Chapter 2

Methodology

This dissertation is a quantitative study based on the econometric models, such as GARCH type models, which are very important in volatility study with high frequency data, and copula method, which is newly popular in dependence studies. The reason why use these models developed for finance study is that the data of this study exhibit the same characteristics of financial time series. The data are high frequency (daily), continuous, with no time trend and exhibit obvious volatility clustering. This research will start with univariate volatility study followed by multivariate DCC-GARCH studies and then copula based dependence studies, comparison between DCC-GARCH studies and copula studies will be made.

2.1 Univariate Model

Volatility models have been very popular in empirical research in finance and econometrics since the early 1990s. The models are based on influential papers by Engle(1982) and Bollerslev (1986). All volatility models start off with a 'mean equation', which is commonly a standard ARIMA or regression model. Then involve adding a 'variance equation' to the original mean equation and which in turn models the conditional variance.

An ARIMA(p, d, q) is expressed as:

$$\left(1-\sum_i^p \theta_i L^i\right)(1-L)^d y_t=\left(1+\sum_i^q \Phi_i L^i\right)\varepsilon_t \tag{2.1}$$

where p, d, and q are integers greater than or equal to zero and refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. When one of the terms is zero, it is usual to drop AR, I or MA.

In this study, volatility models to be estimated are associated with a stationary AR (1) conditional means given by:

$$Y_t = \mu + \theta Y_{t-1} + \varepsilon_t \qquad |\theta| < 1 \tag{2.2}$$

or a MA(1) conditional means given by:

$$Y_t = \mu + \Phi\varepsilon_{t-1} + \varepsilon_t \qquad |\Phi| < 1 \tag{2.3}$$

or a ARMA(1,1) conditional means given by:

$$Y_t = \mu + \theta Y_{t-1} + \Phi\varepsilon_{t-1} + \varepsilon_t \tag{2.4}$$

where, Y_t is Air Pollution Index, ε_t is shock to API.

2.1.1 GARCH Model

The generalised autoregressive conditional heterocedasticity (GARCH) model was developed by Bollerslev (1986). It is rare for the order (p, q) of a GARCH model to be high; indeed the literature suggests that the parsimonious GARCH(1,1) is often adequate for capturing volatility in financial data. In this empirical application, (p, q) tends to be (1, 1). The conditional variance is modeled as:

$$\varepsilon_t = \eta_t\sqrt{h_t}$$

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} \qquad (2.5)$$

where h_t is conditional volatility, conditional on the information of period t-1; η_t is standardized shock to API. $\omega>0$, $\alpha \geq 0$, $\beta \geq 0$ are sufficient to ensure that the conditional variance $h_t>0$; Using results from Ling and Li and Ling and McAleer, the necessary and sufficient condition for the existence of the second moment of ε_t for GARCH (1,1) is $\alpha + \beta < 1$.

GARCH model is lack of asymmetric and leverage. It presumes that the impacts of positive and negative shocks are the same or 'symmetric'. This is because the conditional variance in these equations depends on the magnitude of the lagged residuals, not their sign. In order to accommodate the differential impacts on the conditional variance between positive and negative shocks, Glosten, Jagannathan and Runkle (1992) proposed the following specification for h_t.

2.1.2 GJR-GARCH

The threshold GARCH (TGARCH) (Glosten, Jaganathan, &Runkle, 1993) is a simple extension of the GARCH scheme with extra term(s) added to account for possible asymmetries:

$$\varepsilon_t = \eta_t\sqrt{h_t}$$

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \gamma I(\varepsilon_{t-1}) + \beta h_{t-1} \qquad (2.6)$$

where $\omega>0$, $\alpha \geq 0$, $\alpha + \gamma \geq 1$ and $\beta \geq 0$ are sufficient conditions to ensure that the conditional variance $h_t>0$. $I(\varepsilon_{t-1})$ is an indicator function, taking the values of 1 if $\varepsilon_{t-1}<0$(good news in this study) and 0 if $\varepsilon_{t-1}>0$. The impact of bad news and good news on the conditional variance in this model is different, if $\gamma > 0$, the positive innovations have a higher impact than negative ones. The GJR is asymmetric as long as γ is significant different from zero.

Regularity condition for the existence of the second moment of GJR model is $\left(\alpha + \beta + \frac{\gamma}{2}\right) < 1$. When the conditional shock(η_t) follow a symmetric distribution, the expected short run persistence is $(\alpha + \frac{\gamma}{2})$, and the contribution of shocks to expected long run persistence is $\left(\alpha + \beta + \frac{\gamma}{2}\right)$.

2.1.3 EGARCH

The EGARCH (p, q) model of Nelson (1991) can also accommodate asymmetry and specifies the conditional variance in a different way:

$$\varepsilon_t = \eta_t\sqrt{h_t}$$

$$\log h_t = \omega + \alpha|\varepsilon_{t-1}| + \gamma\varepsilon_{t-1} + \beta \log h_{t-1} \qquad (2.7)$$

EGARCH models the logarithm of conditional volatility, thereby removing the need for constraints on the parameters to ensure a positive conditional variance (Long more& Robinson, 2004). $|\varepsilon_{t-1}|$ and ε_{t-1} capture the size and sign effects of standardized shocks respectively. The presence of leverage effects can be tested by the hypothesis that $\gamma<0$ and $\gamma<\alpha<-\gamma$. The model permits asymmetries via γ and if $\gamma< 0$, negative shocks lead to an increase in volatility. Good news generates less volatility than bad news. The model is asymmetric as long as $\gamma \neq 0$.

EGARCH is asymmetric, can capture leverage, but it does not have statistical properties because we cannot differentiate $|\varepsilon_{t-1}|$.

2.1.4 GARCH in Mean

The ARCH and GARCH framework was further extended to ARCH and GARCH in mean (ARCH-M and GARCH-M) by Engle, Lillen and Robins (1987). The GARCH-M model adds a heteroskedasticity term into the mean equation. It has the specification:

$$Y_t = \mu + \theta Y_{t-1} + \lambda h_t + \varepsilon_t \qquad |\theta| < 1$$

$$\varepsilon_t = \eta_t \sqrt{h_t}$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \tag{2.8}$$

The only difference of GJR in Mean(GJR-M) and EGARCH in Mean (EGARCH-M) with GARCH-M is that they have different variance equations:

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \gamma I(\varepsilon_{t-1}) + \beta h_{t-1} \quad \text{for GJR-GARCH-M} \tag{2.9}$$

and

$$\log h_t = \omega + \alpha|\varepsilon_{t-1}| + \gamma\varepsilon_{t-1} + \beta \log h_{t-1} \text{ for EGARCH-M} \quad (2.10)$$

2.1.5 Models with Dummy

To examine the season dust effect, this study set a dummy variable D, which equals to zero when t is in non dust season and D equals to 1 when t is in dust season. Specifically, to consider the seasonal dust effect, we also employ the all four above mentioned models; the only difference is that we put an additional intercept term D with d as coefficient in all four variance equations.

2.2 Multi-variate Model

In this section, we present the model of the dynamic correlation in urban APIs and regional and national APIs. Let us consider the APIs $Y_t = (Y_{1t}, ..., Y_{kt})'$, for t = 1,..., T. The following mean equation was estimated for each series given as:

$$Y_{it} = \mu_i + aY_{it-1} + \varepsilon_{it}$$

$$\varepsilon_{it} \sim N(0, H_t) \quad (2.11)$$

where Y_{it} is API in series i at time t, ε_{it} is the error term for the API i at time t. Eq. (2.11) was then tested using the test described in Engle (1982) for the existence of ARCH. All estimated series exhibited evidence of ARCH effects. We want to examine the existence of volatility in each series and the dynamic correlations between urban APIs, regional APIs and national APIs, the MGARCH is a good choice. The specific model we use here is DCC, after Engle (2002). The parameterization is given as:

$$H_t \equiv D_t R_t D_t \qquad (2.12)$$

where D_t is the k ×k diagonal matrix of time varying standard deviations from univariate GARCH models with $\sqrt{h_{it}}$ on the ith diagonal, and R_t is the time varying correlation matrix. The log likelihood of this estimator can be written:

$$L = -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + \log(|H_t|) + r_t' H_t^{-1} r_t)$$

$$L = -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + \log(|D_t R_t D_t|) + r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t)$$

$$L = -\frac{1}{2}\sum_{t=1}^{T}(k\log(2\pi) + 2\log(|D_t|) + \log(|R_t|) + \varepsilon_t' R_t^{-1} \varepsilon_t) \qquad (2.11.)$$

where $\varepsilon_t \sim N(0;R_t)$ are the residuals standardized by their conditional standard deviation. We propose to write the elements of D_t as univariate GARCH models, so that

$$h_{it} = \omega_i + \sum_{p=1}^{p_i} a_{ip} r_{it-p}^2 + \sum_{q=1}^{Q_i} \beta_{iq} h_{it-q} \qquad (2.12.)$$

for i =1, 2,⋯⋯k with the usual GARCH restrictions for non-negativity and stationarity being imposed, such as non-negativity of variances and $\sum_{p=1}^{p_i} a_{ip} + \sum_{q=1}^{Q_i} \beta_{iq} < 1$.

The proposed dynamic correlation structure is:

$$Q_t = (1- \sum_{m=1}^{M} \theta 1_m - \sum_{n=1}^{N} \theta 2_n)\bar{Q} + \sum_{m=1}^{M} \theta 1_m \left(\varepsilon_{t-m} \varepsilon_{t-m}'\right) + \sum_{n=1}^{N} \theta 2_n Q_{t-n}$$

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{2.15}$$

Where $\bar{Q}$ is the unconditional covariance of the standardized residuals resulting from the first stage estimation, and

$$Q_t^* = \begin{bmatrix} \sqrt{q_{11}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{q_{kk}} \end{bmatrix} \tag{2.16}$$

so that Q_t^* is a diagonal matrix composed of the square root of the diagonal elements of Q_t. The typical element of R_t will be of the form $\rho_{ijt} = \frac{q_{ijt}}{\sqrt{q_{ii}q_{jj}}}$.

To investigate the seasonal effect of mean and variance, and the effect on the dynamic correlation between local, regional and national APIs, this study tries to set three seasonal dummy in both mean and variance equations, so that equation (2.11) now becomes:

$$Y_{it} = \mu_i + S_2 D_2 + S_3 D_3 + S_4 D_4 + aY_{it-1} + \varepsilon_{it} \tag{2.17}$$

equation (2.14) becomes:

$$h_{it} = \omega_i + S'_2 D_2 + S'_3 D_3 + S'_4 D_4 + \sum_{p=1}^{p_i} a_{ip} r_{it-p}^2 + \sum_{q=1}^{Q_i} \beta_{iq} h_{it-q} \tag{2.18}$$

So that D is seasonal effect vector where D_2, D_3, D_4 equals 1 when t is in summer, autumn, or winter respectively, other equations same. Spring includes March, April and May; summer includes June, July and August; autumn includes September, October and November; winter includes December, January and February.

2.3 Copula Model

Copula methods have advantages over linear correlation in that the copula-based GARCH models allow for better flexibility in joint distributions than bivariate normal or Student-t distributions.

In this study, we are interested in the time varying dependence of air pollution, especially time varying dependence of the propensity of air pollution to improve or deteriorate. So we focus on the conditional Symmetrized Joe-Clayton copula and conditional Gaussian copula of Patton (2006).

The conditional Gaussian copula function is the density of the joint standard uniform variables (u_t, v_t), as the random variables are bivariate normal with a time-varying correlation, ρ_t. Moreover, let $x_t = \phi^{-1}(u_t)$ and $y_t = \phi^{-1}(v_t)$, where $\phi^{-1}(\cdot)$denotes the inverse of the cumulative density function of the standard normal distribution. The density of the time-varying Gaussian copula is then:

$$c_t^{Gau}(u_t, v_t|\rho_t) = \frac{1}{\sqrt{1-\rho_t}} \exp\left\{\frac{2\rho_t x_t y_t - x_t^2 - y_t^2}{2(1-\rho_t^2)} + \frac{x_t^2+y_t^2}{2}\right\}$$

(2.19)

Tail dependence captures the behavior of random variables during extreme events. In our study, it measures the propensity of Shenzhen air pollution to improve or deteriorate simultaneously with regional and national air pollution. The Gumbel, Clayton and SJC copulas efficiently capture the tail dependences arising from the extreme observations caused by the asymmetry.

The density of the time-varying Clayton copula is:

$$c_t^{clay}(u_t, v_t|\theta_t) = (\theta_t + 1)(u_t^{-\theta_t} + v_t^{-\theta_t} - 1)^{-\frac{2\theta_t - 1}{\theta_t}} u_t^{-\theta_t - 1} v_t^{-\theta_t - 1} \qquad (2.20)$$

where $\theta_t \in [0, \infty)$ is the degree of dependence between u_t and v_t, $\theta_t = 0$ implies no dependence and $\theta_t \to \infty$ a fully dependent relationship. The lower-tail dependence measured by the Clayton copula is $\lambda_t^L = 2^{-\frac{1}{\theta_t}}$.

The SJC copula is Patton's (2006a) modification of the Joe–Clayton (JC) copula. It is more general because the symmetry property of the JC copula is only a special case. The density of the JC copula is:

$$C_{JC}(u, v|T^U, T^L) = 1 - \left(1 - \left\{\left[1 - (1 - u)^k\right]^{-\gamma} + \left[1 - (1 - v)^k\right]^{-\gamma} - 1\right\}^{-\frac{1}{\gamma}}\right)^{\frac{1}{k}} \qquad (2.21)$$

where $k = 1/\log_2(2 - T^U)$, $\gamma = -1/\log_2(T^L)$, $T^U \in (0,1)$ and $T^L \in (0,1)$ are the measures of the upper and lower-tail dependencies respectively. The density of the generalized SJC copula is:

$$C_{SJC}(u, v|T^U, T^L) = 0.5\left[C_{JC}(u, v|T^U, T^L) + C_{JC}(1 - u, 1 - v|T^U, T^V) + u + v - 1\right] \qquad (2.22)$$

The SJC copula is symmetric when $T^U = T^L$ and asymmetric otherwise.

The dependent process of the time varying Gaussian copula has the following form:

$$\rho_t = \Lambda_1\left(\omega + \beta\Lambda_1^{-1}(\rho_{t-1}) + \alpha\frac{1}{m}\sum_{i=1}^{m}\phi^{-1}\left(U_{1,t-1}\right)\phi^{-1}\left(U_{2,t-i}\right)\right)$$

$$\Lambda_1(x) = \frac{1-\exp(-x)}{1+\exp(-x)} \tag{2.23}$$

where $\Lambda_1(\cdot)$ is a transformation function which holds the correlation parameter ρ_t in the interval (-1,1), $\phi(\cdot)$ is the standard normal cdf and m is an arbitrary window length.

The upper and lower-tail dependences of the conditional SJC copula is as:

$$T^U = \prod\left(\beta_U^{SJC}T_{t-1}^U + \omega_U^{SJC} + \gamma_U^{SJC}\frac{1}{10}\sum_{i=1}^{10}|u_{t-1} - v_{t-1}|\right)$$

$$T^L = \prod\left(\beta_L^{SJC}T_{t-1}^L + \omega_L^{SJC} + \gamma_L^{SJC}\frac{1}{10}\sum_{i=1}^{10}|u_{t-1} - v_{t-1}|\right) \tag{2.24}$$

where $\prod$ is the logistic transformation to keep T^U and T^L within the (0, 1) interval.