Empirical Finance
Lecture 6: Models of time varying volatility in empirical finance: ARCH and GARCH models.

Module Leader: Dr Stuart Fraser

stuart.fraser@wbs.ac.uk

Room D1.18 (Social Studies)

Introduction

Generalized-Autoregressive-Conditional-Heteroscedastic (GARCH) processes:

- Motivation for GARCH models.
- Examples of different types of GARCH model.
- Some applications of GARCH in finance.
- Identifying, estimating and testing GARCH models.

⇒Seminar 5: Modelling time-varying volatility in the FTSE All-Share Index excess returns.

Motivation for GARCH processes

Two empirical features of financial return series are relevant in this context.

Firstly, there is the observation that over short horizons/holding periods of up to one month:

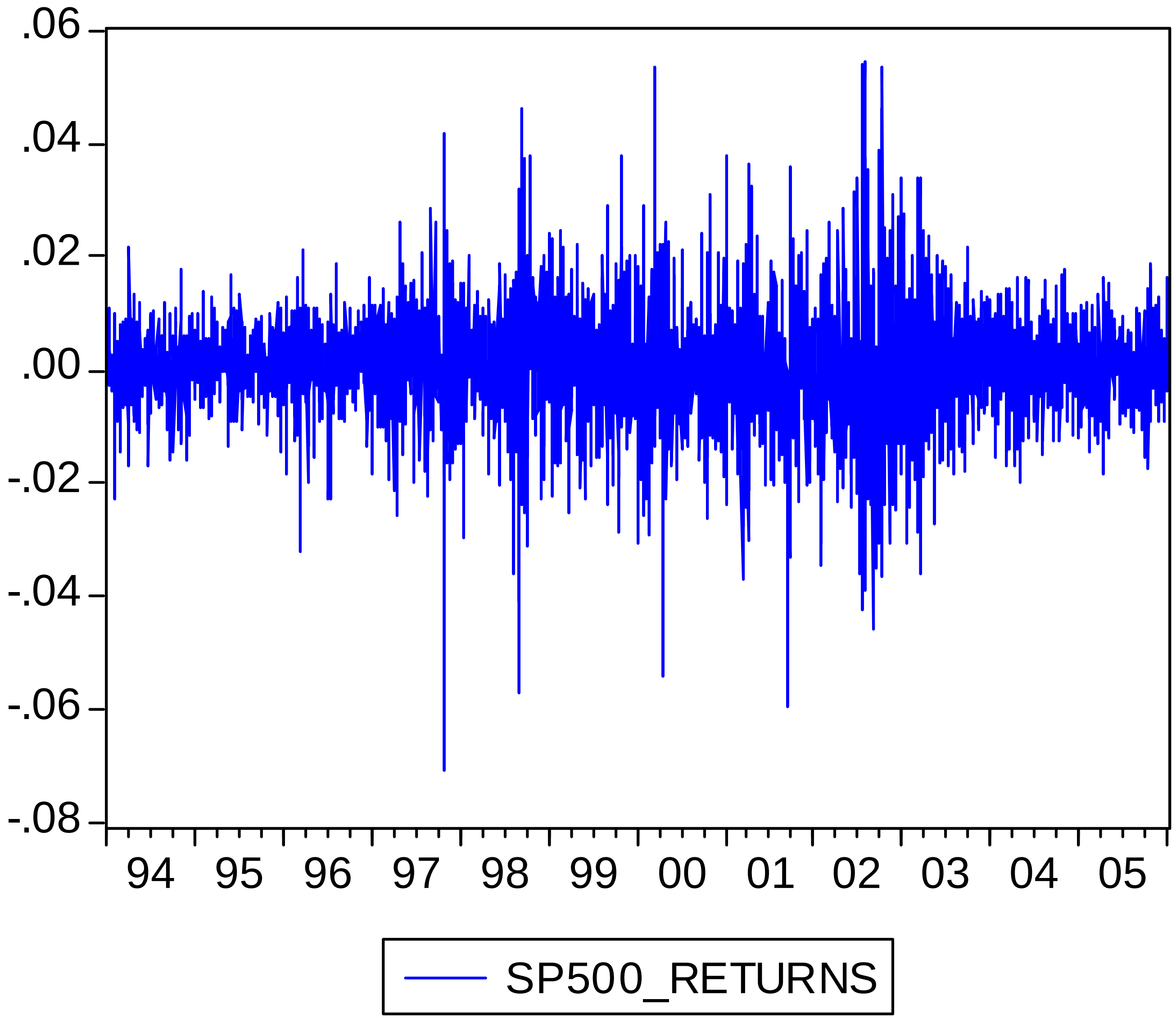
- Return volatility is time varying (there are periods of tranquility and turbulence).
- Volatility clustering: large (small) price changes tend to be followed by further large (small) changes.

  ⇒volatility (risk) is positively correlated.

  ⇒Non-linear dependence in returns.

Secondly, the unconditional distributions of short-horizon returns have fat tails (leptokurtosis).

1. Volatility clustering (see handout for Seminars 1 and 2)

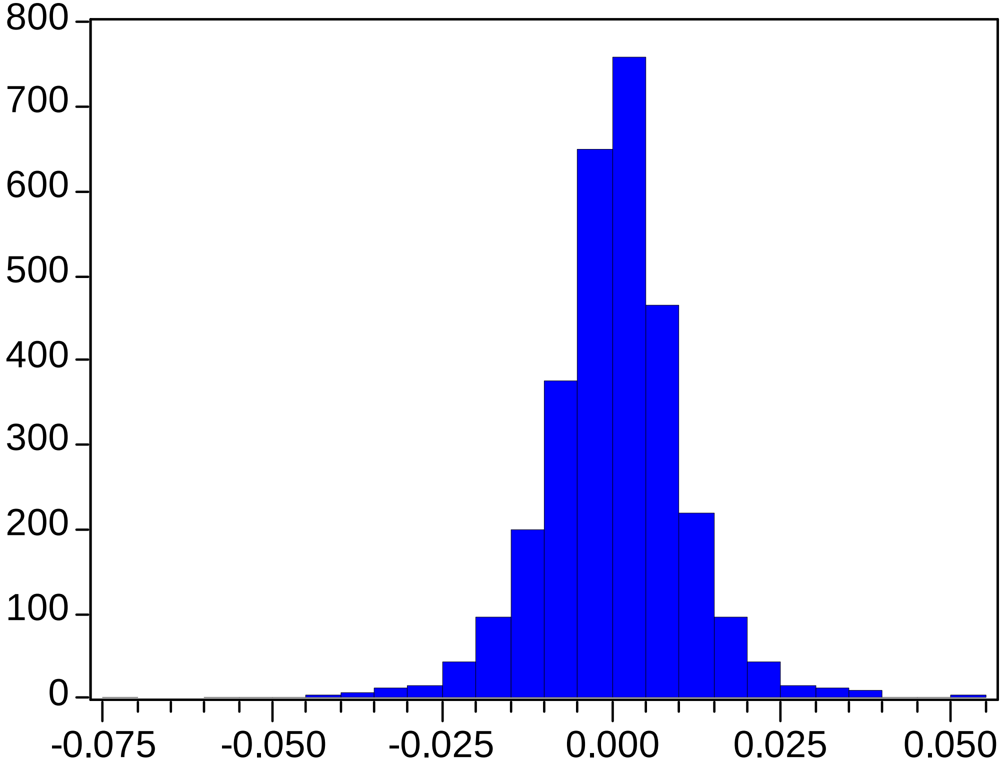


⇒Non –linear dependence in returns

Correlogram of <u>squared</u> SP500 returns

Autocorrel	Partial Correlation		AC	PAC	Q-Stat	Prob
|**	|**	1	0.207	0.207	130.55	0
|**	|*	2	0.2	0.164	252.34	0
|**	|*	3	0.228	0.171	410.19	0
|*	|*	4	0.18	0.093	508.64	0
|**	|*	5	0.205	0.115	636.52	0
|*	|	6	0.155	0.044	709.21	0
|*	|*	7	0.174	0.072	801.5	0
|*	|	8	0.152	0.039	871.82	0
|*	|	9	0.156	0.051	946.1	0
|*	|	10	0.16	0.05	1024.4	0

2. Leptokurtic unconditional return distributions



Series: SP500_RETURNS
Sample 1/03/1994 1/11/2006
Observations 3033

Mean	0.000412
Median	0.000645
Maximum	0.054248
Minimum	-0.070376
Std. Dev.	0.010419
Skewness	-0.138163
Kurtosis	6.818984
Jarque-Bera	1852.783
Probability	0.000000

Volatility clustering

Volatility clustering implies volatility, h(r), is predictable from past information:

$$h_t = f\left(\Omega_{t-1}\right)$$

Ω includes past volatility and other relevant information.

h is the conditional volatility. This measures ex-ante volatility.

A common measure of ex-ante volatility is the conditional variance:

$$h_t = E\left(\left(r_t - \mu\right)^2 \middle| \Omega_{t-1}\right)$$

This is a one-step forecast of the variance conditional on Ω_{t-1}

The ex-post (realized) volatility is:

$$\left(r_t - \mu\right)^2 \cong r_t^2$$

$\mu \cong 0$ for short horizon returns

Alternative volatility/risk measures include:

$\left\|r_t\right\|$	Absolute returns
$\left\|r_t\right\|^{\theta},\ \theta > 0,$	Power returns
$\ln P_t^{high} - \ln P_t^{low}$	Range estimator

General framework for modelling expected returns and time varying volatility

The following general framework sets out a conditional mean equation (to predict expected returns) and a conditional variance equation (to predict risk):

$$r_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = v_t \sigma_t$$

$$\sigma_t^2 = E\left(\varepsilon_t^2 \middle| \Omega_{t-1}\right)$$

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

Mean equation describing equilibrium returns (μ could include the conditional variance to model a time-varying risk premium – see GARCH-M below).

The conditional variance depends on information from previous periods (⇒volatility clustering).

The assumption that the standardized residuals (v) are Gaussian ⇒ the residuals (ε) are conditionally normally distributed.

This assumption is important for estimation of the model (see below).

$$E\left(\varepsilon_t \middle| \Omega_{t-1}\right) = \sigma_t E\left(v_t \middle| \Omega_{t-1}\right) = 0$$

$$\Rightarrow \varepsilon_t \middle| \Omega_{t-1} \sim N\left(0, \sigma_t^2\right)$$

ARCH(q) model (Autoregressive Conditional Heteroscedasticity)

Different assumptions about Ω generate different models of time varying risk.

Suppose $\Omega_{t-1} = \left\{\varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2\right\}$ then

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

The ARCH model was first proposed by Engle (1982) in the context of modeling inflation uncertainty.

This is an ARCH(q) model. A sufficient condition for a positive conditional variance (variances cannot be negative) is that:

$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad i = 1, \ldots, q$$

Engle shared the Nobel Prize in Economics (2003):

"for methods of analyzing economic time series with time-varying volatility (ARCH)"

ARCH(q) model

The ARCH(q) model can be written as an AR(q) model in the squared residuals

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 + u_t,$$
$$u_t = \varepsilon_t^2 - E\left(\varepsilon_t^2 \middle| \Omega_{t-1}\right) = \varepsilon_t^2 - \sigma_t^2$$

Shocks to volatility (u)

The process is stationary if all the roots of the characteristic equation lie outside of the unit circle:

$$1 - \alpha_1 z - \ldots - a_q z^q = 0$$

If the process is stationary then shocks to volatility do not persist ⇒the conditional variance returns eventually to it's long run level.

The long run/unconditional variance can be found from the Wold representation:

$$\varepsilon_t^2 = \frac{\alpha_0 + u_t}{1 - \alpha_1 L - \ldots - \alpha_q L^q}$$
$$\Rightarrow E\left(\varepsilon_t^2\right) = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i}$$

The unconditional variance is finite, positive and constant (i.e., homoscedastic) if:

$$\sum_{i=1}^{q} \alpha_i < 1$$

This condition must hold if the process is stationary.

If this condition holds then volatility is a constant in the long-run.

GARCH(p,q) (Generalised Autoregressive Conditional Heteroscedasticity)

In practice q may need to be set high to capture all the non-linear dependence in returns.

Also with a lot of lags the non-negativity constraints are likely to be violated.

Bollerslev (1986) proposed the GARCH model as a <u>parsimonious</u> alternative to ARCH

For GARCH(p,q): $\Omega_{t-1} = \left\{ \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, \sigma_{t-1}^2, \ldots, \sigma_{t-p}^2 \right\}$

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

The non-negativity constraints (sufficient restrictions) are:

$$\boxed{\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \quad i = 1,.., q, \quad j = 1,.., p}$$

Typically p=q=1 is adequate in most empirical applications.

GARCH(1,1)

The GARCH(1,1) has an ARMA(1,1) representation in the squared residuals.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$
$$\Rightarrow \varepsilon_t^2 - u_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \left(\varepsilon_{t-1}^2 - u_{t-1} \right)$$
$$\Rightarrow \varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 - \beta_1 u_{t-1} + u_t$$

Volatility is stationary if:

$$\alpha_1 + \beta_1 < 1$$

But if

$$\alpha_1 + \beta_1 = 1$$

then shocks to volatility have a permanent effect:
⇒INTEGRATED GARCH (IGARCH) process.

Apparent IGARCH behaviour is found quite often in empirical work (see below).

The unconditional/long-run variance is:

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

The long-run variance converges to a constant iff:

$$\alpha_1 + \beta_1 < 1$$

An implication of IGARCH is that investors should be frequently altering their portfolios following shocks to reflect permanent changes in risk. Since this kind of behaviour isn't observed, IGARCH is incompatible with volatility in the 'real world'. It's possible that shocks to volatility are just highly persistent if not permanent.

Multivariate GARCH (MGARCH)

Generalization of GARCH to systems of n-asset returns.

The conditional volatility is an n×n variance-covariance matrix:

$$H_t = \begin{pmatrix} \sigma^2_{1,t} & \cdots & \sigma_{1n,t} \\ \vdots & \ddots & \vdots \\ \sigma_{n1,t} & \cdots & \sigma^2_{n,t} \end{pmatrix}$$

Conditional covariances off the diagonal

Conditional variances on the diagonal

A widely used formulation of MGARCH is the BEKK model:

$$H_t = W'W + A'H_{t-1}A + B'\Sigma_{t-1}\Sigma'_{t-1}B$$

$$\Sigma_{t-1} = \begin{pmatrix} \varepsilon_{1,t-1} \\ \vdots \\ \varepsilon_{n,t-1} \end{pmatrix}$$

W, A and B are n×n matrices of parameters.

H is positive definite (because the RHS terms are quadratic forms):
⇒The variances are positive
⇒The off-diagonal terms are symmetric: $(\sigma_{ij} = \sigma_{ji})$

ARCH-GARCH processes and fat-tail distributions

Another nice feature of ARCH-GARCH models is that they generate fat-tailed unconditional returns distributions (which are observed empirically – see Seminar 1/2).

For example an ARCH(1) model produces an unconditional return distribution with kurtosis coefficient

$$m_4 = \frac{3\left(1-\alpha_1^2\right)}{1-3\alpha_1^2} > 3$$

(recall the normal distribution has a kurtosis coefficient=3).

Financial applications of volatility/GARCH modelling

Value at Risk (VaR) — NOT TO BE CONFUSED WITH VAR (VECTOR-AUTOREGRESSIVE PROCESS) see lecture 9

VaR measures the £value of market risk on an asset/portfolio of marketable assets.

The maximum the investor can expect to lose in 19/20 days = VaR (at a 5% critical value) $\Rightarrow$ expect to lose more than VaR in 1/20 days.

In the case of a single asset if the return is normally distributed then a 90% confidence interval for the return is

$$\mu \pm 1.65\sigma$$

Returns will be less than μ-1.65σ on 1 in every 20 days (5% of the time).

Assuming $\mu \cong 0$ (reasonable for daily returns) then the downside risk with 5% probability is 1.65σ

Financial applications of volatility/GARCH modelling

VaR

If the value of the asset is £V then: $VaR_t = £V_t \times 1.65\sigma_t$

A forecast of volatility is needed to calculate VaR.

- GARCH provides one option for making this forecast.
- A more commonly used model for VaR volatility is an exponentially weighted moving average (EWMA):

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

Weighted average of lagged ex-ante/forecasted volatility σ_{t-1}^2 and lagged ex-post/realized volatility r_{t-1}^2 (assuming μ=0).

$$\Rightarrow \sigma_t^2 = \lambda^t\sigma_0^2 + (1-\lambda)\sum_{j=1}^{t}\lambda^{j-1}r_{t-j}^2$$

Weights attached to previous volatility decline geometrically/exponentially with the lag.

EWMA used widely by practitioners e.g., JP Morgan (who recommend using λ=0.94)

Financial applications of volatility/GARCH modelling

Dynamic hedge ratios

A common risk-management practice is to take opposite positions in spot and futures markets (a futures hedge).

The finance director's job is to determine the optimal hedge ratio: θ≡number of futures contracts/number of spot contracts.

The optimal value of θ will minimize the risk on the spot-futures portfolios. Choose θ to minimize the portfolio variance:

$$\text{var}\left(r_t^s - \theta_t r_t^f\right) = \text{var}\left(\theta_t r_t^f - r_t^s\right)$$

Short hedge — Long hedge

$$= \sigma_{s,t}^2 + \theta_t^2 \sigma_{f,t}^2 - 2\theta_t \sigma_{sf,t}$$

$$\Rightarrow 2\theta_t \sigma_{f,t}^2 = 2\sigma_{sf,t} \quad \text{(FOC)}$$

$$\Rightarrow \theta_t = \frac{\sigma_{sf,t}}{\sigma_{f,t}^2}$$

Simplest estimate of θ is a static hedge:
Regress r^s on r^f
θ=estimated slope coefficient

If the variances/covariances are time-varying estimate θ using an MGARCH model (n=2) for the spot and futures returns.

See Brooks 8.28/8.29 for applications of MGARCH to estimating dynamic hedge ratios and time-varying CAPM betas.

ARCH-GARCH Modelling Strategy

Analogous to ARMA modeling (Box-Jenkins technique: see lecture 5), ARCH-GARCH modeling involves:

1. Identification of a suitable ARCH-GARCH model
2. Estimation (using Maximum Likelihood)
3. Testing/diagnostic checking of the model to ensure it provides an adequate representation of the actual DGP (see Seminar 5).

Identification of ARCH-GARCH models

First perform a simple ARCH-test (see Appendix) to test for ARCH effects.

If there are ARCH effects present, identify a particular ARCH-GARCH model as follows:

i) The AR and ARMA representations for the squared residuals suggest the ACF and PACFs of the squared residuals can be used to identify a specific ARCH-GARCH model.

An ARCH(q) model is indicated by:

a) An infinite decay in the ACF of the squared residuals.

b) q spikes in the PACF of the squared residuals.

A GARCH(p,q) model is indicated by an infinite decay in both the ACF and the PACF of the squared residuals.

ii) In practice its more common to use information criteria such as the Schwarz Criterion to help select a model (see lecture 5).

iii) Many authors simply assume a GARCH(1,1) specification.

Maximum Likelihood (ML) (see Brooks Chp 8, Appendix)

Suppose we have a sample of independent observations $\left(y_1,\ldots,y_T\right)$ drawn from a known density

$$f\left(y_1,\ldots,y_T \middle| \theta\right) = f\left(y_1 \middle| \theta\right) \times \ldots \times f\left(y_T \middle| \theta\right)$$

The probability of observing different realizations of y for given parameters θ.

However the parameters are unknown and there is a given sample of data. Therefore re-interpret the joint distribution as the *likelihood* function:

The likelihood of observing the (given) data for different values of θ

$$L\left(\theta \middle| y_1,\ldots,y_T\right) \equiv f\left(y_1 \middle| \theta\right) \times \ldots \times f\left(y_T \middle| \theta\right)$$

ML estimator found by maximizing the log-likelihood function

$$\log L(\theta) = \sum_{t=1}^{T} \log f\left(y_t \middle| \theta\right)$$

with respect to θ

ML chooses θ to maximize the likelihood of observing the sample data.

ML estimators are:

1) Consistent.
2) Asymptotically efficient.
3) Asymptotically normally distributed.

ML estimation of ARCH-GARCH model

The ARCH-GARCH likelihood function involves the conditional error density. If v~NID(0,1) then:

$$f\left(\varepsilon_t \middle| \theta\right) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{-\frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2}\right\} = \frac{1}{\sqrt{\sigma_t^2}} f\left(v_t \middle| \theta\right)$$

The log likelihood function is:

$$\log L(\theta) = \sum_{t=1}^{T}\left[\log f\left(v_t \middle| \theta\right) - \frac{1}{2}\log\left(\sigma_t^2\right)\right]$$

$$= -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log\left(\sigma_t^2\right) - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{\sigma_t^2}$$

$\varepsilon_t^2 = (r_t - \mu)^2$ is a function of the unknown parameters of the conditional mean.

σ_t^2 is a function of the unknown parameters of the conditional variance.

The ML estimates of the conditional mean/variance parameters are the values which maximize log $L(\theta)$.

Alternative conditional distributions (see Seminar 5)

What happens if the standardized residuals, *v*, are non-normal (but normality is assumed)?

The ML estimator is still consistent and asymptotically normal but the standard errors are inconsistent.

Eviews gives an option to estimate ARCH-GARCH models with a conditional normal distribution but with a var-cov matrix which is robust to non-normality. This is known as *Quasi*-ML (QML)

Alternatively we can choose a different conditional distribution. *Eviews* allows the choice of a:

i) Student's *t* distribution

ii) Generalized error distribution

Extensions to ARCH-GARCH

Asymmetric GARCH

Standard GARCH models force a symmetric response of the conditional variance to shocks (since σ_t^2 depends on lagged squared residuals).

However, typically bad news (negative shocks) may be expected to increase volatility more than good news (positive shocks) of the same magnitude.

In the context of equity returns this may be due to leverage effects $\Rightarrow$

Negative shocks result in a fall in the value of the firm which increases the debt-equity ratio.

As a result stockholders perceive the firm as being more risky $\Rightarrow$ volatility increases.

Asymmetric GARCH models : Threshold ARCH (TARCH) – Glosten, Jagannathan and Runkle (GJR) model

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

where:

$$I_{t-1} = 1 \text{ if } \varepsilon_{t-1} < 0$$
$$= 0 \text{ if } \varepsilon_{t-1} > 0.$$

Dummy variable

When $\varepsilon_{t-1} > 0$ (good news) the ARCH effect is α_1.

When $\varepsilon_{t-1} < 0$ (bad news) the ARCH effect is $\alpha_1 + \gamma$.

If leverage effects are present then expect $\gamma > 0$.

Asymmetric GARCH models: Exponential GARCH (EGARCH)

$$\log \sigma_t^2 = \alpha + \beta \log \sigma_{t-1}^2 + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \delta \left[\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}} \right]$$

The log transformation ensures that the conditional variance is positive regardless of the parameter values $\Rightarrow$ no need for non-negativity constraints.

Also the effect of past shocks is exponential rather than quadratic (as in GARCH/TARCH)

In the EGARCH model leverage effects $\Rightarrow \gamma<0$

Asymmetry coefficient has opposite sign from TARCH

Bad news

$$\gamma < 0 \text{ and } \varepsilon_{t-1} < 0 \Rightarrow \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} > 0$$

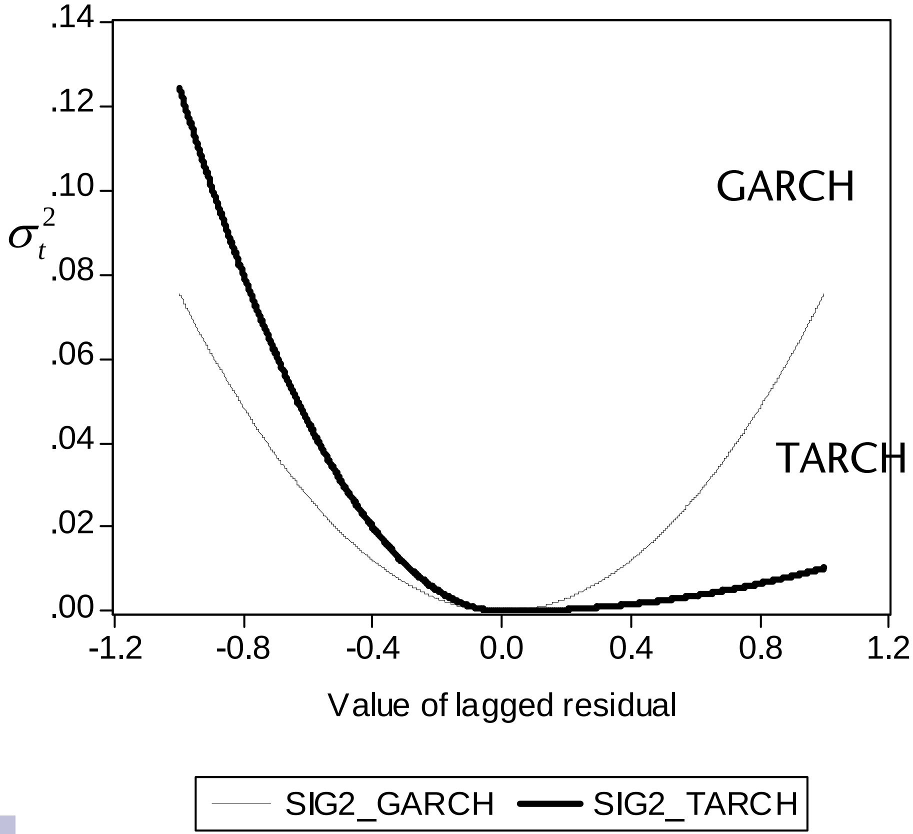
$$\gamma < 0 \text{ and } \varepsilon_{t-1} > 0 \Rightarrow \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} < 0$$

Good news

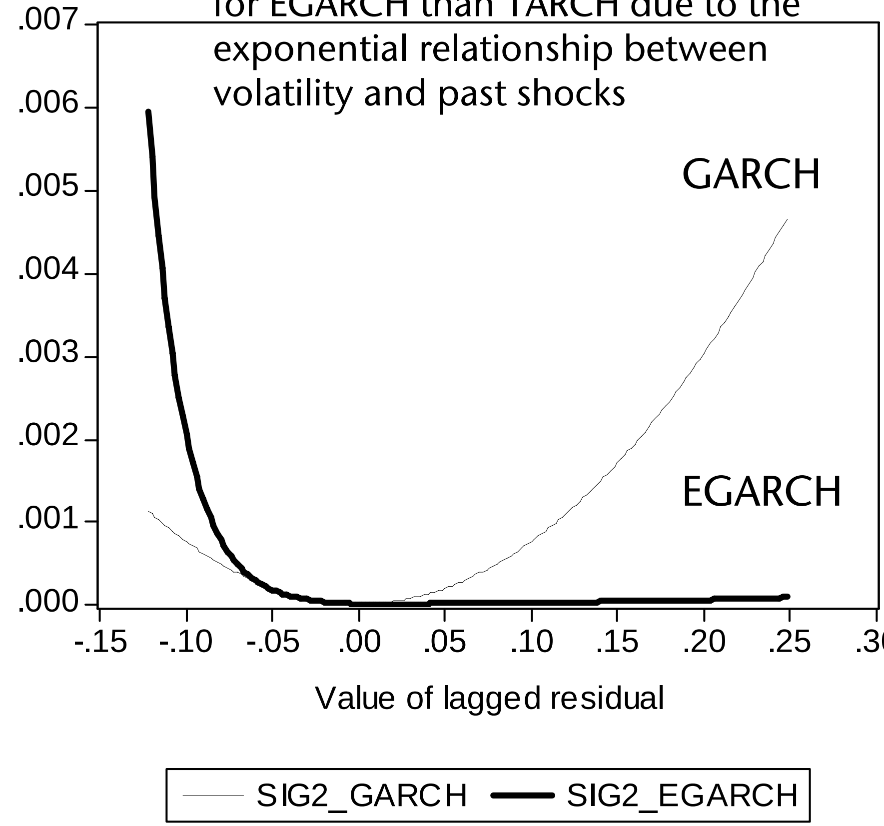
The News Impact Curve (NIC)

The NIC describes the response of volatility to past shocks:

- In the GARCH model the impact of shocks is symmetric about the origin.
- In the asymmetric models (TARCH/EGARCH) negative shocks have a bigger impact on volatility than positive shocks of the same magnitude.



The asymmetry is more pronounced for EGARCH than TARCH due to the exponential relationship between volatility and past shocks



GARCH-in-Mean (GARCH-M) (see Seminar 5)

CAPM applied to the market portfolio gives:

$$r_t = r_f + \lambda\sigma_t^2 + \varepsilon_t$$

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

Applying CAPM to the market portfolio

$$\begin{aligned} E(r) - r_f &= \beta\left[E(r_m) - r_f\right] \\ &= \frac{\operatorname{cov}(r, r_m)}{\sigma_m^2}\left[E(r_m) - r_f\right] \\ &= \lambda \operatorname{cov}(r, r_m) \\ &= \lambda\sigma_m^2, \quad r \equiv r_m \end{aligned}$$

λ Is the MARKET PRICE OF RISK :

$$\left(E(r_m) - r_f\right)/\sigma_m^2$$

The market return therefore varies directly with the conditional variance (which is usually modelled as a GARCH(1,1) process as above).

This is an example of a GARCH-M model: the conditional variance enters the mean equation.

GARCH-M is a model of a time-varying risk premium.

Recall that mistakenly assuming a constant risk premium may give rise to rejection of the EMH (see lecture 2)

Conclusion

GARCH models have assumed a central role in empirical finance for modelling time varying volatility.

There is a proliferation of variations on the ‘plain vanilla’ GARCH model (including models for absolute/power returns, multivariate models…)

Numerous applications in finance discussed well in Brooks Chp 8 and Cuthbertson & Nitzsche Chp 29).

References

Bollerslev (1986), A generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, 31, 307-327.

Brooks (2002), Introductory econometrics for finance, CUP: Cambridge. Chp 8**

Cuthbertson and Nitzsche (2004) Quantitative financial economics: stocks, bonds and foreign exchange, Wiley: Chichester. Chps 29.1-3**

Engle (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, *Econometrica*, 50, 987-1008.

**Key references

Appendix: A simple test for ARCH effects

ARCH LM Test

Step 1: Estimate the mean equation

$$r_t = \hat{\mu} + \hat{\varepsilon}_t$$

Step 2: Regress $\hat{\varepsilon}_t^2$ on q lagged values of itself

$$\hat{\varepsilon}_t^2 = \gamma_0 + \gamma_1 \hat{\varepsilon}_{t-1}^2 + ... + \gamma_q \hat{\varepsilon}_{t-q}^2 + \hat{u}_t$$

Test $H_0 : \gamma_1 = ... = \gamma_q = 0$ using the LM statistic

$$TR^2 \overset{a}{\sim} \chi_q^2$$

Reject the null (no ARCH effects) if the LM stat exceeds the 5% critical value of the χ_q^2 distribution.

ARCH Test example: SP500 returns (In *Eviews*: View/Residual Test/ARCH LM Test…)

ARCH Test:				
F-statistic	79.48879	Prob. F(5,3022)		0
Obs*R-squared	351.9462	Prob. Chi-Square(5)		0
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 02/19/06 Time: 22:23				
Sample (adjusted): 1/11/1994 1/11/2006				
Included observations: 3028 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.78E-05	5.43E-06	8.807116	0
RESID^2(-1)	0.119701	0.018071	6.623788	0
RESID^2(-2)	0.106659	0.01815	5.876659	0
RESID^2(-3)	0.142693	0.018068	7.897765	0
RESID^2(-4)	0.076	0.018149	4.187455	0
RESID^2(-5)	0.114438	0.018071	6.332504	0
R-squared	0.116231	Mean dependent va		0.000109
Adjusted R-squa	0.114768	S.D. dependent var		0.000262
S.E. of regressi	0.000247	Akaike info criterion		-13.77641
Sum squared re	0.000184	Schwarz criterion		-13.76449
Log likelihood	20863.48	F-statistic		79.48879
Durbin-Watson s	2.010396	Prob(F-statistic)		0

The test rejects
⇒evidence of ARCH effects in SP500 returns