

# Financial Time Series Analysis and Econometrics II

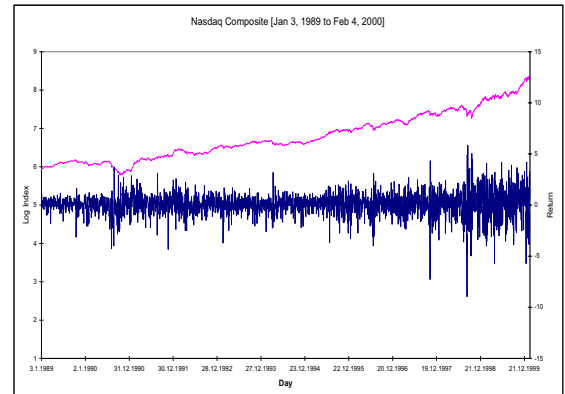
Seppo Pynnönen, 2008\*

\*©Professor Seppo Pynnönen, Department of Mathematics and Statistics, University of Vaasa, Box 700, 65100 Vaasa, phone: (06)324 8259 email: sjp@uwasa.fi, URL: <http://www.uwasa.fi/~sjp/>

## 1. Nonlinear Univariate Times Series

### 1.1 Background

**Example 1.1:** Consider the following daily close-to-close Nasdaq composite share index values [January 3, 1989 to February 4, 2000]

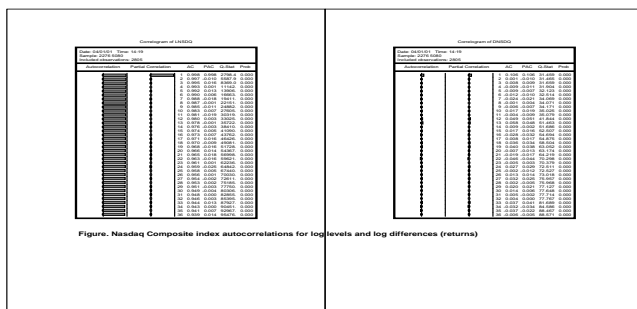


1

Below are autocorrelations of the log-index.

Obviously the persistence of autocorrelations indicate that the series is integrated (see definition below).

The autocorrelations of the return series suggest that the returns are stationary with statistically significant first order autocorrelation.



2

**Definition 1.1:** Time series  $y_t, t = 1, \dots, T$  is covariance stationary if

$$\begin{aligned} E[y_t] &= \mu, \text{ for all } t \\ (1) \quad \text{Cov}[y_t, y_{t+k}] &= \gamma_k, \text{ for all } t \\ \text{Var}[y_t] &= \gamma_0 (< \infty), \text{ for all } t \end{aligned}$$

Any series that are not stationary are said to be nonstationary.

**Definition 1.2:** Time series  $u_t$  is a white noise process if

$$\begin{aligned} E[u_t] &= \mu, \text{ for all } t \\ (2) \quad \text{Cov}[u_t, u_s] &= 0, \text{ for all } t \neq s \\ \text{Var}[u_t] &= \sigma_u^2 < \infty, \text{ for all } t. \end{aligned}$$

We denote  $u_t \sim WN(\mu, \sigma_u^2)$ .

**Remark 1.1:** Usually it is assumed in (2) that  $\mu = 0$ .

**Remark 1.2:** A WN-process is obviously stationary.

3

**Definition 1.3:** Times series  $y_t$  is said to be integrated of order 1, if it is of the form

$$(3) \quad (1 - L)y_t = \delta + \psi(L)u_t,$$

denoted as  $y_t \sim I(1)$ , where

$$(4) \quad \psi(L) = 1 + \psi_1L + \psi_2L^2 + \psi_3L^3 + \dots$$

such that  $\sum_{j=1}^{\infty} |\psi_j| < \infty$ ,  $\psi(1) \neq 0$ , roots of  $\psi(z) = 0$  are outside the unit circle [or the polynomial (4) is of order zero], and  $u_t$  is a white noise series with mean zero and variance  $\sigma_u^2$ .

**Remark 1.3:** If a time series process is of the form of the right hand side of (4), i.e.,

$$(5) \quad x_t = \delta + \psi(L)u_t,$$

where  $\psi(L)$  satisfies the conditions of Def 1.3, it can be shown that  $x_t$  is stationary. We denote  $x_t \sim I(0)$ , i.e, integrated of order zero.

**Remark 1.4:** The assumption  $\psi(1) \neq 0$  is important. It rules out for example trend stationary series

$$(6) \quad y_t = \alpha + \beta t + \psi(L)u_t.$$

Because  $E[y_t] = \alpha + \beta t$ ,  $y_t$  is nonstationary. However,

$$(7) \quad (1 - L)y_t = \beta + \tilde{\psi}(L)u_t,$$

where

$$(8) \quad \tilde{\psi}(L) = (1 - L)\psi(L).$$

Now, although,  $(1 - L)y_t$  is stationary, however,

$$\tilde{\psi}(1) = (1 - 1)\psi(1) = 0,$$

which does not satisfy the rule in Definition 1.3, and hence a trend stationary series is not  $I(1)$ .

**Example 1.2:** (Example 1.1 continued) Below are results after fitting an AR(1) and an MA(1) model to the return series

**Table.** AR(1) estimates.

Dependent Variable: DNSDQ  
Method: Least Squares  
Sample: 2276 5080  
Included observations: 2805  
Convergence achieved after 2 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.086126	0.023048	3.736845	0.0002
AR(1)	0.105933	0.018782	5.640001	0.0000

R-squared	0.011221	Mean dependent var	0.086119
Adjusted R-squared	0.010868	S.D. dependent var	1.097336
S.E. of regression	1.091357	Akaike info criterion	3.013434
Sum squared resid	3338.542	Schwarz criterion	3.017668
Log likelihood	-4224.341	F-statistic	31.80961
Durbin-Watson stat	1.997947	Prob(F-statistic)	0.000000
Inverted AR Roots	.11		

**Table.** MA(1) estimates

Dependent Variable: DNSDQ  
Method: Least Squares  
Sample: 2276 5080  
Included observations: 2805  
Convergence achieved after 4 iterations

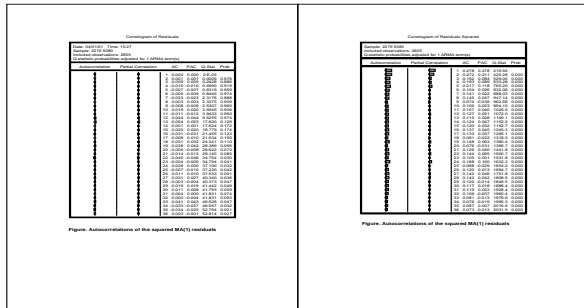
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.086153	0.022811	3.776796	0.0002
MA(1)	0.107093	0.018779	5.702685	0.0000

R-squared	0.011323	Mean dependent var	0.086119
Adjusted R-squared	0.010970	S.D. dependent var	1.097336
S.E. of regression	1.091301	Akaike info criterion	3.013331
Sum squared resid	3338.198	Schwarz criterion	3.017565
Log likelihood	-4224.196	F-statistic	32.10153
Durbin-Watson stat	2.000051	Prob(F-statistic)	0.000000
Inverted MA Roots	-.11		

Both models give virtually equally good fit, MA(1) is just only marginally better.

The residual autocorrelations and related Q-statistics indicate no further autocorrelation left to the series.



The autocorrelations of the squared residuals strongly suggest that there is still left time dependency into the series.

The dependency, however, is nonlinear by nature.

Because squared observations are the building blocks of the variance of the series, the results suggest that the variation (volatility) of the series is time dependent.

This leads to the so called ARCH-family of models.<sup>‡</sup>

<sup>‡</sup>The inventor of this modeling approach is Robert F. Engle (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987–1008.

## 1.2. Conditional Heteroscedasticity

### ARCH-models

The general setup for ARCH models is

$$(9) \quad y_t = \mathbf{x}_t' \boldsymbol{\beta} + u_t$$

with  $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{pt})'$ ,  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)'$ ,  $t = 1, \dots, T$ , and

$$(10) \quad u_t | \mathcal{F}_{t-1} \sim N(0, h_t),$$

where  $\mathcal{F}_t$  is the information available at time  $t$  (usually the past values of  $u_t$ ;  $u_1, \dots, u_{t-1}$ ), and

$$(11) \quad h_t = \text{Var}(u_t | \mathcal{F}_{t-1}) = \omega + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2.$$

Furthermore, it is assumed that  $\omega > 0$ ,  $\alpha_i \geq 0$  for all  $i$  and  $\alpha_1 + \dots + \alpha_q < 1$ .

For short it is denoted  $u_t \sim \text{ARCH}(q)$ .

This reminds essentially an  $\text{AR}(q)$  process for the squared residuals, because defining  $\nu_t = u_t^2 - h_t$ , we can write

$$(12) \quad u_t^2 = \omega + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2 + \nu_t.$$

Nevertheless, the error term  $\nu_t$  is time heteroscedastic, which implies that the conventional estimation procedure used in AR-estimation does not produce optimal results here.

## Properties of ARCH-processes

Consider (for the sake of simplicity) ARCH(1) process

$$(13) \quad h_t = \omega + \alpha u_{t-1}^2$$

with  $\omega > 0$  and  $0 \leq \alpha < 1$  and  $u_t | u_{t-1} \sim N(0, h_t)$ .

(a)  $u_t$  is white noise

(i) Constant mean (zero):

$$(14) \quad E[u_t] = E[\underbrace{E_{t-1}[u_t]}_{=0}] = E[0] = 0.$$

Note  $E_{t-1}[u_t] = E[u_t | \mathcal{F}_{t-1}]$ , the conditional expectation given information up to time  $t-1$ .<sup>§</sup>

<sup>§</sup>The law of iterated expectations: Consider time points  $t_1 < t_2$  such that  $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ , then for any  $t > t_2$

$$(15) \quad E_{t_1}[E_{t_2}[u_t]] = E[E[u_t | \mathcal{F}_{t_2}] | \mathcal{F}_{t_1}] = E[u_t | \mathcal{F}_{t_1}] = E_{t_1}[u_t].$$

(ii) Constant variance: Using again the law of iterated expectations, we get

$$\begin{aligned} \text{Var}[u_t] &= E[u_t^2] = E[E_{t-1}[u_t^2]] \\ &= E[h_t] = E[\omega + \alpha u_{t-1}^2] \\ &= \omega + \alpha E[u_{t-1}^2] \\ &\quad \vdots \\ (16) \quad &= \omega(1 + \alpha + \alpha^2 + \dots + \alpha^n) \\ &\quad + \underbrace{\alpha^{n+1} E[u_{t-n-1}^2]}_{\rightarrow 0, \text{ as } n \rightarrow \infty} \\ &= \omega \left( \lim_{n \rightarrow \infty} \sum_{i=0}^n \alpha^i \right) \\ &= \frac{\omega}{1-\alpha}. \end{aligned}$$

(iii) Autocovariances: Exercise, show that autocovariances are zero, i.e.,  $E[u_t u_{t+k}] = 0$  for all  $k \neq 0$ . (*Hint*: use the law of iterated expectations.)

(b) The unconditional distribution of  $u_t$  is symmetric, but nonnormal:

(i) Skewness: Exercise, show that  $E[u_t^3] = 0$ .

(ii) Kurtosis: Exercise, show that under the assumption  $u_t | u_{t-1} \sim N(0, h_t)$ , and that  $\alpha < \sqrt{1/3}$ , the kurtosis

$$(17) \quad E[u_t^4] = 3 \frac{\omega^2}{(1-\alpha)^2} \cdot \frac{1-\alpha^2}{1-3\alpha^2}.$$

*Hint*: If  $X \sim N(0, \sigma^2)$  then  $E[(X-\mu)^4] = 3(\sigma^2)^2 = 3\sigma^4$ .

Because  $(1-\alpha^2)/(1-3\alpha^2) > 1$  we have that

$$(18) \quad E[u_t^4] > 3 \frac{\omega^2}{(1-\alpha)^2} = 3[\text{Var}(u_t)]^2,$$

we find that the kurtosis of the unconditional distribution exceed that what it would be, if  $u_t$  were normally distributed.

Thus the unconditional distribution of  $u_t$  is nonnormal and has fatter tails than a normal distribution with variance equal to  $\text{Var}[u_t] = \omega/(1-\alpha)$ .

(c) Standardized variables:

Write

$$(19) \quad z_t = \frac{u_t}{\sqrt{h_t}}$$

then  $z_t \sim \text{NID}(0, 1)$ , i.e., normally and independently distributed.

Thus we can always write

$$(20) \quad u_t = z_t \sqrt{h_t},$$

where  $z_t$  independent standard normal random variables (strict white noise).

This gives us a useful device to check after fitting an ARCH model the adequacy of the specification: Check the autocorrelations of the squared standardized series.

### Estimation of ARCH models

Given the model

$$(21) \quad y_t = \mathbf{x}'_t \beta + u_t$$

with  $u_t | \mathcal{F}_{t-1} \sim N(0, h_t)$ , we have  $y_t | \{\mathbf{x}_t, \mathcal{F}_{t-1}\} \sim N(\mathbf{x}'_t \beta, h_t)$ ,  $t = 1, \dots, T$ .

Then the log-likelihood function becomes

$$(22) \quad \ell(\theta) = \sum_{t=1}^T \ell_t(\theta)$$

with

$$\ell_t(\theta) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log h_t - \frac{1}{2} (y_t - \mathbf{x}'_t \beta)^2 / h_t,$$

(23) where  $\theta = (\beta', \omega, \alpha)'$ .

The maximum likelihood (ML) estimate  $\hat{\theta}$  is the value maximizing the likelihood function, i.e.,

$$(24) \quad \ell(\hat{\theta}) = \max_{\theta} \ell(\theta).$$

The maximization is accomplished by numerical methods.

Remark 1.5: OLS estimates of the regression parameters are inefficient (unreliable) compared to the ML estimates.

### Generalized ARCH models

In practice the ARCH needs fairly many lags.

Usually far less lags are needed by modifying the model to

$$(25) \quad h_t = \omega + \alpha u_{t-1}^2 + \delta h_{t-1},$$

with  $\omega > 0$ ,  $\alpha > 0$ ,  $\delta \geq 0$ , and  $\alpha + \delta < 1$ .

The model is called the Generalized ARCH (GARCH) model.

Usually the above GARCH(1,1) is adequate in practice.

Econometric packages call  $\alpha$  (coefficient of  $u_{t-1}^2$ ) the ARCH parameter and  $\delta$  (coefficient of  $h_{t-1}$ ) the GARCH parameter.

Note again that defining  $\nu_t = u_t^2 - h_t$ , we can write

$$(26) \quad u_t^2 = \omega + (\alpha + \delta)u_{t-1}^2 + \nu_t - \delta\nu_{t-1}$$

a heteroscedastic ARMA(1,1) process.

Applying backward substitution, one easily gets

$$(27) \quad h_t = \frac{\omega}{1 - \delta} + \alpha \sum_{j=1}^{\infty} \delta^{j-1} u_{t-j}^2$$

an ARCH( $\infty$ ) process.

Thus the GARCH term captures all the history from  $t - 2$  backwards of the shocks  $u_t$ .

Imposing additional lag terms, the model can be extended to GARCH( $r, q$ ) model

$$(28) \quad h_t = \omega + \sum_{j=1}^r \delta_j h_{t-j} + \sum_{i=1}^q \alpha_i u_{t-i}^2$$

[c.f. ARMA( $p, q$ )].

Nevertheless, as noted above, in practice GARCH(1,1) is adequate.

**Example 1.3: MA(1)-GARCH(1,1) model of Nasdaq returns.** The model is

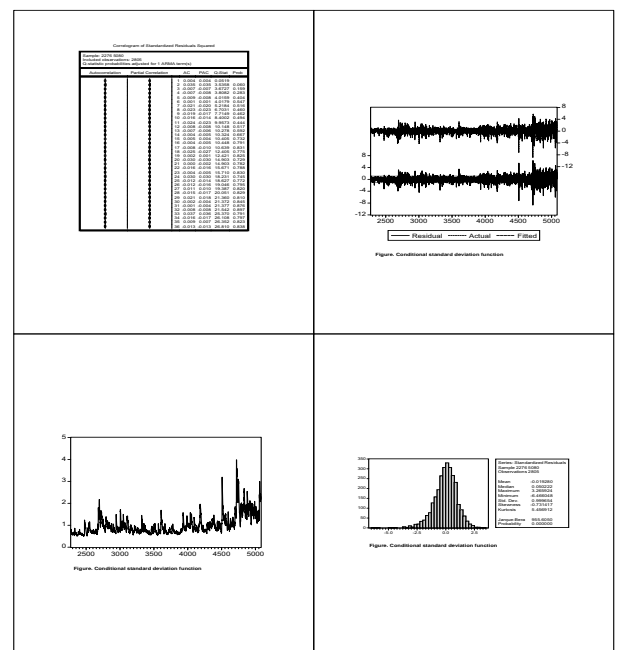
$$(29) \quad \begin{aligned} r_t &= \mu + u_t + \theta u_{t-1} \\ h_t &= \omega + \alpha u_{t-1}^2 + \delta h_{t-1}. \end{aligned}$$

Estimation results (EViews 4.0)

Dependent Variable: DNSDQ  
 Method: ML - ARCH (Marquardt)  
 Sample: 2276 5080  
 Included observations: 2805  
 Convergence achieved after 21 iterations  
 Bollerslev-Wooldridge robust standard errors & covariance  
 MA backcast: 2275, Variance backcast: ON

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.084907	0.017700	4.797124	0.0000
MA(1)	0.171620	0.020952	8.190983	0.0000
Variance Equation				
C	0.027892	0.009213	3.027258	0.0000
ARCH(1)	0.121770	0.020448	5.955103	0.0000
GARCH(1)	0.857095	0.021526	39.81666	0.0000

R-squared	0.007104	Mean dependent var	0.086119
Adjusted R-squared	0.005685	S.D. dependent var	1.097336
S.E. of regression	1.094213	Akaike info criterion	2.695856
Sum squared resid	3352.444	Schwarz criterion	2.706443
Log likelihood	-3775.938	F-statistic	5.008069
Durbin-Watson stat	2.129878	Prob(F-statistic)	0.000507
Inverted MA Roots	-.17		



The autocorrelations of the squared standardized residuals pass the white noise test.

Nevertheless, the normality of the standardized residuals is strongly rejected.

This is why robust standard errors are used in the estimation of the standard errors.

The variance function can be extended by including regressors (exogenous or predetermined variables),  $x_t$ , in it

$$(30) \quad h_t = \omega + \alpha u_{t-1}^2 + \delta h_{t-1} + \pi x_t.$$

Note that if  $x_t$  can assume negative values, it may be desirable to introduce absolute values  $|x_t|$  in place of  $x_t$  in the conditional variance function.

For example, with daily data a Monday dummy could be introduced into the model to capture the non-trading over the weekends in the volatility.

### ARCH-M Model

The regression equation may be extended by introducing the variance function into the equation

$$(31) \quad y_t = \mathbf{x}_t' \beta + \gamma g(h_t) + u_t,$$

where  $u_t \sim \text{GARCH}$ , and  $g$  is a suitable function (usually square root or logarithm).

This is called the ARCH in Mean (ARCH-M) model [Engle, Lilien and Robbins (1987)<sup>¶</sup>].

The ARCH-M model is often used in finance, where the expected return on an asset is related to the expected asset risk.

The coefficient  $\gamma$  reflects the risk-return trade-off.

<sup>¶</sup>Econometrica, 55, 391–407.

### Example 1.4: Does the daily mean return of Nasdaq depend on the volatility level?

Dependent Variable: DNSDQ  
 Method: ML - ARCH (Marquardt)  
 Sample: 2276 5080  
 Included observations: 2805  
 Convergence achieved after 22 iterations  
 Bollerslev-Wooldrige robust standard errors & covariance  
 MA backcast: 2275, Variance backcast: ON

	Coefficient	Std. Error	z-Statistic	Prob.
SQR(GARCH)	0.198064	0.074141	2.671456	0.0076
C	-0.069416	0.061432	-1.129969	0.2585
MA(1)	0.174785	0.020644	8.466806	0.0000

Variance Equation				
C	0.031799	0.009301	3.419007	0.0006
ARCH(1)	0.134070	0.020974	6.392287	0.0000
GARCH(1)	0.842134	0.021350	39.444407	0.0000

R-squared	0.011379	Mean dependent var	0.086119
Adjusted R-squared	0.009613	S.D. dependent var	1.097336
S.E. of regression	1.092049	Akaike info criterion	2.694709
Sum squared resid	3338.007	Schwarz criterion	2.707413
Log likelihood	-3773.330	F-statistic	6.443432
Durbin-Watson stat	2.127550	Prob(F-statistic)	0.000006
Inverted MA Roots	-0.17		

The volatility term in the mean equation is statistically significant indicating that rather than being constant the mean return is dependent on the level of volatility.

Consequently, the data suggests that the best fitting model so far is of the form

$$r_t = \gamma\sqrt{h_t} + u_t + \theta u_{t-1}$$

$$h_t = \omega + \alpha u_{t-1}^2 + \delta h_{t-1}$$

Below are the estimation results for the above model

Dependent Variable: DNSDQ  
 Method: ML - ARCH (Marquardt)  
 Sample: 2276 5080  
 Included observations: 2805  
 Convergence achieved after 16 iterations  
 Bollerslev-Wooldrige robust standard errors & covariance  
 MA backcast: 2275, Variance backcast: ON

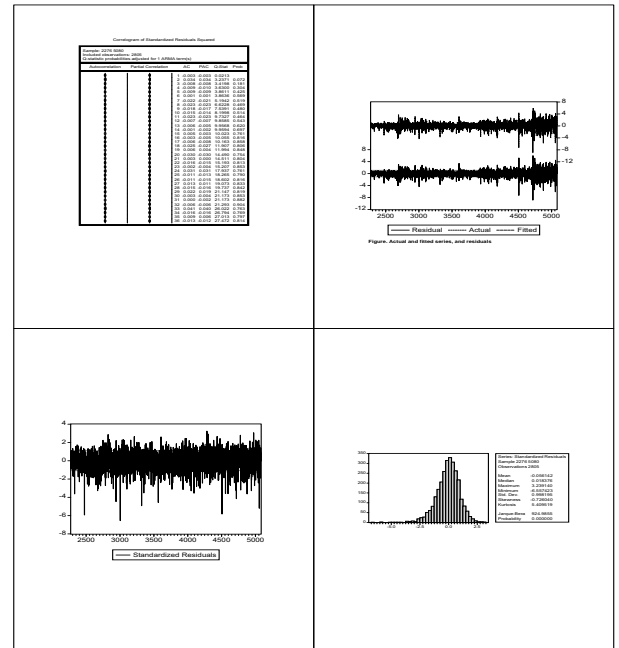
	Coefficient	Std. Error	z-Statistic	Prob.
SQR(GARCH)	0.119204	0.022944	5.195479	0.0000
MA(1)	0.174104	0.020771	8.382098	0.0000

Variance Equation				
C	0.031291	0.009545	3.278211	0.0010
ARCH(1)	0.133713	0.021011	6.363810	0.0000
GARCH(1)	0.843131	0.021785	38.70279	0.0000

R-squared	0.010861	Mean dependent var	0.086119
Adjusted R-squared	0.009448	S.D. dependent var	1.097336
S.E. of regression	1.092141	Akaike info criterion	2.694578
Sum squared resid	3339.759	Schwarz criterion	2.705165
Log likelihood	-3774.146	Durbin-Watson stat	2.132844
Inverted MA Roots	-.17		



Looking at the standardized residuals, the distribution and the sample statistics of the distribution, we observe that the residual distribution is obviously skewed in addition to the leptokurtosis.

The skewness may be due to some asymmetry in the conditional volatility which we have not yet modeled.

In financial data the asymmetry is usually such that downward shocks cause higher volatility in the near future than the positive shocks.

In finance this is called the leverage effect.

An obvious and simple first hand check for the asymmetry is to investigate the cross autocorrelations between standardized and squared standardized GARCH residuals.

Below are the cross autocorrelations between the standardized and squared standardized residuals of the fitted MA(1)-GARCH(1,1) model.

```

=====
Z,Z2(-i)          Z,Z2(+i)          i    lag    lead
=====
****|             |****|             | 0 -0.3916 -0.3916
|             |*|             | 1  0.0342 -0.0782
|             |*|             | 2 -0.0055 -0.0842
|             |             | 3  0.0093 -0.0373
|             |             | 4  0.0066  0.0315
|             |             | 5  0.0134 -0.0046
|             |             | 6 -0.0134 -0.0019
|             |             | 7  0.0113 -0.0004
|             |             | 8 -0.0019  0.0045
|             |             | 9 -0.0034  0.0272
|             |             |10 -0.0205  0.0128
=====

```

The cross autocorrelations correlations are not large, but may indicate the presence of some asymmetry.

## Asymmetric ARCH: TARCh and EGARCH

A kind of stylized fact in stock markets is that downward movements are followed by higher volatility.

EViews includes two models that allow for asymmetric shocks to volatility.

32

## The TARCh model

Threshold ARCH, TARCh (Zakoian 1994, *Journal of Economic Dynamics and Control*, 931–955, Glosten, Jagannathan and Runkle 1993, *Journal of Finance*, 1779–1801) is given by [TARCh(1,1)]

$$(32)h_t = \omega + \alpha u_{t-1}^2 + \gamma u_{t-1}^2 d_{t-1} + \delta h_{t-1},$$

where  $d_t = 1$ , if  $u_t < 0$  (bad news) and zero otherwise.

The impact of good news is  $\alpha$  and bad news  $\alpha + \gamma$ .

Thus,  $\gamma \neq 0$  implies asymmetry.

Leverage exists if  $\gamma > 0$ .

33

### Example 1.5: Estimation results for the MA(1)-TARCh-M model.

```

Dependent Variable: DNSDQ
Method: ML - ARCH (Marquardt)
Sample: 2276 5080
Included observations: 2805
Convergence achieved after 26 iterations
Bollerslev-Wooldrige robust standard errors & covariance
MA backcast: 2275, Variance backcast: ON
=====
                Coefficient  Std. Error  z-Statistic  Prob.
=====
SQR(GARCH)      0.091184    0.023097    3.947880    0.0001
MA(1)           0.184263    0.020899    8.816678    0.0000
=====
                Variance Equation
=====
C                0.037068    0.009513    3.896566    0.0001
ARCH(1)         0.084275    0.025080    3.360240    0.0008
(RESID<0)*ARCH(1) 0.099893    0.040881    2.443502    0.0145
GARCH(1)        0.833239    0.019001    43.85202    0.0000
=====
R-squared        0.009604    Mean dependent var  0.086119
Adjusted R-squared 0.007835    S.D. dependent var  1.097336
S.E. of regression 1.093029    Akaike info criterion 2.686832
Sum squared resid 3344.000    Schwarz criterion  2.699536
Log likelihood   -3762.281    Durbin-Watson stat  2.149776
=====
Inverted MA Roots      -.18
=====

```

The goodness of fit improves and the statistically significant positive asymmetry parameter estimate indicates presence of leverage.

34

Furthermore, as seen below, the first few cross autocorrelations reduce to about one half of the original ones.

They are still statistically significant by slightly exceeding the approximate 95% boundaries  $\pm 2/\sqrt{T} = \pm 2/\sqrt{2805} \approx \pm 0.038$ .

Cross autocorrelations of the standardized and squared standardized MA(1)-TARCh(1,1)-M model.

Z, Z2(-i)	Z, Z2(+i)	i	lag	lead
****		****		0 -0.3543 -0.3543
				1 0.0304 -0.0488
				2 -0.0071 -0.0537
				3 0.0112 -0.0156
				4 0.0069 0.0545
				5 0.0113 0.0080

35

## The EGARCH model

Nelson (1991) (*Econometrica*, 347–370) proposed the Exponential GARCH (EGARCH) model for the variance function of the form (EGARCH(1,1))

$$(33) \quad \log h_t = \omega + \delta \log h_{t-1} + \alpha |z_{t-1}| + \gamma z_{t-1},$$

where  $z_t = u_t / \sqrt{h_t}$  is the standardized shock.

Again, the impact is asymmetric if  $\gamma \neq 0$ , and leverage is present if  $\gamma < 0$ .

## Example 1.6: MA(1)-EGARCH(1,1)-M estimation results.

Dependent Variable: DNSDQ  
 Method: ML - ARCH (Marquardt)  
 Sample: 2276 5080  
 Included observations: 2805  
 Convergence achieved after 28 iterations  
 Bollerslev-Wooldridge robust standard errors & covariance  
 MA backcast: 2275, Variance backcast: ON

	Coefficient	Std. Error	z-Stat	Prob.
SQR(GARCH)	0.084631	0.022593	3.745866	0.0002
MA(1)	0.171543	0.020387	8.414399	0.0000

Variance Equation				
C	-0.197193	0.023051	-8.554804	0.0000
RES /SQR[GARCH](1)	0.251752	0.030816	8.169621	0.0000
RES/SQR[GARCH](1)	-0.071425	0.024034	-2.971755	0.0030
EGARCH(1)	0.958125	0.010941	87.57385	0.0000

R-squared	0.010762	Mean dependent var	0.086119
Adjusted R-squared	0.008995	S.D. dependent var	1.097336
S.E. of regression	1.092390	Akaike info criter	2.682518
Sum squared resid	3340.093	Schwarz criterion	2.695222
Log likelihood	-3756.232	Durbin-Watson stat	2.124928

Inverted MA Roots	-.17
-------------------	------

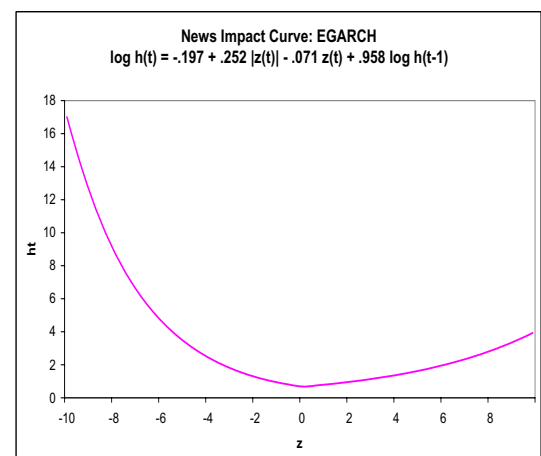
Cross autocorrelations (not shown here) are about the same as with the TARCH model (i.e., disappear). Thus TARCH and EGARCH capture most part of the leverage effect.

## News Impact Curve

The asymmetry of the conditional volatility function can be conveniently illustrated by the news impact curve (NIC).

The curve is simply the graph of  $h_t(z)$ , where  $z$  indicates the shocks (news).

Below is a graph for the NIC of the above estimate EGARCH variance function, where  $h_{t-1}$  is replaced by the median of the estimated EGARCH series.



## The Component ARCH Model

We can write the GARCH(1,1) model as

$$(34) \quad h_t = \bar{\omega} + \alpha(u_{t-1}^2 - \bar{\omega}) + \delta(h_{t-1} - \bar{\omega}),$$

where

$$(35) \quad \bar{\omega} = \frac{\omega}{1 - \alpha - \delta}$$

is the unconditional variance of the series.

Thus the usual GARCH has a mean reversion tendency towards  $\bar{\omega}$ .

A further extension is to allow this unconditional or long term volatility to vary over time.

This leads to so called component ARCH that allows mean reversion to a varying level  $q_t$  instead of  $\bar{\omega}$ .

The model is

$$\begin{aligned} h_t - q_t &= \alpha(u_{t-1}^2 - q_{t-1}) + \delta(h_{t-1} - q_{t-1}) \\ q_t &= \omega + \rho(q_{t-1} - \omega) + \theta(u_{t-1}^2 - h_{t-1}). \end{aligned} \quad (36)$$

An asymmetric version for the model is

$$\begin{aligned} h_t - q_t &= \alpha(u_{t-1}^2 - q_{t-1}) \\ &\quad + \alpha(u_{t-1}^2 - q_{t-1})d_{t-1} + \delta(h_{t-1} - q_{t-1}) \\ q_t &= \omega + \rho(q_{t-1} - \omega) + \theta(u_{t-1}^2 - h_{t-1}). \end{aligned} \quad (37)$$

### Example 1.7: Asymmetric Component ARCH of the Nasdaq composite returns.

```

Dependent Variable: DNSDQ
Method: ML - ARCH (Marquardt)
Sample: 2276 5080
Included observations: 2805
Convergence achieved after 4 iterations
Bollerslev-Wooldrige robust standard errors & covariance
MA backcast: 2275, Variance backcast: ON
=====
                Coefficient  Std. Error z-Statistic  Prob.
=====
SQR(GARCH)      0.097235   0.023997   4.052002   0.0001
MA(1)           0.182908   0.027822   6.574119   0.0000
=====
                Variance Equation
=====
Perm: C          0.926329   0.080794   11.46533   0.0000
Perm: [Q-C]     0.734067   0.077885   9.425047   0.0000
Perm: [ARCH-GARCH] 0.228360   0.048519   4.706581   0.0000
Tran: [ARCH-Q]  0.037121   0.039565   0.938228   0.3481
Tran: (RES<0)*[ARCH-Q] -0.077747   0.076100  -1.021635   0.3070
Tran: [GARCH-Q] -0.688202   0.353158  -1.948710   0.0513
=====
R-squared        0.008632   Mean dependent var   0.086119
Adjusted R-squared 0.006151   S.D. dependent var   1.097336
S.E. of regression 1.093956   Akaike info criterion 2.771137
Sum squared resid 3347.282   Schwarz criterion    2.788076
Log likelihood    -3878.519   Durbin-Watson stat   2.152828
=====
Inverted MA Roots    -.18
=====

```

This model, however, does not fit well into the data. Thus it seems that the best fitting models so far are either the TARARCH or EGARCH.

### 1.3 Regime switching models

A potentially useful approach to model nonlinearities in time series is to assume different behavior (structural break) in different sub-samples (or regimes).

If the dates, the regimes switches have taken place, are known, modeling can be worked out simply with dummy variables.

Consider the following regression model

$$(38) \quad y_t = \mathbf{x}_t' \beta_{S_t} + u_t, \quad t = 1, \dots, T,$$

where

$$(39) \quad u_t \sim \text{NID}(0, \sigma_{S_t}^2),$$

$$(40) \quad \beta_{S_t} = \beta_0(1 - S_t) + \beta_1 S_t,$$

$$(41) \quad \sigma_{S_t}^2 = \sigma_0^2(1 - S_t) + \sigma_1^2 S_t,$$

and

$$(42) \quad S_t = 0 \text{ or } 1, \quad (\text{Regime 0 or 1}).$$

Thus, under regime 1 the coefficient parameter vector is  $\beta_1$  and error variance  $\sigma_1^2$ .

For the sake of simplicity, consider an AR(1) model.

That is,  $\mathbf{x}_t = (1, y_{t-1})'$ .

Usually it is assumed that the possible difference between the regimes is a mean and/or a volatility shift, but no change in the autoregression parameter.

That is,

$$(43) \quad y_t = \mu_{S_t} + \phi_1(y_{t-1} - \mu_{S_{t-1}}) + u_t,$$

with

$$(44) \quad u_t \sim \text{NID}(0, \sigma_{S_t}^2),$$

where  $\mu_{S_t} = \mu_0(1 - S_t) + \mu_1 S_t$  and  $\sigma_{S_t}^2$  as defined above.

If  $S_t$ ,  $t = 1, \dots, T$  is known a priori, then the problem is just a usual dummy variable autoregression problem.

In practice, however, the prevailing regime is not usually directly observable.

Denote then

$$(45) \quad P(S_t = j | S_{t-1} = i) = p_{ij},$$

$i, j = 0, 1$ , called transition probabilities, with  $p_{i0} + p_{i1} = 1$ ,  $i = 0, 1$ .

This kind of process, where the next state depends only on the previous state, is called the Markov process, and the model a Markov switching model in the mean and variance.

Thus, in this model additional parameters to be estimated are the transition probabilities  $p_{ij}$ .

Usually the parameters are estimated (numerically) by the ML method.\*\*

\*\*For a detailed discussion, see Kim Chang-Jin and Charles A. Nelson (1999). *State Space Models with Regime Switching. Classical and Gibbs-Sampling Approaches with Applications*. MIT-Press.

The joint probability density function for  $y_t, S_t, S_{t-1}$ , given past information  $\mathcal{F}_{t-1} = \{y_{t-1}, y_{t-2}, \dots\}$ , is

$$f(y_t, S_t, S_{t-1} | \mathcal{F}_{t-1}) = f(y_t | S_t, S_{t-1}, \mathcal{F}_{t-1}) P(S_t, S_{t-1} | \mathcal{F}_{t-1}),$$

with

$$f(y_t | S_t, S_{t-1}, \mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_{S_t}^2}} \exp \left\{ -\frac{[y_t - \mu_{S_t} - \phi_1(y_{t-1} - \mu_{S_{t-1}})]^2}{2\sigma_{S_t}^2} \right\}.$$

(47)

Then the log-likelihood function to be maximized with respect to the unknown parameters is

$$(48) \quad \ell(\theta) = \sum_{t=1}^T \ell_t(\theta),$$

where

$$(49) \quad \ell_t(\theta) = \log \left[ \sum_{S_t=0}^1 \sum_{S_{t-1}=0}^1 f(y_t | S_t, S_{t-1}, \mathcal{F}_{t-1}) P[S_t, S_{t-1} | \mathcal{F}_{t-1}] \right],$$

$$(50) \quad \theta = (p, q, \mu_0, \mu_1, \phi_1, \sigma_0^2, \sigma_1^2),$$

with

$$(51) \quad p = P[S_t = 0 | S_{t-1} = 0],$$

$$(52) \quad q = P[S_t = 1 | S_{t-1} = 1],$$

being the transition probabilities.

In order to evaluate the log-likelihood function we need to define the joint probabilities  $P[S_t, S_{t-1} | \mathcal{F}_{t-1}]$ .

Because of the Markov property

$$(53) \quad P[S_t | S_{t-1}, \mathcal{F}_{t-1}] = P[S_t | S_{t-1}],$$

we can write

$$(54) \quad P[S_t, S_{t-1} | \mathcal{F}_{t-1}] = P[S_t | S_{t-1}] P[S_{t-1} | \mathcal{F}_{t-1}],$$

and the problem reduces to calculating (estimating) the time dependent state probabilities,  $P[S_{t-1} | \mathcal{F}_{t-1}]$ , and weight them with the transition probabilities to obtain the joint probability.

This can be achieved as follows:

First, let  $P[S_0 = 1 | \mathcal{F}_0] = P[S_0 = 1] = \pi$  be given (then  $P[S_0 = 0] = 1 - \pi$ ).

Then the probabilities  $P[S_{t-1} | \mathcal{F}_{t-1}]$  and the joint probabilities are obtained using the following two steps algorithm

1<sup>0</sup> Given  $P[S_{t-1} = i | \mathcal{F}_{t-1}]$ ,  $i = 0, 1$ , at the beginning of time  $t$  ( $t^{\text{th}}$  iteration),

$$(55) \quad P[S_t = j, S_{t-1} = i | \mathcal{F}_{t-1}] = P[S_t = j | S_{t-1} = i] P[S_{t-1} = i | \mathcal{F}_{t-1}],$$

2<sup>0</sup> Once  $y_t$  is observed, we update the information set  $\mathcal{F}_t = \{\mathcal{F}_{t-1}, y_t\}$  and the probabilities

$$(56) \quad \begin{aligned} P[S_t = j, S_{t-1} = i | \mathcal{F}_t] &= P[S_t = j, S_{t-1} = i | \mathcal{F}_{t-1}, y_t] \\ &= \frac{f(S_t=i, S_{t-1}=j, y_t | \mathcal{F}_{t-1})}{f(y_t | \mathcal{F}_{t-1})} \\ &= \frac{f(y_t | S_t=j, S_{t-1}=i, \mathcal{F}_{t-1}) P[S_t=j, S_{t-1}=i | \mathcal{F}_{t-1}]}{\sum_{s_t, s_{t-1}=0}^1 f(y_t | s_t, s_{t-1}, \mathcal{F}_{t-1}) P[S_t=s_t, S_{t-1}=s_{t-1} | \mathcal{F}_{t-1}]} \end{aligned}$$

with

$$(57) \quad P[S_t = s_t | \mathcal{F}_t] = \sum_{s_{t-1}=0}^1 P[S_t = s_t, S_{t-1} = s_{t-1} | \mathcal{F}_t].$$

Once we have the joint probability for the time point  $t$ , we can calculate the likelihood  $\ell_t(\theta)$ .

The maximum likelihood estimates for  $\theta$  is then obtained iteratively maximizing the likelihood function by updating the likelihood function at each iteration with the above algorithm.

### Steady state probabilities

The probabilities  $\pi = P[S_0 = 1|\mathcal{F}_0]$  is called the steady state probability, and, given the transition probabilities  $p$  and  $q$ , is obtained as

$$(58) \quad \pi = P[S_0 = 1|\mathcal{F}_0] = \frac{1-p}{2-p-q}.$$

Note that in the two state Markov chain

$$(59) \quad P[S_0 = 0|\mathcal{F}_0] = 1 - P[S_0 = 1|\mathcal{F}_0] = \frac{1-q}{2-p-q}.$$

### Smoothed probabilities

Recall that the state  $S_t$  is unobserved.

However, once we have estimated the model, we can make inferences on  $S_t$  using all the information from the sample.

This gives us

$$(60) \quad P[S_t = j|\mathcal{F}_T], \quad j = 0, 1,$$

which are called the smoothed probabilities (for details, see Kim and Nelson 1999, pp. 68–69).

Remark 1.6: In the estimation procedure we derived  $P[S_t = j|\mathcal{F}_t]$  that are usually called the filtered probabilities.

### Expected duration

The expected length the system is going to stay in state  $j$  can be calculated from the transition probabilities. Let  $D$  denote the number of periods the system is in state  $j$ . The probabilities are easily found to be equal to  $P[D = k] = p_{jj}^{k-1}(1 - p_{jj})$ , so that

$$(61) \quad E[D] = \sum_{k=1}^{\infty} kP[D = k] = \frac{1}{1 - p_{jj}}.$$

Note that in our case  $p_{00} = p$  and  $p_{11} = q$ .

Example 1.8: Are there long swings in the dollar/sterling exchange rate?

If the exchange rate  $x_t$  is RW with long swings, it can be modeled as

$$\Delta x_t = \alpha_0 + \alpha_1 S_t + \epsilon_t,$$

so that  $\Delta x_1 \sim N(\mu_0, \sigma_0^2)$  when  $S_t = 0$  and  $\Delta x_t \sim N(\mu_1, \sigma_1^2)$ , when  $S_t = 1$ , where  $\mu_0 = \alpha_0$  and  $\mu_1 = \alpha_0 + \alpha_1$ . Parameters  $\mu_0$  and  $\mu_1$  constitute two different drifts (if  $\alpha_1 \neq 0$ ) in the random walk model.

Estimating the model from quarterly data for sample period 1972I to 1996IV gives

Parameter	Estimate	Std err
$\mu_0$	2.605	0.964
$\mu_1$	-3.277	1.582
$\sigma_0^2$	13.56	3.34
$\sigma_1^2$	20.82	4.79
$p$ (regime 1)	0.857	0.084
$q$ (regime 0)	0.866	0.097

The expected length of stay in regime 0 is given by  $1/(1-p) = 7.0$  quarters, and in regime 1  $1/(1-q) = 7.5$  quarters.

Example 1.9: Suppose we are interested whether the market risk of a share is dependent on the level of volatility on the market. In the CAPM world the market risk of a stock is measured by  $\beta$ .



Consider for the sake of simplicity only the cases of high and low volatility.

The market model is

$$y_t = \alpha_{S_t} + \beta_{S_t}x_t + \epsilon_t,$$

where  $\alpha_{S_t} = \alpha_0(1 - S_t) + \alpha_1 S_t$ ,  $\beta_{S_t} = \beta_0(1 - S_t) + \beta_1 S_t$  and  $\epsilon_t \sim N(0, \sigma_{S_t}^2)$  with  $\sigma_{S_t}^2 = \sigma_0^2(1 - S_t) + \sigma_1^2 S_t$ .

Estimating the model yields

Parameter	Estimate	Std Err	t-value	p-value
$\alpha_0$ (low)	-0.0068	0.0178	-0.39	0.700
$\alpha_1$ (high)	0.0802	0.0508	1.57	0.114
$\beta_0$ (low)	0.9679	0.0215	45.04	0.000
$\beta_1$ (high)	1.8040	0.0690	26.15	0.000
$\sigma_0^2$ (low)	0.5225	0.0198	26.37	0.000
$\sigma_1^2$ (high)	1.7050	0.0711	23.96	0.000

State Prob	
$P(\text{High} \text{High})$	0.96417
$P(\text{Low} \text{High})$	0.03583
$P(\text{High} \text{Low})$	0.01728
$P(\text{Low} \text{Low})$	0.98272
$P(\text{High})$	0.67471
$P(\text{Low})$	0.32529

Log-likelihood -3208.438

The empirical results give evidence that the stock's market risk depends on the level of stock volatility. The expected duration of high volatility is  $1/(1 - .9642) \approx 27$  days, and for low volatility 59 days.

Market returns with high-low volatility probabilities

