Chapter 2

Methodology

2.1 Volatility

This dissertation will be conducted to reveal alternative volatility models. It is useful to start with an explanatory of what volatility is.

Volatility refers to the standard deviation of the continuously compounded time series. Typically, in financial markets, we are interested in losses from the investments that are the deviations from the asset returns. So, volatility is itself a stock variable which would be statistically measured over a period of time.

Volatility is related to risk. It is often use to qualify the risk which is commonly associated with undesirable outcome. One of well-known volatility applications is Value-at-Risk (VaR) which is widely used for risk management.

Before volatility is estimated and modelled, features of it could be mentioned. There are several features about financial market returns and volatility.

1) Volatility is time-varying as the nature of returns fluctuation or does not remain constant through time. In the financial literature, this characteristic is called volatility clustering.
2) Asset returns have fat-tail that their kurtosis excesses 3.
3) Volatility asymmetry which volatility increases if the previous day returns are negative.

4) The returns and volatility of different assets e.g. different company shares, and different markets e.g. stock and bond markets in one or more regions, tend to have correlations or move together.

2.2 Volatility long memory

As mentioned, volatility persistence is a feature of the returns that is important in financial analysis. The long memory characteristic of financial market volatility is also important for volatility forecasting. In time series analysis, volatility has a long memory when the autocorrelation of measures of volatility decline slowly at a hyperbolic rate. Taylor (1986) was the first to show the autocorrelation of absolute returns decays slowly compared with which of squared returns. Granger and Joyeux (1980), and Hosking (1981) showed fractionally integrated in the mean process exhibits long memory property. Baillie (1996), and Baillie, Bollerslev and Mikkelsen (1996) tested for fractional integration in the conditional variance models. Corsi (2009) proposed the Heterogeneous Autoregressive model to present long memory property in Chang et al. (2009). There has been a lot of research investigating whether long memory of volatility performs far better for volatility forecasts.

In order to estimate volatility, since it does not remain constant through time, the conditional volatility models are widely used. The model specifications in modelling volatility are detailed next.

2.3 Model Specifications

Riskmetrics™

RiskMetrics™ of J. P. Morgan (1996) is a standard in the market risk measurement due to its simplicity. Basically, the RiskMetrics™ model is a model where the ARCH and GARCH coefficients are fixed to 0.06 and 0.94 respectively, which is given by

$$\sigma_t^2 = 0.06\varepsilon_{t-1}^2 + 0.94\sigma_{t-1}^2 \tag{2.1}$$

Therefore the RiskMetrics™ model is not required to estimate any unknown parameter. However, it is simply for practitioners to use.

ARCH

Engle (1982) proposed the autoregressive conditional heteroskedasticity of order *q*, or ARCH (*q*), defined as

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{2.2}$$

The parameters $\omega > 0, \alpha_1 > 0$ are sufficient to ensure positive in the conditional variance $h_t > 0$ when q = 1. The α_i represents the ARCH effect that captures the short-run persistence of shocks.

GARCH

Bollerslev (1986) generalized ARCH (q) to the GARCH (p, q) model, given by

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{2.3}$$

The parameters $\omega > 0, \alpha_1 > 0$ and $\beta_1 \geq 0$ are sufficient to ensure positive in the conditional variance, $h_t > 0$. The α_i represents the ARCH effect and β_j represents the GARCH effect that indicates the contribution of shocks to long run persistence ($\alpha_1 + \beta_1$).

GJR

Glosten, Jagannathan, and Runkle (1992) proposed the model to accommodate differential impact on the conditional variance between positive and negative shocks, hereafter the GJR model, given by

$$h_t = \omega + \sum_{i=1}^{q} (\alpha_i + \gamma_i I(\varepsilon_{t-i}))\varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{2.4}$$

where the conditional volatility is positive when parameters satisfy $\alpha_0 > 0, \alpha_i + \gamma_i \geq 0$ and $\beta_j \geq 0,$ for i = 1,…, q and j = 1,…, p. $I(\varepsilon_{t-i})$ is an indicator function that takes value 1 if $\varepsilon_{t-i} < 0$ and 0 otherwise. The impact of positive shocks and negative shocks on conditional variance is allowing asymmetric impact. The expected value of γ_i is greater than zero that means the negative shocks give higher impact than the positive shocks, $\alpha_i + \gamma_i \geq \alpha_i$. It is not possible for leverage, which negative shocks increase risk and positive shocks of equal magnitude decrease risk, to be present in the GJR model.

EGARCH

Nelson (1991) introduced the Exponential GARCH (EGARCH) model which is re-expressed by Bollerslev and Mikkelsen (1996). The EGARCH(*p*,*q*) is given by

$$\log(h_t) = \omega + \sum_{i=1}^{q} \alpha_i |\eta_{t-i}| + \sum_{i=1}^{q} \gamma_i \eta_{t-i} + \sum_{j=1}^{p} \beta_j \log(h_{t-j}) \quad (2.5)$$

where $|\eta_{t-i}|$ and η_{t-i} capture the size and sign effects of the standardized shocks respectively. The positive shocks provide less volatility than the negative shocks when $\gamma_i < 0$. Then the model allows asymmetric and leverage effects.

CCC

Bollerslev (1990) introduced the Constant Conditional Correlation (CCC) model. The CCC model assumes the matrix of conditional correlations. The CCC(1,1) is given by

$$h_{ii,t} = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i h_{ii,t-1}, \quad (2.6)$$

$$h_{ij,t} = \rho_{ij} \sqrt{h_{ii,t} h_{jj,t}} \quad (2.7)$$

where ρ_{ij} is the constant correlation between ε_{it} and ε_{jt}, which can be estimated separately from the conditional variances. The weakness of the CCC model is it cannot capture the spillover effects and asymmetric effects. However the advantage of the CCC model is in the unrestricted applicability for large systems of time series.

DCC

Engle (2002) proposed the Dynamic Conditional Correlation (DCC) model. The DCC model allow for two-stage estimation of the conditional covariance matrix. In the first stage, the univariate volatility models have been estimated and obtain h_t of each of assets. Second stage, asset returns are transformed by the estimated standard deviations from the first state, then used to estimate the parameters of DCC. The DCC model is given by

$$y_t \left| F_{t-1} \sim (0, Q_t), \qquad t = 1,...,T \right. \tag{2.8}$$

$$Q_t = D_t \Gamma_t D_t, \tag{2.9}$$

where $D_t = diag(h_{1t},..., h_{kt})$ is a diagonal matrix of conditional variances, with *m* asset returns, and F_t is the information set available to time *t*. The conditional variance is assumed to follow a univariate GARCH model as in equation (2.2). When the univariate volatility models have been estimated, the standardized residuals, $\eta_{it} = y_{it} / \sqrt{h_{it}}$, are used to estimate the dynamic conditional correlations, as follows:

$$Q_t = (1 - \phi_1 - \phi_2) S + \phi_1 \eta_{t-1} \eta'_{t-1} + \phi_2 Q_{t-1} \tag{2.10}$$

$$\Gamma_t = \left\{ (diag(Q_t)^{-1/2} \right\} Q_t \left\{ (diag(Q_t)^{-1/2} \right\} \tag{2.11}$$

where S is the unconditional correlation matrix of the ε. Equation (2.11) is used to standardize the matrix estimated in equation (2.10) to satisfy the definition of a correlation matrix. θ_1 and θ_2 are scalar parameters. In financial time series, $\theta_1 = 0$ and $\theta_2 = 1$ imply the long run conditional correlation matrix is constant which news has little practical effect in changing the purportedly dynamic conditional correlations.

VARMA-GARCH

Ling and McAleer (2003) proposed the VARMA-GARCH model, which assumes symmetry in the effects of positive and negative shocks on conditional volatility. The VARMA-GARCH model is given by

$$y_t = E(y_t | F_{t-1}) + \varepsilon_t \tag{2.12}$$

$$\varepsilon_t = D_t \eta_t , \tag{2.13}$$

$$H_t = W + \sum_{i=1}^{r} A_i \vec{\varepsilon}_{t-1} + \sum_{j=1}^{s} B_j H_{t-j} \tag{2.14}$$

where $y_t = (y_{1t},..., y_{mt})', \eta_t = (\eta_{1t},...,\eta_{mt})'$ is a sequence of independently and identically distributed random vectors.

where $H_t = (h_{1t},...,h_{mt})', \vec{\varepsilon}_t = (\varepsilon_{1t}^2,...,\varepsilon_{mt}^2)'$, and W, $A_i \; \forall i = 1,...,r$, and $B_j \forall j = 1,...,s$ are $m \times m$ matrices. As in the univariate GARCH model, VARMA-

GARCH assumes that negative and positive shocks have identical impacts on the conditional variance.

VARMA-AGARCH

McAleer et al. (2009) extended the VARMA-GARCH model into the VARMA-AGARCH model which assumed asymmetric impacts of positive and negative shocks proposed the following specification of conditional variance. The VARMA-AGARCH model is given by

$$H_t = W + \sum_{i=1}^{r} A_i \vec{\varepsilon}_{t-1} + \sum_{i=1}^{r} C_i I_{t-i} \vec{\varepsilon}_{t-i} + \sum_{j=1}^{s} B_j H_{t-j} \qquad (2.15)$$

where C_i are $m \times m$ matrices for i = 1,…,r and I_t = diag(I_{1t}, …, I_{mt}), so that

$$I = \begin{cases} 0, & \varepsilon_{i,t} > 0 \\ 1, & \varepsilon_{i,t} \le 0 \end{cases}. \qquad (2.16)$$

The VARMA-AGARCH model is reduced to the VARMA-GARCH model when $C_i = 0$ for all i.

FIGARCH

Baillie (1996), and Baillie, Bollerslev and Mikkelsen (1996) investigated a model with long-memory input for the conditional variance h_t, by inserting the additional filter $(1-L)^d$ and short-memory filter (the autoregressive moving-average (ARMA)) and then making the GARCH more general known as the fractional integration (FI) GARCH model. Fractional integration achieves long memory parsimoniously by imposing a set of infinite-dimensional restrictions on the infinite variable lags. Those restrictions are transmitted by the fractional difference operators. The FIGARCH (1,*d*,1) model is defined by

$$h_t = \omega + [1 - \beta_1 L - (1 - \phi_1 L)(1 - L)^d]\varepsilon_t^2 + \beta_1 h_{t-1} \tag{2.17}$$

where the differencing parameter *d* is between 0 and 1. The filter then represents fractional differencing which is defined by the binomial expansion as

$$(1-L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + ... \tag{2.18}$$

FIEGARCH

Bollerslev and Mikkelsen (1996) purposed the fractionally integrated GARCH (FIEGARCH) specifications.

From the EGARCH model of Nelson (1991), the returns are assumed to have conditional distributions that are normal with constant mean and with variances. The FIEGARCH(*p*,*d*,*q*) model is specified as:

$$\log(h_t) = \omega_t + \phi(L)^{-1}(1-L)^d[1+\alpha(L)]g(z_{t-1}) \tag{2.19}$$

$$g(z_{t-1}) = \theta_1 z_{t-1} + \theta_2(|z_{t-1}| - E[|z_{t-1}|]) \tag{2.20}$$

where ω_t and h_t denote conditional means and conditional variance respectively. The standardized residuals are $z_t = e_t / \sqrt{h_t}$.

ARFIMA-GARCH

Ling and Li (1997a) proposed a fractionally integrated autoregressive model with conditional heteroskedasticity, ARFIMA(*p*,*d*,*q*)-GARCH(*r*,*s*). In this model, the dependent variable exhibits significant autocorrelation between observations widely separate in time. The specifications in the conditional mean equation, y_t displays long memory, or long-term dependence which could be modelled by a fractionally integrated ARMA process, or ARFIMA process initially introduced by Granger (1980) and Granger and Joyeux (1980). The ARFIMA(*p*,*d*,*q*) is given by

$$\phi(L)(1-L)^d (y_t - \mu_t) = \theta(L)\varepsilon_t \tag{2.21}$$

This is discrete time process y_t with standard normal distribution z_t and the GARCH model as

$$h_t = \omega + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{s} \beta_j h_{t-j} \tag{2.22}$$

If $|d| < 0.5$, and $\sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j < 1$, then { y_t } is invertible and stationary.

Palma and Zevallos (2001) showed that in ARFIMA-GARCH model, the data have long memory if 0 < *d* < 0.5. The squared data have intermediate-memory if 0 < *d* < 0.25 and long memory if 0.25 < *d* < 0.5.

ARFIMA-FIEGARCH

The combination of ARFIMA filters and conditionally heteroskedastic input with long-range dependency such as FIEGARCH model gives the ARFIMA-FIEGARCH model. Robinson and Hidalgo (1997), and Palma and Zevallos (2001) showed the similar type of result for this context which the squares of the input sequence $\{\varepsilon_t\}$ has a long memory with filter parameter $d^* = d_\varepsilon + d_y < 0.5$, then the process $\{y_t\}$ has long memory. d_ε is the differencing parameter of long-memory input, FIEGARCH, and d_y is the differencing parameter of long-memory filter, ARFIMA, where $0 < d_\varepsilon, d_y < 0.5$.

There are some problems with the fractionally integrated models. Fractional integration is easy for mathematic but hard for a clear economic interpretation. It is non trivial to estimate and not easily extendible to multivariate process. The application of the fractional difference operator requires a very long build up period which results in a loss of many observations. Also, these classes of fractionally integrated models are able to reproduce only the unifractial type of scaling. In order to simplify additive volatility models which can reproduce multiscaling process, Corsi (2009) introduced the Heterogeneous Autoregressive (HAR) model mentioned next.

HAR

Corsi (2009) proposed the Heterogeneous Autoregressive (HAR) model as an alternative model for realized volatilities. This model comes from the basic idea of “Heterogeneous Market Hypothesis” of Müller et al. (1993), which recognize the presence of heterogeneity derived from the difference in the time horizon or the different autoregressive structures presented at each time scale (see McAleer and Medeiros (2008) for details). The basic idea is that agents with different time horizons perceive, react and cause different types of volatility components. In this case, the three volatility components are the short-term with daily frequency, the medium-term made of portfolio manager who rebalance their positions weekly, and the long-term with characteristic time of months. The HAR(*h*) model is based on the following process in the mean equation (see Chang et al. (2009)).

$$y_{t,h} = \frac{y_t + y_{t-1} + y_{t-2} + \ldots + y_{t-h+1}}{h} \tag{2.23}$$

where typical values of *h* in financial market are 1 for daily, 5 for weekly, and 20 for monthly data that referred to HAR(1), HAR(1,5), and HAR(1,5,20), respectively. The models of HAR(1), HAR(1,5), and HAR(1,5,20) are given by

$$y_t = \phi_1 + \phi_2 y_{t-1} + \varepsilon_t \tag{2.24}$$

$$y_t = \phi_1 + \phi_2 y_{t-1} + \phi_3 y_{t-1,5} + \varepsilon_t \tag{2.25}$$

$$y_t = \phi_1 + \phi_2 y_{t-1} + \phi_3 y_{t-1,5} + \phi_4 y_{t-1,20} + \varepsilon_t \tag{2.26}$$

2.4 Model Estimations and Distributions

Following the GARCH specification, the return process can be written as

$$y_t = \mu_t + \varepsilon_t \tag{2.27}$$

$$\varepsilon_t = z_t \sqrt{h_t} \tag{2.28}$$

The process is assumed no serial correlation in daily returns. Otherwise, an AR(1) or MA(1) or ARMA(1,1) could be added in equation (2.27).

Equation (2.27) is estimated with some appropriate specification for volatility, h_t, as given above. Typically, the GARCH models are estimated using maximum likelihood (ML) approach. The likelihood function is a function of the parameters set. We assume the process in equation (2.27) estimated under the normal distributional assumptions and the student-*t*-distribution.

If the mean equation is expressed as in equation (2.27) and equation (2.28), the log-likelihood function of the standard normal distribution is given by

$$L_{norm} = -\frac{1}{2}\sum_{t=1}^{T}\left[\ln(2\pi) + \ln(h_t) + z_t^2\right] \tag{2.29}$$

where T is the number of observations.

The log-likelihood function of the student-*t* distribution is given by

$$L_{stud} = T\left\{\ln\Gamma\left(\frac{\upsilon+1}{2}\right) - \ln\Gamma\left(\frac{\upsilon}{2}\right) - \frac{1}{2}\ln\left[\pi(\upsilon-2)\right]\right\} - \frac{1}{2}\sum_{t=1}^{T}\left[\ln(h_t) + (1+\upsilon)\ln\left(1+\frac{z_t^2}{\upsilon-2}\right)\right] \tag{2.30}$$

where υ is the degree of freedom, $2 < \upsilon \le \infty$ and $\Gamma(.)$ is the gamma function.

In the absence of normality of z_t, however, the GARCH models are estimated using maximum likelihood approach to obtain Quasi-Maximum Likelihood Estimators (QMLE) with the normal likelihood function. The log-likelihood function of the standard normal distribution is given by

$$L_{norm} = -\frac{1}{2}\sum_{t=1}^{T}\left[\ln(h_t) + z_t^2\right] \tag{2.31}$$

where T is the number of observations.

Even though z_t is not normally distributed, QMLE is consistent. The log-moment condition for consistency of QMLE is discussed in McAleer et al. (2007).

2.5 Forecasts

Estimating is useful to understand the mechanism and is a critical issue in forecasting, indeed, forecasting is crucial to implement the volatility models. Forecasts can be distinguished between in-sample and out-of-sample. In-sample forecast is based on parameters estimated using all data in the sample and assumed to be stable across time. Out-of-sample forecast is a robustness designed whether forecast is closer to reality.

In this section, the forecasts of the conditional volatility models are provided. For GARCH model, the optimal *h*-step-ahead forecast of the conditional variance is given by

$$h_{t+h|t} = \omega + \sum_{i=1}^{q}\alpha_i\varepsilon_{t+h-i|t}^2 + \sum_{j=1}^{p}\beta_j h_{t+h-j|t} \tag{2.32}$$

Equation (2.32) is computed recursively. Similarly, one can obtain the *h*-step-ahead forecast of the conditional variance of an ARCH and FIGARCH models.

2.6 Value-at-Risk, Daily capital charge, and the number of violations

In financial risk management, Value-at-Risk (VaR) is widely used measure of loss on a portfolio. This is the area where volatility models play an important role and an application. The VaR measures are required only if the banks decide their own models for calculating VaR related to capital requirement.

Value-at-Risk

A VaR threshold is the lower bound of a confidence interval for the mean. Suppose the daily returns y_t following the conditional mean and a random component ε_t, i.e. $\varepsilon_t \sim D(\mu_t, \sigma_t)$ with the unconditional mean μ_t and the standard deviation σ_t. Then we can estimate VaR with various methods. The VaR threshold for y_t can be calculated by

$$VaR_t = E(y_t | I_{t-1}) - \alpha\sigma_t \qquad (2.33)$$

where α is the critical value from the distribution of ε_t to get the appropriate confidence level. σ_t can be replaced by any estimate of the conditional variance to get an appropriate VaR (see McAleer and da Veiga (2008a) for more details).

Daily capital charge and the number of violations

Basel II Accord requires banks hold their capital reserves appropriate to the risk the banks expose themselves to their investment practices. In other words, the greater risk to which the banks are exposed, the greater the amount of capital the banks will need to hold capital charges. The Basel Accord imposes penalties in the form of higher multiplicative factor *k* on banks. In order to optimize the capital charges or minimize problem, the number of violations and the VaR forecasts are taken into account and defined by (see McAleer (2009), and McAleer, Jiménez-Martin and Peréz-Amaral (2009) for more detail).

$$\underset{\{k,VaR\}}{Minimize}\, DCC_t = \sup\left\{-(3+k)\overline{VaR}_{60}, -VaR_{t-1}\right\} \tag{2.34}$$

where

DCC = daily capital charges, which is the higher of $-(3+k)\overline{VaR}_{60}$ or $-VaR_{t-1}$,

VaR_t = Value-at-Risk for day *t*, as in (2.33),

$\overline{VaR}_{60}$ = mean VaR over the previous 60 working days,

k = the Basel Accord violation penalty, that is greater than or equal zero but less than or equal one ($0 \le k \le 1$), in Table 2.1.

Banks can control their daily capital charges by a good quality of volatility in VaR and the value of *k* arising from the violation penalty. Table 2.1 shows the Basel Accord penalty zones.

Table 2.1 Basel Accord penalty zones

Zone	**Number of Violations**	***k***
Green	0 to 4	0.00
Yellow	5	0.40
	6	0.50
	7	0.65
	8	0.75
	9	0.85
Red	10+	1.00

Note: The number of violations is given for 250 business days.