GARCH-models

Ass.Prof. Andriy Stavytskyy

References

- Bollerslev, T.P. (1986), Generalized Autoregresive Conditional Heteroscedasticity, Journal of Econometrics.
- Enders, W. (2003). Applied Econometric Time Series, second edition. Wiley
- Engle, R. F. (2001), GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics, Journal of Economic Perspectives.
- Engle, R.F. (1982), Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K, Econometrica.
- Greene W. H., Econometric Analysis, 7th ed., Prentice Hall, 2011.
- Hamilton, J. (1994). Time Series Analysis, Princeton University Press
- Hayashi, F. (2001). Econometrics, Princeton University Press
- Tsay, R. (2002). Analysis of Financial Time Series, Wiley

Agenda

- Non-linear models
- Autoregressive Conditionally Heteroscedastic Models
- Different modifications of GARCH-models
- Example: Application of GARCH Models in stock markets
- Example: Application of GARCH Models for investment decisions
- Example: Application of GARCH Model to forecast the Gold Futures Prices

Non-linear models

Traditional model

$y_t = \beta_0 + \beta_1 x_{1t} + ... + \beta_{k-1} x_{k-1t} + u_t$,

or more compactly

$y = X\beta + u$.

We also assumed $u_t \sim N(0,\sigma^2)$.

The linear structural (and time series) models cannot explain a number of important features common to much financial data

- leptokurtosis
- Volatility clustering or volatility pooling
- leverage effects

Traditional approach

- Until the early 1980s econometrics had focused almost solely on modeling the means of series - i.e., their actual values.

$$y_t = E_t(y_t \mid x) + \varepsilon_t , \quad \varepsilon_t \sim N(0,\sigma^2)$$

- For an AR(1) process:

$$E_{t-1}(y_t|x) = E_{t-1}(y_t) = \alpha + \beta\, y_{t-1}$$

- Note:

$$E(y_t) = \alpha/(1-\beta)$$

$$Var(y_t) = \sigma^2/(1-\beta^2)$$

Conditional VS. Unconditional moments

- The conditional moment is time varying, though the unconditional moment is not!

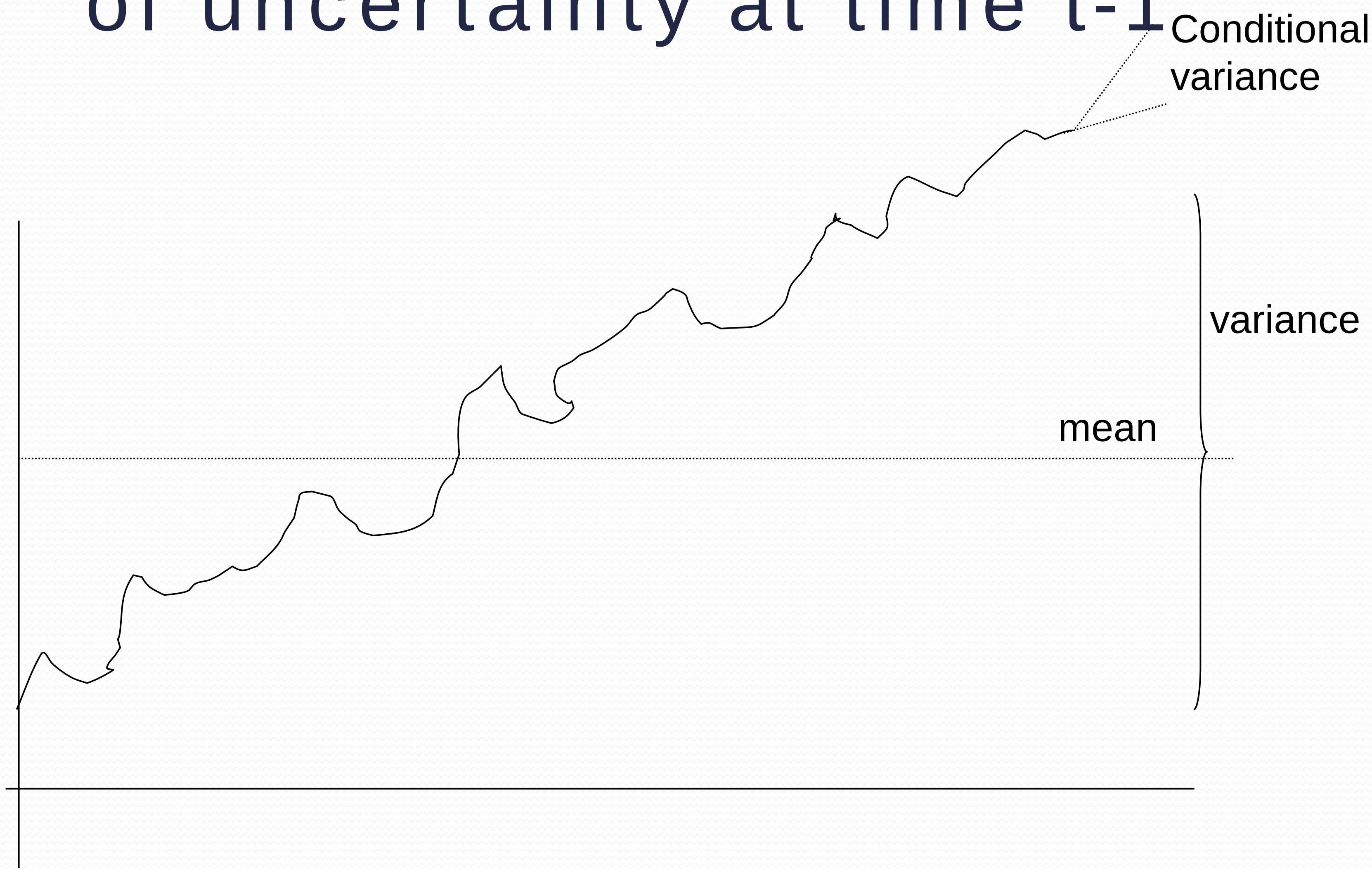
- Unconditional variance:

$$Var(y_t) = E[(y_t - E[y_t])^2] = \sigma^2/(1-\beta^2)$$

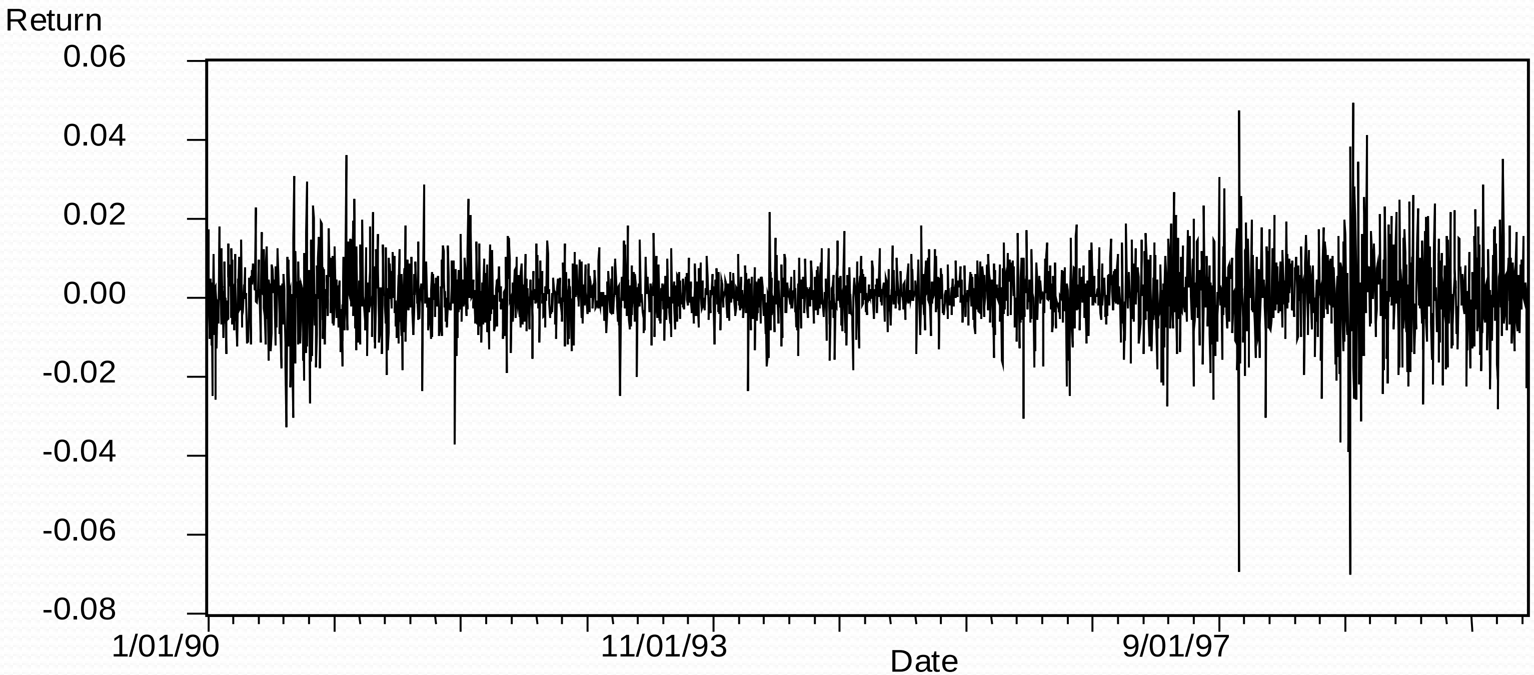
- Conditional variance:

$$Var_{t-1}(y_t) = E_{t-1}[(y_t - E_{t-1}[y_t])^2] = E_{t-1}[\varepsilon_t^2]$$

$Var_{t-1}(y_t)$ is the true measure of uncertainty at time t-1

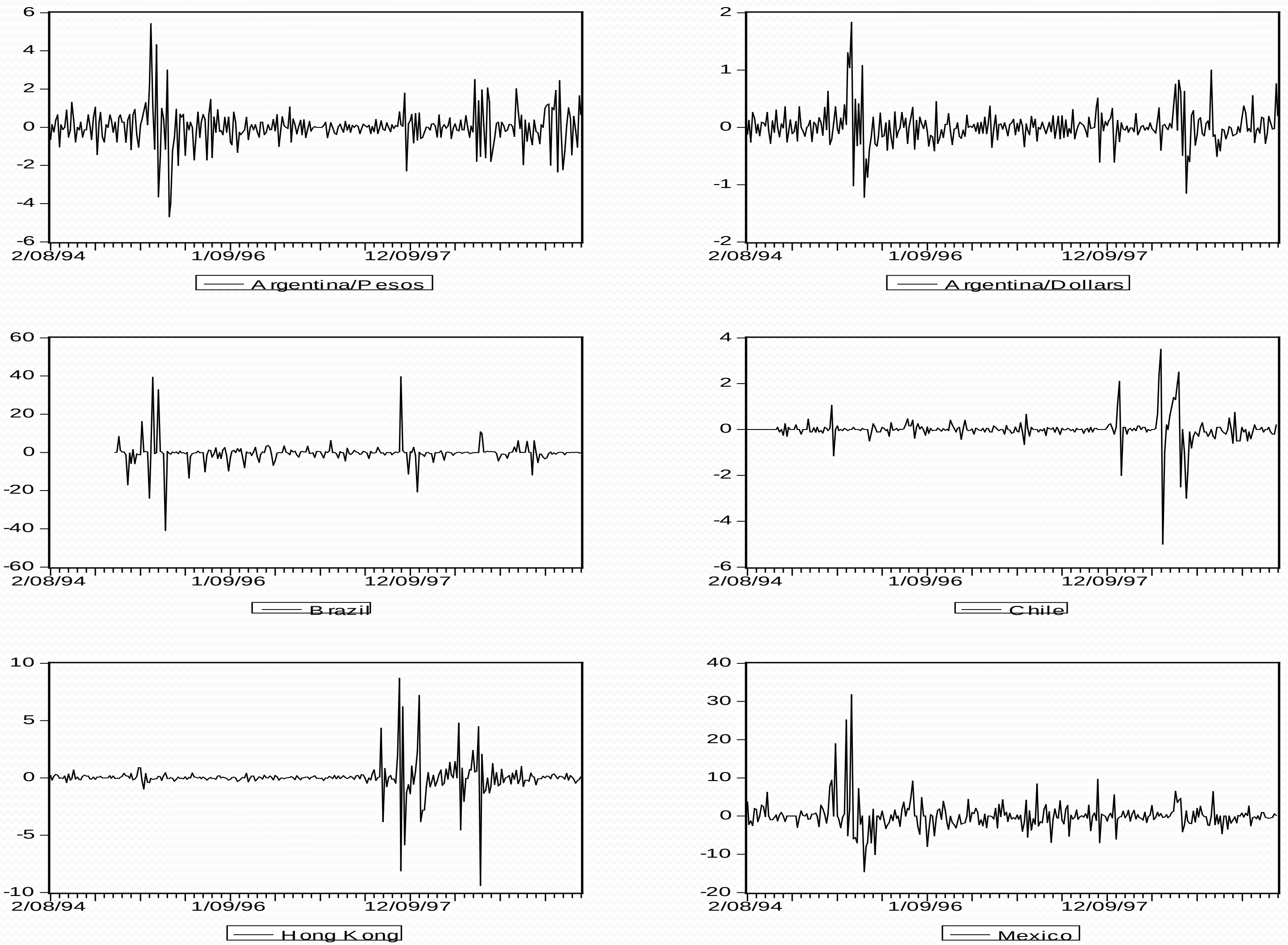


A Sample Financial Asset Returns Time Series



Daily S&P 500 Returns for January 1990 – December 1999

Changes in interest rates



Argentina, Brazil, Chile, Mexico and HK

Non-linear Models: A Definition

- Let's define a non-linear data generating process as one that can be written

$$y_t = f(u_t, u_{t-1}, u_{t-2}, \ldots)$$

where u_t is an iid error term and f is a non-linear function.

- A slightly more specific definition as

$$y_t = g(u_{t-1}, u_{t-2}, \ldots) + u_t\sigma^2(u_{t-1}, u_{t-2}, \ldots)$$

where g is a function of past error terms only and σ^2 is a variance term.

- Models with nonlinear $g(\bullet)$ are "non-linear in mean", while those with nonlinear $\sigma^2(\bullet)$ are "non-linear in variance".

Types of non-linear models

- Many apparently non-linear relationships can be made linear by a suitable transformation.
- On the other hand, it is likely that many relationships in finance are intrinsically non-linear.

- There are many types of non-linear models, e.g.
  - ARCH / GARCH
  - switching models
  - bilinear models

Testing for Non-linearity

- The “traditional” tools of time series analysis (acf’s, spectral analysis) may find no evidence that we could use a linear model, but the data may still not be independent.
- Portmanteau tests for non-linear dependence have been developed.
- The simplest is Ramsey’s RESET test, which took the form:

$$\hat{u}_t = \beta_0 + \beta_1 \hat{y}_t^2 + \beta_2 \hat{y}_t^3 + \ldots + \beta_{p-1} \hat{y}_t^p + v_t$$

- Many other non-linearity tests are available, e.g. the “BDS test” and the bispectrum test.
- One particular non-linear model that has proved very useful in finance is the ARCH model due to Engle (1982).

Heteroscedasticity

- An example of a structural model is

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + u_t$$

with $u_t \sim N\left(0, \sigma^2\right)$

- The assumption that the variance of the errors is constant is known as homoscedasticity, i.e.

$$\text{Var}(u_t) = \sigma_u^2.$$

What if the variance of the errors is not constant?

- heteroscedasticity
- would imply that standard error estimates could be wrong.

- Is the variance of the errors likely to be constant over time?
  - Not for financial data.

Autoregressive Conditionally Heteroscedastic Models

Autoregressive Conditionally Heteroscedastic (ARCH) Models –1

- So use a model which does not assume that the variance is constant.
- Recall the definition of the variance of u_t:

$$\sigma_t^2 = Var(u_t / u_{t-1}, u_{t-2}, \ldots) = E[(u_t - E(u_t))^2 / u_{t-1}, u_{t-2}, \ldots]$$

- We usually assume that $E(u_t) = 0$, so

$$\sigma_t^2 = Var(u_t / u_{t-1}, u_{t-2}, \ldots) = E[(u_t^2 / u_{t-1}, u_{t-2}, \ldots]$$

- What could the current value of the variance of the errors plausibly depend upon?
  - Previous squared error terms.
- This leads to the autoregressive conditionally heteroscedastic model for the variance of the errors:

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

- This is known as an ARCH(1) model.

Autoregressive Conditionally Heteroscedastic (ARCH) Models –2

- The full model would be $y_t = \beta_1 + \beta_2 x_{2t} + \ \dots \ + \beta_k x_{kt} + u_t$

$$u_t \sim N(0, \sigma_t{}^2)$$
$$\sigma_t{}^2 = \alpha_0 + \alpha_1 u^2{}_{t-1}$$

- We can easily extend this to the general case where the error variance depends on q lags of squared errors (ARCH(q) model):

$$\sigma_t{}^2 = \alpha_0 + \alpha_1 u^2{}_{t-1} + \alpha_2 u^2{}_{t-2} + \dots + + \alpha_q u^2{}_{t-q}$$

Ways of Writing ARCH Models - 1

Instead of calling the variance $\sigma_t\ ^2$, in the literature it is usually used h_t, so the model is

$$y_t = \beta_1 + \beta_2 x_{2t} + \ \dots \ + \beta_k x_{kt} + u_t$$

$$u_t \sim N(0, h)$$

$$h = \alpha_0 + \alpha_1 u^2{}_{t-1} + \alpha_2 u^2{}_{t-2} + {}_{\dots} + +\alpha_q u^2{}_{t-q}$$

Ways of Writing ARCH Models - 2

- We can also write $y_t = \beta_1 + \beta_2 x_{2t} + \ \dots \ + \beta_k x_{kt} + u_t$

,

$$u_t = v_t \sigma_t$$
$$v_t \sim N(0, 1)$$

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 u^2{}_{t-1} + \alpha_2 u^2{}_{t-2} + {}_{\dots} + + \alpha_q u^2{}_{t-q}}$$

- The two are different ways of expressing exactly the same model. The first form is easier to understand while the second form is required for simulating from an ARCH model, for example.

Testing for "ARCH Effects"

- First, run any postulated linear regression of the form given in the equation above, e.g.

$$y_t = \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t$$

saving the residuals, u_t .

- Then square the residuals, and regress them on *q* own lags to test for ARCH of order *q*, i.e. run the regression

$$u_t^2 = \gamma_0 + \gamma_1 u_{t-1}^2 + \gamma_2 u_{t-2}^2 + \ldots + \gamma_q u_{t-q}^2 + v_t$$

where v_t is iid.
Obtain R^2 from this regression

- The test statistic is defined as $T \bullet R^2$ (the number of observations multiplied by the coefficient of multiple correlation) from the last regression, and is distributed as a $\chi^2(q)$.

Testing for "ARCH Effects" – 2

- The null and alternative hypotheses are

$$H0: \gamma_1 = 0 \text{ and } \gamma_2 = 0 \text{ and } \gamma_3 = 0 \text{ and ... and } \gamma_q = 0$$

$$H1: \gamma_1 \neq 0 \text{ or } \gamma_2 \neq 0 \text{ or } \gamma_3 \neq 0 \text{ or ... or } \gamma_q \neq 0.$$

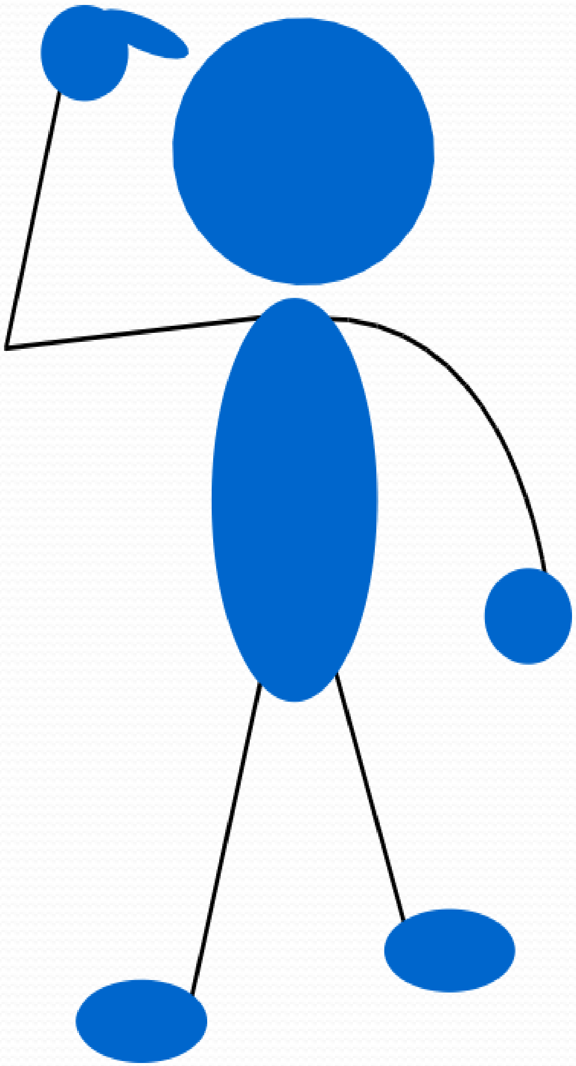
- If the value of the test statistic is greater than the critical value from the χ^2 distribution, then reject the null hypothesis.

Problems with ARCH(q) Models

- How do we decide on **q**?
- The required value of **q** might be very large
- Non-negativity constraints might be violated.
  - *When we estimate an ARCH model, we require* $\alpha_i > 0 \ \forall \ i=1,2,\ldots,q$ *(since variance cannot be negative)*

- A natural extension of an ARCH(q) model which gets around some of these problems is a GARCH model.

Generalised ARCH (GARCH) Models



- Due to Bollerslev (1986).
- Allow the conditional variance to be dependent upon previous own lags
- The variance equation is now

$$u^2_t = \alpha_0 + \alpha_1 u^2_{t-1} + \beta \sigma^2_{t-1}$$

- This is a GARCH(1,1) model, which is like an ARMA(1,1) model for the variance equation.

Variance – 1

- We could also write

$$\sigma_{t-1}^2 = \alpha_0 + a_1 u_{t-2}^2 + \beta \sigma_{t-2}^2$$
$$\sigma_{t-2}^2 = \alpha_0 + a_1 u_{t-3}^2 + \beta \sigma_{t-3}^2$$

- Substituting :

$$\sigma_t^2 = \alpha_0 + a_1 u_{t-1}^2 + \beta \left(\alpha_0 + a_1 u_{t-2}^2 + \beta \sigma_{t-2}^2 \right) =$$
$$= \alpha_0 + a_1 u_{t-1}^2 + \alpha_0 \beta + a_1 \beta u_{t-2}^2 + \beta \sigma_{t-2}^2$$

Variance – 2

- In general

$$\sigma_t^2 = \alpha_0 + a_1 u_{t-1}^2 + \alpha_0 \beta + a_1 \beta u_{t-2}^2 + \beta^2 \left(\alpha_0 + a_1 u_{t-3}^2 + \beta \sigma_{t-3}^2 \right)$$

$$\sigma_t^2 = \alpha_0 + a_1 u_{t-1}^2 + \alpha_0 \beta + a_1 \beta u_{t-2}^2 + \alpha_0 \beta^2 + \alpha_1 \beta^2 u_{t-3}^2 + \beta^3 \sigma_{t-3}^2$$

$$\sigma_t^2 = \alpha_0 \left(1 + \beta + \beta^2 \right) + a_1 u_{t-1}^2 \left(1 + \beta L + \beta^2 L^2 \right) + \beta^3 \sigma_{t-3}^2$$

$$\ldots$$

$$\sigma_t^2 = \alpha_0 \left(1 + \beta + \beta^2 + \ldots \right) + a_1 u_{t-1}^2 \left(1 + \beta L + \beta^2 L^2 + \ldots \right) + \beta^\infty \sigma_0^2$$

- So the GARCH(1,1) model can be written as an infinite order ARCH model.

GARCH(p,q)

- We can again extend the GARCH(1,1) model to a GARCH(p,q):

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \ldots + \beta_p \sigma_{t-p}^2$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

Generalised ARCH (GARCH)

- In general a GARCH(1,1) model will be sufficient to capture the volatility clustering in the data.

- GARCH is better than ARCH:
  - more parsimonious - avoids overfitting (the previous page show that a GRACH(1,1) is equivalent to an ARCH(∞), but with only 3 parameters)
  - less likely to breech non-negativity constraints

The Unconditional Variance under the GARCH Specification

The unconditional variance of u_t is given by $\mathrm{Var}(u_t) = \dfrac{\alpha_0}{1-(\alpha_1+\beta)}$

$$\alpha_1 + \beta < 1$$

when

$$\alpha_1 + \beta \geq 1$$

is termed “non-stationarity” in variance

$$\alpha_1 + \beta = 1$$

is termed intergrated GARCH

For non-stationarity in variance, the conditional variance forecasts will not converge on their unconditional value as the horizon increases.

Estimation of ARCH / GARCH Models

- Since the model is no longer of the usual linear form, we cannot use OLS.
- We use another technique known as maximum likelihood.
- The method works by finding the most likely values of the parameters given the actual data.
- More specifically, we form a log-likelihood function and maximise it.

Estimation of ARCH / GARCH Models: ML

- The steps involved in actually estimating an ARCH or GARCH model are as follows

- Specify the appropriate equations for the mean and the variance - e.g. an AR(1)- GARCH(1,1) model:

$$y_t = \mu + \phi y_{t-1} + u_t \quad , u_t \sim \mathrm{N}(0,\sigma_t^2)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2$$

- Specify the log-likelihood function to maximise:

$$L = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}(y_t - \mu - \phi y_{t-1})^2 / \sigma_t^2$$

- The computer will maximise the function and give parameter values and their standard errors

Estimation of GARCH Models Using Maximum Likelihood

- Now we have $y_t = \mu + \phi y_{t-1} + u_t$, $u_t \sim N(0, \sigma^2{}_t)$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}{}^2$$

- Unfortunately, the LLF for a model with time-varying variances ***cannot be maximised analytically***, except in the simplest cases. So, a numerical procedure is used to maximise the ***log-likelihood function.*** A potential problem: local optima or multimodalities in the likelihood surface.

$$L = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(\sigma_t{}^2) - \frac{1}{2}\sum_{t=1}^{T}(y_t - \mu - \phi y_{t-1})^2 / \sigma_t{}^2$$

- The way we do the optimisation is:
  - Set up LLF.
  - Use regression to get initial guesses for the mean parameters.
  - Choose some initial guesses for the conditional variance parameters.
  - Specify a convergence criterion - either by criterion or by value.

Non-Normality and Maximum Likelihood

- Recall that the conditional normality assumption for u_t is essential.
- We can test for normality using the following representation

$$u_t = v_t \sigma_t$$

$$v_t \sim N(0,1)$$

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 \sigma_{t-1}^2}$$

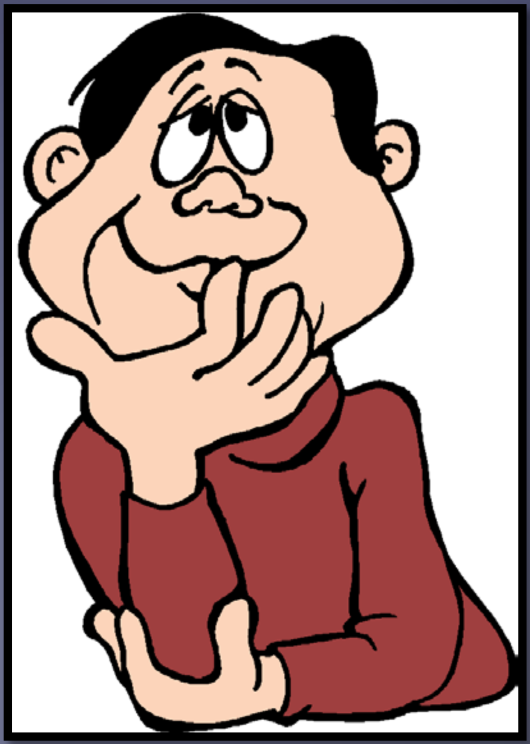
$$v_t = \frac{u_t}{\sigma_t}$$

- The sample counterpart is $\hat{v}_t = \dfrac{\hat{u}_t}{\hat{\sigma}_t}$
- Are the $\hat{v}_t$ normal? Typically $\hat{v}_t$ are still leptokurtic, although less so than the $\hat{u}_t$ Is this a problem?
- Not really, as we can use the ML with a robust variance/covariance estimator. ML with robust standard errors is called Quasi- Maximum Likelihood or QML.

Problems with GARCH(p,q) Models

- Non-negativity constraints may still be violated
- GARCH models cannot account for leverage effects

Different modifications of GARCH-models



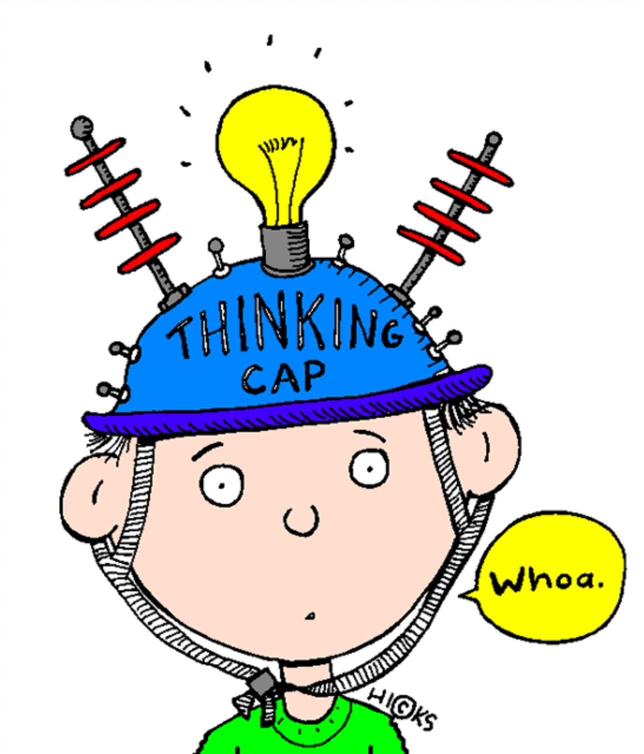
EXOGENOUS VARIABLES IN A GARCH MODEL

- Include predetermined variables into the variance equation
- Easy to estimate and forecast one step
- Multi-step forecasting is difficult
- Timing may not be right

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta h_{t-1} + \gamma z_{t-1}$$

Extensions to the Basic GARCH Model

- Since the GARCH model was developed, a huge number of extensions and variants have been proposed.
- Three of the most important examples are:
  - EGARCH,
  - GJR,
  - GARCH-M.



The EGARCH Model

- Suggested by Nelson (1991). The variance equation is given by

$$\log(\sigma_t^{\,2}) = \omega + \beta \log(\sigma_{t-1}^{\,2}) + \gamma \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^{\,2}}} + \alpha\left[\frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^{\,2}}} - \sqrt{\frac{2}{\pi}}\right]$$

- Advantages of the model
  - Since we model the $\log(\sigma_t^2)$, then even if the parameters are negative, σ_t^2 will be positive.
  - We can account for the leverage effect: if the relationship between volatility and returns is negative, γ, will be negative.

The GJR Model

- Due to Glosten, Jaganathan and Runkle

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma u_{t-1}^2 I_{t-1}$$

where $I_{t-1} = 1$ if $u_{t-1} < 0$ and $I_{t-1} = 0$ otherwise.

- For a leverage effect, we would see $\gamma > 0$.
- We require $\alpha_1 + \gamma \geq 0$ and $\alpha_1 \geq 0$ for non-negativity.

An Example of the use of a GJR Model

- Using monthly S&P 500 returns, December 1979- June 1998
- Estimating a GJR model, we obtain the following results.

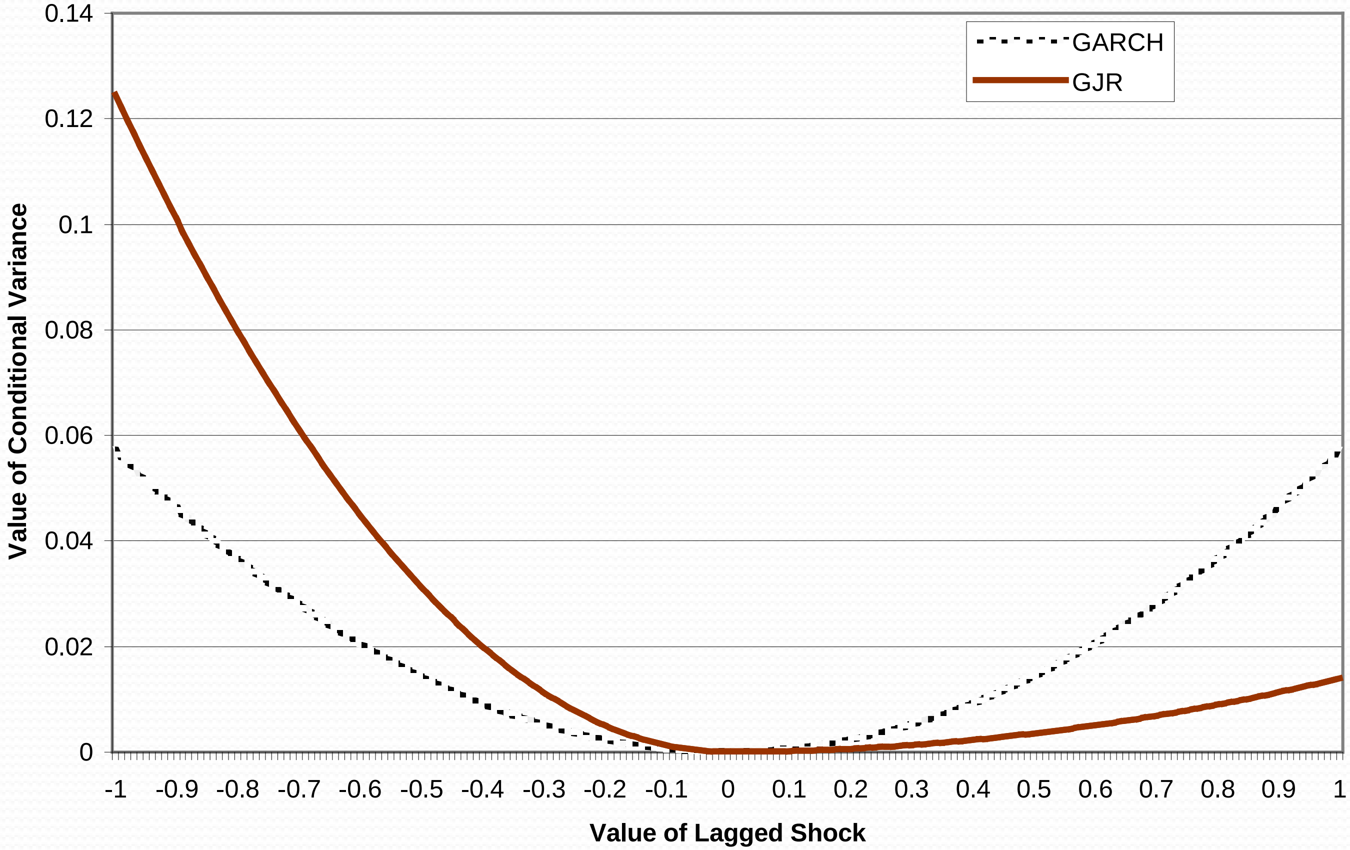
$$y_t = 0.172$$
$$(3.198)$$

$$\sigma_t^{\ 2} = 1.243 + 0.015 u_{t-1}^2 + 0.498 \sigma_{t-1}^{\ \ 2} + 0.604 u_{t-1}^2 I_{t-1}$$
$$(16.372) \quad (0.437) \quad (14.999) \quad (5.772)$$

News Impact Curves

The news impact curve plots the next period volatility (h_t) that would arise from various positive and negative values of u_{t-1}, given an estimated model.

News Impact Curves for GARCH and GJR Model Estimates



Integrated GARCH (IGARCH)

- This model was originally described in Engle and Bollerslev (1986).
- Model restricts the parameters of the GARCH model to sum to one and drop the constant term:

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

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

The Power ARCH (PARCH)

- Taylor (1986) and Schwert (1989) introduced the standard deviation GARCH model, where the standard deviation is modeled rather than the variance:

$$\sigma_t^{\delta} = \alpha_0 + \sum_{j=1}^{q} \beta_j \sigma_{t-1}^{\delta} + \sum_{i=1}^{p} \alpha_i \left(\left| u_{t-1} \right| - \gamma_i u_{t-1} \right)^{\delta}$$

$$\delta > 0, \left| \gamma_i \right| \leq 1, i = 1, \ldots, r, \gamma_i = 0 \; for \, i > r, r \leq p.$$

Component GARCH (CGARCH) – 1

- The model allows mean reversion to a varying level:

$$\sigma_t^2 - m_t = \alpha\left(u_{t-1}^2 - m_{t-1}\right) + \beta\left(\sigma_{t-1}^2 - m_{t-1}\right)$$

$$m_t = \alpha_0 + \rho\left(m_{t-1} - \alpha_0\right) + \phi\left(u_{t-1}^2 - \sigma_{t-1}^2\right)$$

$$\sigma_t^2 - m_t$$

- The first equation describes the transitory component, which converges to zero with powers of (α+β).
- The second equation describes the long run component m_t, which converges to α_0 with powers of ρ. ρ is typically between 0.99 and 1 so that approaches very slowly.

Component GARCH (CGARCH) – 2

- Combination of equations:

$$\sigma_t^2 = (1-a-\beta)(1-\rho)\alpha_0 + (\alpha+\phi)u_{t-1}^2$$
$$-(\alpha\rho + (\alpha+\beta)\phi)u_{t-2}^2 + (\beta-\phi)\sigma_{t-1}^2$$
$$-(\beta\rho - (\alpha+\beta)\phi)\sigma_{t-2}^2$$

- The component model is a (nonlinear) restricted GARCH(2, 2) model.

GARCH-in Mean

- We expect a risk to be compensated by a higher return. So why not let the return of a security be partly determined by its risk?

- Engle, Lilien and Robins (1987) suggested the ARCH-M specification. A GARCH-M model would be

$$y_t = \mu + \delta\sigma_{t-1} + u_t \quad , u_t \sim \mathrm{N}(0,\sigma_t^2)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2$$

where δ can be interpreted as a sort of risk premium.

- It is possible to combine all or some of these models together to get more complex “hybrid” models - e.g. an ARMA-EGARCH(1,1)-M model.

Threshold ARCH (TARCH)

- Rabemananjara, R. and J.M. Zakoian (1993), "***Threshold ARCH Models and Asymmetries in Volatilities***" Journal of Applied Econometrics.

TARCH-model – 1

- Large events to have an effect but no effect from small events

$$\sigma^2{}_t = \alpha_0 + \sum_{i=1}^{q} (\alpha^+{}_i I(\varepsilon_{t-i} > \kappa) + \alpha^-{}_i I(\varepsilon_{t-i} < \kappa)) u^2{}_{t-i}) +$$

$$+ \sum_{j=1}^{p} \beta_j \sigma^2{}_{t-j}$$

TARCH-model – 2

- There are two variances:

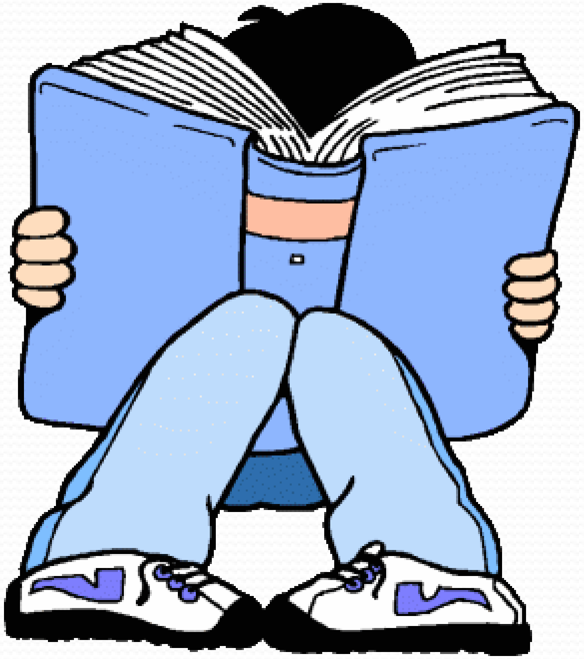
$$\sigma_{t-i}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha^{+}{}_{i} u_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma^2{}_{t-j}, if\,(u_{t-i} > \kappa)$$

$$\sigma_{t-i}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha^{-}{}_{i} \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma^2{}_{t-j}, if\,(u_{t-i} < \kappa)$$

- Many other versions are possible by adding minor asymmetries or non-linearities in a variety of ways.

Switching ARCH (SWARCH)

- Hamilton, J. D. and R. Susmel (1994), "***Autoregressive Conditional Heteroskedasticity and Changes in Regime***," Journal of Econometrics.



SWARCH – 1

- Simplest case: 2-state process.
- Assume the existence of an unobserved variable, s_t, that can take two values: one or two (or zero or one).
- Postulate a Markov transition matrix, P, for the evolution of the unobserved variable:

$$p(s_t = 1 \mid s_{t-1} = 1) = p$$
$$p(s_t = 2 \mid s_{t-1} = 1) = (1-p)$$
$$p(s_t = 1 \mid s_{t-1} = 2) = q$$
$$p(s_t = 2 \mid s_{t-1} = 2) = (1-q)$$

SWARCH – 2

- Reformulate ARCH(q) equation to make the conditional variance dependent on s_t –i.e., the state of the economy.

$$\sigma^2{}_t = \alpha_{0\,s_t, s_{t-1}} + \sum_{i=1}^{q} \alpha_{i, s_t, s_{t-1}} u^2{}_{t-i}$$

- A parsimonious formulation:

$$\sigma^2{}_t / \gamma_{s_t} = \alpha_0 + \sum_{i=1}^{q} \alpha_i u^2{}_{t-i} / \gamma_{s_{t-i}}$$

SWARCH – 3

- For a SWARCH(1) with 2 states (1 and 2) we have 4 possible σ_t^2:

$$\sigma^2_t = \alpha_0\gamma_1 + \alpha_1\varepsilon^2_{t-1}\gamma_1 / \gamma_1, s_t = 1, s_{t-1} = 1$$

$$\sigma^2_t = \alpha_0\gamma_1 + \alpha_1\varepsilon^2_{t-1}\gamma_1 / \gamma_2, s_t = 1, s_{t-1} = 2$$

$$\sigma^2_t = \alpha_0\gamma_2 + \alpha_1\varepsilon^2_{t-1}\gamma_2 / \gamma_1, s_t = 2, s_{t-1} = 1$$

$$\sigma^2_t = \alpha_0\gamma_2 + \alpha_1\varepsilon^2_{t-1}\gamma_2 / \gamma_2, s_t = 2, s_{t-1} = 2$$

SWARCH – 4

- Estimation of the model will estimate the volatility parameters and the transition probabilities. As a byproduct of the estimation, we will also have an estimate for the latent variable –i.e., the “state.”

Use of GARCH-models

What Use Are GARCH-type Models?

- GARCH can model the volatility clustering effect since the conditional variance is autoregressive. Such models can be used to forecast volatility.

- We could show that

$$\mathrm{Var}\,(y_t \mid y_{t-1}, y_{t-2}, \ldots) = \mathrm{Var}\,(u_t \mid u_{t-1}, u_{t-2}, \ldots)$$

- So modelling σ_t^2 will give us models and forecasts for y_t as well.

- Variance forecasts are additive over time.

Forecasting Variances using GARCH Models

- Producing conditional variance forecasts from GARCH models uses a very similar approach to producing forecasts from ARMA models.
- It is again an exercise in iterating with the conditional expectations operator.
- What is needed is to generate forecasts of

$$\sigma^2_{T+1} \mid \Omega T,\ \sigma^2_{T+2} \mid \Omega T,\ \ldots,\ \sigma^2_{T+s} \mid \Omega T$$

  where ΩT denotes all information available up to and including observation T.
- Adding one to each of the time subscripts of the above conditional variance equation, and then two, and then three would yield the following equations

$$\sigma^2_{T+1\ =}\ \alpha_0 + \alpha_1 +$$

Forecasting Variances

- Any s-step ahead forecast ($s \geq 2$) would be produced by

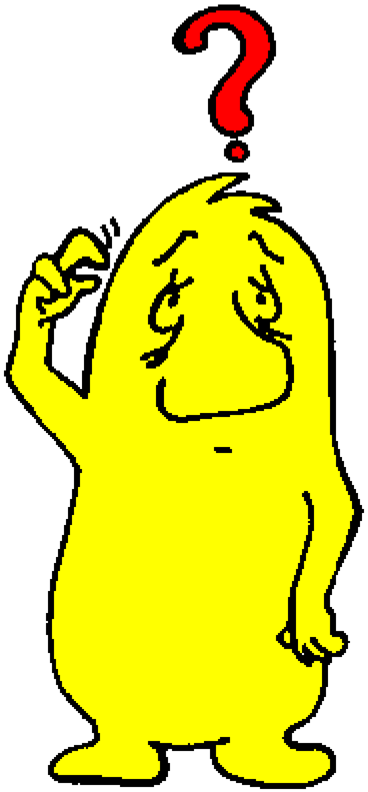
$$h_{s,T}^{f} = \alpha_0 \sum_{i=1}^{s-1} (\alpha_1 + \beta)^{i-1} + (\alpha_1 + \beta)^{s-1} h_{1,T}^{f}$$

What Use Are Volatility Forecasts?

- Option pricing
- Dynamic hedge ratios (the size of the futures position to the size of the underlying exposure, i.e. the number of futures contracts to buy or sell per unit of the spot good.
- Stock pricing

Questions & Answers

- Lots of ARCH models. Which one to use?
- Choice of p and q. How many lags to use?
- It turns out that the GARCH(1,1) is a great starting model.
- Add a leverage effect for financial series and it's even better.



Testing Hypotheses about Non-linear Models

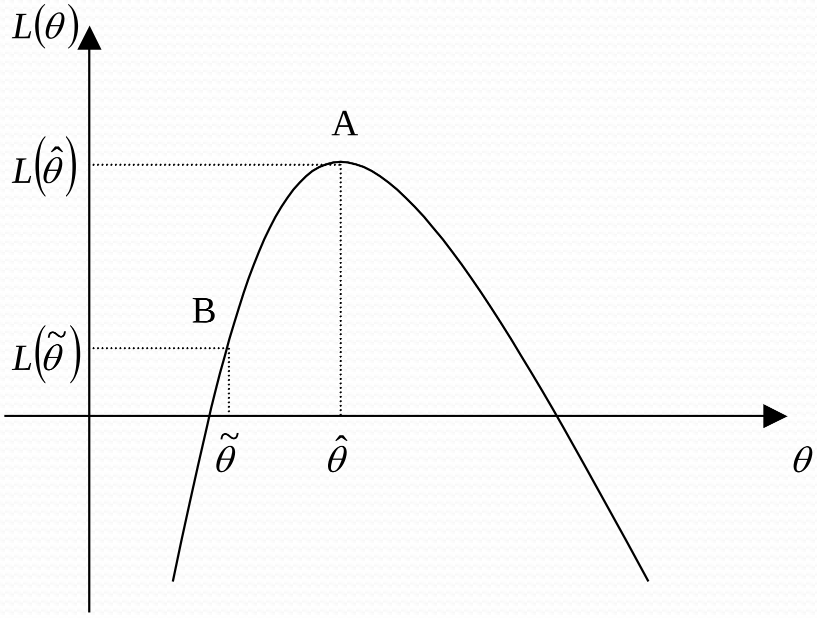
- Usual t- and F-tests are still valid in non-linear models, but they are not flexible enough.

- There are three hypothesis testing procedures based on maximum likelihood principles:
  - Wald,
  - Likelihood Ratio,
  - Lagrange Multiplier.

- Consider a single parameter, θ to be estimated,
- Denote the MLE as $\hat{\theta}$ and a restricted estimate as $\tilde{\theta}$

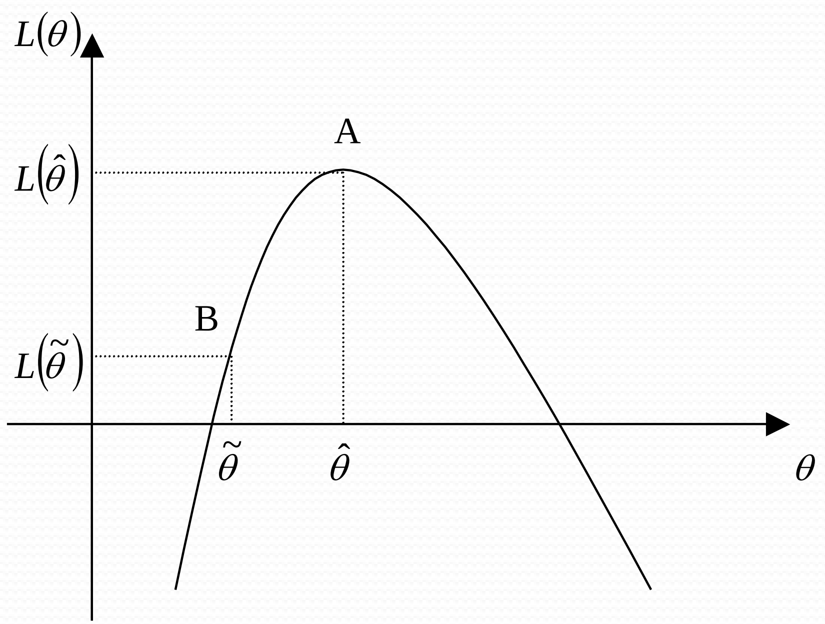
Comparison of Testing Procedures – 1

- Denoting the maximised value of the LLF by unconstrained ML as L($\hat{\theta}$) and
  the constrained optimum as $L(\tilde{\theta})$



Comparison of Testing Procedures – 2

- The vertical distance forms the basis of the LR test.
- The Wald test is based on a comparison of the horizontal distance.
- The LM test compares the slopes of the curve at A and B.



Likelihood Ratio Tests – 1

- Estimate under the null hypothesis and under the alternative.
- Then compare the maximised values of the LLF.
- So we estimate the unconstrained model and achieve a given maximised value of the LLF, denoted Lu
- Then estimate the model imposing the constraint(s) and get a new value of the LLF denoted Lr.

- Note, $Lr \leq Lu$ comparable to $RRSS \geq URSS$

- The LR test statistic is given by

$$LR = -2(Lr - Lu) \sim \chi^2(m)$$

where m = number of restrictions

Likelihood Ratio Tests: Example

- We estimate a GARCH model and obtain a maximised LLF of 66.85. We are interested in testing whether $\beta = 0$ in the following equation.

$$y_t = \mu + \phi y_{t-1} + u_t, \quad u_t \sim N(0, \sigma^2_t)$$

$$\sigma^2_t = \alpha_0 + \alpha_1 u^2_{t-1} + \beta\sigma^2_{t-1}$$

- We estimate the model imposing the restriction and observe the maximised LLF falls to 64.54. Can we accept the restriction?

- LR = -2(64.54-66.85) = 4.62.
- The test follows a $\chi 2(1) = 3.84$ at 5%, so reject the null.

Example: Application of GARCH Models in stock markets

Goal

Goal

- to consider the out of sample forecasting performance of GARCH and EGARCH Models for predicting stock index volatility;
- compare GARCH with implied volatility (the markets expectation of the “average” level of volatility).

Used data

- weekly closing prices (Wednesday to Wednesday, and Friday to Friday) for the S&P100 Index option and the underlying 11 March 83 - 31 Dec. 89

Source

- Day & Lewis (1992)

The Model – 1

The "Base" Models

- For the conditional mean $$R_{Mt} - R_{Ft} = \lambda_0 + \lambda_1 \sqrt{h_t} + u_t$$

- And for the variance $$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}$$

or

$$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 \left(\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left[\left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| - \left(\frac{2}{\pi} \right)^{1/2} \right] \right)$$

where

- R_{Mt} - denotes the return on the market portfolio
- R_{Ft} - denotes the risk-free rate
- h_t - denotes the conditional variance from the GARCH-type models
- σ_t^2 denotes the implied variance from option prices.

The Model: add in a lagged value of the implied volatility

Add in a lagged value of the implied volatility parameter to equations:

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + \delta\sigma_{t-1}^2$$

$$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1 (\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma \left[\left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| - \left(\frac{2}{\pi} \right)^{1/2} \right]) + \delta \ln(\sigma_{t-1}^2)$$

Tests – 1

H0 : $\delta = 0$ (base model).

H0 : $\alpha_1 = 0$ and $\beta_1 = 0$

H0 : $\alpha_1 = 0$ and $\beta_1 = 0$ and $\theta = 0$ and $\gamma = 0$.

Tests – 2

- If this second set of restrictions holds, then

$$h_t^2 = \alpha_0 + \delta\sigma_{t-1}^2$$

$$\ln(h_t^2) = \alpha_0 + \delta \ln(\sigma_{t-1}^2)$$

- We can test all of these restrictions using a likelihood ratio test.

In-sample Likelihood Ratio Test Results: GARCH

Equation for Variance specification	λ_0	λ_1	$\alpha_0 \times 10^{-4}$	α_1	β_1	δ	Log-L	χ^2
$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1}$	0.0072 (0.005)	0.071 (0.01)	5.428 (1.65)	0.093 (0.84)	0.854 (8.17)	-	767.321	17.77
$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + \delta \sigma_{t-1}^2$	0.0015 (0.028)	0.043 (0.02)	2.065 (2.98)	0.266 (1.17)	-0.068 (-0.59)	0.318 (3.00)	776.204	-
$h_t^2 = \alpha_0 + \delta \sigma_{t-1}^2$	0.0056 (0.001)	-0.184 (-0.001)	0.993 (1.50)	-	-	0.581 (2.94)	764.394	23.62

In-sample Likelihood Ratio Test Results: EGARCH

	λ_0	λ_1	$\alpha_0 \times 10^{-4}$	β_1	θ	γ	δ	Log-L	χ^2
$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1\left(\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma\left[\left\lvert\frac{u_{t-1}}{\sqrt{h_{t-1}}}\right\rvert - \left(\frac{2}{\pi}\right)^{1/2}\right]\right)$	-0.0026 (-0.03)	0.094 (0.25)	-3.62 (-2.90)	0.529 (3.26)	-0.273 (-4.13)	0.357 (3.17)	-	776.436	8.09
$\ln(h_t) = \alpha_0 + \beta_1 \ln(h_{t-1}) + \alpha_1\left(\theta \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \gamma\left[\left\lvert\frac{u_{t-1}}{\sqrt{h_{t-1}}}\right\rvert - \left(\frac{2}{\pi}\right)^{1/2}\right]\right) + \delta \ln(\sigma_{t-1}^2)$	0.0035 (0.56)	-0.076 (-0.24)	-2.28 (-1.82)	0.373 (1.48)	-0.282 (-4.34)	0.210 (1.89)	0.351 (1.82)	780.480	-
$\ln(h_t^2) = \alpha_0 + \delta \ln(\sigma_{t-1}^2)$	0.0047 (0.71)	-0.139 (-0.43)	-2.76 (-2.30)	-	-	-	0.667 (4.01)	765.034	30.89

Models conclusions

- Implied volatility has *extra incremental power* for modelling stock volatility beyond GARCH.
- But the models *do not represent a true test* of the predictive ability of implied volatility. So the authors conduct an out of sample forecasting test.
- There are 729 data points. They use the first 410 to estimate the models, and then make a 1-step ahead forecast of the following week's volatility.
- Then they roll the sample forward one observation at a time, constructing a new one step ahead forecast at each step.

Out-of-Sample Forecast Evaluation

- They evaluate the forecasts in two ways:
- The first is by regressing the realised volatility series on the forecasts plus a constant: $\sigma_{t+1}^2 = b_0 + b_1\sigma_{ft}^2 + \xi_{t+1}$

  where σ_{t+1}^2 is the "actual" value of volatility, and σ_{ft}^2 is the value forecasted for it during period t.
- Perfectly accurate forecasts imply $b_0 = 0$ and $b_1 = 1$.

- "True" value of volatility at time t:
  - the square of the weekly return on the index (SR);
  - the variance of the week's daily returns multiplied by the number of trading days in that week (VW).

Out-of Sample Model Comparisons

Forecasting Model	Proxy for *ex post* volatility	b_0	b_1	R^2
Historic	SR	0.0004	0.129	0.094
		(5.60)	(21.18)	
Historic	WV	0.0005	0.154	0.024
		(2.90)	(7.58)	
GARCH	SR	0.0002	0.671	0.039
		(1.02)	(2.10)	
GARCH	WV	0.0002	1.074	0.018
		(1.07)	(3.34)	
EGARCH	SR	0.0000	1.075	0.022
		(0.05)	(2.06)	
EGARCH	WV	-0.0001	1.529	0.008
		(-0.48)	(2.58)	
Implied Volatility	SR	0.0022	0.357	0.037
		(2.22)	(1.82)	
Implied Volatility	WV	0.0005	0.718	0.026
		(0.389)	(1.95)	

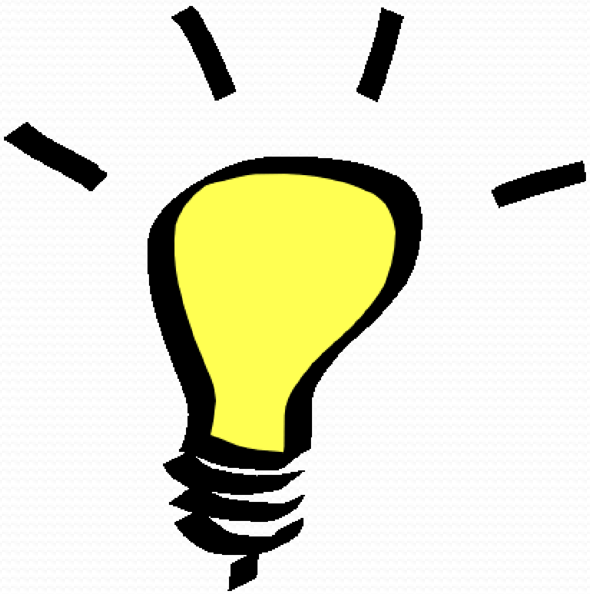
Do the IV Forecasts Encompass those of the GARCH Models?

$$\sigma_{t+1}^2 = b_0 + b_1\sigma_{It}^2 + b_2\sigma_{Gt}^2 + b_3\sigma_{Et}^2 + b_4\sigma_{Ht}^2 + \xi_{t+1}$$

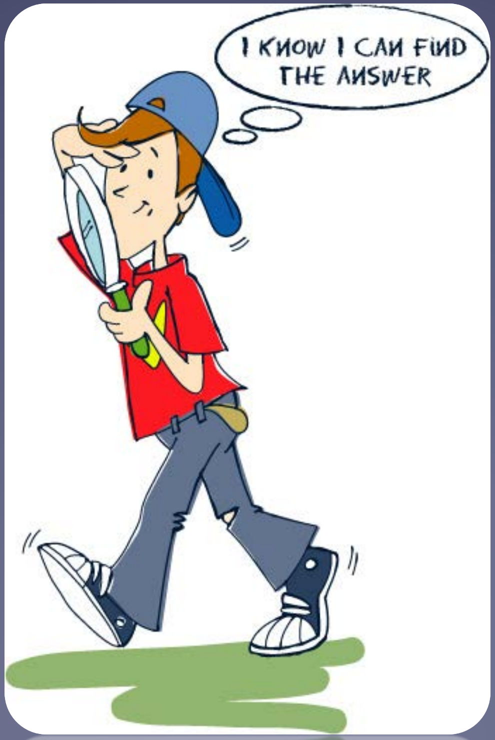
Forecast comparison	b_0	b_1	b_2	b_3	b_4	R^2
Implied vs. GARCH	-0.00010 (-0.09)	0.601 (1.03)	0.298 (0.42)	-	-	0.027
Implied vs. GARCH vs. Historical	0.00018 (1.15)	0.632 (1.02)	-0.243 (-0.28)	-	0.123 (7.01)	0.038
Implied vs. EGARCH	-0.00001 (-0.07)	0.695 (1.62)	-	0.176 (0.27)	-	0.026
Implied vs. EGARCH vs. Historical	0.00026 (1.37)	0.590 (1.45)	-0.374 (-0.57)	-	0.118 (7.74)	0.038
GARCH vs. EGARCH	0.00005 (0.37)	-	1.070 (2.78)	-0.001 (-0.00)	-	0.018

Conclusions

- Within sample results suggest that IV contains extra information not contained in the GARCH / EGARCH specifications.
- Out of sample results suggest that nothing can accurately predict volatility!



Example: Application of GARCH Models for investment decisions



OBJECTIVE

Objective

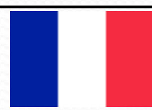
- To establish a variance forecasting model

Why?

- Important for risk managers (VaR)
- Used to price options
- Volatility + Return = investment decision

DATA SET – 1

- Source DataStream
- Period 3/27/1998 - 3/28/2008 (10 years)
- Granularity 1 day

DATA SET – 2

- Local Instruments
  - Change in Exchange Rates
    - EUR / USD / JPY / GBP
  - Change in short-term interest rates
    - T-Bill (US) / BTAN (FR)

- Global Instruments
  - Change in Short-term Eurodollar rate
  - Change in the Term Structure spread

BACKGROUND INFORMATION

- Realized / observed volatility is measured by squared returns
- Volatility displays a positive correlation with its own past
- Simple Model $$\sigma^2_{t+1} = \frac{1}{m}\sum_{\tau=1}^{m} u^2_{t+1-\tau}$$
- PB : Equal weights on the past m observations

FORECASTING MODELS

- Flexible model GARCH (1,1)

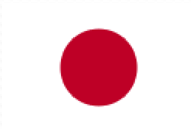
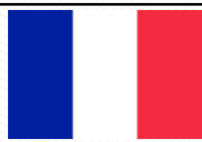
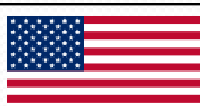
$$\sigma^2_{t+1} = \alpha_0 + \alpha_1 u_t^{\,2} + \beta\,\sigma_t^{\,2}$$

- Extended to Local and Global Instruments

$$\sigma^2_{t+1} = \alpha_0 + \alpha_1 u_t^{\,2} + \beta\,\sigma_t^{\,2} + \sum_{i=1}^{n} (\gamma_i^L Z_i^L + \gamma_i^G Z_i^G)$$

- Models to be tested
  - GARCH (1,1)
  - GARCH (1,1) + Local
  - GARCH (1,1) + Global
  - GARCH (1,1) + Local + Global

Best models for each country

RESULTS

- Best model for German Market
- R^2 of 15.56%
- Final equation
  - Simple GARCH +
  - % Change in €/£ Exchange (L) +
  - % Change in Term Structure Spread (G)
- No universal model
  - Different countries = different models

Example: Application of GARCH Model to forecast the Gold Futures Prices

Gold: An Investment Tool

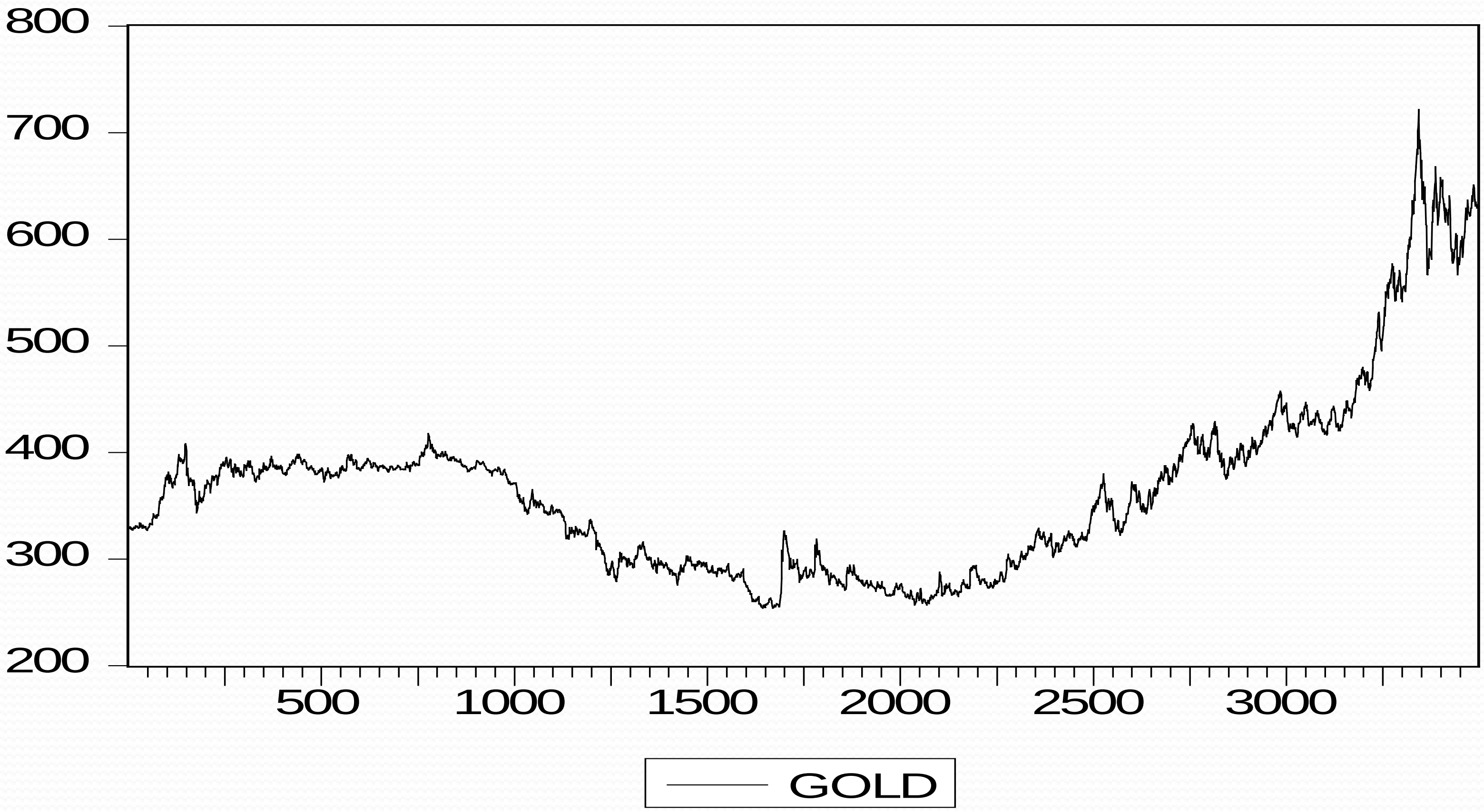
- Equities and Commodities
- Gold forms 45% of total futures trading globally
- An effective hedging tool
- Higher liquidity than other real assets
- Oil price impact on Gold
- Resale value of Gold
- Forecasting the future Gold prices

Data and Methodology

- Daily prices from NYMEX and COMEX
- 14 years, appx. 3500 data points

- ARIMA
- GARCH

Price Graph of Gold Prices



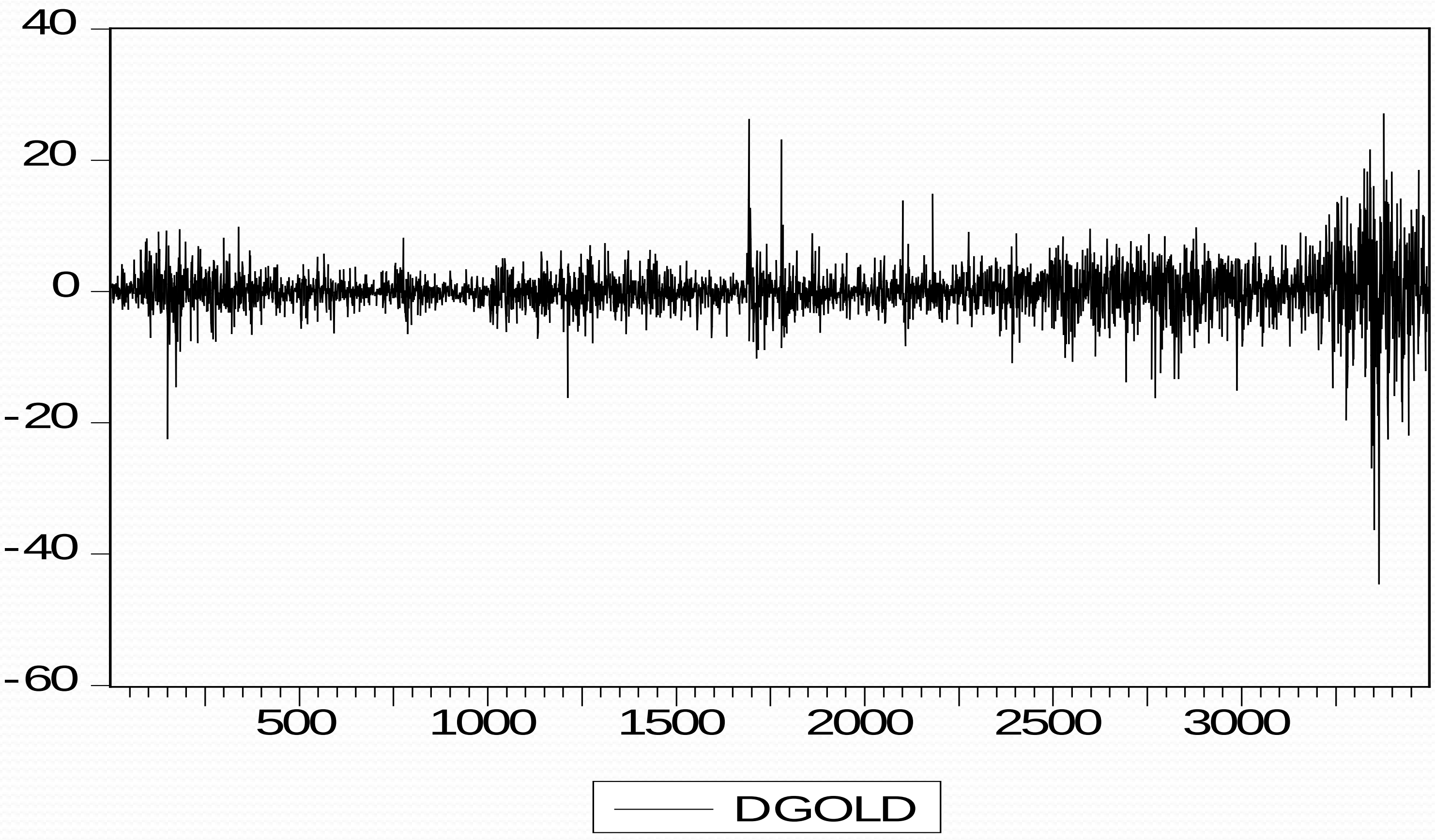
Stationarity

- Gold Price series is not stationary.
- First Difference of the price series.

Dickey Fuller Test

	Dickey Fuller Test (Gold prices, D(Gold))			
Intercept	ADF Test Statistic	-23.863	1% Critical Value*	-3.4353
			Akaike info criterion	5.482112
			Schwarz criterion	5.492703
Trend and Intercept	ADF Test Statistic	-23.9215	1% Critical Value*	-3.9662
			Akaike info criterion	5.481955
			Schwarz criterion	5.494311
None	ADF Test Statistic	-23.8299	1% Critical Value*	-2.5664
			Akaike info criterion	5.481968
			Schwarz criterion	5.490794

Graph



Normality

Mean	0.086144
Median	0
Skewness	-0.74033
Kurtosis	18.11785
Jarque Barra	33582.55
Probability	0

Kurtosis high, indicating a 'fat tail' distribution or a leptokurtic distribution.

Heteroscedasticity

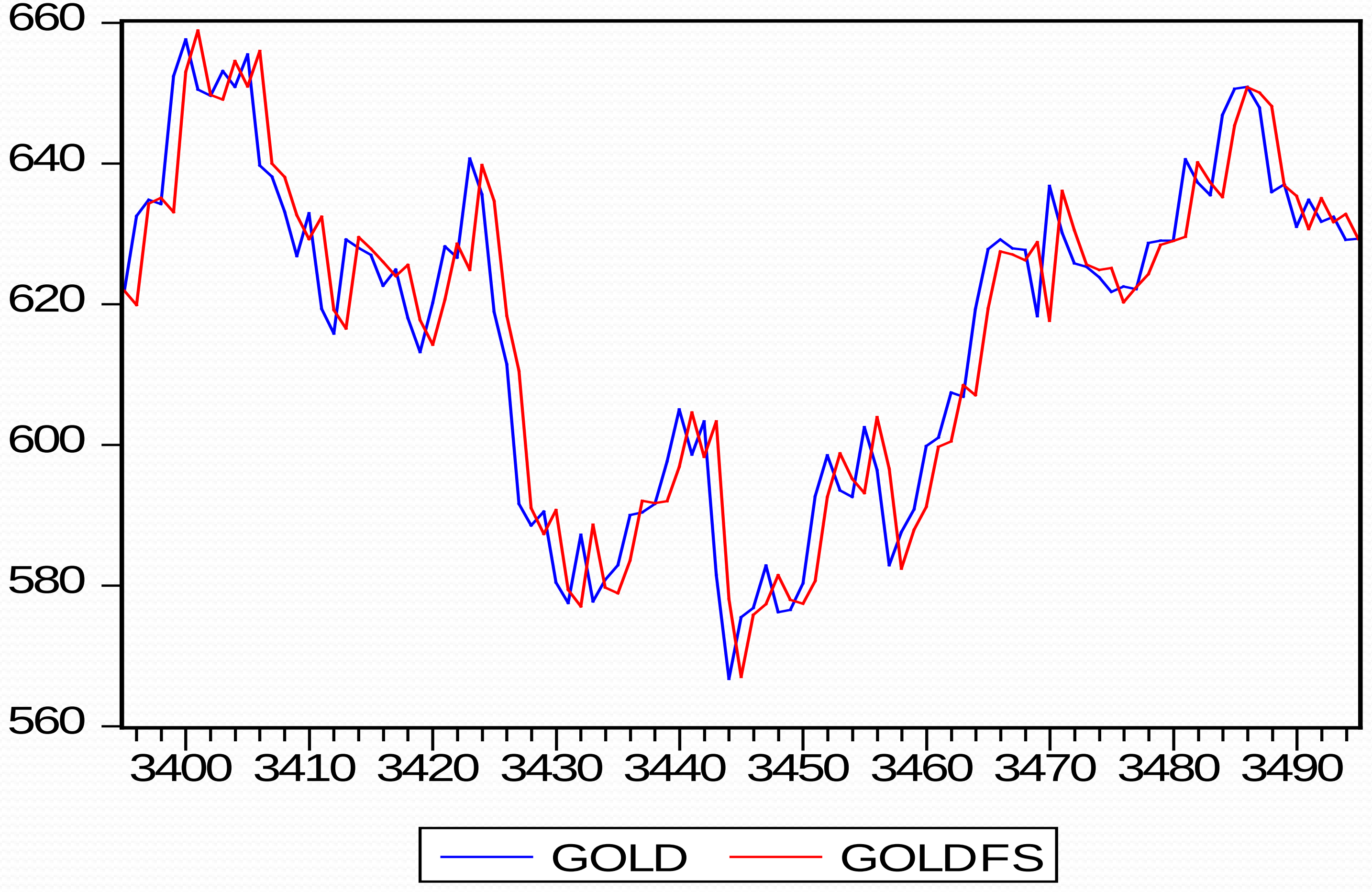
ARCH Test:

F-statistic	74.38698	Probability	0
Obs*R-squared	72.86433	Probability	0

GARCH (1,1)

		Variance Equation		
C	9.54E-08	9.48E-08	1.00621	0.3143
ARCH(1)	0.04162	0.01339	3.10826	0.0019
GARCH(1)	0.960488	0.012698	75.64137	0

Performance of out of sample forecast



Review

Problem

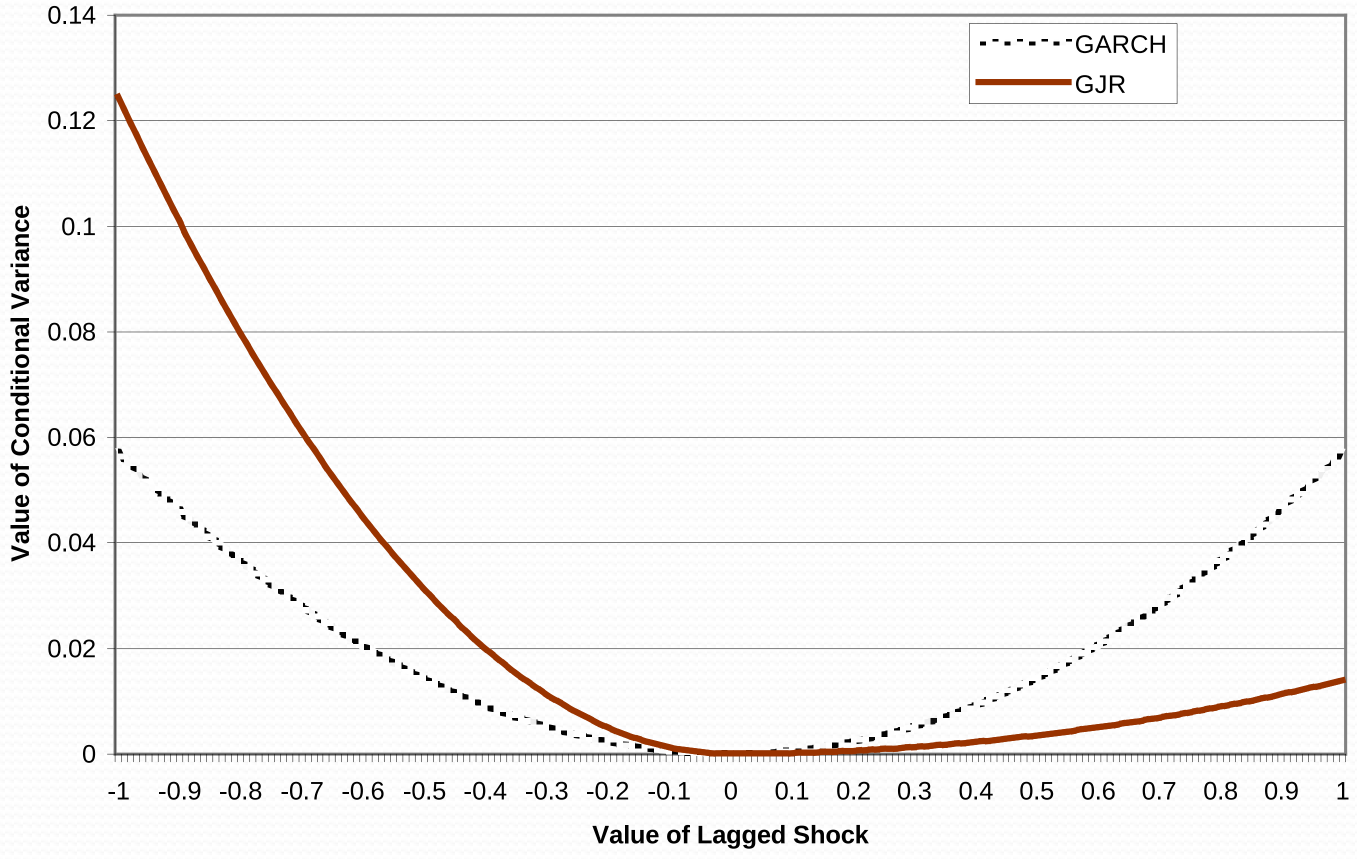
- the linear structural (and time series) models cannot explain a number of important features common to much financial data

Types of GARCH

GARCH(p,q) – the most common non-linear model

- EGARCH,
- GJR,
- GARCH-M etc.

News Impact Curves



Estimation of ARCH / GARCH Models

- Since the model is no longer of the usual linear form, we cannot use OLS.
- We use another technique known as maximum likelihood.
- The method works by finding the most likely values of the parameters given the actual data.
- More specifically, we form a log-likelihood function and maximise it.

Forecasting GARCH-type Models?

- We could show that

$$\mathrm{Var}\,(y_t \mid y_{t-1}, y_{t-2}, \ldots) = \mathrm{Var}\,(u_t \mid u_{t-1}, u_{t-2}, \ldots)$$

- So modelling σ_t^2 will give us models and forecasts for y_t as well.

Questions?

Thank you for your attention!