Modelling Volatility

Brooks, C., Introductory Econometrics for Finance, ch 8.

Topics

- Non-linearity and non-linear models
- Conditional variance
- ARHC model
- ARCH test
- GARCH model
- Parameter estimation using maximum likelihood ML
- Estimation of ARCH/GARCH models in Eviews
- Forecasting variances using GARCH models
- Testing hypotheses about non-linear models
  - LR test
- Extensions to the Basic GARCH Model
  - EGARCH
  - TGARCH
  - IGARCH
  - GARCH-X
  - ARCH-M

Non-linearity

Motivation: the linear time series models cannot explain a number of important features common to much financial data.

- Volatility clustering

  Large changes tend to be followed by large changes and small changes tend to be followed by small changes

- Leptokurtosis

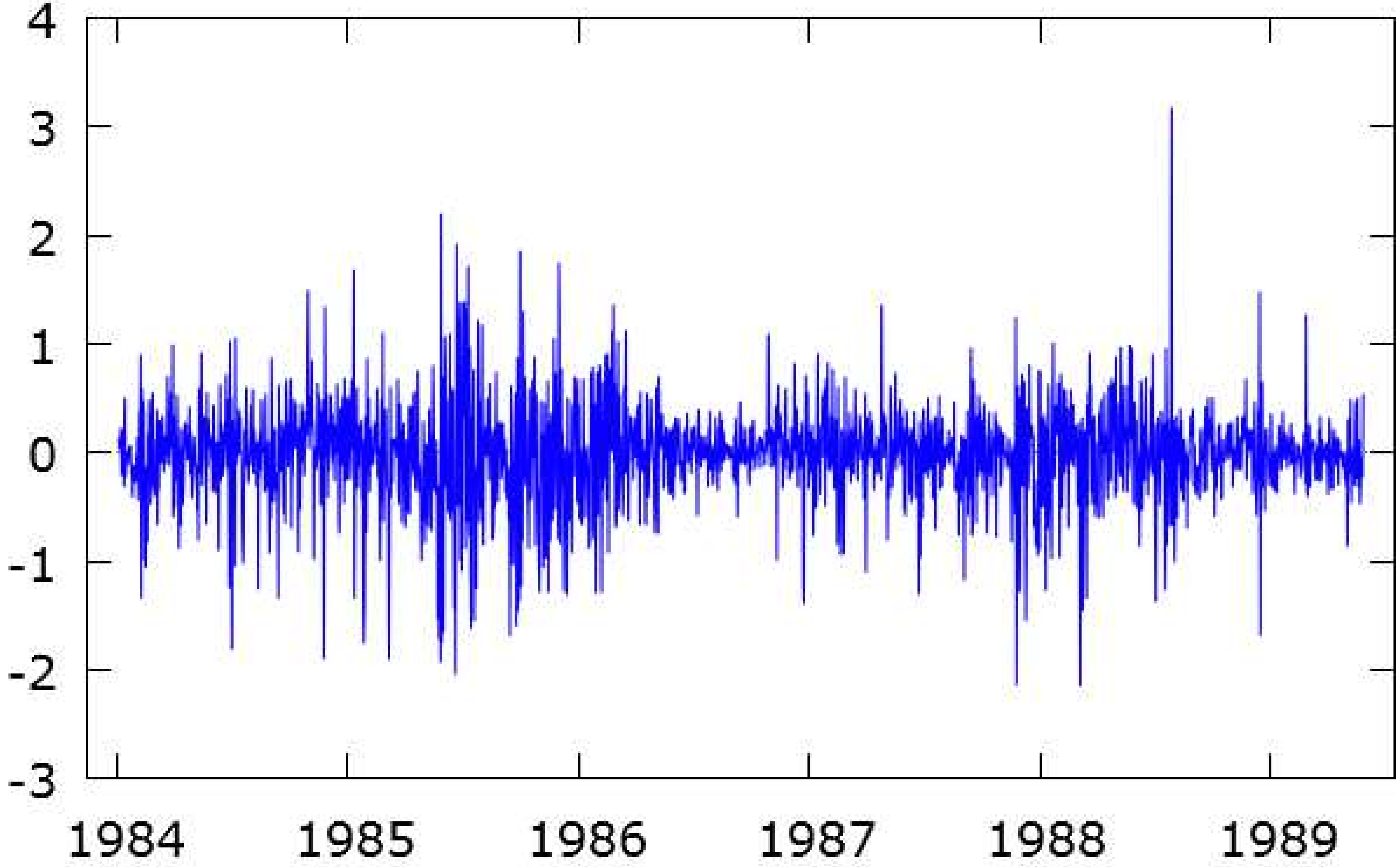
  A distribution with positive excess kurtosis is called leptokurtic. The shape of a distribution is more peaked than that of a normal distribution is called leptokurtic

- Leverage effect

  The leverage effect refers to the generally negative correlation between an asset return and its changes of volatility. Volatility increases when the stock price falls.

An Example: Volatility Clustering

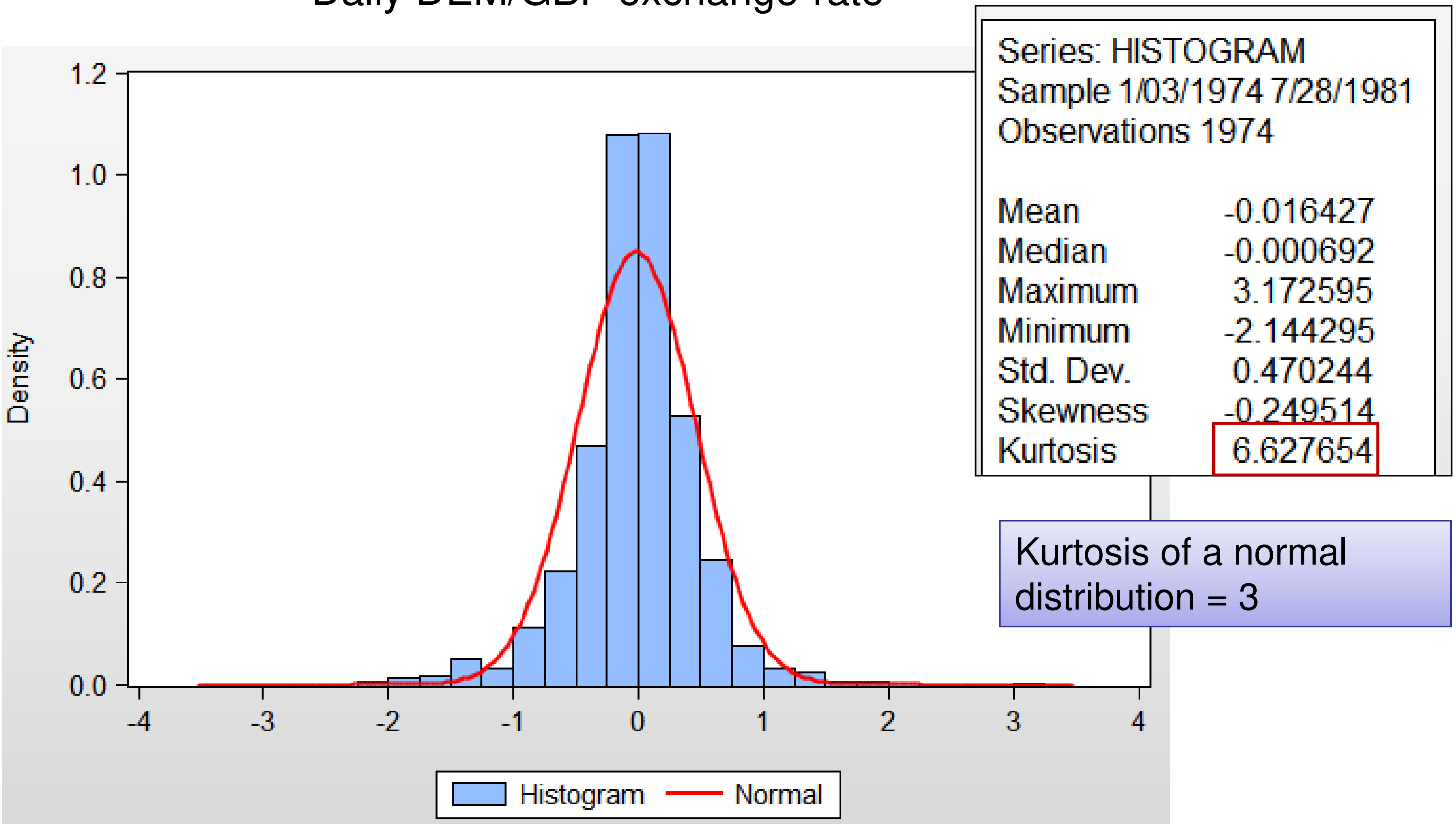
Daily DEM/GBP exchange-rate



Bollerslev, T. & Ghysels, E. (1996), Periodic Autoregressive Conditional Heteroscedasticity ,*Journal of Business & Economic Statistics,* Vol. 14, 139-151

An Example: Leptokurtosis

Daily DEM/GBP exchange-rate

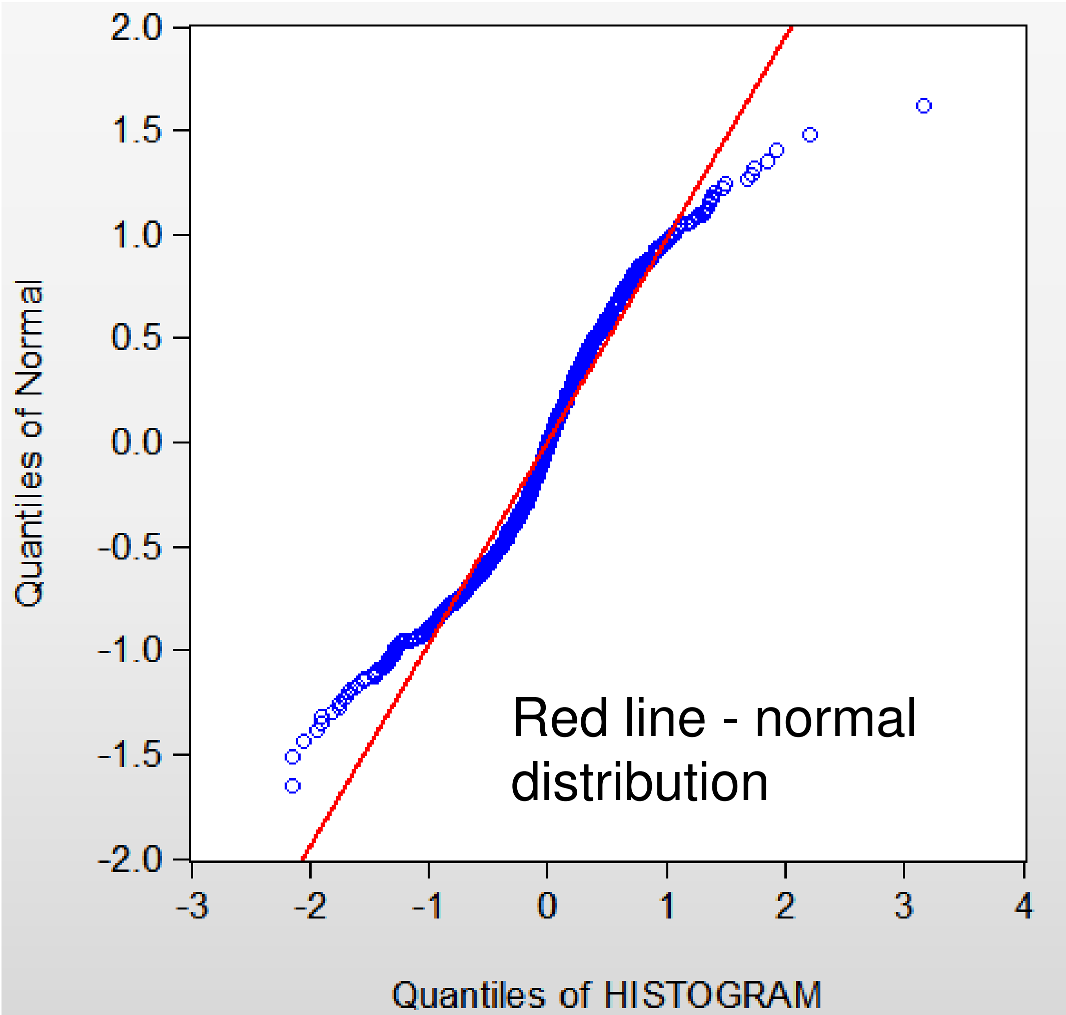


Kurtosis of a normal distribution = 3

An Example: Leptokurtosis

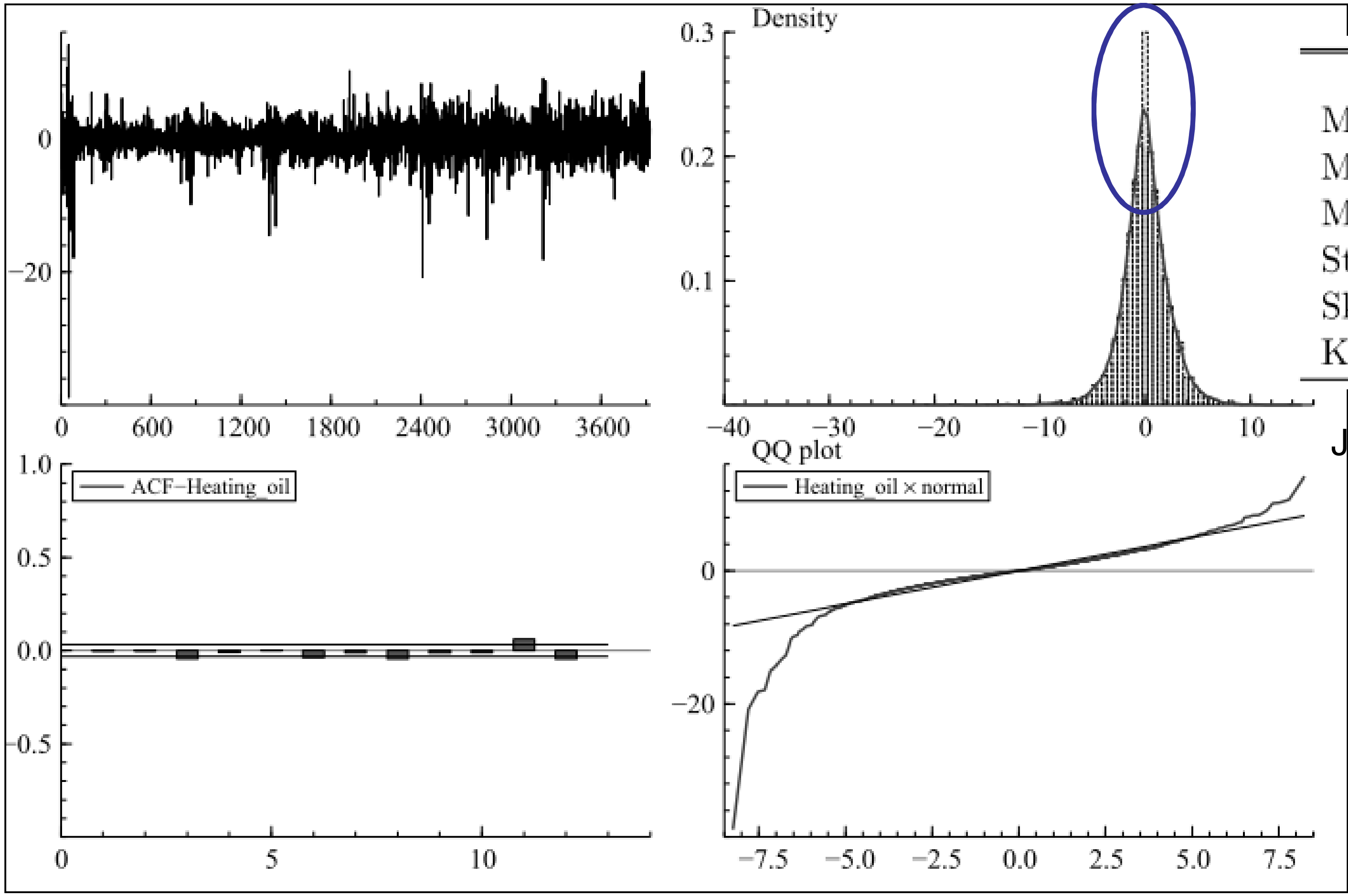
Daily DEM/GBP exchange-rate

Q-Q plot



An Example: Returns on Heating Oil Futures

Daily data on futures prices on heating oil traded in the New York Mercantile Exchange between November 1 1990 and November 22 2005.



	Heating oil
Minimum	-39.09
Maximum	13.99
Mean	0.016
Stand. dev.	2.38
Skewness	-1.75
Kurtosis	27.40

Jarque-Bera $9.9 \cdot 10^4$

$(p = 0.00)$

Marzo, M. & Zagaglia, P. Conditional Leptokurtosis in Energy Prices: Multivariate Evidence from Futures Markets *Dipartimento Scienze Economiche, Universita' di Bologna,* **2007**

Non-linear Models: A Definition

Campbell, Lo and MacKinlay (1997) define a non-linear data generating process as one that can be written

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

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

They also give a slightly more specific definition as

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

$g(\bullet)$ Generally non-linear in mean

$\sigma^2(\bullet)$ Generally non-linear in variance

Types of Non-linear Models

- Autoregressive conditional heteroskedastic model ARCH, GARCH
- Other models
  - Bi-linear model BL
  - Exponential autoregressive model EAR
  - Threshold autoregressive model TAR
  - Smooth threshold autoregressive model STAR
  - AR(p) model with time-dependent coefficients
  - Mixed non-linear model BL-ARCH, STAR-ARCH

Gooijer, J. G. D. & Kumar, K. Some recent developments in non-linear time series modeling, testing, and forecasting. *International Journal of Forecasting*, 1992, 8, 135 – 156.

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:

$$y_t = \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \ldots + \alpha_k x_{kt} + u_t$$

The original model

Test regression

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \alpha_k x_{kt} + \gamma_1 \hat{y}_t^2 + \gamma_2 \hat{y}_t^3 + \ldots + \gamma_{p-1} \hat{y}_t^p + v_t$$

- Many other non-linearity tests are available, e.g. the "BDS test" and the bispectrum test.

Models for Volatility

- Historical volatility
  - Calculating the variance over some historical period, and this then becomes the volatility forecast for all future periods
- Exponentially weighted moving average
  - A simple extension of the historical volatility: more recent observations to have a stronger impact on the forecast of volatility than older data points.
- Autoregressive volatility models
  - A time series of observations on some volatility proxy are obtained. Proxies: squared returns or range estimators.
- Autoregressive conditionally heteroscedastic (ARCH) models

Unconditional and conditional variance

Unconditional variance

$$\text{Var}(u_t) = \sigma^2$$ Constant over time

A conditional variance is the variance of a random variable given the value(s) of one or more other variables X

$$\text{Var}(u_t \mid x_{t-1}, x_{t-2}, \ldots) = \sigma_t^2$$

Changes over time

Conditional Variance and Exogeneous Variable

$$y_{t+1} = u_{t+1} x_t$$

y_{t+1} is the variable of interest

u_t a white noise disturbance term, with constant variance σ^2

x_t an indenpendent variable that can be observed at period *t*

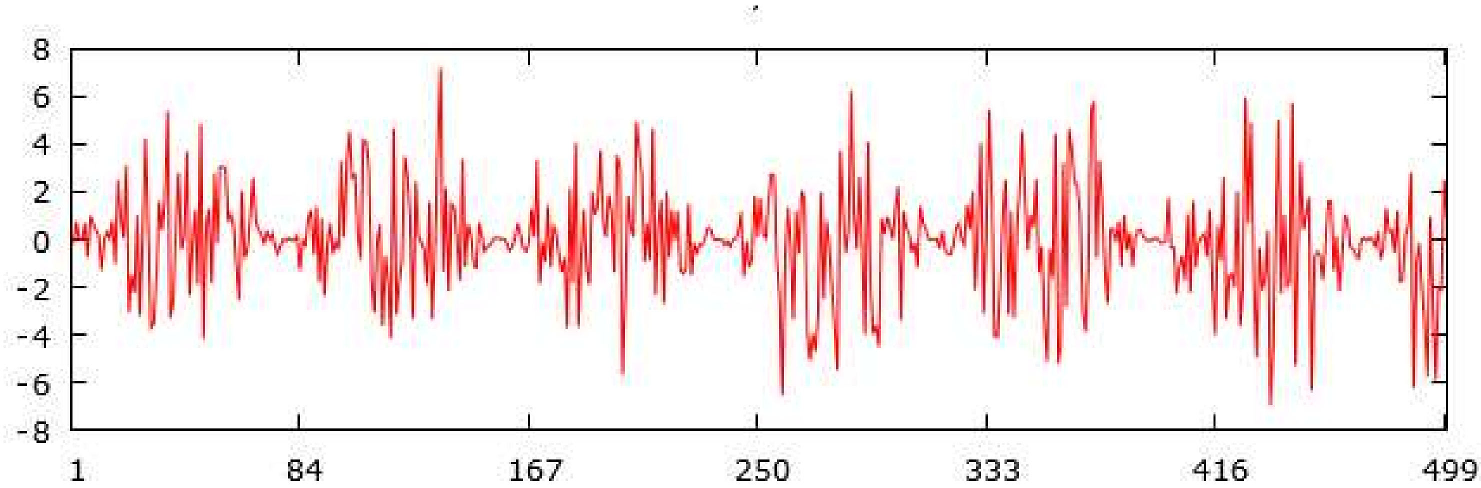
If $x_t = x_{t-1} = x_{t-2} = \ldots = const$ then Var(Y)= constant

If $\{x_t\} \neq const$ then *Y* variance depends on *t*

Conditional variance $\mathrm{Var}\left(y_{t+1} \mid x_t\right) = x_t^2 \sigma^2$

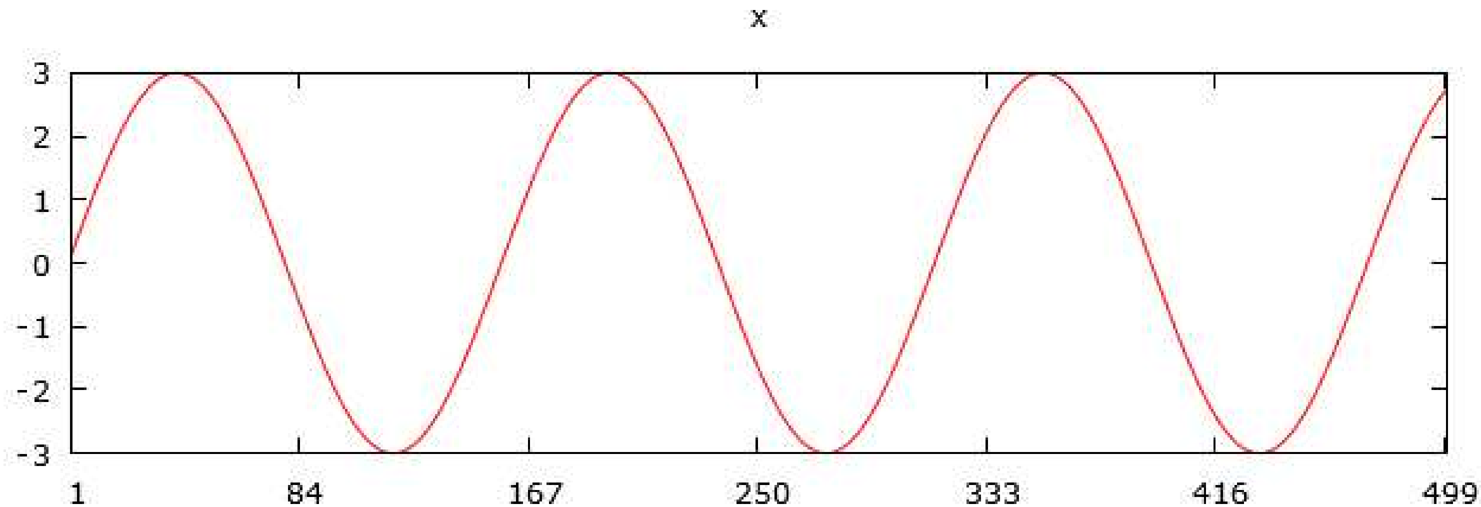
Conditional Variance and Exogeneous Variable: Simulation

$$y_t = u_t x_{t-1}$$



Exogeneous variable

$$x_t = 3\sin(0.04t)$$

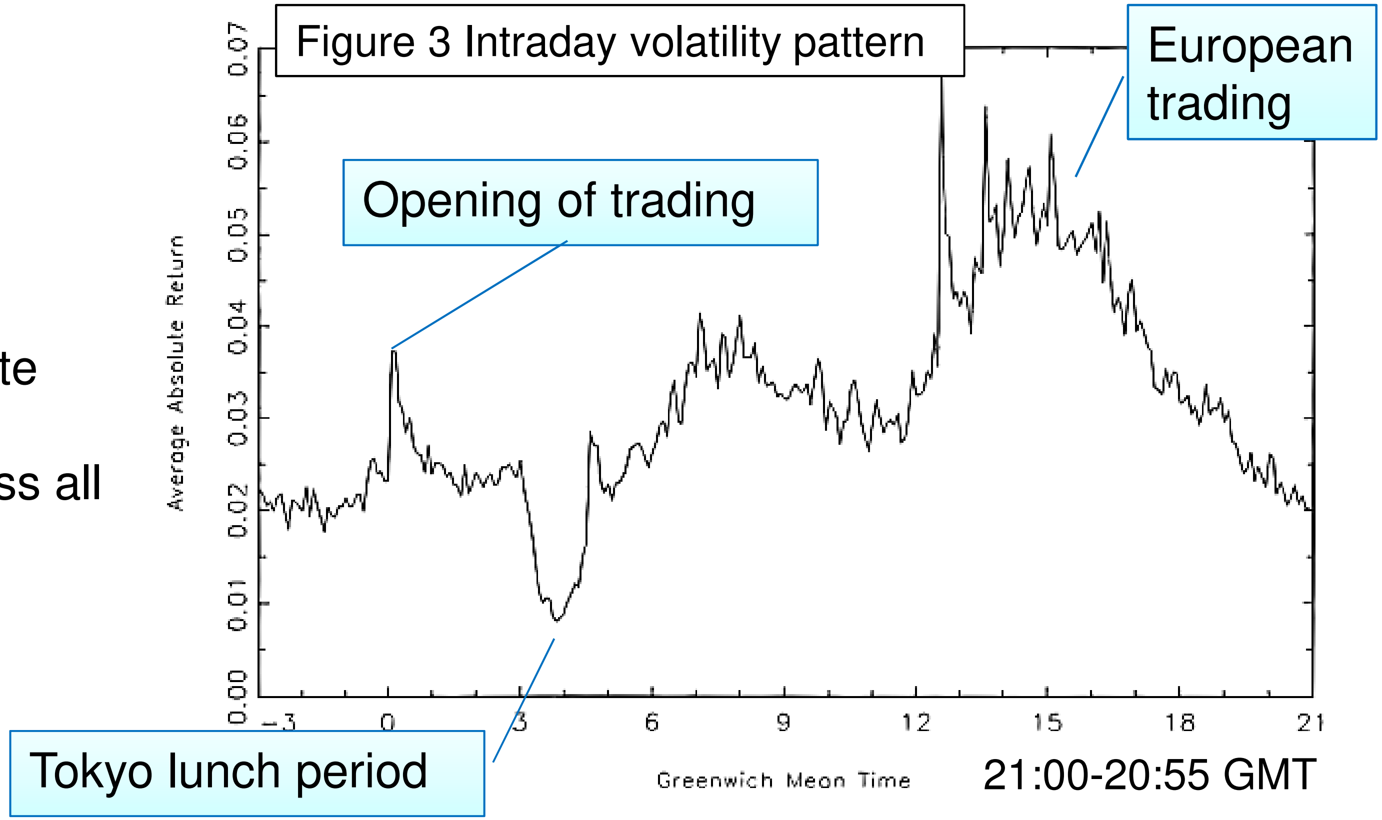


An Example: DEM–USD Exchange Rate Volatility

Andersen and Bollerslev (1988): five-minute returns for the deutsche mark–dollar spot exchange rate from October 1, 1992, through September 30, 1993.

Calendar effects

The average absolute return for each five-minute interval across all 260 weekdays.

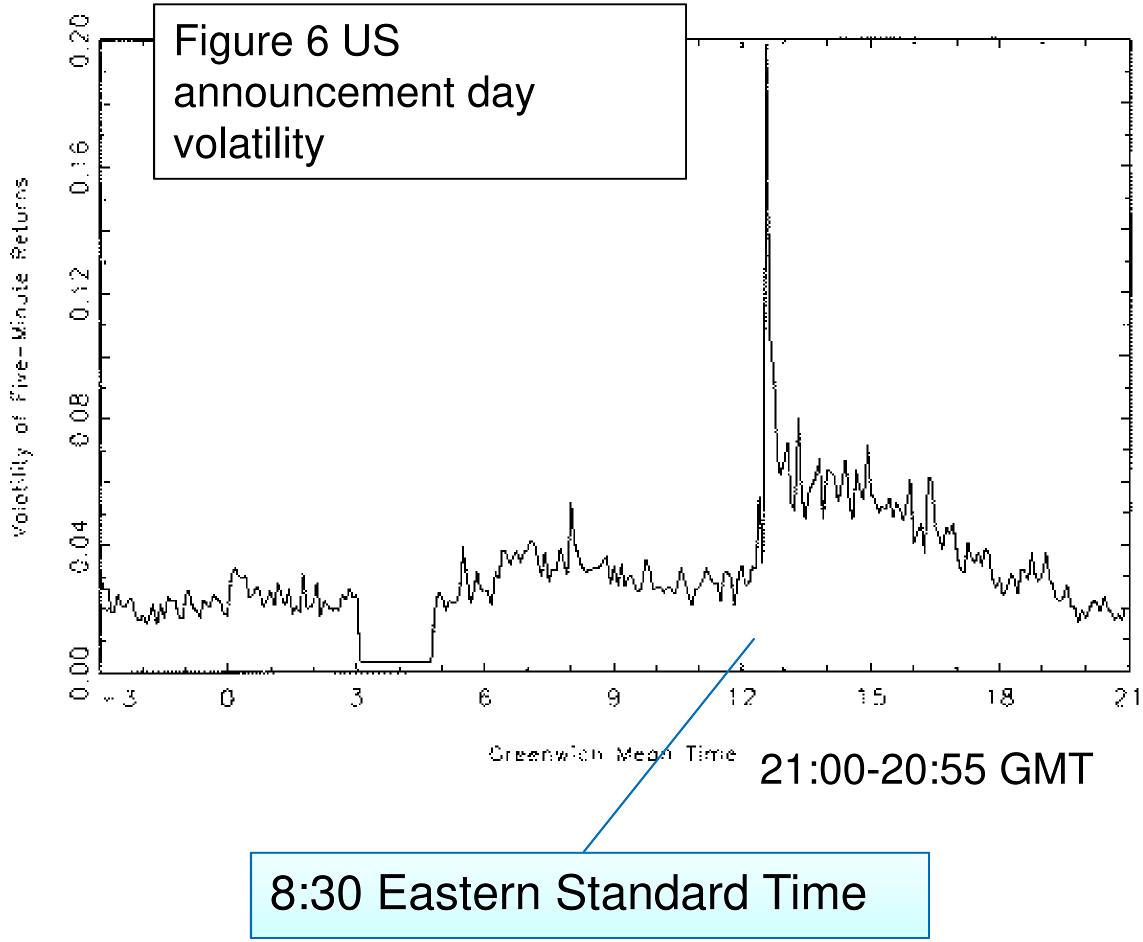


Andersen, T. G. & Bollerslev, T. (1998), Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies The Journal of Finance, Blackwell Publishers, Inc., Vol. 53 , 219-265

An Example: DEM–USD Exchange Rate Volatility (cont'd)

Andersen and Bollerslev (1988): five-minute returns for the deutsche mark–dollar spot exchange rate from October 1, 1992, through September 30, 1993.

U.S. macroeconomic announcements are the source of the observed volatility spikes



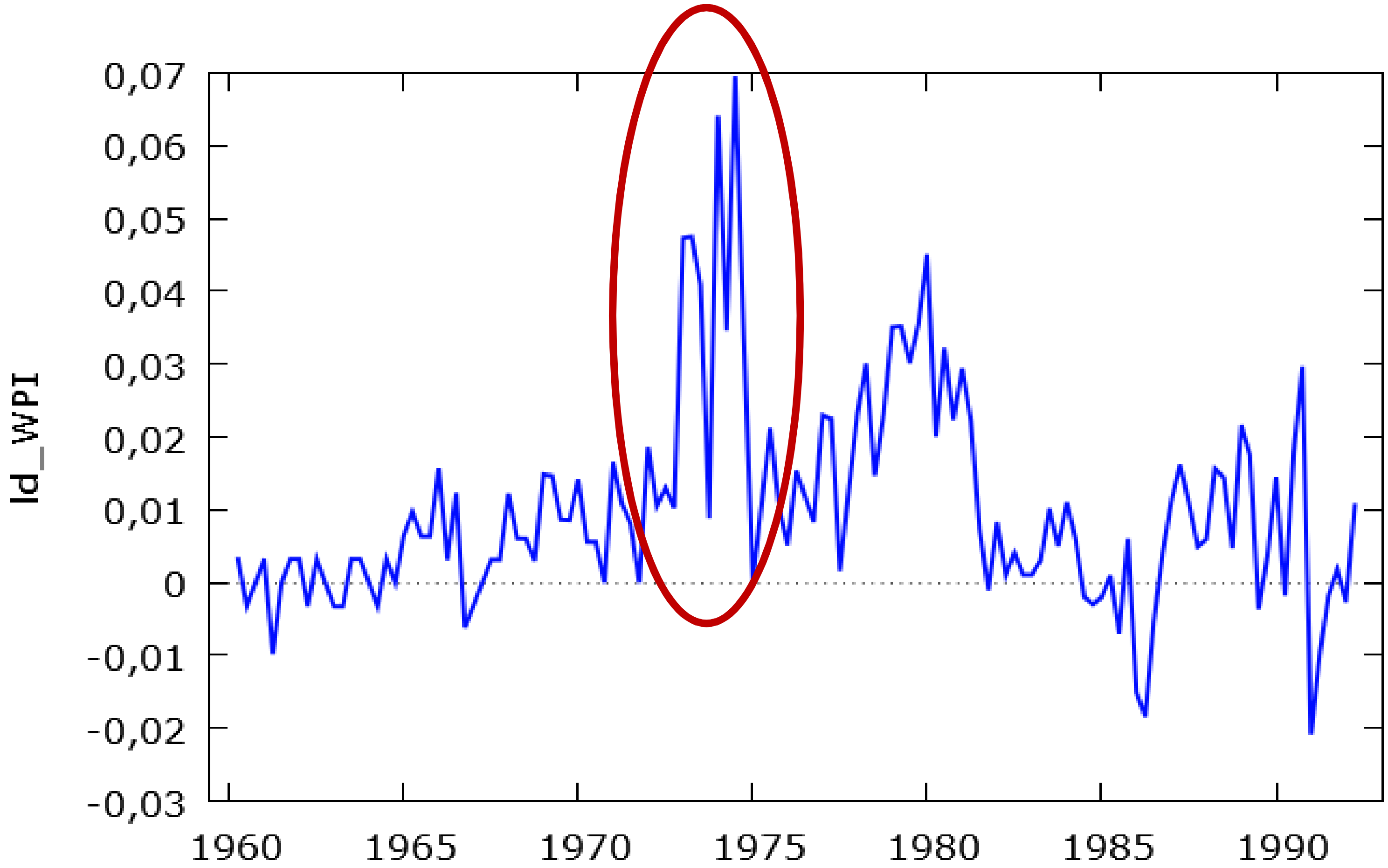
A Problem with Exogeneous Variable

A major difficulty with this strategy is

- it assumes a specific cause X for the changing variance

An Example: US Wholesale Price Index

Reason of large volatility in the middle of 1970s?
Was it the oil price shocks? A change in the conduct of monetary policy?
Something other?



Autoregressive Conditionally Heteroscedastic (ARCH) Models

The conditional variance of u_t:

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

We usually assume that $E[u_t] = 0$

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

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:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$ An ARCH(1) model

Autoregressive Conditionally Heteroscedastic (ARCH) Models (cont'd)

- The full model would be

$$y_t = \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t, \quad u_t \sim N(0, \sigma_t^2)$$

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

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

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_q u_{t-q}^2$$

An ARCH(q) model

- Instead of calling the variance σ_t^2, in the literature it is usually called h_t, so the model is

$$y_t = \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t, \quad u_t \sim N(0, h_t)$$

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_q u_{t-q}^2$$

Another Way of Writing ARCH Models

- For illustration, consider an ARCH(1). Instead of the above, we can write

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

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2}, \qquad v_t \sim N(0,1)$$

This is known as Engle spetsification

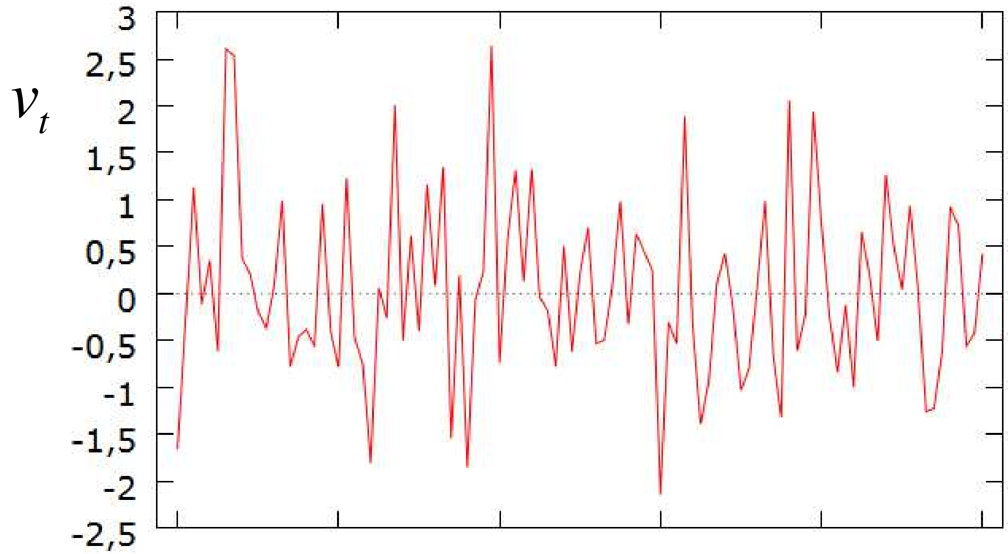
- The two are different ways of expressing exactly the same model. The first form is easier to understand while the second form is required for simulating from an ARCH model, for example.

Engle, R. F. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation *Econometrica*, Vol. 50, pp. 987-1007.

An Example: ARCH(1)

White noise

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

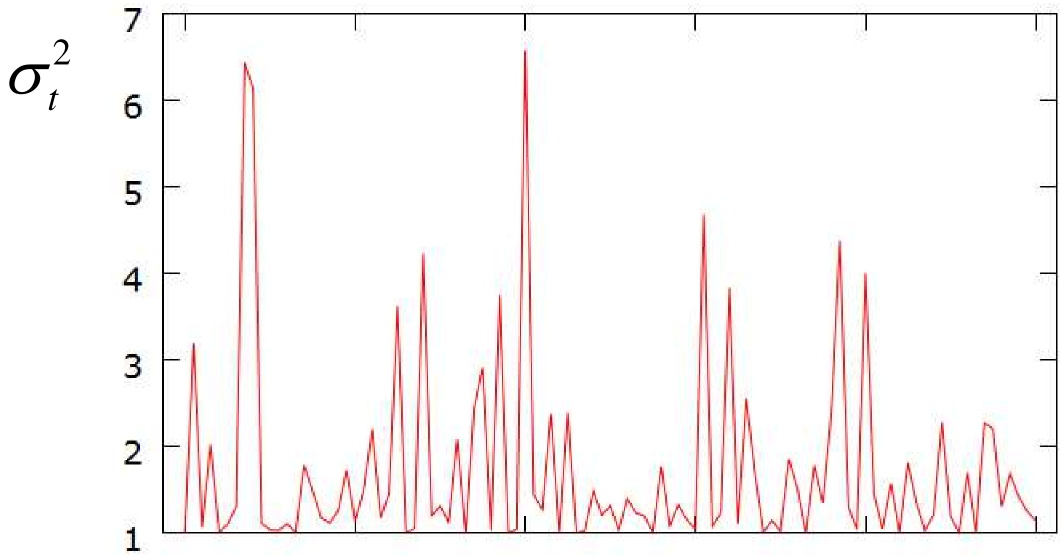


ARCH(1)

Conditional variance

$$\sigma_t^2 = 1 + 0.8u_{t-1}^2$$

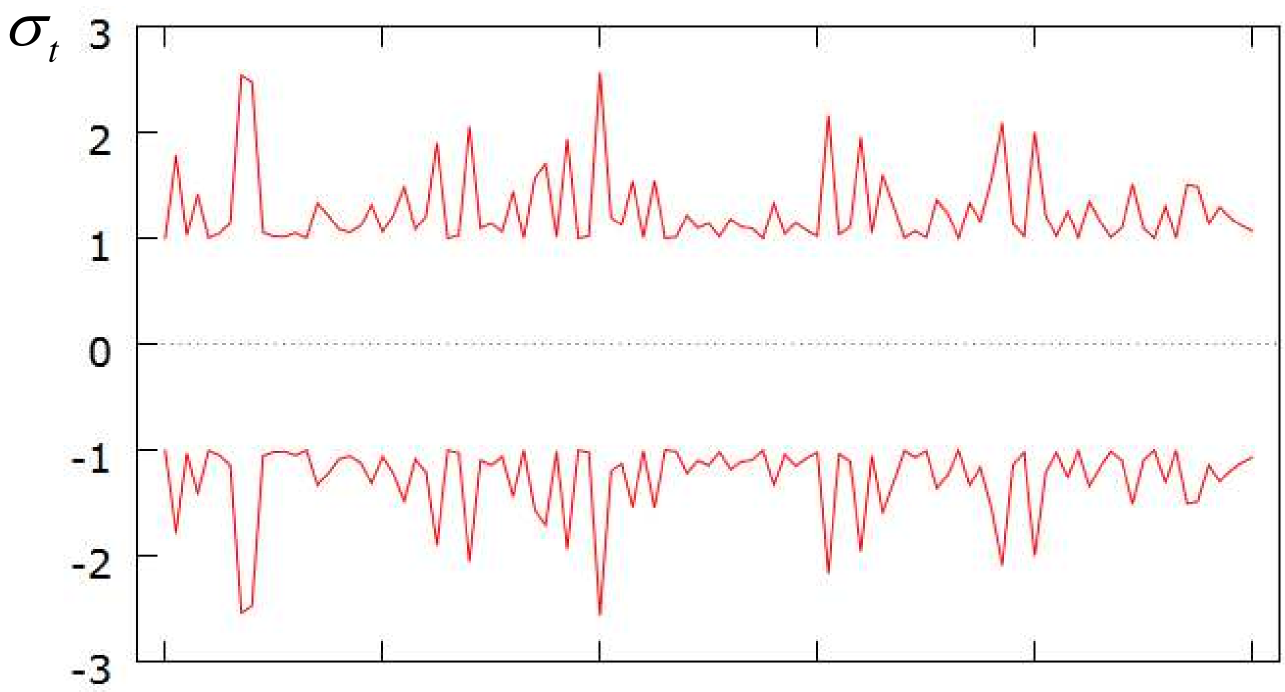
where $u_t = v_t \sigma_t$



An Example: ARCH(1) (cont'd)

Conditional standard deviation

$$\sigma_t = \pm\sqrt{\sigma_t^2}$$



ARMA-ARCH(1) Model

Conditional mean $y_t = \beta_1 + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + u_t, \quad u_t \sim N(0, \sigma_t^2)$

Conditional variance $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$

There is no restrictions for the conditional mean model.
Its not specified.

So process Y may be an ARMA process

For example ARMA(1,1) – ARCH(1) model

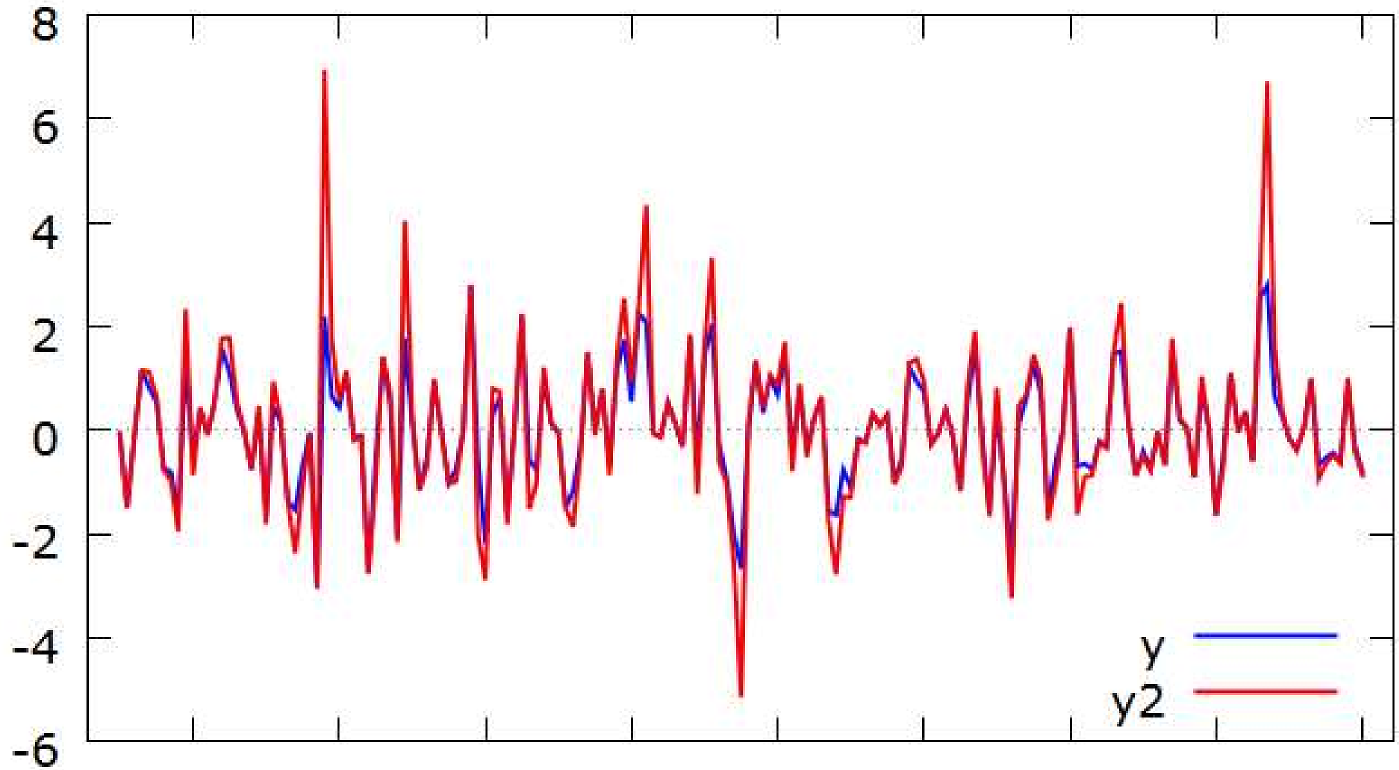
$$y_t = c + \phi_1 y_{t-1} + u_t + \theta_1 u_{t-1}, \quad u_t \sim N(0, \sigma_t^2)$$ ARMA(1,1)

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2$$ ARCH(1)

An Exmple: AR(1) - ARCH(1)

$$y_t = 0.1 y_{t-1} + v_t, \quad v_t \sim N(0,1)$$

AR(1)



$$y_{2t} = 0.1 y_{2\,t-1} + u_t, \quad u_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = 1 + 0.8 u_{t-1}^2$$

AR(1) – ARCH(1)

Testing for "ARCH Effects"

1. 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 $\hat{u}_t$

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

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \gamma_2 \hat{u}_{t-2}^2 + \ldots + \gamma_q \hat{u}_{t-q}^2 + v_t$$

   v_t is iid

   Obtain R^2 from this regression.

3. The null and alternative hypotheses are

$$\mathrm{H}_0: \quad \gamma_1 = \gamma_2 = \ldots = \gamma_q = 0$$

There is no ARCH effect

$$\mathrm{H}_1: \quad \gamma_1 \neq 0 \quad \text{or} \quad \gamma_2 \neq 0 \quad \text{or} \ldots \text{or} \quad \gamma_q \neq 0$$

Testing for "ARCH Effects" (cont'd)

4. There are two test statistics

- The *F*-statistic is an omitted variable test for the joint significance of all lagged squared residuals. The exact finite sample distribution of the *F*-statistic under H_0 is not known.

- The Engle's LM teststatistic is the number of observations *T* multiplied by the R^2 from the last regression, and is asymptotically distributed as a χ^2

$$LM = T \cdot R^2 \sim \chi^2(q)$$

5. If the value of the test statistic is greater than the critical value, then reject the null hypothesis.

In the presence of ARCH, OLS is consistent, but inefficient.

Testing for “ARCH Effects” (cont'd)

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \gamma_2 \hat{u}_{t-2}^2 + \ldots + \gamma_q \hat{u}_{t-q}^2 + v_t$$

How many lags?

One way to choose *q* is to compare loglikelihood values for different choices of q. You can use

- the likelihood ratio test
- or information criteria

to compare loglikelihood values.

ARCH Test, an Example

Pound-dollars returns are estimated as ARMA(1,1) model

In EViews after the estimation:
View->Residual Test-> Heteroskedasticity Tests -> Choose ARCH

Number of lags 5

Test Equation

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \gamma_2 \hat{u}_{t-2}^2 + + \gamma_2 \hat{u}_{t-3}^2 + \gamma_4 \hat{u}_{t-4}^2 + \gamma_4 \hat{u}_{t-5}^2 + v_t$$

Heteroskedasticity Test: ARCH

F-statistic	5.909063	Prob. F(5,1814)	0.0000
Obs*R-squared	29.16797	Prob. Chi-Square(5)	0.0000

The ARCH effect exists

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 12/05/16 Time: 15:17
Sample (adjusted): 7/14/2002 7/07/2007
Included observations: 1820 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.154689	0.011369	13.60633	0.0000
RESID^2(-1)	0.118068	0.023475	5.029627	0.0000
RESID^2(-2)	-0.006579	0.023625	-0.278463	0.7807
RESID^2(-3)	0.029000	0.023617	1.227920	0.2196
RESID^2(-4)	-0.032744	0.023623	-1.386086	0.1659
RESID^2(-5)	-0.020316	0.023438	-0.866798	0.3862

R-squared	0.016026	Mean dependent var	0.169496
Adjusted R-squared	0.013314	S.D. dependent var	0.344448
S.E. of regression	0.342147	Akaike info criterion	0.696140
Sum squared resid	212.3554	Schwarz criterion	0.714293
Log likelihood	-627.4872	Hannan-Quinn criter.	0.702837
F-statistic	5.909063	Durbin-Watson stat	1.995904
Prob(F-statistic)	0.000020		

ARCH Model, Non-negativity Constraints

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \ldots + \alpha_q u_{t-q}^2$$

Variance σ_t^2 cannot be negative.

Squares u_t^2 are always positive.

So parameters α_i must be positive $\alpha_i > 0 \quad \forall \quad i = 1, 2, \ldots$

If some α_i *is* negative, then a sufficiently large realization of u_{t-i} can render a negative value for the conditional variance.

ARCH Model, Stationarity

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2$$ ARCH(q) model

Stationarity condition $$\sum_{i=1}^{q} \alpha_i < 1$$ Same as for AR model

If a process is stationary

Unconditional mean $$E\left[u_t\right] = 0$$

Conditional mean $$E\left[u_t \mid u_{t-1}, u_{t-2}, \ldots .\right] = 0$$

Unconditional variance $$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i}$$

Problems with ARCH(q) Models

- Non-negativity and stationarity constraints might be violated.
- How do we decide on *q*?
  - One possibility: LR test
- The required value of *q* might be very large

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

Generalised ARCH (GARCH) Models

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

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

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

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

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

- Substituting into (1) for σ_{t-1}^2

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

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

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

Generalised ARCH (GARCH) Models (cont'd)

- Now substituting into (2) for σ_{t-2}^2

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_0\beta + \alpha_1\beta u_{t-2}^2 + \beta^2\left(\alpha_0 + \alpha_1 u_{t-3}^2 + \beta\sigma_{t-3}^2\right) =$$

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

$$= \alpha_0(1 + \beta + \beta^2) + \alpha_1 u_{t-1}^2(1 + \beta L + \beta^2 L^2) + \beta^3\sigma_{t-3}^2$$

The lag operator $Lu_t = u_{t-1}$

- An infinite number of successive substitutions would yield

$$\sigma_t^2 = \alpha_0(1 + \beta + \beta^2 + \ldots) + \underbrace{\alpha_1 u_{t-1}^2(1 + \beta L + \beta^2 L^2 + \ldots)}_{\text{ARCH}(\infty)} + \beta^\infty\sigma_0^2$$

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

Generalised ARCH (GARCH) Models (cont'd)

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

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

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

$\sum_{i=1}^{p} \alpha_i \sigma_{t-i}^2$ Autoregressive component

$\sum_{i=1}^{q} \beta_i u_{t-i}^2$ Moving average component

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

Notation

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

GARCH (*p*,*q*)

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

ARCH (1), GARCH (0,1)

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

ARCH (2), GARCH(0, 2)

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \beta_1 \sigma_{t-1}^2$$

GARCH (1,2)

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \alpha_3 u_{t-3}^2 + $$
$$+ \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2$$

GARCH (2,3)

Why is GARCH Better than ARCH?

- More parsimonious - avoids overfitting
- Less likely to breech non-negativity constraints

The Unconditional Variance under the GARCH Specification

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

The unconditional variance of u_t is given by

$$\text{Var}(u_t) = \frac{\alpha_0}{1-(\alpha_1 + \beta)}$$

when $\alpha_1 + \beta < 1$

$\alpha_1 + \beta \geq 1$ is termed "non-stationarity" in variance

$\alpha_1 + \beta = 1$ is termed integrated GARCH (IGARCH)

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.
  - Heteroskedasticity!
- 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 (cont'd)

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

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

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

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

2. Specify the log-likelihood function to maximize

$$L = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}\frac{y_t - \mu - \phi_{t-1}}{\sigma_t^2}$$

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

Parameter Estimation using Maximum Likelihood

Consider the bivariate regression case with homoscedastic errors for simplicity:

$$y_t = \beta_1 + \beta_2 x_t + u_t$$

Assuming that $u_t \sim N(0, \sigma^2)$

then $y_t \sim N(\beta_1 + \beta_2 x_t, \sigma^2)$

so that the probability density function for a normally distributed random variable with this mean and variance is given by

$$f(y_t \mid \beta_1 + \beta_2 x_t, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \frac{(y_t - \beta_1 - \beta_2 x_t)^2}{\sigma^2} \right\} \quad (1)$$

Successive values of y_t would trace out the familiar bell-shaped curve.

Assuming that u_t are iid, then y_t will also be iid.

Parameter Estimation using Maximum Likelihood (cont'd)

Then the joint pdf for all the y’s can be expressed as a product of the individual density functions

$$f(y_1, y_2, \ldots, y_T \mid \beta_1 + \beta_2 x_t, \sigma^2) = f(y_1 \mid \beta_1 + \beta_2 x_1, \sigma^2) \cdot f(y_2 \mid \beta_1 + \beta_2 x_2, \sigma^2) \cdot$$

$$\cdot \ldots \cdot f(y_T \mid \beta_1 + \beta_2 x_T, \sigma^2) = \prod_{t=1}^{T} f(y_t \mid \beta_1 + \beta_2 x_t, \sigma^2) \qquad (2)$$

Substituting into equation (2) for every y_t from equation (1),

$$f(y_1, y_2, \ldots, y_T \mid \beta_1 + \beta_2 x_t, \sigma^2) = \frac{1}{\sigma^T \left(\sqrt{2\pi}\right)^T} \exp\left\{ -\frac{1}{2} \sum_{t=1}^{T} \frac{\left(y_t - \beta_1 - \beta_2 x_t\right)^2}{\sigma^2} \right\} \qquad (3)$$

Parameter Estimation using Maximum Likelihood (cont'd)

The typical situation we have is that the x_t and y_t are given and we want to estimate β_1 β_2, σ^2. If this is the case, then f(●) is known as the likelihood function, denoted $LF(\beta_1$ β_2, $\sigma^2)$, so we write

$$LF(\beta_1, \beta_2 \sigma^2) = \frac{1}{\sigma^T \left(\sqrt{2\pi}\right)^T} \exp\left\{ -\frac{1}{2} \sum_{t=1}^{T} \frac{\left(y_t - \beta_1 - \beta_2 x_t\right)^2}{\sigma^2} \right\} \quad (4)$$

Maximum likelihood estimation involves choosing parameter values (β_1 β_2, σ^2), that maximise this function.

We want to differentiate (4) w.r.t. β_1 β_2, σ^2 but (4) is a product containing T terms.

Parameter Estimation using Maximum Likelihood (cont'd)

Since $\max_{x} f(x) = \max_{x} \ln(f(x))$ we can take logs of (4)

Then, using the various laws for transforming functions containing logarithms, we obtain the log-likelihood function, LLF:

$$LLF = \ln\left[\frac{1}{\sigma^T\left(\sqrt{2\pi}\right)^T}\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\beta_1-\beta_2 x_t\right)^2}{\sigma^2}\right\}\right]$$

$$= -T\ln\sigma - \frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\beta_1-\beta_2 x_t\right)^2}{\sigma^2} \qquad (5)$$

Parameter Estimation using Maximum Likelihood (cont'd)

$$LLF = \ln\left[\frac{1}{\sigma^T\left(\sqrt{2\pi}\right)^T}\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\beta_1-\beta_2x_t\right)^2}{\sigma^2}\right\}\right]$$

$$= -T\ln\sigma - \frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\beta_1-\beta_2x_t\right)^2}{\sigma^2} \qquad (5)$$

Differentiating (5) w.r.t. β_1 , β_2, σ^2

$$\frac{\partial LLF}{\partial\beta_1} = -\frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\beta_1-\beta_2x_t\right)\cdot 2-1}{\sigma^2} \qquad (6)$$

$$\frac{\partial LLF}{\partial\beta_2} = -\frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\beta_1-\beta_2x_t\right)\cdot 2-x_t}{\sigma^2} \qquad (7)$$

$$\frac{\partial LLF}{\partial\sigma^2} = -\frac{T}{2}\frac{1}{\sigma^2} + \frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\beta_1-\beta_2x_t\right)^2}{\sigma^4} \qquad (8)$$

Parameter Estimation using Maximum Likelihood (cont'd)

Setting (6)-(8) to zero to minimise the functions, and putting hats above the parameters to denote the maximum likelihood estimators, we get

from (6) $$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x} \qquad (9)$$

from (7) $$\hat{\beta}_2 = \frac{\sum y_t x_t - T\overline{xy}}{\sum x_t^2 - T\bar{x}^2} \qquad (10)$$

from (8) $$\hat{\sigma}^2 = \frac{1}{T}\sum u_t^2 \qquad (11)$$

(9) & (10) are identical to OLS estimators.
(11) is different. The OLS estimator was $$\hat{\sigma}^2 = \frac{1}{T-2}\sum u_t^2$$

Therefore the ML estimator of the variance of the disturbances is biased, although it is consistent.

But how does this help us in estimating heteroscedastic models?

Estimation of GARCH Models Using Maximum Likelihood

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

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

$$L = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}\frac{y_t - \mu - \phi_{t-1}}{\sigma_t^2}$$

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

The way we do the optimisation is:

1. Set up LLF.
2. Use regression to get initial guesses for the mean parameters.
3. Choose some initial guesses for the conditional variance parameters.
4. 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

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

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

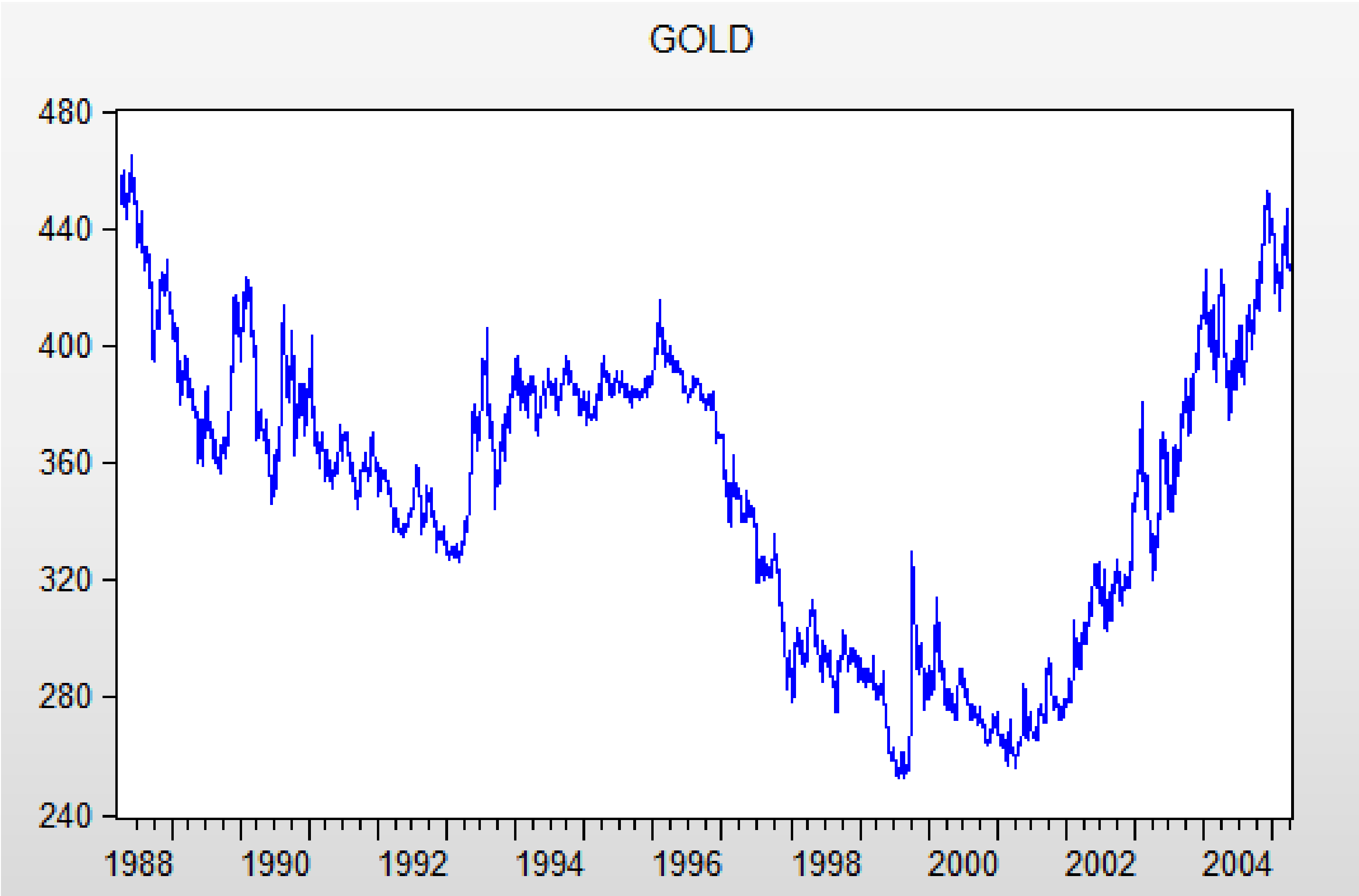
$$v_t = \frac{u_t}{\sigma_t}$$

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 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.

An Example: Price of Gold

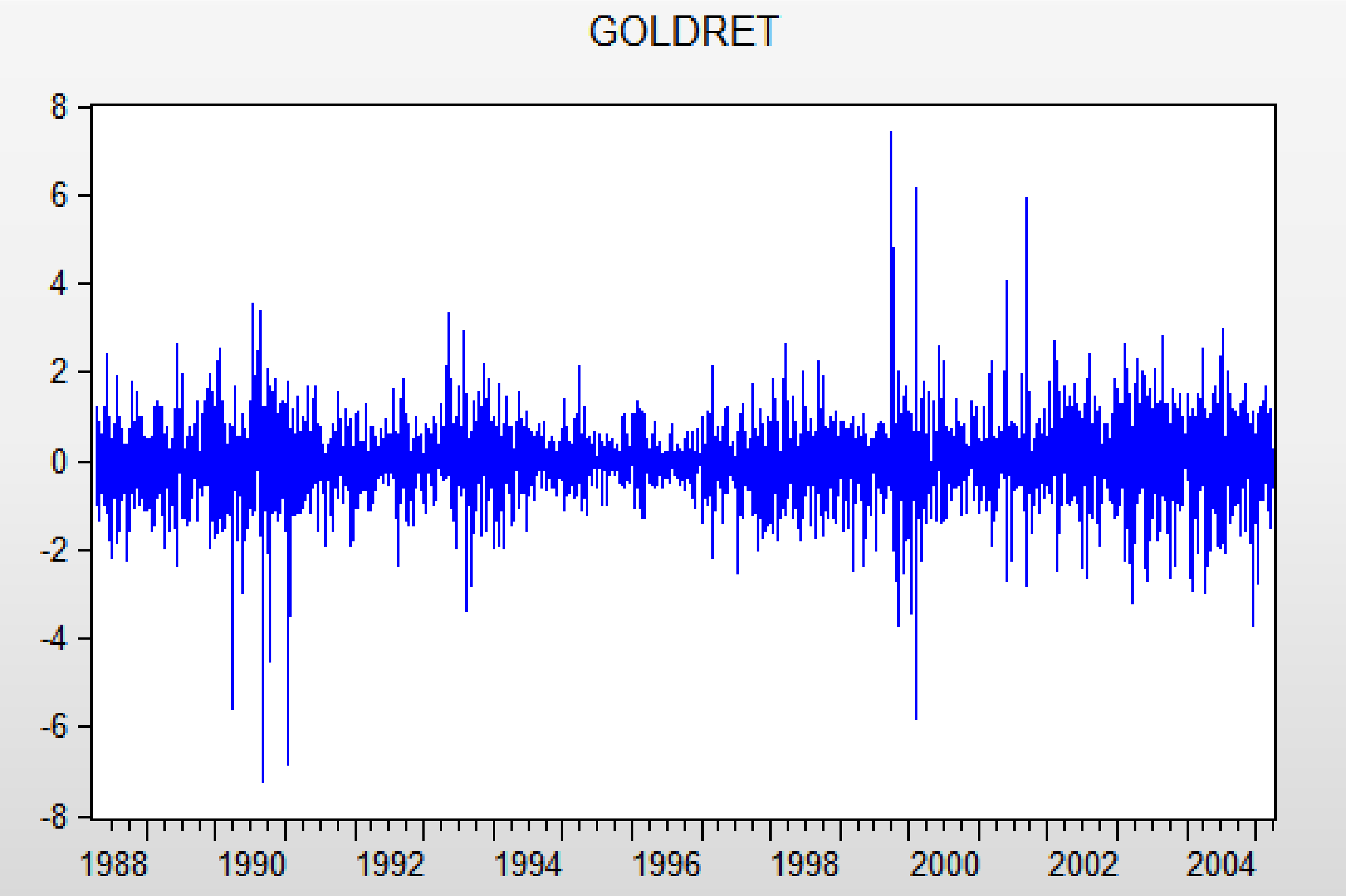
Gold bullion daily closing prices (US dollars per troy ounce) from 4 April 1988 to 5 April 2005



Xekalaki, E., Degiannakis, S. (2010) ARCH Models for Financial Applications, Wiley

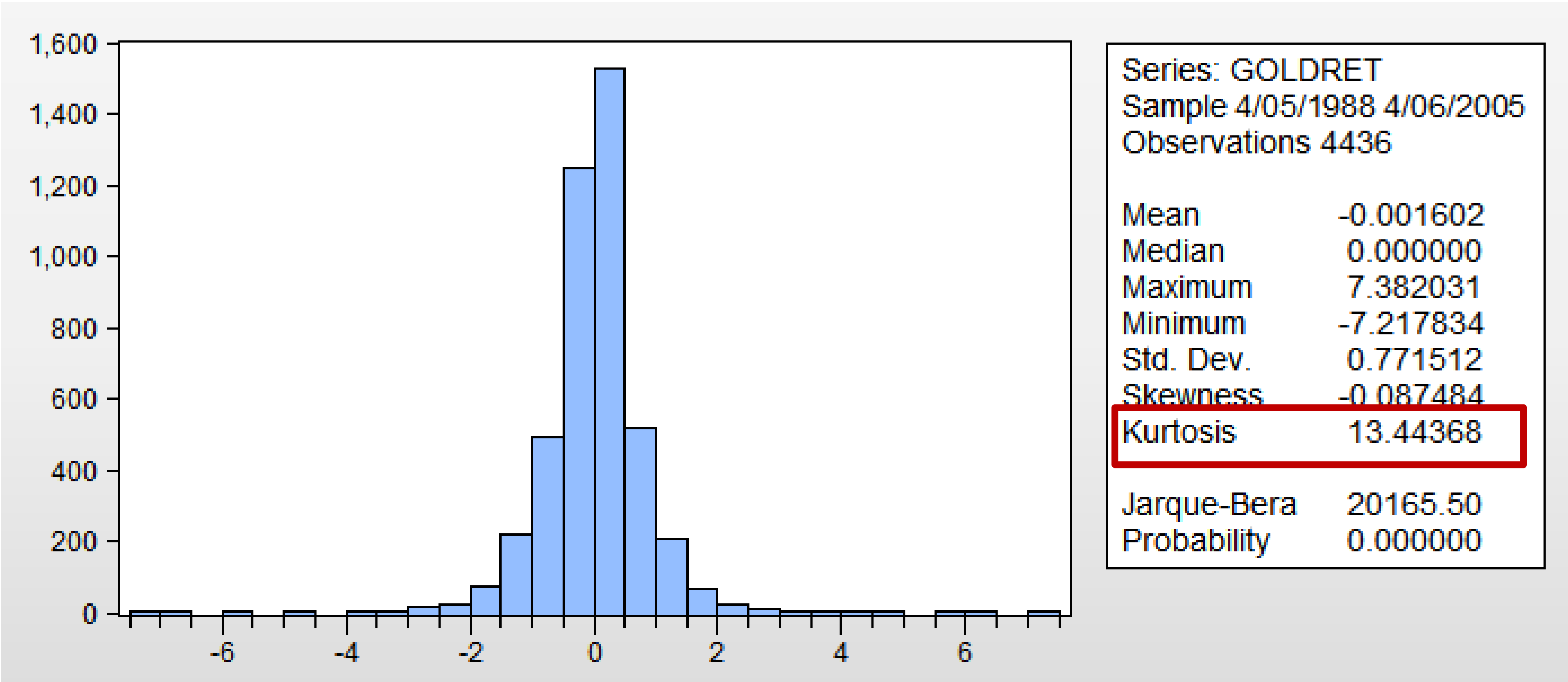
An Example: Price of Gold (cont'd)

Gold bullion continuously compounded daily returns



An Example: Price of Gold (cont'd)

Gold bullion continuously compounded daily returns, descriptive statistics



An Example: Price of Gold (cont'd)

They uesed following specification

$$y_t = c + \phi y_{t-1} + u_t$$

AR(1) model for conditional mean

$$u_t = v_t \sigma_t$$

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

GARCH(*p*,*q*) model for conditional variance

$$v_t \sim i.i.d\ N(0,1)$$

white noise

9 GARCH (p,q) models were analysed:
autoregressive orders $p = 0, 1, 2$
movig average orders $q = 1, 2, 3$

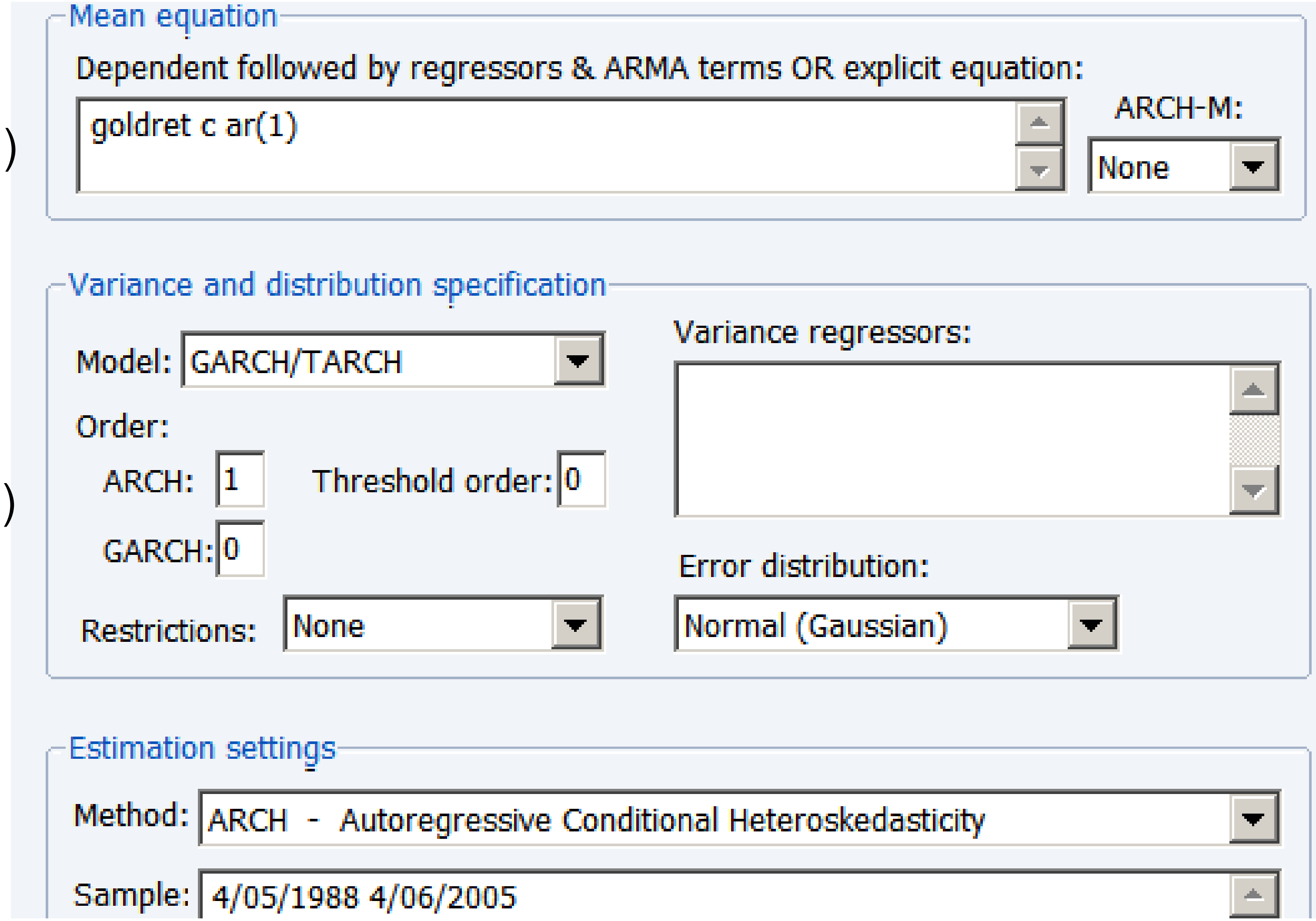
They applied Schwarz's criterion to select the model that fits the data best

An Example: Price of Gold (cont'd)

Quick -> Estimate Equation -> Method ARCH

AR(1)

GARCH (0,1)



An Example: Price of Gold (cont'd)

AR(1)

$$y_t = c_0 + c_1 y_{t-1} + u_t$$

GARCH (0,1)

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

Dependent Variable: GOLDRET
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 12/11/12 Time: 21:13
Sample (adjusted): 4/07/1988 4/06/2005
Included observations: 4435 after adjustments
Convergence achieved after 12 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(3) + C(4)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.002551	0.010526	-0.242329	0.8085
AR(1)	0.002478	0.018233	0.135909	0.8919
Variance Equation				
C	0.478507	0.005792	82.61536	0.0000
RESID(-1)^2	0.182573	0.010840	16.84201	0.0000

R-squared	0.000006	Mean dependent var	-0.001379
Adjusted R-squared	-0.000220	S.D. dependent var	0.771456
S.E. of regression	0.771540	Akaike info criterion	2.252077
Sum squared resid	2638.853	Schwarz criterion	2.257846
Log likelihood	-4989.980	Hannan-Quinn criter.	2.254111
Durbin-Watson stat	1.999444		

Inverted AR Roots	.00

An Example: Price of Gold (cont'd)

AR(1)

$$y_t = c_0 + c_1 y_{t-1} + u_t$$

GARCH (1,2)

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

Dependent Variable: GOLDRET
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 12/12/12 Time: 16:40
Sample (adjusted): 4/07/1988 4/06/2005
Included observations: 4435 after adjustments
Convergence achieved after 24 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*RESID(-2)^2 + C(6)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.010218	0.009080	-1.125381	0.2604
AR(1)	-0.013620	0.018462	-0.737743	0.4607
Variance Equation				
C	0.001566	0.000278	5.637386	0.0000
RESID(-1)^2	0.130440	0.009269	14.07245	0.0000
RESID(-2)^2	-0.095633	0.009139	-10.46459	0.0000
GARCH(-1)	0.964531	0.001837	525.1995	0.0000

R-squared	-0.000399	Mean dependent var	-0.001379
Adjusted R-squared	-0.000625	S.D. dependent var	0.771456
S.E. of regression	0.771697	Akaike info criterion	2.134172
Sum squared resid	2639.921	Schwarz criterion	2.142826
Log likelihood	-4726.525	Hannan-Quinn criter.	2.137223
Durbin-Watson stat	1.966558		

Inverted AR Roots	-.01

An Example: Price of Gold (cont'd)

	Schwarz's information criterion	
ARCH(1)	2,257846	
ARCH(2)	2,25448	
ARCH(3)	2,233224	
GARCH(1,1)	2,150997	
GARCH(1,2)	2,142827	The best model
GARCH(1,3)	2,14469	
GARCH(2,1)	2,144333	
GARCH(2,2)	2,143513	
GARCH(2,3)	2,144932	

An Example: Price of Gold (cont'd)

GARCH(1,2)

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + $$
$$+\beta_1 \sigma_{t-1}^2$$

Variance Equation				
C	0.001566	0.000278	5.637386	0.0000
RESID(-1)^2	0.130440	0.009269	14.07245	0.0000
RESID(-2)^2	-0.095633	0.009139	-10.46459	0.0000
GARCH(-1)	0.964531	0.001837	525.1995	0.0000

$$\hat{\alpha}_1 = 0{,}1304 \quad \hat{\alpha}_2 = -0{,}0956$$

$$\hat{\beta}_1 = 0{,}9645$$

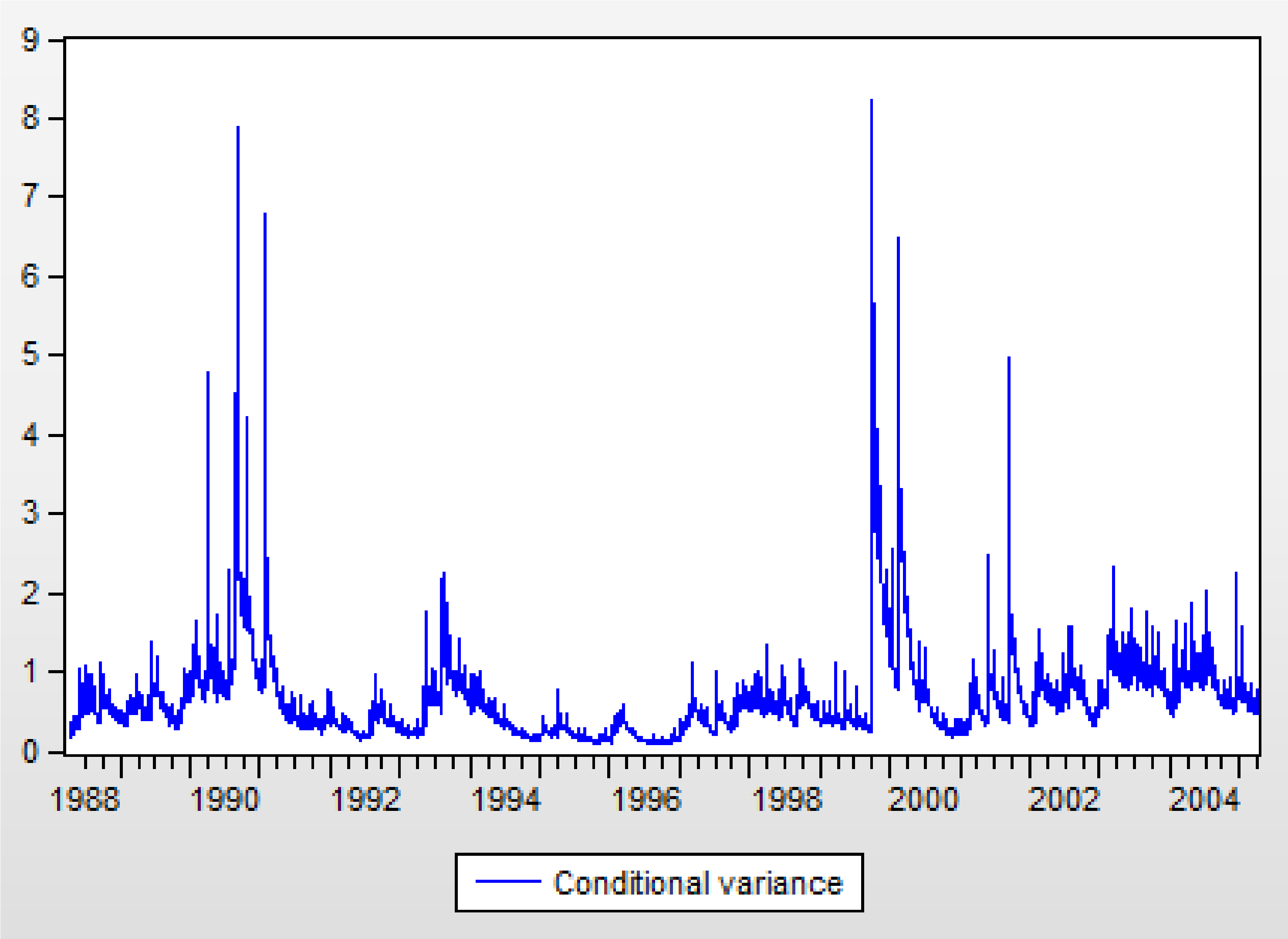
$$\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\beta}_2 = 0{,}9993$$

Examination of many financial time series suggests that,

- parameters α are small,
- the sum of all parameters *β* is near the value 1
- the sum of all parameters ≈ 1

An Example: Price of Gold (cont'd)

The graph of conditional variance



Residual Diagnostics after the Estimation

- Correlogram-Q-statistics

  Displays the correlogram (autocorrelations and partial autocorrelations) of the standardized residuals. This can be

  used to test for remaining serial correlation in the mean equation and to check the specification of the mean equation. If the mean equation is correctly specified, all Q -statistics should not be significant.

- Correlogram Squared Residuals

  Displays the correlogram of the squared standardized residuals.

  This view can be used to test for remaining ARCH in the variance equation and to check the specification of the variance equation. If the variance equation is correctly specified, all Q-statistics should not be significant,

Residual Diagnostics after the Estimation (cont'd)

- Histogram–Normality Test

Displays descriptive statistics and a histogram of the standardized residuals. You can use the Jarque-Bera statistic to test the null of whether the standardized residuals are normally distributed. If the standardized residuals are normally distributed, the Jarque-Bera statistic should not be significant.

- ARCH LM Test

Carries out Lagrange multiplier tests to test whether the standardized residuals exhibit additional ARCH. If the variance equation is correctly specified, there should be no ARCH left in the standardized residuals.

An Example: Correlogram of Standardized Residuals

Sample: 4/07/1988 4/06/2005
Included observations: 4435
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.027	0.027	3.2036	
		2	0.026	0.025	6.2471	0.012
		3	0.005	0.004	6.3534	0.042
		4	0.011	0.010	6.8537	0.077
		5	0.007	0.007	7.0911	0.131
		6	-0.007	-0.008	7.2967	0.199
		7	-0.012	-0.012	7.9056	0.245
		8	0.009	0.009	8.2337	0.312
		9	0.017	0.017	9.4790	0.304
		10	-0.000	-0.002	9.4798	0.394
		11	-0.008	-0.009	9.7849	0.460
		12	-0.019	-0.019	11.451	0.406
		13	-0.021	-0.020	13.431	0.338
		14	-0.004	-0.002	13.510	0.409
		15	-0.007	-0.005	13.700	0.472

White noise. The mean equation is correctly specified

An Example: Correlogram of Standardized Residuals Squared

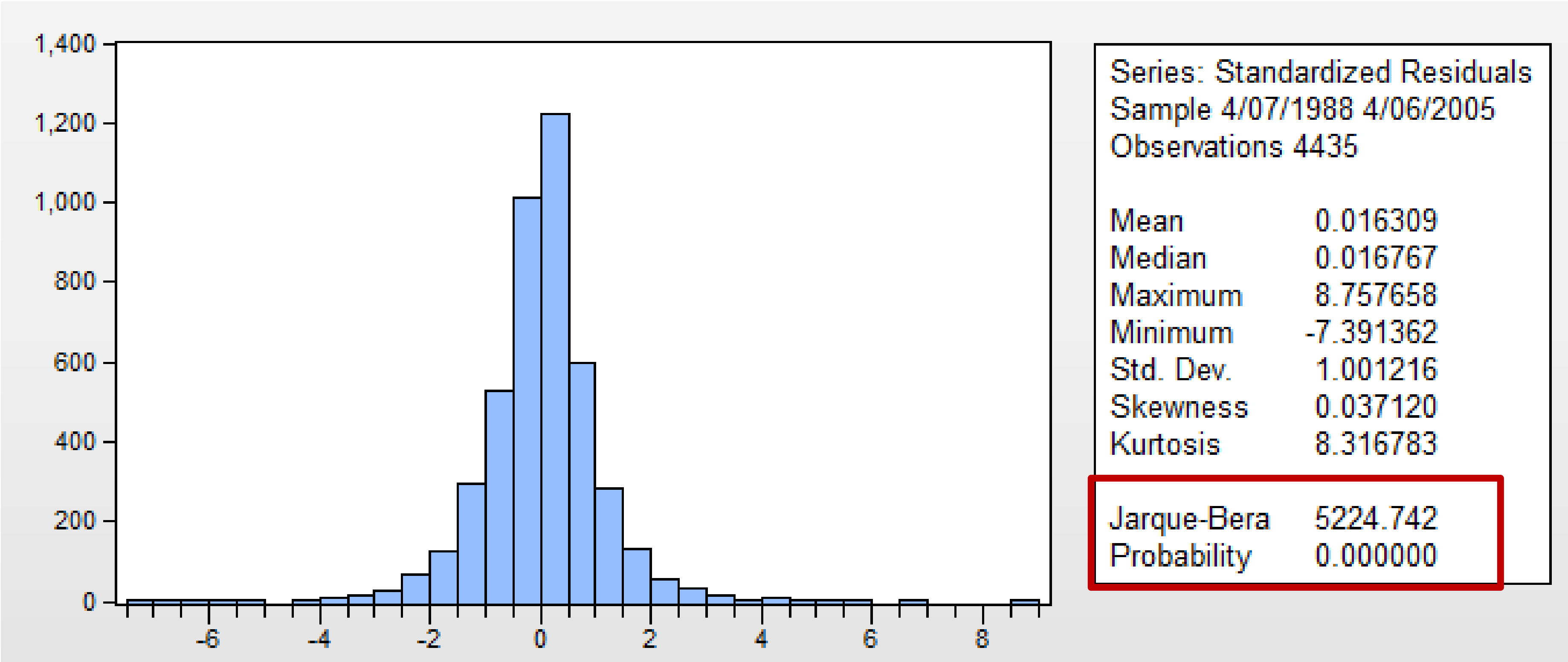
Correlogram of Standardized Residuals Squared

Date: 12/12/12 Time: 16:43
Sample: 4/07/1988 4/06/2005
Included observations: 4435
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.007	0.007	0.1988	
		2	-0.007	-0.007	0.4021	0.526
		3	0.031	0.031	4.6591	0.097
		4	0.012	0.012	5.3162	0.150
		5	0.005	0.005	5.4257	0.246
		6	-0.010	-0.011	5.8935	0.317
		7	0.000	-0.000	5.8937	0.435
		8	-0.003	-0.003	5.9304	0.548
		9	-0.011	-0.010	6.4703	0.595
		10	-0.005	-0.005	6.5986	0.679
		11	0.000	0.001	6.5995	0.763
		12	-0.006	-0.006	6.7804	0.817
		13	0.004	0.005	6.8610	0.867
		14	-0.005	-0.005	6.9755	0.903
		15	-0.003	-0.003	7.0137	0.934

White noise. The variance equation is correctly specified

An Example: Normality of Standardized Residuals



There is no normal distribution

An Example: ARCH LM Test

Heteroskedasticity Test: ARCH

F-statistic	1.080507	Prob. F(5,4424)	0.3689
Obs*R-squared	5.403264	Prob. Chi-Square(5)	0.3687

Test Equation:
Dependent Variable: WGT_RESID^2
Method: Least Squares
Date: 12/06/16 Time: 14:51
Sample (adjusted): 4/14/1988 4/06/2005
Included observations: 4430 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.955217	0.052561	18.17351	0.0000
WGT_RESID^2(-1)	0.006527	0.015035	0.434151	0.6642
WGT_RESID^2(-2)	-0.007098	0.015034	-0.472137	0.6369
WGT_RESID^2(-3)	0.031009	0.015027	2.063557	0.0391
WGT_RESID^2(-4)	0.011673	0.015034	0.776415	0.4375
WGT_RESID^2(-5)	0.005240	0.015035	0.348487	0.7275

R-squared	0.001220	Mean dependent var	1.002718
Adjusted R-squared	0.000091	S.D. dependent var	2.713374
S.E. of regression	2.713251	Akaike info criterion	4.835525
Sum squared resid	32568.29	Schwarz criterion	4.844188
Log likelihood	-10704.69	Hannan-Quinn criter.	4.838580
F-statistic	1.080507	Durbin-Watson stat	1.999860
Prob(F-statistic)	0.368916		

There is no ARCH effect
The variance equation is correctly specified

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.
Consider the following GARCH(1,1) model:

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

What is needed is to generate forecasts of

$$\sigma_{T+1}^2 \mid \Omega_T$$
$$\sigma_{T+2}^2 \mid \Omega_T$$
$$\ldots$$
$$\sigma_{T+s}^2 \mid \Omega_T$$

where Ω_T denotes all information available up to and including observation T

Forecasting Variances using GARCH Models (cont'd)

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

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

$$\sigma_{T+1}^2 = \alpha_0 + \alpha_1 u_T^2 + \beta\sigma_T^2$$

$$\sigma_{T+2}^2 = \alpha_0 + \alpha_1 u_{T+1}^2 + \beta\sigma_{T+1}^2$$

$$\sigma_{T+3}^2 = \alpha_0 + \alpha_1 u_{T+2}^2 + \beta\sigma_{T+2}^2$$

The one step ahead forecast for σ^2 is easy to calculate since, at time *T*, the values of all the terms on the RHS are known.

$$\sigma_{1,T}^{f\,2} = \sigma_{T+1}^2 = \alpha_0 + \alpha_1 u_T^2 + \beta\sigma_T^2$$ The one step ahead forecast

Forecasting Variances using GARCH Models (cont'd)

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

How the 2-step ahead forecast for σ^2 calculated?

$$\sigma_{2,T}^{f\,2} = \alpha_0 + \alpha_1 E[u_{T+1}^2 \,|\, \Omega_T] + \beta \sigma_{1,T}^{f\,2}$$

$E[u_{T+1}^2 \,|\, \Omega_T]$ The conditional expectation at time *T*

$$E[u_{T+1}^2 \,|\, \Omega_T] = \sigma_{T+1}^2$$

$$\downarrow$$

$$E[u_{T+1}^2 \,|\, \Omega_T] = \sigma_{1,T}^{f\,2}$$

$$\sigma_{2,T}^{f\,2} = \alpha_0 + \alpha_1 \sigma_{1,T}^{f\,2} + \beta \sigma_{1,T}^{f\,2} = \alpha_0 + \sigma_{1,T}^{f\,2}(\alpha_1 + \beta)$$

Forecasting Variances using GARCH Models (cont'd)

By similar arguments, the 3-step ahead forecast will be given by

$$\sigma_{3,T}^{f\,2} = \alpha_0 + \alpha_1 E[u_{T+2}^2 \mid \Omega_T] + \beta\sigma_{2,T}^{f\,2} =$$

$$= \alpha_0 + \alpha_1\sigma_{2,T}^{f\,2} + \beta\sigma_{2,T}^{f\,2} = \alpha_0 + \sigma_{2,T}^{f\,2}(\alpha_1 + \beta) =$$

$$= \alpha_0 + (\alpha_1 + \beta)\left[\alpha_0 + (\alpha_1 + \beta)\sigma_{1,T}^{f\,2}\right] =$$

$$= \alpha_0 + \alpha_0(\alpha_1 + \beta) + (\alpha_1 + \beta)^2 \sigma_{1,T}^{f\,2}$$

Any *s*-step ahead forecast (s>1) would be produced by

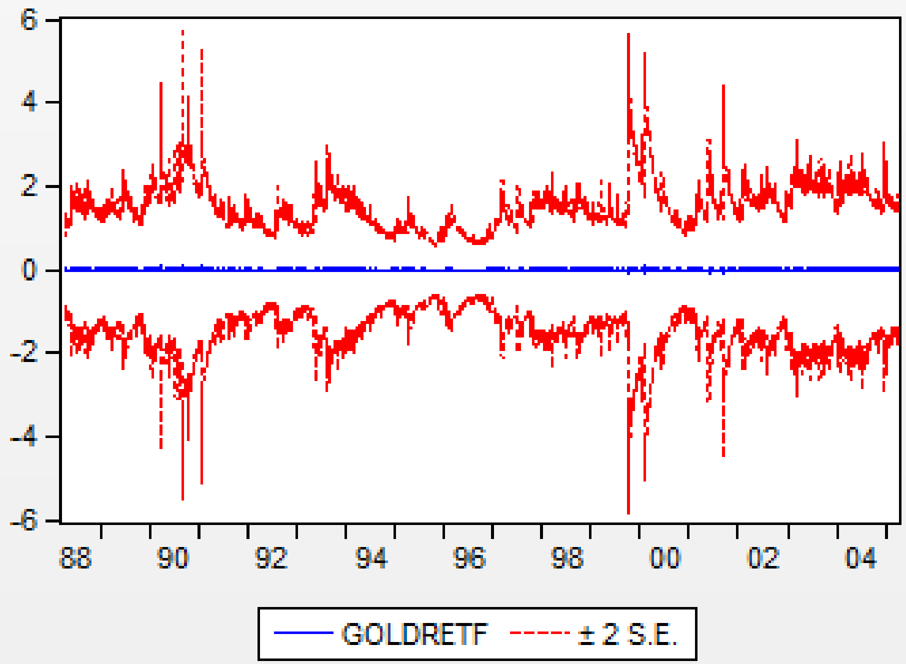
$$\sigma_{s,T}^{f\,2} = \alpha_0 \sum_{i=1}^{s-1} (\alpha_1 + \beta)^{i-1} + (\alpha_1 + \beta)^{s-1} \sigma_{1,T}^{f\,2}$$

An Example: Price of Gold (cont'd)

Static forecast (in-sample) is one-step forecast

mean

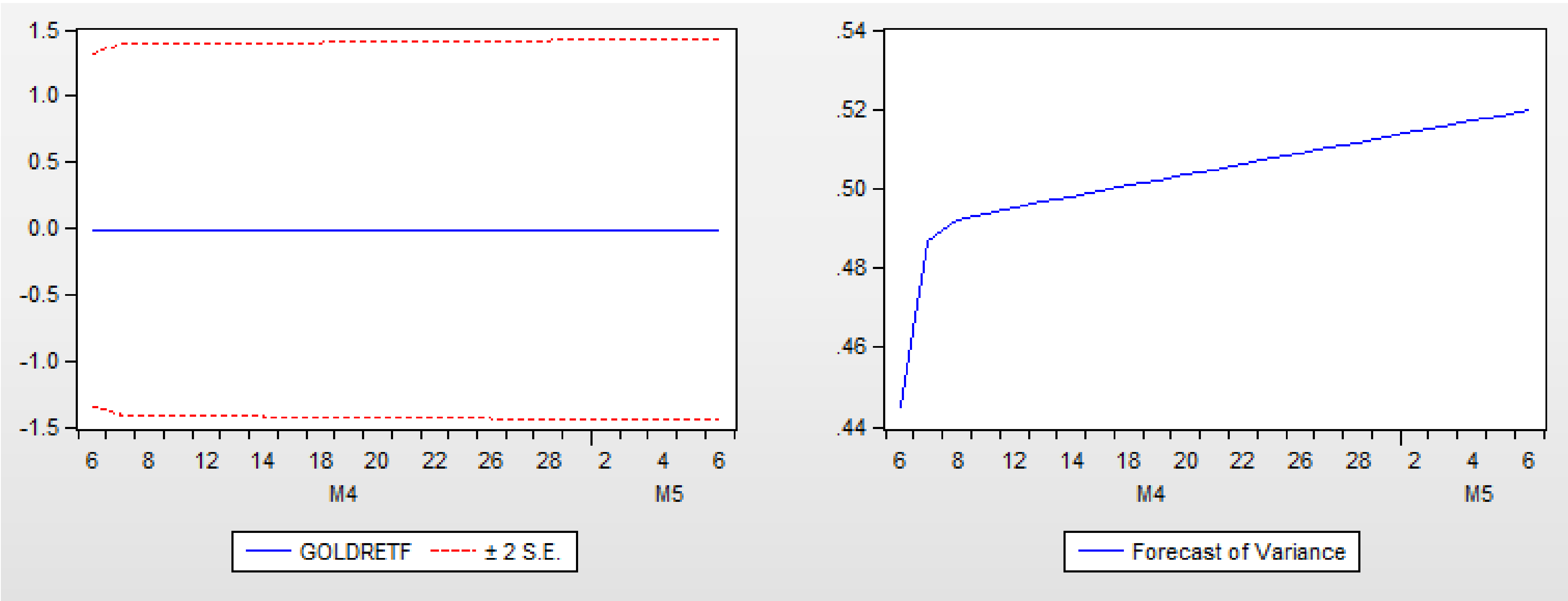
double standard deviation



Forecast: GOLDRETF	
Actual: GOLDRET	
Forecast sample: 4/05/1988 4/06/2005	
Adjusted sample: 4/07/1988 4/06/2005	
Included observations: 4435	
Root Mean Squared Error	0.771523
Mean Absolute Error	0.515789
Mean Abs. Percent Error	93.52696
Theil Inequality Coefficient	0.981444
Bias Proportion	0.000135
Variance Proportion	0.972553
Covariance Proportion	0.027312

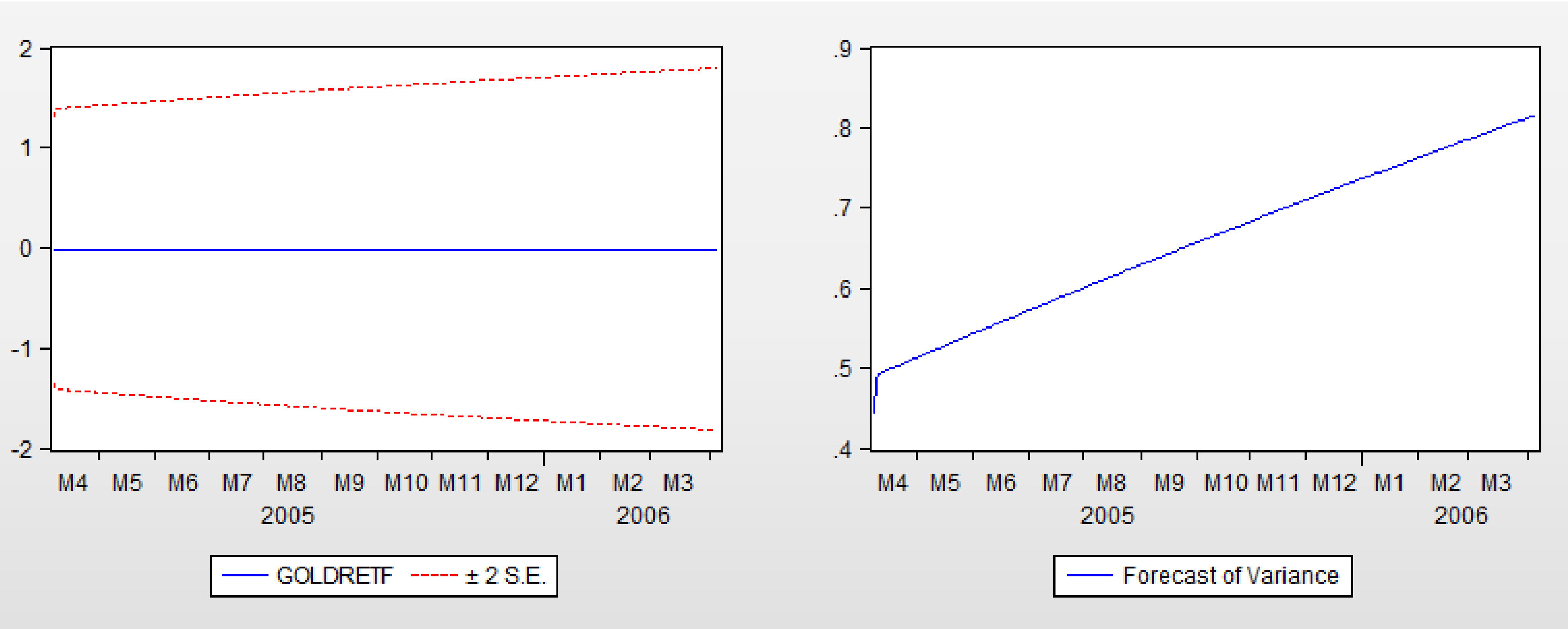
An Example: Price of Gold (cont'd)

Dynamic forecast (out-of-sample)
1 month ahead



An Example: Price of Gold (cont'd)

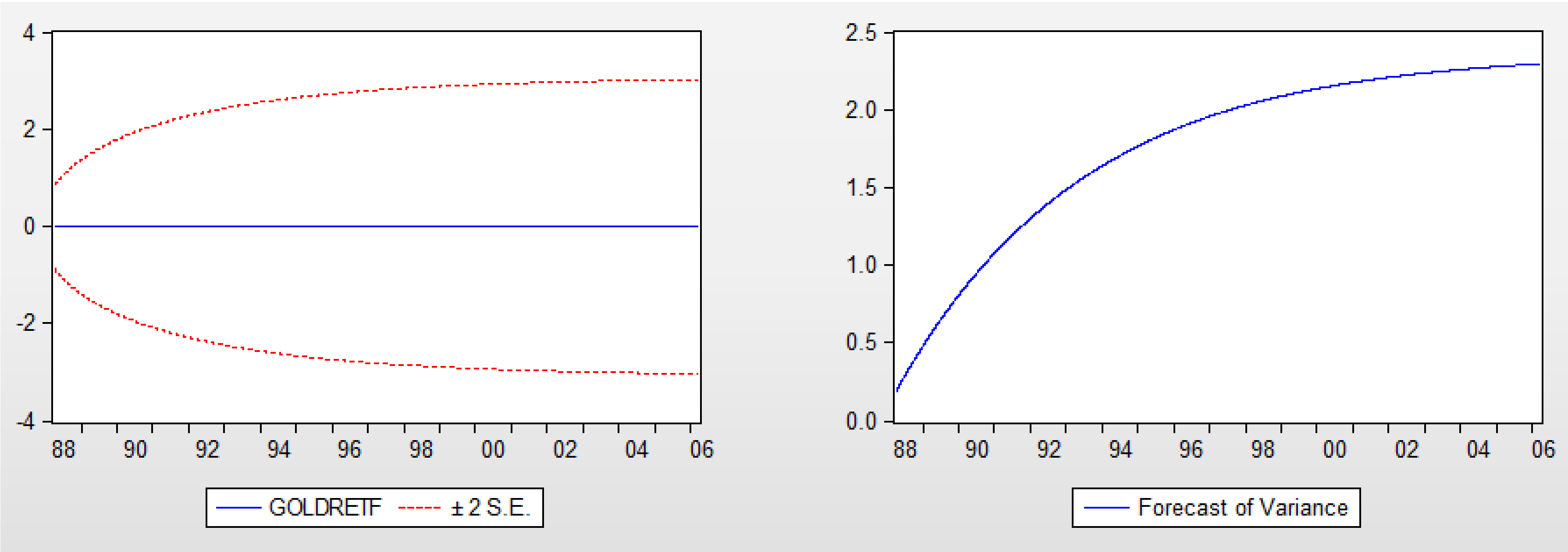
Dynamic forecast (out-of-sample)
1 year ahead



An Example: Price of Gold (cont'd)

Dynamic forecast (out-of-sample)
18 years ahead

1988 - 2006



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:
  - Likelihood Ratio test,
  - Wald test,
  - Lagrange Multiplier test.

Likelihood Ratio Tests

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

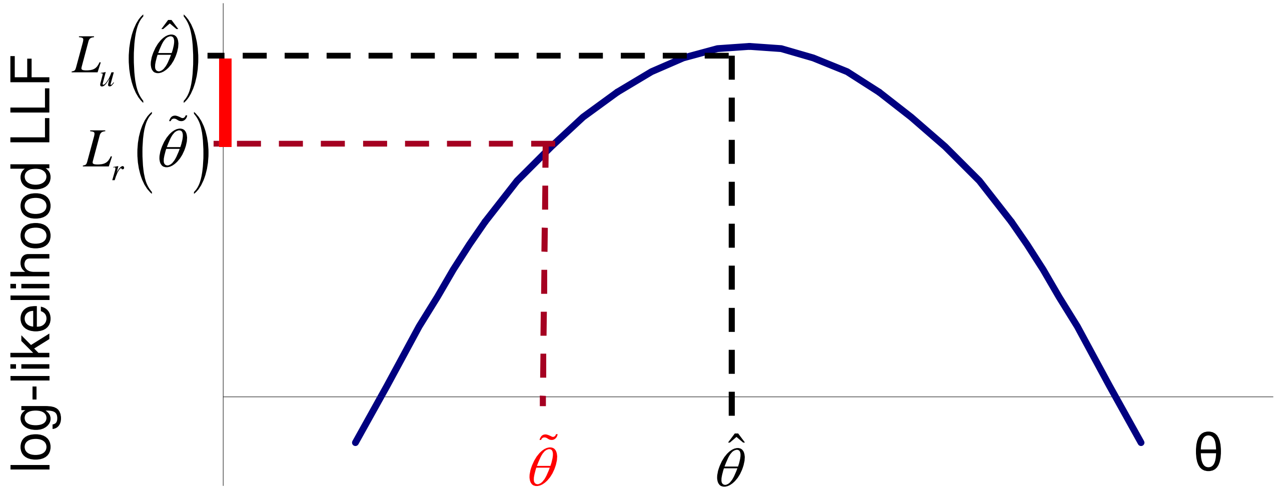
- So we estimate the unconstrained model and achieve a given maximised value of the LLF, denoted L_u
- Then estimate the model imposing the constraint(s) and get a new value of the LLF denoted L_r.
- Which will be bigger?

$$L_u \geq L_r$$

- The LR test statistic is given by

$$LR = 2(L_u - L_r)$$

If you reject H_0: the constraints are not valid



Likelihood Ratio Tests (cont'd)

Example: We estimate a GARCH model and obtain a maximised LLF of 66.85. We are interested in testing whether *β*= 0 in the following equation.

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

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

We estimate the model imposing the restriction.

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

The maximised LLF falls to 64.54. Can we accept the restriction?

$$LR = 2 \cdot (66.85 - 64.54) = 4.62$$

The test follows a $\chi^2(1) = 3.84$ at 5%, so reject the null.

Extensions to the Basic GARCH Model

- Since the GARCH model was developed, a huge number of extensions and variants have been proposed. The most important examples are
  - EGARCH (exponential GARCH)
  - GJR or TARCH (threshold GARCH)
  - GARCH-M
- Problems with GARCH(p,q) Models:
  - Non-negativity constraints may still be violated
  - GARCH models cannot account for leverage effects
- Possible solutions: the exponential GARCH (EGARCH) model or the GJR model, which are asymmetric GARCH models.

The EGARCH Model

Suggested by Nelson (1991): exponential GARCH model EGARCH.
The variance equation is given by

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

Asymmetric term

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 γ is negative, the relationship between volatility and y is negative

The EGARCH Model (cont'd)

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

If "bad news" $u_{t-1} < 0$ and $\gamma < 0$

$y_{t-1} = bx_{t-1} + u_{t-1} \quad \downarrow$ $\qquad \ln\left(\sigma_t^2\right) \quad \uparrow$

Testing asymmetry: look at statistical significance of parameter γ

$$H_0 : \gamma = 0;$$

$$H_1 : \gamma \neq 0$$

An Example: Exhange Rate Japan Yen to USD

Dependent Variable: RJPY
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 12/03/12 Time: 13:21
Sample (adjusted): 7/08/2002 7/07/2007
Included observations: 1826 after adjustments
Convergence achieved after 12 iterations
Presample variance: backcast (parameter = 0.7)
LOG(GARCH) = C(2) + C(3)*ABS(RESID(-1)/@SQRT(GARCH(-1))) + C(4)*RESID(-1)/@SQRT(GARCH(-1)) + C(5)*LOG(GARCH(-1))

	Variable	Coefficient	Std. Error	z-Statistic	Prob.
	C	0.003756	0.010025	0.374722	0.7079
		Variance Equation			
ω	C(2)	-1.262782	0.194243	-6.501047	0.0000
α	C(3)	0.214215	0.034226	6.258919	0.0000
γ	C(4)	-0.046461	0.024983	-1.859751	0.0629
β	C(5)	0.329164	0.112572	2.924037	0.0035

R-squared	-0.000031	Mean dependent var	0.001328
Adjusted R-squared	-0.000031	S.D. dependent var	0.439632
S.E. of regression	0.439639	Akaike info criterion	1.183216
Sum squared resid	352.7398	Schwarz criterion	1.198303
Log likelihood	-1075.276	Hannan-Quinn criter.	1.188781
Durbin-Watson stat	1.981879		

EGARCH

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

No asymmetry

Threshold GARCH or TARCH Model

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{j=1}^{p} \alpha_j u_{t-j}^2 + \sum_{j=1}^{r} \gamma_j u_{t-j}^2 I_{t-j}^-$$

Dummy variable for asymmetry

$$I_t^- = \begin{cases} 1, & u_t < 0 \\ 0, & u_t > 0 \end{cases}$$

"bad news", shock is negative

"hgood news", shock is positive

Innovations have an asymmetric impact on volatility:

$u_{t-j} < 0$, then $\alpha_j + \gamma_j$

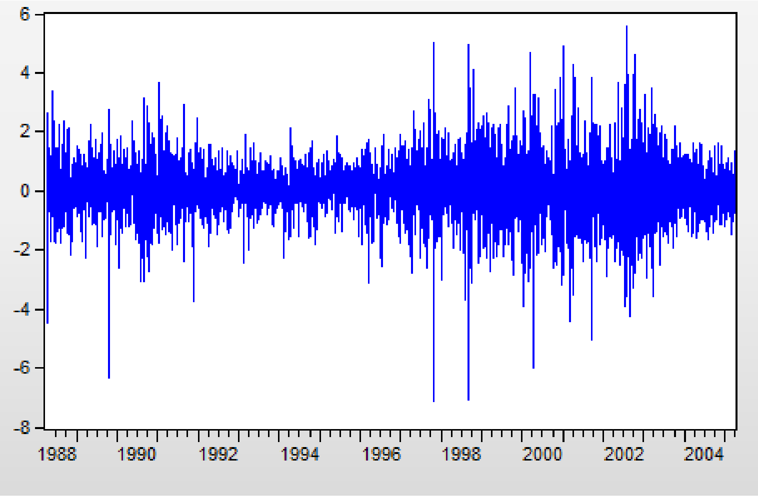
$u_{t-j} > 0$, then α_j

Testing asymmetry: $H_0 : \gamma_i = 0; \quad H_1 : \gamma_i \neq 0$

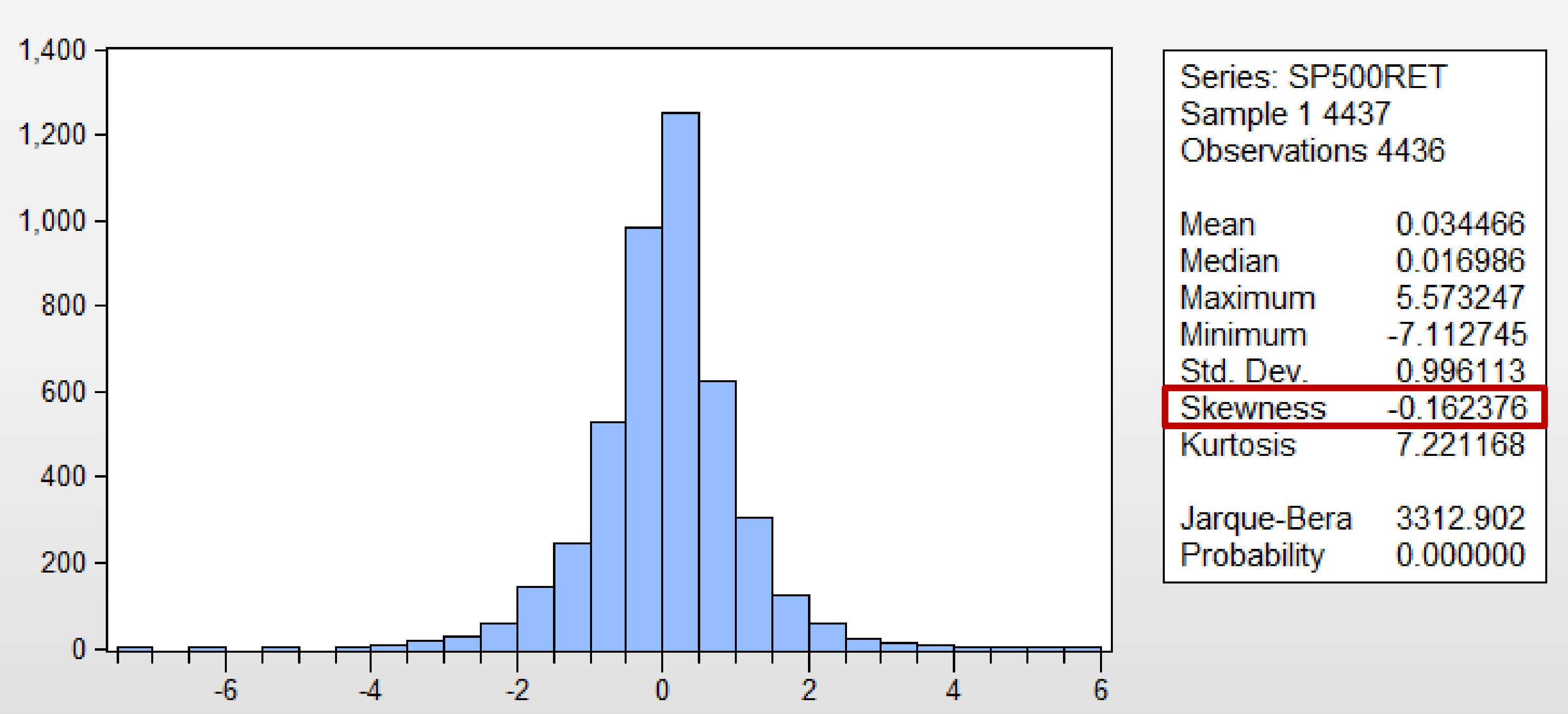
If γ_j >0, then the negative shock increases volatility, This is a leverage effect.

Also GJR model (Brooks 8.12)

An Example: S&P500



Log return



An Example: S&P500 (cont'd)

TARCH(1,3)

$$\sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \gamma u_{t-1}^2 I_{t-1}^- + \alpha_2 u_{t-2}^2 + \alpha_3 u_{t-3}^2 + \beta_1 \sigma_{t-1}^2$$

γ → RESID(-1)^2*(RESID(-1)<0)

Is statistically significant, asymmetry exists, negative shocks or "bad news" increase volatility.

GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*RESID(-1)^2*(RESID(-1)<0) + C(6)*RESID(-2)^2 + C(7)*RESID(-3)^2 + C(8)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.026507	0.012351	2.146032	0.0319
AR(1)	0.018256	0.014977	1.218932	0.2229
Variance Equation				
C	0.012764	0.002758	4.627366	0.0000
RESID(-1)^2	-0.039896	0.016672	-2.393040	0.0167
RESID(-1)^2*(RESID(-1)<0)	0.099200	0.014188	6.991801	0.0000
RESID(-2)^2	0.037327	0.027537	1.355518	0.1753
RESID(-3)^2	0.008303	0.021880	0.379467	0.7043
GARCH(-1)	0.930797	0.008931	104.2215	0.0000

R-squared	-0.000599	Mean dependent var	0.034262
Adjusted R-squared	-0.000825	S.D. dependent var	0.996133
S.E. of regression	0.996543	Akaike info criterion	2.585098
Sum squared resid	4402.407	Schwarz criterion	2.596637
Log likelihood	-5724.454	Hannan-Quinn criter.	2.589167
Durbin-Watson stat	2.046566		

Inverted AR Roots	.02

GARCH Model with Exogenous Variables

Engle ja Patton (2001) used GARCH(1,1)-X model on Dow Jones Industrial Index. Data Aug 23, 1988 – Aug 22, 2000, 3131 observations

They used the lagged level of the three month US Treasury bill rate as exegenous regressor x

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

Table 5. Results from the GARCH(1,1)-X model.

	Coefficient	Robust standard error
Constant	0.0608	0.0145
ω	−0.0010	0.0016
α	0.0464	0.0040
β	0.9350	0.0065
φ	0.0031	0.0005

Engle, R. & Patton, A. (2001), What good is a volatility model? *Quantitative Finance*, Vol. 1 , 237-245

GARCH in Mean (GARCH-M)

- Many classic areas of finance suggest that the mean of a relationship will be affected by the volatility or uncertainty of a series.
- Engle, Lilien and Robins(1987) allow for this explicitly using an ARCH framework, the ARCH-M specification
- Typically either the variance or the standard deviation are included in the mean relationship.

$$y_t = bx_t + \delta\sigma_t^2 + u_t, \qquad u_t \sim N(0, \sigma_t)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \qquad \text{GARCH(p,q)}$$

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.