$

Which implies:

$$E(Y_{t+1}|Y_t, \dots Y_{t-p}) = E(Y_{t+1}|Y_t) = 0$$

$$\rightarrow Thus\ not\ autocorrelated$$

$$E(Y_{t+1}) = 0$$

But this does not imply the following:

$$Var(Y_{t+1}|X_t, X_{t-1}, \dots) = 0$$

In fact, even if the variance function could be predictable based on the past, the series' mean could remain unpredictable! – thus **conditional heteroskedasticity** could be present!

Conditional dependence of volatility

- Notice that if ε_{t-1} is large, we expect the conditional variance in the next period to be large as well.
- In order for both the unconditional and the conditional variance to be positive, we need the restriction to hold that: $\underline{\alpha_0 > 0; \quad \alpha_1 > 0}$
- Also : in order for the conditional variance to be a finite AR-process, $0 < \alpha_1 < 1$ must hold.

- **<u>In summary then</u>**, both the conditional and unconditional expected errors are zero, the errors are serially uncorrelated {E$(\varepsilon_t, \varepsilon_{t-1}|\mu_t) = 0$}, while the **conditional variance** of the error terms are dependent on its own lagged values → implying the existence of persistent volatility in the error terms.

ARCH(p)

- Extending the ARCH theory to a p[th] order follows simply as extending the autoregressive residual process (defined as h_t^2 now):

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^{p} \alpha_i {\varepsilon_{t-i}}^2}$$

- With the **unconditional variance** (if $\varepsilon_t \rightarrow$ made stationary):
- $Var(\varepsilon_t) = Var(h_t . \eta_t) = \alpha_0 / (1 - \sum_{i=1}^{p} \alpha_i) \;\; \rightarrow \textit{which is } \boldsymbol{constant}$

$$\textit{With constraints}: \boldsymbol{\alpha_0 > 0, \alpha_i > 0} \;\; (\forall\, \boldsymbol{i}),\;\; \boldsymbol{0 < \sum_{i=1}^{p} \alpha_i < 1}$$

- **Conditional variance**
- $Var_{t-1}(\varepsilon_t | \varepsilon_{t-1}, \dots, \varepsilon_{t-p}) = 1. h_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i E_{t-1}({\varepsilon_{t-i}}^2) = \alpha_0 + \sum_{i=1}^{p} \alpha_i {\varepsilon_{t-i}}^2$

$$\rightarrow \textit{which is } \boldsymbol{conditionally\ heteroskedastic}$$

Formalized Notation of ARCH(p)

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

$\{x_t$ can be an ARMA process with exogenous variables $\rightarrow$ called the mean equation$\}$

$$\varepsilon_t = h_t . \eta_t \;\; ; \;\; h_t = \sqrt{\alpha_0 + \sum_{i=1}^{p} \alpha_i {\varepsilon_{t-i}}^2} \;\; ;$$

$$h_t \sim ARCH(p) \rightarrow \text{variance equation}$$

$$\eta_t \sim WN(0,1)$$

$$\textit{With constraints :} \; \alpha_0 > 0, \;\; \alpha_i > 0, \qquad 0 < \sum_{i=1}^{p} \alpha_i < 1$$

$$\boldsymbol{\varepsilon_t} = \textbf{ ordinary residuals } \textit{from the ARIMA model}$$

$$\eta_t = \varepsilon_t / h_t \sim N(0,1) = \textbf{standardized residuals } \textit{from the ARCH process}$$

$$\textit{Also, the ARMA and ARCH orders } (\textit{of } x_t \textit{ and } \mathrm{h}_t) \textit{ need not equate}!$$

Unpacking the previous slide

- x_t is the ***unconditional*** *(Long run)* ***mean*** of y_t
- ε_t is the ***ordinary residual*** [$\boldsymbol{\varepsilon_t} = y_t - x_t$ $\rightarrow$ hence the **ordinary** residuals], which is serially **uncorrelated**, BUT **dependent** on previous lags through its second moment…
  - This **autoregressive dependence** is clear from the **conditional variance** term : **[$\boldsymbol{h_t}$]** , which is an autoregressive function of the lags of $\boldsymbol{\varepsilon_t^2}$.

    → $\boldsymbol{h_t}$ thus is called the ***conditional variance*** as it is the part of the variance which is conditional on past trends in volatility (error variance).
- [$\boldsymbol{\eta_t}$] is then called the ***standardized residuals***, as it has controlled for the conditional heteroskedasticity in the ordinary residuals : $\eta_t = \boldsymbol{\varepsilon_t / h_t}$

$\boldsymbol{\eta_t} \sim N(0,1) \rightarrow$ thus implying we should have WN std resids after fitting ARCH!

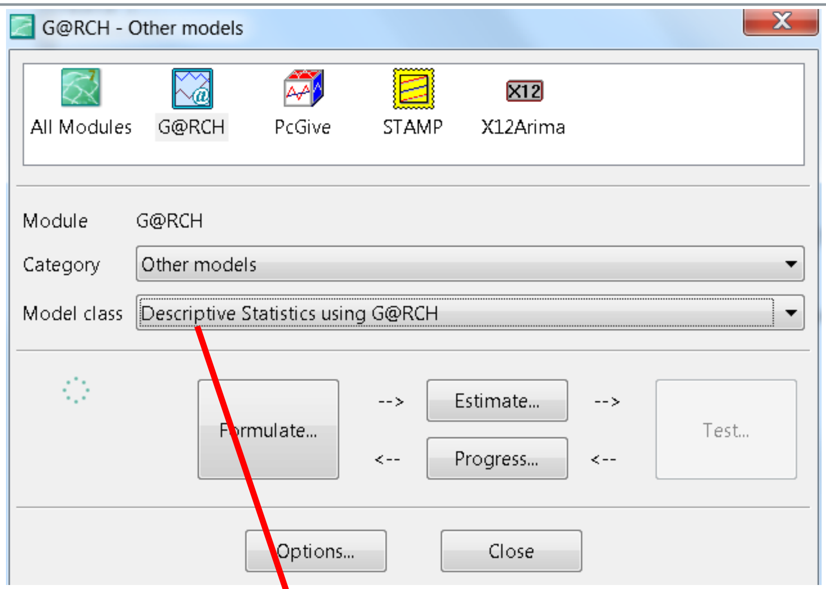
Uncorrelated, but dependent…

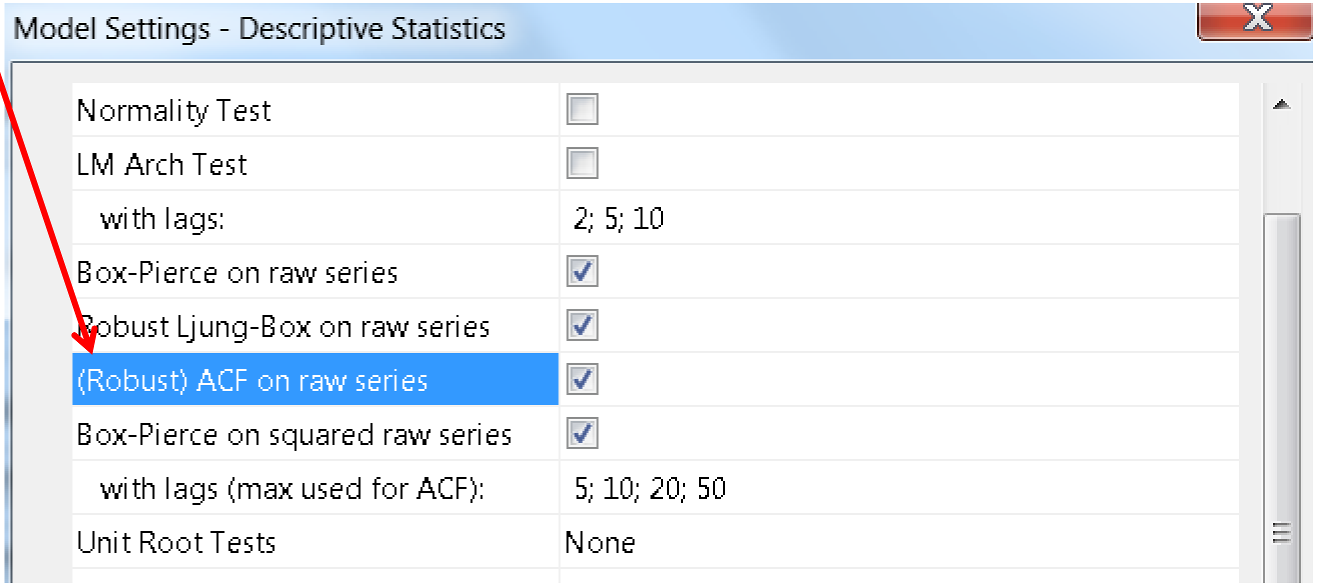
- Notice that the above framework implies the series can be **serially uncorrelated** and have stationary ordinary residuals that are unconditionally homoskedastic, but at times display heteroskedasticity conditional on past shocks…
- Thus the ordinary residuals, although **uncorrelated**, are **dependent** on past shocks.
- This implies that the ARCH model is able to capture periods of tranquility and volatility in our series $\{y_t\}$!
- Thus, after fitting the conditional variance equation, we should have :
- Stationary **ordinary residuals** (serially uncorrelated, but dependent)
- White Noise **standardized residuals** → which implies after taking into account the conditional heteroskedasticity, our standardized residuals are WN (serially uncorrelated and independent).

Can we trust ACF tests in the presence of conditional heteroskedasticity?

- Standard Portmanteau tests like the Box-Pierce or LBQ-stats used for assessing serial dependence – will be fatally biased in the presence of heteroskedasticity.
- This follows as the critical values calculated for these tests do not account for the presence of dependence on past second order moments (conditional heteroskedasticity) and thus are not longer $\chi^2 -$ distributed.
- As a result, Francq and Zakoian (2009) proposed a corrected Portmanteau test in presence of GARCH effects: The robust LBQ stats.

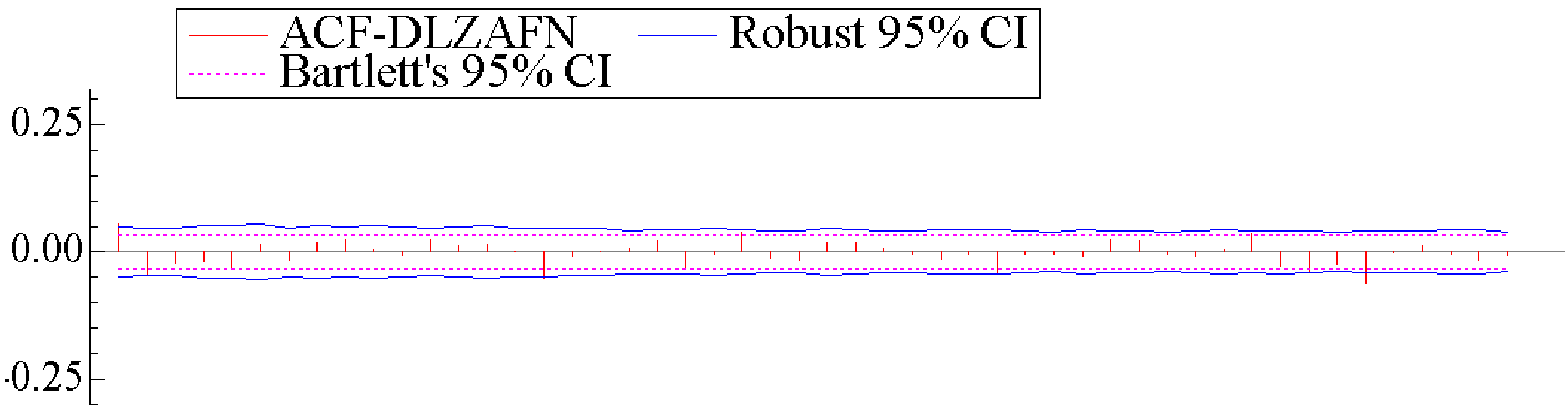
Robust ACF tests in OxMetrics (for interest sake, if you ever use OX again…)





Robust ACF

- As seen, ordinary ACF tests over-reject the null of no serial autocorrelation. F&Z show that it rejects more than double the amount using simulated data.



Testing for conditional heteroskedasticity

- A formal way of GARCH testing is to use the Lagrange Multiplier (LM) test on the squared residuals.
- Engle (1982) suggests basing the LM test on the auxiliary regression:

$$\widehat{\varepsilon_t^2} = \alpha_0 + \alpha_1 \widehat{\varepsilon_{t-1}^2} + \cdots + \alpha_p \widehat{\varepsilon_{t-p}^2}$$

With the R^2 of the regression then used to construct the test stat TR^2. This statistic is then assymptotically distributed as a χ^2 under the null of no ARCH effect up to lag p .

Caution here is that if the mean equation is misspecified, we might mistake persistence in the residuals for persistence in the squared resids…

Fractional Integration

- As finding the right mean equation is vital prior to focussing on the second order moments, let's consider the mean equation some more.
- In many studies it has been shown that series can and do at times exhibit significant serial autocorrelation between periods widely separated in time.
- This is defined as the data having **long memory.**
- Such series are then best modelled using Fractional integration (so as to account for longer term persistence), as first differencing completely removes past dependence structures and focuses on short term dynamics.

Fractional Integration

- Fractional Integration can be achieved by using the ARFIMA class models developed by Granger and Joyeux (1980).
- With $A(L) = AR\ terms\ \ \&\ \ B(L) = MA\ terms$, fractional integration implies:

$$A(L)(1-L)^{\delta}(y_t - \mu_t) = B(L).\,\varepsilon_t$$

With $(1-L)^{\delta}$ being the long memory of the process, and defined as:

$$(1-L)^{\delta} = \sum_{k=0}^{\infty} \frac{\Gamma(\delta+1)}{\Gamma(k+1)\Gamma(\delta-k+1)} L^k$$

$$= 1 - \sum_{k=1}^{\infty} c_k(\delta) L^k$$

$$\left(0 < \delta < 1\ , \qquad c_k = \frac{1}{k}\delta(1-\delta)^{\mathrm{k}-1}, \qquad \Gamma = Gamma\ function \right)$$

Fractional Integration

- Standard unit root tests are unable to effectively distinguish between truly $I(1)$ series and series displaying structural breaks, or strong dependence on observations in the past – i.e. not necessarily immediate past but longer term memory.
- The KPSS unit root test allows for the series to be fractionally integrated and should be used in the presence of potential long memory

```
KPSS Test without trend and 2 lags
H0: DLZAFN is I(0)

KPSS Statistics: 0.0671028
Asymptotic critical values of Kwiatkowski et al. (1992), JoE, 54,1, p. 159-178

    1%    5%   10%
 0.739 0.463 0.347
```

Fractional Integration

- Although δ can take on any values, the series is both stationary and invertible if $\delta \in (-0.5\ , 0.5)$. If it exceeds 0.5, it is nonstationary as it then possesses infinite variance (see Granger and Joyeux). Thus:
- $\delta \in [0; 0.5) \rightarrow$ the autocorrelations decay hyperbolically to zero (as opposed to the geometric decay of ordinary ARMA process) – thus accounting for LT memory
- $\delta \in [-0.5\ ; 0) \rightarrow$ the process exhibits LR negative dependence (anti-persistence).
- $\delta \rightarrow 0$, implies the process exhibits only short term memory (stationary and invertible ARMA)
- $\delta \in [0.5; 1) \rightarrow$ the proc has strong persistence, probably requiring FD.

Fractional Integration

- A long memory process is thus $I(d)$, and should be accounted for by fitting an ARFIMA model.

- With financial returns series, however, this assumption might be a bit strong – as it effectively then assumes long memory in returns, and by definition thus violates weak market efficiency…

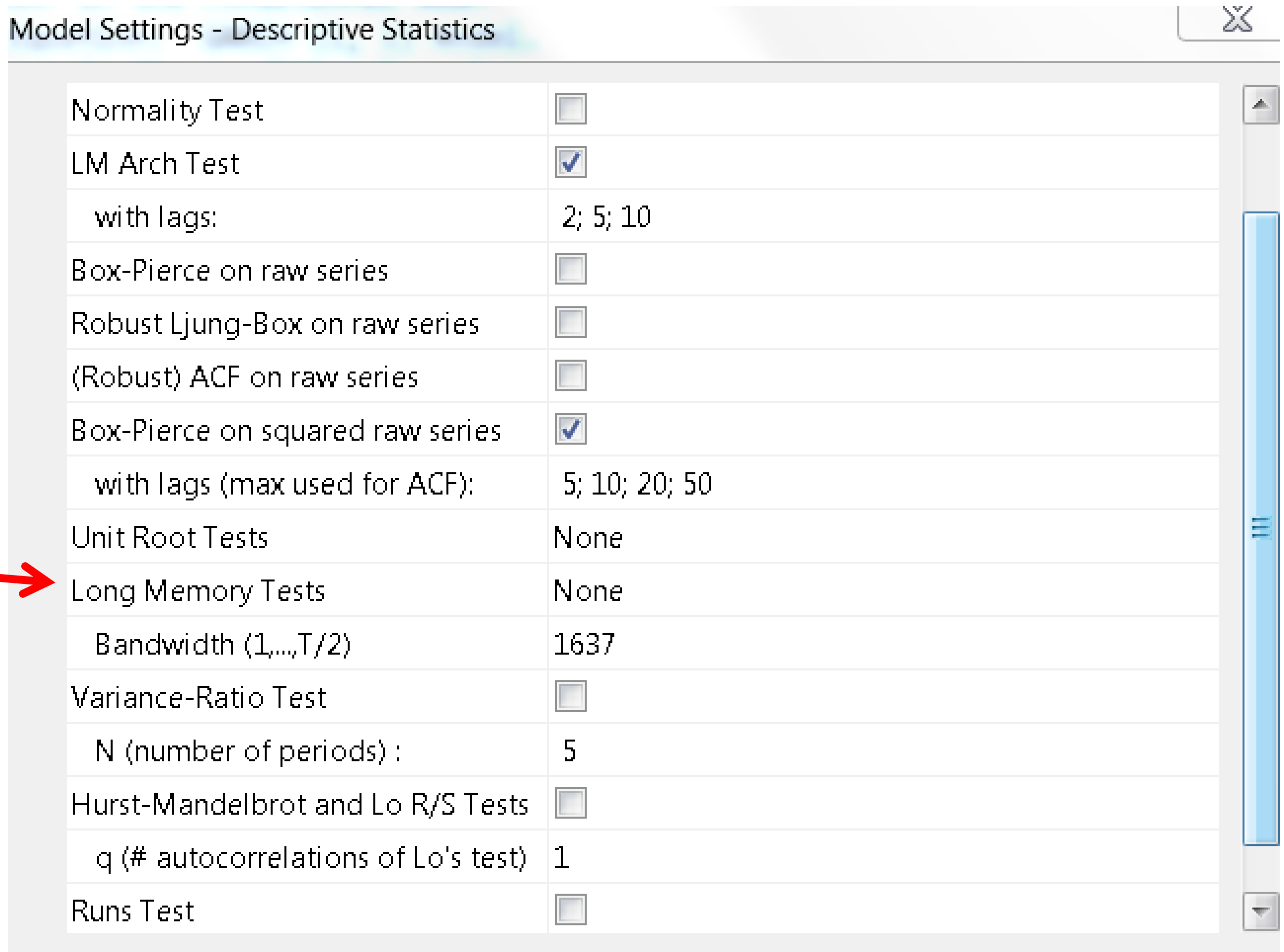
- Here is the KPSS and Long memory test for SA Financials series:

```
KPSS Test without trend and 2 lags
H0: ZAFN is I(0)

KPSS Statistics: 98.3249
Asymptotic critical values of Kwiatkowski et al. (1992), JoE, 54,1, p. 159-178

    1%    5%   10%
 0.739 0.463 0.347
```

Long Memory test:

```
----  Log Periodogram Regression ----
d parameter              1.00482 (0.0170425) [0.0000]
No of observations: 3275; no of periodogram points: 1637
```

Test for long memory (again in Ox)



Model Settings - Descriptive Statistics

Normality Test	☐
LM Arch Test	☑
with lags:	2; 5; 10
Box-Pierce on raw series	☐
Robust Ljung-Box on raw series	☐
(Robust) ACF on raw series	☐
Box-Pierce on squared raw series	☑
with lags (max used for ACF):	5; 10; 20; 50
Unit Root Tests	None
Long Memory Tests	None
Bandwidth (1,...,T/2)	1637
Variance-Ratio Test	☐
N (number of periods) :	5
Hurst-Mandelbrot and Lo R/S Tests	☐
q (# autocorrelations of Lo's test)	1
Runs Test	☐

Fractional Integration

- So for now let's assume the standard dlog suffices for our series.
- Typically, for financial returns series – the data exhibits strong remaining first order persistence, requiring an AR(1) to be fitted. This will then be the μ_t mean process.

- Let's now clean the residuals of the mean process for remaining serial autocorrelation…

Generalizing the ARCH process… GARCH

- Bollerslev (1986) generalized the approach proposed by Engle by allowing the conditional variance term (h_t) to display an ARMA (p,q) process.
- That implies, a GARCH (p,q) series would have the following form:

$$y_t = \alpha + \mu_t + \varepsilon_t$$

$$\varepsilon_t = h_t . \eta_t \;\; ;$$

$$h_t = \sqrt{\alpha_0 + \underline{\sum_{i=1}^{p} \boldsymbol{\beta_i h_{t-i}}^{\mathbf{2}}} + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}{}^2} \quad ; \quad \eta_t \sim WN \sim N(0,1)$$

With constraints : $\alpha_0 > 0, \;\; \alpha_i > 0, \;\; \beta_i > 0,$

$$0 < \sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i < 1.$$

Also, the ARMA and GARCH orders (of x_t and h_t respectively) need not equate!

It follows directly that GARCH(0,1) is equivalent to an ARCH(1) specification

Generalizing the ARCH process… GARCH

- Bollerslev (1986) generalized the approach proposed by Engle by allowing the conditional variance term (h_t) to display an ARMA (p,q) process.
- That implies, a GARCH (p,q) series would have the following form:

Some Return series (Dlog.ret.resids). *(NB: No Unit Root)*

$$y_t = \alpha + \mu_t + \varepsilon_t$$

Some ARIMA model, e.g. *(NB: No remaining AC or PAC)*

$$\varepsilon_t = h_t . \eta_t \ ;$$

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^{p} \boldsymbol{\beta_i h_{t-i}}^{\mathbf{2}} + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^{\ 2}} \ ; \quad \eta_t \sim WN \sim N(0,1)$$

(NB: No remaining AC or PAC)

With constraints : $\alpha_0 > 0, \ \alpha_i > 0, \ \beta_i > 0,$

$$0 < \sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i < 1.$$

Also, the ARMA and GARCH orders (of x_t and h_t respectively) need not equate!

It follows directly that GARCH(0,1) is equivalent to an ARCH(1) specification

Generalizing the ARCH process… GARCH

- Following similar reasoning as for deriving the ARCH model, the **unconditional mean** of ε_t follows as:
- $E_t(\varepsilon_t) = E(\mathrm{h}_t.\eta_t) = E(\mathrm{h}_t)\mathrm{E}(\eta_t) = 0 \;\; (using\ independence)$
- **Unconditional variance:**

$$E(\varepsilon_t{}^2) = E\left(\mathrm{h}_t{}^2\right).1 = E\left(\alpha_0 + \sum_{i=1}^{p} \beta_i \mathrm{h}_{t-i}{}^2 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}{}^2\right)$$

$$= \alpha_0 + \left(\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i\right)\varepsilon_{t-i}{}^2 \quad ;\{as\ E(\varepsilon_{t-1}{}^2) = E\left(\mathrm{h}_{t-1}{}^2\right)\}$$

$$\{as : E(\varepsilon_t{}^2) = E(\varepsilon_{t-1}{}^2)\} \rightarrow E(\varepsilon_t{}^2) = \frac{\alpha_0}{1 - \left(\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i\right)}$$

With $0 < \left(\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i\right) < 1$,

Which is a constant.

Generalizing the ARCH process… GARCH

- **The Autocorrelation function:**

$$E\left(\varepsilon_t, \varepsilon_{t-j}\right) = 0 \quad , \qquad \forall j$$

- **Conditional variance:**

$$E_{t-1}({\varepsilon_t}^2) = E_{t-1}\left({\mathrm{h}_t}^2{\eta_t}^2\right) = E_{t-1}\left({\mathrm{h}_t}^2\right).1 = \boldsymbol{{h_t}^2}$$
$$= \alpha_0 + \sum_{i=1}^{p} \beta_i {\mathrm{h}_{t-i}}^2 + \sum_{i=1}^{q} \alpha_i {\varepsilon_{t-i}}^2$$

- Which is a non-constant, conditional heteroskedastic process with an ARMA form.

Interpreting the coefficients

- Suppose we have the GARCH(1,1) process :

$$h_t^2 = \alpha_0 + \alpha(\varepsilon_{t-1}^2) + \beta(h_{t-1}^2)$$

$\alpha \rightarrow$ extent to which a shock today feeds into tomorrow's volatility, or the response of h_t to new information on an unexpected shock

We can rewrite the top part as: (adding and subtracting $\alpha. h_{t-1}^2$)

$$h_t^2 = \alpha_0 + \alpha.(\varepsilon_{t-1}^2 - h_{t-1}^2) + (\alpha + \beta).(h_{t-1}^2)$$

<u>So that from this form we can interpret the two coefficients as:</u>

- $\alpha \rightarrow$ Impact of the unanticipated shock part (*remember:* $E(\varepsilon_t^2) = E(\eta_t^2).E(h_t^2) = 1.E(h_t^2)$*, so that the difference* $(\varepsilon_{t-1}^2 - h_{t-1}^2)$ *is the unexpected volatility part…)*
- $(\alpha + \beta) \rightarrow$ Degree of Autoregressive decay, i.e. the rate at which the effect of a previous shock dies down on the variance process.
  - Typically we find that $(\alpha + \beta) \rightarrow 1$, implying financial time-series show a slow decay / strong persistence in the volatility process

Parsimony of GARCH

- Using the GARCH approach now let's the modeller fit a more parsimonious conditional variance equation if the ARCH fit requires a large order [i.e. a large *p* for ARCH(*p*)].
- Also note that fitting a GARCH(p,q) process effectively implies fitting an ARMA process on the conditional variance ($\mathrm{h}_t{}^2$) of the series y_t.
- This implies that the correlogram of the series y_t should display stationarity… While the squared residual correlogram (again, representing the volatility **conditional** on past shocks) would display an ARMA structure!
  - The same then applies to fitting the GARCH orders as before → by viewing the correlograms!
- Borreslev (1986) proved the ACF of the squared residuals, resulting from a GARCH(p,q) process, acts like that of an ARMA(**m**,p) process: with $\boldsymbol{m} \rightarrow \max(p, q)$

Determining the order of GARCH

- To check whether conditional heteroskedasticity is present in our series, it seems to follow simply that we fit an ARMA model to series $[y_t]$ to make it stationary, after which we square the residuals, and then take the correlogram of $[{\varepsilon_t}^2]$. It seems then plausible to follow the Borreslev $ARMA(m, p)$ specification, determining the order graphically…
- This is not entirely correct though (although for simplicity is mostly used as a guideline), as the initial model on $[y_t]$ was fitted assuming ***constant conditional variance***. Thus using the residuals of such a model seems at odds with finding the correct order of the ***conditionally heteroskedastic*** variance!
- What we therefore do is, fit subsequent heteroskedasticity models and the ARMA models simultaneously, using Max Likelihood techniques, and then testing which model fits the data closest using some identification criteria like AIC / SBIC…

Correlogram of sqrd ordinary residuals of J203's ARIMA(1,1,1) model:

- The correlogram clearly shows the conditional heteroskedasticity of the squared ordinary resids…
- It also shows a slow decay of the shocks, thus the series probably has a $(\alpha + \beta) \rightarrow 1$

ε^2

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.315	0.315	51.940	0.000
		2	0.227	0.142	79.028	0.000
		3	0.286	0.204	122.01	0.000
		4	0.366	0.252	192.70	0.000
		5	0.286	0.108	235.85	0.000
		6	0.144	-0.05...	246.76	0.000
		7	0.329	0.204	304.21	0.000
		8	0.320	0.106	358.54	0.000
		9	0.167	-0.06...	373.47	0.000
		1...	0.081	-0.10...	376.94	0.000
		1...	0.080	-0.13...	380.36	0.000
		1...	0.234	0.087	409.57	0.000
		1...	0.130	0.026	418.69	0.000
		1...	0.118	0.049	426.15	0.000
		1...	0.122	-0.00...	434.10	0.000
		1...	0.124	-0.01...	442.33	0.000
		1...	0.121	0.045	450.28	0.000

Fitting a GARCH(1,1)

- Although Bollerslev suggests we should be using the correlogram to guide our ordering selection, the model most typically used in financial time-series analysis is the GARCH(1,1) model (see Hansen and Lunde (2004) who survey many types of GARCH models and make this conclusion for financial time-series on aggregate – but they do suggest adding leverage effects too, which we will do in the next session…)

How does the CPU estimate GARCH?

- Think about this – the CPU needs to first estimate a mean equation, then get the residuals, and then fit an autoregressive series on these residuals to find the parameter values for the GARCH process…
- As such, OLS won't cut it… But Engle proved that efficient estimates could be obtained by using a Maximum Likelihood approach to estimate the mean and the variance equations simultaneously!
- Assuming **normality** of the residuals, this implies setting up a Likelihood system as:

Estimation of a GARCH system

- Suppose Y_t follows an AR(1)-GARCH(1,1) system:

$$y_t = c + \phi y_{t-1} + \varepsilon_t, \qquad with \;\; \varepsilon_t \sim N(0, h_t)$$

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

- After specifying the c and ϕ parameters using OLS… We then specify the LogLikelihood function L to be maximized, assuming residuals are Normally distributed, to obtain the $\alpha's$ above:

$$L = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{N}\log(h_t^2) - 1/2\;\frac{1}{2}\sum_{t=1}^{N}[(\varepsilon_{t-1}^2)/h_t^2]$$

- The parameter values that maximize L will be obtained by MLE approach, then be used in the GARCH system…

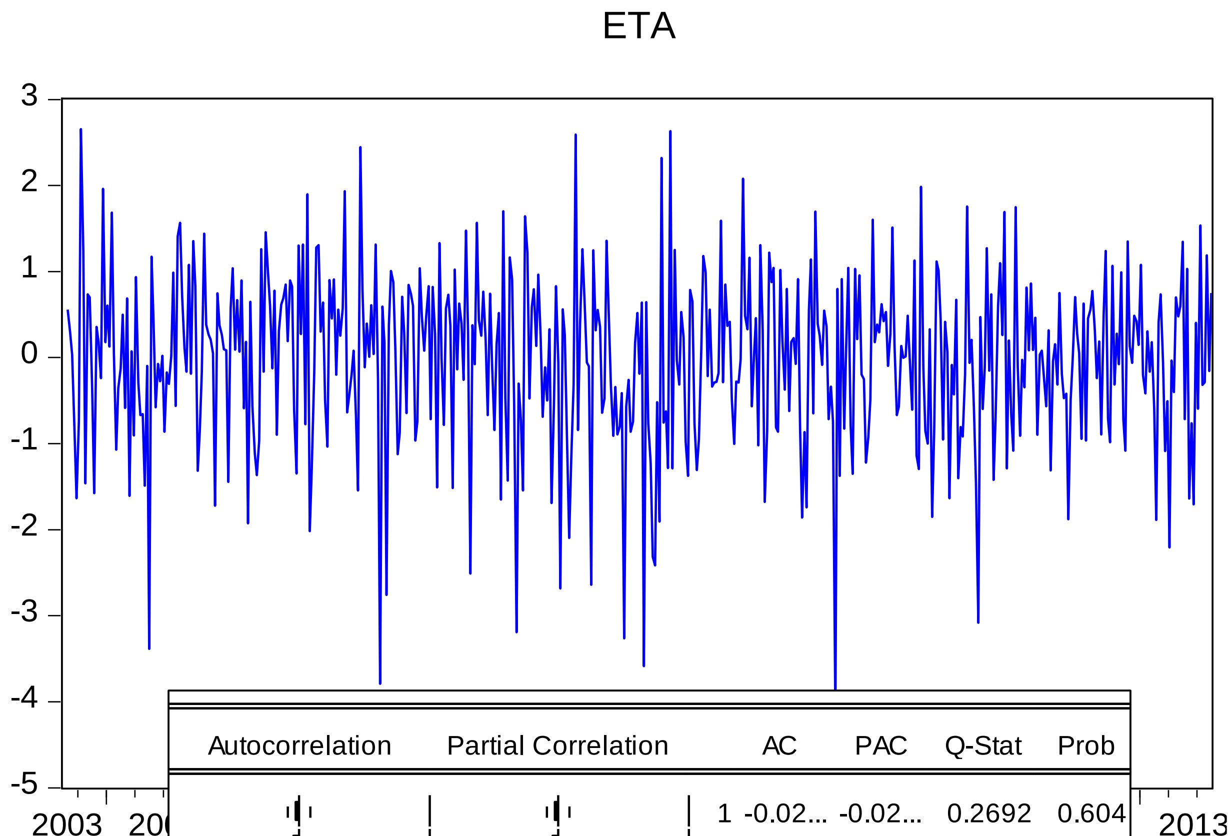
Determining the order of GARCH

After fitting subsequent ARCH / GARCH models, we test its validity by checking:

- The **coefficients** (both for statistical significance and whether they adhere to their constraints – collectively and individually),
- And also **NB :** graphing the squared **standardized** residuals, $[\eta_t{}^2]$, and checking whether all conditional heteroskedasticity has been removed (as η_t is the true stochastic process of the system, it should represent a WN series)
  - This requires testing for WN on the **Standardized residuals.**
- The parameters should also **adhere to its restrictions** (here $\alpha + \beta < 1$ and each is > 0)
- Note from this output, $\alpha + \beta = 0.95$, showing strong autoregressive persistence in the volatility process – **underlining the strong market momentum experience on the JSE over the last ten years**

η_t^2 graphed…

- From this graph we can clearly see that η_t^2 is a closer approximation to WN than the ordinary residual series that we had earlier (when we did not control for conditional heteroskedasticity.
- Graphing its correlogram also suggests it is WN. →



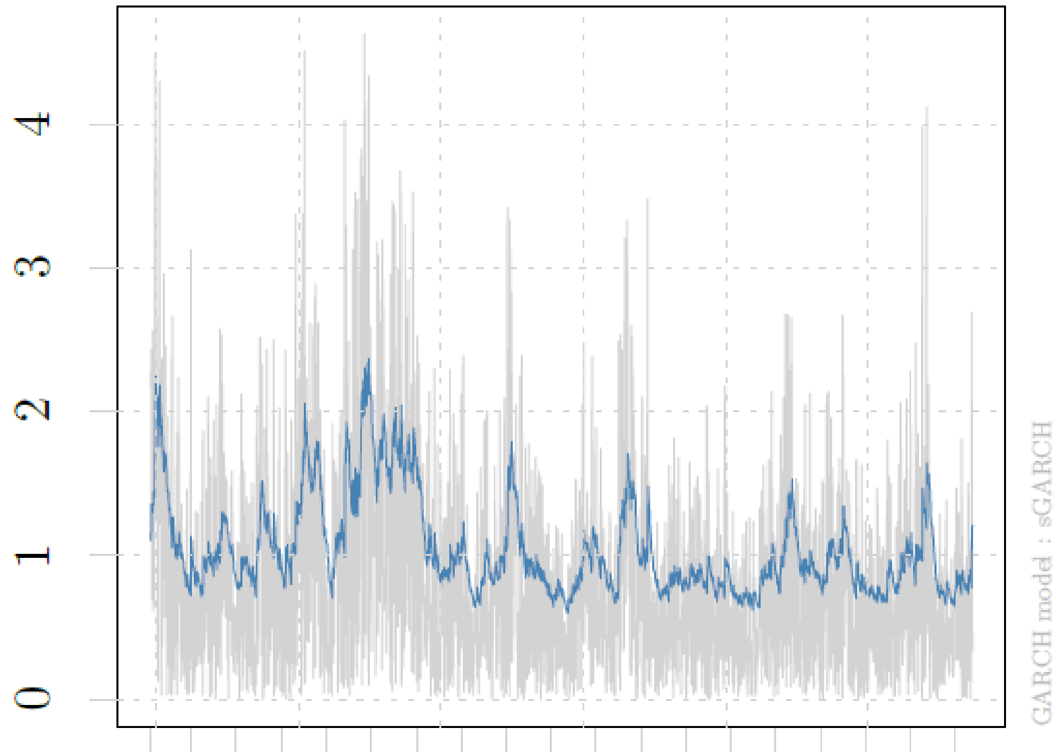
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.02...	-0.02...	0.2692	0.604
		2	-0.02...	-0.02...	0.6667	0.717
		3	0.006	0.005	0.6863	0.876
		4	-0.01...	-0.01...	0.7759	0.942
		5	-0.04...	-0.04...	2.0253	0.846
		6	-0.00...	-0.00...	2.0395	0.916
		7	0.052	0.049	3.4784	0.838
		8	-0.00...	-0.00...	3.4924	0.900
		9	0.031	0.032	3.9932	0.912

Goodness of fit measures for GARCH

- Comparing different ARCH/GARCH fittings to the model in an effort to control for conditional heteroskedasticity, we can use the standard **AIC** and **SBC** measures {which we want ideally as small as possible ($-\infty$)}

- However, as the sum of squared residuals (SSR) $= \sum {\varepsilon_t}^2$: ordinary goodness of fit measures may not apply!
- As the conditional heteroskedasticity has now been controlled for, we should rather use : $SSR' = \sum {\eta_t}^2$, the sum of squares of the **standardized** resids (which we'd like to minimize).
- Another measure that we can use is Maximum Likelihood (L).
  - **It follows that models with large L-values, tend to have low SSR' values.**

Conditional SD (vs —returns—)



After fitting ARCH / GARCH, how is my model affected?!

- Since the mean of each ordinary residual is zero $\{E_t(\varepsilon_t) = 0\}$, the optimal $j - step\ ahead$ forecast of $[y_t]$ is not affected by the presence of the ARCH/GARCH error specifications (These forecasts we make simply by using our model μ_t)
- The size of the confidence interval surrounding the forecast is, however, affected (and at times strongly affected) by the conditional volatility, allowing us to account for periods of greater uncertainty in our forecasts.
- Also the t-tests for the coefficients will be more accurate, as the conditional heteroskedasticity, if left unaccounted for, biases the significance tests toward rejection of the Null (significance)…
  - **Graphically, this means…**

Forecasting Conditional Variance

- Fitting a GARCH model on the conditional variance of a time series also allows us to forecast conditional volatility very simply (and thus forecast market momentum).
- This is especially useful for traders in financial markets who are interested in forecasting the rate of return to holding an asset for a short period of time.
- In particular, if an investor holds an asset from time $t \rightarrow t+1$, the long run mean (μ_t) would be used together with the conditional variance $E_t({\varepsilon_{t+1}}^2)$, as the unconditional (long run) variance would be of limited use in this case!!
- Let's see how we can forecast conditional volatility one period ahead $[E_t({\varepsilon_{t+1}}^2)]$ when we are at time t.

Forecasting Conditional Variance

- The one-step-ahead forecast follows simply from the definition of GARCH :

$$E_t({\varepsilon_{t+1}}^2) = E_t\left({h_{t+1}}^2{\eta_{t+1}}^2\right) = E_t\left({h_{t+1}}^2\right).1$$
$$= \widehat{\boldsymbol{h_{t+1}}^{\boldsymbol{2}}} = \alpha_0 + \sum_{i=1}^{p} \beta_i {h_{t+1-i}}^2 + \sum_{i=1}^{q} \alpha_i {\varepsilon_{t+1-i}}^2$$

Which, for a GARCH(1,1) process, is:

$$E_t(\widehat{{\varepsilon_{t+1}}^2}) = \widehat{\boldsymbol{h_{t+1}}^{\boldsymbol{2}}} = \alpha_0 + \beta.{h_t}^2 + \alpha.{\varepsilon_t}^2$$

Which is easy to calculate at time (t), for we have ${h_t}^2$ & ${\varepsilon_t}^2$

Forecasting Conditional Variance

- The 2-step-ahead forecast follows by using simple iterations :

$$E_t({\varepsilon_{t+2}}^2) = E_t\left({\mathrm{h}_{t+2}}^2{\eta_{t+2}}^2\right) = E_t\left({\mathrm{h}_{t+2}}^2\right).1$$

$$= \widehat{\boldsymbol{h_{t+2}}^{\mathbf{2}}} = \alpha_0 + \sum_{i=1}^{p} \beta_i {\mathrm{h}_{t+2-i}}^2 + \sum_{i=1}^{q} \alpha_i {\varepsilon_{t+2-i}}^2$$

Because : $E_t({\varepsilon_{t+k}}^2) = E_t\left({\mathrm{h}_{t+k}}^2\right).1$ $(\textit{due to independence of } \mathrm{h}_t \,\&\, \eta_t)$

$$\widehat{\boldsymbol{h_{t+2}}^{\mathbf{2}}} = \alpha_0 + \left(\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i\right) {\mathrm{h}_{t+2-i}}^2.$$

If we let $\left(\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i\right) = \sum C_i$, it follows simply that :

$$\widehat{\boldsymbol{h_{t+2}}^{\mathbf{2}}} = \alpha_0 + \sum_{i=1}^{\max(p,q)} C_i {\mathrm{h}_{t+2-i}}^2$$

Forecasting Conditional Variance

Thus using iteration: $\widehat{\boldsymbol{h}_{\boldsymbol{t+j}}}^{\mathbf{2}} = \alpha_0 + \sum_{i=1}^{\max(p,q)} C_i \mathrm{h}_{t+j-i}{}^2$

$$= \alpha_0[1 + (C_1) + (C_i)^2 + \cdots + (C_j)^{j-1}] + C_j{}^j \mathrm{h}_{t+j-i}{}^2$$

Rewritten, the $\boldsymbol{j-}$ period ahead forecast of the volatility process using a GARCH(1,1) process and recursive substitution & the law of iterated expectations, is:

$$E_t(h_{t+j}) = (\alpha + \beta)^j . \left(h_t^2 - \frac{\alpha_0}{1 - \alpha - \beta} \right) + \frac{\alpha_0}{1 - \alpha - \beta}$$

Which, if $j \to \infty$, $E_t(h_{t+j}) \to \frac{\alpha_0}{1-\alpha-\beta}$, which means the iterated forecast collapses to the unconditional (LR) variance, which is a ***very desirable*** attribute of the GARCH models… (which others, such as the EWMA, does not exhibit)

Forecasting Conditional Variance

Note:

- we can continue this recursive forecasting process indefinitely, **but** it should be obvious that the further we forecast the less accurate it becomes:

Interpreting the Forecast:

- If say : $\widehat{\boldsymbol{h_{t+2}}}^{\boldsymbol{2}}$ is very large, it would mean the unconditional forecast of $[y_t]$ for period $(t+2)$, found using the ARMA-model $[\mu_t]$, has a large probability of being inaccurate,: this is **because** we expect the volatility in error terms seen recently **to continue** [i.e. we expect the market momentum to persist into the next period…]

Frequency and DGP…

- Higher frequency data (say daily) tend to be more volatile and has more noise that might not correspond to LR fundamental behaviour.
- This needs to be taken into consideration when looking for GARCH effects.
- Also, it can happen that GARCH effects are picked up as a result of breaks in the unconditional variance and not necessarily as a result of conditional heteroskedasticity (taking several periods and testing for GARCH on each, as opposed to the entire sample, might solve this problem)...
- Thus a good robustness check is required in establishing GARCH effects for a sample.
  - This might be especially prevalent in monthly / annual data

Pro's and cons of fitting ARCH / GARCH

Strengths of using ARCH / GARCH:

- ARCH provides us with a means of controlling for conditional heteroskedasticity, which is particularly useful when studying financial data / data that exhibit periods of momentum / volatility clustering.
- It allows us to forecast volatility into a future period, allowing the modeler to adjust the confidence interval band and be mindful of potential future volatility clustering.

Weaknesses of using ARCH / GARCH

- Positive and negative shocks have the same effect on forecasting volatility : Not necessarily the case in practice : negative & positive momentum in stock markets, e.g., behave differently!
- There are strict parameter constraints that must suffice
- It does not provide us with any insight into the **source** of the conditional variance (merely a forecast based on historic data)
- ARCH models, more often than not, tend to over-predict volatility, as they respond slower to large and isolated shocks. They also only predict shock persistence, not the initial shock.

Summary

- 1. Modeling the mean effect (ARIMAX specs, or whichever) and testing for ARCH effects
  - Ho: no ARCH eects versus Ha : ARCH eects
- Use Q-statistics of squared residuals
- 2. Order determination
  - Use PACF of the squared residuals. (In practice, simply try some reasonable order).
- 3. Estimation: Conditional MLE
- 4. Model checking:
  - Q-stat of standardized residuals and squared standardized residuals. Skewness & Kurtosis of standardized residuals.
  - **R provides many plots for model checking and for presenting the results, mas we will see in the tuts**.

Many Packages in R

- We will use RUGARCH and RMGARCH packages in R. But there are many others. The following code will, e.g., fit a GARCH(11) model:

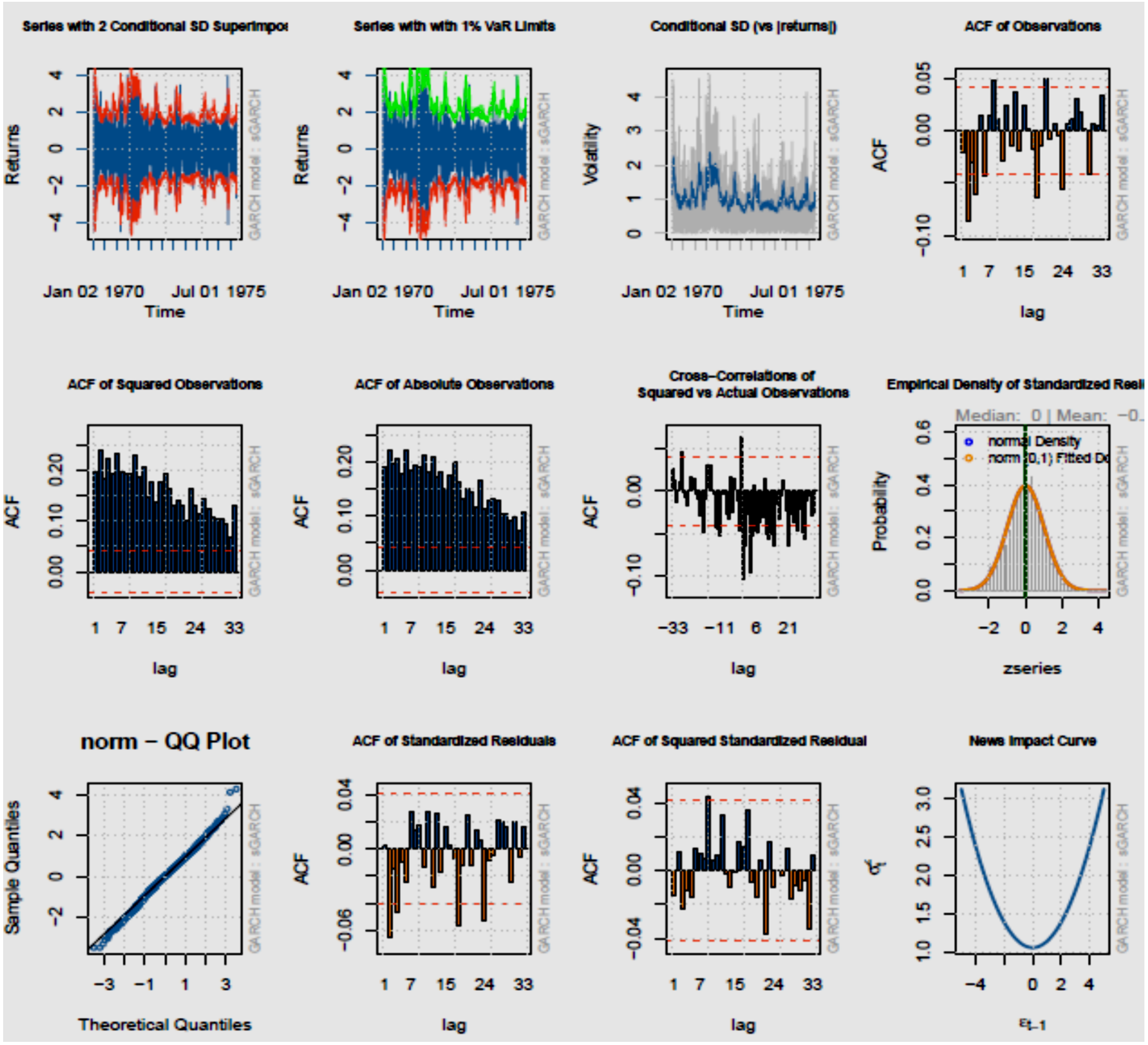
```
> library(fGarch)
> # Create log returns for your series. Test for remaining UR, and thereafter control for remaining serial persistence by fitting, e.g., an ARIMA model. After doing so, get the residuals (log.ret.resids) and do the following:
> acf(log.ret.resids)  # Should show WN
> acf(log.ret.resids^2) # Should show persistence in 2nd moment…
> pacf(log.ret.resids^2) # Same here. Note these two stats motivate fitting garch model…
> Box.test(log.ret.resids ^ 2, lag = 10,type = 'Ljung') # More formal test…
> m1=garchFit(~garch(1,1), data = log.ret.resids, trace = F)
> summary(m1)
```

Plotting in fgarch in R

- > plot(m1)
- Make a plot selection (or 0 to exit):
- 1: Time Series
- 2: Conditional SD
- 3: Series with 2 Conditional SD Superimposed
- 4: ACF of Observations
- 5: ACF of Squared Observations
- 6: Cross Correlation
- 7: Residuals
- 8: Conditional SDs
- 9: Standardized Residuals
- 10: ACF of Standardized Residuals
- 11: ACF of Squared Standardized Residuals
- 12: Cross Correlation between r^2 and r
- 13: QQ-Plot of Standardized Residuals

- We will see many other plotting capabilities in RUGARCH as well..

Standard RUGARCH plots in R



UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

GARCH model Variants

Strong persistence in conditional volatility

- Most volatility models show strong persistence (i.e. $\alpha + \beta \to 1$).
- As a result, the integrated GARCH (IGARCH) model has been proposed when the strength of persistence resembles a unit root process in the conditional variance equation (hence the "integrate" part).
- From last week, remember that the conditional forecast of volatility (j - periods ahead) is:

$$E_t(h_{t+j}) = (\alpha + \beta)^j . \left(h_t^2 - \frac{\alpha_0}{1 - \alpha - \beta}\right) + \frac{\alpha_0}{1 - \alpha - \beta}$$

IGARCH

- If now we have $(\alpha + \beta \rightarrow 1)$, the conditional estimate would be (we omit the proof):

$$E_t\left(h_{t+j}\right) = h_t^2 + j.\alpha_0$$

Which very closely resembles a random walk process with a drift…

As such, the unconditional variance does not converge and therefore the process does not have a stationary covariance series. Nelson did show, however, that a such a process can still be strictly stationary otherwise…

IGARCH

- The IGARCH model then restricts that parameters of the GARCH model to sum to one, and it drops the coefficient:

$$h_t^2 = \sum_{i=1}^{p} \beta_i \mathrm{h}_{t+1-i}{}^2 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t+1-i}{}^2$$

With : $\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i = 1$

- Notice that this constraint implies the GARCH process acts like an autoregressive series with a **unit root (but it isn't, as it is a deterministic solution with no residuals…)**
- The IGARCH process then accounts for this unit root **explicitly**

Despite many series showing very strong volatility persistence, in practice the IGARCH form is regarded as a highly unlikely volatility process design and as such is not often used…

RiskMetrics

- J.P. Morgan released a technical not in October 1994 describing its own internal market risk management methodology, termed RiskMetrics^{TM}.
- The simplicity of the approach is regarded as its strength, with its obvious appeal to non-technical practitioners.
- The approach is very simply an IGARCH(1,1) model, with the parameters fixed:

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

Typically, $\alpha = 0\ and\ \lambda = 0.94$ for daily and 0.97 for weekly data.

This is the basic model. JP Morgan (and other institutions) often use variants thereof.

Dealing with asymmetries…

- Up to now, the conditional variance equation has only considered the magnitudes of past residuals and ignored the signs.
- But as Black showed that leverage effects matter for mean equations (where negative returns show greater persistence), Glosten, Jagannathan and Rungle (1989) showed that volatility models show similar asymmetry.
- As such they proposed the **GJR-GARCH** model (which is a special case of the TARCH model), which explicitly controls for **sign** in past residuals by introducing an indicator variable into the variance equation.

GJR-GARCH

$$h_t^{\,2} = \alpha_0 + \alpha_1(\varepsilon_{t-1}{}^2) + \varphi.\,\mathrm{I}_{t-1}(\varepsilon_{t-1}{}^2) + \beta h_{t-1}{}^2$$

Where : $\mathrm{I} \begin{array}{ll} = 1 & if\ \varepsilon_{t-1} < 0 \\ = 0 & if\ \varepsilon_{t-1} \geq 0 \end{array}$

This implies that for a negative shock in $t-1$, (ε_{t-1}), the impact on the conditional variance in t is:

$$h_t^{\,2} = (\alpha_1 + \varphi.\,\mathrm{I}_{t-1})(\varepsilon_{t-1}{}^2) + \beta h_{t-1}{}^2$$

which is **larger** if $\phi \rightarrow$ is positive.

When regressing the model, a significant $t-$statistic for φ implies the data contains a **leverage effect**. Typically we find that ϕ would be positive, indicating increased volatility persistence if the past shock was negative…

TARCH model

- The more general form of the GJR model is the TARCH model.
- It basically suggests that there exists a **threshold effect**, where if residuals are larger than T, the conditional variance persistence increases.
- The GJR-GARCH is therefore a TARCH with a threshold zero $(T = 0)$

EGARCH model

- Another model that controls for asymmetry and does not need to impose non-negativity constraints in the variance equations, is Nelson's (1991) EGARCH model:

$$\ln(h_t^2) = \beta_0 + \beta_1 \ln(h_{t-1}^2) + \gamma_1 . \frac{\varepsilon_{t-1}}{h_t} + \gamma_2 \{ \left| \frac{\varepsilon_{t-1}}{h_t} \right| - \sqrt{\frac{2}{\pi}} \}$$

With $E\left(\frac{\varepsilon_{t-1}}{h_t}\right) = \sqrt{\frac{2}{\pi}}$ if normal distribution is assumed.

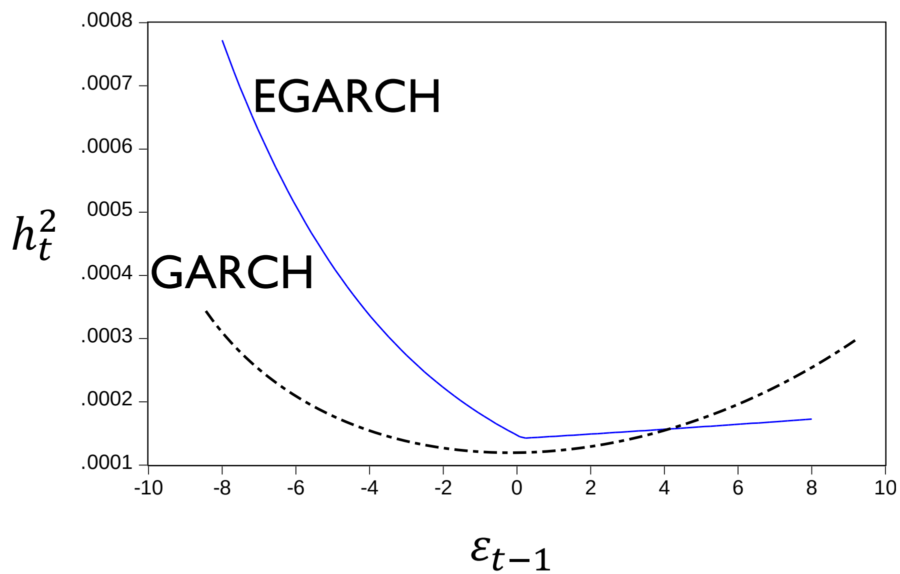
So that EGARCH always produces a **positive** conditional variance requiring no restrictions on parameters (except that $|\beta_1| < 1$).

Now, as $\left|\frac{\varepsilon_{t-1}}{h_t}\right|$ and $\frac{\varepsilon_{t-1}}{h_t}$ are included, h_t^2 will be asymmetrically distributed across positive / negative residuals (so that if $\gamma < 0$, we see leverage effects…)

News Impact Curve (NIC)

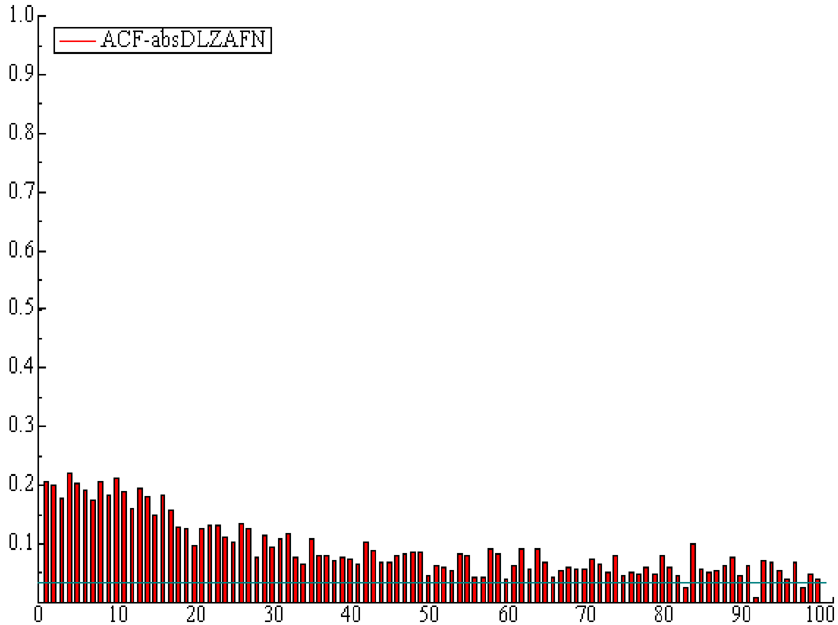
- To get a visual indication of the asymmetric impact of the volatility process to negative and positive shocks, we can draw a NIC. The curve is drawn for the estimated coefficient values of a series below.
- This NIC is for the JSE ALSI (J203) using a EGARCH(1,1) approach:

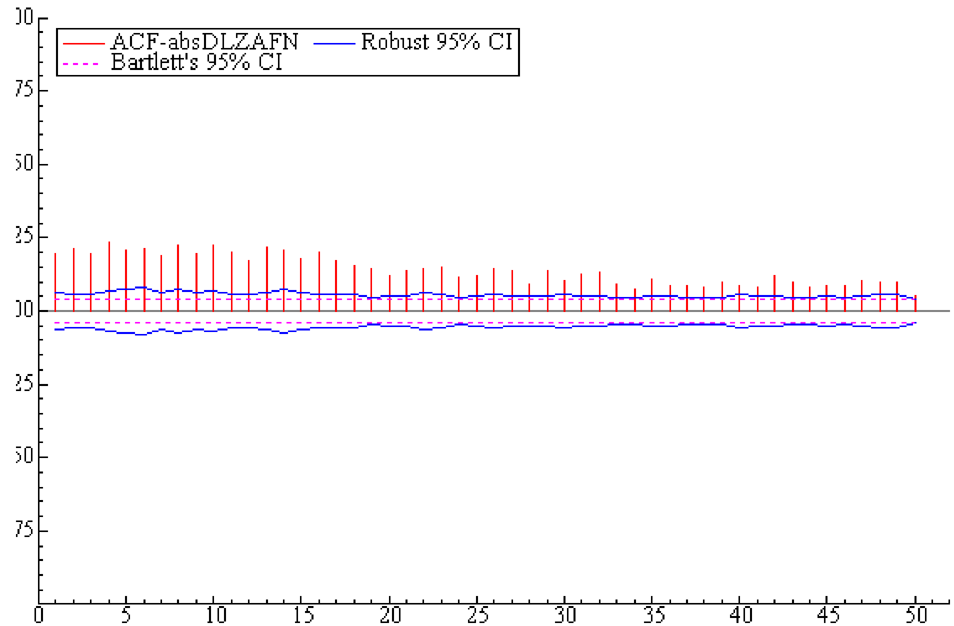
EGARCH thus accounts for both Leverage and Level impact of shocks to volatility persistence



APARCH

- Ding, Granger and Engle (1993), propose a Box-Cox transformation on the std residual series – motivated by Taylor's finding that absolute returns are positively autocorrelated at long lags – indicating longer memory in the series. See below the ACF of **abs(DLZAFN)**:





APARCH

- APARCH models thus look as follows:

$$\sigma_t^{\delta} = \omega + \sum_{i=1}^{q} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^{\delta} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^{\delta}$$

With $\delta > 0$ and $\gamma \in (-1\,;1)$.

The closer $\delta \to 1$, the **longer** the memory process…

```
Mean Equation:  ARMA (1, 0) model.
No regressor in the conditional mean
Variance Equation:  APARCH (1, 1) model.
No regressor in the conditional variance
Normal distribution.

Strong convergence using numerical derivatives
Log-likelihood = 6246.64
Please wait : Computing the Std Errors ...

 Robust Standard Errors (Sandwich formula)
                  Coefficient  Std.Error  t-value  t-prob
Cst(M)               0.000241 0.00042121   0.5723  0.5672
AR(1)                0.031908   0.020288    1.573  0.1159
ARCH(Alpha1)         0.055532  0.0099121    5.603  0.0000
GARCH(Beta1)         0.956288  0.0074677    128.1  0.0000
APARCH(Gamma1)       0.310889   0.096865    3.210  0.0013
APARCH(Delta)        1.224143    0.24735    4.949  0.0000

No. Observations :       2553  No. Parameters  :         6
Mean (Y)         :    0.00011  Variance (Y)    :   0.00055
Skewness (Y)     :   -0.28606  Kurtosis (Y)    :   7.52398
Log Likelihood   :   6246.636
```

APARCH

- The APARCH model nests several other GARCH variants, including:
- ARCH ($\delta = 2,\ \gamma_i = 0 \;\&\; \beta_j = 0$)
- GARCH ($\delta = 2,\ \gamma_i = 0$)
- GJR-GARCH ($\delta = 2$)
- Log-ARCH ($\delta \rightarrow 0$)
- From the output on the previous page (ZA-FN returns), we see:
  - $\delta = 1.22$, it is not significantly different from 1, but indeed significantly different from 2 (considering S.E. = 0.24).
  - This implies – i.s.o. modelling the conditional variance (GARCH), it is more relevant to model the conditional S.D. (as shown by viewing the Absolute Return ACF).
  - This implies greater correlation between Absolute as opposed to squared residuals, i.e. indicates the presence of **long-memory…**
  - The significant and positive γ also shows the presence of Leverage!

FIGARCH

- As discussed for the mean equation earlier, we can model the second order persistence as a fractional integration process.
- This entails that shocks to the second order moment process decay at an exponential rate.
- FIGARCH modelling thus, similar to ARFIMA, replaces the FD operator $(1-L)$ by: $(1-L)^d$.
- Thus, FIGARCH(p,d,q) model is given by:

$$\sigma_t^2 = \underbrace{\omega[1-B(L)]^{-1}}_{\omega^*} + \{1-[1-B(L)]^{-1}\phi(L)(1-L)^d\}\varepsilon_t^2 + B(L)\sigma_t^2$$

Stationary: $0 < \delta < 1$ &:

Covariance stationary if $|\delta| < 0.5$

FIGARCH

- Benefit of the FIGARCH lies in the fact that, in contrast to GARCH ($d = 0$) and IGARCH ($d = 1$) type models: the shocks to the conditional second order moments **do not die out exponentially**, but rather past shocks decay at a **slow, hyperbolic rate** – implying long memory and thus fatter tails.
  - Thus not controlling for long memory – underestimates impact of shocks
- We can also combine FIGARCH and APARCH: FIAPARCH, to control for:
  - Asymmetry
  - Fat tails
  - Long memory

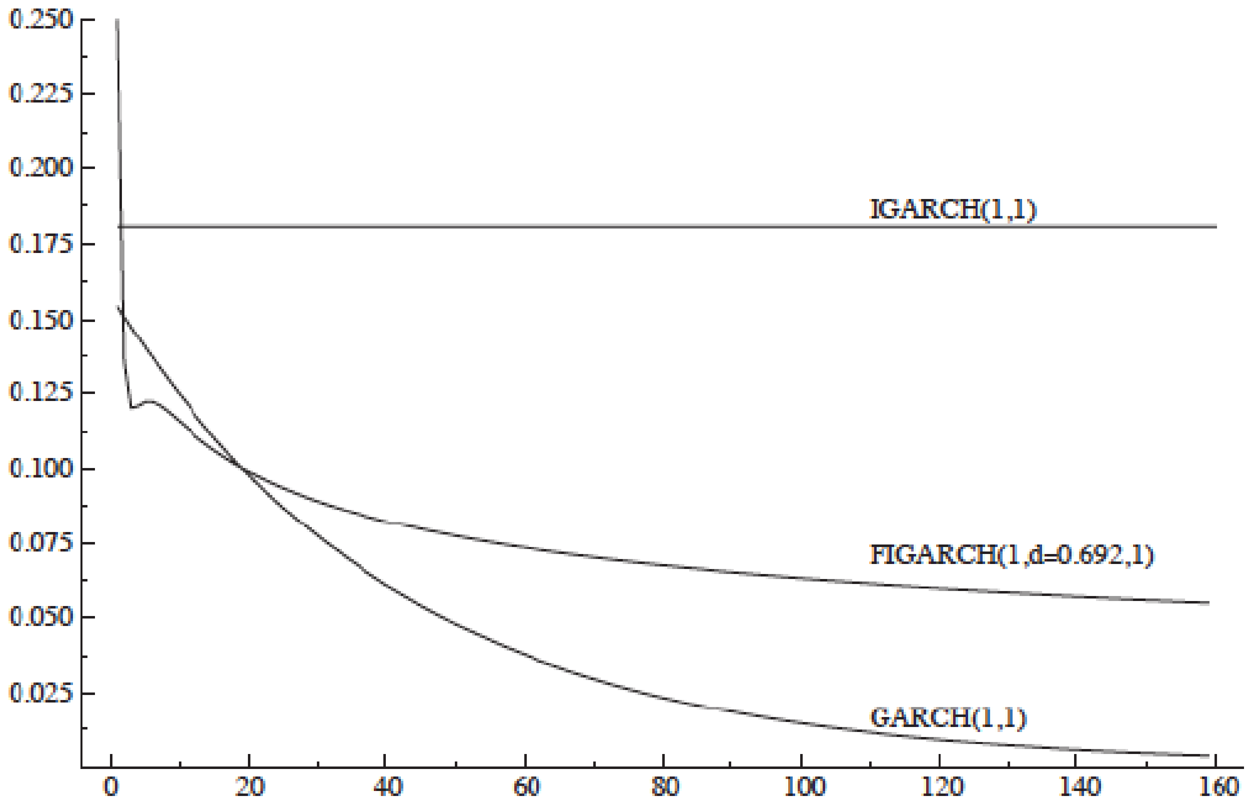


Fig. 1. Cumulative impulse response functions for GARCH, IGARCH and FIGARCH models.

As is clear from the figure (BBM, 1996), using monte-carlo estimates – we see GARCH shocks dying out quickly, IGARCH shocks impacting infinite periods, FIGARCH shocks dying out more slowly (closer to what is seen from ACFs

UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

GARCH-M

Adding GARCH into the **mean equation**

- ***(See: Engle, R.F., D. Lilien and R. Robins (1987), "Estimating Time Varying Risk Premia in the Term Structure: the ARCH-M Model," Econometrica.)***

- From standard MPT theories, investors are risk averse, and as such require larger compensation for holding assets with higher associated risk (or variance of return)
- Initially introduced by the above authors, the GARCH-M model adds to our understanding of the volatility process by including it into the mean equation.
- Thus it approximates how investors price volatility.

GARCH-M model

- The standard GARCH-M specification is:

$$y_t = \mu + \delta h_{t-1} + \varepsilon_t$$
$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta h_{t-1}^2$$

  - So that if $\delta > 0$ and statistically significant, then we can deduce that increased risk (as proxied for by the h_t process) is "priced into" the mean equation.
  - **Thus the parameter can be interpreted as the additional return required by investors for each additional unit of risk (measured by one S.D. above).**

Thus $\delta =$ **risk premium** for an investor to hold y.

- Engle, *et al* initially used the ARCH-m model to study y_t as the difference in return of a 6-month T-bill vs two 3-month T-bills (rolled over for the 6month period). The excess return to holding the longer term bond was given as y_t.
- Note too that EVIEWS allows the modeller to include the ARCH-m factor as variance (h_t^2), S.D. (h_t), or as $\log(h_t^2)$ - depending on preference.

Other models…

- There are a host of GARCH models available that fit the specific characteristics of nearly any type of financial time-series .
- These include NARCH, CGARCH, OGARCH, QGARCH, VARCH, etc. – which will not be discussed in this course.

Including regressors into the volatility equations…

- Of course, similar to adding variables into the mean equation, we can add variables into the **variance equations.**
- This would enable us to test whether a certain variable significantly impacts the volatility process underlying a series.

$$y_t = \mu + \varepsilon_t$$

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta h_{t-1}^2 + \boldsymbol{\gamma} . \boldsymbol{x_t}$$

We can then interpret the significance and sign of γ, which could add to the understanding of the volatility process.

Break in volatility

- Likewise, we can add **indicator variables** that proxy for periods of increased levels of volatility. Thus we can test whether there has been a structural break in the volatility process after the global financial crisis: GFC, by testing the hypothesis that: $H_0: \gamma \neq 0$ in:

$$y_t = \mu + \varepsilon_t$$

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta h_{t-1}^2 + \gamma . GFC$$

$$GFC = \begin{matrix} 1 & \forall \; periods \; in \; GFC \\ 0 & \forall \; periods \; outside \; GFC \end{matrix}$$

Interesting application

- Day and Lewis (1992) studied the out-of-sample forecasting ability of GARCH / EGARCH models at forecasting volatility of stock indices.
- They compare their autoregressive estimates to the *implied volatilities* as given by the aggregate level of volatility implied by options prices.
- Note that options prices has, as input, the strike price, duration, etc. – **and volatility,** which needs to be estimated by the option writer.
- Thus we can compare this implied volatility (as implied by the options prices which we use to extract the implied volatility) to that which is estimated from the EGARCH model.
- They then test whether in their GARCH model (and EGARCH), $\delta \neq 0$ for: $h_t^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta h_{t-1}^2 + \delta(Implied\ Vol)$

Interesting Application

- Christiansen also used univariate volatility models to study volatility spill-over effects (***Christiansen (2007): Volatility Spill-over effects in European Bond Markets)***.
- The approach is simple and intuitive and can basically be described as follows:
- The paper studies shocks to European bond markets in 3 different effect classifications:
  1. Local market (own country)
  2. Regional Market (Europe aggregate)
  3. Global Market (as proxied for by the US bond market).

  She then aims to study whether shocks are transmitted from:

$$US \rightarrow EU \rightarrow EU\ member\ country$$

Study is in two-steps

- Step 1: Fit a univariate AR(1)-GARCH(1,1) model for the US Bond Returns:

$$R_{US,t} = c_0 + c_1 R_{US,t-1} + \varepsilon_{US,t}$$

$$h^2_{US,t} = \omega_{US} + \alpha_{US}\varepsilon_{US,t} + \beta_{US} h^2_{US,t-1}$$

- Step 2: Fit a univariate AR(1)-GARCH(1,1) model for the EU.

$$R_{EU,t} = c_0 + c_1 R_{EU,t-1} + c_{2,t-1} R_{US,t-1} + c_{3,t-1}\varepsilon_{US,t} + \varepsilon_{EU,t}$$

$$h^2_{EU,t} = \omega_{EU} + \alpha_{EU}\varepsilon_{EU,t} + \beta_{EU} h^2_{EU,t-1}$$

- Step 2: Fit a univariate AR(1)-GARCH(1,1) model for the Germany:

$$R_{GE,t} = c_0 + c_1 R_{GE,t-1} + c_{2,t-1} R_{US,t-1} + c_{3,t-1} R_{EU,t-1} + c_{4,t-1}\varepsilon_{US,t} + c_{5,t-1}\varepsilon_{EU,t} + \varepsilon_{GE,t}$$

$$h^2_{GE,t} = \omega_{GE} + \alpha_{GE}\varepsilon_{GE,t} + \beta_{GE} h^2_{GE,t-1}$$

Study is in two-steps

- <u>Step 1: Fit a univariate AR(1)-GARCH(1,1) model for the US Bond Returns:</u>

$$R_{US,t} = c_0 + c_1 R_{US,t-1} + \varepsilon_{US,t}$$

$$h^2_{US,t} = \omega_{US} + \alpha_{US}\varepsilon_{US,t} + \beta_{US} h^2_{US,t-1}$$

- <u>Step 2: Fit a univariate AR(1)-GARCH(1,1) model for the EU.</u>

$$R_{EU,t} = c_0 + c_1 R_{EU,t-1} + c_{2,t-1} R_{US,t-1} + c_{3,t-1}\varepsilon_{US,t} + \varepsilon_{EU,t}$$

$$h^2_{EU,t} = \omega_{EU} + \alpha_{EU}\varepsilon_{EU,t} + \beta_{EU} h^2_{EU,t-1}$$

- <u>Step 2: Fit a univariate AR(1)-GARCH(1,1) model for the Germany:</u>

$$R_{GE,t} = c_0 + c_1 R_{GE,t-1} + c_{2,t-1} R_{US,t-1} + c_{3,t-1} R_{EU,t-1} + c_{4,t-1}\varepsilon_{US,t} + c_{5,t-1}\varepsilon_{EU,t} + \varepsilon_{GE,t}$$

$$h^2_{GE,t} = \omega_{GE} + \alpha_{GE}\varepsilon_{GE,t} + \beta_{GE} h^2_{GE,t-1}$$

Volatility Spill-over

- From the previous slide then, the idiosyncratic shocks ($\varepsilon_{US,t}, \varepsilon_{EU,t}, \varepsilon_{GE,t}$) are assumed independent, **but** this does not apply to the returns:

$$\epsilon_{US,t} = \varepsilon_{US,t}$$

$$\epsilon_{EU,t} = c_{3,t-1}\varepsilon_{US,t} + \varepsilon_{EU,t}$$

$$\epsilon_{EU,t} = c_{4,t-1}\varepsilon_{US,t} + c_{5,t-1}\varepsilon_{EU,t} + \varepsilon_{GE,t}$$

So that she is able to test for the significance of $c_{3,t-1}$, $c_{4,t-1}$ and $c_{5,t-1}$ above and so establish whether there is significant volatility spill-over effects from $US \rightarrow \underrightarrow{EU \rightarrow} Germany$

Reading…

- Go read this interesting article on the importance of including GARCH estimates in autoregressive models (and the problems with falsely rejected $H_0's$ if you don't):
- **Hamilton, J.D. (2008), "Macroeconomics and ARCH, Working paper, UCSD**