10 Chapter 10: ARCH and GARCH models

In contrast to the traditional time series analysis which focuses on modeling the conditional first moment, ARCH and GARCH models specifically take the dependency of the conditional second moment into modeling consideration and accommodate the increasingly important need to explain and model risk and uncertainty in, for example, financial time series.

The ARCH models were introduced by Engle (1982) in order to model varying (conditional) variance or volatility of time series. It is often found in economy and finance that the larger values of time series also lead to larger instability (i.e. larger variances), which is termed as *(conditional) heteroscedasticity.*

10.1 Some real data examples

Fig. 10.1 show quarterly returns of the UK and Canadian Real Estate Markets, from 1971 to 2006 and 1961 to 2005 respectively. It is clear that the returns of real estate market exhibit the largest variation around the peaks.

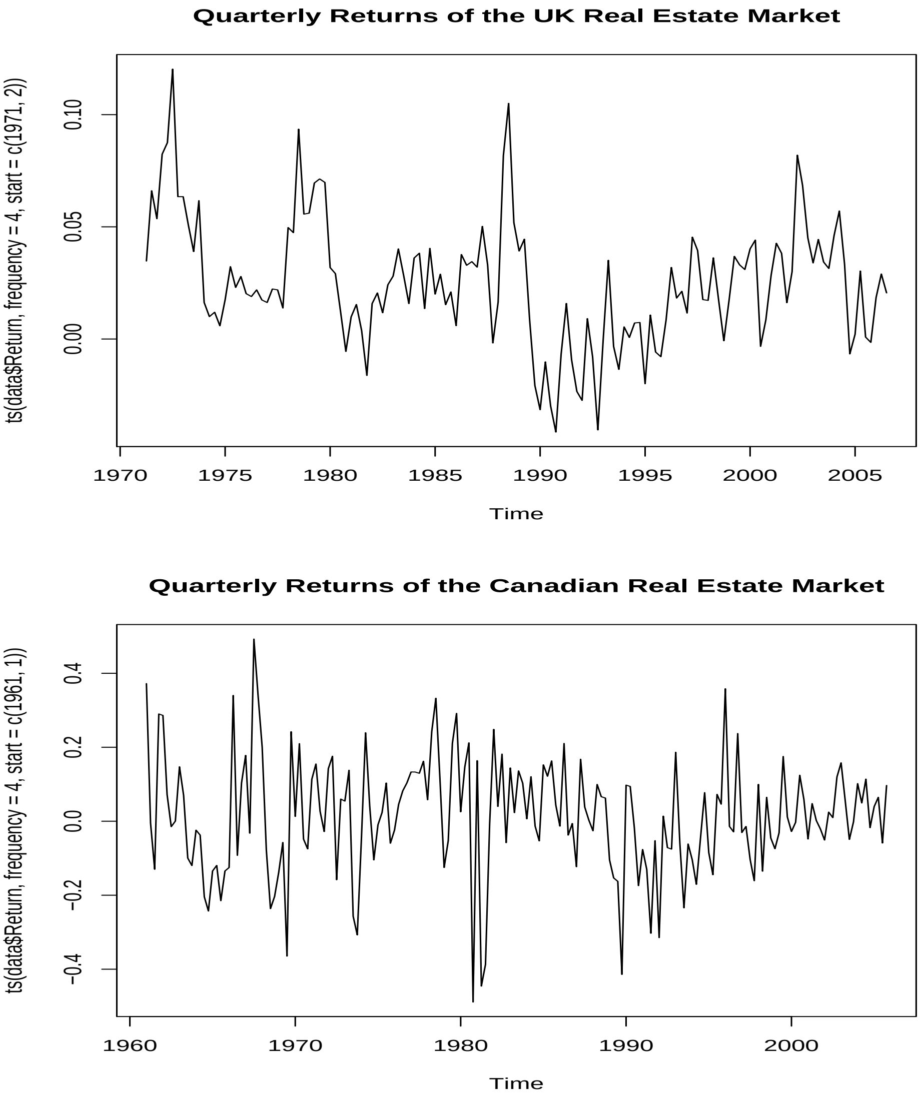

Figure 10.1: Quarterly Returns of the UK and Canadian Real Estate Markets.

Fig. 10.2 show yearly sunspot numbers from 1700 to 2005. Similarly the largest variation occurs around the peaks.

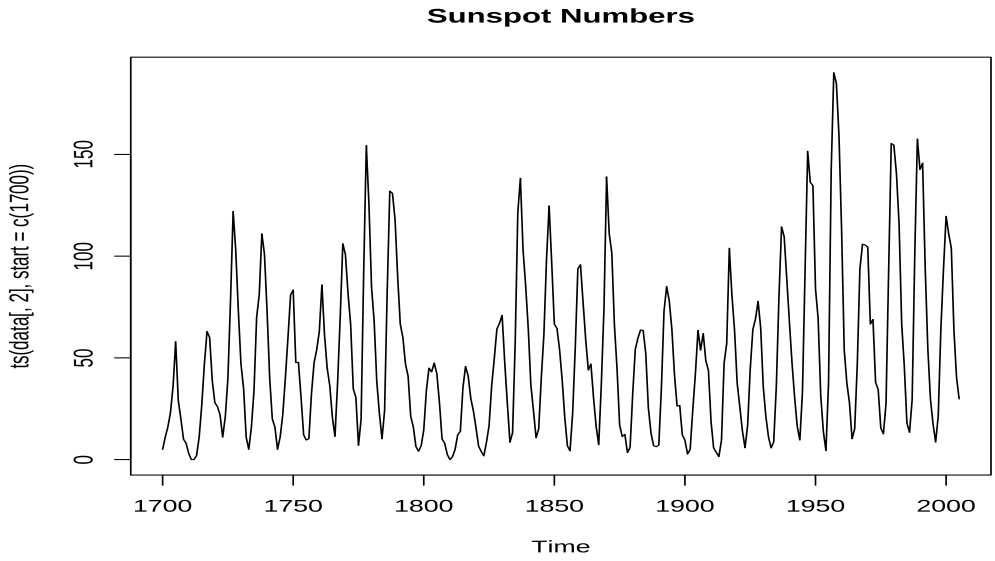

Figure 10.2: Yearly Sunspot numbers.

Standard examples of financial time series are: stock prices, interest rates and foreign exchange rates.

Since financial data typically have the correlation coefficient close to 1 at lag 1, e.g. the exchange rate between the US dollar and pounds sterling hardly changes from today to tomorrow), it is much more interesting and also practically more relevant to model the returns of a financial time series rather than the series itself. Let $\{Y_t\}$ be a stock price series, for example. The **returns** are typically defined as

$$X_t = \log Y_t - \log Y_{t-1} \quad \text{or} \quad X_t = \frac{Y_t - Y_{t-1}}{Y_{t-1}}, \tag{10.1}$$

which measure the relative changes of price. Note that the two forms above are approximately the same as

$$\log Y_t - \log Y_{t-1} = \log\left(1 + \frac{Y_t - Y_{t-1}}{Y_{t-1}}\right) \approx \frac{Y_t - Y_{t-1}}{Y_{t-1}}. \tag{10.2}$$

Rydberg (2000, "Realistic Statistical Modelling of Financial Data") summarizes some important **stylized features** of financial return series, which have been repeatedly observed in all kinds of assets including stock prices, interest rates, and foreign exchange rates.

(i) *Heavy tails.* It has been generally accepted that the distribution and the return X_t has tails heavier than the tails of a normal distribution. Typically, it is assumed that X_t only has a finite number of finite moments, although it is still an ongoing debate how many moments actually exist. Nevertheless, it seems a general agreement nowadays that the daily return has a finite second moment (i.e. $EX_t^2 < \infty$). This also serves as a prerequisite for ARCH/GARCH modeling.

(ii) *Volatility clustering.* The term volatility clustering refers to the fact that large price changes occur in clusters. Indeed, large volatility changes tend to be followed by large volatility changes, and periods of tranquility alternate with periods of high volatility.

(iii) *Asymmetry.* There is evidence that the distribution of stock returns is slightly negatively skewed. One possible explanation could be that trades react more strongly to negative information than positive information.

(iv) *Aggregational Gaussianity.* When the sampling frequency increases, the central limit law sets in and the distribution of the returns over a long time-horizon tends toward a normal distribution. Note that a return over k days is simply the aggregation of k daily returns:

$$\log Y_k - \log Y_0 = \sum_{t=1}^{k} (\log Y_t - \log Y_{t-1}) = \sum_{t=1}^{k} X_t. \tag{10.3}$$

(v) *Long range dependence.* The returns themselves of all kinds of assets hardly show any serial correlation, which, however, does not mean that they are independent. In fact, both squared returns and absolute returns often exhibit persistent autocorrelations, indicating possible long-memory dependence in those transformed return series.

10.2 Models

Engle (1982) defines an autoregressive conditional heteroscedastic (ARCH) model as

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= c_0 + b_1 X_{t-1}^2 + \ldots + b_p X_{t-p}^2, \end{aligned} \tag{10.4}$$

where $c_0 \geq 0,\ b_j \geq 0,\ \{\varepsilon_t\} \sim \text{IID}(0,1)$, and ε_t is independent of $\{X_{t-j},\ j \geq 1\}$. We write $\{X_t\} \sim \text{ARCH}(p)$.

It is easy to see that

$$EX_t = E\left(X_t | X_{t-1}, \ldots, X_{t-p}\right) = 0,$$

$$\text{Var}\left(X_t | X_{t-1}, \ldots, X_{t-p}\right) = \sigma_t^2,$$

$$\text{Cov}\left(X_t, X_k\right) = 0 \qquad \text{for all } t \neq k. \tag{10.5}$$

Basic idea: the predictive distribution of X_t based on its past is a scale-transform of the distribution of ε_t with the scaling constant σ_t depending on the past of the process. For example, if $\varepsilon_t \sim N(0,1)$, the predictive distribution is $N(0, \sigma_t^2)$ with the variance σ_t^2 depending on the conditions based on which the prediction was made. Further, a large predictive variance will be caused by the large absolute values of observations in the immediate past.

Bollerslev (1986) introduced a Generalized autoregressive conditional heteroscadastic (GARCH) model by replacing the second equation in 10.4 by

$$\sigma_t^2 = c_0 + a_1 \sigma_{t-1}^2 + \ldots + a_q \sigma_{t-q}^2 + b_1 X_{t-1}^2 + \ldots + b_p X_{t-p}^2, \tag{10.6}$$

where $c_0 \geq 0,\ a_j \geq 0$, and $b_j \geq 0$. We write $\{X_t\} \sim \text{GARCH(p, q)}$.

A general form of ARCH (∞) model:

$$Y_t = \rho_t \xi_t, \quad \rho_t = a + \sum_{j=1}^{\infty} b_j Y_{t-j}, \tag{10.7}$$

where $\{\xi_t\}$ is a sequence of non-negative i.i.d. random variables, $a \geq 0$, and $b_j \geq 0$.

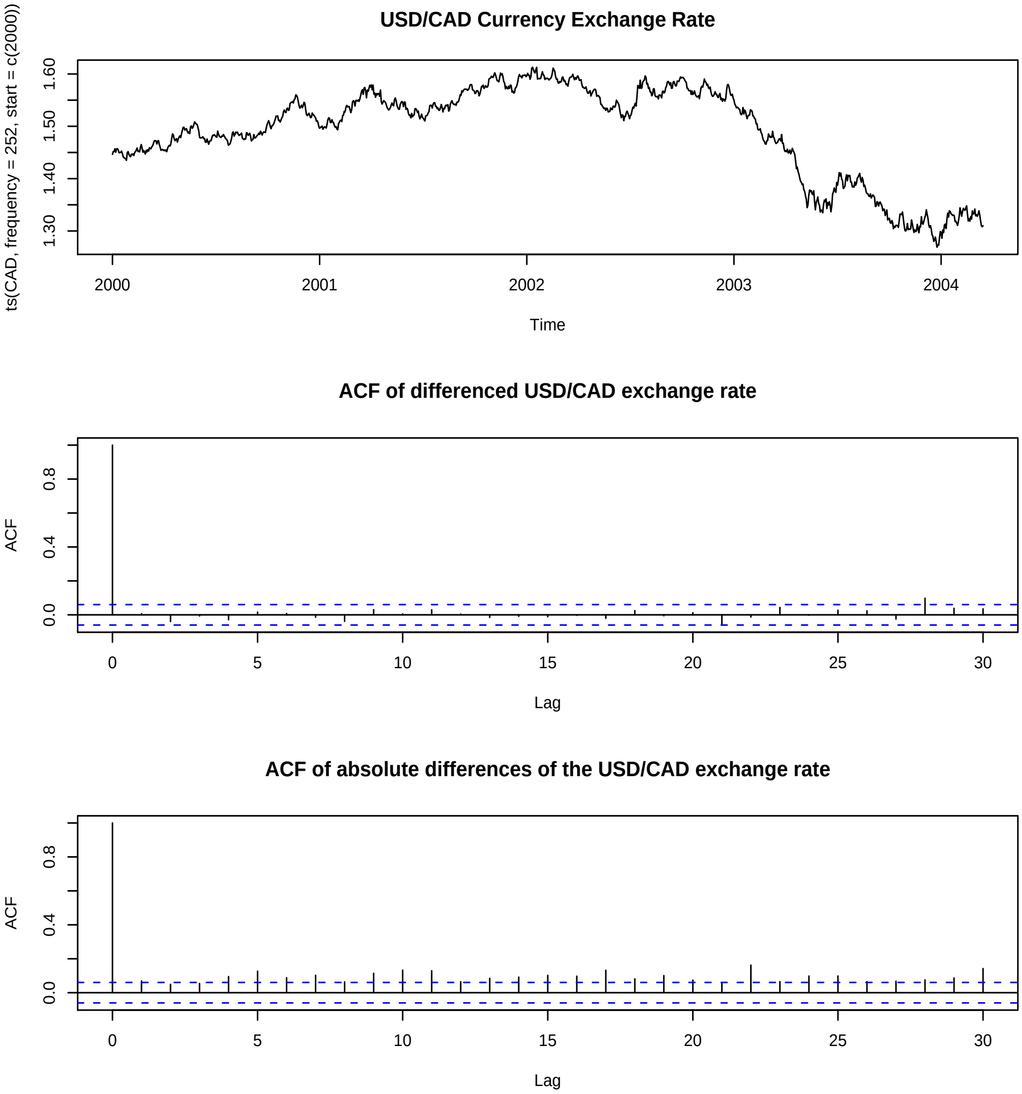

Figure 10.3: Time Series plot (top) and ACF plots for the difference $\{X_t - X_{t-1}\}$ (middle) and the absolute differences $\{|X_t - X_{t-1}|\}$ (bottom) of the USD/CAD currency exchange rate January 3, 2000 to March 31, 2004.

The above model includes the classical ARCH model 10.4 as a special case if we let $Y_t = X_t^2$.

It also contains the GARCH model if the coefficients $\{a_j\}$ fulfil certain conditions, for example, that all the roots of equations $1 - a_1 z - \ldots - a_q z^q = 0$ are greater than 1.

Theorem 1 (Giraitis, Kokoszka and Leipus (2000)).

(i) Under the condition $\sum_{j=1}^{\infty} b_j < 1$, the above ARCH($\infty$) model has a unique strictly stationary solution $\{Y_t,\ t = 0, \pm 1, \pm 2, \ldots\}$ for which

$$EY_t = \frac{a}{1 - \sum_{j=1}^{\infty} b_j}. \tag{10.8}$$

Further, the unique solution is $Y_t \equiv 0$ for all t, if $a = 0$.

(ii) Suppose that $E\xi_t^2 < \infty$ and

$$\max\left\{1, \sqrt{E\xi_t^2}\right\} \sum_{j=1}^{\infty} b_j < 1. \tag{10.9}$$

Then model has a unique strictly stationary solution $\{Y_t\}$ with $EY_t^2 < \infty$.
It follows from the theorem that the ARCH model 10.4 admits a strictly stationary solution if $\sum_{j=1}^{q} b_j < 1$.

Example. ARCH (1)

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t, \\ \{\varepsilon_t\} &\sim IID(0,1), \\ \sigma_t^2 &= a + bX_{t-1}^2, \end{aligned} \tag{10.10}$$

where $a > 0$ and $0 < b < 1$. Then

$$\begin{aligned} X_t^2 &= a\varepsilon_t^2 + bX_{t-1}^2\varepsilon_t^2 \\ &= a\varepsilon_t^2 + ab\varepsilon_t^2\varepsilon_{t-1}^2 + b^2 X_{t-2}^2\varepsilon_t^2\varepsilon_{t-1}^2 \\ &= \ldots \\ &= a\sum_{j=0}^{k} b^j \varepsilon_t^2\varepsilon_{t-1}^2 \ldots \varepsilon_{t-j}^2 + b^{k+1}X_{t-k-1}^2\varepsilon_t^2\varepsilon_{t-1}^2\ldots\varepsilon_{t-k}^2 \end{aligned} \tag{10.11}$$

Thus if $EX_0^2 < \infty$ and $\{\varepsilon_t\}$ i.i.d.,

$$E\left\{b^{k+1}X_{t-k-1}^2\varepsilon_t^2\varepsilon_{t-1}^2\ldots\varepsilon_{t-k}^2\right\} \to 0 \tag{10.12}$$

$$E\left\{a\sum_{j=0}^{k} b^j\varepsilon_t^2\varepsilon_{t-1}^2\ldots\varepsilon_{t-j}^2\right\} \to \frac{a}{1-b} \tag{10.13}$$

as $k \to \infty$. Hence

$$X_t^2 = a\sum_{j=0}^{\infty} b^j\varepsilon_t^2\varepsilon_{t-1}^2\ldots\varepsilon_{t-j}^2. \tag{10.14}$$

Consequently, $EX_t^2 = a/(1-b)$, and

$$X_t = \varepsilon_t\sqrt{a\left(1 + \sum_{j=1}^{\infty} b^j\varepsilon_{t-1}^2\ldots\varepsilon_{t-j}^2\right)}. \tag{10.15}$$

This also shows that $\{X_t\}$ is strictly stationary.

10.3 Basic properties of ARCH models

$X_t = \sigma_t\varepsilon_t$, and $\sigma_t^2 = c_0 + b_1X_{t-1}^2 + \ldots + b_pX_{t-p}^2$, where $c_0 \geq 0,\ b_j \geq 0$ are constants, $\{\varepsilon_t\} \sim IID(0,1)$, and ε_t is independent of $\{X_{t-k}, k \geq 1\}$ for all t.

Theorem 2.

(i) The sufficient and necessary condition for the above model defining a unique strictly stationary process $\{X_t, t = 0, \pm1, \pm2, \ldots\}$ with $EX_t^2 < \infty$ is $\sum_{j=1}^{p} b_j < 1$. Further,

$$EX_t = 0 \quad \text{and} \quad EX_t^2 = \frac{c_0}{1 - \sum_{j=1}^{p} b_j} \tag{10.16}$$

and $X_t \equiv 0$ for all t if $c_0 = 0$.

(ii) If $E\varepsilon_t^4 < \infty$ and

$$\max\left\{1, \sqrt{E\varepsilon_t^4}\right\} \sum_{j=1}^{p} b_j < 1, \tag{10.17}$$

the strictly stationary solution has the finite fourth moment, namely $EX_t^4 < \infty$.

The sufficiency of (i) and (ii) follows from Theorem 1 immediately. The necessity of (i) follows from Theorem 1 of Bollerslev (1986), which shows that $\sum_j b_j < 1$ is necessary for the model having a (weakly) stationary solution. (Note that Theorem 1 in Bollerslev's paper does not depend on the assumed normality.)

Remark.

(i) Stationary ARCH process $\{X_t\}$ is also WN $\left(0, \frac{c_0}{1-\sum_{j=1}^{p} b_j}\right)$.

(ii) It follows from the model that

$$X_t^2 = c_0 + b_1 X_{t-1}^2 + \ldots + b_p X_{t-p}^2 + e_t, \tag{10.18}$$

where

$$e_t = \left(\varepsilon_t^2 - 1\right) \left\{ c_0 + \sum_{j=1}^{p} b_j X_{t-j}^2 \right\}. \tag{10.19}$$

Further,

$$E\left(e_t | X_{t-k}, X_{t-k-1}, \ldots\right) = 0 \quad \text{for} \quad k \geq 1. \tag{10.20}$$

For any $k > p$,

$$\operatorname{Var}\left(X_{t+k} | X_{t-m}, m \geq 0\right) = c_0 + \sum_{j=1}^{p} b_j \operatorname{Var}\left(X_{t+k-j} | X_{t-m}, m \geq 0\right), \tag{10.21}$$

which reflects a 'volatility cluster' in financial time series analysis.

(iii) Under the additional condition 10.17, $\{e_t\} \sim$WN$(0, \sigma_e^2)$ with

$$\sigma_e^2 = \operatorname{Var}(\varepsilon_t^2) E \left\{ c_0 + \sum_{j=1}^{p} b_j X_{t-p}^2 \right\}^2 < \infty. \tag{10.22}$$

Note that the condition $\sum_{j=1}^{p} b_j < 1$ implies that $1 - \sum_{j=1}^{p} b_j z^j \neq 0$ for all $|z| \leq 1$. Hence $\{X_t^2\}$ is a causal AR(p) process. Therefore, the ACF (also ACVF) of the process $\{X_t^2\}$ can be easily calculated in terms of its MA(∞)-representation, which implies the fact that for all τ

$$\operatorname{Corr}\left(X_t^2, X_{t+\tau}^2\right) > 0, \tag{10.23}$$

although $\operatorname{Corr}(X_t, X_{t+\tau}) = 0$.

(iv) The stationary ARCH process $\{X_t\}$ has heavier tails than those of the white noise $\{\varepsilon_t\}$ based on which $\{X_t\}$ is defined.

Let $\kappa_\varepsilon = E(\varepsilon_t^4) / \left(E(\varepsilon_t^2)\right)^2$ the kurtosis of the distribution of ε_t . Then

$$\begin{aligned} E\left(X_t^4 \middle| X_{t-1}, \ldots, X_{t-p}\right) &= \sigma_t^4 E\varepsilon_t^4 = \kappa_\varepsilon \sigma_t^4 \left(E\varepsilon_t^2\right)^2 \\ &= \kappa_\varepsilon \left\{ E\left(X_t^2 \middle| X_{t-1}, \ldots, X_{t-p}\right) \right\}^2. \end{aligned} \tag{10.24}$$

Now it follows from Jensen's inequality that

$$E(X_t^4) = \kappa_\varepsilon E\left\{E\left(X_t^2 \middle| X_{t-1},\ldots,X_{t-p}\right)\right\}^2 \geq \kappa_\varepsilon \left(EX_t^2\right)^2. \tag{10.25}$$

Hence $\kappa_x \equiv E\left(X_t^4\right)/\left(EX_t^2\right)^2 \geq \kappa_\varepsilon$. In the case that ε_t is normal and $\kappa_x \geq \kappa_\varepsilon = 3$, X_t has leptokurtosis (i.e. fat tails).

Example. ARCH(1) model:

$$X_t = \sigma_t \varepsilon_t, \quad \text{and} \quad \sigma_t^2 = c_0 + b_1 X_{t-1}^2, \tag{10.26}$$

where $\{\varepsilon_t\} \sim IID(0,1), c_0 > 0$ and $b_1 \, \epsilon \, (0,1)$. Then $EX_t^2 = c_0/(1-b_1)$, and for $k \geq 1$:

$$\begin{aligned} \operatorname{Var}\left(X_{t+k}|X_{t-j},\ j\geq 0\right) &= \operatorname{Var}\left(X_{t+k}|X_t\right) = c_0 + b_1 \operatorname{Var}\left(X_{t+k-1}|X_t\right) \\ &= \frac{c_0\left(1-b_1^k\right)}{1-b_1} + b_1^k X_t^2. \end{aligned} \tag{10.27}$$

which indicates that a large value of $|X_t|$ will lead to large predictive risk (i.e. conditional variance) in a sustained period in the immediate future.

Suppose $\varepsilon_t \sim N(0,1)$. Then the condition 10.17 reduces to $3b_1^2 < 1$. Under this condition, $\operatorname{Corr}\big(X_t^2, X_{t+r}^2\big) = b_1^{|\tau|}$, and $\{X_t^2\}$ follows a causal AR(1) equation

$$X_t^2 = c_0 + b_1 X_{t-1}^2 + e_t, \tag{10.28}$$

where $e_t = \left(\varepsilon_t^2 - 1\right)\left(c_0 + b_1 X_{t-1}^2\right)$. Hence

$$\begin{aligned} EX_t^4 &= E(\sigma_t\varepsilon_t)^4 = E(\sigma_t^4)E(\varepsilon_t^4) = 3E(\sigma_t^4) = 3[E(c_0^2 + b_1^2 X_{t-1}^4 + 2c_0 b_1 X_{t-1}^2)] \\ &= 3[(c_0^2 + b_1^2 EX_{t-1}^4 + 2c_0 b_1 EX_{t-1}^2)], \end{aligned} \tag{10.29}$$

substituting $EX_{t-1}^2 = c_0/(1-b_1)$, we get

$$EX_t^4 = \frac{3c_0^2(1+b_1)}{(1-3b_1^2)(1-b_1)}. \tag{10.30}$$

Finally

$$\frac{EX_t^4}{(EX_t^2)^2} = \frac{3c_0^2(1+b_1)(1-b_1)^2}{c_0^2(1-3b_1^2)(1-b_1)} = \frac{3(1-b_1^2)}{1-3b_1^2} > 3. \tag{10.31}$$

Hence X_t has leptocurtosis (fat tails).

(i) Large values of $|X_t|$ lead to large σ_t.

(ii) $\{X_t\} \sim$WN.

(iii) Heavy tails: heavier for larger b_1 .