Conditional Heteroscedastic Models

What is stock volatility?

A simple answer: conditional variance of stock returns

Why is volatility important?

Has many important applications

- Option (derivative) pricing
- Risk management, e.g. value at risk (VaR)
- Asset allocation
- Interval forecasts

Black-Scholes European call option:

$$c_t = P_t\Phi(x) - Kr^{-\ell}\Phi(x - \sigma_t\sqrt{\ell})$$

$$x = \frac{\ln(P_t/Kr^{-\ell})}{\sigma_t\sqrt{\ell}} + \frac{1}{2}\sigma_t\sqrt{\ell}$$

- P_t: current price of the stock,

- r: risk-free interest rate,
- K: strike price
- ℓ: time to expiration
- σ_t: conditional standard deviation of the log return of the specified stock,
- $\Phi(x)$: normal CDF

A key characteristic: Not directly observable!!

How to calculate volatility?

1. Use high-frequency data: French, Schwert & Stambaugh (1987); see Section 3.12.
2. Implied volatility of options data
3. Econometric modeling

We focus on the latter

Basic idea

Shocks of asset returns are NOT serially correlated, but dependent.

See ACF of squared and absolute returns of some stocks

Basic structure

$$r_t = \mu_t + a_t, \mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i},$$

Volatility models are concerned with time-evolution of

$$\sigma_t^2 = \mathrm{Var}(r_t | F_{t-1}) = \mathrm{Var}(a_t | F_{t-1}).$$

the conditional variance of a return.

Two general categories

- "Fixed function"
- Stochastic function

of the available information.

Univariate volatility models

1. Autoregressive conditional heteroscedastic (ARCH) model of Engle (1982),
2. Generalized ARCH (GARCH) model of Bollerslev (1986),
3. GARCH-M models
4. IGARCH models
5. Exponential GARCH (EGARCH) model of Nelson (1991),
6. Conditional heteroscedastic ARMA (CHARMA) model of Tsay (1987),
7. Random coefficient autoregressive (RCA) model of Nicholls and Quinn (1982)
8. Stochastic volatility (SV) models of Melino and Turnbull (1990), Harvey, Ruiz and Shephard (1994), and Jacquier, Polson and Rossi (1994).
9. Threshold GARCH models

ARCH model

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2,$$

where $\{\epsilon_t\}$ is a sequence of iid r.v. with mean 0 and variance 1, $\alpha_0 > 0$ and $\alpha_i \geq 0$ for $i > 0$.

Distribution of ϵ_t: Standard normal, standardized Student-t or generalized error dist (GED).

Properties of ARCH models

Consider an ARCH(1) model

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

where $\alpha_0 > 0$ and $\alpha_1 \geq 0$.

1. $\mathrm{E}(a_t) = 0$

2. $\mathrm{Var}(a_t) = \alpha_0/(1-\alpha_1)$ if $0 < \alpha_1 < 1$

3. Under normality,

$$m_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)},$$

   provided $0 < \alpha_1^2 < 1/3$.

The 3rd property implies heavy tails.

Advantages

- Simplicity
- Generates volatility clustering
- Heavy tails (high kurtosis)

Weaknesses

- Symmetric btw $+$ & $-$ returns
- Restrictive
- Provides no explanation
- Not sufficiently adaptive in prediction

Building an ARCH Model

1. Modeling the mean effect and testing

   Use Q-statistics of squared residuals; McLeod and Li (1983) & Engle (1982)

2. Order determination

   Use PACF of the squared residuals

3. Estimation: Conditional MLE

4. Model checking: Q-stat of standardized residuals and squared standardized residuals. Skewness & Kurtosis of standardized residuals.

5. Software: RATS, E-views, S-plus, SCA, OX, etc.

Estimation:

Under normality, the likelihood function of an ARCH(m) model is $f(a_1, \cdots, a_T|\boldsymbol{\alpha})$

$$= f(a_T|F_{T-1})f(a_{T-1}|F_{T-2}) \cdots f(a_{m+1}|F_m)f(a_1, \cdots, a_m|\boldsymbol{\alpha})$$

$$= \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp[-\frac{a_t^2}{2\sigma_t^2}] \times f(a_1, \cdots, a_m|\boldsymbol{\alpha}),$$

where $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \cdots, \alpha_m)'$ and $f(a_1, \cdots, a_m|\boldsymbol{\alpha})$ is the joint pdf of $a_1, \cdots, a_m$.

A conditional log likelihood fun. is

$$\ell(a_{m+1}, \cdots, a_T|\boldsymbol{\alpha}, a_1, \cdots, a_m)$$

$$-\sum_{t=m+1}^{T}\left[\frac{1}{2}\ln(\sigma_t^2)+\frac{1}{2}\frac{a_t^2}{\sigma_t^2}\right],$$

where $\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2$ can be evaluated recursively.

Under t-dist, the cond. log likelihood fun. is

$$\ell(a_{m+1},\cdots,a_T|\boldsymbol{\alpha},A_m)$$

$$=-\sum_{t=m+1}^{T}\left[\frac{v+1}{2}\ln\left(1+\frac{a_t^2}{(v-2)\sigma_t^2}\right)+\frac{1}{2}\ln(\sigma_t^2)\right],$$

if v is given.

Otherwise, it becomes $\ell(a_{m+1},\cdots,a_T|\boldsymbol{\alpha}.v,A_m)$

$$=(T-m)[\ln(\Gamma((v+1)/2))-\ln(\Gamma(v/2))-0.5\ln((v-2)\pi)]$$

$$+\ell(a_{m+1},\cdots,a_T|\boldsymbol{\alpha},A_m).$$

Example: Monthly log returns of Intel stock

Under normality,

$$r_t = 0.021 + a_t, \quad \sigma_t^2 = 0.01 + 0.444a_{t-1}^2.$$

Model checking:

Standardized shocks $\{\tilde{a}_t\}$

$Q(10) = 12.53(0.25)$

For $\{\tilde{a}_t^2\}$

$Q(10) = 17.23(0.07)$.

Implications

- Expected monthly log return is about 2.1%, which is remarkable.
- $\hat{\alpha}_1^2 = 0.444^2 < 1/3$ so that 4th moment exists.

t-innovation: t_5-dist

$$r_t = 0.022 + a_t, \quad \sigma_t^2 = 0.01 + 0.303 a_{t-1}^2.$$

Model checking:

Standardized shocks: $Q(10) = 13.66(0.19)$

Squared standardized shocks: $Q(10) = 23.83(0.008)$

Refinement:

$$r_t = 0.023 + a_t. \quad \sigma_t^2 = 0.01 + 0.226 a_{t-1}^2 + 0.108 a_{t-2}^2.$$

The model passes Q-stat.

Comparison:

- Using a heavy-tailed dist for ϵ_t reduces the ARCH effect
- The difference between the three models is small for this particular instance.

GARCH Model

$$a_t = \sigma_t \epsilon_t,$$

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

where $\{\epsilon_t\}$ is defined as before, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\Sigma_{i=1}^{\max(m,s)}(\alpha_i + \beta_i) < 1$.

Re-parameterization:

Let $\eta_t = a_t^2 - \sigma_t^2$. $\{\eta_t\}$ un-correlated series.

The GARCH model becomes

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

This is an ARMA form for the squared series a_t^2.

Use it to understand properties of GARCH models, e.g. moment equations, forecasting, etc.

Focus on a GARCH(1,1) model

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

- Weak stationarity: $0 \le \alpha_1, \beta_1 \le 1, (\alpha_1 + \beta_1) < 1$.

- Volatility clusters

- Heavy tails: if $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, then
$$\frac{E(a_t^4)}{[E(a_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3.$$

- For 1-step ahead forecast,
$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2.$$

  For multi-step ahead forecasts, use $a_t^2 = \sigma_t^2 \epsilon_t^2$ and rewrite the model as
$$\sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2 + \alpha_1 \sigma_t^2 (\epsilon_t^2 - 1).$$

  2-step ahead volatility forecast
$$\sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(1).$$

  In general, we have
$$\sigma_h^2(\ell) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(\ell - 1), \quad \ell > 1.$$

This result is exactly the same as that of an ARMA(1,1) model with AR polynomial $1 - (\alpha_1 + \beta_1)B$.

Example: Monthly excess returns of S&P 500 index starting from 1926 for 792 observations.

The fitted of a Gaussian AR(3) model

$$r_t = .088r_{t-1} - .023r_{t-2} - .123r_{t-3} + .007 + a_t,$$

$\hat{\sigma}_a^2 = 0.00333.$

For the GARCH effects, use a GARCH(1,1) model.

A joint estimation:

$$\begin{aligned} r_t &= 0.021r_{t-1} - 0.034r_{t-2} - 0.013r_{t-3} + 0.0085 + a_t \\ \sigma_t^2 &= .0001 + .848\sigma_{t-1}^2 + 0.122a_{t-1}^2. \end{aligned}$$

Implied unconditional variance of a_t is

$$\frac{0.000099}{1 - 0.8476 - 0.1219} = 0.00325$$

close to expected value.

A simplified model:

$$r_t = 0.0065 + a_t, \sigma_t^2 = .0001 + .822\sigma_{t-1}^2 + .135a_{t-1}^2.$$

Model checking:

For $\tilde{a}_t$: $Q(10) = 10.32(0.41)$ and $Q(20) = 22.66(0.31)$.

For $\tilde{a}_t^2$: $Q(10) = 8.83(0.55)$ and $Q(20) = 15.82(0.73)$.

Forecast: 1-step ahead forecast:

$$\sigma_h^2(1) = 0.00014 + 0.822\sigma_h^2 + 0.1352a_h^2$$

Horizon	1	2	3	4	5	∞
Return	.0065	.0065	.0065	.0065	.0065	.0065
Volatility	.0031	.0031	.0031	.0031	.0031	.0032

t-innovation: t_5 dist

$$r_t = 0.0085 + a_t, \sigma_t^2 = .0002 + .127a_{t-1}^2 + .822\sigma_{t-1}^2.$$

Estimation of degrees of freedom:

$$r_t = 0.0083 + a_t,$$

$$\sigma_t^2 = .0002 + .123a_{t-1}^2 + .819\sigma_{t-1}^2,$$

where the estimated degrees of freedom is 6.51.

Forecasting evaluation

Not easy to do; see Andersen and Bollerslev (1998).

IGARCH model

An IGARCH(1,1) model:

$$a_t = \sigma_t \epsilon_t, \sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1) a_{t-1}^2.$$

For the monthly excess returns of the S&P 500 index, we have

$$r_t = .007 + a_t, \sigma_t^2 = .0001 + .806\sigma_{t-1}^2 + .194a_{t-1}^2$$

For an IGARCH(1,1) model,

$$\sigma_h^2(\ell) = \sigma_h^2(1) + (\ell - 1)\alpha_0, \quad \ell \geq 1,$$

where h is the forecast origin.

Effect of $\sigma_h^2(1)$ on future volatilities is persistent, and the volatility forecasts form a straight line with slope α_0. See Nelson (1990) for more info.

Special case: $\alpha_0 = 0$.

used in RiskMetrics to VaR calculation.

The GARCH-M model

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

where c is referred to as risk premium, which is expected to be positive.

Example: A GARCH(1,1)-M model for the monthly excess returns of S&P 500 index from January 1926 to December 1991. The fitted model is

$r_t = 0.0028 + 1.99\sigma_t^2 + a_t, \sigma_t^2 = .00016 + .133 a_{t-1}^2 + .814 \sigma_{t-1}^2.$

Std err of risk premium is 0.753.

The EGARCH model

Asymmetry in responses to + & − returns:

$$g(\epsilon_t) = \theta\epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)],$$

with $E[g(\epsilon_t)] = 0$.

To see asymmetry of $g(\epsilon_t)$, rewrite it as

$$g(\epsilon_t) = \begin{cases} (\theta + \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \text{if } \epsilon_t \geq 0, \\ (\theta - \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \text{if } \epsilon_t < 0. \end{cases}$$

An EGARCH(m, s) model:

$$a_t = \sigma_t \epsilon_t, \ \ln(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 B + \cdots + \beta_s B^s}{1 - \alpha_1 B - \cdots - \alpha_m B^m} g(\epsilon_{t-1}).$$

Some features of EGARCH models:

- uses log trans. to relax the positiveness constraint
- asymmetric responses

Consider an EGARCH(1,0) model

$$a_t = \sigma_t \epsilon_t, \qquad (1 - \alpha B) \ln(\sigma_t^2) = (1 - \alpha)\alpha_0 + g(\epsilon_{t-1}),$$

Under normality, $E(|\epsilon_t|) = \sqrt{2/\pi}$ and the model becomes

$$(1 - \alpha B) \ln(\sigma_t^2) = \begin{cases} \alpha_* + (\theta + \gamma)\epsilon_{t-1} & \text{if } \epsilon_{t-1} \geq 0, \\ \alpha_* + (\theta - \gamma)\epsilon_{t-1} & \text{if } \epsilon_{t-1} < 0 \end{cases}$$

where $\alpha_* = (1 - \alpha)\alpha_0 - \sqrt{\frac{2}{\pi}}\gamma$.

A nonlinear fun. similar to that of the TAR model of Tong (1978, 1990).

Specifically, we have

$$\sigma_t^2 = \sigma_{t-1}^{2\alpha} \exp(\alpha_*) \begin{cases} \exp[(\theta + \gamma)\frac{a_{t-1}}{\sqrt{\sigma_{t-1}^2}}] & \text{if } a_{t-1} \geq 0, \\ \exp[(\theta - \gamma)\frac{a_{t-1}}{\sqrt{\sigma_{t-1}^2}}] & \text{if } a_{t-1} < 0. \end{cases}$$

The coefs $(\theta + \gamma)$ & $(\theta - \gamma)$ show the asymmetry in response to positive and negative a_{t-1}. The model is, therefore, nonlinear if $\gamma \neq 0$.

See Nelson (1991) for an exmaple.

Another example: Monthly log returns of IBM stock from January 1926 to December 1997 for 864 observations.

An AR(1)-EGARCH(1,0):

$$\begin{aligned} r_t &= 0.0105 + 0.092 r_{t-1} + a_t, \quad a_t = \sigma_t \epsilon_t \\ \ln(\sigma_t^2) &= -5.496 + \frac{g(\epsilon_{t-1})}{1 - .856B}, \\ g(\epsilon_{t-1}) &= -.0795\epsilon_{t-1} + .2647[|\epsilon_{t-1}| - \sqrt{2/\pi}], \end{aligned}$$

Model checking:

For $\tilde{a}_t$: $Q(10) = 6.31(0.71)$ and $Q(20) = 21.4(0.32)$

For $\tilde{a}_t^2$: $Q(10) = 4.13(0.90)$ and $Q(20) = 15.93(0.66)$

Discussion:

Using $\sqrt{2/\pi} \approx 0.7979$, we obtain

$$\ln(\sigma_t^2) = -1.0 + 0.856 \ln(\sigma_{t-1}^2) + \begin{cases} 0.1852\epsilon_{t-1} & \text{if } \epsilon_{t-1} \geq 0 \\ -0.3442\epsilon_{t-1} & \text{if } \epsilon_{t-1} < 0. \end{cases}$$

Taking anti-log transformation, we have

$$\sigma_t^2 = \sigma_{t-1}^{2\times 0.856} e^{-1.001} \times \begin{cases} e^{0.1852\epsilon_{t-1}} & \text{if } \epsilon_{t-1} \geq 0 \\ e^{-0.3442\epsilon_{t-1}} & \text{if } \epsilon_{t-1} < 0. \end{cases}$$

For a standardized shock with magnitude 2, (i.e. two standard deviations), we have

$$\frac{\sigma_t^2(\epsilon_{t-1} = -2)}{\sigma_t^2(\epsilon_{t-1} = 2)} = \frac{\exp[-0.3442 \times (-2)]}{\exp(0.1852 \times 2)} = e^{0.318} = 1.374.$$

Therefore, the impact of a negative shock of size two-standard deviations is about 37.4% higher than that of a positive shock of the same size.

Forecasting: some recursive formula available

The CHARMA model

Make use of "interaction" btw past shocks

A CHARMA model is defined as

$$r_t = \mu_t + a_t, a_t = \delta_{1t}a_{t-1} + \delta_{2t}a_{t-2} + \cdots + \delta_{mt}a_{t-m} + \eta_t,$$

where $\{\eta_t\}$ is iid $N(0, \sigma_\eta^2)$, $\{\boldsymbol{\delta}_t\} = \{(\delta_{1t}, \cdots, \delta_{mt})'\}$ is a sequence of iid random vectors $D(\mathbf{0}, \boldsymbol{\Omega})$, $\{\boldsymbol{\delta}_t\} \perp \{\eta_t\}$.

The model can be written as

$$a_t = \boldsymbol{a}'_{t-1}\boldsymbol{\delta}_t + \eta_t,$$

with conditional variance

$$\sigma_t^2 = \sigma_\eta^2 + \boldsymbol{a}'_{t-1}\text{Cov}(\boldsymbol{\delta}_t)\boldsymbol{a}_{t-1}$$

$$= \sigma_\eta^2 + (a_{t-1}, \cdots, a_{t-m})\boldsymbol{\Omega}(a_{t-1}, \cdots, a_{t-m})'.$$

Example: Monthly excess returns of S&P 500 index (26-91).

A fitted model is

$$r_t = 0.0068 + a_t,$$

$$\sigma_t^2 = .00136 + (a_{t-1}, a_{t-2}, a_{t-3})\widehat{\boldsymbol{\Omega}}(a_{t-1}, a_{t-2}, a_{t-3})'$$

where, std errors in parentheses,

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 0.121(.036) & -0.062(.028) & 0 \\ -0.062(.028) & 0.191(.025) & 0 \\ 0 & 0 & 0.299(0.042) \end{bmatrix}.$$

Effects of explanatory variables

Can be used in the same manner, i.e. with random coefs.

RCA model

A time series r_t is a RCA(p) model if

$$r_t = \phi_0 + \sum_{i=1}^{p} (\phi_i + \delta_{it}) r_{t-i} + a_t.$$

For the model, we have

$$\begin{aligned} \mu_t &= E(a_t|F_{t-1}) = \sum_{i=1}^{p} \phi_i a_{t-i}, \\ \sigma_t^2 &= \sigma_a^2 + (r_{t-1}, \cdots, r_{t-p}) \boldsymbol{\Omega}_\delta (r_{t-1}, \cdots, r_{t-p})'. \end{aligned}$$

Stochastic volatility model

A (simple) SV model is

$$a_t = \sigma_t \epsilon_t, (1 - \alpha_1 B - \cdots - \alpha_m B^m) \ln(\sigma_t^2) = \alpha_0 + v_t$$

where ϵ_t's are iid $N(0,1)$, v_t's are iid $N(0, \sigma_v^2)$, $\{\epsilon_t\}$ and $\{v_t\}$ are independent.

Long-memory SV model

A simple LMSV is

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t = \sigma \exp(u_t/2), \qquad (1 - B)^d u_t = \eta_t$$

where $\sigma > 0$, ϵ_t's are iid $N(0,1)$, η_t's are iid $N(0, \sigma_\eta^2)$ and independent of ϵ_t, and $0 < d < 0.5$.

The model says

$$\begin{aligned}
\ln(a_t^2) &= \ln(\sigma^2) + u_t + \ln(\epsilon_t^2) \\
&= [\ln(\sigma^2) + E(\ln \epsilon_t^2)] + u_t + [\ln(\epsilon_t^2) - E(\ln \epsilon_t^2)] \\
&\equiv \mu + u_t + e_t.
\end{aligned}$$

Thus, the $\ln(a_t^2)$ series is a Gaussian long-memory signal plus a non-Gaussian white noise; see Breidt, Crato and de Lima (1998).

Application

see Examples 3.5 & 3.6

Computing: to be discussed in class.