Lecture 4
Modelling volatility

Zorica Mladenović

Literature:
Tsay (2010), Ch. 3 and
Brooks (2019), Ch. 9.

Structure

- **Characteristics of financial data**
- **ARCH models**
- **How to test for the ARCH effect?**
- **GARCH models**
- **Forecasting volatility**
- **Some modifications of GARCH models**
- **Note on estimation**
- **Examples in EVIEWS**

Key characteristics of financial data (Tsay, 2010)

Linear models (ARMA or ARIMA) cannot explain a number of relevant features often found in financial data.

1. **Leptokurtosis (high kurtosis)**

   • Time series tend to have heavy/fat tails, i.e. empirical distribution is highly peaked at the mean.

2. **Volatility clustering**

   • "... Large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes..." (Mandelbrot, 1963)

   • Turbulent (high variability) period is followed by quiet (low variability) period; these subperiods are recurrent but not in a periodic way.

   • It is described by conditional variance of time series (which is not directly observable)

3. **Leverage effect**

   • There is a tendency for changes in prices to be negatively correlated with changes in volatility.

   • Volatility tends to rise more with a large price fall ("bad news") than with a price increase ("good news") of the same magnitude.

Additional characteristics (Bollerslev, Engle and Nelson, 1994)
in Handbook of Econometrics, Vol. IV, eds. Engle and McFadden

4. Non-trading periods

• Information that accumulates when financial markets are closed is reflected in prices after the markets reopen.

5. Forecastable events

• Forecastable releases of important information are associated with high ex ante volatility.

6. Volatility and autocorrelation

• There are evidences of strong inverse relation between volatility and autocorrelation.

7. Co-movements in volatiities

• "... It seems fair to say that when stock volatilities change, they all tend to change in the same direction." (Black, 1976)

Autoregressive conditional heteroskedasticity model (ARCH)

Engle (1982) – The Nobel prize winner in 2003

We start by defining conditional mean, μ_t, and conditional variance, σ_t^2,

$$\mu_t = E\left[r_t | F_{t-1}\right]$$

$$\sigma_t^2 = E\left[\left(r_t - \mu_t\right)^2 | F_{t-1}\right]$$

F_{t-1} **is information set available at time** $t-1$

For r_t it is assumed to follow ARMA specification:

$$r_t = \mu_t + a_t$$

so that μ_t:

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j}$$

What about error term a_t?

$$a_t = \sigma_t \epsilon_t,\ \epsilon_t : \mathbf{iid}(0,1)$$

For conditional variance σ_t^2 it is assumed:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2$$

$\alpha_0 > 0,\ \alpha_1, \ldots, \alpha_m \geq 0$

This is $ARCH(m)$ model.

Distributions for ϵ_t:

$\mathcal{N}(0,1)$ **distribution**

Standardized t distribution

Generalized error distribution.

Model implies the following:

Large past squared shocks $a_{t-1}^2, \ldots, a_{t-m}^2$ cause a large conditional variance σ_t^2.

ARCH (1) model

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2, \quad a_t = \sigma_t \epsilon_t$$
$$\alpha_0 > 0,\ \alpha_1 \geq 0$$

The unconditional mean of a_t is 0:

$$E(a_t) = E\left[E(a_t|F_{t-1})\right] = E\left[\sigma_t E(\epsilon_t)\right] = 0$$

The unconditional variance of a_t is $\frac{\alpha_0}{1-\alpha_1}$:

$$\mathbf{var}(a_t) = E(a_t^2) = E\left[E(a_t^2|F_{t-1})\right] = E\left[\alpha_0 + \alpha_1 a_{t-1}^2\right]$$

$$\mathbf{var}(a_t) = \alpha_0 + \alpha_1 E\left(a_{t-1}^2\right)$$

$$\mathbf{var}(a_t) = \alpha_0 + \alpha_1 \underbrace{E\left(a_{t-1}^2\right)}_{\mathbf{var}(a_t)}$$

$$\mathbf{var}(a_t) = \frac{\alpha_0}{1-\alpha_1}$$

$$\Longrightarrow 0 \leq \alpha_1 < 1$$

For the fourth moment to be finite, $E(a_t^4)$, (proof is skipped):

$\Longrightarrow 0 \leq \alpha_1^2 < \frac{1}{3}$

Characteristics of ARCH models (Tsay, 2010, 2013)

The key advantages (Tsay, 2013):

- **Models can produce volatility clusters.**
- **The shocks a_t have heavy tails.**

Weakness of ARCH models (Tsay, 2010):

- **Positive and negative shocks have the same effects on volatility.**
- **They are restrictive in terms of parameter values given the conditions needed for the fourth moment to be finite.**
- **They seem as models without clear economic implication.**
- **They have a tendency to overpredict the volatility, because they respond slowly to large isolated shocks.**

How to test for ARCH effect?

Identification of the conditional mean equation is conducted by standard procedures (the Box-Jenkins approach).

Let $\widehat{a}_t$ be residuals from the mean equation.

They are used to investigate the presence of conditional heteroskedasticity or ARCH effects.

Two test-statistics are implemented:

a) *The Box-Ljung Q^2 statistic*

b) *The Engle LM statistic.*

a) **The Box-Ljung Q^2 statistic is just the Box-Ljung statistic, but applied on squared residuals, $\widehat{a}_t^2$.**

Under the null hypothesis of constant variance, it is assumed that first m elements of ACF of a_t^2 are zero:

$$H_0 : \rho_1^\star = \rho_2^\star = \ldots = \rho_m^\star = 0,$$

where $\rho_l^\star$, $l = 1, \ldots, m$ is the lag-l ordinary autocorrelation coefficient of a_t^2.

b) The Engle statistic (Lagrange multiplier test) has the similar idea as Q^2. The baseline model is:

$$\widehat{a}_t^2 = d_0 + d_1\widehat{a}_{t-1}^2 + d_2\widehat{a}_{t-2}^2 + \ldots + d_m\widehat{a}_{t-m}^2 + e_t$$

- **The null hypothesis of constant variance implies no autocorrelation in $\widehat{a}_t^2$:**

  $$H_0 : d_1 = d_2 = \ldots = d_m = 0.$$

- **The validity of the null is checked by F-test that compares residual sum of squares of 2 models: restricted and unrestricted. Both models are estimated by the OLS. Numbers of degrees of freedom are: m and $T - 2m - 1$.**

  Notation: $ARCH\, F(m)$**.**

- **Asymptotically, this test-statistic has χ^2 distribution with m d.o.f. In this case it is calculated as T (or $T - m$) times R^2 of the baseline model.**

  Notation: $ARCH\, \chi^2(m)$**.**

- **Consequence: The order of ARCH model is suggested by the number of significant SPACF coefficients of $\widehat{a}_t^2$.**

Generalized autoregressive conditional heteroskedasticity model (GARCH)
Bollerslev (1986)

$$a_t = \sigma_t \epsilon_t, \; \epsilon_t : \mathbf{iid}(0,1)$$

Conditional variance is defined as:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

$\alpha_0 > 0, \; \alpha_1, \ldots, \alpha_m \geq 0, \beta_1, \ldots, \beta_s \geq 0, \; \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$

Additionally, $\alpha_i = 0$ **for** $i > m$**,** $\beta_j = 0$ **for** $j > s$**.**

Notation: $GARCH(m, s)$**.**

Given above conditions model enables constant unconditional variance.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

Let $\xi_t = a_t^2 - \sigma_t^2$. Then,

$$\sigma_t^2 = a_t^2 - \xi_t \text{ and } \sigma_{t-i}^2 = a_{t-i}^2 - \xi_{t-i},\ i = 0, 1, \ldots, s.$$

$GARCH(m, s)$ **is:**

$$a_t^2 = \alpha_0 + \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2 + \xi_t - \sum_{j=1}^{s} \beta_j \xi_{t-j}$$

New error term ξ_t has zero mean and zero autocovariance coefficients. However, it can be shown that it is not iid sequence.

We have reached $ARMA(m^*, s)$ model for a_t^2, and $m^* = max(m, s)$.

Conclusion:

$GARCH$ **model is** $ARMA$ **model for squared error term a_t^2.**

Consequence for unconditional variance:

$$E(a_t^2) = \alpha_0 + \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i) E(a_{t-i}^2) + E(\xi_t) - \sum_{j=1}^{s} \beta_j E(\xi_{t-j})$$

$$\mathbf{var}(a_t^2) = E\left(a_t^2\right) = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} (\alpha_i + \beta_i)}$$

GARCH(1,1) model

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

$$\alpha_0 > 0,\ \alpha_1 \geq 0,\ \beta_1 \leq 1,\ \alpha_1 + \beta_1 < 1$$

Two issues are considered:

1. Relation with ARCH model.

2. Forecasting volatility.

1. Relation with ARCH model

GARCH(1,1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

When we replace σ_{t-1}^2, σ_{t-2}^2, etc.

$$\begin{aligned}
\sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\
&= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \underbrace{\left(\alpha_0 + \alpha_1 a_{t-2}^2 + \beta_1 \sigma_{t-2}^2\right)}_{\sigma_{t-1}^2} \\
&= \alpha_0(1+\beta_1) + \alpha_1 a_{t-1}^2 + \alpha_1 \beta_1 a_{t-2}^2 + \beta_1^2 \underbrace{\left(\alpha_0 + \alpha_1 a_{t-3}^2 + \beta_1 \sigma_{t-3}^2\right)}_{\sigma_{t-2}^2} \\
&= \ldots \\
&= \alpha_0(1+\beta_1+\beta_1^2+\beta_1^3+\ldots) + \alpha_1 a_{t-1}^2 + \alpha_1\beta_1 a_{t-2}^2 + \alpha_1\beta_1^2 a_{t-3}^2 + \ldots
\end{aligned}$$

For $|\beta_1| < 1$**:**

$$\sigma_t^2 = \frac{\alpha_0}{1-\beta_1} + \alpha_1 a_{t-1}^2 + \alpha_1 \beta_1 a_{t-2}^2 + \alpha_1 \beta_1^2 a_{t-3}^2 + \ldots$$

- $GARCH(1,1)$ **has** $ARCH(\infty)$ **representation.**

- **Given:** $0 < \beta_1 < 1$**, the impact of** a_{t-i}^2 **on volatility diminishes quickly with increasing** i**.**

- $GARCH(1,1)$ **model is more parsimonious specification than** $ARCH$ **model. With only three parameters** $GARCH$ **captures the influence of all shocks.**

2. Forecasting volatility

Relevant for:

- **Measuring risk, as in Riskmetrics**
- **Estimating value at risk (VaR)**
- **GARCH models may underpredict VaR, especially if its derived under the assumption of normality!!! Combination with EVT approach is useful.**

Based on h observations future values of volatility are forecasted.

The forecast origin is h.

$GARCH(1,1)$ **model:**

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad a_t = \sigma_t \epsilon_t$$

True value of conditional variability in $h+1$ is:

$$\begin{aligned}\sigma_{h+1}^2 &= \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2 \\ &= \alpha_0 + \alpha_1 \left(\epsilon_h^2 \sigma_h^2\right) + \beta_1 \sigma_h^2 + \alpha_1 \sigma_h^2 - \alpha_1 \sigma_h^2 \\ &= \alpha_0 + \left(\alpha_1 + \beta_1\right) \sigma_h^2 + \alpha_1 \left(\epsilon_h^2 - 1\right) \sigma_h^2\end{aligned}$$

One-step ahead forecast $\widehat{\sigma}_h^2(1)$ of σ_{h+1}^2 is:

$$\begin{aligned}\widehat{\sigma}_h^2(1) &= E\left(\sigma_{h+1}^2 | F_h\right) \\ &= \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2\end{aligned}$$

True value of conditional variability in $h+2$ is:

$$\sigma^2_{h+2} = \alpha_0 + \alpha_1 a^2_{h+1} + \beta_1 \sigma^2_{h+1}$$

$$= \alpha_0 + (\alpha_1 + \beta_1)\, \sigma^2_{h+1} + \alpha_1 \left(\epsilon^2_{h+1} - 1\right) \sigma^2_{h+1}$$

Two-step ahead forecast $\widehat{\sigma}^2_h(2)$ of σ^2_{h+2} is:

$$\begin{aligned}
\widehat{\sigma}^2_h(2) &= E\left(\sigma^2_{h+2}|F_h\right) \\
&= \alpha_0 + (\alpha_1 + \beta_1)\, \widehat{\sigma}_h(1) + \alpha_1 E\left[(\epsilon^2_{h+1} - 1)|F_h\right] \sigma^2_{h+1} \\
&= \alpha_0 + (\alpha_1 + \beta_1)\, \widehat{\sigma}_h(1) \text{ \textbf{because} } E\left[(\epsilon^2_{h+1} - 1)|F_h\right] = 0
\end{aligned}$$

Three-step ahead forecast $\widehat{\sigma}^2_h(3)$ of σ^2_{h+3} is:

$$\begin{aligned}
\widehat{\sigma}^2_h(3) &= E\left(\sigma^2_{h+3}|F_h\right) \\
&= \alpha_0 + (\alpha_1 + \beta_1)\, \widehat{\sigma}_h(2) \\
&= \alpha_0 + (\alpha_1 + \beta_1)\left[\alpha_0 + (\alpha_1 + \beta_1)\, \widehat{\sigma}_h(1)\right] \\
&= \alpha_0 \underbrace{\left(1 + (\alpha_1 + \beta_1)\right)}_{\frac{\left[1 - (\alpha_1 + \beta_1)^2\right]}{1 - (\alpha_1 + \beta_1)}} + (\alpha_1 + \beta_1)^2\, \widehat{\sigma}_h(1)
\end{aligned}$$

In general, for any forecasting horizon l,

l-step ahead forecast $\widehat{\sigma}_h^2(l)$ of σ_{h+l}^2:

$$\widehat{\sigma}_h^2(l) = \alpha_0 \left[1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 + \ldots + (\alpha_1 + \beta_1)^{(l-2)}\right]$$

$$+ (\alpha_1 + \beta_1)^{(l-1)}\, \widehat{\sigma}_h^2(1)$$

$$= \alpha_0 \frac{\left[1 - (\alpha_1 + \beta_1)^{(l-1)}\right]}{1 - (\alpha_1 + \beta_1)} + (\alpha_1 + \beta_1)^{(l-1)}\, \widehat{\sigma}_h^2(1)$$

Since $\alpha_1 + \beta_1 < 1$, for relatively long horizon l ($l \to \infty$) forecasts converge to unconditional variance of a_t:

$$\widehat{\sigma}_h^2(l) \to \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}$$

- **If $\alpha_1 + \beta_1 = 1$, then**

  GARCH becomes integrated GARCH model, IGARCH(1,1).

  It does not have stable unconditional variance (Why?)

- **Volatility forecast is:**

$$\widehat{\sigma}_h^2(l) = \alpha_0(l-1) + \widehat{\sigma}_h^2(1)$$

- **This forecast depends on the one-step forecast and further linearly increases according to α_0 and forecast horizon l.**

- **The one-step ahead forecast has long-lasting effect on multi-step forecast.**

- **When $\alpha_0 = 0$, then $\widehat{\sigma}_h^2(l) = \widehat{\sigma}_h^2(1)$. This is commonly used in practical work.**

Some modifications of GARCH models

1. ARCH-M, 2. TGARCH, and 3. EGARCH.

1. ARCH-M model: ARCH in mean

Only mean equation is modified in the following way: new explanatory variable is added being conditional variance from the volatility equation:

$$r_t = \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{j=0}^{q} \theta_j a_{t-j} + \zeta \sigma_t^2$$

- **Parameter of σ_t^2, ζ measures risk premium.**

- **If $\zeta > 0$, then increased risk leads to a rise in the mean return.**

- **Instead of σ_t^2 the following variables may appear: σ_t or $\ln(\sigma_t)$.**

- **All three measures may be included not only in time t, but for example in time t-1.**

2. TGARCH model: Threshold *GARCH* model

Different name:

The GJR model (Glosten, Jagannathan and Runkle, 1993).

- **This model accounts for the possibility that volatility reacts asymmetrically to negative and positive random shocks.**

- **Negative shock in the previous period ($a_{t-1} < 0$) is likely to influence current volatility more than positive shock of the same magnitude ($a_{t-1} > 0$).**

- **Volatility equation for $TGARCH(1,1)$ model is:**

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_1^\star N_{t-1} a_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

$$N_{t-1} = \begin{cases} 1, & a_{t-1} < 0 \\ 0, & a_{t-1} \geq 0 \end{cases}$$

- **Impact of the positive shock: α_1.**

- **Impact of the negative shock: $\alpha_1 + \alpha_1^\star$.**

- **It is expected that: $\alpha_1^\star > 0$.**

Tests for asymmetries in volatilities (Engle and Ng, 1993)

1. **Negative sign bias test**
2. **Negative size bias test**
3. **Joint test for sign and size bias**

They are based on estimated residuals:

- **Ordinary residuals,** $\widehat{a}_t$**, or**
- **Standardized residuals,** $\dfrac{\widehat{a}_t}{\widehat{\sigma}_t}$**.**

1. Negative sign bias test

$$S_{t-1}^{-} = \begin{cases} 1, & \widehat{a}_{t-1} < 0 \\ 0, & \widehat{a}_{t-1} \geq 0 \end{cases}$$

The following equation is set:

$$\widehat{a}_t^2 = \gamma_0 + \gamma_1 S_{t-1}^{-} + \textbf{error term}$$

The negative sign bias test is test for the significance of γ_1**.**

2. Negative size bias test

$$S_{t-1}^{-} = \begin{cases} 1, & \widehat{a}_{t-1} < 0 \\ 0, & \widehat{a}_{t-1} \geq 0 \end{cases}$$

The following equation is considered:

$$\widehat{a}_t^2 = \gamma_0 + \gamma_2 \underbrace{S_{t-1}^{-}\widehat{a}_{t-1}}_{\textbf{Interactive dummy}} + \textbf{error term}$$

The negative size bias test is test for the significance of γ_2.

$$S_{t-1}^{-}\widehat{a}_{t-1} = \begin{cases} \widehat{a}_{t-1}, & \widehat{a}_{t-1} < 0 \\ 0, & \widehat{a}_{t-1} \geq 0 \end{cases}$$

3. Sign and size bias test

$$S^{+}_{t-1} = 1 - S^{-}_{t-1}$$

Relevant model is:

$$\widehat{a}_t^2 = \gamma_0 + \gamma_1 S^{-}_{t-1} + \gamma_2 S^{-}_{t-1}\widehat{a}_{t-1} + \gamma_3 S^{+}_{t-1}\widehat{a}_{t-1} + \textbf{error term}$$

The joint significance test, $H_0 : \gamma_1 = \gamma_2 = \gamma_3 = 0$**, is:**

$$TR^2 : \chi^2_3$$

and R^2 **is coefficient of determination from the above estimated regression.**

Individual t**-ratio may also be used. If the parameter given below is significant:**

- $\gamma_1 \Longrightarrow$ **Significant negative sign bias exists.**
- $\gamma_2 \Longrightarrow$ **Significant negative size bias exists.**
- $\gamma_3 \Longrightarrow$ **Significant positive size bias exists.**

3. EGARCH model: Exponential GARCH model (Nelson, 1991)

- **Model for volatility (one of many forms) is:**

$$\ln(\sigma_t^2) = \beta_0 + \alpha_1 \left(\frac{a_{t-1}}{\sigma_{t-1}} \right) + \alpha_1^{\star} \left| \frac{a_{t-1}}{\sigma_{t-1}} \right| + \beta_1 \ln(\sigma_{t-1}^2)$$

- **The logarithmic transformation guarantees that variance will never be negative.**

- **Standardized shock is introduced:** $\frac{a_t}{\sigma_t}$.

- **Relative size of the shock is relevant, and also the sign.**

- **The impact of the positive standardized shock:**

  $\alpha_1 + \alpha_1^{\star}$.

- **The impact of the negative standardized shock:**

  $-\alpha_1 + \alpha_1^{\star}$.

Note on the estimation of GARCH models

- **Method of the maximum likelihood (ML) is usually based on the maximization of log-likelihood function of the form (assuming that the error term is normally distributed):**

$$-\frac{T}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\ln(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}\frac{a_t^2}{\sigma_t^2}$$

- **The presence of time-varying variance σ_t^2 implies that the maximum of the function cannot be found analytically, but numerically.**

- **Numerical solutions are based on the application of optimization methods. Given that model is correctly specified consistent estimators are reached.**

- **In the case error term does not have normal distribution, quasi maximum likelihood (QML) is applied with different standard errors derived from the Bollerslev-Wooldridge approach. Consistent estimators are provided by this modification.**

Note on the optimization methods

- **Methods available by softwares (like EVIEWS) are based on the determination of the first and second derivatives of the log-likelihood function with respect to the parameters values at each iteration (the gradient and Hessian matrix respectively).**

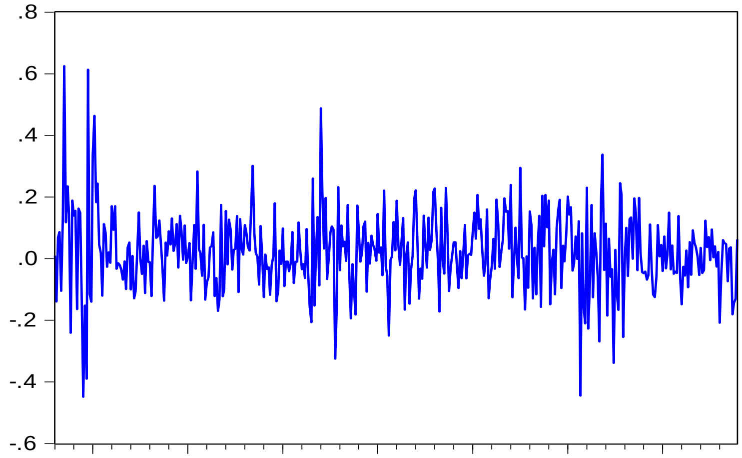
- **They are basically modification and substantial computational improvement of the iterative Gauss-Newton algorithm.**

- **The BHHH (Bernt, Hall, Hall, Hausman) algorithm combines calculation of the first derivatives numerically and calculation of the approximation of the second derivatives. Such an approach increases the calculation speed.**

- **Empirical results may differ across different softwares.**

ARCH modeling – EVIEWS

1. The monthly returns of Intel stock – ARCH1.wf1; Tsay (2010), p. 123 - 124. The example is modified by extending the sample (1973-2008) and estimated models (see also Tsay, 2013, *An Introduction to Analysis of Financial Data with R*, Wiley) This new sample containing simple returns is considered here.



The squared mean corrected values of series exhibit strong autocorrelation, suggesting the presence of ARCH structure.

Dependent Variable: RT
Method: Least Squares
Sample: 1973M01 2008M12
Included observations: 432

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.022139	0.006181	3.581808	0.0004

R-squared	0.000000	Mean dependent var	0.022139
Adjusted R-squared	0.000000	S.D. dependent var	0.128466
S.E. of regression	0.128466	Akaike info criterion	-1.263994
Sum squared resid	7.113006	Schwarz criterion	-1.254576
Log likelihood	274.0227	Hannan-Quinn criter.	-1.260276
Durbin-Watson stat	1.956371		

ACF and PACF of residuals

	AC	PAC	Q-Stat	Prob
1	0.022	0.022	0.2047	0.651
2	0.007	0.006	0.2255	0.893
3	0.080	0.080	3.0417	0.385
4	-0.054	-0.058	4.3330	0.363
5	-0.014	-0.013	4.4246	0.490
6	0.044	0.039	5.2716	0.509
7	-0.109	-0.103	10.547	0.160
8	-0.082	-0.080	13.540	0.095
9	-0.008	-0.011	13.571	0.138
10	0.029	0.052	13.938	0.176
11	-0.057	-0.058	15.380	0.166
12	0.046	0.037	16.319	0.177

ACF and PACF of squared residuals

	AC	PAC	Q-Stat	Prob
1	0.168	0.168	12.217	0.000
2	0.136	0.111	20.286	0.000
3	0.207	0.176	39.092	0.000
4	0.193	0.134	55.353	0.000
5	0.114	0.037	61.031	0.000
6	0.083	0.002	64.046	0.000
7	0.112	0.037	69.627	0.000
8	0.051	-0.021	70.800	0.000
9	0.077	0.031	73.424	0.000
10	0.053	0.002	74.691	0.000
11	0.042	0.001	75.490	0.000
12	0.158	0.134	86.607	0.000

Dependent Variable: RT
Method: ML - ARCH (BHHH) - Normal distribution
Included observations: 432
Convergence achieved after 11 iterations
Bollerslev-Wooldridge robust standard errors & covariance
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	**0.019366**	0.005707	3.393108	0.0007
	Variance Equation			
C	0.010984	0.001117	9.836414	0.0000
RESID(-1)^2	0.382927	0.101562	3.770381	0.0002

R-squared	-0.000467	Mean dependent var	0.022139
Adjusted R-squared	-0.000467	S.D. dependent var	0.128466
S.E. of regression	**0.128496**	Akaike info criterion	-1.326643
Sum squared resid	7.116328	Schwarz criterion	-1.298390
Log likelihood	289.5549	Hannan-Quinn criter.	-1.315489

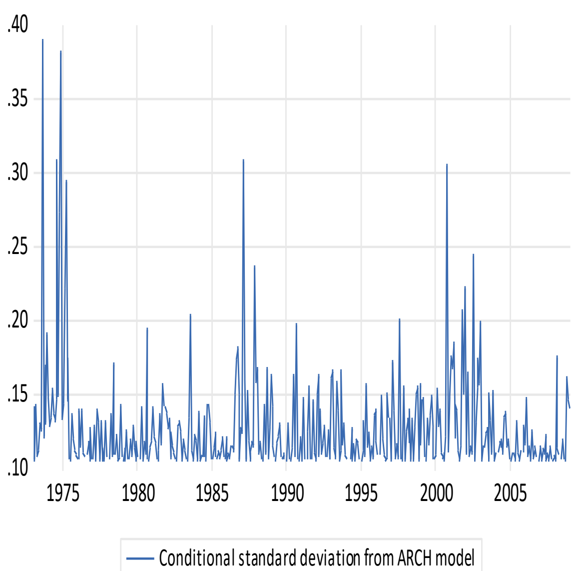
Interpretation of estimated parameters:

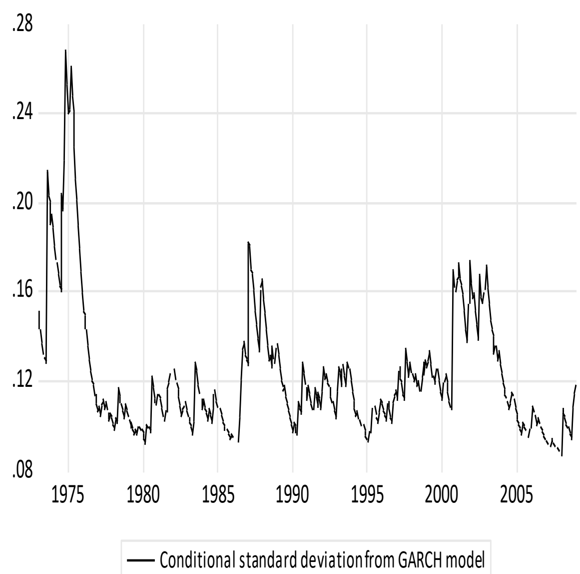
- **Expected monthly simple return for Intel stock is about 1.94%.**
- **The unconditional variance is 0.010984/(1-0.382927)=0.017800, which is close to (S.E. of regression)^2 = (0.128496)^2**

However, ARCH(1) model does not perform well overall. The statistics reported are based on standardized residuals: Q(12)=13.03(0.37), Q^2(10)=13.46(0.20), Q^2(12)=24.80(0.02), Q^2(24)=36.45(0.05), ARCH(12)=26.68(0.01), ARCH(24)=25.96(0.36).

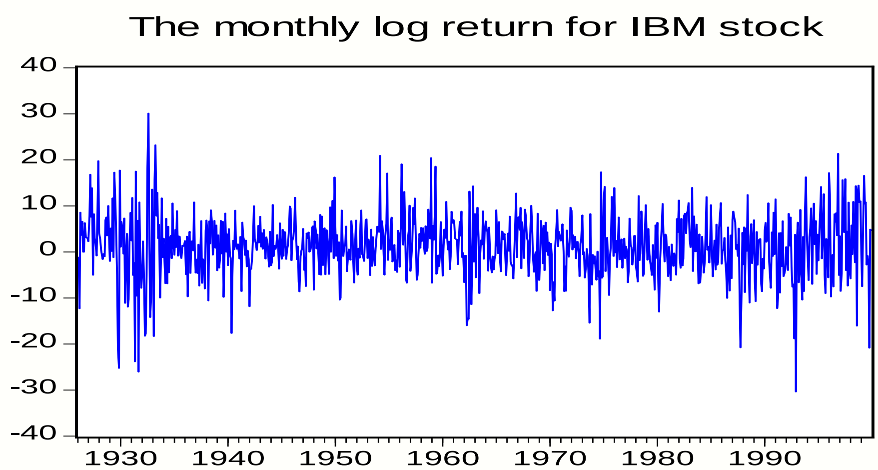
We can notice that significant autocorrelation exists at lag 12, which is probably due to seasonality. If GARCH(1,1) model is used instead of ARCH(1) this autocorrelation almost disappears.

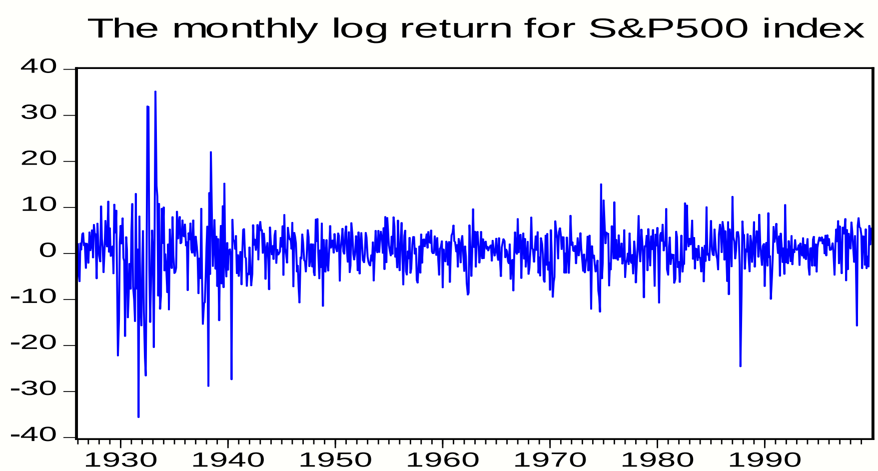
Models estimate volatility differently (graphed as conditional standard deviation).





2. The monthly log returns for IBM stock and S&P500 index, January 1926 – December 1999, 888 observations – ARCH2.wf1; Tsay (2010), p. 158-159.





2.1 GARCH(1,1) model for the monthly log return for S&P500 (SP)

Dependent Variable: SP
Method: ML ARCH - Normal distribution (OPG - BHHH / Marquardt
steps)

Sample: 1926M01 1999M06
Included observations: 882
Convergence achieved after 42 iterations
Coefficient covariance computed using Bollerslev-Wooldridge QML
sandwich with expected Hessian
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	0.679662	0.146596	4.636310	0.0000
		Variance Equation		
C	0.629930	0.350780	1.795795	0.0725
RESID(-1)^2	0.115497	0.028551	4.045335	0.0001
GARCH(-1)	0.867804	0.033304	26.05669	0.0000

R-squared	-0.000672	Mean dependent var	0.533113
Adjusted R-squared	-0.000672	S.D. dependent var	5.655301
S.E. of regression	5.657202	Akaike info criterion	5.960518
Sum squared resid	28195.47	Schwarz criterion	5.982206
Log likelihood	-2624.589	Hannan-Quinn criter.	5.968811
Durbin-Watson stat	1.844350		

2.2 GARCH(1,1) model for SP with exogenous variable in the volatility equation (exogenous variable is the lag-one squared mean corrected return on IBM stock)

Dependent Variable: SP
Method: ML ARCH - Normal distribution (Marquardt / EViews legacy)

Sample (adjusted): 1926M02 1999M06
Included observations: 881 after adjustments
Convergence achieved after 6 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1) + C(5)*(IBM(-1)
-1.24)^2

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.526561	0.134108	3.926389	0.0001
		Variance Equation		

C	1.045882	0.157732	6.630772	0.0000
RESID(-1)^2	0.124571	0.016049	7.761867	0.0000
GARCH(-1)	0.851386	0.014558	58.48363	0.0000
(IBM(-1)-1.24)^2	-0.008718	7.97E-14	-1.09E+11	0.0000

R-squared	-0.000001	Mean dependent var	0.531196
Adjusted R-squared	-0.000001	S.D. dependent var	5.658227
S.E. of regression	5.658229	Akaike info criterion	5.963644
Sum squared resid	28173.69	Schwarz criterion	5.990778
Log likelihood	-2621.985	Hannan-Quinn criter.	5.974020
Durbin-Watson stat	1.844174		

2.3. Fitted conditional variances for SP from July to December 1999 using models with (GF2) and without (GF1) the past log returns of IBM stock.

	GF1	GF2
1999M07	**26.35839**	**21.81075**
1999M08	**25.29412**	**21.25707**
1999M09	**22.77741**	**19.27321**
1999M10	**21.87333**	**18.76497**
1999M11	**22.96281**	**16.59937**
1999M12	**20.72576**	**15.29261**

The inclusion of the lag-1 return of IBM stock reduces the volatility of the S&P500 index return.