Volatility models: ARCH model

(Hamilton: Chapters 21, Tsay: Chapter 3)

Isabel Casas

Lecture Outline

- Why do we care about volatility?
- Empirical regularities of asset returns
- Engle's ARCH model
- Testing for ARCH effects
- Estimating ARCH models
- Bollerslev's GARCH model

SYDDANSK UNIVERSITET

What do you think?

What do we call volatility?

1. The process unconditional variance
2. The process conditional variance

- What does large volatility mean?
- What is volatility clustering?
- Can we observe volatility?
- Should we take it into account then?

Syddansk Universitet

Why Do We Care About Volatility?

1. Many derivative securities depend explicitly on volatility

Example: Black-Scholes call option price:

$$C_t^{BS}(\sigma_t) = S_0\Phi(d_1) - Kr^{-rT}\Phi(x - \sigma_t\sqrt{T})$$

$$\begin{aligned}
d_1 &= \frac{\ln(S_0/K) + (r + \sigma_t^2/2)}{\sigma\sqrt{T}} \\
S_0 &= \text{current stock price} \\
T &= \text{time to maturity} \\
r &= \text{risk free interest rate} \\
\sigma_t &= \sqrt{Var(r_t|F_t)}
\end{aligned}$$

Syddansk Universitet

Why Do We Care About Volatility?

Note: derivation of Black-Scholes formula assumes constant volatility!
However, Black-Scholes implied volatility is time-varying

$$\sigma_t^{implied} : C_t^{observed} - C_t^{BS}(\sigma_t^{implied}) = 0$$

If Black-Scholes assumptions were correct then

$$\sigma_t^{implied} = \bar{\sigma} = \text{ constant}$$

Note: $\sigma_t^{implied}$ is an observable time series of volatility estimateds based on a model for option prices.

Why Do We Care About Volatility?

2. Risk management measures such as value-at-risk (VaR) and expected shortfall (ES) depend explicitly on volatility

Let F denote the distribution of dollar losses on a portfolio of assets. Let r_t denote the daily continuously compounded return on the portfolio and let W_0 denote the initial value of the portfolio. Then the daily dollar return is

$$W_0 \exp(r_t)$$

By convention, the loss distribution F is the distribution of

$$L_t = -(W_0 \exp(r_t) - W_0) = -W_0(\exp(r_t) - 1) = -W_0 r_t$$

where $r_t = \exp(r_t) - 1$ is the simple return.

Why Do We Care About Volatility?

Value-at-Risk (VaR). For $0.95 \leq q > 1$, say, VaR_q is the qth quantile of the distribution F

$$VaR_q = F^{-1}(q)$$

where F^{-1} is the inverse of F.

Expected Shortfall (ES). ES_q is the expected loss size, given that VaR_q is exceeded:

$$ES_q = E[L|L > VaR_q]$$

Note: ES_q is related to VaR_q via

$$ES_q = VaR_q + E[L - VaR_q|L > VaR_q]$$

SYDDANSK UNIVERSITET

Why Do We Care About Volatility?

Example: VaR and ES for normal distribution: $L \sim N(\mu, \sigma^2)$

VaR:

$$\begin{aligned} VaR_q =& \mu + \sigma z_q \\ z_q =& q \cdot 100\% \text{ quantile for N (0, 1)} \end{aligned}$$

ES:

$$\begin{aligned} ES_q =& \mu + \sigma \frac{\phi(z)}{1 - \Phi(z)} = \mu + \sigma \frac{\phi(z_q)}{1 - q} \\ z =& (VaR_q - \mu)/\sigma \end{aligned}$$

SYDDANSK UNIVERSITET

Why Do We Care About Volatility?

3. Portfolio allocation in a Markowitz mean-variance framework depends explicitly on volatility (also covariance/correlation)

Let $r_t \sim N(\mu, \sigma^2)$ be an $n \times 1$ vector of simple monthly returns, and let w be an $n \times 1$ vector of portfolio weights that sum to one. Let $R_{p,t} = w' r_t$ denote the simple return on the portfolio. The Markowitz mean-variance efficient portfolio w with target expected return μ_p^0 solves

$$\min_w var(R_{p,t}) = w'\Sigma w \text{ s.t. } w'\mu = \mu_p^0 \text{ and } w'\mathbf{1} = \mathbf{1}$$

where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \dots \sigma_{2n} & \\ \vdots & & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \dots & \sigma_n^2 \end{pmatrix}$$

Why Do We Care About Volatility?

4. Modeling the volatility of a time series can improve the efficiency in parameter estimation (e.g. feasible GLS) and the accuracy in interval foresting (i.e, provide correct standard error bands for forecasts)

Empirical Regularities of Asset Returns

1. Thick tails
   - Excess kurtosis decreases with aggregation
2. Volatility clustering.
   - Large changes followed by large changes; small changes followed by small changes
3. Leverage effects
   - changes in prices often negatively correlated with changes in volatility
4. Non trading periods
   - Volatility is smaller over periods when markets are closed than when they are open

Empirical Regularities of Asset Returns

5. Forecastable events
   - Forecastable releases of information are associated with high ex ante volatility
6. Volatility and serial correlation
   - Inverse relationship between volatility and serial correlation of stock indices
7. Volatility co-movements
   - Evidence of common factors to explain volatility in multiple series

Empirical Regularities of Asset Returns

- (Log) prices are nonstationary and show dynamic properties in line with processes that are integrated of order one.
- Therefore, we focus our analysis on (log) price returns.
- Changes of compounded return are usually not autocorrelated
- Daily price variations (high frequency) exhibit:
  - positive autocorrelation
  - periods of higher and smaller price variations alternate (empirical volatilities tend to cluster)
  - So although prices are hardly predictable, the variance of the forecast error in time dependent and can be eastimated by means of observed past variations

SYDDANSK UNIVERSITET

Engle's ARCH(m) Model

The ARCH(m) model for $r_t = \ln P_t - \ln P_{t-1}$ (compounded returns) is:

$$
\begin{aligned}
r_t =& \mu_t + \epsilon_t, \qquad \epsilon_t | F_{t-1} \sim IID(0, {\sigma_t}^2) \\
\sigma_t^2 =& a_0 + a_1 \epsilon_{t-1}^2 + \ldots + a_m \epsilon_{t-m}^2, \qquad a_0 > 0, a_i \geq 0 \\
\epsilon_t =& z_t \sigma_t \qquad z_t \sim IID(0,1)
\end{aligned}
$$

- Large past shocks $\{\epsilon_{t-i}^2\}_{i=1}^m$ imply a large conditional variance σ_t^2
- As a consequence, the shocks ϵ_t are large too
- This means that large shocks *tend* to be followed by large shocks (volatility clustering)

Engle's ARCH(m) Model

Remarks:

- In practice, it is often assumed that z_t follows a standard normal or a standarised t-Student.
- Some authors denote the conditional variance by h_t, and therefore $\epsilon_T = \sqrt{h_t} z_t$

SYDDANSK UNIVERSITET

Engle's ARCH(p) Model

The conditional mean and variance of r_t

$$\mu_t = E(r_t | F_{t-1})$$

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

Properties of ARCH errors

1. Conditional moments of shocks ϵ_t

$$
\begin{aligned}
E[\epsilon_t|F_{t-1}] =& E[z_t\sigma_t|F_{t-1}] = \sigma_t E[z_t|F_{t-1}] = 0 \\
var(\epsilon_t|F_{t-1}) =& E[\epsilon_t^2|F_{t-1}] = \sigma_t^2 E[z_t^2|F_{t-1}] = \sigma_t^2 \\
E[\epsilon_t^m|F_{t-1}] =& 0 \text{ for } m \text{ odd} \\
E[\epsilon_t\epsilon_{t-j}] = 0& \text{ for } j = 1, 2, \ldots
\end{aligned}
$$

SYDDANSK UNIVERSITET

Properties of ARCH Errors

2. The error ϵ_t is stationary with mean zero and constant unconditional variance

$$
\begin{aligned}
E[\epsilon_t] =& E[E[z_t \sigma_t | F_{t-1}]] \\
=& E[\sigma_t E[z_t | F_{t-1}]] = 0
\end{aligned}
$$

$$
\begin{aligned}
var(\epsilon_t) =& E[\epsilon_t^2] = E[E[z_t^2 \sigma_t^2 | F_{t-1}]] \\
=& E[\sigma_t^2 E[z_t^2 | F_{t-1}]] = E[\sigma_t^2]
\end{aligned}
$$

Properties of ARCH Errors

$$
\begin{aligned}
E[\sigma_t^2] =& E[a_0 + a(L)\epsilon_t^2] \\
=& a_0 + a_1 E[\epsilon_{t-1}^2] + \ldots + a_p E[\epsilon_{t-p}^2] \\
=& a_0 + a_1 E[\sigma_t^2] + \ldots + a_p E[\sigma_{t-p}^2]
\end{aligned}
$$

Assuming stationarity $(E(\sigma_t^2) = E(\sigma_{t-1}^2) = \bar{\sigma}^2)$

$$
E[\sigma_t^2] = \bar{\sigma}^2 = \frac{a_0}{1 - a_1 - \ldots - a_p} = \frac{a_0}{a(1)}
$$

Since the variance is positive then $a_0 > 0$ and the polynomial $a(1) > 0$

Properties of ARCH Errors

3. ϵ_t is leptokurtic (under normality assumption)

$$
\begin{aligned}
E[\epsilon_t^4] =& E[\sigma_t^4 E[z_t^4|F_{t-1}]] = E[\sigma_t^4]3 \\
\geq& (E[\sigma_t^2])^2 3 = (E[\epsilon_t^2])^2 3 \quad \text{by Jensen's inequality} \\
\Rightarrow& \frac{E[\epsilon_t^4]}{(E[\epsilon_t^2])^2} > 3
\end{aligned}
$$

- The excess kurtosis is positive and the tail distribution of ϵ_t is heavier than the normal distribution

$$\text{kurt}(\epsilon_t) > 3 = \text{kurt(normal)}$$

- This means that the Gaussian ARCH model produces more outliers than the Gaussian white noise.

Syddansk Universitet

Properties of ARCH Errors

4. σ_t^2 is a serially correlated random variable

$$\sigma_t^2 = a_0 + a(L)\epsilon_t^2,$$
$$E[\sigma_t^2] = \frac{a_0}{a(1)} = \bar{\sigma}^2$$

Using $a_0 = 1 - a(L)\bar{\sigma}^2$:

$$\sigma_t^2 - \bar{\sigma}^2 = a(L)(\epsilon_t^2 - \bar{\sigma}^2)$$

5. ϵ_t^2 has a stationary AR(p) representation.

$$\sigma_t^2 + \epsilon_t^2 = a_0 + a(L)\epsilon_t^2 + \epsilon_t^2$$
$$\Rightarrow \epsilon_t^2 = a_0 + a(L)\epsilon_t^2 + (\epsilon_t^2 - \sigma_t^2)$$

where $(\epsilon_t^2 - \sigma_t^2) = v_t$ is a conditionally heteroskedastic MDS, i.e $E(v_t|F_{t-1}) = 0$

SYDDANSK UNIVERSITET

Properties of ARCH Errors

6. ϵ_t^2 exhibits em volatility mean reversion.

Example: Consider ARCH(1) with $0 < a < 1$

$$\sigma_t^2 = a_0 + a\epsilon_{t-1}^2 = (1-a)\bar{\sigma}^2 + a\epsilon_{t-1}^2$$

$$\Downarrow$$

$$(\epsilon_t^2 - \bar{\sigma}^2) = a(\epsilon_{t-1}^2 - \bar{\sigma}^2) + v_t$$

$$\Downarrow$$

$$E[\epsilon_t^2 | F_{t-1}] - \bar{\sigma}^2 = a^k [E(\epsilon_{t-k}^2) - \bar{\sigma}^2] \rightarrow 0 \text{ as } k \rightarrow \infty$$

Building a volatility model

Building a volatility model for asset return series consists of four steps:

1. Specify the mean equation by fitting an ARMA model for the return series.
   - Note that the residuals should not have any autocorrelation
2. Use residuals to test for ARCH effects
3. Specify a volatility model if ARCH effects are statistically significant, and perform a joint estimation of the mean and volatilty equations
4. Check the fitted model and refine if necessary

Testing for ARCH effects

Let $\epsilon_t = r_t - \mu_t$ be the residuals of the mean equation.

Q: Do the residuals square follow an AR(m) model?

$$\epsilon_t^2 = a_0 + a_1\epsilon_{t-1}^2 + \ldots + a_m\epsilon_{t-m}^2 + \nu_t$$

This is written into a test:

$H_0 : a_1 = \ldots = a_m = 0$ vs. $H_1 :$ at least one coef is nonzero

SYDDANSK UNIVERSITET

Testing for ARCH effects

Two tests for conditional heteroskedasticiy are available:

1. Apply the Ljung-Box statistics S_{LB} to ϵ_T^2.
2. The Lagrange multiplier test.

$$SSR_0 = \sum_{t=m+1}^{T} (\epsilon_t^2 - \bar{\omega})^2 \quad \bar{\omega} = \frac{1}{T} \sum_{t=1} \epsilon_t^2$$

$$SSR_1 = \sum_{t=m+1}^{T} \hat{\nu}_t^2 \quad \hat{\nu}^2 \text{ is the least-squares residual}$$

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - 2m - 1)} \sim \chi_m^2$$

The null hypothesis is rejected if $F > \chi_m^2(\alpha)$

SYDDANSK UNIVERSITET

Testing for ARCH effects

We consider the monthly log stock returns of Intel Corporation from 1973–2008 (Intel.txt)

The series does not have a significant serial correlation so we can directly test for ARCH effects

```
> intel = read.table(file="../data/Intel.txt",header = T)
> #We convert simple returns in log returns
> r_t= log(intel[,2]+1)
> Box.test(r_t, lag=12, type="Ljung")

        Box-Ljung test

data:  r_t
X-squared = 18.2635, df = 12, p-value = 0.1079

> #The test results in no autocorrelation of the returns
```

Testing for ARCH effects

```
> #Remove the mean of the returns and test for ARCH effects
> epsilon_t= r_t - mean(r_t)
> Box.test(epsilon_t^2, lag=12, type="Ljung")

        Box-Ljung test

data:  epsilon_t^2
X-squared = 89.8509, df = 12, p-value = 5.274e-14
```

There is no function to calculate the LM test in R for ARCH effects but you can program it easily.

Weakness of ARCH models

1. The model assume that positive and negative effects have the same effects on volatility
2. It is restrictive in the size of parameters a_i to ensure the fourth moment is positive
3. It does not give indication of the causes of heteroskedasticity
4. It is likely to overpredict the volatility because it responds very slowly to large isolated shocks of the returns

SYDDANSK UNIVERSITET

Specifying an ARCH model

Once we have found that there are ARCH effects, which is the order of the ARCH model?
The model looks like

$$\sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + \ldots + \epsilon_{t-m}^2$$

- For a given sample ϵ_t^2 is an unbiased estimate of σ_t^2
- We expect that ϵ_t^2 is linearly related to $\epsilon_{t-1}^2 \ldots$
- A single ϵ_t^2 is not an efficiente estimate of σ_t^2, but it can be an approximation (proxy)

Specifying an ARCH model

Define $\nu_t = \epsilon_t^2 - \sigma_t^2$:

$$\epsilon_t^2 = a_0 + a_1\epsilon_{t-1}^2 + \ldots + \epsilon_{t-m}^2 + \nu_t$$

which is an AR(m) model for ϵ_t^2.

- So we can estimate the order by plotting the PACF of ϵ_t^2
- The only problem is the ν_t is not an iid series, so the least-squares estimates of the prior model are consistent but inefficient.

Estimation of an ARCH model

Depending on the distribution of ϵ_t, we can obtain several likelihood function for the sample $\{r_t\}_{t=1}$

Exact likelihood:

$$f_\epsilon(\epsilon_t, \ldots, \epsilon_T; \boldsymbol{\theta}) = \prod_{t=m+1}^{T} f_{\epsilon_t|F_{t-1}}(\epsilon_t|\epsilon_{t-1}, \ldots \epsilon_1; \boldsymbol{\theta}) \quad \underbrace{f_{\epsilon_1,\ldots,\epsilon_m}(\epsilon_1, \ldots, \epsilon_m; \boldsymbol{\theta})}_{\text{Might be difficult to fin}}$$

$$\star = \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\sigma_t^2}\right)$$

$$\cdot \frac{(det(\boldsymbol{\Omega_m^{-1}}))^{1/2}}{(2\pi)^{m/2}} \exp\left(-\frac{1}{2}\tilde{\boldsymbol{\epsilon}}'\boldsymbol{\Omega_m^{-1}}\tilde{\boldsymbol{\epsilon}}\right)$$

$\star$ $-z_t$ follows a standarised normal distribution

Syddansk Universitet

Estimation of an ARCH model

Conditional likelihood:

$$
\begin{aligned}
f_\epsilon(\epsilon_t, \ldots, \epsilon_T; \boldsymbol{\theta}) &= \prod_{t=m+1}^{T} f_{\epsilon_t|F_{t-1}}(\epsilon_t|\epsilon_{t-1}, \ldots \epsilon_1; \boldsymbol{\theta}) \\
\star &= \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\sigma_t^2}\right) \\
\star\star &= \prod_{t=m+1}^{T} \frac{\Gamma(\nu+1)/2}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}} \frac{1}{\sigma_t} \left[1 + \frac{\epsilon_t^2}{(\nu-2)\sigma_t^2}\right]^{-(\nu+1)/2}
\end{aligned}
$$

$\star\star$ – z_t follows a starndarised Student-t distribution with ν degrees of freedom.

Estimation of an ARCH model

```
> library(tseries)
> #Fit an ARCH(1), trace=F to supress numerical output of gradient
> intel.arch<-garch(r_t, order=c(0,1), grad="numerical", trace=F)
> coef(intel.arch)

        a0         a1
0.01144064 0.36708117

> confint(intel.arch)

         2.5 %     97.5 %
a0 0.009388698 0.01349258
a1 0.209865558 0.52429678
```

Syddansk Universitet

Estimation of an ARCH model

A more complete R function:

```
> library(fGarch)
> intel.arch2<-garchFit(r_t~garch(1,0), trace=F)
> coef(intel.arch2)

          mu        omega       alpha1
-0.001550561  0.146527491  0.370867049

> intel.arch2.2<-garchFit(r_t~garch(1,0), trace=F,
+                             include.mean=F)
> coef(intel.arch2.2)

    omega    alpha1
0.1464835 0.3713363
```

Model checking

After estimating the model obtain the residuals

$$\hat{\epsilon}_t = \frac{(r_t - \hat{\mu}_t)}{\hat{\sigma}_t}$$

- Ljung-Box test of $\hat{\epsilon}_t^2$ independence
- $\hat{\epsilon}_t$ should have heavier tails than the standard normal (QQ-plot)
- Try the summary(intel.arch2) in your computer $\Rightarrow$ the Ljung-Box test of the squared residuals have a large p-value (independence)
- Try plot(intel.arch2) and choose 13 to see the QQ-plot

SYDDANSK UNIVERSITET

Predict and ARCH model

```
> #1 to 5-step-ahead predictions
> predict(intel.arch2, 5)

  meanForecast meanError standardDeviation
1 -0.001550561 0.5005455         0.5005455
2 -0.001550561 0.4893329         0.4893329
3 -0.001550561 0.4851086         0.4851086
4 -0.001550561 0.4835326         0.4835326
5 -0.001550561 0.4829468         0.4829468
```

t-innovations

For comparison we also fit an ARCH(1) model with z_t a standarised Student-t distribution

```
>  options(width=60)
> intel.t<-garchFit(r_t~garch(1,0), cond.dist="std",
+                     trace=F)
> round(coef(intel.t),4)

    mu   omega alpha1   shape
0.0113 0.1548 0.5491 3.4435
```

With a skew Student-t distribution

```
> intel.st<-garchFit(r_t~garch(1,0), cond.dist="sstd",
+                     trace=F)
> round(coef(intel.st),4)

    mu   omega alpha1    skew   shape
0.0038 0.1536 0.5303 0.9603 3.5016
```

t-innovations

Fit an ARMA(1,0) + GARCH(1,1)

```
> intel.arma<-garchFit(r_t~arma(1,0)+garch(1,1), trace=F)
> coef(intel.arma)
```

```
         mu          ar1        omega       alpha1
-0.006097039  0.051377920  0.011189123  0.157402965
      beta1
 0.799951837
```

SYDDANSK UNIVERSITET