Module 9.3: Time Series Analysis
Fall Term 2022

Week 5:

Volatility Modelling

HSLU Lucerne University of Applied Sciences and Arts

Outline in Weeks

1. Introduction; Descriptive Modelling
2. Returns; Autocorrelation; Stationarity
3. ARMA Models
4. Unit Roots; Regressions between Time Series
5. Volatility Modelling
6. Value at Risk
7. Cointegration
8. Panel Data

Outline

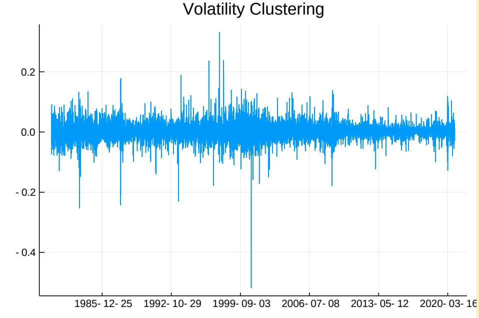
Goal

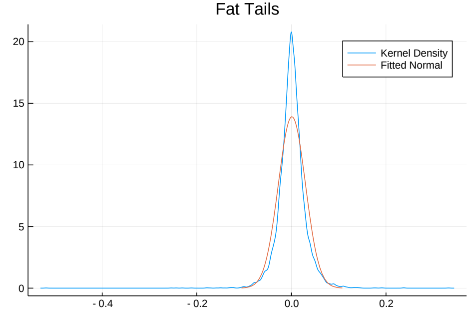
- Recall these stylized facts about asset returns:
  1. Lack of autocorrelation (efficient market hypothesis)
  2. Volatility clustering
  3. Distribution has heavy tails
  4. Leverage effects
- Goal today: model the last 3 of these, starting with the volatility clustering.

Volatility

- The *volatility* of an investment is a measure of its *risk*. Usually defined as the standard deviation of the return on the investment.
- Volatility is an important ingredient in:
  - portfolio selection;
  - risk management;
  - option pricing.
- Daily financial returns display *volatility clustering*: periods of high volatility alternate with more tranquil periods.
- In other words: large (in absolute value) returns tend to be followed by large (in absolute value) returns.
- This forms the basis for the *autoregressive-conditional heteroskedasticity* model (ARCH; Engle, 1982) and the *generalized ARCH* model (GARCH; Bollerslev, 1986).

Example: Daily Returns on Apple Stock





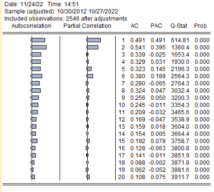
Reminder: Parameters vs. sample values

- We usually write σ for the standard deviation of, e.g., a normally distributed variable.
- σ is a *parameter* and therefore unknown.
- The best we can hope for is to *estimate* it, usually with the *sample standard deviation* s.
- With stock returns, the standard deviation (or *volatility*) changes over time, due to *volatility clustering*.
- We write σ_t for the volatility in period t.
- Note that σ_t is *unobserved*. The best we can do is *estimate* it. We'll write $\widehat{\sigma}_t$ for this estimate.
- Today, we'll mostly discuss different methods of estimation.

Detecting Volatility Clustering (I)

- Since volatility clustering means that large returns tend to be followed by large returns, it is possible to detect it by inspecting the correlogram of the squared returns.

Example: correlogram of squared S&P500 returns.



Date: 11/24/22 Time: 14:51
Sample (adjusted): 10/30/2012 10/27/2022
Included observations: 2546 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.491	0.491	614.81	0.000
		2	0.541	0.395	1360.4	0.000
		3	0.339	-0.025	1653.4	0.000
		4	0.329	0.031	1930.0	0.000
		5	0.323	0.145	2196.3	0.000
		6	0.380	0.189	2564.3	0.000
		7	0.280	-0.065	2764.3	0.000
		8	0.324	0.047	3032.4	0.000
		9	0.256	0.050	3200.3	0.000
		10	0.245	-0.011	3354.3	0.000
		11	0.209	-0.032	3465.6	0.000
		12	0.169	-0.047	3538.9	0.000
		13	0.159	0.018	3604.0	0.000
		14	0.154	0.005	3664.4	0.000
		15	0.192	0.078	3758.7	0.000
		16	0.128	-0.063	3800.8	0.000
		17	0.141	-0.011	3851.9	0.000
		18	0.088	-0.002	3871.8	0.000
		19	0.062	-0.052	3881.6	0.000
		20	0.108	0.075	3911.7	0.000

- Clearly, there is a lot of predictability in squared returns (unlike returns themselves).

Detecting Volatility Clustering (II)

- Besides relying on the Q-tests from the correlogram, another formal test is Engle's *ARCH-LM* test (essentially a Breusch-Godfrey test applied to the squared residuals).
- EViews only offers it for residuals, not for a series itself. Hence, we start by regressing the returns on an intercept.
- The ARCH-LM test is based on the auxiliary regression

$$\hat{u}_t^2 = \gamma_0 + \gamma_1 \hat{u}_{t-1}^2 + \ldots + \gamma_m \hat{u}_{t-m}^2 + e_t.$$

- The lag length m is chosen by the user, e.g., 5 for daily data.
- The test statistic is $T \cdot R_{aux}^2$ and has a $\chi^2(m)$ distribution under $H_0 : \gamma_1 = \cdots = \gamma_m = 0$ (no volatility clustering).

Example: ARCH-LM test for the S&P500

Heteroskedasticity Test: ARCH

F-statistic	302.2045	Prob. F(5,2535)	0.0000
Obs*R-squared	948.9587	Prob. Chi-Square(5)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/24/22 Time: 15:11
Sample (adjusted): 11/06/2012 10/27/2022
Included observations: 2541 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.14E-05	8.68E-06	3.614959	0.0003
RESID^2(-1)	0.299355	0.019651	15.23388	0.0000
RESID^2(-2)	0.396998	0.020528	19.33915	0.0000
RESID^2(-3)	-0.087729	0.021921	-4.001999	0.0001
RESID^2(-4)	-0.015277	0.020529	-0.744168	0.4568
RESID^2(-5)	0.145365	0.019653	7.396390	0.0000

R-squared	0.373459	Mean dependent var	0.000120
Adjusted R-squared	0.372223	S.D. dependent var	0.000528
S.E. of regression	0.000419	Akaike info criterion	-12.71671
Sum squared resid	0.000444	Schwarz criterion	-12.70292
Log likelihood	16162.59	Hannan-Quinn criter.	-12.71171
F-statistic	302.2045	Durbin-Watson stat	2.054495
Prob(F-statistic)	0.000000		

The null of no volatility clustering is clearly rejected (p-value is zero, $T \cdot R^2_{aux} = 948.96$ much larger than critical value 11.07).

Outline

Historical Volatility

- A first simple estimator is *historical volatility*, i.e., the sample standard deviation of the most recent m observations (often $m = 250$, one year).
- If $r_t = \ln P_t - \ln P_{t-1}$ denotes the daily log-return, then

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

  (Typically the average return is relatively close to zero). This is an estimate of the squared volatility over day $t+1$, made at the end of day t.
- 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.

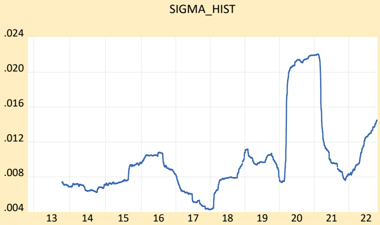
RiskMetrics

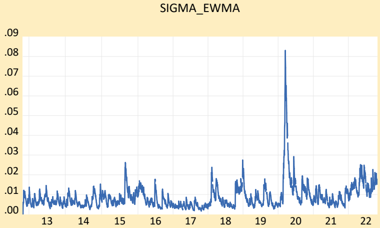
- Problems with historical volatility are addressed by replacing equally weighted moving average by an *exponentially* weighted moving average (EWMA), also used in JPMorgan's *RiskMetrics* system:

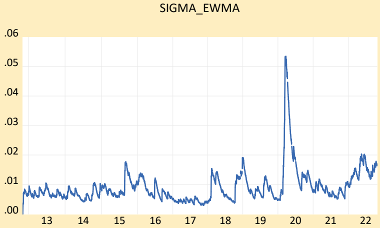
$$\begin{aligned}\widehat{\sigma}^2_{t+1,EWMA} &= (1-\lambda)\sum_{j=0}^{\infty}\lambda^j r^2_{t-j} \\ &= \lambda\widehat{\sigma}^2_{t,EWMA} + (1-\lambda)r_t^2, \qquad 0<\lambda<1.\end{aligned}$$

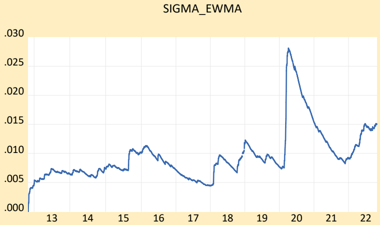
- This means that observations further in the past get a smaller weight.
- In practice we do not have $r_{t-\infty}$, but the second equation can be started up by an initial estimate / guess $\sigma^2_{0,EWMA}$.
- The larger λ, the stronger the persistence of shocks (large returns).
- For daily data, RiskMetrics recommends $\lambda = 0.94$.

Example: S&P500 volatility, historical and EWMA ($\lambda = 0.8, 0.94, 0.99$)









Outline

The ARCH Model

- The first-order *autoregressive-conditional heteroskedasticity* (ARCH(1)) model, due to Engle (1982), for a return r_t with mean zero is

$$\sigma_{t+1}^2 = \omega + \alpha r_t^2 .$$

- In practice, we need to allow for $\mathbb{E}[r_{t+1}] = \mu_{t+1} \neq 0$. Then $r_{t+1} = \mu_{t+1} + u_{t+1}$, and the model becomes

$$\sigma_{t+1}^2 = \omega + \alpha u_t^2 .$$

The ARCH Model

- When trying to estimate ARCH models one might find that more lags are needed, leading to ARCH(q):
$$\sigma_{t+1}^2 = \omega + \alpha_1 u_t^2 + \ldots + \alpha_q u_{t-q+1}^2.$$
- *Note*: Variances must be positive, therefore we need to impose $\omega > 0$, $\alpha_i \geq 0$, $i = 1, \ldots, q$.
- It can be shown that an ARCH(q) models corresponds to an AR(q) for the squared returns. Thus, we could determine the order from the correlogram of the squared returns: SPACF should cut off after q lags.
- In the example above, we might conclude that we need an ARCH(6) model.

The GARCH 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)), due to Bollerslev (1986):

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

- *Advantage*: Flexible structure with only 3 parameters to estimate.

The GARCH Model

- The GARCH(1,1) model is stationary if the unconditional (“average”) variance $\sigma^2 = \mathbb{E}[\sigma_t^2]$ is positive, constant and finite.
- This requires

$$\begin{aligned} \sigma^2 = \mathbb{E}[\sigma_{t+1}^2] &= \omega + \alpha \mathbb{E}[u_t^2] + \beta \mathbb{E}[\sigma_t^2] \\ &= \omega + \alpha\sigma^2 + \beta\sigma^2. \end{aligned}$$

- Hence, provided that $\alpha + \beta < 1$ (the *stationarity condition*),

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

- The nonstationary model with $\alpha + \beta = 1$ is called *integrated GARCH* (IGARCH): infinite variance, no mean-reversion in volatility.
- Notice that an IGARCH with $r_t = u_t$, $\omega = 0$, $\beta = \lambda$, and $\alpha = (1 - \lambda)$ is just the RiskMetrics model.

The GARCH Model

Some other properties:

- The ACF and PACF of r_t^2 in case of stationary GARCH(1,1) are both exponentially decaying, no cut-off point.
- The *standardized returns*

$$z_{t+1} = \frac{r_{t+1} - \mu_{t+1}}{\sigma_{t+1}}$$

satisfy $E(z_{t+1}) = 0$ and $\mathrm{var}(z_{t+1}) = 1$. Therefore the model may be formulated as

$$\begin{aligned} r_{t+1} &= \mu_{t+1} + u_{t+1} = \mu_{t+1} + \sigma_{t+1} z_{t+1}, \\ \sigma_{t+1}^2 &= \omega + \alpha u_t^2 + \beta \sigma_t^2 . \end{aligned}$$

- Often it is assumed that z_t are i.i.d. as $N(0,1)$.
- Even if $z_t \sim N(0,1)$, it can be shown that varying σ_t implies that r_t has non-normal distribution, with higher kurtosis.

The GARCH(p, q) Model

- The GARCH(1, 1) model can be extended to the GARCH(p, q) model

$$\sigma_{t+1}^2 = \omega + \alpha_1 u_t^2 + \cdots + \alpha_q u_{t-q+1}^2 + \beta_1 \sigma_t^2 + \cdots + \beta_p \sigma_{t-p+1}^2$$

  although in practice, this is rarely necessary.
- The model is stationary if $\sum_{i=1}^{p} \beta_i + \sum_{i=1}^{q} \alpha_i < 1$, and the unconditional variance is

$$\frac{\omega}{1 - \sum_{i=1}^{p} \beta_i - \sum_{i=1}^{q} \alpha_i}.$$

Outline

Estimation of GARCH Models

- GARCH cannot be estimated by ordinary least-squares (because σ_t^2 is not observed).
- Such models are estimated by *maximum likelihood*: the joint density of the observations $\{r_1, \ldots, r_T\}$ is maximized with respect to the parameters.
- Maximization of $\log L$ can be done by numerical optimization algorithms. By default, EViews does this under the assumption of normality.
- If we are not sure that the z_t's are normally distributed, then we may still use the same estimation technique. This is called *quasi-maximum likelihood estimator*.
- However, we need to construct standard errors via a more robust method (*Bollerslev-Wooldridge standard errors*).

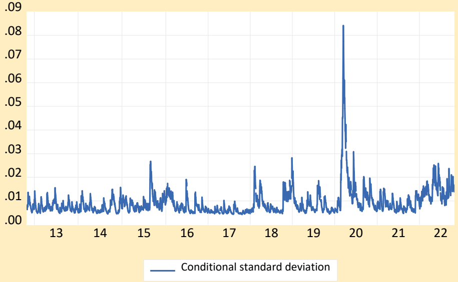
Example: EViews output, estimated GARCH model for S&P500

Dependent Variable: R
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
Date: 11/24/22 Time: 15:40
Sample (adjusted): 10/30/2012 10/27/2022
Included observations: 2546 after adjustments
Convergence achieved after 18 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000799	0.000135	5.911544	0.0000
Variance Equation				
C	4.50E-06	4.30E-07	10.46617	0.0000
RESID(-1)^2	0.223750	0.016516	13.54710	0.0000
GARCH(-1)	0.741309	0.016652	44.51892	0.0000

R-squared	-0.001398	Mean dependent var	0.000390
Adjusted R-squared	-0.001398	S.D. dependent var	0.010946
S.E. of regression	0.010953	Akaike info criterion	-6.742939
Sum squared resid	0.305331	Schwarz criterion	-6.733760
Log likelihood	8587.761	Hannan-Quinn criter.	-6.739609
Durbin-Watson stat	2.287783		

Example: Estimated GARCH volatility of S&P500



Outline

Testing GARCH Models

- Diagnostic tests are based on the *standardized residuals* $\hat{z}_t := \hat{u}_t/\hat{\sigma}_t$. If μ_t and σ_t are correctly specified, we should find no autocorrelation in $\hat{z}_t$ and $\hat{z}_t^2$.
- Therefore, the model can be tested using Q-statistics for $\hat{z}_t$ or $\hat{z}_t^2$.
- Lagrange-Multiplier (LM) test against ARCH, which is obtained by $T \cdot R^2$ in the regression

$$\hat{z}_t^2 = \gamma_0 + \gamma_1 \hat{z}_{t-1}^2 + \ldots + \gamma_m \hat{z}_{t-m}^2 + e_t.$$

- To test for normality of z_t, we can use the Jarque-Bera test based on the skewness and kurtosis of $\hat{z}_t$.

Example: Correlogram of standardized residuals for the S&P500

Date: 11/24/22 Time: 15:48
Sample (adjusted): 10/30/2012 10/27/2022
Included observations: 2546 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.032	-0.032	2.6093	0.106
		2	0.010	0.009	2.8792	0.237
		3	-0.007	-0.007	3.0169	0.389
		4	-0.014	-0.014	3.4899	0.479
		5	-0.022	-0.023	4.7188	0.451
		6	-0.022	-0.023	5.9352	0.430
		7	0.027	0.025	7.7406	0.356
		8	-0.028	-0.027	9.8064	0.279
		9	0.008	0.005	9.9642	0.353
		10	-0.009	-0.009	10.169	0.426
		11	0.021	0.020	11.328	0.416
		12	-0.033	-0.032	14.092	0.295
		13	-0.030	-0.033	16.403	0.228
		14	-0.027	-0.030	18.323	0.192
		15	-0.014	-0.014	18.815	0.222
		16	0.000	-0.002	18.815	0.278
		17	0.022	0.021	20.073	0.271
		18	0.021	0.017	21.170	0.271
		19	-0.021	-0.020	22.254	0.272
		20	0.018	0.015	23.127	0.283

*Probabilities may not be valid for this equation specification.

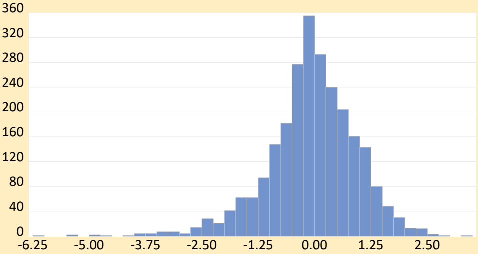
Example: Correlogram of squared standardized residuals for the S&P500

Date: 11/24/22 Time: 15:48
Sample (adjusted): 10/30/2012 10/27/2022
Included observations: 2546 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.001	0.001	0.0011	0.973
		2	-0.000	-0.000	0.0015	0.999
		3	-0.003	-0.003	0.0182	0.999
		4	0.024	0.024	1.4430	0.837
		5	-0.028	-0.028	3.4029	0.638
		6	-0.008	-0.008	3.5668	0.735
		7	-0.021	-0.021	4.6583	0.702
		8	-0.017	-0.017	5.3624	0.718
		9	-0.008	-0.006	5.5064	0.788
		10	0.038	0.037	9.1425	0.519
		11	-0.004	-0.004	9.1864	0.605
		12	-0.017	-0.017	9.8904	0.626
		13	0.000	-0.001	9.8906	0.703
		14	-0.018	-0.020	10.676	0.711
		15	-0.002	-0.001	10.686	0.774
		16	0.003	0.004	10.712	0.827
		17	0.002	0.002	10.720	0.871
		18	-0.007	-0.005	10.850	0.901
		19	-0.007	-0.009	10.990	0.924
		20	0.020	0.017	11.993	0.916

*Probabilities may not be valid for this equation specification.

Example: Normality test of standardized residuals for the S&P500



Series: Standardized Residuals
Sample 10/30/2012 10/27/2022
Observations 2546

Mean	-0.060225
Median	-0.034548
Maximum	3.439897
Minimum	-6.012431
Std. Dev.	0.998517
Skewness	-0.737406
Kurtosis	5.246133
Jarque-Bera	765.9409
Probability	0.000000

Outline

Asymmetry and the News Impact Curve

- The *news impact curve* (NIC) is the effect of u_t on σ^2_{t+1}, keeping σ^2_t and the past fixed.
- For GARCH(1,1), this is the parabola $NIC(u_t|\sigma^2_t = \sigma^2) = A + \alpha u^2_t$, with $A = \omega + \beta\sigma^2$. This has a minimum at $u_t = 0$, and is symmetric around that minimum.
- For equity, a large negative shock is expected to increase volatility more than a large positive shock, because of *leverage effect*:

$$
\begin{array}{ll}
 & \downarrow \text{value of firm's stock} \\
\Rightarrow & \downarrow \text{equity value of the firm} \\
\Rightarrow & \uparrow \text{debt-to-equity ratio} \\
\Rightarrow & \text{shareholders (as residual claimants) perceive future cashflows} \\
 & \text{as more risky.}
\end{array}
$$

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

GJR-GARCH (or TARCH, threshold GARCH)

The GJR-GARCH(1,1) model is

$$\sigma_{t+1}^2 = \omega + \alpha u_t^2 + \gamma u_t^2 \mathrm{I}_t + \beta \sigma_t^2.$$

where

$$\mathrm{I}_t = \begin{cases} 1 & \text{if} \quad u_t < 0 \\ 0 & \text{if} \quad u_t \geq 0 \end{cases},$$

and u_t/σ_t has a symmetric distribution.
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 $\sigma^2 = \omega / \left[1 - \alpha - \frac{1}{2}\gamma - \beta\right]$.

EGARCH

The EGARCH(1,1) model is

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

with $z_t = u_t/\sigma_t$ as usual. If $z_t \sim$ i.i.d. $N(0,1)$ then $E|z_t| = \sqrt{2/\pi}$.
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;
- $\gamma z_t + \alpha(|z_t| - E|z_t|)$ is an i.i.d. mean-zero shock to log-volatility;
- if $|\beta| < 1$, $\log \sigma_t^2$ is stationary with mean $\omega/(1-\beta)$.

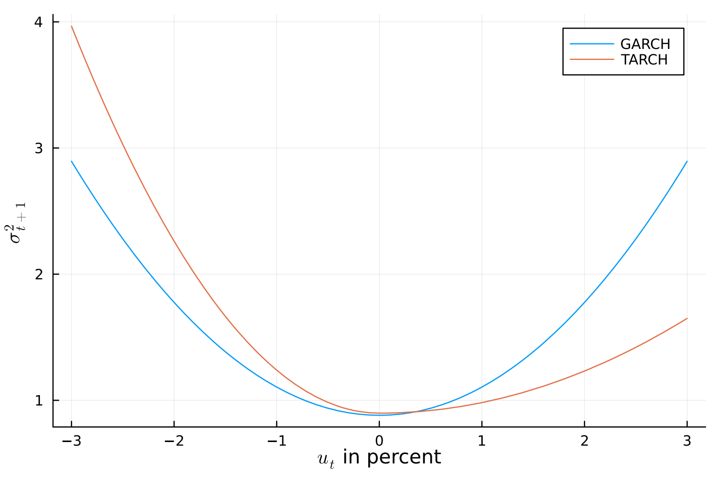
Example: EViews output, estimated TARCH model for S&P500

Dependent Variable: R
Method: ML ARCH - Normal distribution (BFGS / Marquardt steps)
Date: 11/24/22 Time: 15:56
Sample (adjusted): 10/30/2012 10/27/2022
Included observations: 2546 after adjustments
Convergence achieved after 26 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-1)^2*(RESID(-1)<0) + C(5)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000472	0.000146	3.229615	0.0012
Variance Equation				
C	4.18E-06	3.67E-07	11.39238	0.0000
RESID(-1)^2	0.083481	0.008795	9.491794	0.0000
RESID(-1)^2*(RESID(-1)<0)	0.257376	0.028027	9.183082	0.0000
GARCH(-1)	0.755584	0.014405	52.45111	0.0000

R-squared	-0.000056	Mean dependent var	0.000390
Adjusted R-squared	-0.000056	S.D. dependent var	0.010946
S.E. of regression	0.010946	Akaike info criterion	-6.767193
Sum squared resid	0.304922	Schwarz criterion	-6.755719
Log likelihood	8619.636	Hannan-Quinn criter.	-6.763031
Durbin-Watson stat	2.290853		

Example: NIC of GARCH and TARCH models for S&P500



Outline

Volatility Forecasting

- GARCH models directly provide forecasts of next day's volatility:

$$\widehat{\sigma}^2_{t+1} = \hat{\omega} + \hat{\alpha}\hat{u}_t^2 + \hat{\beta}\widehat{\sigma}_t^2.$$

- This can also be expressed as

$$\widehat{\sigma}^2_{t+1} = \widehat{\sigma}^2 + \hat{\alpha}(\hat{u}_t^2 - \widehat{\sigma}^2) + \hat{\beta}(\widehat{\sigma}_t^2 - \widehat{\sigma}^2).$$

  with $\widehat{\sigma}^2 = \hat{\omega}/(1 - \hat{\alpha} - \hat{\beta})$; see the exercises.
- So the forecast differs from the average variance $\widehat{\sigma}^2$ if $\hat{u}_t^2$ or $\widehat{\sigma}_t^2$ differ from $\widehat{\sigma}^2$.

Multi-Period Forecasts

- Regarding multi-period forecasts of the stationary GARCH(1, 1) model, it can be shown that
$$\widehat{\sigma}^2_{t+s} = \widehat{\sigma}^2 + (\hat{\alpha} + \hat{\beta})^{s-1}(\widehat{\sigma}^2_{t+1} - \widehat{\sigma}^2),$$
see the exercises.
- This implies $\widehat{\sigma}^2_{t+s} \to \widehat{\sigma}^2$ as $s \to \infty$.
- For RiskMetrics ($\alpha + \beta = 1$, $\omega = 0$), this simplifies to $\widehat{\sigma}^2_{t+s} = \widehat{\sigma}^2_{t+1}$ for all s.

Outline

Learning Goals

Students

- can use appropriate tests to detect volatility clustering,
- are able to estimate, interpret, and forecast the various models (historical volatility, RiskMetrics, (G)ARCH, TARCH, EGARCH), and to apply diagnostic tests to the standardized residuals,
- and understand the concept of leverage, and the NIC.

Homework

- Exercise 5
- Questions 1 and 3 from Chapter 9 of Brooks (2019)