GARCH Models

Topics:

- Volatility: historical, RiskMetrics
- The ARCH and GARCH models
- Estimation of GARCH models
- Testing of GARCH models
- Asymmetry and the news impact curve
- GARCH-in-mean
- Non-Gaussian Likelihoods for GARCH models.
- Specification Testing in GARCH models
- Volatility forecasting

Volatility: Introduction

Half of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2003 was awarded to Robert F. Engle III *"for methods of analyzing economic time series with time-varying volatility (ARCH)"*

Volatility: Introduction

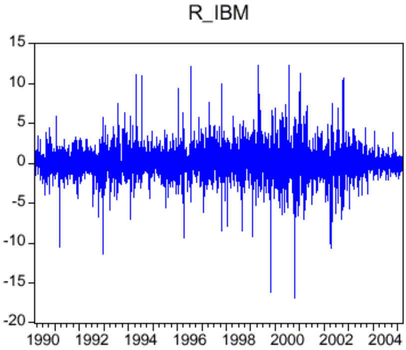
The *volatility* of an investment is a measure of its *risk*. Usually defined as the variance of the returns on the investment.
Volatility is an important ingredient in:

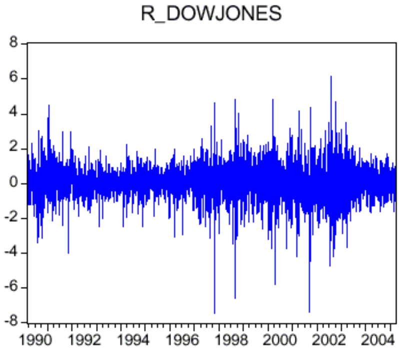
- portfolio selection;
- risk management;
- option pricing.

Empirical Evidence: Daily financial returns display *volatility clustering*: periods of high volatility alternate with more tranquil periods.

- This forms the basis for the *Autoregressive-Conditional Heteroskedasticity model* (Robert Engle) and the *Generalized Autoregressive-Conditional Heteroskedasticity* model (Tim Bollerslev).

Daily log-returns on IBM stock price and Dow Jones index, March 1990 – March 2005





Historical Volatility

- A first simple estimator is *historical volatility*, i.e., the sample variance of the most recent m observations (often $m = 250$, one year).
- Denote R_t the daily log-return, that is if P_t is the investment in period $t : R_t = \Delta \log(P_t)$.
- *Historical volatility* is defined as

$$\sigma^2_{t,HIST} = \frac{1}{m} \sum_{j=0}^{m-1} R^2_{t-1-j}.$$

  (Typically its sample average is very close to zero). This is an estimate of the volatility over day t, made at the end of day $t-1$.)
- **Main disadvantages:**
  - either noisy (small m), or reacts slowly to new information (large m);
  - "ghosting" feature: large shock leads to higher volatility for exactly m periods, then drops out.

Problems with historical volatility are addressed by replacing equally weighted moving average by an *exponentially weighted moving average* (EWMA), also used in JPMorgan's *RiskMetrics* system:

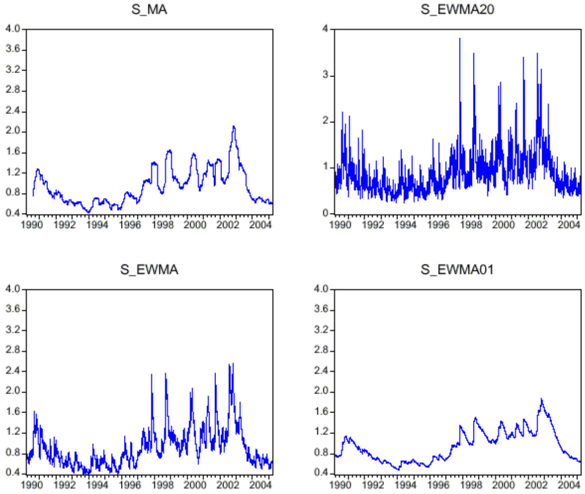
$$\begin{aligned}\sigma^2_{t,EWMA} &= (1-\lambda)\sum_{j=0}^{\infty}\lambda^j R^2_{t-1-j} \\ &= \lambda\sigma^2_{t-1,EWMA} + (1-\lambda)R^2_{t-1}, \qquad 0<\lambda<1.\end{aligned}$$

This means that observations further in the past get a smaller weight.
Remarks:

- In practice we do not have $R_{t-\infty}$, but the second equation can be started up by an initial estimate / guess $\sigma^2_{0,EWMA}$.
- For daily data, RiskMetrics recommends $\lambda = 0.94$.

Volatility: historical, RiskMetrics

Historical (MA) and EWMA volatilities ($\lambda = 0.8, 0.94, 0.99$) of DJ index



The ARCH(1) model

The $AR(1)$ model $Y_t = c + \phi Y_{t-1} + \varepsilon_t$ (assuming that ε_t is a martingale difference sequence) can be formulated as $E_{t-1}(Y_t) = c + \phi Y_{t-1}$ where $E_{t-1}(.) = E(.|F_{t-1})$ where F_{t-1} is the information until and including period $t-1$. The first-order *autoregressive-conditional heteroskedasticity* (ARCH(1)) model for a return R_t is an AR(1) for R_t^2:

$$E_{t-1}(R_t^2) = \omega + \alpha R_{t-1}^2,$$

assuming that $E_{t-1}(R_t) = 0$. This model can be written as

$$\begin{aligned} \sigma_t^2 &= var_{t-1}(R_t) \\ &= E_{t-1}(R_t^2) \\ &= \omega + \alpha R_{t-1}^2. \end{aligned}$$

where $var_{t-1}(.) = var(.|F_{t-1})$. In practice we need to allow for $E_{t-1}(R_t) = \mu_t \neq 0$. Then $R_t = \mu_t + u_t$, and the model is AR(1) for u_t^2:

$$\sigma_t^2 = var_{t-1}(R_t) = E_{t-1}(u_t^2) = \omega + \alpha u_{t-1}^2.$$

Properties of ARCH(1)

It is possible to show that u_t is stationary if and only if $0 \leq \alpha < 1$. The variance is

$$\sigma^2 = var(u_t) = \frac{\omega}{1 - \alpha}$$

Defining $\eta_t = u_t^2 - \sigma_t^2$, the ARCH(1) model is an AR(1) model for u_t^2 :

$$u_t^2 = \omega + \alpha u_{t-1}^2 + \eta_t,$$

with $E_{t-1}(\eta_t) = 0$, hence $E(\eta_t) = 0$ and $cov(\eta_t, \eta_{t-l}) = 0$ for $l \geq 1$. Assuming that $var(\eta_t)$ is finite u_t^2 is stationary.

The ARCH(q) model

When trying to estimate ARCH models one might find that more lags are needed, leading to ARCH(q):

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

Note: Variances must be positive, a sufficient condition is $\omega > 0$, $\alpha_i \geq 0, i = 1, \ldots, q$.

- A necessary and sufficient condition for stationarity of u_t is that $\sum_{i=1}^{q} \alpha_i < 1$ [and $\omega > 0, \alpha_i \geq 0, i = 1, \ldots, q$.].
- Corresponds to an $AR(q)$ model for u_t^2 :

$$u_t^2 = \omega + \alpha_1 u_{t-1}^2 + ... + \alpha_q u_{t-q}^2 + \eta_t,$$

  with $\eta_t = u_t^2 - \sigma, E_{t-1}(\eta_t) = 0$, hence $E(\eta_t) = 0$ and $cov(\eta_t, \eta_{t-l}) = 0$ for $l \geq 1$ and assuming that $var(\eta_t)$ is finite.
- Note that the roots of $A(z) = 1 - \alpha_1 z - \alpha_2 z^2 - ... - \alpha_q z^q$ must be outside the unit circle for u_t^2 to be stationary.
- If $\omega > 0, \alpha_i \geq 0, i = 1, \ldots, q$ this is equivalent to $\sum_{i=1}^{q} \alpha_i < 1$.
- We can identify the model by checking the ACF and PACF of u_t^2

The GARCH(1,1) model

- A simpler structure than ARCH(q) is an ARMA(1,1) for R_t^2 or u_t^2, which leads to the *generalized ARCH* model of orders (1,1) (GARCH(1,1)):

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2, \qquad \omega > 0, \alpha \geq 0, \beta \geq 0.$$

- For $\beta < 1$ this is equivalent to an ARCH(∞) model

$$\sigma_t^2 = \frac{\omega}{1-\beta} + \sum_{j=0}^{\infty} \beta^j \alpha u_{t-1-j}^2.$$

- **Advantage:** We have less parameters to estimate.

The GARCH(1,1) model

- The GARCH(1,1) model is stationary if and only if $\alpha + \beta < 1$ and $\omega > 0, \alpha \geq 0, \beta \geq 0$. Thus

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}.$$

- It corresponds to an ARMA(1,1) model for u_t^2 :

$$u_t^2 = \omega + (\alpha + \beta)u_{t-1}^2 + \eta_t - \beta\eta_{t-1}^2$$

with $\eta_t = u_t^2 - \sigma_t^2 . E_{t-1}(\eta_t) = 0$, hence $E(\eta_t) = 0$ and $cov(\eta_t, \eta_{t-l}) = 0$ for $l \geq 1$. We assume that $var(\eta_t)$ is finite.

- The ACF and PACF of u_t^2 in case of stationary GARCH(1,1) are both exponentially decaying, no cut-off point;

GARCH(q,s) and IGARCH

Further generalisation GARCH (q,s)

$$\sigma_t^2 = \omega + \sum_{\ell=1}^{q} \alpha_l u_{t-\ell}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \omega > 0, \alpha_l \geq 0, \beta_j \geq 0$$

Remarks:

- Corresponds to $ARMA(p,s)$ for u_t^2 with $p = \max\{q,s\}$
- Stationary if and only if $\sum_{l=1}^{q} \alpha_l + \sum_{j=1}^{s} \beta_j < 1$ [and $\omega > 0, \alpha_l \geq 0, \beta_j \geq 0$].
- The acf and pacf of u_t^2 in case of stationary GARCH(q,s) are both exponentially decaying, no cut-off point;
- $\omega > 0, \alpha_l \geq 0, \beta_j \geq 0$ sufficient (not necessary) for $\sigma_t^2 > 0$.
- Model with unit root ($\alpha_1 + \beta_1 = 1$ in case of GARCH(1,1)): *Integrated GARCH (IGARCH)*: infinite variance, no mean-reversion in volatility.
- IGARCH(1,1) with $\omega = 0$ and $R_t = u_t$ leads to RiskMetrics / EWMA.

$$\sigma_t^2 = (1 - \beta_1) R_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

The GARCH model

Some other properties:

- Let $\sigma_t = \sqrt{\sigma_t^2}$.The *standardized returns*

$$\varepsilon_t = \frac{R_t - \mu_t}{\sigma_t} = \frac{R_t - E_{t-1}(R_t)}{\sqrt{var_{t-1}(R_t)}}$$

  satisfy $E_{t-1}(\varepsilon_t) = 0$ and $var_{t-1}(\varepsilon_t) = 1$. Therefore the model may be formulated as

$$\begin{aligned} R_t &= \mu_t + u_t = \mu_t + \underbrace{\sigma_t \varepsilon_t}_{u_t}, \\ \sigma_t^2 &= \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2. \end{aligned}$$

- Often it is assumed that ε_t are independent and identically distributed as $N(0, 1)$.

Kurtosis of GARCH models

For any weakly stationary process x_t with constant $E(x_t^4)$ define kurtosis as

$$k_x = \frac{E[(x_t - E(x_t))^4]}{var(x_t)^2}.$$

Measures the fatness of the tails of a probability density function.

- If $x_t \sim N(\mu_x, \sigma_x^2)$, then $k_x = 3$. Excess Kurtosis $K_x = k_x - 3$.
- *Empirical evidence:* financial data often displays more kurtosis than that permitted under the assumption of normality (fat tails.)
- $u_t = \sigma_t \varepsilon_t$ with $E(\varepsilon_t) = 0$ $var(\varepsilon_t) = 1$ and assuming that ε_t are i.i.d.:

  $$k_u = \frac{E[\sigma_t^4 {\varepsilon_t}^4]}{var(\sigma_t \varepsilon_t)^2} = \frac{E[\sigma_t^4]}{E[\sigma_t^2]^2} k_\varepsilon \geq k_\varepsilon.$$

  because by Jensen inequality $E(h(X)) \geq h(E(X))$ for a random variable X if $h(.)$ is a convex function.
- Thus if $\varepsilon_t \sim N(0,1)$ $k_u \geq 3$.
- Even if $\varepsilon_t \sim N(0,1)$, varying σ_t implies that R_t has non-normal distribution, with higher kurtosis.

Estimation of GARCH models

- GARCH cannot be estimated by ordinary least-squares. These models are estimated by *conditional maximum likelihood*.
- General Gaussian GARCH model

$$R_t = \mu_t(\phi) + \underbrace{\sigma_t(\phi,\psi)\varepsilon_t}_{u_t}$$

- Under the assumption that $\varepsilon_t \sim$ i.i.d. $N(0,1)$, and ϕ are the regression/ARMA parameters and ψ are the GARCH parameters.
- **Example:** AR(1)-GARCH(1,1) so

$$\mu_t = \phi_0 + \phi_1 R_{t-1},\ \sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta\sigma_{t-1}^2$$

  hence $\phi = (\phi_0, \phi_1)'$ and $\psi = (\omega, \alpha, \beta)'$
- σ_t^2 depends on ϕ through $u_{t-1}^2 = (R_{t-1} - \mu_{t-1}(\phi))^2$.
- Conditional density $R_t|F_{t-1} \sim N(\mu_t, \sigma_t^2)$

$$f(R_t|F_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2(\phi,\psi)}} \exp\left(-\frac{1}{2}\frac{(R_t - \mu_t(\phi))^2}{\sigma^2(\phi,\psi)}\right),$$

Estimation of GARCH models

If we assume that $\mu_t(\phi)$ and $\sigma_t^2(\phi,\psi)$ depend on at most m lags of R_t, u_t and σ_t^2

Conditional joint density

$$f(R_T, ..., R_{m+1}|F_m) = (\prod_{t=m+1}^{T} f(R_t|F_{t-1}))$$

- The conditional log-likelihood:

$$\begin{aligned} \mathcal{L}^*(\gamma|F_m) &= \sum_{t=m+1}^{T} \ell_t(\gamma) \\ \ell_t(\gamma) &= \log f(R_t|F_{t-1}), \\ \gamma &= (\phi', \psi')' \end{aligned}$$

  Starting values for σ_t^2 and $u_t, t = 1, ..., m$

$$\begin{aligned} u_t &= 0, \\ \sigma_t^2 &= \frac{1}{T-m} \sum_{t=m+1}^{T} (R_t - \mu_t(\phi))^2 \end{aligned}$$

- Maximization of $\mathcal{L}^*(\gamma|F_m)$ can be done by numerical optimization algorithms.

Asymptotic Results

- Under general conditions (stationarity, existence of moments)

$$\sqrt{T}(\hat{\gamma}_{CML} - \gamma_0) \xrightarrow{D} N(0, A_0^{-1})$$

where $A_0 = E[-\frac{\partial^2 \ell_t(\gamma_0)}{\partial\gamma\partial\gamma'}]$ if the assumption of gaussianity holds (where γ_0) are the true values of the parameters.

- If we are not sure that $R_t|F_{t-1}$ are normally distributed and you use it anyway then we may still use the same estimation technique.
- This is called *quasi-maximum likelihood estimator (QML).*
- *QML* is consistent under correct specification of both the conditional mean and the conditional variance.
- In this case

$$\sqrt{T}(\hat{\gamma}_{CML} - \gamma_0) \xrightarrow{D} N(0, \mathcal{V}),$$
$$\mathcal{V} = A_0^{-1} B_0 A_0^{-1}$$

where $B_0 = E[\frac{\partial \ell_t(\gamma_0)}{\partial\gamma} \frac{\partial \ell_t(\gamma_0)}{\partial\gamma'}]$.

Asymptotic Results

- Standard errors of $\hat{\gamma}_{CML}$ can be obtained from the sandwich estimator for the asymptotic variance of $\hat{\theta}$:

$$\hat{\mathcal{V}} = \hat{A}_0^{-1}\hat{B}_0\hat{A}_0^{-1} \xrightarrow{p} \mathcal{V}$$

where

$$\begin{aligned}
\hat{A}_0 &= \frac{1}{T}\sum\nolimits_{t=m+1}^{T}[-\frac{\partial^2 \ell_t(\hat{\gamma}_{CML})}{\partial\theta\partial\theta'}], \\
\hat{B}_0 &= \frac{1}{T}\sum\nolimits_{t=m+1}^{T}[\frac{\partial \ell_t(\hat{\gamma}_{CML})}{\partial\theta}\frac{\partial \ell_t(\hat{\gamma}_{CML})}{\partial\theta'}]
\end{aligned}$$

This in turn can be used to construct t-tests.

- The standard errors computed via this robust method are known (in this context) as *Bollerslev-Wooldridge standard errors*.

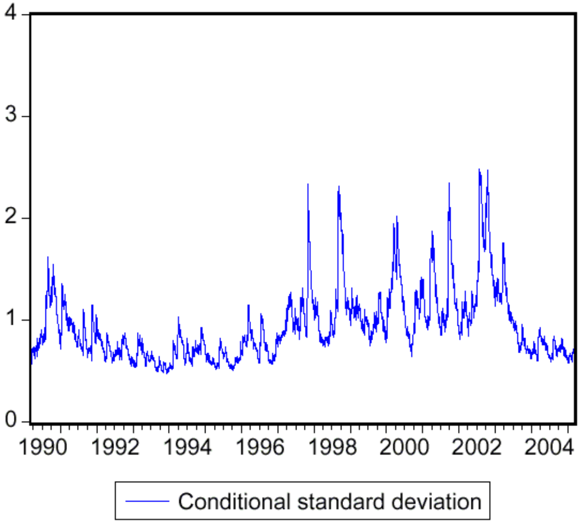
Example:

Dependent Variable: Y
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 12/03/1990 11/03/2005
Included observations: 3915 after adjustments
Convergence achieved after 15 iterations
Bollerslev-Wooldrige robust standard errors & covariance
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.053344	0.012855	4.149691	0.0000
Variance Equation				
C	0.008021	0.002392	3.353844	0.0008
RESID(-1)^2	0.058619	0.010235	5.727463	0.0000
GARCH(-1)	0.934423	0.009392	99.48822	0.0000

Estimation of GARCH models

- Fitted values for the variances



Asymmetry and the news impact curve

- The *news impact curve* (NIC) is the effect of u_t on σ_{t+1}^2, keeping σ_t^2 and the past fixed.
- For GARCH(1,1), this is the parabola $NIC(u_t|\sigma_t^2 = \sigma^2) = A + \alpha u_t^2$, with $A = \omega + \beta\sigma^2$. This has a minimum at $u_t = 0$, and is symmetric around that minimum.
- Often this is unrealistic: a large negative shock (stock market crash) is expected to increase volatility much more than a large positive shock.

Asymmetry and the news impact curve

This is known as *leverage effect*:

- Example:

$$\begin{aligned} &\downarrow \quad \text{value of firm's stock} \\ \Rightarrow \quad &\downarrow \text{equity value of the firm} \\ \Rightarrow \quad &\uparrow \text{debt-to-equity ratio } (\textit{leverage}) \\ \Rightarrow \quad &\text{shareholders perceive} \\ &\text{future cashflows more risky.} \end{aligned}$$

- Two popular proposals to deal with this issue
  - Nelson's exponential GARCH (EGARCH);
  - Glosten, Jagannathan and Runkle's GJR-GARCH (also known as threshold GARCH, TGARCH.)

EGARCH

The EGARCH(1,1) model is

$$\log \sigma_t^2 = (1-\beta)\omega + \gamma\varepsilon_{t-1} + \alpha(|\varepsilon_{t-1}| - E|\varepsilon_{t-1}|) + \beta \log \sigma_{t-1}^2,$$

with $\varepsilon_t = u_t/\sigma_t$ as usual. If $\varepsilon_t \sim$ i.i.d. $N(0,1)$ then $E|\varepsilon_t| = \sqrt{2/\pi}$.

Properties:

- NIC is asymmetric if and only if $\gamma \neq 0$; leverage effect if $\gamma < 0$;
- σ_t^2 is positive for all parameter values;
- $\gamma\varepsilon_{t-1} + \alpha(|\varepsilon_{t-1}| - E|\varepsilon_{t-1}|)$ is an i.i.d. mean-zero shock to log-volatility;
- if $|\beta| < 1$, $\log \sigma_t^2$ is stationary with mean ω;

GJR-GARCH (or TARCH, threshold GARCH)

The GJR-GARCH(1,1) model is

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \gamma u_{t-1}^2 \mathrm{I}_{t-1} + \beta \sigma_{t-1}^2 .$$

where

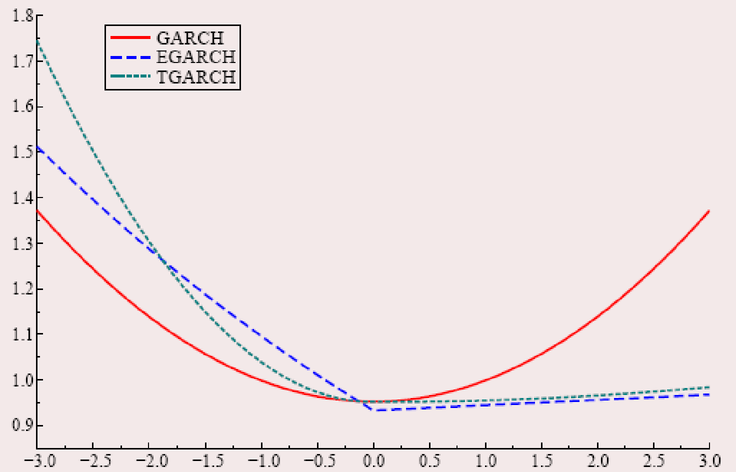
$$\mathrm{I}_{t-1} = \begin{cases} 1 & \text{if} \quad u_{t-1} < 0 \\ 0 & \text{if} \quad u_{t-1} \geq 0 \end{cases} .$$

Distribution of u_t is *symmetric*.

Properties:

- NIC is asymmetric if and only if $\gamma \neq 0$; leverage effect if $\gamma > 0$;
- σ_t^2 is positive if $\omega > 0$, $\alpha \geq 0$, $\gamma \geq 0$, $\beta \geq 0$;
- u_t^2 is stationary if $0 \leq \alpha + \frac{1}{2}\gamma + \beta < 1$, with unconditional variance $\sigma^2 = \omega / \left[1 - \alpha - \frac{1}{2}\gamma - \beta\right]$.
- For both EGARCH and TGARCH, distribution of one-period returns $r_t = u_t$ is symmetric, but multi-period returns $r_t[k] = r_t + ... + r_{t-k+1}$ have an asymmetric distribution.

Example: Estimated NIC for S&P 500 index



GJR-GARCH (or TARCH, threshold GARCH)

Output:

Dependent Variable: Y
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 12/03/1990 11/03/2005
Included observations: 3915 after adjustments
Convergence achieved after 17 iterations
Bollerslev-Wooldrige robust standard errors & covariance
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-1)^2*(RESID(-1)<0) +
C(5)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.031402	0.012706	2.471433	0.0135
Variance Equation				
C	0.012293	0.002949	4.168243	0.0000
RESID(-1)^2	0.009436	0.009670	0.975799	0.3292
RESID(-1)^2*(RESID(-1)<0)	0.091836	0.017677	5.195257	0.0000
GARCH(-1)	0.931539	0.009270	100.4872	0.0000

GARCH-in-mean

- Most models in finance suppose that investors should obtain a higher return for taking additional risks.
- We can model this by letting the return depend on the risk,. We would obtain

$$\begin{aligned} R_t &= \mu_t + \underbrace{\sigma_t \varepsilon_t}_{u_t}, \varepsilon_t \sim N(0,1) \\ \mu_t &= E(R_t|F_{t-1}) = \mu + \delta\sigma_t^2 \\ \sigma_t^2 &= \omega + \alpha u_{t-1}^2 + \beta\sigma_{t-1}^2. \end{aligned}$$

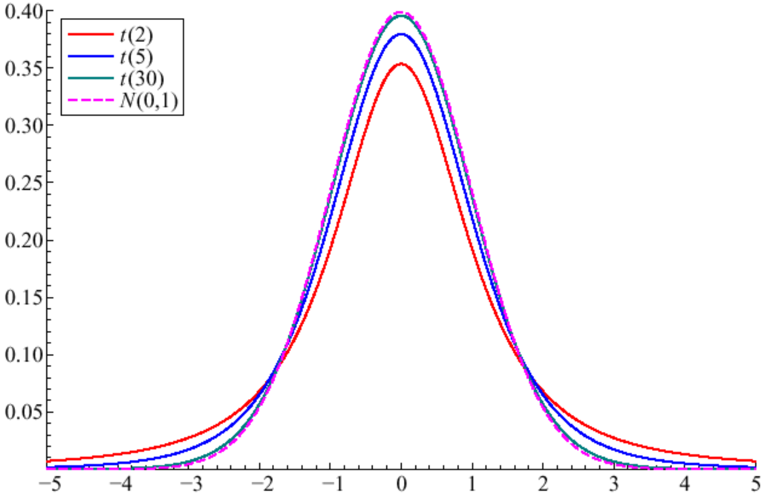
  known as GARCH-M model.
- If $\delta > 0$, then $\uparrow \delta \Rightarrow \uparrow R_t$ thus δ can be interpreted as a risk premium.

Non-Gaussian Likelihoods for GARCH models.

- The Student's $t(d)$ distribution is well known from linear regression as the distribution of t-statistics, where the degrees of freedom d is given by $T - k$.
- The same family of distributions can be defined for any (non-integer) $d > 0$.
- Small values of d correspond to fat tails: for $d = 1$ we obtain the Cauchy distribution, which has no finite mean or variance.
- As $d \to \infty$, we approach the $N(0, 1)$ distribution (in fact it is close to $N(0, 1)$ if $d = 30$).
- For $d > 2$, the variance of a $t(d)$ random variable X is $d/(d-2)$;
- the distribution of $\varepsilon = X/\sqrt{d/(d-2)}$ is called *standardized* $t(d)$, denoted $\tilde{t}(d)$.
- For $d > 4$ the excess kurtosis is $6/(d-4)$. The distributions are symmetric around 0 (hence mean and skewness are 0).

Non-Gaussian Likelihoods for GARCH models.

Student's t densities:



Non-Gaussian Likelihoods for GARCH models.

- The GARCH model $R_t = \mu_t + \sigma_t \varepsilon_t$, $\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2$, may be extended to $\varepsilon_t \sim \tilde{t}(d)$, where d is an extra parameter that can be estimated by maximum likelihood.
- In practice this GARCH-t model often gives a substantially better fit than the Gaussian model.
- **Potential problem:** If ε_t does not have the $\tilde{t}(d)$ distribution the Quasi-Maximum Likelihood estimator is *not consistent* if the true distribution of ε_t is not symmetric about zero even if the the conditional mean and conditional variances are well specified.
- Recall that if it is used the incorrect assumption that $\varepsilon_t \sim N(0, 1)$ to construct the likelihood function the Quasi-Maximum Likelihood estimator is still consistent provided that the conditional mean and conditional variances are well specified..

Estimation of GARCH-t :

Dependent Variable: Y1
Method: ML - ARCH (Marquardt) - Student's t distribution
Sample: 2/01/1981 12/12/2005
Included observations: 6507
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000489	9.34E-05	5.237805	0.0000
Variance Equation				
C	5.85E-07	1.38E-07	4.223580	0.0000
RESID(-1)^2	0.045023	0.004596	9.795426	0.0000
GARCH(-1)	0.949536	0.004807	197.5429	0.0000
T-DIST. DOF	5.829108	0.375539	15.52196	0.0000
Log likelihood	21678.18	Durbin-Watson stat		1.973611

Specification Testing

Suppose you estimated an ARMA model and you would like to test if the model is well specified:

- Diagnostic tests are based on the *residuals* $\hat{u}_t$.
- If there are no ARCH effects we should find no autocorrelation in $\hat{u}_t^2$.
- Therefore, the model can be tested using Q-statistics for $\hat{u}_t^2$. The latter statistic tests if there are ARCH effects.
- Suppose that we would like to test that there are no ARCH effects.
- We can consider the model $E_{t-1}(u_t^2) = \gamma_0 + \gamma_1 u_{t-1}^2 + \ldots + \gamma_m u_{t-m}^2$ and if there are no ARCH effects we must have $H_0 : \gamma_1 = ... = \gamma_m = 0$
- Lagrange-Multiplier (LM) test against ARCH, which is obtained by $LM = T \cdot R^2 \xrightarrow{D} \chi^2(m)$ in the regression

$$\hat{u}_t^2 = \hat{\gamma}_0 + \hat{\gamma}_1 \hat{u}_{t-1}^2 + \ldots + \hat{\gamma}_m \hat{u}_{t-m}^2 + e_t.$$

Specification Testing in GARCH models

Suppose now that you estimated a GARCH(q,s) model.
How do you test if the model is well specified?

- Diagnostic tests are based on the *standardized residuals* $\hat{\varepsilon}_t = \hat{u}_t / \hat{\sigma}_t$. If μ_t and σ_t are correctly specified, we should find no autocorrelation in $\hat{\varepsilon}_t$.
- We can apply the Q-statistics and Lagrange-Multiplier test for serial correlation in $\hat{\varepsilon}_t$.

How can we test if the model assumed for the conditional variance is well specified?

- We would like to test $H_0 : E_{t-1}(u_t^2) = \sigma_t^2$ which is equivalent to $E_{t-1}(\varepsilon_t^2 - 1) = 0$ with $\varepsilon_t = u_t / \sigma_t$.
- It is standard practice to apply the tests for ARCH effects described above, though they are *not valid* after the estimation of the GARCH model.

A *valid* Lagrange-Multiplier (LM) test against ARCH is constructed in the following way:

- We can consider the model $E_{t-1}(\varepsilon_t^2) = \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \ldots + \gamma_m \varepsilon_{t-m}^2$ and if there are no further ARCH effects we must have $H_0 : \gamma_1 = ... = \gamma_m = 0$.
- Let ψ be the parameters of the conditional variance σ_t^2 and define the vector

$$x_t = \frac{1}{\hat{\sigma}_t^2} \frac{\partial \hat{\sigma}_t^2}{\partial \psi'},$$

  where $\frac{\partial \hat{\sigma}_t^2}{\partial \psi'}$ is the derivative of σ_t^2 with respect to ψ estimated under H_0.
- Lagrange-Multiplier (LM) test against ARCH, which is obtained by $LM = T \cdot R^2 \xrightarrow{D} \chi^2(m)$ in the regression

$$\hat{\varepsilon}_t^2 = \gamma_0 + \gamma_1 \hat{\varepsilon}_{t-1}^2 + \ldots + \gamma_m \hat{\varepsilon}_{t-m}^2 + \delta' x_t + e_t.$$

Volatility forecasting

- GARCH specification of

$$\begin{aligned}\sigma_h^2(1) &= var(u_{h+1}|F_h)\\ &= E(u_{h+1}^2|F_h)\end{aligned}$$

by construction gives one-step ahead forecasts of u_{h+1}^2.

- Multi-step forecasts involves (using tower property) for $\ell \geq 1$

$$\begin{aligned}\sigma_h^2(\ell) &= var(u_{h+\ell}|F_h)\\ &= E(u_{h+\ell}^2|F_h)\\ &= E(E(u_{h+\ell}^2|F_{h+\ell-1})|F_h)\\ &= E(\sigma_{h+\ell}^2|F_h)\end{aligned}$$

For example for the GARCH(1,1) model we have

$$\begin{aligned}\sigma_h^2(\ell) &= E(\sigma_{h+\ell}^2|F_h)\\ &= E(\omega + \alpha u_{h+\ell-1}^2 + \beta\sigma_{h+\ell-1}^2|F_h)\\ &= \omega + (\alpha+\beta)\sigma_h^2(\ell-1)\end{aligned}$$

Volatility forecasting

- If $\alpha + \beta < 1$, where $\sigma^2 = \omega/(1 - \alpha - \beta)$

$$\sigma_h^2(\ell) = \sigma^2 + (\alpha + \beta)^{\ell-1}(\sigma_h^2(1) - \sigma^2) \to \sigma^2$$

  as $\ell \to \infty$.
- If $\alpha + \beta = 1$ (IGARCH) then $\sigma_h^2(\ell) = \sigma_h^2(1) + \omega(\ell - 1)$