Univariate volatility models

Kevin Kotzé

Contents

Introduction

- Many valuation methods for derivatives depend on volatility
- Measures such as value-at-risk (VaR) and expected shortfall (ES) depend on volatility
- Portfolio allocation in Markowitz mean-variance framework depends on volatility
- Volatility or risk affects the spread between long and short-term interest rates
- More accurate measure of volatility would allow us to identify a miss-priced asset
- Facilitate more efficient allocation of capital
- Volatility models could be used to analyse time-varying risk premiums

Introduction

- While volatility in asset returns is not directly observable it has the following characteristics:
  - It is usually high for certain periods of time and low for other periods
  - It evolves over time in a continuous manner (volatility jumps are rare)
  - It does not diverge to infinity, as it varies within a fixed range
  - It seems to react differently to large price increases and decreases (asymmetric effects)
- If we don't not account for these features of the data in the model, it is incorrectly specified and the parameter estimates may be inaccurate

Introduction

- Three types of volatility measures for securities include:
    - Volatility as the conditional standard deviation of daily returns
    - Implied volatility that makes use of prices from options markets
    - Realised volatility that relies on high frequency financial data to calculate intra-day returns and daily volatility measures

- In what follows we focus our attention on the first case

Stylised facts

- Certain periods exhibit higher volatility than others
  - The variance is not constant it is heteroscedastic

- Large changes are associated with other large changes
  - These volatility clusters suggest that variance in t is dependent on variance in $t-1$
  - Conditional variance is dependent on time
  - Assumption of i.i.d. returns is violated
  - But volatility in 2002 would not appear to depend on the volatility in 1999 (no long-run dependence)
  - Unconditional volatility is independent of time

- These features are also displayed when taking the square or absolute value of the variable

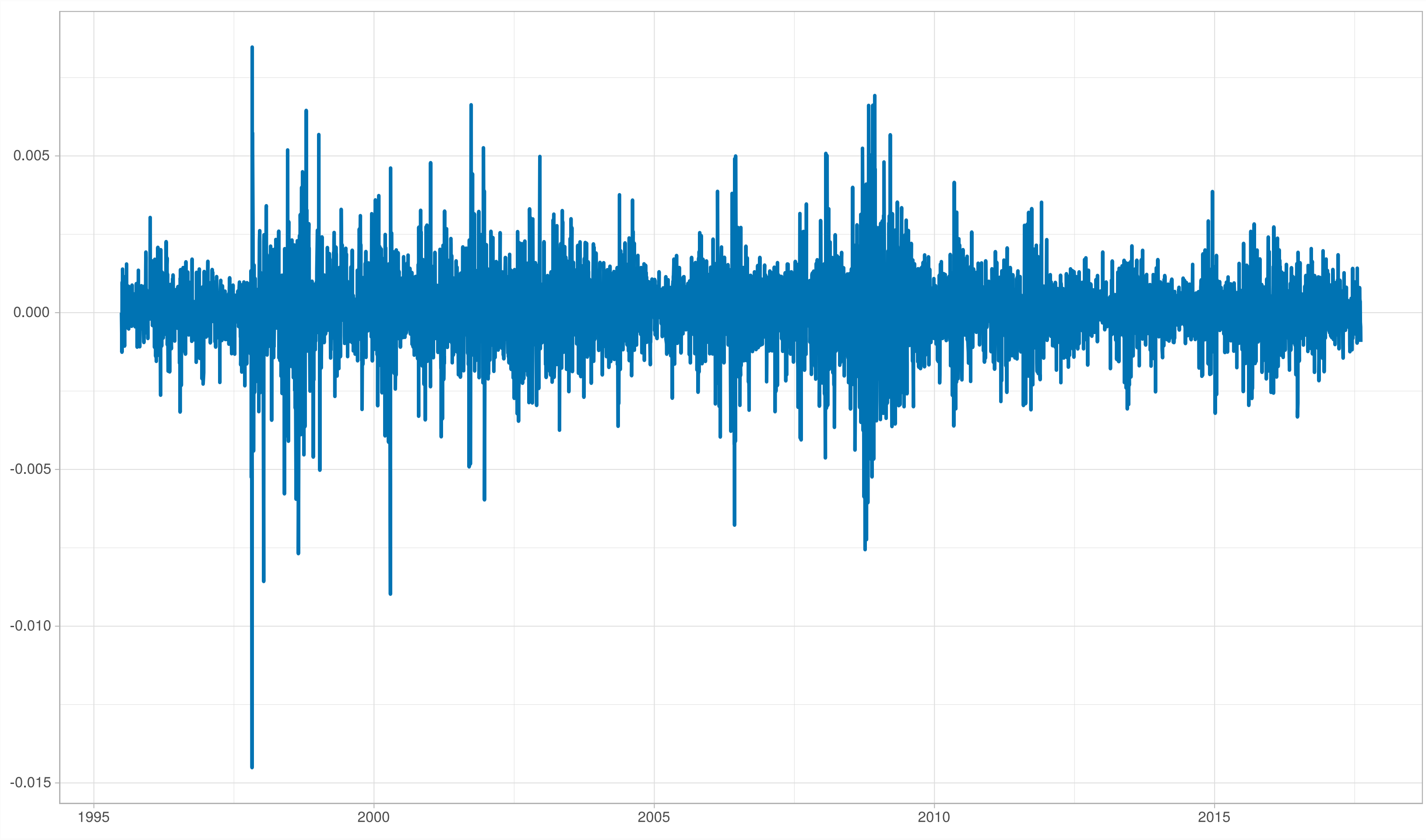

Figure : FTSE/JSE All Share (daily returns)

Stylised facts

- During certain periods the distribution of returns have fat tails:
  - This decreases with aggregation
  - May be partially attributed to volatility clustering although many assets have non-Gaussian returns

- Changes in prices are negatively related to changes in volatility:
  - Leverage effects occur in equity markets
  - Volatility rises in response to lower than expected returns

Structure of a model

- Consider the conditional mean and variance of y_t given the information set, I_{t-1}
- The first two moments are given by

$$\mu_t = \mathbb{E}\left[y_t | I_{t-1}\right]$$
$$\sigma_t^2 = \text{var}\left[y_t | I_{t-1}\right] = \mathbb{E}\left[(y_t - \mu_t)^2 | I_{t-1}\right]$$

- Usually assume that y_t follows a stationary $ARMA(p, q)$ model where $y_t = \mu_t + a_t$ and

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j}$$

- One could also include explanatory variables in the above expression

Structure of a model

- Note that now we have

$$\sigma_t^2 = \mathsf{var}(y_t|I_{t-1}) = \mathsf{var}(a_t|I_{t-1})$$

- a_t is referred to as the shock or innovation to variable y_t at time t
- μ_t is the mean equation for y_t
- σ_t^2 is the volatility equation for y_t
- Modelling conditional heteroscedasticity amounts to describing the evolution of the conditional variance over time

Constructing volatility models

- Construction of a volatility model consists of the following four steps:

1. Specify a mean equation after testing for serial dependence in the data. If necessary, make use of an econometric model (e.g. an ARMA model) for the return series to remove any linear dependence
2. Use the residuals from the mean equation to test for ARCH effects
3. Specify a volatility equation, if ARCH effects are statistically significant perform a joint estimation of the mean and volatility equations
4. Check the fitted model carefully and refine it if necessary

Testing for ARCH effects

- Assume that $a_t = y_t - \mu_t$ are the residuals from the mean equation
- The squared series a_t^2 is then used to check for conditional heteroscedasticity
- McLeod & Li (1983) apply the usual Ljung-Box statistics $Q(m)$ to a_t^2
- Null hypothesis of the test statistic is that the first m lags of ACF of the a_t^2 series are zero
- Engle (1982) uses a Lagrange multiplier test that is equivalent to an F-test for $\alpha_i = 0$, $(i = 1, \ldots, m)$ in the linear regression

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_m a_{t-m}^2 + v_t, \quad t = m+1, \ldots, T$$

ARCH Model

- ARCH models first introduced by Engle (1982) for modelling inflation
- Implications for financial risk-modelling soon became apparent
- Some researchers (notably Stock & Watson) still use of a variant of these models to forecast inflation
- Model has been extended & modified in many ways
- Surveys of the literature include, Engle & Bollerslev (1986), Bollerslev (1992), and Bera & Higgins (1993), amongst others

ARCH Model

- Basic idea of ARCH models is that
  - Shock to an asset return is serially uncorrelated but dependent
  - Dependence of a_t is described by a quadratic function of lagged values
- Specifically, an $ARCH(m)$ model assumes

$$
\begin{aligned}
a_t &= \sigma_t \varepsilon_t, \\
\sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_m a_{t-m}^2
\end{aligned}
$$

- where ε_t is a sequence of **i. i. d.** random variables with moments $[0, 1]$
- Can follow standard normal, student- t, generalised error distribution (GED), skew-distribution, etc.
- $\alpha_0 > 0$, and $\alpha_i \geq 0$ for $i > 0$
- Hence, large shocks tend to be followed by another large shock
- This may describe the volatility clustering

ARCH: Properties of ARCH Models

- Consider the $ARCH(1)$ model

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

- Note that the unconditional mean of a_t remains zero, since

$$\mathbb{E}\left[a_t\right] = \mathbb{E}\left\{\mathbb{E}\left[a_t | I_{t-1}\right]\right\} = \mathbb{E}\left\{\sigma_t \mathbb{E}\left[\varepsilon_t\right]\right\} = 0$$

- Secondly, the unconditional variance of a_t can then be derived as

$$\begin{aligned} \text{var}\left[a_t\right] &= \mathbb{E}\left[a_t^2\right] = \mathbb{E}\left\{\mathbb{E}\left[a_t^2 | I_{t-1}\right]\right\} \\ &= \mathbb{E}\left[\alpha_0 + \alpha_1 a_{t-1}^2\right] = \alpha_0 + \alpha_1 \mathbb{E}\left[a_{t-1}^2\right] \end{aligned}$$

ARCH: Properties of ARCH Models

- Since a_t is a stationary process with $\mathbb{E}\left[a_t\right] = 0$, **var**$\left[a_t\right] =$ **var**$\left[a_{t-1}\right] = \mathbb{E}\left[a_{t-1}^2\right]$
- Therefore, we have **var**$\left[a_t\right] = \alpha_0 + \alpha_1$**var**$\left[a_t\right]$ and **var**$\left[a_t\right] = \alpha_0/(1 - \alpha_1)$
- Now since the variance of a_t must be positive we require $0 \leq \alpha_1 < 1$
- When calculating the higher-order moments the excess kurtosis of a_t is positive and the tails are heavier than a normal distribution
- This is consistent with what is observed for asset price returns
- Similar findings may be provided for the $ARCH(m)$ model

ARCH: Advantages & Weaknesses

- Key advantages of using an ARCH model include
  - The model can produce volatility clusters
  - The shocks a_t in the model have heavy tails

- Weaknesses of these models include
  - Assumes positive and negative shocks have the same effect
  - Somewhat restrictive - parameters need to be within particular intervals
  - Does not provide insight into the source of variations
  - Over-predicts volatility as it responds slowly to large isolated shocks

ARCH: Making use of an ARCH Model

- If an ARCH effect is significant we can use the PACF of a_t^2 to determine the ARCH order
- Although a_t^2 is not an efficient estimate of σ_t^2 it is informative when specifying the order m
- Several likelihood functions are used in ARCH estimation, depending on the distribution of ε_t
- For complex models: starting values, optimisation algorithm, use of analytic or numerical derivatives, convergence criteria, etc. matter
- For a properly specified ARCH model the standardised residuals are given by $\tilde{a}_t = a_t/\sigma_t$
- Use Ljung-Box statistics for $\tilde{a}_t$ to check remaining correlation in mean equation, while the test on $\tilde{a}_t^2$ is used for the volatility equation
- Could use other diagnostic tests for appropriateness of these two equations
- Could forecast volatility by exploiting the recursive specification

GARCH modelling

- Although the ARCH model has a simple functional form it often requires a large m
- To simplify the model Bollerslev (1986) proposed the generalised ARCH (GARCH) model

$$
\begin{aligned}
a_t &= \sigma_t \varepsilon_t \\
\sigma_t^2 &= \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2
\end{aligned}
$$

- where ε_t is a sequence of $\mathbf{i.i.d.}$ random variables with moments $[0, 1]$
- Can follow standard normal, student- t, generalised error distribution (GED), skew-distribution, etc.
- $\alpha_0 > 0$, and $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i,j=1}^{m,s} \left(\alpha_i + \beta_j \right) < 1$
- Incorporates AR and MA components in the volatility equation, which is more parsimonious and requires less restrictions

GARCH model: Strengths & Weakness

- Consider the simple $GARCH(1,1)$ model

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad 0 \leq \alpha_1, \quad \beta_1 \leq 1, \quad (\alpha_1 + \beta_1) < 1$$

- Note that a large a_{t-1}^2 or σ_{t-1}^2 gives rise to a large σ_t^2
- Can be shown that the tails of the distribution for a $GARCH(1,1)$ process are heavier than a normal distribution
- Forecasts can utilise the recursive nature of the model
- Has similar weaknesses to that of the ARCH model

GARCH model: Forecasting

- As the volatility of time series variable is not directly observable, comparing the forecasting performance of different models can be problematic
- Some compare the volatility forecasts for $\sigma^2_{h=H}$ with the shock $a^2_{h=H}$
- However, we usually find a low correlation between $a^2_{h=H}$ and $\sigma^2_{h=H}$
- Although a^2_{t+1} is a consistent estimate of σ^2_{h+1} it is not always an accurate estimate of σ^2_{h+1}
- A single observation of a random variable with a known mean value cannot provide an accurate estimate of its variance

GARCH model: Practical points

- Estimate y_t using best ARMA model
- Obtain square of fitted residuals a_t^2
- Make use of the ACF and PACF to determine the order for s and m
- Use Q-statistics to test for groups of significant coefficients
- Rejecting the null that a_t^2 is serially uncorrelated is equivalent to rejecting the null of no ARCH & GARCH errors

GARCH model: Practical points

- In many $GARCH(1,1)$ applications, the estimated α_1 is close to zero and the estimated β_1 is close to unity
  - In this case β_1 becomes unidentified if $\alpha_1 = 0$
  - The distribution of ML estimates can be ill-behaved when parameters are nearly unidentified

- Ma, Nelson and Startz (2007) show that in a GARCH model where α_1 is close to zero
  - Estimated standard error for β_1 is usually spuriously small
  - t-statistics for testing hypotheses about the true value of β_1 are severely size distorted
  - Concentrated log-likelihood as a function of β_1 exhibits multiple maxima

Practical points on estimation

- To guard against spurious inference Ma, Nelson and Startz (2007) recommend:
    - Compare estimates from pure $ARCH(m)$ models, which do not suffer from the identification problem, with estimates from the $GARCH(1,1)$
    - If the volatility dynamics from these models are similar then the spurious inference problem is not likely to be present
    - However, if they differ then the value of β_1 may be spuriously identified and there may not be any ARCH / GARCH effects

- Alternatively, use the Engle (1982) or McLeod & Li (1983) test that may only be applied to ARCH models

The IGARCH model: Nelson (1990)

- The conditional volatility in financial returns is highly persistent
- The sum of α and β coefficients is often close to 1
- By restricting both coefficients to $\alpha + \beta = 1$ we have a parsimonious model that has interesting properties

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

where the solution for σ_t^2 is a slowly decaying exponential smoothing function and not a unit root (Nelson, 1990)

$$\sigma_t^2 = (1 - \beta_1) \left[a_{t-1}^2 + \beta_1 a_{t-2}^2 + \beta_1^2 a_{t-3}^3 + \ldots \right]$$

- The conditional variance is a decaying function of current and past a_t^2 and β_1 is the discount factor

The ARCH-M model

- Engle, Lilien and Robins (1987) extended the ARCH to allow the mean to depend on its own conditional variance
- Modern finance theory suggests that volatility may be related to risk premia on assets
- Higher volatility should result in higher risk premia since risk averse agents need compensation for risky assets
- Risk premia will be an increasing function of the conditional variance of returns

The GARCH-M model

$$
\begin{aligned}
y_t &= \mu + c\sigma_t^2 + a_t \\
a_t &= \sigma_t \varepsilon_t \\
\sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\end{aligned}
$$

- Risk-premium parameter is c
- Positive c suggests returns are positively related to past volatility

Models with Asymmetry: EGARCH

- Nelson (1991) proposed a model where the conditional variance is in log-linear form:

$$\begin{aligned} a_t &= \sigma_t \varepsilon_t, \\ \log(\sigma_t^2) &= \alpha_0 + \frac{1 + \beta_1 L + \ldots + \beta_{s-1} L^{s-1}}{1 - \alpha_1 L - \ldots - \alpha_m L^m} g(\varepsilon_{t-1}) \end{aligned}$$

- where the asymmetry in $g(\varepsilon_t)$ is given by

$$g(\varepsilon_t) = \begin{cases} (\theta + \gamma)\varepsilon_t - \gamma \mathbb{E}(|\varepsilon_t|) & \text{if } \varepsilon_{t-1} \geq 0, \\ (\theta - \gamma)\varepsilon_t - \gamma \mathbb{E}(|\varepsilon_t|) & \text{if } \varepsilon_{t-1} < 0 \end{cases}$$

Models with Asymmetry: EGARCH

- The alternative form for the $EGARCH(m, s)$ model may be expressed as

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{m} \alpha_i \frac{|a_{t-i}| + \gamma_i a_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{s} \beta_j \log(\sigma_{t-j}^2)$$

- A positive a_{t-i} contributes $\alpha_i(1 + \gamma_i)|\varepsilon_{t-i}|$ to the log volatility
- A negative a_{t-i} gives $\alpha_i(1 - \gamma_i)|\varepsilon_{t-i}|$
- Where $\varepsilon_{t-i} = a_{t-i}/\sigma_{t-i}$
- The γ_i parameter signifies the leverage effect of a_{t-i}, which is expected to be negative

Models with Asymmetry: TGARCH

- "Bad" news may effect volatility more than "good" news
- Glosten, Jaganathan, and Runkle (1994) propose Threshold GARCH (often termed GJR-GARCH)
- Consider the model

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

- where N_{t-i} is an indicator for negative a_{t-i} and

$$N_{t-i} = \begin{cases} 1 & \text{if } a_{t-i} < 0, \\ 0 & \text{if } a_{t-i} \geq 0 \end{cases}$$

- A positive a_{t-i} contributes $\alpha_i a_{t-i}^2$ to σ_t^2
- A negative a_{t-i} has a larger impact $(\alpha_i + \gamma_i) a_{t-i}^2$ with $\gamma_i > 0$

Asymmetric power ARCH model

- The general $APARCH(m, s)$ model of Ding *et al.* (1993) could be written as

$$y_t = \mu_t + a_t, \quad a_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim D(0,1)$$
$$\sigma_t^\delta = \omega + \sum_{i=1}^{m} \alpha_i \left(|a_{t-i}| + \gamma_i a_{t-i}\right)^\delta + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^\delta$$

- where δ is a positive real number
- When $\delta = 2$ the APARCH model reduces to a TGARCH model
- When $\delta = 1$ the model uses volatility directly in the volatility equation
- When $\delta \to 0$ the model reduces to the EGARCH model

Non-symmetric GARCH model

- The features of the $NGARCH(m, s)$ model are discussed in Engle & Ng (1993) and Duan (1995)
- Takes the form

$$
\begin{aligned}
y_t = \mu_t + a_t, \quad a_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim D(0,1) \\
\sigma_t^\delta = \beta_0 + \beta_1 \sigma_{t-1}^2 + \beta_2 (a_{t-1} - \theta \sigma_{t-1})^2
\end{aligned}
$$

- where μ_t is the conditional mean
- $D(0,1)$ denotes a distribution, β_i contains non-negative parameters
- Can be shown that if $\theta > 0$ and $\beta_2 > 0$, then ε_{t-1} is negatively related to σ_t^2
- Therefore, θ is a leverage parameter and should be positive

Stochastic Volatility model

- Introduce a stochastic innovation to the conditional variance equation of a_t
- Additional stochastic term is used to explain the unexpected shocks to the volatility process

$$\begin{aligned} a_t &= \sigma_t \varepsilon_t \\ (1 - \alpha_1 L - \ldots - \alpha_m L^m) \log(\sigma_t^2) &= \alpha_0 + v_t \end{aligned}$$

- where ε_t are **i. i. d.** $\mathcal{N}(0, 1)$, the v_t are **i. i. d.** $\mathcal{N}(0, \sigma_v^2)$
- ε_t and v_t are independent, while α_0 is a constant
- All the polynomials, $1 - \sum_{i=1}^{m} \alpha_i L^i$, ensure stationarity
- Although its more flexible, parameter estimation is more difficult
- Extensions include the long-memory SV model and the SV-M model

Conclusion

- Statistical models that describe volatility have been applied to a wide range of analyses
- Many financial decisions are partially based upon the amount of risk that may be incurred
- Not surprising to note that an analysis into the evolution of volatility constitutes an integral part of most asset pricing and portfolio analysis theories
- During the course of this lecture, we considered the use of various ARCH, GARCH and stochastic volatility models
- Particular attention was directed towards the relative strength and weakness of each of the characterisations