Università di Pavia

# GARCH Models
# Estimation and Inference

**Eduardo Rossi**

# LIKELIHOOD FUNCTION

The procedure most often used in estimating $\theta_0$ in ARCH models involves the maximization of a likelihood function constructed under the auxiliary assumption of an *i.i.d.* distribution for the standardized innovation $z_t(\theta)$.

$$\begin{aligned} z_t(\theta) &\equiv \varepsilon_t(\theta)/\sigma_t(\theta) \\ z_t(\theta) &\sim i.i.d.(0,1) \\ z_t(\theta) &\sim f(z_t(\theta);\eta) \end{aligned}$$

$\eta$ is the nuisance parameter, $\eta \in H \subseteq R^k$.
Let $(y_T, y_{T-1}, \ldots, y_1)$ be a sample realization from an ARCH model, and $\psi' \equiv (\theta', \eta')$, the combined $(m+k)\times 1$ parameter vector to be estimated for the conditional mean, variance and density functions.

The log-likelihood function for the $t$-th observation is then given by

$$l_t\left(y_t;\psi\right)=\log\left\{f\left[z_t\left(\theta\right);\eta\right]\right\}-\frac{1}{2}\log\left[\sigma_t^2\left(\theta\right)\right]\qquad t=1,2,\ldots$$

The term $-\frac{1}{2}\ln\left[\sigma_t^2\left(\theta\right)\right]$ is the Jacobiam that arises in the transformation from the standardized innovations, $z_t\left(\theta\right)$ to the observables $y_t$

$$f\left(y_t;\psi\right)=f\left(z_t\left(\theta\right);\eta\right)|J|,\text{ where } J=\frac{\partial z_t}{\partial y_t}=\frac{1}{\sigma_t\left(\theta\right)}$$

# LIKELIHOOD FUNCTION

The log-likelihood function for the full sample:

$$\log L_T\left(y_T, y_{T-1}, \ldots, y_1; \psi\right) = \sum_{t=1}^{T} l_t\left(y_t; \psi\right).$$

The maximum likelihood estimator for the true parameters $\psi_0' \equiv (\theta_0', \eta_0')$, say $\widehat{\psi}_T$ is found by the maximization of the log-likelihood:

$$\widehat{\psi}_T = \arg\max_{\psi} \log L_T(\psi)$$

# Likelihood function

Assuming the conditional density and the $\mu_t(\theta)$ and $\sigma_t^2(\theta)$ functions to be differentiable for all $\psi \in \Theta \times H \equiv \Psi$, the MLE $\widehat{\psi}$ is the solution to

$$S_T\left(y_T, y_{T-1}, \ldots, y_1; \widehat{\psi}\right) \equiv \sum_{t=1}^{T} s_t\left(y_t; \widehat{\psi}\right) = 0$$

where

$$s_t \equiv \frac{\partial l_t(y_t, \psi)}{\partial \psi}$$

is the score vector for the $t$-th observation.

For the conditional mean and variance parameters in $\theta$

$$\frac{\partial l_t(y_t, \psi)}{\partial \theta} = f\left[z_t(\theta); \eta\right]^{-1} f'\left[z_t(\theta); \eta\right] \frac{\partial z_t(\theta)}{\partial \theta} - \frac{1}{2}\left[\sigma_t^2(\theta)\right]^{-1} \frac{\partial \sigma_t^2}{\partial \theta}$$

# Likelihood function

where $f'\left[z_t(\theta);\eta\right] \equiv \dfrac{\partial f\left(z_t(\theta);\eta\right)}{\partial z_t}$ and

$$\begin{aligned}
\frac{\partial z_t(\theta)}{\partial\theta} &= \frac{\partial}{\partial\theta}\left(\frac{\varepsilon_t(\theta)}{\sqrt{\sigma_t^2}}\right) = \frac{\partial}{\partial\theta}\left(\frac{y_t-\mu_t(\theta)}{\sqrt{\sigma_t^2}}\right) \\
&= \frac{-\frac{\partial\mu_t}{\partial\theta}\sqrt{\sigma_t^2}-\frac{1}{2}\left(\sigma_t^2\right)^{-1/2}\frac{\partial\sigma_t^2}{\partial\theta}\varepsilon_t(\theta)}{\sigma_t^2} \\
&= -\frac{\partial\mu_t}{\partial\theta}\left(\sigma_t^2(\theta)\right)^{-1/2}-\frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-3/2}\frac{\partial\sigma_t^2}{\partial\theta}\varepsilon_t(\theta).
\end{aligned}$$

where

$$\varepsilon_t(\theta) \equiv y_t-\mu_t(\theta).$$

In practice the solution to the set of $m+k$ non-linear equations is found by numerical optimization techniques. With the normal distribution:

$$f\left[z_t\left(\theta\right);\eta\right]=\left(2\pi\right)^{-1/2}\exp\left\{-\frac{z_t\left(\theta\right)^2}{2}\right\}$$

the log-likelihood is:

$$l_t=-\frac{1}{2}\log\left(2\pi\right)-\frac{1}{2}z_t\left(\theta\right)^2-\frac{1}{2}\log\left(\sigma_t^2\right)$$

# GAUSSIAN LIKELIHOOD - THE SCORE

It follows that the score vector takes the form:

$$\begin{aligned}
s_t &= -z_t \frac{\partial z_t}{\partial \theta} - \frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-1} \frac{\partial \left(\sigma_t^2(\theta)\right)}{\partial \theta} \\
&= -z_t \left( -\frac{\partial \mu_t}{\partial \theta}\left(\sigma_t^2(\theta)\right)^{-1/2} - \frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-3/2} \frac{\partial \sigma_t^2}{\partial \theta} \varepsilon_t(\theta) \right) \\
&\quad -\frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-1} \frac{\partial \left(\sigma_t^2(\theta)\right)}{\partial \theta} \\
&= \frac{\varepsilon_t(\theta)}{\sqrt{\sigma_t^2}} \frac{\partial \mu_t(\theta)}{\partial \theta} \sigma_t^2(\theta)^{-1/2} + \frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-3/2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\varepsilon_t^2(\theta)}{\sqrt{\sigma_t^2(\theta)}} \\
&\quad -\frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-1} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \\
&= \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\varepsilon_t^2(\theta)}{\left(\sigma_t^2(\theta)\right)^2} - \frac{1}{2} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \left(\sigma_t^2(\theta)\right)^{-1}
\end{aligned}$$

# GAUSSIAN LIKELIHOOD - THE SCORE

$$s_t = \frac{\partial \mu_t(\theta)}{\partial \theta} \frac{\varepsilon_t(\theta)}{\sigma_t^2} + \frac{1}{2} \left( \sigma_t^2(\theta) \right)^{-1} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \left[ \frac{\varepsilon_t^2(\theta)}{(\sigma_t^2(\theta))} - 1 \right]$$

Several other conditional distributions have been employed in the literature to capture the degree of tail fatness in speculative prices.

# NONNORMAL DISTRIBUTIONS - STUDENT'S T DENSITY

Standardized Student's t density:

$$f(z_t;\nu) = c(\nu)\left[1 + \frac{z_t^2}{\nu - 2}\right]^{-\frac{(\nu+1)}{2}}$$

$\nu$ degrees of freedom, with

$$c(\nu) = \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)\sqrt{\pi(\nu-2)}}$$

$\nu > 2$. The condition for a finite moment of order $n$ is $n < \nu$. In particular the kurtosis is finite when $\nu > 4$ and then

$$k = \frac{3(\nu-2)}{(\nu-4)}$$

## NONNORMAL DISTRIBUTIONS - STUDENT'S T DENSITY

As $\nu \to \infty$ the density function converges to $N(0,1)$. The gamma function is defined as

$$\Gamma(u) = \int_0^\infty x^{u-1} e^{-x} dx, \qquad u > 0$$

In the EGARCH(p,q) model:

$$E[|z_t|] = \frac{2\sqrt{\nu-2}\,\Gamma[(\nu+1)/2]}{\sqrt{\pi}(\nu-1)\Gamma[\nu/2]}.$$

The log-likelihood:

$$l_t = -\frac{1}{2}\log(h_t) + \log(c(\nu)) - \frac{\nu+1}{2}\log\left(1 + \frac{z_t^2}{\nu-2}\right)$$

# NONNORMAL DISTRIBUTIONS - GED

$$\begin{aligned} f\left(z_t;\upsilon\right) &= \frac{\upsilon \exp\left[-\left(\frac{1}{2}\right)\left|z_t/\lambda\right|^{\upsilon}\right]}{\lambda 2^{(1+1/\upsilon)}\Gamma\left(1/\upsilon\right)} \qquad -\infty < z_t < \infty, 0 < \upsilon \leq \infty \\ &= C(\upsilon)\exp\left[-\frac{1}{2}\left|\frac{z_t}{\lambda}\right|^{\upsilon}\right] \end{aligned}$$

$$\begin{aligned} \lambda &\equiv \left[2^{(-2/\upsilon)}\Gamma\left(1/\upsilon\right)/\Gamma\left(3/\upsilon\right)\right]^{1/2} \\ C(\upsilon) &= \frac{\upsilon}{2}\left[\frac{\Gamma(3/\upsilon)}{\Gamma(1/\upsilon)^3}\right]^{1/2} \end{aligned}$$

$\upsilon$ is a tail-thickness parameter

$$\begin{aligned} l_t\left(\psi\right) &= \log\left\{f\left[z_t\left(\theta\right);\upsilon\right]\right\} - \frac{1}{2}\log\left[\sigma_t^2\left(\theta\right)\right] \qquad t = 1, 2, \ldots \\ &= \log\left(C(\upsilon)\right) - \frac{1}{2}\left\{\log\left[\sigma_t^2\left(\theta\right)\right] + \left|\frac{z_t}{\lambda}\right|^{\upsilon}\right\} \end{aligned}$$

# NONNORMAL DISTRIBUTIONS - SKEW-T

Skew-T density:

$$f(z_t;\nu,\lambda) = \begin{cases} bc\left(1+\frac{1}{\nu-2}\left(\frac{bz_t+a}{1-\lambda}\right)^2\right)^{-(\nu+1)/2} & \text{for } z_t < -\frac{a}{b} \\ bc\left(1+\frac{1}{\nu-2}\left(\frac{bz_t+a}{1+\lambda}\right)^2\right)^{-(\nu+1)/2} & \text{for } z_t \geq -\frac{a}{b} \end{cases}$$

$$c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi(\nu-2)}}$$

$$b = \sqrt{1+3\lambda^2-a^2}$$

$$a = 4\lambda c\left(\frac{\nu-2}{\nu-1}\right)$$

This density is defined for $2 < \nu < \infty$ and $-1 < \lambda < +1$.

## NONNORMAL DISTRIBUTIONS - SKEW-T

This density encompasses a large set of conventional densities, allowing us to use standard ML tests:

1. if $\lambda$=0, the Skew-$t$ reduces to the traditional Student's $t$ distribution.
2. If $\lambda$=0 and $\nu$=$\infty$ we have the normal density.

Let

$$d_t = (bz_t + a)(1 - \lambda s)$$

where $s$ is a sign dummy

$$s = \begin{cases} 1 & z_t < -a/b \\ -1 & z_t \geq -a/b \end{cases}$$

# NONNORMAL DISTRIBUTIONS - SKEW-T

The log-likelihood contribution is

$$l_t = \log\left(b\right) + \log\left(c\right) - \frac{1}{2}\log\left(\sigma_t^2\left(\theta\right)\right) - \frac{(\nu+1)}{2}\log\left(1 + \frac{d_t^2}{\nu-2}\right)$$

When $\theta = (\alpha', \beta')'$ where $\alpha$ are the conditional mean parameters and $\beta$ are the conditional variance parameters, the score takes the form:

$$s_t = \begin{pmatrix} \frac{\partial l_t}{\partial \alpha} \\ \frac{\partial l_t}{\partial \beta} \end{pmatrix}$$

where

$$\frac{\partial l_t}{\partial \alpha} = \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\varepsilon_t(\alpha)}{\sigma_t^2(\beta)}$$

$$\frac{\partial l_t}{\partial \beta} = \frac{1}{2}\left(\sigma_t^2(\beta)\right)^{-1} \frac{\partial \sigma_t^2(\beta)}{\partial \beta}\left[\frac{\varepsilon_t^2(\alpha)}{\left(\sigma_t^2(\beta)\right)} - 1\right].$$

# Quasi-maximum likelihood

Under regularity conditions, the QML estimator is asymptotically normal distributed with

$$\sqrt{T}\left(\widehat{\theta}_n - \theta_0^*\right) \xrightarrow{d} N\left(0, A^{-1}BA^{-1}\right)$$

The matrices $A$ and $B$ are, respectively, equal to:

$$A = -\frac{1}{T}E_0\left[\frac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta'}\right]$$

$$B = \frac{1}{T}E_0\left[\frac{\partial \log L(\theta)}{\partial\theta}\frac{\partial \log L(\theta)}{\partial\theta'}\right]$$

# QUASI-MAXIMUM LIKELIHOOD

The matrices $A$ and $B$ are not, in general, equal when specification errors are present. Thus comparing estimates of the matrices $A$ and $B$ can be useful for detecting specification errors.

# QUASI-MAXIMUM LIKELIHOOD

The second derivatives matrix of the $t$-th log-likelihood function is equal to:

$$
\begin{aligned}
\frac{\partial^2 l_t}{\partial\theta\partial\theta'} = \; & \frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-2}\frac{\partial\sigma_t^2(\theta)}{\partial\theta}\frac{\partial\sigma_t^2(\theta)}{\partial\theta'} - \frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-1}\frac{\partial^2\sigma_t^2(\theta)}{\partial\theta\partial\theta'} \\
& - \frac{\varepsilon_t^2(\theta)}{\left(\sigma_t^2(\theta)\right)^3}\frac{\partial\sigma_t^2(\theta)}{\partial\theta}\frac{\partial\sigma_t^2(\theta)}{\partial\theta'} \\
& + \frac{1}{2}\frac{\varepsilon_t^2(\theta)}{\left(\sigma_t^2(\theta)\right)^2}\frac{\partial^2\sigma_t^2(\theta)}{\partial\theta\partial\theta'} - \frac{\partial\mu_t(\theta)}{\partial\theta}\frac{\partial\sigma_t^2(\theta)}{\partial\theta'}\frac{\varepsilon_t(\theta)}{\left(\sigma_t^2(\theta)\right)^2} \\
& + \frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)}\frac{\partial^2\mu_t(\theta)}{\partial\theta\partial\theta'} - \left(\sigma_t^2(\theta)\right)^{-1}\frac{\partial\mu_t(\theta)}{\partial\theta}\frac{\partial\mu_t(\theta)}{\partial\theta'} \\
& - \frac{\varepsilon_t(\theta)}{\left(\sigma_t^2(\theta)\right)^2}\frac{\partial\sigma_t^2(\theta)}{\partial\theta}\frac{\partial\mu_t(\theta)}{\partial\theta'}.
\end{aligned}
$$

# Quasi-maximum likelihood

given that $E_{t-1}\left[\frac{\varepsilon_t(\theta)}{\left(\sigma_t^2(\theta)\right)^{1/2}}\right]=0$ and $E_{t-1}\left[\frac{\varepsilon_t^2(\theta)}{\left(\sigma_t^2(\theta)\right)}\right]=1$, we have that

$$\begin{aligned}
A_t &= E_0\left[\frac{\partial^2 l_t}{\partial\theta\partial\theta'}\right] \\
&= E_0\left[\frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-2}\frac{\partial\sigma_t^2(\theta)}{\partial\theta}\frac{\partial\sigma_t^2(\theta)}{\partial\theta'}+\left(\sigma_t^2(\theta)\right)^{-1}\frac{\partial\mu_t(\theta)}{\partial\theta}\frac{\partial\mu_t(\theta)}{\partial\theta'}\right]
\end{aligned}$$

The information matrix for time $t$ is

$$
\begin{aligned}
B_t &= E_0\left[\frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \theta'}\right] \\
&= E_0\left[\frac{\partial \mu_t(\theta)}{\partial \theta}\frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2}\frac{\partial \sigma_t^2(\theta)}{\partial \theta}\frac{\varepsilon_t^2(\theta)}{\left(\sigma_t^2(\theta)\right)^2} - \frac{1}{2}\frac{\partial \sigma_t^2(\theta)}{\partial \theta}\left(\sigma_t^2(\theta)\right)^{-1}\right] \times \\
&\quad \left[\frac{\partial \mu_t(\theta)}{\partial \theta}\frac{\varepsilon_t(\theta)}{\sigma_t^2(\theta)} + \frac{1}{2}\frac{\partial \sigma_t^2(\theta)}{\partial \theta}\frac{\varepsilon_t^2(\theta)}{\left(\sigma_t^2(\theta)\right)^2} - \frac{1}{2}\frac{\partial \sigma_t^2(\theta)}{\partial \theta}\left(\sigma_t^2(\theta)\right)^{-1}\right]'
\end{aligned}
$$

## Quasi-maximum likelihood

$$
\begin{aligned}
B_t \quad = \quad & E_0\left[\begin{array}{c}\frac{1}{4}\frac{1}{\left(\sigma_t^2(\theta)\right)^2}\frac{\partial\sigma_t^2(\theta)}{\partial\theta}\frac{\partial\sigma_t^2(\theta)}{\partial\theta'}\left(K_t\left(\theta\right)-1\right)\\ +\left(\sigma_t^2\left(\theta\right)\right)^{-1}\frac{\partial\mu_t(\theta)}{\partial\theta}\frac{\partial\mu_t(\theta)}{\partial\theta'}\end{array}\right]+\\
& E_0\left[\frac{1}{2}\frac{1}{\left(\sigma_t^2\left(\theta\right)\right)^3}\left(\frac{\partial\sigma_t^2\left(\theta\right)}{\partial\theta}\frac{\partial\mu_t\left(\theta\right)}{\partial\theta'}+\frac{\partial\mu_t\left(\theta\right)}{\partial\theta}\frac{\partial\sigma_t^2\left(\theta\right)}{\partial\theta'}\right)M_{3t}\left(\theta\right)\right].
\end{aligned}
$$

where

$$
\begin{aligned}
M_{3t}\left(\theta\right) \quad &= \quad E_{t-1}\left[\varepsilon_t^3\left(\theta\right)\right]\\
K_t\left(\theta\right) \quad &= \quad \frac{E_{t-1}\left[\varepsilon_t^4\left(\theta\right)\right]}{\left(\sigma_t^2\left(\theta\right)\right)^2}
\end{aligned}
$$

Whenever it is possible to decompose the parameter vector in $\theta = (\alpha', \beta')'$, the hessian matrix for the $t$-th observation is:

$$A_t = \begin{bmatrix} E\left[\left(\sigma_t^2(\beta)\right)^{-1} \frac{\partial \mu_t(\alpha)}{\partial \alpha} \frac{\partial \mu_t(\alpha)}{\partial \alpha'}\right] & 0 \\ 0 & E\left[\frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-2} \frac{\partial \sigma_t^2(\beta)}{\partial \beta} \frac{\partial \sigma_t^2(\beta)}{\partial \beta'}\right] \end{bmatrix}$$

# QUASI-MAXIMUM LIKELIHOOD

The information matrix depends on $M_{3t}$ and $K_t$:

$$B_t = \begin{bmatrix} E\left[\left(\sigma_t^2\left(\theta\right)\right)^{-1}\frac{\partial\mu_t\left(\alpha\right)}{\partial\alpha}\frac{\partial\mu_t\left(\alpha\right)}{\partial\alpha'}\right] & E\left[\frac{1}{2\left(\sigma_t^2\left(\alpha\right)\right)^3}\frac{\partial\mu_t\left(\alpha\right)}{\partial\alpha}\frac{\partial\sigma_t^2\left(\beta\right)}{\partial\beta'}M_{3t}\left(\theta\right)\right] \\ E\left[\frac{1}{2\left(\sigma_t^2\left(\alpha\right)\right)^3}\frac{\partial\sigma_t^2\left(\beta\right)}{\partial\beta}\frac{\partial\mu_t\left(\alpha\right)}{\partial\alpha'}M_{3t}\left(\theta\right)\right] & E\left[\frac{1}{4\left(\sigma_t^2\left(\theta\right)\right)^2}\frac{\partial\sigma_t^2\left(\beta\right)}{\partial\beta}\frac{\partial\sigma_t^2\left(\beta\right)}{\partial\beta'}\left(K_t\left(\theta\right)-1\right)\right] \end{bmatrix}$$

When the true conditional distribution is normal, i.e. $z_t \sim N(0,1)$ and $M_{3t}(\theta) = 0$ and $K_t(\theta) = 3$, the blocks

$$\begin{aligned} B_{12,t} &= \mathbf{0} \\ B_{21,t} &= \mathbf{0} \\ B_{22,t} &= E\left[\frac{1}{2\left(\sigma_t^2(\theta)\right)^2}\frac{\partial \sigma_t^2(\beta)}{\partial \beta}\frac{\partial \sigma_t^2(\beta)}{\partial \beta'}\right] \end{aligned}$$

and the expressions for $A_t$ and $B_t$ coincide.

The asymptotic variance-covariance matrices of the QML estimator $\widehat{\beta}_T$ reduces to:

$$Var^{asy}\left[\sqrt{T}\left(\widehat{\beta}_T - \beta\right)\right] = \left[\frac{1}{T}\sum_t E\left[\frac{1}{2}\left(\sigma_t^2(\theta)\right)^{-2}\frac{\partial \sigma_t^2(\beta)}{\partial \beta}\frac{\partial \sigma_t^2(\beta)}{\partial \beta'}\right]\right]^{-1}.$$

# ASYMPTOTIC RESULTS

For GARCH(1,1):

- Lee and Hansen (1994) and Lumsdaine (1996) proved that the local QMLE is consistent and asympotically normal, assuming that $E(\log(\alpha_1 z_t^2 + \beta_1)) < 0$ (necessary and sufficient condition for strict stationarity).
- Lee and Hansen (1994) required that $E_{t-\tau}[z_t^{2+k}] < \infty$ uniformly with $k > 0$.
- Lumsdaine (1996) required that $E[z_t^{32}] < \infty$.
- Lee and Hansen (1994) showed that the global QMLE is consistent if $\epsilon_t$ is covariance stationary.

# Asymptotic results

For GARCH(p,q):

- Ling and Li (1997) proved that the local QMLE is consistent and asympotically normal if $E[\epsilon_t^4] < \infty$.
- Ling and McAleer (2002) proved the consistency of the global QMLE under only the second-order moment condition.
- Ling and McAleer (2002) derived the asymptotic normality of the global QMLE under the $6th$ moment condition.

# Testing for ARCH disturbances

Test for the presence of ARCH effect. This can be done with a LM test. The test is based upon the score under the null and information matrix under the null. The null hypothesis is

$$\alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$$

Consider the ARCH model with

$$\sigma_t^2 = \sigma^2 (z_t \alpha)$$

where $\sigma^2(\cdot)$ is a differentiable function.

$$z_t = \left(1, \widehat{\epsilon}_{t-1}^2, \ldots, \widehat{\epsilon}_{t-q}^2\right)$$

$\alpha = (\alpha_1, \ldots, \alpha_q)'$ where $\widehat{\epsilon}_t$ are the OLS residuals.

# TESTING FOR ARCH DISTURBANCES

Under the null, $\sigma_t^2$ is a constant $\sigma_t^2 = \sigma_0^2$. The derivative of $\sigma_t^2$ with respect to $\alpha$ is

$$\frac{\partial \sigma_t^2}{\partial \alpha} = \sigma^{2\prime} z_t'$$

where

$$\sigma^{2\prime} = \frac{\partial \sigma^2 \left( z_t \alpha \right)}{\partial \sigma^2}$$

is the scalar derivative of $\sigma^2 \left( z_t \alpha \right)$.

The log-likelihood function is

$$\log L_T = \sum_{t=1}^{T} l_t \left( \alpha \right) = \sum_{t=1}^{T} -\frac{1}{2} \left[ \ln \left( \sigma_t^2 \right) + \frac{\widehat{\epsilon}_t^2}{\sigma_t^2} \right]$$

the derivative of $l_t$ with respect to $\alpha$ is:

$$\frac{\partial l_t}{\partial \alpha} = \frac{\sigma^{2\prime} z_t'}{2 \sigma_t^2} \left[ \frac{\widehat{\epsilon}_t^2}{\sigma_t^2} - 1 \right]$$

# Testing for ARCH disturbances

the score under the null is

$$\frac{\partial \log L_T}{\partial \alpha}\Big|_0 = \frac{\sigma^{2\prime}}{2\sigma_0^2}\sum_t z_t' \left(\frac{\widehat{\epsilon}_t^2}{\sigma_0^2} - 1\right) = \frac{\sigma^{2\prime}}{2\sigma_0^2} Z' f^0$$

where

$$f^0 = \left[\left(\frac{\widehat{\epsilon}_1^2}{\sigma_0^2} - 1\right), \ldots, \left(\frac{\widehat{\epsilon}_T^2}{\sigma_0^2} - 1\right)\right]'$$

and

$$Z' = (z_1', \ldots, z_T')$$

is a $((q+1) \times T)$ matrix. The second derivatives matrix is

$$\begin{aligned}
\frac{\partial^2 l_t}{\partial \alpha \partial \alpha'} &= -\frac{\sigma^{2\prime} z_t' \sigma^{2\prime} z_t}{2\sigma_t^4}\left[\frac{\widehat{\epsilon}_t^2}{\sigma_t^2} - 1\right] + \frac{\sigma^{2\prime} z_t'}{2\sigma_t^2}\left[\frac{-\sigma^{2\prime} z_t + \widehat{\epsilon}_t^2}{\sigma_t^4}\right] \\
&= -\left(\frac{\sigma^{2\prime}}{\sigma_t^2}\right)^2 \frac{\widehat{\epsilon}_t^2}{\sigma_t^2} z_t' z_t + \frac{1}{2}\left(\frac{\sigma^{2\prime}}{\sigma_t^2}\right)^2 z_t' z_t
\end{aligned}$$

# TESTING FOR ARCH DISTURBANCES

This yields the information matrix under the null:

$$
\begin{aligned}
A_{\alpha\alpha,0} &= -\frac{1}{T}E\left[\frac{\partial^2 \log L_T}{\partial\alpha\partial\alpha'}\right]|_0 = -\frac{1}{T}E\left[E\left[\sum\frac{\partial^2 l_t}{\partial\alpha\partial\alpha'}\,|\Phi_{t-1}\right]\right]|_0 \\
&= -\frac{1}{T}E\left[\sum E\left[\frac{\partial^2 l_t}{\partial\alpha\partial\alpha'}\,|\Phi_{t-1}\right]\right]|_0 \\
&= -\frac{1}{T}E\left[\sum E\left[-\left(\frac{\sigma^{2\prime}}{\sigma_t^2}\right)^2\frac{\widehat{\epsilon}_t^2}{\sigma_t^2}z_t'z_t+\frac{1}{2}\left(\frac{\sigma^{2\prime}}{\sigma_t^2}\right)^2 z_t'z_t\,|\Phi_{t-1}\right]\right]|_0 \\
&= \frac{1}{T}\sum_{t=1}^{T}\left\{-\frac{1}{2}\left(\frac{\sigma^{2\prime}}{\sigma_0^2}\right)^2 E\left[z_t'z_t\right]+\left(\frac{\sigma^{2\prime}}{\sigma_0^2}\right)^2 E\left[z_t'z_t\right]\right\} = \\
&= \frac{1}{2}\left(\frac{\sigma^{2\prime}}{\sigma_0^2}\right)^2\frac{1}{T}\sum_{t=1}^{T}E\left[z_t'z_t\right].
\end{aligned}
$$

# TESTING FOR ARCH DISTURBANCES

The LM statistic is given by

$$\xi_{LM} = \frac{1}{T}\left(\frac{\partial \log L_T}{\partial \alpha}|_0\right)' A^{-1}_{\alpha\alpha,0}\left(\frac{\partial \log L_T}{\partial \alpha}|_0\right)$$

the LM statistic is

$$\begin{aligned}\xi_{LM} &= f^{0\prime} Z\frac{\sigma^{2\prime}}{2\sigma_0^2}\left[\frac{1}{2}\left(\frac{\sigma^{2\prime}}{\sigma_0^2}\right)^2\sum_{t=1}^{T} E\left[z_t' z_t\right]\right]^{-1}\frac{\sigma^{2\prime}}{2\sigma_0^2}Z' f^0 \\ &= f^{0\prime} Z\left(\sum_{t=1}^{T} E\left[z_t' z_t\right]\right)^{-1} Z' f^0/2\end{aligned}$$

it can be consistently estimated by

$$\xi_{LM} = \frac{f^{0\prime} Z\left(Z' Z\right)^{-1} Z' f^0}{2}.$$

# Testing for ARCH disturbances

When we assume normality $p\lim\left(f^{0\prime}f^0/T\right) = 2$. Thus an asymptotically equivalent statistic would be

$$\xi^* = \frac{Tf^{0\prime}Z\left(Z'Z\right)^{-1}Z'f^0}{\left(f^{0\prime}f^0\right)} = TR^2$$

where $R^2$ is the squared multiple correlation between $f^0$ and $Z$. Since adding a constant and multiplying by a scalar will not change the $R^2$ of a regression, this is also the $R^2$ of the regression of $\widehat{\epsilon}_t^2$ on an intercept and $q$ lagged values of $\widehat{\epsilon}_t^2$.

The statistic will be asymptotically distributed as chi square with $q$ degrees of freedom when the null hypothesis is true.

The test procedure is to run the OLS regression and save the residuals. Regress the squared residuals on a constant and $q$ lags and test $TR^2$ as a $\chi_q^2$. This will be an asymptotically locally most powerful test.

# TEST FOR ASYMMETRIC EFFECTS

Engle and Ng (1993) put forward three diagnostic tests for volatility models:

1. the *Sign Bias Test*
2. the *Negative Size Bias Test*
3. the *Positive Size Bias Test*.

# TEST FOR ASYMMETRIC EFFECTS

These tests examine whether we can predict the squared normalized residual by some variables observed in the past which are not included in the volatility model being used. If these variables can predict the squared normalized residual, then the variance model is misspecified.

- The *sign bias test* examines the impact of positive and negative return shocks on volatility not predicted by the model under consideration.
- The *negative size bias test* focuses on the different effects that large and small negative return shocks have on volatility which are not predicted by the volatility model.
- The *positive size bias test* focuses on the different impacts that large and small positive return shocks may have on volatility, which are not explained by the volatility model.

# TEST FOR ASYMMETRIC EFFECTS

To derive the optimal form of these tests, we assume that the volatility model under the null hypothesis is a special case of a more general model of the following form:

$$\log\left(\sigma_t^2\right) = \log\left(\sigma_{0t}^2\left(\delta_0' z_{0t}\right)\right) + \delta_a' z_{at}$$

where $\sigma_{0t}^2\left(\delta_0' z_{0t}\right)$ is the volatility model hypothesized under the null, $\delta_0$ is a $(k \times 1)$ vector of parameters under the null, $z_{0t}$ is a $(k \times 1)$ vector of explanatory variables under the null, $\delta_a$ is a $(m \times 1)$ vector of additional parameters, $z_{at}$ is a $(m \times 1)$ vector of missing explanatory variables:

$$H_0 : \delta_a = \mathbf{0}$$

# TEST FOR ASYMMETRIC EFFECTS

This form encompasses both the GARCH and EGARCH models. For the GARCH(1,1) model

$$\sigma^2_{0t}\left(\delta_0' z_{0t}\right) = \delta_0' z_{0t}$$

$$z_{0t} \equiv \left[1, \sigma^2_{t-1}, \varepsilon^2_{t-1}\right]'$$

$$\delta_0 \equiv [\omega, \beta, \alpha]'$$

$$\delta_a = [\beta^*, \phi^*, \psi^*]'$$

$$z_{at} = \left[\log\left(\sigma^2_{t-1}\right), \frac{\varepsilon_{t-1}}{\sigma_{t-1}}, \left(\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - \sqrt{2/\pi}\right)\right]'$$

# TEST FOR ASYMMETRIC EFFECTS

The encompassing model is

$$\begin{aligned}\log\left(\sigma_t^2\right) &= \log\left[\omega+\beta\sigma_{t-1}^2+\alpha\varepsilon_{t-1}^2\right]+\beta^*\log\left(\sigma_{t-1}^2\right)\\ &\quad +\phi^*\frac{\varepsilon_{t-1}}{\sigma_{t-1}}+\psi^*\left(\frac{\left|\varepsilon_{t-1}\right|}{\sigma_{t-1}}-\sqrt{2/\pi}\right)\end{aligned}$$

when $\alpha=\beta=0$ is an EGARCH(1,1) while with $\beta^*=\phi^*=\psi^*=0$ is a GARCH(1,1) model.

The null hypothesis is $\delta_a=0$. Let $\upsilon_t$ be the normalized residual corresponding to observation $t$ under the volatility model hypothesized:

$$\upsilon_t \equiv \frac{\varepsilon_t}{\sigma_t}$$

# TEST FOR ASYMMETRIC EFFECTS

The LM test statistic for $H_0 : \delta_a = 0$ is a test of $\delta_a = 0$ in the auxiliary regression

$$v_t^2 = z_{0t}^{*\prime}\delta_0 + z_{at}^{*\prime}\delta_a + u_t$$

where

$$\begin{aligned} z_{0t}^{*} &\equiv \sigma_{0t}^{-2}\left(\frac{\partial \sigma_t^2}{\partial \delta_0}\right) \\ z_{at}^{*} &\equiv \sigma_{0t}^{-2}\left(\frac{\partial \sigma_t^2}{\partial \delta_a}\right) \end{aligned}$$

Both $\dfrac{\partial \sigma_t^2}{\partial \delta_0}$ and $\dfrac{\partial \sigma_t^2}{\partial \delta_a}$ are evaluated at $\delta_a = 0$ and $\widehat{\delta}_0$ (the maximum likelihood estimator of $\delta_0$ under $H_0$)

# Test for Asymmetric Effects

The derivatives are

$$\frac{\partial \sigma_t^2}{\partial \delta_0} = \frac{\partial [\sigma_{0t}^2\,(\delta_0' z_{0t})\, e^{\delta_a' z_{at}}]}{\partial \delta_0} = \frac{\partial \sigma_{0t}^2}{\partial \delta_0} e^{\delta_a' z_{at}}$$

$$\frac{\partial \sigma_t^2}{\partial \delta_a} = \frac{\partial [\sigma_{0t}^2\,(\delta_a' z_{0t})\, e^{\delta_a' z_{at}}]}{\partial \delta_a} = \sigma_{0t}^2\,(\delta_0' z_{0t})\, e^{\delta_a' z_{at}}\, z_{at}$$

$$\left.\frac{\partial \sigma_t^2}{\partial \delta_0}\right|_{\delta_0=\widehat{\delta}_0, \delta_a=\mathbf{0}} = \left.\frac{\partial \sigma_{0t}^2}{\partial \delta_0}\right|_{\delta_0=\widehat{\delta}_0} e^{\delta_a' z_{at}}$$

$$\left.\frac{\partial \sigma_t^2}{\partial \delta_a}\right|_{\delta_0=\widehat{\delta}_0, \delta_a=\mathbf{0}} = \sigma_{0t}^2 \left(\widehat{\delta}_0' z_{0t}\right) z_{at}$$

# TEST FOR ASYMMETRIC EFFECTS

If the parameters restrictions are met, the right-hand side variables should have no explanatory variables power at all. Thus, the test is often computed as

$$\xi_{LM} = TR^2$$

where $R^2$ is the squared multiple correlation of auxiliary regression, and $T$ is the number of observations in the sample.

Under the encompassing model, $\left(\frac{\partial \sigma_t^2}{\partial \delta_a}\right)$ evaluated under the null is equal to $\sigma_{0t}^2 z_{at}$, hence $z_{at}^* = z_{at}$. The regression actually involves regressing $v_t^2$ on a constant $z_{0t}^*$ and $z_{at}$. The variables in $z_{at}$ are $S_{t-1}$, $S_{t-1}^- \varepsilon_{t-1}$ and $S_{t-1}^+ \varepsilon_{t-1}$.

# TEST FOR ASYMMETRIC EFFECTS

The optimal form for conducting the *sign bias test* is:

$$v_t^2 = a + b_1 S_{t-1}^{-} + \gamma' z_{0t}^{*} + e_t$$

where

$$S_{t-1}^{-} = \begin{cases} 1 & \varepsilon_{t-1} < 0 \\ 0 & otherwise \end{cases}$$

the regression for the *negative size bias test* is:

$$v_t^2 = a + b_2 S_{t-1}^{-} \varepsilon_{t-1} + \gamma' z_{0t}^{*} + e_t$$

the *positive size bias test statistic*:

$$v_t^2 = a + b_3 S_{t-1}^{+} \varepsilon_{t-1} + \gamma' z_{0t}^{*} + e_t$$

$$S_{t-1}^{+} = \begin{cases} 1 & \varepsilon_{t-1} > 0 \\ 0 & otherwise \end{cases}$$

# TEST FOR ASYMMETRIC EFFECTS

The t-ratios for $b_1$, $b_2$ and $b_3$ are the sign bias, the negative size bias, and the positive size bias test statistics, respectively.

The joint test is the LM test for adding the three variables in the variance equation under the maintained specification:

$$v_t^2 = a + b_1 S_{t-1}^- + b_2 S_{t-1}^- \varepsilon_{t-1} + b_3 S_{t-1}^+ \varepsilon_{t-1} + \gamma' z_{0t}^* + e_t$$

The test statistics is $TR^2$. If the volatility model is correct then $b_1 = b_2 = b_3 = 0$, $\gamma = 0$ and $e_t$ is i.i.d. If $z_{0t}^*$ is not included the test will be conservative; the size will be less than or equal to the nominal size, and the power may be reduced.