ECOM90024 – FORECASTING IN ECONOMICS & BUSINESS

LECTURE 10: GARCH MODELS AND FORECASTING

TODAY’S LECTURE

- The GARCH model

- Forecasting with a conditional volatility model

THE GARCH MODEL

- In the previous lecture we observed that due to the constraints on its parameter values, the ARCH model typically requires many parameters in order to adequately describe the dependence in the conditional volatility process of an asset return.
- One way to reduce the number of parameters required is to specify a generalized ARCH (GARCH) model developed by Bollerslev (1986) which introduces lags of σ_t^2 into the conditional variance equation. For instance, a GARCH(1,1) is specified as

$$\varepsilon_t = \sigma_t \upsilon_t$$

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

$$\upsilon_t \sim_{i.i.d.} N(0,1)$$

- Where $\alpha_0 > 0$, $\alpha_1 \geq 0$, and $\beta_1 \geq 0$

THE GARCH MODEL

- To understand the stochastic properties of the GARCH model, let's first define,

$$\eta_t = \varepsilon_t^2 - \sigma_t^2$$

- Computing it's expected value, we obtain

$$E[\eta_t] = E[\varepsilon_t^2] - E[\sigma_t^2] = E[\sigma_t^2 \upsilon_t^2] - E[\sigma_t^2] = 0$$

- Where the last equality is obtained using the fact that $\upsilon_t \sim_{i.i.d.} N(0,1)$
- Now let's consider the j-th autocovariance of η_t

$$cov(\eta_t, \eta_{t-j}) = E[\eta_t \eta_{t-j}] = E[(\varepsilon_t^2 - \sigma_t^2)(\varepsilon_{t-j}^2 - \sigma_{t-j}^2)]$$

THE GARCH MODEL

- Again, substituting in $\varepsilon_t = \sigma_t \upsilon_t$, we obtain

$$cov(\eta_t, \eta_{t-j}) = E\left[(\varepsilon_t^2 - \sigma_t^2)(\varepsilon_{t-j}^2 - \sigma_{t-j}^2)\right] = E\left[(\sigma_t^2 \upsilon_t^2 - \sigma_t^2)(\sigma_{t-j}^2 \upsilon_{t-j}^2 - \sigma_{t-j}^2)\right]$$

- Multiplying it all out,

$$cov(\eta_t, \eta_{t-j}) = E\left[(\sigma_t^2 \upsilon_t^2)(\sigma_{t-j}^2 \upsilon_{t-j}^2) - (\sigma_t^2 \upsilon_t^2)\sigma_{t-j}^2 - (\sigma_{t-j}^2 \upsilon_{t-j}^2)\sigma_t^2 + \sigma_t^2 \sigma_{t-j}^2\right]$$

- Passing the expectations operator through these additive components and again using the fact that $\upsilon_t \sim_{i.i.d.} N(0,1)$, we obtain,

$$cov(\eta_t, \eta_{t-j}) = E\left[\sigma_t^2 \sigma_{t-j}^2\right] - E\left[\sigma_t^2 \sigma_{t-j}^2\right] - E\left[\sigma_t^2 \sigma_{t-j}^2\right] + E\left[\sigma_t^2 \sigma_{t-j}^2\right] = 0$$

- Therefore we have shown that $E[\eta_t] = 0$ and $cov(\eta_t, \eta_{t-j}) = 0$, which means that it is a white noise series.

THE GARCH MODEL

- Going back to our GARCH(1,1) model, recall that the conditional variance equation is specified as

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

- If we substitute $\sigma_t^2 = \varepsilon_t^2 - \eta_t$, we can rewrite it as

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \varepsilon_{t-1}^2 + \eta_t - \beta_1 \eta_{t-1}$$

- Collecting the terms, we can see that we have an equation that has the same structure as an ARMA(1,1) model!

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + \eta_t - \beta_1 \eta_{t-1}$$

THE GARCH MODEL

- It is then straightforward to show that the unconditional mean of ε_t^2 (or equivalently the unconditional variance of ε_t) is given by,

$$E[\varepsilon_t^2] = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}$$

- We can see from this equation that in order for the innovations ε_t to have a finite and positive variance, we have to impose the restriction that $(\alpha_1 + \beta_1) < 1$

- Since $\alpha_1 \geq 0$, and $\beta_1 \geq 0$, we can see from the equation

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + \eta_t - \beta_1\eta_{t-1}$$

- That a GARCH(1,1) specification characterizes ε_t^2 in such a way that a large realization of ε_{t-1}^2 will tend to be followed by a large ε_t^2, thereby generating the volatility clustering that we so often see in financial data.

THE GARCH MODEL

- Now let's compute the kurtosis of our GARCH(1,1) process, first we start with the conditional fourth central moment of the innovation ε_t

$$E\left[\varepsilon_t^4 \middle| \Omega_{t-1}\right] = E\left[\sigma_t^4 \upsilon_t^4 \middle| \Omega_{t-1}\right] = 3\sigma_t^4$$

- Again, the unconditional fourth central moment is then obtained via the law of iterated expectations

$$E[\varepsilon_t^4] = E\left[E\left[\varepsilon_t^4 \middle| \Omega_{t-1}\right]\right] = 3E[\sigma_t^4]$$

- To compute $E[\sigma_t^4]$ we first recognize that

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

THE GARCH MODEL

- It follows that

$$\sigma_t^4 = (\alpha_0 + (\alpha_1 v_{t-1}^2 + \beta_1)\sigma_{t-1}^2)^2 = \alpha_0^2 + 2\alpha_0(\alpha_1 v_{t-1}^2 + \beta_1)\sigma_{t-1}^2 + (\alpha_1 v_{t-1}^2 + \beta_1)^2\sigma_{t-1}^4$$

- The unconditional expectation is then given by

$$E[\sigma_t^4] = \alpha_0^2 + 2\alpha_0 E[(\alpha_1 v_{t-1}^2 + \beta_1)\sigma_{t-1}^2] + E[(\alpha_1 v_{t-1}^2 + \beta_1)^2\sigma_{t-1}^4]$$

- Since $v_t \sim_{i.i.d.} N(0,1)$, we have that

$$E[(\alpha_1 v_{t-1}^2 + \beta_1)\sigma_{t-1}^2] = E[(\alpha_1 v_{t-1}^2 + \beta_1)]E[\sigma_{t-1}^2]$$

$$E[(\alpha_1 v_{t-1}^2 + \beta_1)^2\sigma_{t-1}^4] = E[(\alpha_1 v_{t-1}^2 + \beta_1)^2]E[\sigma_{t-1}^4]$$

THE GARCH MODEL

- Looking at the first term, we can see easily that since $\varepsilon_{t-1}^2 = \sigma_{t-1}^2 \upsilon_{t-1}^2$ and $\upsilon_t \sim_{i.i.d.} N(0,1)$

$$E[\sigma_{t-1}^2] = E[\varepsilon_{t-1}^2]$$

- So that,

$$E[(\alpha_1 \upsilon_{t-1}^2 + \beta_1)]E[\sigma_{t-1}^2] = (\alpha_1 + \beta_1)\frac{\alpha_0}{1-(\alpha_1+\beta_1)}$$

- Going to the second term, we have that

$$E[(\alpha_1 \upsilon_{t-1}^2 + \beta_1)^2]E[\sigma_{t-1}^4] = E[\alpha_1^2 \upsilon_{t-1}^4 + 2\alpha_1\beta_1 \upsilon_{t-1}^2 + \beta_1^2]E[\sigma_{t-1}^4]$$

- Again, using the fact that $\upsilon_t \sim_{i.i.d.} N(0,1)$, we obtain

$$E[(\alpha_1 \upsilon_{t-1}^2 + \beta_1)^2]E[\sigma_{t-1}^4] = (3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2)\, E[\sigma_{t-1}^4]$$

THE GARCH MODEL

- Putting it all together, we obtain

$$E[\sigma_t^4] = \alpha_0^2 + 2\alpha_0(\alpha_1 + \beta_1)\frac{\alpha_0}{1-(\alpha_1+\beta_1)} + (3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2)\, E[\sigma_{t-1}^4]$$

- Again, if we presume our process to be fourth order stationary, $E[\sigma_t^4] = E[\sigma_{t-1}^4]$, we can rewrite it as,

$$E[\sigma_t^4] = \alpha_0^2\left(1 + \frac{2(\alpha_1+\beta_1)}{1-(\alpha_1+\beta_1)}\right)\frac{1}{(1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2)}$$

- Simplification yields,

$$E[\sigma_t^4] = \frac{\alpha_0^2(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)(1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2)}$$

THE GARCH MODEL

- Therefore

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

- Recall that the kurtosis is given by

$$kurt_\varepsilon = \frac{E[\varepsilon_t^4]}{(E[\varepsilon_t^2])^2}$$

- The square of the variance is simply

$$(E[\varepsilon_t^2])^2 = \left(\frac{\alpha_0}{1-(\alpha_1+\beta_1)}\right)^2 = \frac{\alpha_0^2}{(1-\alpha_1-\beta_1)^2}$$

THE GARCH MODEL

- Putting it all together, we obtain,
- $$kurt_\varepsilon = \frac{E[\varepsilon_t^4]}{(E[\varepsilon_t^2])^2} = \left[\frac{3\alpha_0^2(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)(1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2)}\right] \times \left[\frac{(1-\alpha_1-\beta_1)^2}{\alpha_0^2}\right]$$

- Cancelling terms, we end up with,

$$kurt_\varepsilon = \frac{3(1+\alpha_1+\beta_1)(1-\alpha_1-\beta_1)}{(1-3\alpha_1^2-2\alpha_1\beta_1-\beta_1^2)} = \frac{3(1-(\alpha_1+\beta_1)^2)}{1-(\alpha_1+\beta_1)^2-2\alpha_1^2} > 3$$

Hence we have shown that the GARCH(1,1) model is also leptokurtic

- Where the last inequality is obtained from the fact that $\alpha_1 > 0$ which means that the denominator is smaller than the numerator. From this expression we can see that it must be the case that the GARCH parameters α_1 and β_1 must be such that

$$1-(\alpha_1+\beta_1)^2-2\alpha_1^2 > 0$$

THE GARCH MODEL

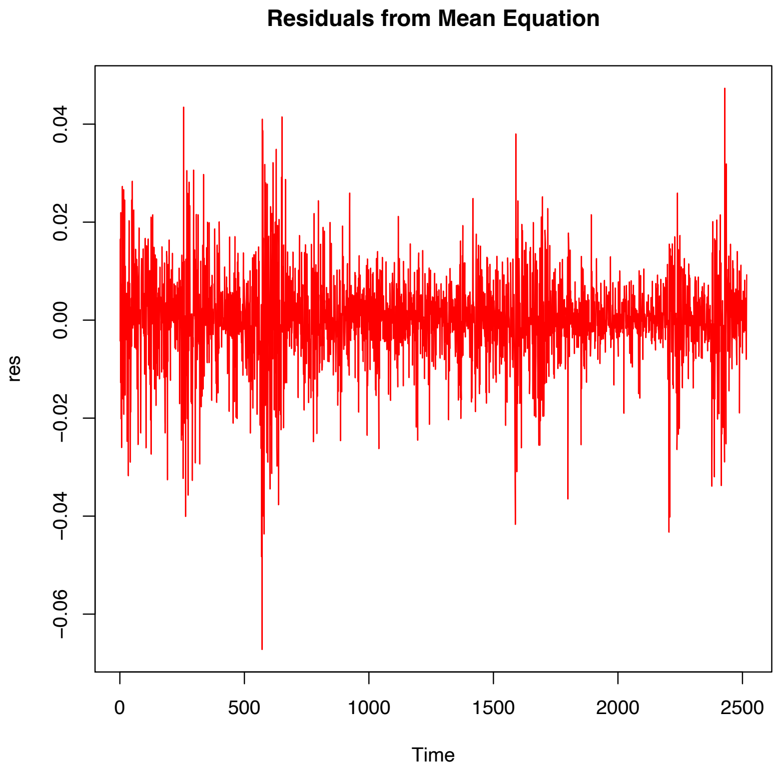
- GARCH models have the ability to account for conditional volatility dependence with much less terms than an ARCH model. In many practical applications, a GARCH order of 2 or lower is usually sufficient!

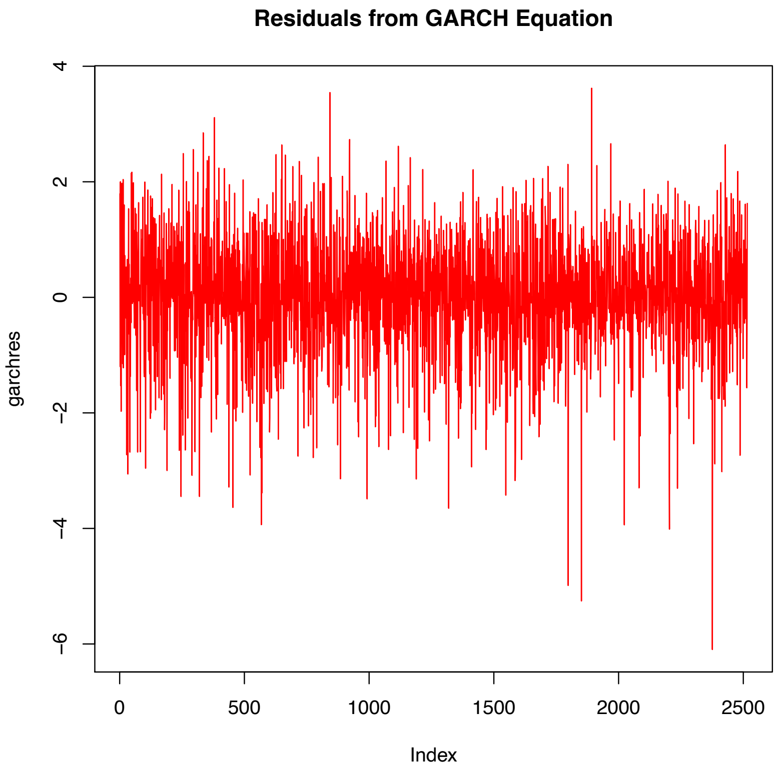
- Let’s go back to our S&P data. In the previous lecture we saw that an ARCH(10) was require to account for the conditional volatility dependence in the data. Let’s use the same data and estimate a GARCH(1,1)

```
Call:
garch(x = res, order = c(1, 1))

Model:
GARCH(1,1)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.0964 -0.5076  0.0169  0.5783  3.6198 

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)    
a0 3.197e-06   3.329e-07    9.602   <2e-16 ***
a1 1.493e-01   1.185e-02   12.602   <2e-16 ***
b1 8.176e-01   1.292e-02   63.267   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

THE GARCH MODEL

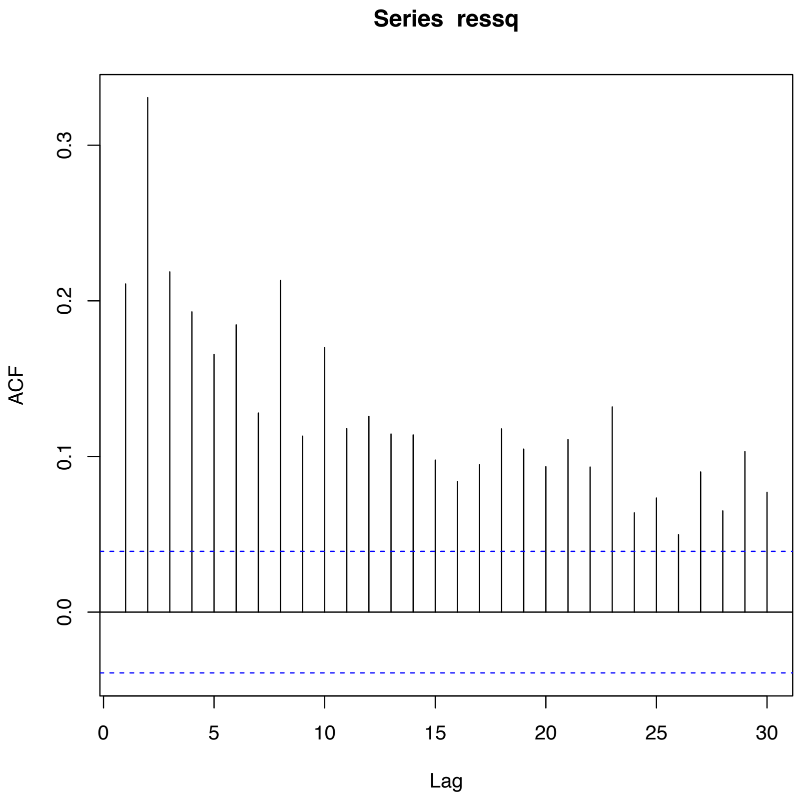
- Here are the residuals generated from the GARCH(1,1) model,

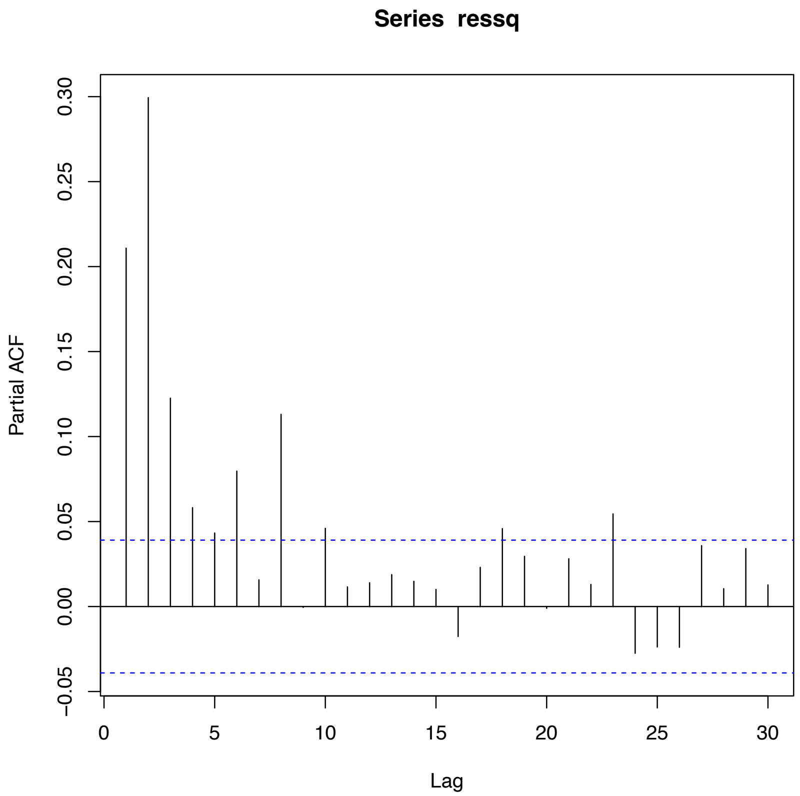




THE GARCH MODEL

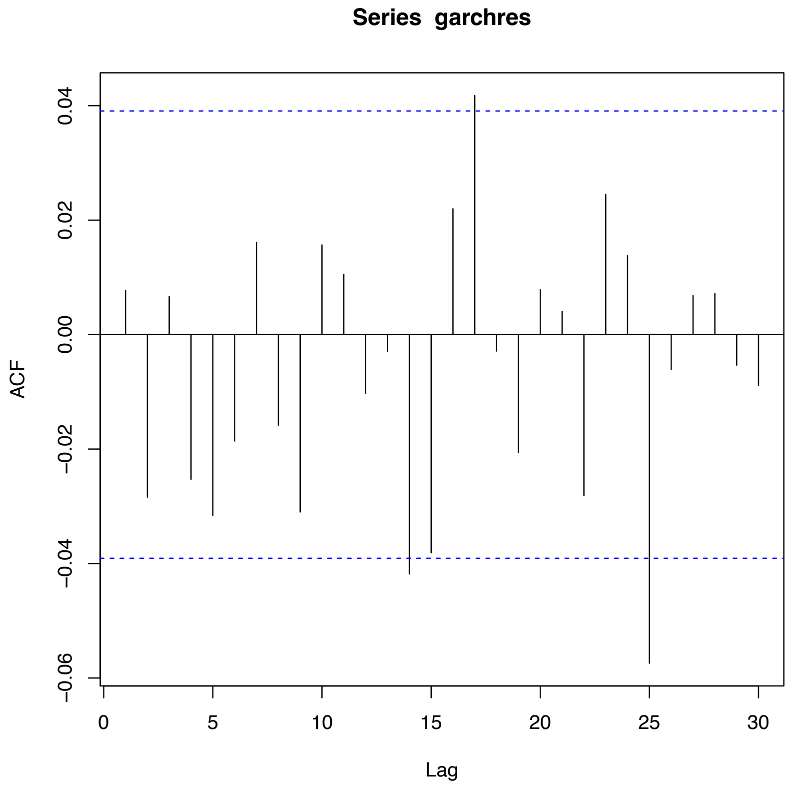
- Recall that there was a lot of dependence in the squared residuals from the mean equation,

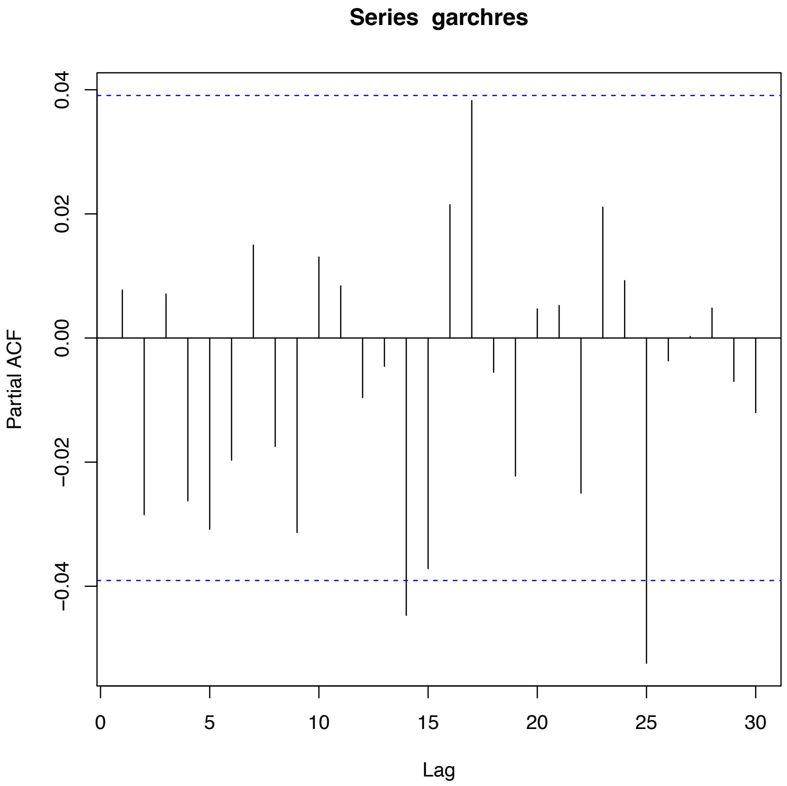




THE GARCH MODEL

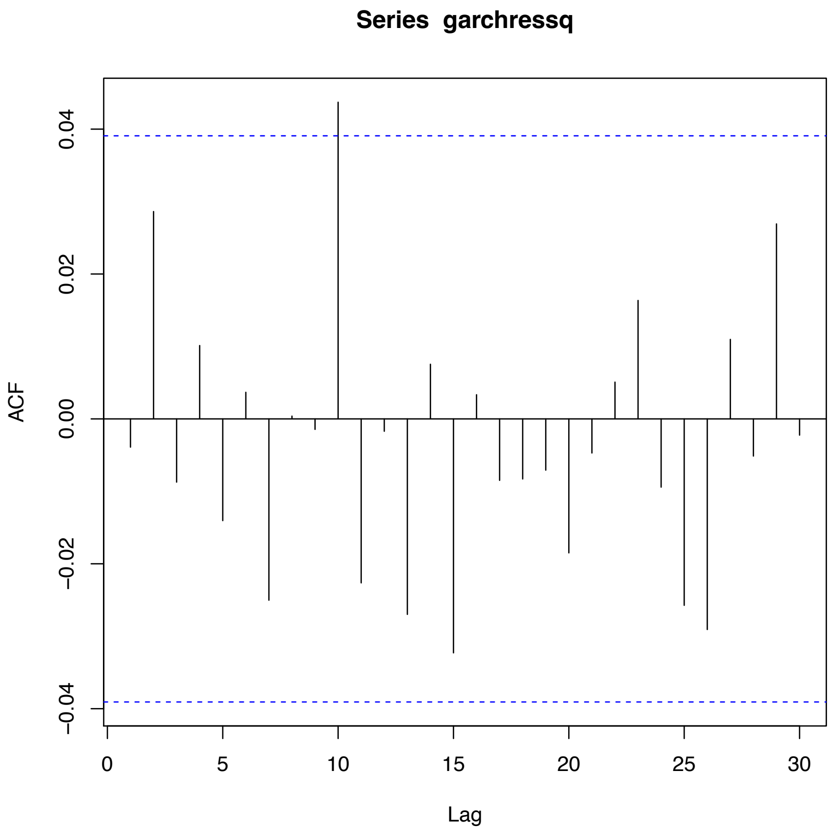
- Here are the sample ACF and PACF of the residuals from the GARCH(1,1) model,

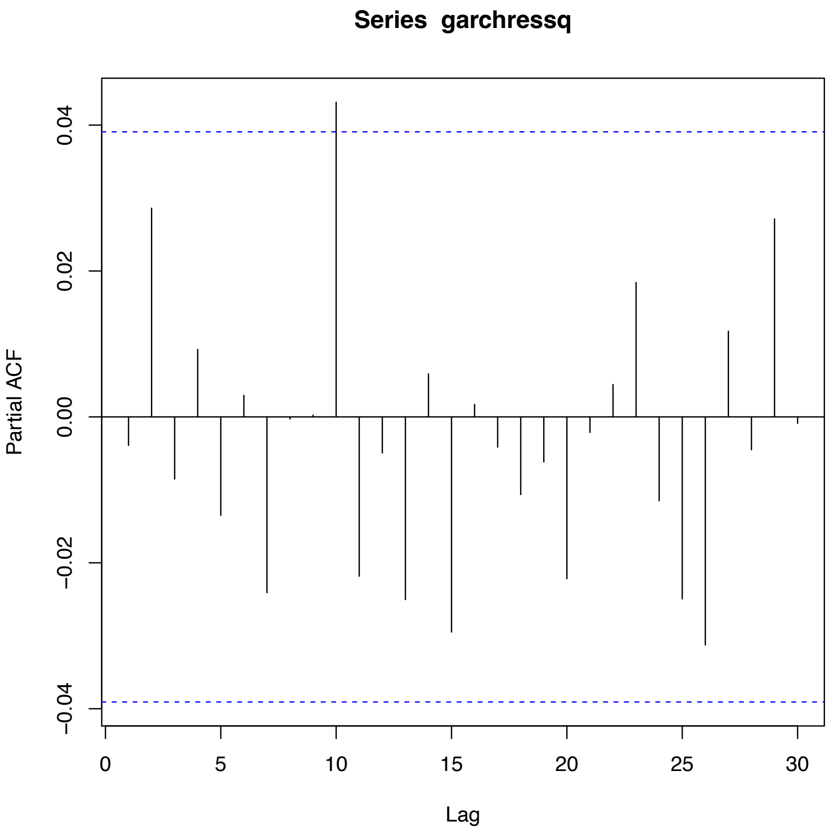




THE GARCH MODEL

- Here are the sample ACF and PACF of the squared residuals from the GARCH(1,1) model,





THE GARCH MODEL

- Performing our Box tests on the squared residuals from our GARCH(1,1) estimation we can clearly see that the dependence in the conditional variance has been removed!

```
> Box.test(garchressq, lag = m, type = "Box-Pierce")

	Box-Pierce test

data:  garchressq
X-squared = 35.606, df = 51, p-value = 0.9499

> Box.test(garchressq, lag = m, type = "Ljung-Box")

	Box-Ljung test

data:  garchressq
X-squared = 35.971, df = 51, p-value = 0.945
```

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Once we have estimated a conditional volatility model, we can proceed to compute forecasts of the conditional variance in a recursive fashion.
- Let's suppose that we have an ARCH(1), then the 1-step ahead forecast is given by,

$$\sigma_{t+1|t}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2$$

- It follows that the 2-step ahead forecast is given by,

$$\sigma_{t+2|t}^2 = \alpha_0 + \alpha_1 E\left[\varepsilon_{t+1}^2 \middle| \Omega_t\right] = \alpha_0 + \alpha_1 \sigma_{t+1|t}^2 = \alpha_0 + \alpha_1(\alpha_0 + \alpha_1 \varepsilon_t^2) = \alpha_0 + \alpha_1 \alpha_0 + \alpha_1^2 \varepsilon_t^2$$

- And the 3-step ahead forecast is given by

$$\sigma_{t+3|t}^2 = \alpha_0 + \alpha_1 \sigma_{t+2|t}^2 = \alpha_0 + \alpha_1(\alpha_0 + \alpha_1 \alpha_0 + \alpha_1^2 \varepsilon_t^2) = \alpha_0 + \alpha_1 \alpha_0 + \alpha_1^2 \alpha_0 + \alpha_1^3 \varepsilon_t^2$$

- Therefore the h-step ahead forecast is given by

$$\sigma_{t+h|t}^2 = \alpha_1^h \varepsilon_t^2 + \alpha_0 \sum_{i=0}^{h-1} \alpha_1^i$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- If we let the forecast horizon become arbitrarily large $h \to \infty$, the conditional variance forecast converges to the unconditional variance

$$\sigma^2_{t+h|t} = \alpha_1^h \varepsilon_t^2 + \alpha_0 \sum_{i=0}^{h-1} \alpha_1^i \to \frac{\alpha_0}{1-\alpha_1}$$

- Now let's suppose that the conditional variance equation follows a GARCH(1,1) process, the 1-step ahead forecast will be given by

$$\sigma^2_{t+1|t} = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2$$

- The 2-step ahead forecast will be given by

$$\sigma^2_{t+2|t} = \alpha_0 + \alpha_1 E\left[\varepsilon^2_{t+1} \middle| \Omega_t\right] + \beta_1 \sigma^2_{t+1|t}$$

- Now we have that

$$E\left[\varepsilon^2_{t+1} \middle| \Omega_t\right] = E\left[\sigma^2_{t+1} \upsilon^2_{t+1} \middle| \Omega_t\right] = E\left[\sigma^2_{t+1} \middle| \Omega_t\right] = \sigma^2_{t+1|t}$$

- Where the second last equality is obtained from the fact that $\upsilon_t \sim_{i.i.d.} N(0,1)$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- It follows that

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

- Simplification yields,

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

- The 3-step ahead forecast is then given by

$$\sigma^2_{t+3|t} = \alpha_0 + \alpha_1 E\left[\varepsilon^2_{t+2}\middle|\Omega_t\right] + \beta_1\sigma^2_{t+2|t} = \alpha_0 + (\alpha_1 + \beta_1)\sigma^2_{t+2|t}$$

- Substituting in for $\sigma^2_{t+2|t}$, we obtain

$$\sigma^2_{t+3|t} = \alpha_0 + (\alpha_1 + \beta_1)[\alpha_0(1 + \alpha_1 + \beta_1) + (\alpha_1 + \beta_1)(\alpha_1\varepsilon_t^2 + \beta_1\sigma_t^2)]$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Simplification yields,

$$\sigma^2_{t+3|t} = \alpha_0(1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2) + (\alpha_1 + \beta_1)^2(\alpha_1\varepsilon_t^2 + \beta_1\sigma_t^2)$$

- From this we can see that the h-step ahead forecast will be given by

$$\sigma^2_{t+h|t} = \alpha_0 \sum_{i=0}^{h-1} (\alpha_1 + \beta_1)^i + (\alpha_1 + \beta_1)^{h-1}(\alpha_1\varepsilon_t^2 + \beta_1\sigma_t^2)$$

- Since $(\alpha_1 + \beta_1) < 1$, again, as $h \to \infty$, the forecast converges to the unconditional variance of the process!

$$\sigma^2_{t+h|t} \to \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Now that we know how to compute forecasts of the conditional variance, we can then proceed to develop forecast intervals that respond to shocks! Let’s start with a simple process,

$$R_t = \mu + \varepsilon_t$$

- Recall that if $\varepsilon_t \sim_{i.i.d} (0, \sigma^2)$, the h-step ahead forecast is then given by,

$$R_{t+h|t} = E\left[R_{t+h} \middle| \Omega_t\right] = \mu$$

- So that the forecast error is given by

$$R_{t+h} - E\left[R_{t+h} \middle| \Omega_t\right] = \varepsilon_{t+h}$$

- And the forecast error variance is given by

$$\sigma_h^2 = \sigma^2$$

- So that the 95% forecast interval is given by

$$\mu \pm 1.96\sigma$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Now let's suppose that our time series R_t has an ARCH(1) structure,

$$R_t = \mu + \varepsilon_t$$

$$\varepsilon_t = \sigma_t \upsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$
$$\upsilon_t \sim_{i.i.d.} (0,1), \quad \alpha_i \geq 0 \; \forall \; i$$

- Again, the h-step ahead forecast and forecast error are given by,

$$R_{t+h|t} = E\left[R_{t+h} \middle| \Omega_t\right] = \mu$$

$$R_{t+h} - E\left[R_{t+h} \middle| \Omega_t\right] = \varepsilon_{t+h}$$

- So we can see that the presence of ARCH effects do not impact the computation of the point forecasts.

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Because the innovations have an ARCH structure, their conditional variance is time varying. Therefore, we can compute the conditional forecast error variance as

$$E\left[\varepsilon_{t+h}^2 \middle| \Omega_t\right] = E\left[\sigma_{t+h}^2 \upsilon_{t+h}^2 \middle| \Omega_t\right] = \sigma_{t+h|t}^2$$

- Since the innovations are specified as an ARCH(1), we have that

$$\sigma_{t+h|t}^2 = \alpha_1^h \varepsilon_t^2 + \alpha_0 \sum_{i=0}^{h-1} \alpha_1^i$$

- Then our h-step ahead 95% forecast interval can then be computed as

$$\mu \pm 1.96\sigma_{t+h|t}$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Now let's suppose that R_t has a time varying conditional mean in the form of an AR(1)

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

$$\varepsilon_t = \sigma_t \upsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$
$$\upsilon_t \sim_{i.i.d.} (0,1), \quad \alpha_i \geq 0 \ \forall \ i$$

- We know from previous lectures that the 1-step ahead point forecast is given by

$$E\left[R_{t+1} \middle| \Omega_t\right] = \phi R_t$$

- And that the forecast error is given by

$$R_{t+1} - E\left[R_{t+1} \middle| \Omega_t\right] = \varepsilon_{t+1}$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- It follows that

$$E\left[\varepsilon_{t+1}^2 \middle| \Omega_t\right] = E\left[\sigma_{t+1}^2 \upsilon_{t+1}^2 \middle| \Omega_t\right] = \sigma_{t+1|t}^2$$

- Where

$$\sigma_{t+1|t}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2$$

- Then the 1-step ahead 95% forecast interval can then be computed as

$$\phi R_t \pm 1.96\sigma_{t+1|t}$$

- We also know that the 2-step ahead point forecast is given by

$$E\left[R_{t+2} \middle| \Omega_t\right] = \phi^2 R_t$$

- And the forecast error is given by

$$f_{t+2|t} = R_{t+2} - E\left[R_{t+2} \middle| \Omega_t\right] = \varepsilon_{t+2} + \phi\varepsilon_{t+1}$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- We can see clearly that

$$E\left[f_{t+2|t}\middle|\Omega_t\right] = 0$$

- The conditional forecast error variance is then computed as

$$E\left[f_{t+2|t}^2\middle|\Omega_t\right] = E\left[(\varepsilon_{t+2} + \phi\varepsilon_{t+1})^2\middle|\Omega_t\right]$$

- Multiplying out the square, we obtain

$$E\left[f_{t+2|t}^2\middle|\Omega_t\right] = E\left[\varepsilon_{t+2}^2 + 2\phi\varepsilon_{t+2}\varepsilon_{t+1} + \phi^2\varepsilon_{t+1}^2\middle|\Omega_t\right] = E\left[\varepsilon_{t+2}^2\middle|\Omega_t\right] + 2\phi E\left[\varepsilon_{t+2}\varepsilon_{t+1}\middle|\Omega_t\right] + \phi^2 E\left[\varepsilon_{t+1}^2\middle|\Omega_t\right]$$

- Let's focus our attention on the cross product term. We can see that since $\varepsilon_t = \sigma_t \upsilon_t$ we can write,

$$\phi E\left[\varepsilon_{t+2}\varepsilon_{t+1}\middle|\Omega_t\right] = \phi E\left[\sigma_{t+2}\upsilon_{t+2}\sigma_{t+1}\upsilon_{t+1}\middle|\Omega_t\right] = \phi E[\upsilon_{t+2}]E[\upsilon_{t+1}]E\left[\sigma_{t+2}\sigma_{t+1}\middle|\Omega_t\right] = 0$$

- Where the last two equalities are obtained from the fact that $\upsilon_t \sim_{i.i.d.} (0,1)$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Then the conditional 2-step ahead forecast error variance is given by

$$E\left[f_{t+2|t}^2\middle|\Omega_t\right] = E\left[\varepsilon_{t+2}^2\middle|\Omega_t\right] + \phi^2 E\left[\varepsilon_{t+1}^2\middle|\Omega_t\right] = \sigma_{t+2|t}^2 + \phi^2\sigma_{t+1|t}^2$$

- Where

$$\sigma_{t+1|t}^2 = \alpha_0 + \alpha_1\varepsilon_t^2$$

$$\sigma_{t+2|t}^2 = \alpha_0 + \alpha_1\alpha_0 + \alpha_1^2\varepsilon_t^2$$

- Therefore the 95% 2-step ahead forecast interval is computed as

$$\phi^2 R_t \pm 1.96\sqrt{\sigma_{t+2|t}^2 + \phi^2\sigma_{t+1|t}^2}$$

FORECASTING WITH A CONDITIONAL VOLATILITY MODEL

- Now recall that the h-step ahead forecast error variance for a standard AR(1) is given by

$$\sigma_h^2 = \sigma^2(1 + \phi_1^2 + \phi_1^4 + \cdots + \phi_1^{2h-2})$$

- From this we can see that the h-step ahead forecast error variance for an AR(1) with time varying conditional volatility will then be given by

$$\sigma_{t+h|t}^2 + \phi_1^2 \sigma_{t+h-1|t}^2 + \phi_1^4 \sigma_{t+h-2|t}^2 + \cdots + \phi_1^{2h-2} \sigma_{t+1|t}^2$$

- Since the innovations have been specified as an ARCH(1),

$$\sigma_{t+h|t}^2 = \alpha_1^h \varepsilon_t^2 + \alpha_0 \sum_{i=0}^{h-1} \alpha_1^i$$

APPROACH TO ESTIMATING & FORECASTING AN ARMA MODEL WITH ARCH EFFECTS

1. Estimate ARMA(p, q)

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

2. Extract residuals from Step 1 and test for ARCH effects. If ARCH effects are detected, specify an ARCH or GARCH model for the innovations,

$$\varepsilon_t = \sigma_t \upsilon_t$$

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

$$\upsilon_t \sim_{i.i.d.} N(0,1)$$

3. Check that the ARCH/GARCH specification is adequate by generating the squared standardized residuals from the ARCH/GARCH and testing whether any dependence remains (i.e. are there any ARCH effects that are unaccounted for)

APPROACH TO ESTIMATING & FORECASTING AN ARMA MODEL WITH ARCH EFFECTS

4. If the ARCH/GARCH model is adequate, proceed to generate the h-step ahead point forecasts from the ARMA(p, q) model.

5. To generate the h-step ahead forecast intervals, first derive the expression for the h-step ahead forecast error and associated conditional variance. This will be a function of the ARMA coefficients and the conditional variances from the ARCH/GARCH model, $\boldsymbol{\theta} = \{c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2_{t+1|t}, \dots \sigma^2_{t+h|t}\}$

6. Using your estimates of the ARMA coefficients and the forecasts of the conditional variances, compute the 95% forecast interval as

$$E\left[Y_{t+h} \middle| \Omega_t\right] \pm 1.96\sigma_h$$

THE STANDARD GARCH MODEL

- Recall that a standard GARCH(1,1) process is specified as:

$$\varepsilon_t = \sigma_t \upsilon_t$$

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

$$\upsilon_t \sim_{i.i.d.} N(0,1)$$

- Looking closer at the conditional variance equation we can see that this specification assumes that prior shocks have a symmetric effect on the conditional variance.

- When we consider the observed history of the effect of negative shocks on the volatility of financial returns versus positive shocks, this assumption of symmetry may not be appropriate!

- Financial theory also tells us that negative shocks should have a larger effect on the conditional variance via the leverage effect. In short, A negative shock raises the debt-equity ratio, thereby increasing leverage and consequently risk!

THRESHOLD GARCH (TARCH OR GJR-GARCH)

- Glosten, Jaganathan and Runkle (1993) proposed the following threshold GARCH specification:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \lambda_1 \varepsilon_{t-1}^2 I_{t-1}$$

- Where I_{t-1} is an indicator variable defined as:

$$I_{t-1} = \begin{cases} 1 & \varepsilon_{t-1} < 0 \\ 0 & \varepsilon_{t-1} \geq 0 \end{cases}$$

- The leverage effect in stock returns would lead us to expect that $\lambda_1 > 0$ so that a negative shock $\varepsilon_{t-1} < 0$ would have a larger effect on volatility than positive shock of the same magnitude.

EXPONENTIAL GARCH (E-GARCH)

- The E-GARCH specification by Nelson (1991) provides another way to capture asymmetry. This approach specifies the conditional variance equation in terms of logs:

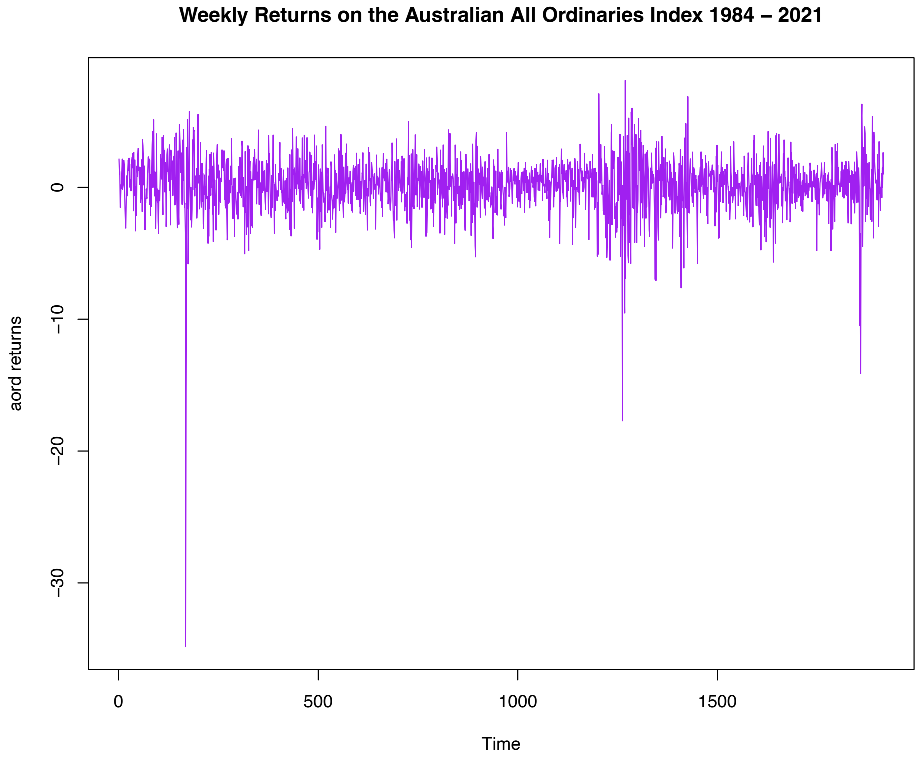
$$\ln(\sigma_t^2) = \alpha_0 + \lambda_1 \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \alpha_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \beta_1 \ln(\sigma_{t-1}^2)$$

- In this specification, the parameter α_1 captures potential asymmetry in the effect of ε_{t-1} on $\ln(\sigma_t)$. If the leverage effect exists, then $\alpha_1 < 0$ so that negative shocks are associated with a larger effect than positive shocks of the same magnitude.

- This specification also has the added advantage that the conditional variance is guaranteed to be positive without having to restrict the parameters to be positive.

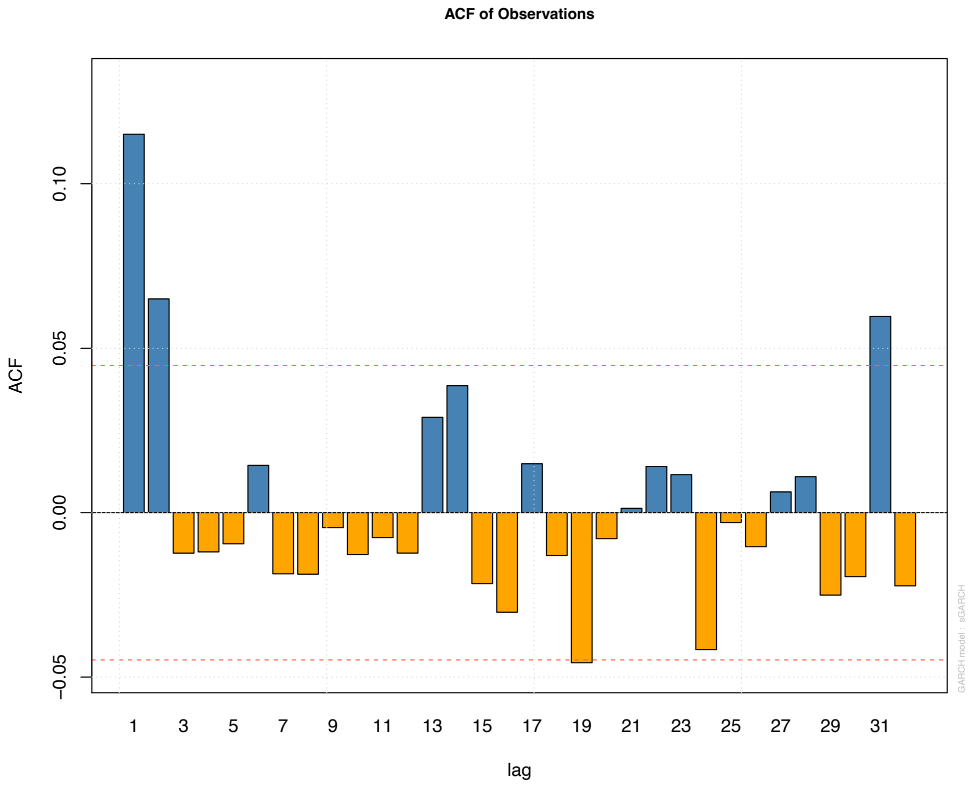
EXAMPLE: ALL ORDINARIES INDEX

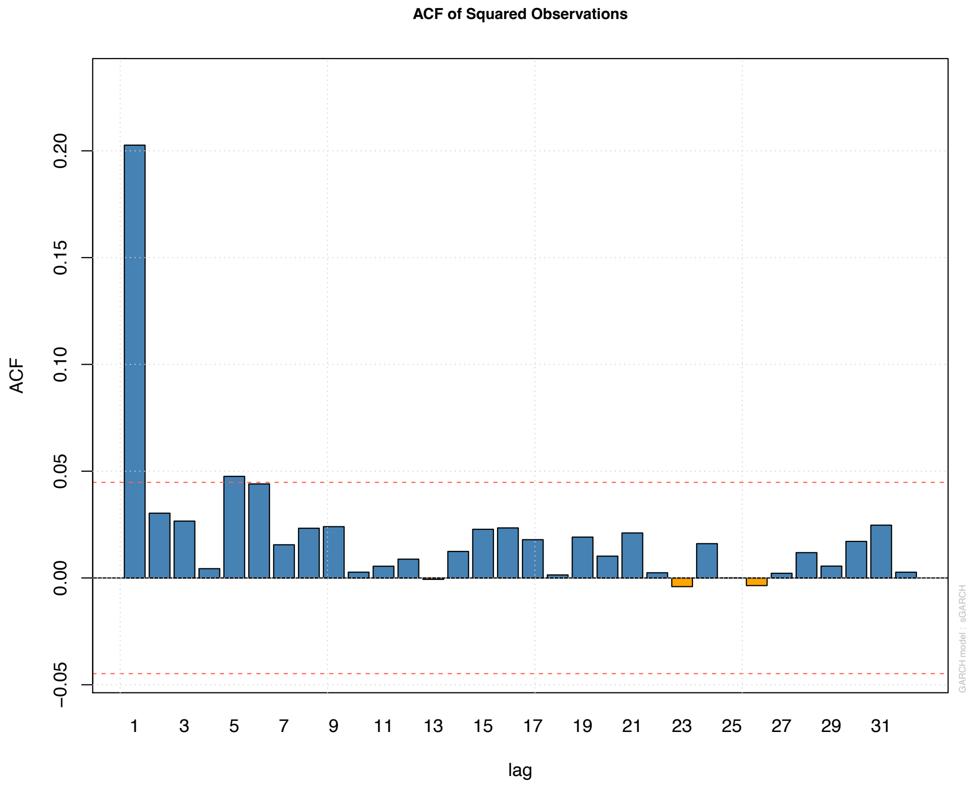
- Let's consider the weekly returns on the Australian All Ordinaries index:



EXAMPLE: ALL ORDINARIES INDEX

- Looking at the sample ACF of the returns and squared returns:





EXAMPLE: ALL ORDINARIES INDEX - GARCH

- Our sample ACF and PACF plots indicate to us that there shouldn't be too many higher order terms in the conditional mean equation or the conditional variance equation.

- Using the BIC, we choose an AR(1) – GARCH(1,1).

- We would use the robust standard errors if there still exists heteroscedasticity in our model. This will occur if our conditional variance model is inadequately specified. If our model has adequately accounted for the heteroscedasticity, we can use the regular standard errors.

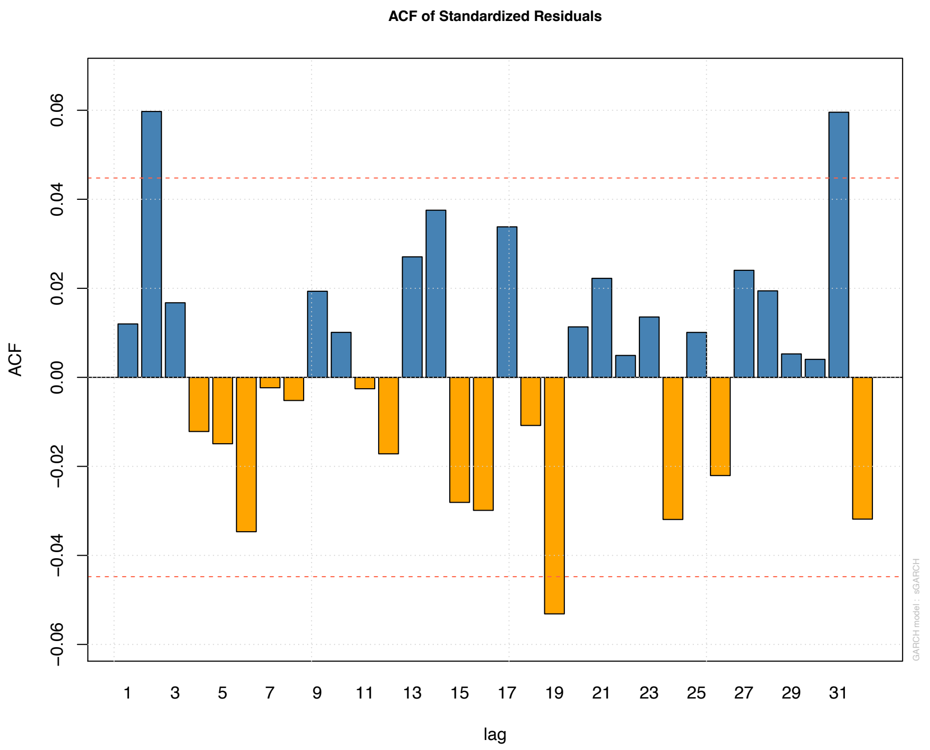
- Performing an ARCH-LM test on the standardized residuals, we see that no heteroscedasticity remains:

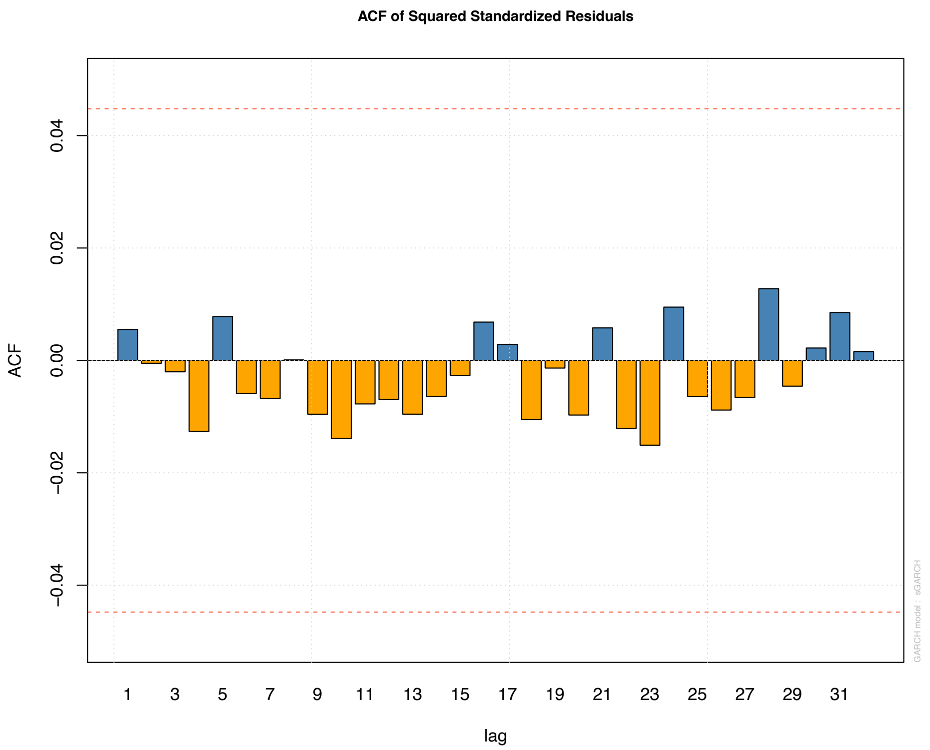
```
> ArchTest(std.res, lags=m)

	ARCH LM-test; Null hypothesis: no ARCH effects

data:  std.res
Chi-squared = 5.2451, df = 44, p-value = 1
```

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

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

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.036180    0.025619   1.4122 0.157883
omega   0.087547    0.030173   2.9015 0.003714
alpha1  0.127919    0.018963   6.7458 0.000000
beta1   0.871038    0.018413  47.3062 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.036180    0.029474   1.2275 0.219633
omega   0.087547    0.073474   1.1915 0.233444
alpha1  0.127919    0.039725   3.2201 0.001281
beta1   0.871038    0.027044  32.2087 0.000000
```

EXAMPLE: ALL ORDINARIES INDEX - GARCH

- Once we have estimated our GARCH model, we can have a look at the sample ACFs of our standardized residuals and squared standardized residuals. There doesn't appear to be much dependence left!





EXAMPLE: ALL ORDINARIES INDEX - GARCH

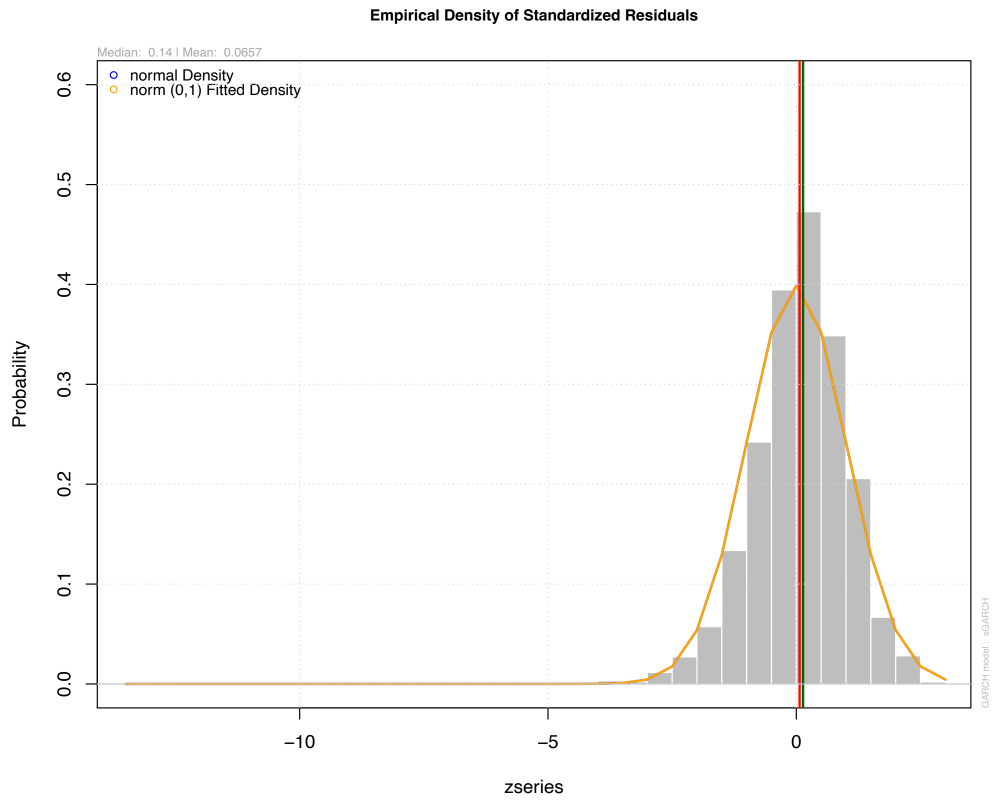
- This is confirmed by our Box-Pierce and Ljung Box tests:

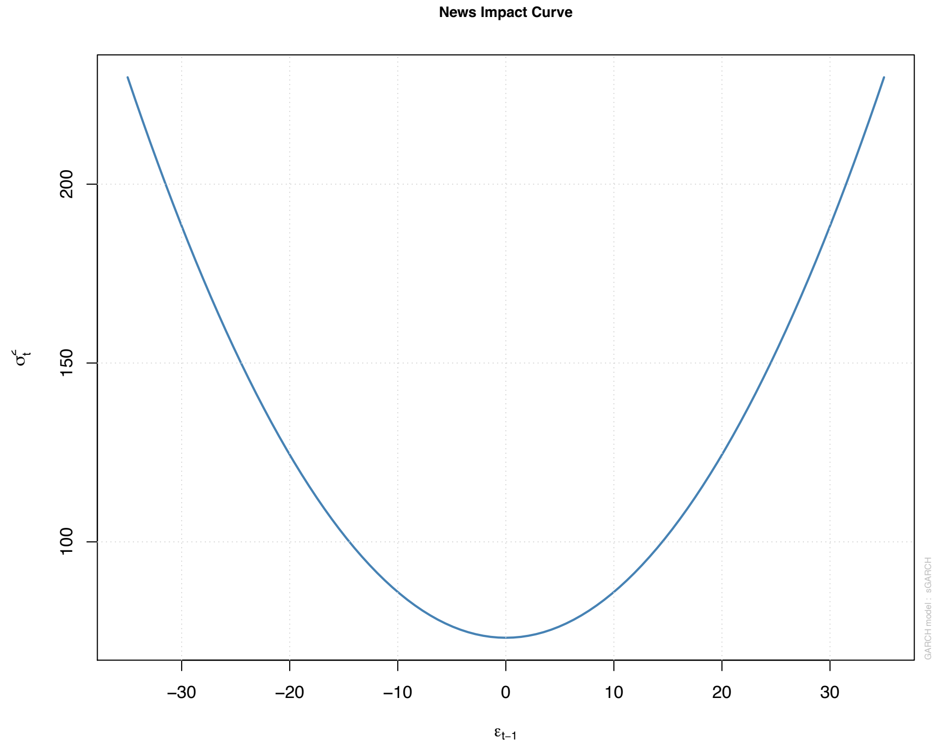
```
> AutocorTest(std.res, lag=m, type="Ljung-Box")

        Box-Ljung test

data:  std.res
X-squared = 47.986, df = 44, p-value = 0.3144

> AutocorTest(std.res, lag=m, type="Box-Pierce")

        Box-Pierce test

data:  std.res
X-squared = 47.451, df = 44, p-value = 0.3338
```

```
> AutocorTest(std.res^2, lag=m, type="Ljung-Box")

        Box-Ljung test

data:  std.res^2
X-squared = 5.2622, df = 44, p-value = 1

> AutocorTest(std.res^2, lag=m, type="Box-Pierce")

        Box-Pierce test

data:  std.res^2
X-squared = 5.1952, df = 44, p-value = 1
```

EXAMPLE: ALL ORDINARIES INDEX - GARCH

- The empirical density of our standardized residuals appear to be much more highly peaked than a standard normal. Also notice that the news impact curve is symmetric!





EXAMPLE: ALL ORDINARIES INDEX – GJR GARCH

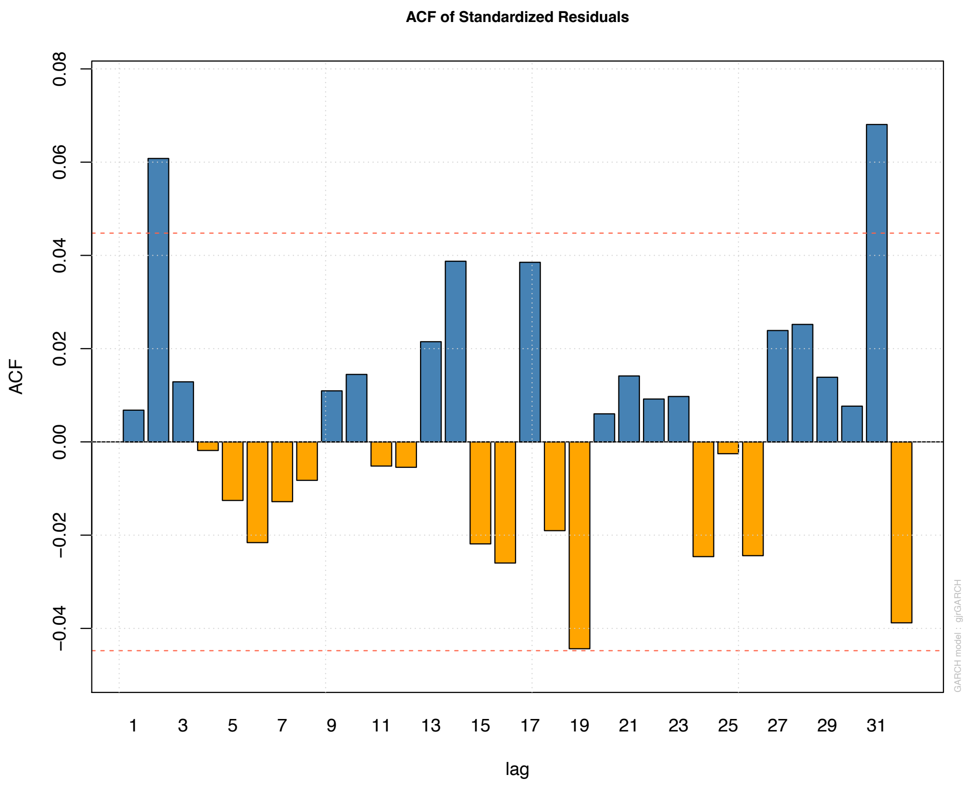
- Let's now estimate a GJR-GARCH(1,1) model to see whether a leverage effect exists.
- Since we can never be 100% sure that our model is correctly specified, we should always use the robust standard errors.
- Note that the coefficient multiplying the threshold variable is not statistically significant!

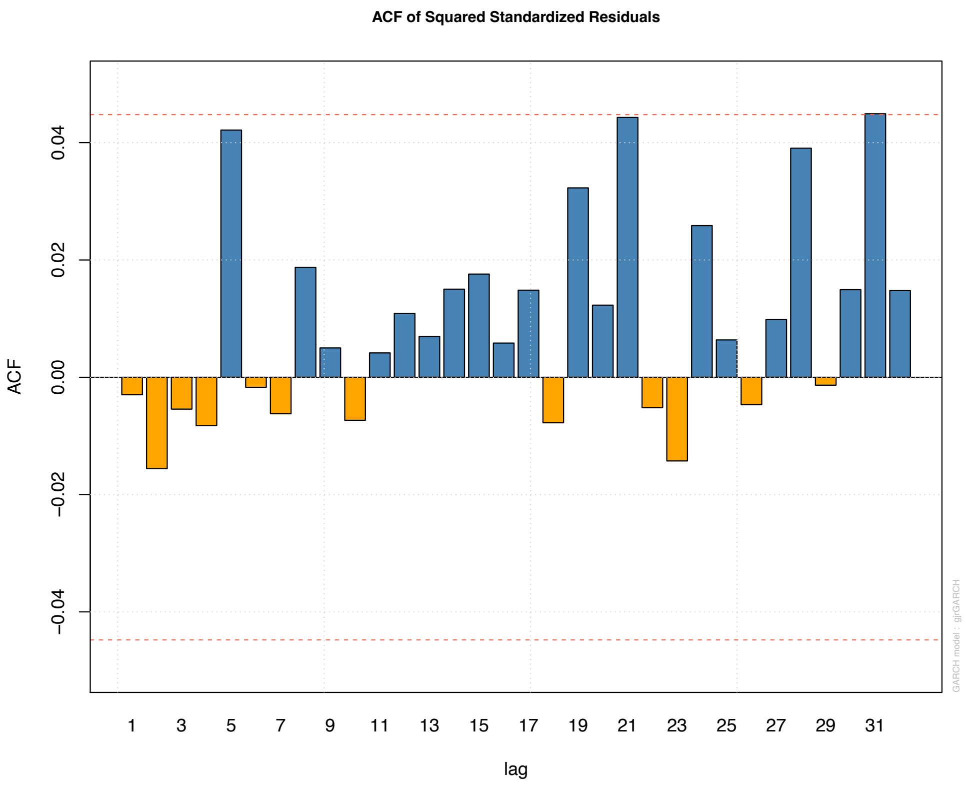
```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

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

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.050817    0.026382   1.9262 0.054074
omega   0.935918    0.156137   5.9942 0.000000
alpha1  0.035625    0.023346   1.5260 0.127010
beta1   0.542452    0.052241  10.3836 0.000000
gamma1  0.542471    0.069545   7.8003 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.050817    0.028579   1.7781 0.075385
omega   0.935918    0.538568   1.7378 0.082248
alpha1  0.035625    0.030023   1.1866 0.235385
beta1   0.542452    0.202823   2.6745 0.007484
gamma1  0.542471    0.346792   1.5643 0.117758
```

```
> ArchTest(std.gjr.res, lags=m)

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  std.gjr.res
Chi-squared = 27.058, df = 44, p-value = 0.9791
```

EXAMPLE: ALL ORDINARIES INDEX – GJR GARCH

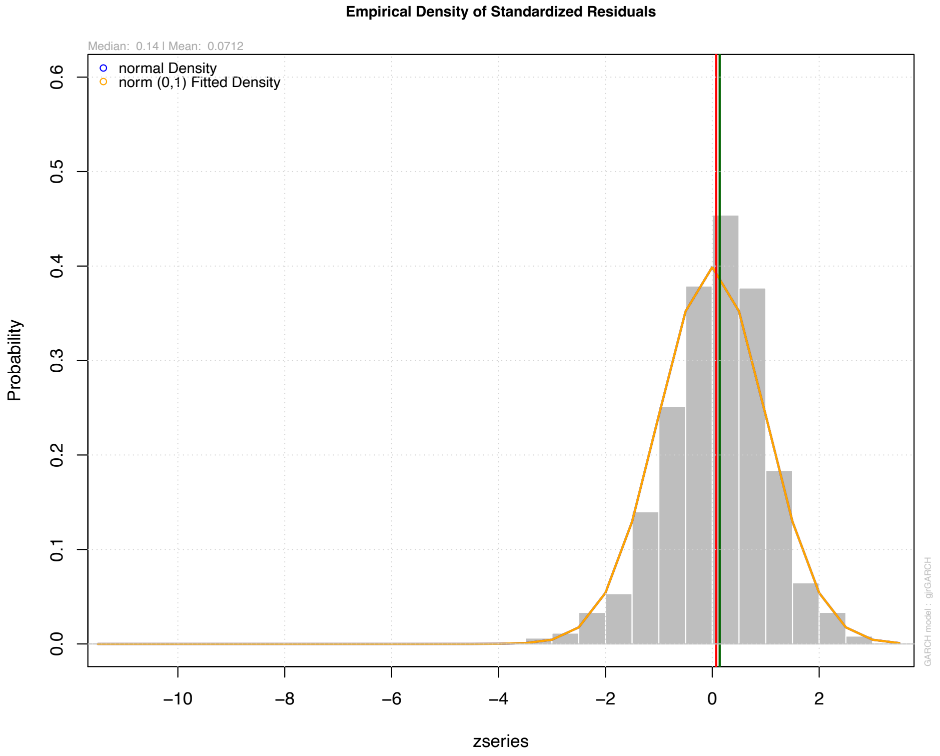
- The sample ACFs of our standardized residuals and squared standardized residuals again do not appear to have much structure.

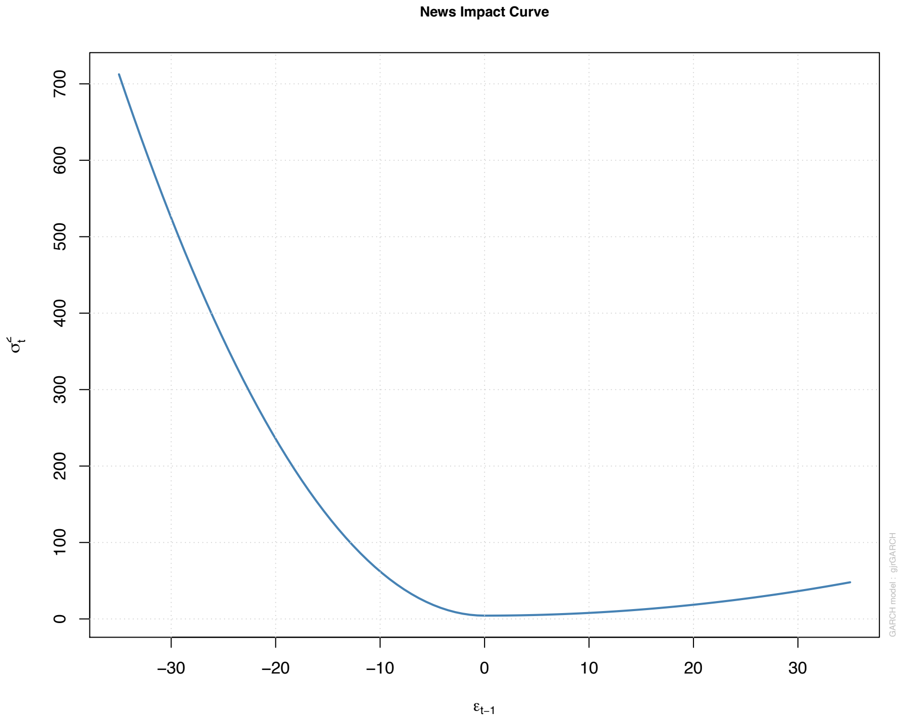




EXAMPLE: ALL ORDINARIES INDEX – GJR GARCH

- Again, the empirical density of the standardized residuals is highly peaked. However now notice the asymmetry in the news impact curve!





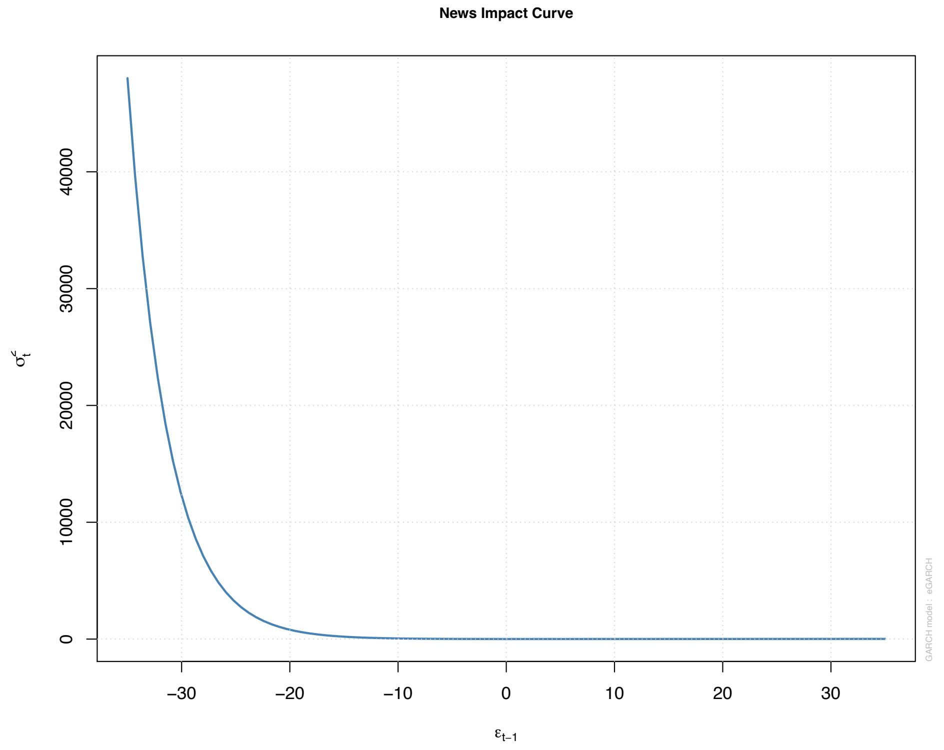
EXAMPLE: ALL ORDINARIES INDEX – E-GARCH

- We can also try an E-GARCH(1,1) specification.
- Again, we see that the coefficient multiplying the term that captures asymmetry is statistically significant!
- The residuals are very similar to the GJR-GARCH model.

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

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

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.061149    0.020746   2.9475 0.003204
omega   0.283736    0.041489   6.8388 0.000000
alpha1 -0.243365    0.026949  -9.0306 0.000000
beta1   0.809201    0.029226  27.6876 0.000000
gamma1  0.330932    0.034841   9.4984 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.061149    0.018675   3.2744 0.001059
omega   0.283736    0.141950   1.9988 0.045625
alpha1 -0.243365    0.102997  -2.3628 0.018136
beta1   0.809201    0.084020   9.6310 0.000000
gamma1  0.330932    0.100416   3.2956 0.000982
```

EXAMPLE: ALL ORDINARIES INDEX – E-GARCH

- The asymmetry in the news impact curve is much more extreme compared to the GJR-GARCH model!

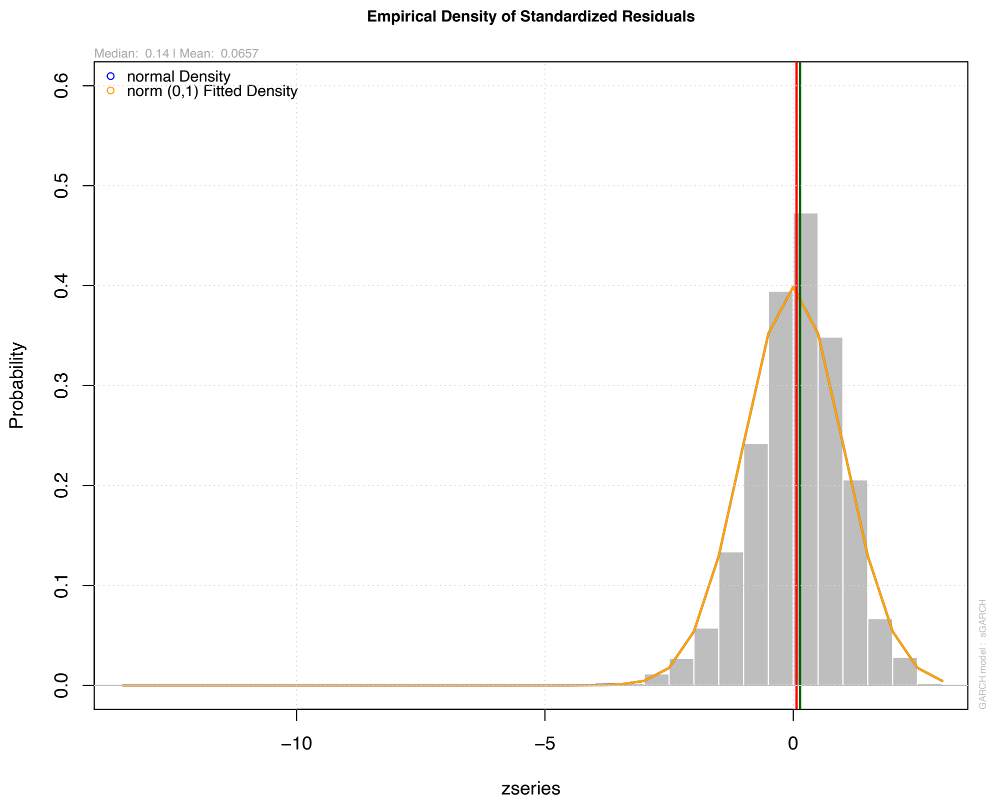


STUDENT – T DISTRIBUTIONAL ASSUMPTION

- Our GARCH specifications so far have assumed that the distribution of shocks is normal (i.e. that $\upsilon_t \sim_{i.i.d.} N(0,1)$).

- Although the GARCH model allows for some degree of leptokurtosis in the unconditional distribution of our time series of interest, in many practical instances, it is not able to capture the full extent of the leptokurtosis that is present in the data.

- We can see this clearly when the empirical distribution of our standardized residuals does not conform to the standard normal density function.



STUDENT – T DISTRIBUTIONAL ASSUMPTION

- One way to allow for a greater degree of leptokurtosis would be to assume that our shocks follow a heavier tailed distribution such as a Student-t!

- The Student-t distribution is given by:

$$f(\varepsilon_t|\varepsilon_{t-1};\boldsymbol{\theta}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\sigma_t(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{\varepsilon_t^2}{\sigma_t(\nu-2)}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

- Where $\Gamma()$ is the gamma function and σ_t^2 is our conditional variance.

- The Student-t distribution has a shape parameter ν which determines its excess kurtosis over the normal distribution:

$$Excess\ \ kurtosis = \frac{6}{\nu-4}$$

STUDENT - T DISTRIBUTIONAL ASSUMPTION

- Let’s now re-estimate our GJR-GARCH (1,1) model with Student-t innovations.
- Notice that we now have an estimate for the shape parameter ν!
- Also notice that the leverage effect parameter is much smaller in magnitude. Why do you think this is the case?

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

Conditional Variance Dynamics
creenshot -----------------------------
GARCH Model     : gjrGARCH(1,1)
Mean Model      : ARFIMA(1,0,0)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.058073    0.024367   2.3833 0.017160
omega   0.338609    0.092545   3.6589 0.000253
alpha1  0.032779    0.019656   1.6676 0.095390
beta1   0.801468    0.038831  20.6400 0.000000
gamma1  0.165621    0.050792   3.2608 0.001111
shape   8.036520    1.176146   6.8329 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
ar1     0.058073    0.025483   2.2789 0.022674
omega   0.338609    0.150345   2.2522 0.024309
alpha1  0.032779    0.020672   1.5856 0.112821
beta1   0.801468    0.058177  13.7765 0.000000
gamma1  0.165621    0.074193   2.2323 0.025596
shape   8.036520    1.961669   4.0968 0.000042
```

STUDENT - T DISTRIBUTIONAL ASSUMPTION

- Looking at the empirical density of the standardized residuals we can see that they conform to the shape of a standardized Student-t distribution.

- This improved fit is also reflected in the information criteria:

MODEL	AIC	BIC
GARCH(1,1)	4.201363	4.212967
GJR-GARCH(1,1)	4.157604	4.172108
GJR-GARCH(1,1) - STD	4.068946	4.086351