Module 9.3: Time Series Analysis Fall Term 2023

Week 5:

Volatility Modelling



Outline in Weeks

- Introduction; Descriptive Modelling
- Returns; Autocorrelation; Stationarity
- ARMA Models
- Unit Roots; Regressions between Time Series
- Volatility Modelling
- Value at Risk
- Cointegration

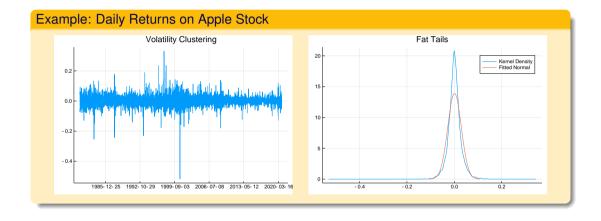
- Introduction
- 2 Historical, RiskMetrica
- The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- 8 Epilogue

Goal

- Recall these stylized facts about asset returns:
 - Lack of autocorrelation (efficient market hypothesis)
 - Volatility clustering
 - Distribution has heavy tails
 - 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).



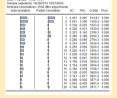
Reminder: Parameters vs. sample values

- ullet We usually write σ for the standard deviation of, e.g., a normally distributed variable.
- \bullet 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 $\hat{\sigma}_t$ 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.



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

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

- 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).

Introduction

0000000

A CHARLES

	Heteroskedasticity Te	Heteroskedasticity Test: ARCH				
Obs N-squared 946.9567 Flob. Clin-Square(5) 0.0	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

Madable

Variable	Coefficient	Sta. Error	t-Statistic	Prob.
С	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 depend	tent var	0.000120
Adjusted R-squared	0.372223	S.D. depende	ent var	0.000528
S.E. of regression	0.000419	Schwarz criterion		-12.71671
Sum squared resid	0.000444			-12.70292
Log likelihood	16162.59	Hannan-Quin		-12.71171
F-statistic	302.2045	Durbin-Watso	on stat	2.054495
Prob(F-statistic)	0.000000			

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

- Introduction
- 2 Historical, RiskMetrics
- 3 The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- 8 Epilogue

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

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

(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:

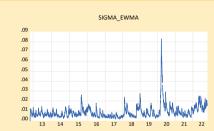
$$\widehat{\sigma}_{t+1,EWMA}^2 = (1-\lambda) \sum_{j=0}^{\infty} \lambda^j r_{t-j}^2$$

$$= \lambda \widehat{\sigma}_{t,EWMA}^2 + (1-\lambda) r_t^2, \qquad 0 < \lambda < 1.$$

- 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 $\sigma_{0,EWMA}^2$.
- ullet 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 ($\lambda = 0.8, 0.94, 0.99$)









- Introduction
- 2 Historical, RiskMetrics
- 3 The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- Epilogue

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 $\omega > 0$, $\alpha_i \ge 0$, $i = 1, \dots, q$.
- 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_t² or u_t², which leads to the generalized ARCH model of orders (1,1) (GARCH(1,1)), due to Bollerslev (1986):

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

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

$$\sigma^{2} = \mathbb{E}[\sigma_{t+1}^{2}] = \omega + \alpha \mathbb{E}[u_{t}^{2}] + \beta \mathbb{E}[\sigma_{t}^{2}]$$
$$= \omega + \alpha \sigma^{2} + \beta \sigma^{2}.$$

• 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 $var(z_{t+1}) = 1$. 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

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

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

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

- Introduction
- 2 Historical, RiskMetrics
- The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- 8 Epilogue

Estimation of GARCH Models

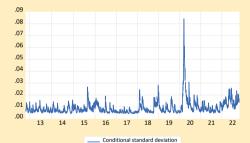
- GARCH cannot be estimated by ordinary least-squares (because σ_t^2 is not observed).
- Such models are estimated by *maximum likelihood*: the joint density of the observations $\{r_1, \ldots, r_T\}$ is maximized with respect to the parameters.
- Maximization of $\log L$ can be done by numerical optimization 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
Hefnod ML RACH - Normal distribution (BFGS / Marquardt steps)
Date: 11/24/22 Time: 15.40
Sample digulated; 10/30/2012 10/027/2022
Included observations: 2546 after adjustments
Convergence achieved after 18 feat glosser for the convergence achieved after 18 feat glosser for the convergence achieved after 18 feat glosser for product of gradients
Codefficient covariance computed using outer product of gradients
CARCHs (27) = COVENSIDE CVEY - CARCARCH (-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
С	0.000799	0.000135	5.911544	0.0000
	Variance	Equation		
С	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 depend	lent 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			





- Introduction
- 2 Historical, RiskMetrics
- The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- 8 Epilogue

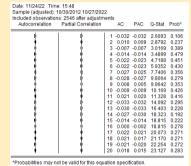
Testing GARCH Models

- 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 .
- ullet Lagrange-Multiplier (LM) test against ARCH, which is obtained by $T \cdot R^2$ in the regression

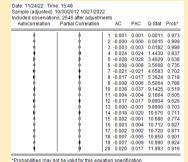
$$\hat{z}_t^2 = \gamma_0 + \gamma_1 \hat{z}_{t-1}^2 + \ldots + \gamma_m \hat{z}_{t-m}^2 + e_t.$$

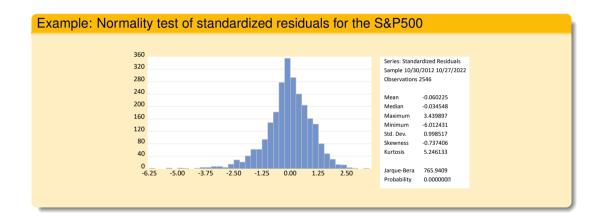
• To test for normality of z_t , we can use the Jarque-Bera test based on the skewness and kurtosis of \hat{z}_t .

Example: Correlogram of standardized residuals for the S&P500









- 1 Introduction
- 2 Historical, RiskMetrics
- The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- Epilogue

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.
- For equity, a large negative shock is expected to increase volatility more than a large positive shock, because of *leverage effect*:
 - ↓ value of firm's stock
 - \Rightarrow \downarrow equity value of the firm
 - \Rightarrow \uparrow debt-to-equity ratio
 - \Rightarrow shareholders (as residual claimants) perceive future cashflows as more risky.
- Multiple extensions exist to deal with this issue. Here we focus on Glosten, Jagannathan and Runkle's GJR-GARCH model.

GJR-GARCH (or TARCH, threshold GARCH)

The GJR-GARCH(1,1) model is

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

where

$$I_t = \left\{ \begin{array}{ll} 1 & \text{if} \quad u_t < 0 \\ 0 & \text{if} \quad u_t \ge 0 \end{array} \right.,$$

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 $\omega > 0$, $\alpha \ge 0$, $\gamma \ge 0$, $\beta \ge 0$;
- u_t^2 is stationary if $0 \le \alpha + \frac{1}{2}\gamma + \beta < 1$, 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:

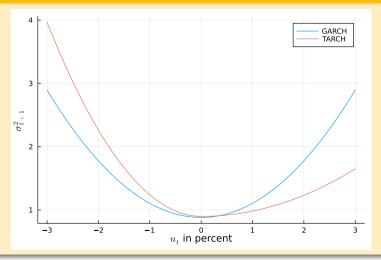
- NIC is asymmetric if and only if $\gamma \neq 0$; leverage effect if $\gamma < 0$;
- σ_t^2 is positive for all parameter values;
- $\gamma z_t + \alpha(|z_t| E|z_t|)$ is an i.i.d. mean-zero shock to log-volatility;
- if $|\beta| < 1$, $\log \sigma_t^2$ is stationary with mean $\omega/(1-\beta)$.

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: 10:00/20/12 10:027/20/22
Included observations: 25:46 after adjustments
Convergence achieved after 26 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backsat (grammeter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)*2 + C(4)*RESID(-1)*2"(RESID(-1)*0) +
C(5)*CARCH(+1)

variable	Coemcient	Std. Error Z-Statistic		Prob.	
С	C 0.000472 0.000146 3.229619		3.229615	0.0012	
	Variance	Equation			
С	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	0.014405 52.45111 Mean dependent var S.D. dependent var Akaike info criterion Schwarz criterion Hannan-Quinn criter.		0.0001	
GARCH(-1)	0.755584			0.000	
R-squared	-0.000056			0.00039	
Adjusted R-squared	-0.000056			0.010946	
S.E. of regression	0.010946			-6.76719	
Sum squared resid	0.304922			-6.755719	
Log likelihood	8619.636			-6.76303	
Durbin-Watson stat	2.290853				

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



- Introduction
- 2 Historical, RiskMetrics
- 3 The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- Epilogue

Volatility Forecasting

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

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

Multi-period forecasts can be constructed recursively. In principle, one would use

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

but \hat{u}_{t+1}^2 is unobserved.

- Solution: replace \hat{u}_{t+1}^2 with its estimate, $\hat{\sigma}_{t+1}^2$.
- Result:

$$\widehat{\sigma}_{t+2}^2 = \widehat{\omega} + \left(\widehat{\alpha} + \widehat{\beta}\right)\widehat{\sigma}_{t+1}^2,$$

- Introduction
- 2 Historical, RiskMetrics
- The ARCH and GARCH Models
- Estimation of GARCH Models
- Testing GARCH Models
- 6 Asymmetry and the News Impact Curve
- Volatility Forecasting
- 8 Epilogue

oduction Historical, RiskMetrics ARCH/GARCH Estimation Testing Asymmetry Forecasting Epilogue

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)