Econ 623 Econometrics II

Topic 7: ARCH Models

1 Forms and Properties of ARCH-type Models

1.1 ARCH(q) model (Engle, 1982)

It models the conditional variance as a deterministic function of past returns. Suppose y_t is the return of a financial asset.

$$\begin{cases} y_t = \mu + \sigma_t \varepsilon_t \\ \sigma_t^2 = \alpha + \beta_1 y_{t-1}^2 + \cdots + \beta_q y_{t-q}^2 \end{cases} .$$

A special ARCH(q) model that we will consider here is,

$$\begin{cases} y_t = \sigma_t \varepsilon_t, \ \varepsilon_t \overset{iid}{\sim} N(0,1) \\ \sigma_t^2 = \alpha + \beta_1 y_{t-1}^2 + \cdots + \beta_q y_{t-q}^2 \end{cases} .$$

Properties:

- $E(y_t|I_{t-1}) = 0$
- $E(y_t^2|I_{t-1}) = \sigma_t^2 = \alpha + \beta_1 y_{t-1}^2 + \cdots + \beta_q y_{t-q}^2$
- $y_t^2 \sim AR(q)$
- Since σ_t^2 is time variant, the model is called the Autoregressive Conditional Heteroskedasticity (ARCH) model.
- According to the model, the conditional variance is completely determined by the lagged squared returns.

- The kurtosis of ARCH is bigger than 3. In particular, when $q = 1$, the kurtosis of y_t is given by $3(1-\beta_1^2)/(1-3\beta_1^2) > 1$ as long as $\beta_1 \neq 0$. So ARCH models have fatter tails than the normal distribution.

- If $\beta_1, \cdots, \beta_q, \alpha > 0$, and $\beta_1 + \cdots + \beta_q < 1$, the ARCH(q) model is covariance stationary. If so, $E(y_t^2) = \alpha/(1-\beta_1-\cdots-\beta_q)$

- $E(y_t\ y_s) = 0$ for any $t \neq s$.

- $E(y_t^2\ y_s^2) \neq 0$ for some $t \neq s$.

- In most empirical studies, we need many lags in the variance equation (ie q must be large).

1.2 GARCH(p,q) (Bollerslev, 1986)

It models the conditional variance as a deterministic function of past returns and conditional variance

$$\begin{cases} y_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} \beta_j y_{t-j}^2 + \sum_{j=1}^{p} \alpha_j \sigma_{t-j}^2 \end{cases},$$

Properties:

- $E(y_t|I_{t-1}) = 0$, $E(y_{t-1}^2|I_t) = \sigma_t^2$
- $y_t^2 = \alpha_0 + (\alpha_1 + \beta_1)y_{t-1}^2 + \cdots + (\alpha_m + \beta_m)y_{t-m}^2 + w_t - \alpha_1 w_{t-1} - \cdots - \alpha_p w_{t-p}$, where $m = \max\{p, q\}$, $w_t = y_t^2 - \sigma_t^2$
- $y_t^2 \sim ARMA(\max\{p, q\}, p)$
- According to the model, the conditional variance is the weighted average of the lagged squared returns (ARCH term) and the lagged conditional variance (GARCH term).

- If $\alpha_1, \cdots, \alpha_p > 0$, $\beta_1, \cdots, \beta_q > 0$, $\alpha > 0$ and $(\alpha_1 + \beta_1) + \cdots + (\alpha_m + \beta_m) < 1$, the GARCH(p,q) model is covariance stationary. $E(y_t^2) = \alpha/(1 - \alpha_1 - \cdots - \alpha_m - \beta_1 - \cdots - \beta_m)$

- In the GARCH (1,1) model, $E(y_t^2|I_t) = \alpha_0 + (\alpha_1 + \beta_1)y_{t-1}^2 + w_t - \alpha_1 w_{t-1}$, so $(\alpha_1 + \beta_1)$ governs the persistence of volatilities.

- To describe the real financial time series, a parsimonious GARCH model is often doing the job as well as an ARCH model with a long lag length. In the empirical study, the GARCH(1,1) specification has been found to be adequate in most application. In financial data, often $(\alpha_1 + \beta_1)$ is very found to be close to 1 in the GARCH(1,1) model. Therefore, the volatility process is persistent, suggesting volatility is predictable.

1.3 IGARCH (Engle and Bollerslev, 1986)

Volatility is as persistent as a unit root

$$\begin{cases} y_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} \beta_j y_{t-j}^2 + \sum_{j=1}^{p} \alpha_j \sigma_{t-j}^2 \\ \sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j = 1 \end{cases},$$

Properties:

- A shock to the conditional variance remains forever.
- Unconditional expected value of y_t^2 is infinite. Hence, the IGARCH model is not covariance stationary.
- It is parsimonious compared with GARCH(p,q).

1.4 GJR-GARCH (Glosten, Jaganathan and Runkle, 1993)

Conditional variance responds to good news in different ways to bad news.

$$\begin{cases} y_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \beta_i y_{t-i}^2 + \gamma y_{t-1}^2 d_{t-1} + \sum_{i=1}^{p} \alpha_i \sigma_{t-i}^2 \\ d_{t-1} = 1 \text{ if } y_{t-1} < 0,\ d_{t-1} = 0 \text{ if } y_{t-1} \geq 0 \end{cases},$$

Properties:

- d_t is a decision rule variable. The threshold is $y_{t-1} = 0$.
- The model can lead to the leverage effect. When $\gamma > 0$, bad news ($y_{t-1} < 0$) has larger impact on volatility than good news ($y_{t-1} \geq 0$).
- It is also known as the threshold ARCH (TARCH) model.

1.5 EGARCH (Nelson, 1991)

Model the log-conditional volatility to ensure volatility is positive.

$$\begin{cases} y_t = \sigma_t \varepsilon_t \\ \ln(\sigma_t^2) = \alpha_0 + \sum_{j=1}^{p} \alpha_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^{q} (\beta_i \left| \frac{y_{t-i}}{\sigma_{t-i}} \right| + \gamma_i \frac{y_{t-i}}{\sigma_{t-i}}) \end{cases},$$

Properties:

- Conditional variance depends on both the size and the sign of lagged residuals.
- γ_i captures the asymmetric effect. When $\gamma_j < 0$, bad news ($y_{t-i} < 0$) has larger impact on volatility than good news ($y_{t-i} \geq 0$).
- In the volatility equation, $\ln(\sigma_t^2)$ but σ_t^2 is modeled such that no restriction on the parameters is needed.

1.6 Absolute Value ARCH (Schwert, 1989)

$$\sigma_t = \alpha_0 + \sum_{i=1,q} \beta_i |y_{t-i}| + \sum_{i=1}^{p} \alpha_i \sigma_{t-i}$$

1.7 NARCH (non-linear ARCH) (Higgin and Bera, 1992)

$$\sigma_t^\gamma = \alpha_0 + \sum_{i=1}^{q} \beta_i |y_{t-i}|^\gamma + \sum_{i=1,p} \alpha_i {\sigma_{t-i}}^\gamma$$

- Nest the standard GARCH model. If $\gamma = 2$, it is a GARCH; if $\gamma = 1$, it is an absolute value ARCH.

1.8 Quadratic ARCH (Sentana, 1991)

$$\sigma_t^2 = \alpha_0 + \beta y_{t-1}^2 + \delta y_{t-1} + \alpha h_{t-1}$$

- $\delta < 0$ means positive returns increase volatility less than negative returns.

1.9 Switching ARCH (Cai, 1994; Hamilton and Susmel, 1994)

$$\begin{cases} y_t = \sigma_t \varepsilon_t \\ \sigma_t^2 = \gamma(S_t) + \sum_{i=1,q} \beta_i y_{t-i}^2 \\ \gamma(S_t) = \gamma_0 + \gamma_1 S_t \\ P(S_t = i | S_t = j) = p_{ij} \end{cases},$$

- There are two ARCH models and the economy switches from one to another following a Markov chain.

1.10 ARCH in Mean (Engle, Lilien and Robin(1987))

Allow inter-temporal risk-return trade-offs

$$\begin{cases} y_t = \delta\sigma_t^2 + \sigma_t\varepsilon_t \\ \sigma_t^2 = \alpha + \beta_1 y_{t-1}^2 + \cdots + \beta_q y_{t-q}^2 \end{cases},$$

- This model is the extension of the Capital Asset Pricing Model (CAPM).

1.11 Non-Normal GARCH

- t-GARCH (Bollerslev, 1987): $\varepsilon_t \overset{iid}{\sim} t_{(\eta)}$ with η being the degrees of freedom. Let $f(z)$ be the density of ε_t. Then

$$f(z) = \frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\sqrt{\pi(\eta-2)}\Gamma(\eta/2)}\left(1+\frac{z^2}{\eta-2}\right)^{-(\eta+1)/2}.$$

  where $\Gamma(\cdot)$ is the gamma function, and η is the degrees of freedom. The mean of this distribution is 0 and the variance is 1. In the t distribution, η is a tail-thickness parameter. When $\eta = \infty$, it becomes a standard normal distribution. When η is finite, it has fatter tails than the normal distribution. The smaller the η, the fatter the tails.

- GED-GARCH (Nelson, 1991): $\varepsilon_t \overset{iid}{\sim}$ Generalized Error Distribution (GED). Let $f(z)$ be the density of ε_t. Then

  $$f(z) = \frac{\nu \exp[-0.5|z/\alpha|^{\nu}]}{\alpha 2^{(1+1/\nu)\Gamma(1/\nu)}},$$

  where

  $$\alpha = \sqrt{2^{-2/\nu}\Gamma(1/\nu)/\Gamma(3/n)}.$$

  In GED, ν is a tail-thickness parameter. When $\nu = 2$, it becomes a standard normal distribution. When $\nu < 2$, it has fatter tails than the normal distribution.

- Skewed t-GARCH (Hansen, 1991): $\varepsilon_t \overset{iid}{\sim}$ skewed $t_{(\eta,\lambda)}$. Let $f(z)$ be the density of ε_t. Then

$$f(z) = \begin{cases} bc\left(1 + \left(\frac{1}{\eta-2}\left(\frac{bz+a}{1-\lambda}\right)^2\right)\right)^{-(\eta+1)/2} & \text{if } \ z < -a/b, \\ bc\left(1 + \left(\frac{1}{\eta-2}\left(\frac{bz+a}{1+\lambda}\right)^2\right)\right)^{-(\eta+1)/2} & \text{if } \ z \geq -a/b, \end{cases},$$

where $\eta > 2$, $-1 < \lambda < 1$. The constant a, b, c are given by

$$a = 4\lambda c\left(\frac{\eta-2}{\eta-1}\right),$$

$$b = \sqrt{1+3\lambda^2 - a^2},$$

$$c = \frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\sqrt{\pi(\eta-2)}\Gamma\left(\frac{\eta}{2}\right)}$$

In the skewed t, λ captures asymmetry. If $\lambda = 0$, it becomes a $t_{(\eta)}$ distribution. $\lambda < 0$ gives a negatively skewed distribution (ie the left tail is longer than the right tail).

2 Estimation of ARCH-type Models

2.1 MLE for Gaussian GARCH

$y_t = \sigma_t \varepsilon_t$, $\varepsilon_t \overset{iid}{\sim} N(0,1)$

$\Longrightarrow \frac{y_t}{\sigma_t} = \varepsilon_t \overset{iid}{\sim} N(0,1)$

$\Longrightarrow \ln f(y_1, ..., y_T; \theta) = \sum_{t=1}^{T} [\ln \phi(\varepsilon_t) - 0.5 \ln \sigma_t^2] = \sum_{t=1}^{T} [\ln \phi(\frac{y_t}{\sigma_t}) - 0.5 \ln \sigma_t^2]$

- σ^2 is obtained from the recursive volatility equation in the model. Bollerslev (1986) suggested making the initial values of y_t^2 and σ_t^2 at the unconditional mean of y_t^2. That is, $y_0^2, y_{-1}^2, \cdots, \sigma_0^2, \sigma_{-1}^2, \cdots$ are all assumed to be $\frac{1}{T} \sum_{t=1}^{T} y_T^2$.

- The second term is the Jacobian in the transformation.

- The resulting estimator is consistent, asymptotically normal and asymptotically efficient.

2.2 MLE for non-Gaussian ARCH

$y_t = \sigma_t \varepsilon_t$, $\varepsilon_t \overset{iid}{\sim} f(z)$

$\Longrightarrow \frac{y_t}{\sigma_t} = \varepsilon_t \overset{iid}{\sim} f(z)$

$\Longrightarrow \ln f(y_1, ..., y_T; \theta) = \sum_{t=1}^{T}[\ln f(\varepsilon_t) - 0.5 \ln \sigma_t^2] = \sum_{t=1}^{T}[\ln f(\frac{y_t}{\sigma_t}) - 0.5 \ln \sigma_t^2]$

- As before, σ^2 is obtained from the recursive volatility equation in the model and we can assume $y_0^2, y_{-1}^2, \cdots, \sigma_0^2, \sigma_{-1}^2, \cdots$ to be $\frac{1}{T}\sum_{t=1}^{T} y_T^2$.
- The second term is the Jacobian in the transformation.
- The resulting estimator is consistent, asymptotically normal and asymptotically efficient.

2.3 Quasi-ML for Non-Gaussain GARCH Models

- Even in the case where $\varepsilon_t \overset{iid}{\sim} f(z)$ where f is different from the Gaussian distribution, calculate the likelihood function by assuming that $\varepsilon_t \overset{iid}{\sim} N(0,1)$. The resulting estimator is consistent, asymptotically normal but asymptotically inefficient. The asymptotic variance of QMLE is larger than that of MLE.

3 Testing in ARCH

3.1 Testing for ARCH effects: LM Test

- Calculate R^2 from the regression of y_t^2 on a constant and $y_{t-1}^2, ..., y_{t-q}^2$. $T \cdot R^2$ is the statistic of the LM test. If there exists one q such that the LM statistic is too big (compared with the χ^2 distribution with q degrees of freedom), we have to reject the hypothesis of no ARCH effect in y_t. See Engle (1984).

3.2 Choosing the order parameter p, q in ARCH models: LR test

$$\begin{cases} H_0 : GARCH(p,q) \\ H_1 : GARCH(p+i, q+j) \end{cases} .$$

- Calculate the likelihood values for both models. Calculate the likelihood ratio (LR) statistic $-2(\ln L(H_0) - \ln L(H_1))$. If LR statistic is too big (compared with χ^2 distribution with $(i+j)$ degrees of freedom), we have to reject the $GARCH(p,q)$ model.

4 Volatility Forecasting with ARCH

$E(\sigma^2_{t+h}|I_t) = E(y^2_{t+h}|I_t)$, where $I_t = y_1, \cdots, y_T$.

4.1 Forecasting Volatility with ARCH Models (no GARCH term)

$\sigma^2_{T+1} = \alpha_0 + \beta_1 y^2_T + \cdots + \beta_q y^2_{T+1-q}$

$\Longrightarrow E(\sigma^2_{T+1}|I_T) = \alpha_0 + \beta_1 y^2_T + \cdots + \beta_q y^2_{T+1-q}$

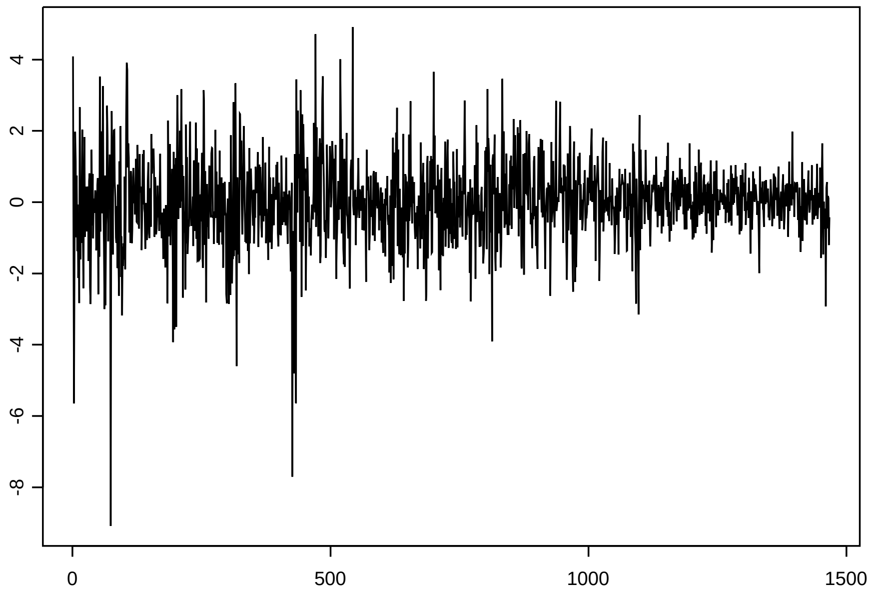
$\Longrightarrow E(\sigma^2_{T+2}|I_T) = \alpha_0 + \beta_1 \widehat{y}^2_{T+1} + \beta_2 y^2_T + \cdots + \beta_q y^2_{T+2-q}$

$\Longrightarrow E(\sigma^2_{T+2}|I_T) = \alpha_0 + \beta_1 E(\sigma^2_{T+1}|I_T) + \beta_2 y^2_T + \cdots + \beta_q y^2_{T+2-q}$

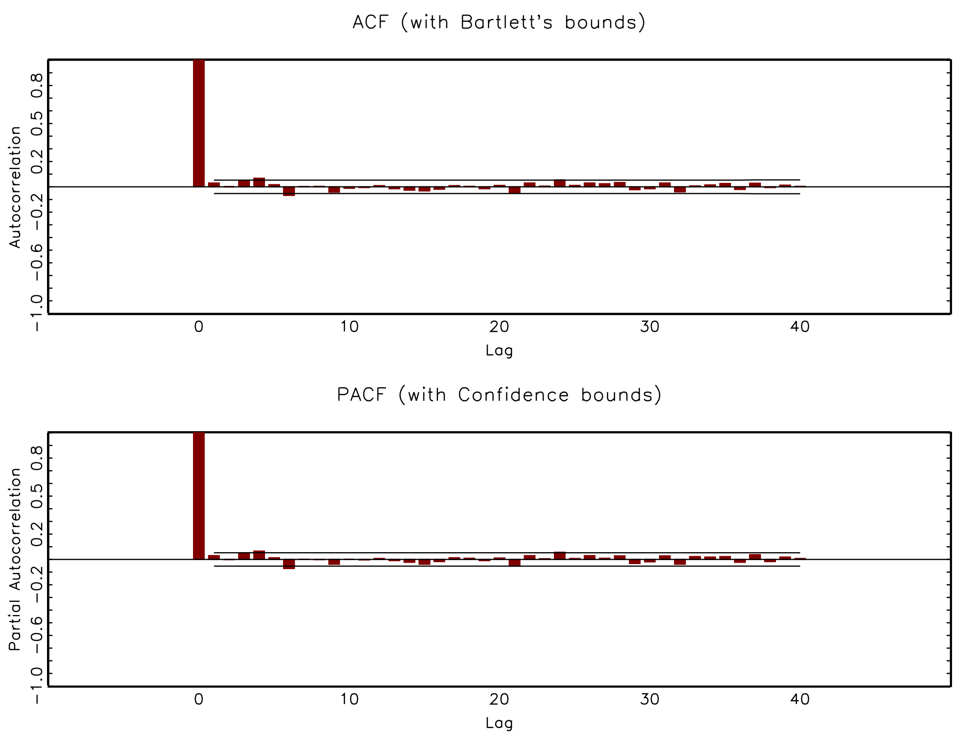
4.2 Forecasting Volatility with GARCH Models

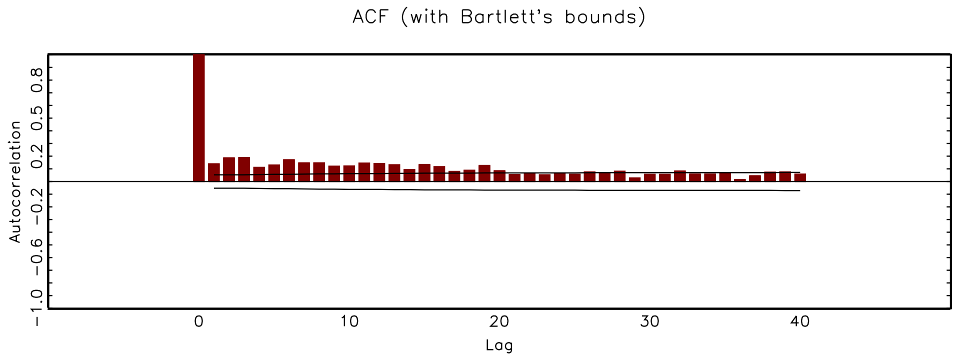
Following Baillie and Bollerslev (1992), the asymptotically optimal h-day ahead forecast of the volatility can be calculated by iterating on

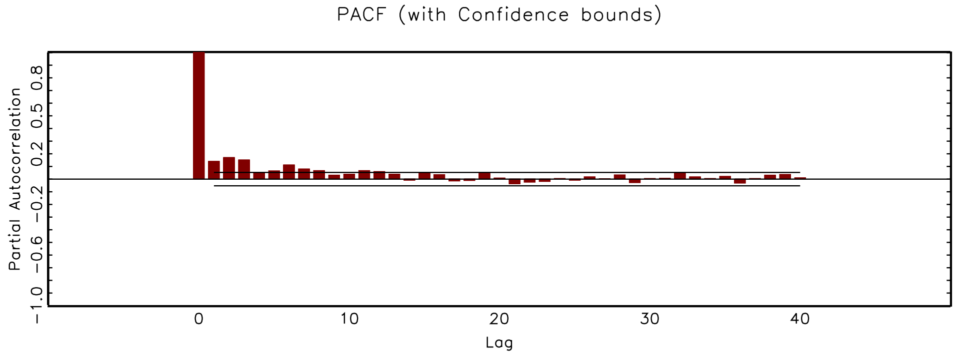
$$\begin{aligned}
\hat{\sigma}^2_{T+h} &= \lambda + \sum_{i=1}^{m} (\alpha_i + \beta_i)\hat{\sigma}^2_{T+h-i} - \beta_h \widehat{w}_T - ... - \beta_m \widehat{w}_{T+h-m}, \text{ for } h = 1, ...p \\
&= \lambda + \sum_{i=1}^{m} (\alpha_i + \beta_i)\hat{\sigma}^2_{T+h-i}, \text{ for } h = p+1, ..., \\
\hat{\sigma}^2_\tau &= y^2_\tau, \text{ for } 0 < \tau \leq T, \\
\hat{\sigma}^2_\tau &= y^2_\tau = T^{-1} \sum_{i=1}^{T} y^2_i, \text{ for } \tau \leq 0, \\
\widehat{w}_\tau &= y^2_\tau - E(y^2_\tau | I_{\tau-1}), \text{ for } 0 < \tau \leq T, \\
\widehat{w}_\tau &= 0, \text{ for } \tau \leq 0.
\end{aligned}$$



Daily returns on STI from Jan 03, 2000 to Oct 28, 2005




ACF (with Bartlett's bounds)
Autocorrelation
Lag


PACF (with Confidence bounds)
Partial Autocorrelation
Lag