Financial Econometrics

Chapter 4: Volatility

In Choi

Sogang University

References

- Chapter 3 of Tsay.
- Bollerslev, T., Engle, R. F., and Nelson, D. B. (1994), "ARCH model" in Handbook of Econometrics IV, 2959–3038, ed. Engle, R. F., and McFadden, D. C. Amsterdam: Elsevier Science.
- Andersen, T., T. Bollerslev, P.F. Christoffersen and F.X. Diebold (2006), "Volatility and correlation forecasting" in Handbook of Economic Forecasting I, 777-878, ed. Elliott, G., C.W.J. Granger and A. Timmerman. Amsterdam: Elsevier Science.

Why volatility?

- Important for option pricing (see the Black–Scholes option pricing formula)
- Important for risk management. Volatility modeling provides a simple approach to calculating value at risk of a financial position.
- Important for investment in options and futures
- Modeling the volatility of a time series can improve the efficiency in parameter estimation and the accuracy in interval forecast.

Volatility models

Univariate volatility models (a partial list)

- Autoregressive conditional heteroskedastic (ARCH) model of Engle (1982)
- The generalized ARCH (GARCH) model of Bollerslev (1986)
- The exponential GARCH (EGARCH) model of Nelson (1991)
- The stochastic volatility (SV) models of Melino and Turnbull (1990), Harvey, Ruiz, and Shephard (1994), and Jacquier, Polson, and Rossi (1994)

Characteristics of volatility

1. There exist volatility clusters.
2. Volatility evolves over time in a continuous manner—that is, volatility jumps are rare.
3. Volatility varies within some fixed range. Statistically speaking, this means that volatility is often stationary.
4. Volatility seems to react differently to a big price increase or a big price drop (asymmetry in volatility).

Conditional expectation

For two continuous random variables, X and Y, we say that the conditional distribution of Y given $X = x$ is

$$f_{Y|X}\left(y|x\right) = \frac{f\left(x,y\right)}{f_X\left(x\right)}$$

where $f(x, y)$ is the joint distribution of X and Y and $f_X(x)$ is the marginal distribution of X.

Remark

(i) $f_{Y|X}(y|x)$ *is a function of* x *and possibly a different probability distribution for each* x.

(ii) When we wish to describe the entire family of distribution we use the phrase "the distribution of $Y \mid X$*".*

(iii) If X *and* Y *are independent,*

$$f_{Y|X}(y|x) = f_Y(y)$$

Conditional expectation

A conditional mean is the mean of the conditional distribution and is defined by

$$E\left[Y|X=x\right]=\begin{cases}\int_y yf_{Y|X}\left(y|x\right)dy & \text{if } y \text{ is continuous}\\ \sum_y yf_{Y|X}\left(y|x\right) & \text{if } y \text{ is discrete}\end{cases}$$

Conditional expectation

Remark

(i) Note that

$$E\left[Y|X=x\right]=E\left[Y\right]$$

if X and Y are independent.
(ii) $E(Y \mid X)$ is a random variable whose value depends on X.

Conditional expectation

Example

Define the joint pdf of (X, Y) by

$$
\begin{aligned}
f(0,10) &= f(0,20) = \frac{2}{18}, \\
f(1,10) &= f(1,30) = \frac{3}{18}, \\
f(1,20) &= \frac{4}{18}, \\
f(2,30) &= \frac{4}{18}.
\end{aligned}
$$

Conditional expectation

Example

The marginal pdf's are

$$
\begin{aligned}
f_X(0) &= f(0,10) + f(0,20) = \frac{4}{18} \\
f_X(1) &= f(1,10) + f(1,30) + f(1,20) = \frac{10}{18} \\
f_X(2) &= f(2,30) = \frac{4}{18}.
\end{aligned}
$$

Example

The conditional probability distribution of Y given that $X = 0$ is

$$f_{Y|X}(10 \quad | \quad 0) = \frac{f(0,10)}{f_X(0)} = \frac{2/18}{4/18} = \frac{1}{2}$$
$$f_{Y|X}(20 \quad | \quad 0) = \frac{f(0,20)}{f_X(0)} = \frac{2/18}{4/18} = \frac{1}{2}.$$

Example

The conditional mean given that $X = 0$ is

$$E(Y \mid X = 0) = 10 \times \frac{1}{2} + 20 \times \frac{1}{2} = 25.$$

In addition, $E(Y \mid X)$ is a random variable that takes different values depending on the value of X. (Try to tabulate its distribution!)

(i) **Law of Iterated Expectations**

$$E[Y] = E\left[E\left[Y|X\right]\right]$$

(ii)

$$E\left[g(Y)f(X)|X\right] = f(X)E\left[g(Y)|X\right]$$

Structure of volatility models

- Main motivation: The return data is either serially uncorrelated or with minor lower order serial correlations, but it is dependent.
- If r_t is *iid*,

$$\begin{aligned} & E[g(r_t) - Eg(r_t)][g(r_{t-h}) - Eg(r_{t-h})] \\ = \; & E[g(r_t) - Eg(r_t)] \times E[(g(r_{t-h}) - Eg(r_{t-h})] = 0 \end{aligned}$$

for any function $g(\cdot)$ and $h > 0$. But if r_t is not *iid*, the first equality does not hold.

- If r_t is just serially uncorrelated, we have

$$E[r_t - E(r_t)][r_{t-h} - E(r_{t-h})] = 0$$

for any $h > 0$. But this does not imply $E[g(r_t) - Eg(r_t)][g(r_{t-h}) - Eg(r_{t-h})] = 0$ for any arbitrary function $g(\cdot)$.

Structure of volatility models

- Let

$$\mu_t = E(r_t \mid F_{t-1}),\ \sigma_t^2 = Var(r_t \mid F_{t-1}) = E[(r_t - \mu_t)^2 \mid F_{t-1}],$$

where F_{t-1} denotes the information set available at time $t-1$. Typically, F_{t-1} consists of all linear functions of the past returns. Thus, we may consider the conditional variance as

$$E[(r_t - \mu_t)^2 \mid F_{t-1}] = E[(r_t - \mu_t)^2 \mid r_{t-1}, r_{t-2}, \ldots].$$

Structure of volatility models

- Assume

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

(r_t follows ARMA(p,q)). Then,

$$\sigma_t^2 = Var(a_t \mid F_{t-1}) \text{ (conditional variance of } a_t).$$

The conditional heteroskedastic models are concerned with the evolution of σ_t^2.

Structure of volatility models

- Two general categories of the conditional heteroskedastic models
  1. An exact function to govern the evolution of σ_t^2 (ARCH, GARCH).
  2. Stochastic equation to describe σ_t^2 (stochastic volatility model).
- Assume that the model for the conditional mean is given. Then, a_t is referred to as the shock or mean-corrected return of an asset return at time t.

The ARCH model

- The ARCH model:

$$a_t = \sigma_t \epsilon_t, \ \sigma_t^2 = \alpha_o + \alpha_1 a_{t-1}^2 + ... + \alpha_m a_{t-m}^2.$$

1. ϵ_t is a sequence of iid r.v. with mean 0 and variance 1.
2. $\alpha_0 > 0$ and $\alpha_i \geq 0$ for all $i > 0$.
3. The coefficients α_i satisfy some regularity conditions to ensure that the unconditional variance of a_t is finite.
4. ϵ_t is often assumed to follow the standard normal or a standardized Student-t distribution.

The ARCH model

- Large past squared shocks a_{t-i}^2 imply a large conditional variance σ_t^2. This means that, under the ARCH framework, large shocks tend to be followed by another large shock.

Consider the ARCH(1) model

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

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

- $E(a_t) = E[E(a_t \mid F_{t-1})] = E[\sigma_t E(\epsilon_t)] = 0.$
- For $h \geq 1, \ E(a_{t+h} a_t) = E[E(a_{t+h} a_t \mid F_{t+h-1})] = E[a_t \sigma_{t+h} E(\epsilon_{t+h})] = 0.$

- Assume that $Var(a_t)$ does not change over time. Since

$$
\begin{aligned}
& Var(a_t) \\
= \; & E(a_t^2) \\
= \; & E[E(a_t^2 \mid F_{t-1})] \\
= \; & E[\sigma_t^2 E(\epsilon_t^2 \mid F_{t-1})] \\
= \; & E(\sigma_t^2) \\
= \; & \alpha_o + \alpha_1 E(a_{t-1}^2) \\
= \; & \alpha_o + \alpha_1 Var(a_{t-1}),
\end{aligned}
$$

$$
Var(a_t) = \frac{\alpha_o}{1 - \alpha_1}.
$$

We also require that $0 \leq \alpha_1 < 1$ for $Var(a_t) > 0$.

- If ϵ_t follows a normal distribution and if $E(a_t^4)$ does not change over time,

$$E(a_t^4) = \frac{3\alpha_o^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}.$$

Thus we should have $0 \leq \alpha_1^2 < \frac{1}{3}$. The kurtosis of a_t is

$$\frac{E(a_t^4)}{[Var(a_t)]^2} = \frac{3\alpha_o^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} \times \frac{(1-\alpha_1)^2}{\alpha_o^2} = 3\frac{1-\alpha_1^2}{1-3\alpha_1^2} > 3,$$

implying that the tail distribution of a_t is heavier than that of a normal distribution.

- Weakness
  1. The model assumes that positive and negative shocks have the same effects on volatility.
  2. The ARCH model is rather restrictive. For instance, α_1^2 of an ARCH(1) model must be in the interval [0, 1/3] if the series is to have a finite fourth moment.
  3. The ARCH model is not structural model for the source of variations of a financial time series.
  4. ARCH models are likely to overpredict the volatility because they respond slowly to large isolated shocks to the return series.

- Steps to follow
  1. Fit an ARMA model and obtain ARMA residual a_t.
  2. Select the ARCH order.
  3. Estimate the selected ARCH model by the maximum likelihood estimation.
  4. Model checking: The standardized shocks $\frac{a_t}{\sigma_t}$ are iid random variables. Thus, use Q-stat of standardized residuals $\frac{a_t}{\sigma_t}$.

- Order Determination
  Let $\eta_t = a_t^2 - \sigma_t^2$. Then,

$$a_t^2 = \alpha_o + \alpha_1 a_{t-1}^2 + ... + \alpha_m a_{t-m}^2 + \eta_t,$$

  where η_t is a white noise process. Thus, we may use information criteria or PACF to determine the order of the ARCH process.

- Maximum likelihood estimation

  Assume $\epsilon_t \sim iidN(0,1)$. Then, the joint pdf of $a_1, ...a_T$ is (recall: $f(x,y) = f(x \mid y)f(y)$)

$$f(a_1, ..., a_T) = f(a_T \mid F_{T-1})f(a_{T-1} \mid F_{T-2})...f(a_{m+1} \mid F_m)f(a_1, ..., a_m)$$

  Ignoring the joint pdf of $a_1, ..., a_m$, the conditional (on $a_1, ..., a_m$) pdf of $a_{m+1}, ..., a_T$ is

$$\Pi_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp(-\frac{a_t^2}{2\sigma_t^2}).$$

- The conditional log-likelihood function is

$$\begin{aligned} l(a_{m+1}, ..., a_T \quad &| \quad a_1, ..., a_m, \alpha_o, ..., \alpha_m) \\ &= \sum_{t=m+1}^{T} \left(-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma_t^2) - \frac{a_t^2}{2\sigma_t^2} \right). \end{aligned}$$

The maximum likelihood estimators of $\alpha_o, ..., \alpha_m$ maximize this function. These are the parameter values that are most probable given observations.

- We may use t-distribution instead of normal. The degree of freedom is either specified or estimated along with other parameters.

- Consider an ARCH(m) model

$$\sigma_t^2 = \alpha_o + \alpha_1 a_{t-1}^2 + ... + \alpha_m a_{t-m}^2.$$

The 1-step ahead forecast of σ_t^2 is

$$\sigma_t^2(1) = \alpha_o + \alpha_1 a_t^2 + ... + \alpha_m a_{t+1-m}^2.$$

The 2-step ahead forecast is

$$\sigma_t^2(2) = \alpha_o + \alpha_1 a_t^2(1) + ... + \alpha_m a_{t+2-m}^2.$$

The l-step ahead forecast is defined similarly.

The GARCH model

- The GARCH model

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_o + \alpha_1 a_{t-1}^2 + \ldots + \alpha_m a_{t-m}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_s \sigma_{t-s}^2, \end{aligned}$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$ (This ensures that the unconditional variance of a_t is finite).

The GARCH model

- Let $\eta_t = a_t^2 - \sigma_t^2$ so that $\sigma_t^2 = a_t^2 - \eta_t$. Then, the GARCH model is rewritten as

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

The GARCH model

Example

Assume m=1 and s=2. Let $\eta_t = a_t^2 - \sigma_t^2$ so that $\sigma_t^2 = a_t^2 - \eta_t$. Then, the GARCH model is rewritten as

$$
\begin{aligned}
\sigma_t^2 &= \alpha_o + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 \\
&\Rightarrow \\
a_t^2 - \eta_t &= \alpha_o + \alpha_1 a_{t-1}^2 + \beta_1 \left(a_{t-1}^2 - \eta_{t-1} \right) + \beta_2 \left(a_{t-2}^2 - \eta_{t-2} \right) \\
&\Rightarrow \\
a_t^2 &= \alpha_o + (\alpha_1 + \beta_1)\, a_{t-1}^2 + \beta_1 a_{t-2}^2 + \eta_t - \beta_1 \eta_{t-1} - \beta_2 \eta_{t-2}.
\end{aligned}
$$

The GARCH model

- This is an ARMA model for a_t^2.
- Zero-mean

$$\begin{aligned} E(\eta_t) &= E(a_t^2 - \sigma_t^2) = E(E(a_t^2 - \sigma_t^2 \mid F_{t-1})) \\ &= E(\sigma_t^2 E(\epsilon_t^2 \mid F_{t-1})) - E(\sigma_t^2) = 0 \end{aligned}$$

- Constant variance

$$\begin{aligned} E(\eta_t^2) &= E(a_t^2 - \sigma_t^2)^2 = E(E(a_t^2 - \sigma_t^2)^2 \mid F_{t-1})) \\ &= E(E(a_t^4 - 2a_t^2\sigma_t^2 + \sigma_t^4 \mid F_{t-1})) \\ &= E(E(a_t^4 \mid F_{t-1})) - E(E(\sigma_t^4 \mid F_{t-1})) \\ &= E(\sigma_t^4 E(\epsilon_t^4)) - E(\sigma_t^4) \\ &= 2E(\sigma_t^4) \\ &= m \text{ (a constant), if } E(\sigma_t^4) \text{ is a constant.} \end{aligned}$$

The GARCH model

- For $h \geq 1$,

$$\begin{aligned} E(\eta_{t+h}\eta_t) &= E\left[E((a_{t+h}^2 - \sigma_{t+h}^2)(a_t^2 - \sigma_t^2) \mid F_{t+h-1})\right] \\ &= E\left[(a_t^2 - \sigma_t^2)E((a_{t+h}^2 - \sigma_{t+h}^2) \mid F_{t+h-1})\right]. \end{aligned}$$

But

$$\begin{aligned} E((a_{t+h}^2 - \sigma_{t+h}^2) \mid F_{t+h-1}) &= E((\sigma_{t+h}^2\epsilon_{t+h}^2 - \sigma_{t+h}^2) \mid F_{t+h-1}) \\ &= \sigma_{t+h}^2 E(\epsilon_{t+h}^2) - \sigma_{t+h}^2 \\ &= 0, \end{aligned}$$

which gives $E(\eta_{t+h}\eta_t) = 0$.

The GARCH model

- Hence,

$$E(a_t^2) = \frac{\alpha_o}{1 - \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)}.$$

- Consider the GARCH(1,1) model

$$\sigma_t^2 = \alpha_o + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2,\ 0 \le \alpha_1, \beta_1 \le 1,\ (\alpha_1 + \beta_1) < 1.$$

1. A large a_{t-1}^2 or σ_{t-1}^2 gives rise to a large σ_t^2. (volatility clustering)
2. The excess kurtosis of a_t is greater than 3.
3. Order for the GARCH model can be determined by using information criteria for the ARMA model of a_t^2.

The GARCH model

- Hence,

$$E(a_t^2) = \frac{\alpha_o}{1 - \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)}.$$

- Consider the GARCH(1,1) model

$$\sigma_t^2 = \alpha_o + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \; 0 \leq \alpha_1, \beta_1 \leq 1, \; (\alpha_1 + \beta_1) < 1.$$

1. A large a_{t-1}^2 or σ_{t-1}^2 gives rise to a large σ_t^2. (volatility clustering)
2. The excess kurtosis of a_t is greater than 3.
3. Order for the GARCH model can be determined by using information criteria for the ARMA model of a_t^2.

- Consider the GARCH(1,1) model

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

The 1-step forecast is

$$\sigma_t^2(1) = E(\sigma_{t+1}^2 \mid F_t) = \alpha_o + \alpha_1 a_t^2 + \beta_1 \sigma_t^2.$$

- For multi-step forecast, write

$$\sigma_{t+1}^2 = \alpha_o + (\alpha_1 + \beta_1)\,\sigma_t^2 + \alpha_1\sigma_t^2(\epsilon_t^2 - 1).$$

Then

$$\begin{aligned}\sigma_t^2(2) &= E(\sigma_{t+2}^2 \mid F_t) = E(\alpha_o + (\alpha_1 + \beta_1)\,\sigma_{t+1}^2 + \alpha_1\sigma_{t+1}^2(\epsilon_{t+1}^2 - 1) \\ &= \alpha_o + (\alpha_1 + \beta_1)\,\sigma_t^2(1).\end{aligned}$$

In general,

$$\sigma_t^2(l) = \alpha_o + (\alpha_1 + \beta_1)\,\sigma_t^2(l-1).$$

The integrated GARCH model

- The impact of past squared shocks $\eta_{t-i} = a_{t-i}^2 - \sigma_{t-i}^2$ for $i > 0$ on a_t^2 is persistent.
- The IGARCH(1,1) model

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_o + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1) a_{t-1}^2, \; 0 < \beta_1 < 1. \end{aligned}$$

- The IGARCH phenomenon (persistence of volatility) might be caused by occasional level shifts in volatility.

The GARCH-M model

- The GARCH-M model assumes that the return of a security may depend on its volatility.
- The GARCH(1,1)-M model

$$
\begin{aligned}
r_t &= \mu + c\sigma_t^2 + a_t, \\
a_t &= \sigma_t \epsilon_t, \\
\sigma_t^2 &= \alpha_o + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \; 0 < \beta_1 < 1.
\end{aligned}
$$

c : risk-premium parameter
r_t is serially correlated.

The exponential GARCH model

- The exponential GARCH model allows for asymmetric effects between positive and negative asset returns.
- An EGARCH(m, s) model can be written as

$$
\begin{aligned}
a_t &= \sigma_t \epsilon_t, \\
\ln(\sigma_t^2) &= \alpha_0 + \frac{1 + \beta_1 B + \ldots + \beta_s B^s}{1 - \alpha_1 B - \ldots - \alpha_m B^m} g(\epsilon_{t-1}), \\
g(\epsilon_t) &= \theta \epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)].
\end{aligned}
$$

Here $g(\epsilon_t)$ is asymmetric with respect to ϵ_t.

The exponential GARCH model

Example Let $m = 1$ and $s = 0$. Assume ϵ_t are iid standard normal. Then,

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

In this case, $E(|\epsilon_t|) = \sqrt{2/\pi}$ and

$$(1 - \alpha B)\ln(\sigma_t^2) = \begin{cases} \left((1-\alpha)\alpha_0 - \sqrt{2/\pi}\gamma\right) + (\theta + \gamma)\epsilon_{t-1}, & \epsilon_{t-1} \geq 0 \\ \left((1-\alpha)\alpha_0 - \sqrt{2/\pi}\gamma\right) + (\theta - \gamma)\epsilon_{t-1}, & \epsilon_{t-1} < 0 \end{cases}$$

The stochastic volatility model

- The model:

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

where $\epsilon_t \sim iid\ N(0,1)$, $v_t \sim iid\ N(0, \sigma_v^2)$, ϵ_t and v_t are independent, α_o is a constant, and all zeros of the polynomial $1 - \alpha_1 z - ... - \alpha_m z^m = 0$ are greater than one in modulus.

The stochastic volatility model

- Introducing the innovation v_t substantially increases the flexibility of the model in describing the evolution of σ_t^2, but it also increases the difficulty in parameter estimation.
- Quasi-likelihood or Monte Carlo method can be used to estimate the model.

See Andersen, T., T. Bollerslev, P.F. Christoffersen and F.X. Diebold (2006), "Volatility and correlation forecasting" in Handbook of Economic Forecasting I, 777-878, ed. Elliott, G., C.W.J. Granger and A. Timmerman. Amsterdam: Elsevier Science.

Brownian motion A Brownian motion or Wiener process is a stochastic process $[W(t); t \geq 0]$ with the following three properties

(i) $P[W(0) = 0] = 1$.

(ii) If $0 \leq t_0 \leq t_1 \leq ... \leq t_k$,

$$\begin{aligned} & P[W(t_i) - W(t_{i-1}) \in H_i, i = 1, ..., k] \\ = & \prod_{i=1}^{k} P[W(t_i) - W(t_{i-1}) \in H_i] \end{aligned}$$

$(W(t_k) - W(t_{k-1})$ is not affected by $W(t_1) - W(t_0), ..., W(t_{k-1}) - W(t_{k-2})$.)

(iii) $P[W(t) - W(s) \in H] = \frac{1}{\sqrt{2\pi(t-s)}} \int_H e^{-\frac{x^2}{2(t-s)}} dx$.

Realized volatility

- The model we have considered is

$$r_t = \mu_t + \sigma_t \epsilon_t,$$

where μ_t and σ_t are conditional mean and variance, respectively.

- Its continuous-time version is

$$dp(t) = \mu(t)dt + \sigma(t)dW(t). \ t \in [0, T].$$

- For small $\Delta > 0$,

$$r(t, \Delta) \equiv p(t) - p(t - \Delta) \simeq \mu(t - \Delta)\Delta + \sigma(t - \Delta)\Delta W(t),$$

where $\Delta W(t) \equiv W(t) - W(t - \Delta) \sim N(0, \Delta)$.

Realized volatility

- In addition,

$$
\begin{aligned}
r^2(t,\Delta) &= \mu^2(t-\Delta)\Delta^2 + 2\mu(t-\Delta)\Delta\sigma(t-\Delta)\Delta W(t) \\
&\quad + \sigma^2(t-\Delta)\left[\Delta W(t)\right]^2 .
\end{aligned}
$$

- The conditional variance of $r(t,\Delta)$ is

$$
Var\left[r(t,\Delta) \mid F_{t-\Delta}\right] \simeq E\left[r^2(t,\Delta) \mid F_{t-\Delta}\right] \simeq \sigma^2(t-\Delta)\Delta
$$

- Thus

$$
\begin{aligned}
RV(t,\Delta) &= \sum_{j=1}^{1/\Delta} E\left[r^2(t-1+j\Delta,\Delta) \mid F_{t-1+j\Delta}\right] \\
&\simeq \sum_{j=1}^{1/\Delta} \sigma^2(t-1+j\Delta)\Delta \simeq \int_{t-1}^{t} \sigma^2(s)ds.
\end{aligned}
$$

- As $\Delta \longrightarrow 0$,

$$RV(t, \Delta) \xrightarrow{p} \int_{t-1}^{t} \sigma^2(s) ds.$$

- It has been known that $RV(t, \Delta)$ has a long memory. it is well-fitted by ARFIMA (autoregressive fractionally integrated moving average) model.

Realized volatility

- No ARCH structure is allowed for the daily returns.
- This approach is applied to high-frequency data. For example, estimate the daily volatility by using intraday data having 5 minutes intervals.

Empirical examples

The data used are the monthly log returns of IBM stock and S&P 500 index from January 1926 to December 1999.

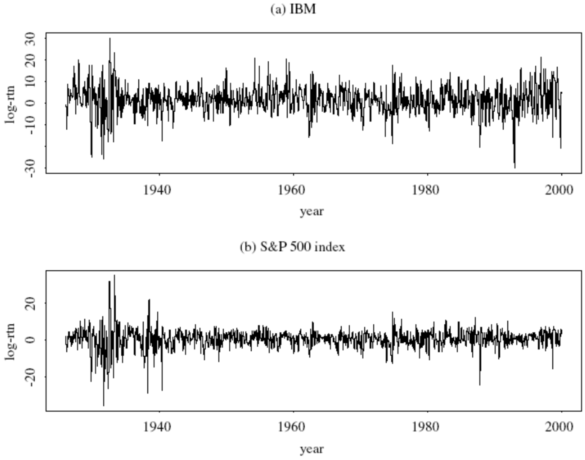

Figure 3.11. Time plots of monthly log returns for IBM stock and S&P 500 index. The sample period is from January 1926 to December 1999. The returns are in percentages and include

Empirical examples

- GARCH(1,1) modelling of the IBM stock returns

$$\begin{aligned} r_t &= 1.23 + 0.099 r_{t-1} + a_t, \ \ a_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= 3.206 + 0.103 a_{t-1}^2 + 0.825 \sigma_{t-1}^2. \end{aligned}$$

All the coefficient estimates are statistically significant.

Empirical examples

- Using the standardized residuals $\tilde{a}_t = a_t/\sigma_t$, we obtain $Q(10) = 7.82(0.553)$ and $Q(20) = 21.22(0.325)$, where p value is in parentheses. There are no serial correlations in the residuals of the mean equation.

Empirical examples

- The Ljung–Box statistics of the $\tilde{a}_t^2$ series show $Q(10) = 2.89(0.98)$ and $Q(20) = 7.26(0.99)$, indicating that the standardized residuals have no conditional heteroskedasticity.

Empirical examples

- To study the summer effect on stock volatility of an asset, define an indicator variable

$$u_t = \begin{cases} 1 & \text{if } t \text{ is June, July, or August} \\ 0 & \text{Otherwise} \end{cases}$$

and modify the volatility equation as

$$\sigma_t^2 = \alpha_o + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + u_t(\alpha_{o0} + \alpha_{10} a_{t-1}^2 + \beta_{10} \sigma_{t-1}^2).$$

Empirical examples

- The estimation results are:

$$\begin{aligned} r_t &= 1.21 + 0.099 r_{t-1} + a_t, \;\; a_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= 4.539 + 0.113 a_{t-1}^2 + 0.816 \sigma_{t-1}^2 - 5.154 u_t. \end{aligned}$$

  The summer effect on stock volatility is statistically significant at the 1% level. Furthermore, the volatility of IBM monthly log stock returns is indeed lower during the summer.

Empirical examples

- For the monthly log return series of S&P 500 index, fit a GARCH(1,1) model

$$r_t = 0.609 + a_t,\ \ a_t = \sigma_t \epsilon_t,\ \ \sigma_t^2 = 0.717 + 0.147 a_{t-1}^2 + 0.839 \sigma_{t-1}^2.$$

  Based on the standardized residuals $\tilde{a}_t = a_t/\sigma_t$, we have $Q(10) = 11.51(0.32)$ and $Q(20) = 23.71(0.26)$, where the number in parentheses denotes p value. For the $\tilde{a}_t^2$ series, we have $Q(10) = 9.42(0.49)$ and $Q(20) = 13.01(0.88)$. Therefore, the model seems adequate at the 5% significance level.