Econometrics

S Volatility Modelling

Dr. Rachida Ouysse
School of Economics[1]

UNSW SYDNEY Australia's Global University

UNSW AUSTRALIA | AGSM Business School

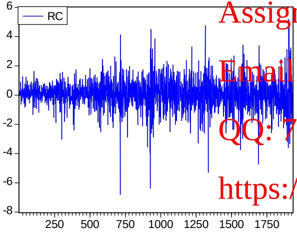
Lecture Plan

- Motivation for modeling return volatility
- Measures of return volatility
- Conditional volatility via smoothing
- ARCH
  - Conditional variance is a function of info set;
  - It captures "clustering" in return series;
  - It explains non-normality of return, to some extent;
  - It can be used to improve interval forecasts and VaR (Value at Risk);
  - Estimation and testing.

Introduction and Motivation

eg. Volatility in NYSE index return

- Clustering.
- Squared returns are strongly autocorrelated.



Correlogram of RC2

Sample: 1 1931
Included observations: 1930

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.179	0.179	62.044	0.000
		2	0.168	0.140	116.47	0.000
		3	0.175	[illegible]	176.00	0.000
		4	0.109	0.044	198.89	0.000
		5	0.184	0.130	264.80	0.000
		6	0.113	0.035	289.54	0.000
		7	0.136	0.067	325.24	0.000
		8	0.112	0.030	349.60	0.000
		9	0.108	0.039	372.11	0.000
		10	0.095	0.015	389.80	0.000
		11	0.107	0.042	411.90	0.000
		12	0.102	0.028	432.21	0.000
		13	0.066	-0.006	440.72	0.000
		14	0.053	-0.016	446.26	0.000
		15	0.041	-0.016	449.58	0.000
		16	0.097	0.052	467.86	0.000
		17	0.040	-0.018	470.96	0.000
		18	0.068	0.024	479.94	0.000
		19	0.063	0.014	487.77	0.000
		20	0.045	0.005	491.70	0.000

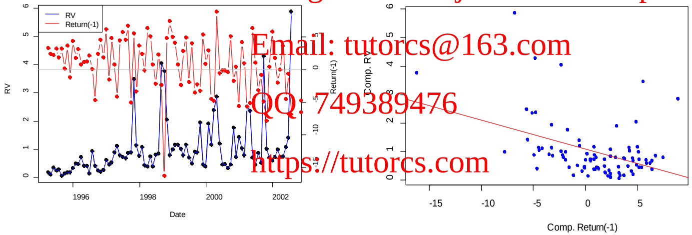
Motivation

eg. Volatility in NYSE index return

- Monthly realised var

  RV =sample mean of squared daily returns in a month

- RV is negatively correlated to lagged monthly return. Corr$(RV, Return(-1)) = -0.419$.



Motivation

- Importance of retur[illegible]

  Asset pricing, risk management and portfolio selection

  Substantial dependence structure in volatility

- Clustering:
  - strong autocorrelations in squared returns,
  - large variations tend to be followed by large variations

- Asymmetry:
  - negative returns tend to cause more volatility than positives

- ARMA are unable to capture these features

  Conditional variance is constant in ARMA.

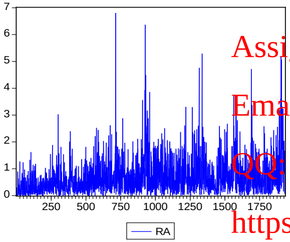
  Amend ARMA with a suitable conditional variance: ARCH and GARCH models.

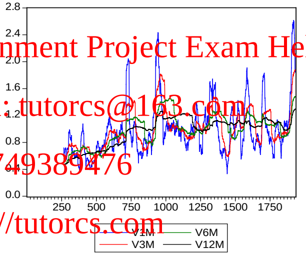
Volatility

Measures of return vola[illegible]ency of variation)

- Historical volatility: [illegible]iance or Stddev

eg. NYSE composite return: Sample Stddev





Realized Volatility

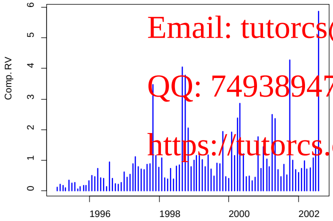
Measures of return vola

- **Realised volatility:** riance = Sample mean of squared higher frequency returns

  (eg. daily RV = Sample mean of squared 5-min returns)

eg. NYSE composite return: Monthly realised variance

RV = Sample mean of squared daily returns in a month



Realized Volatility

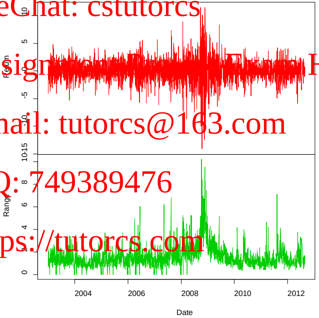
Measures of return vola

- **Range (high/low):**

  $100 \times$ ln(high/low) in a time interval (eg, a day)

eg. BHP

daily return and range



Implied Volatility

Implied volatility:

standard deviation derived from options prices

- Option of an asset: the right to buy/sell the asset at a future time (maturity) at a fixed price (strike).
- Given theprice of an option, maturity, strike and risk-free interest rate, the std deviation can be recovered from Black-Scholes formula, known as IV.
- IV represents market's opinions on the return's std deviation.

Black-Scholes formula:

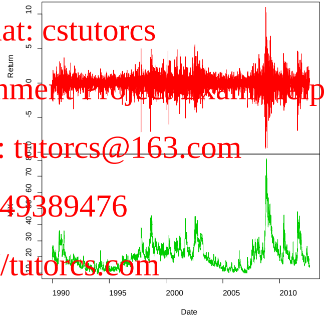
price of an option =f(stdev, maturity,strike, r_f−rate)

Implied Volatility

- **Implied volatility**

eg. **VIX**: index of IVs ptions on the SP500 index

SP500 daily return & VIX



Conditional Volatility

– Conditional vari[illegible]turn

- $\sigma^2_{t+1|t} = \mathrm{Var}(r_t$ [illegible]

  where $r_{t+1} = 100\ln(P_{t+1}/P_t)$ is the return and Ω_t is the information set at the end of period t .

- It should capture “clustering” or autocorrelations in squared returns, and facilitate predicting the return volatility

- Knowing it helps to
  - assess the risk of an asset via value-at-risk;
  - price options;
  - form mean-variance efficient portfolios.

Conditional Volatility

Exponentially weighted moving average (EWMA)

- The squared returns $\{r_t^2, r_{t-1}^2, \cdots, r_1^2\}$ carry info about the volatility as $E(r_t^2) \equiv$ variance.
- A weighted average of squared returns is an approximation to the conditional variance. Recent observations should weigh more.
- EWMA: for $0 < \lambda < 1$,

$$\sigma_{t+1|t}^2 = (1-\lambda)(r_t^2 + \lambda r_{t-1}^2 + \lambda^2 r_{t-2}^2 + \cdots)$$

  - weights decay exponentially;
  - weights sum up to 1;
  - RiskMetrics recommend $\lambda = 0.94$

EWMA

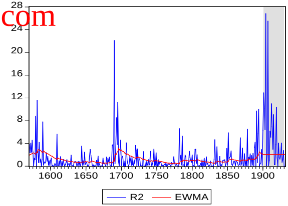
- EWMA: al[illegible] formulation
  - $\sigma^2_{1|0} = r_1^2$
  - $\sigma^2_{t+1|t} = (1-\lambda) r_t^2 + \lambda \sigma^2_{t|t-1}$, for $t = 1,2,3,\dots$

– Quick and easy;

– Can be used as 1-step ahead prediction.

eg. NYSE Composite return:

$\lambda = 0.94$



ARCH (autoregressive conditional heteroskedasticity) Engle (1982) – Nobel price winner 1993

Autoregressive conditional heteroskedasticity (ARCH) models are a class of models where the conditional variance evolves according to an autoregressive process.

First define the conditional variance of the error term u_t to be

$$\sigma_t^2 = \mathrm{var}\left(\mu_t \,|\mu_{t-1}, \mu_{t-2}, \dots\right) = E\left(\left(\mu_t - E\left(\mu_t\right)\right)^2 |\mu_{t-1}, \mu_{t-2}, \dots\right)$$

As it is usually assumed that $E(\mu_t) = 0$

$$\sigma_t^2 = \mathrm{var}\left(\mu_t \,|\mu_{t-1}, \mu_{t-2}, \dots\right) = E\left(\mu_t^2 \,|\mu_{t-1}, \mu_{t-2}, \dots\right) = E_{t-1}\left(\mu_t^2\right)$$

The ARCH(1) model assumes

$$\sigma_t^2 = E_{t-1}\left(\mu_t^2\right) = \alpha_0 + \alpha_1 \mu_{t-1}^2$$

The conditional variance captures 'clustering': large past shock leads to large conditional variance.

ARCH (autoregressive conditional heteroskedasticity)

Extensions

- An ARCH(q) mod

$$\sigma_t^2 = \alpha_0 + \alpha_1\mu_{t-1}^2 + \alpha_2\mu_{t-2}^2 + \ldots + \alpha_q\mu_{t-q}^2$$

- Under ARCH, the conditional mean equation can take any form. An example of a full model would be

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + \mu_t \qquad \mu_t \sim N\left(0, \sigma_t^2\right)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1\mu_{t-1}^2$$

Alternative notation

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + \mu_t$$
$$\mu_t = \nu_t\sigma_t \qquad \nu_t \sim N(0,1)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1\mu_{t-1}^2$$

Properties of ARCH(1)

- ARCH(1): $\mu_t|\Omega_{t-1} \sim N(0, \sigma_t^2)$, $\Omega_{t-1} = \{y_{t-1}, \mu_{t-1}, y_{t-2}, \mu_{t-2} \cdots\}$ is the info set at the end of period $t-1$:

  $\sigma_t^2 = \alpha_0 + \alpha_1 \mu_{t-1}^2, \;\; \alpha_0 > 0, \;\; 0 \le \alpha_1 < 1$

  - Its conditional variance is time varying: Var$(\mu_t|\Omega_{t-1}) = \sigma_t^2$, CI$(95\%)$ =?
  - It is WN:(Use LIE) $E(\mu_t) = 0$, Var$(\mu_t) = \frac{\alpha_0}{1-\alpha_1}$, Cov$(\mu_t, \mu_{t-j}) = 0$

    But it is NOT independent WN or iid WN. Why?

Proof of properties

Definition (Law of Iterated Expectations)

For a random variable Y and information sets Ω_1 and Ω_2, the the LIE states that

$$E\left(Y|\Omega_1\right) = E\left(E\left(Y|\Omega_2\right)|\Omega_1\right),$$

where information set Ω_1 is included in information set Ω_2.
Example: $E\left(Y_t|\Omega_{t-2}\right) = E\left(E\left(Y_t|\Omega_{t-1}\right)|\Omega_{t-2}\right)$
Special Case: If Ω_1 is empty set, then $E\left(Y\right) = E\left(E\left(Y|\Omega_2\right)\right)$.

$$\mu_t = \nu_t \sigma_t = \nu_t \sqrt{\alpha_0 + \alpha_1 \mu_{t-1}^2}, \text{ where } \nu_t \text{ is } N(0,1)$$

1. Unconditional Expectation of μ_t: We have that $\mu_t|\Omega_{t-1} \sim N(0, \sigma_t^2)$:

$$\begin{aligned} E\left(\mu_t\right) &= E\left[E\left[\mu_t|\Omega_{t-1}\right]\right] & (1) \\ E\left[\mu_t|\Omega_{t-1}\right] &= 0 & (2) \\ E\left(\mu_t\right) &= 0. & (3) \end{aligned}$$

Proof of properties

$\mu_t = \nu_t \sigma_t = \nu_t \sqrt{\alpha_0 + \ldots}$ …ere ν_t is $N(0,1)$

2. Unconditional varia… We have that

$$
\begin{aligned}
E\left(\mu_t^2\right) &= E\left[E\left[\mu_t^2 | \Omega_{t-1}\right]\right] && (4)\\
&= E\left[E\left[\nu_t^2\left(\alpha_0 + \alpha_1 \mu_{t-1}^2\right) | \Omega_{t-1}\right]\right] && (5)\\
&= E\left[\left(\alpha_0 + \alpha_1 \mu_{t-1}^2\right) E\left[\nu_t^2 | \Omega_{t-1}\right]\right] && (6)\\
&= E\left[\alpha_0 + \alpha_1 \mu_{t-1}^2\right] = E\left[\alpha_0 + \alpha_1 E\left[\mu_{t-1}^2 | \Omega_{t-2}\right]\right] && (7)\\
&= \alpha_0 + \alpha_1 E\left[\alpha_0 + \alpha_1 \mu_{t-2}^2\right] && (8)\\
&= \cdots = \alpha_0\left(1 + \alpha_1 + \alpha_1^2 + \cdots + \alpha_1^{t-1}\right) + \alpha_1^t E\left[\mu_0^2\right] && (9)
\end{aligned}
$$

As $t \to \infty$, the unconditional variance converges if $\alpha_1 < 1$ to:
$E\left(\mu_t^2\right) = \frac{\alpha_0}{1-\alpha_1}$. $\longrightarrow$ Unconditionally, the process μ_t is **homoskedastic.**

Properties of ARCH(1)

- It can be alternatively expressed as: $\mu_t = \sigma_t v_t,\ v_t \sim iidN(0,1)$, where $v_t = \mu_t/\sigma_t$ is the **standardised shock**.
- When model is correct, v_t^2 should have no autocorrelation
- The unconditional distribution of μ_t is **NOT normal**, with heavy tails (kurtosis > 3).

MLE of ARCH(1)

- An example: $AR(1)$ -

$$y_t = c + \phi_1 y_{t-1} + \mu_t, \mu_t|\Omega_{t-1} \sim N(0, \sigma_t^2), \quad (10)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 \mu_{t-1}^2, \quad (11)$$
$$\alpha_0 > 0, 0 \leq \alpha_1 < 1. \quad (12)$$

- Likelihood of $\{y_1, y_2, \ldots, y_{T-1}, y_T\}$

$$L(\Theta) = f(y_T|\Omega_{T-1}) f(y_{T-1}|\Omega_{T-2}) \cdots f(y_2|\Omega_1) f(y_1)$$
$$f(y_t|\Omega_{t-1}) = (2\pi\sigma_t^2)^{-1/2} \exp\{-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma_t^2}\}. \quad (13)$$

- ML Estimator maximises the Log likelihood function

$$\ln L(\Theta) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=2}^{T}\left[\ln(\sigma_t^2) + \frac{(y_t - c - \phi_1 y_{t-1})^2}{\sigma_t^2}\right].$$

MLE of ARCH(1)

- ML estimators are generally consistent with an asymptotic normal distribution.

- The above holds even when the conditional normality $\mu_t|\omega_{t-1} \sim N(0, \sigma_t^2)$ is **incorrectly** assumed, as long as the conditional mean and conditional variance are correctly specified.

- With robust quasi ML standard errors, inference is standard.

Example

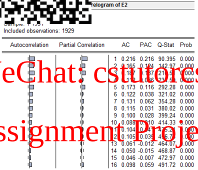
eg. NYSE composite re)-ARCH(5)

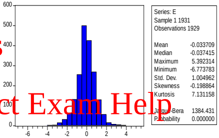
Dependent Variable: RC
Method: ML - ARCH (Marquardt) - Normal distribution

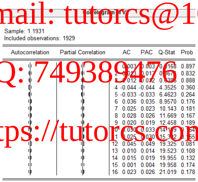
Sample (adjusted): 3 1931
Included observations: 1929 after adjustments
Convergence achieved after 17 iterations
Bollerslev-Wooldrige robust standard errors & covariance
Variance backcast: OFF
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*RESID(-2)^2 + C(6)*RESID(-3)^2 + C(7)*RESID(-4)^2 + C(8)*RESID(-5)^2

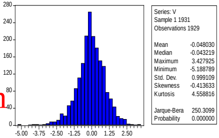
	Coefficient	Std. Error	z-Statistic	Prob.
C	0.073023	0.020198	3.615442	0.0003
AR(1)	0.109876	0.026312	4.175893	0.0000
Variance Equation				
C	0.333844	0.034675	9.627824	0.0000
RESID(-1)^2	0.163226	0.048011	3.399769	0.0007
RESID(-2)^2	0.253623	0.056578	4.482731	0.0000
RESID(-3)^2	0.084144	0.029243	2.877400	0.0040
RESID(-4)^2	0.160562	0.036483	4.400981	0.0000
RESID(-5)^2	0.059128	0.023724	2.492331	0.0127

R-squared	0.001836	Mean dependent var	0.035168
Adjusted R-squared	-0.001801	S.D. dependent var	1.006452
S.E. of regression	1.007357	Akaike info criterion	2.664282
Sum squared resid	1949.371	Schwarz criterion	[illegible]
Log likelihood	-2561.700	F-statistic	0.504830
Durbin-Watson stat	2.069097	Prob(F-statistic)	0.831451
Inverted AR Roots	.11		









Example

eg. NYSE composite re)-ARCH(5)

- Squared residuals (E2) of AR(1) have strong autocorrelation.
  Squared standardised residuals (V2) are not autocorrelated
- Residuals (E) of AR(1) have larger kurtosis.
  Standardised residuals (V) larger negative skewness.
- Normality is rejected for both E and V.

Two essential checks for the 'adequacy' of a model

- Adequate mean equation: E (residuals) has no autocorrelation;
- Adequate variance equation: V2 has no autocorrelation

Comments and limitations of ARCH

Advantages of ARCH

- It is able to capture 'clustering' in return series or the autocorrelation in squared returns.
- It facilitates volatility forecasting.
- It explains, partially, non-normality in return series.

Limitations of ARCH

- In ARCH(q), the q may be selected by AIC, SIC or LR test. The correct value of q might be very large. The model might not be parsimonious. (eg. ARCH(1) would not work for the composite return)
- The conditional variance σ_t^2 cannot be negative: Requires non-negativity constraints on the coefficients. Sufficient (but not necessary) condition is: $\alpha_i \geq 0$ for all $i = 0, 1, 2, \cdots q$. Especially for large values of q this might be violated.