Chapter 3

Methodology

The theoretical model we will not explain more in this part. Thus, we only focus on econometric models which will find the appropriate forecasting models for People's Republic of China's exchange rates return in percentage and that of Thailand's and study the dependence structure between these two countries' exchange rates return in percentage. The details of the econometric models which this study uses are as follows.

3.1 Econometric model

The research data is mainly based on People's Republic of China's exchange rates and Thailand's exchange rates. And collect 1402 observations as the research sample. The main methodology employed in this study based on the sample we collected and it has three steps: First, we will do data processing. Put our data into rate of return and use ADF and PP test to check is there any unit root in our data. Based on two-sample K-S test, it will provides the information if we can use the parametric copulas. Second, in order to through the past exchange rates return in percentage to calculate the future rate of return, and capture the exchange rates characteristic which will help governments modify the policies and guide investors understand the foreign exchange markets and the relationship between foreign exchanges. We will find the appropriate forecasting model for each country. This part will base on linear and nonlinear model selection use AIC, BIC and MAPE. Last, based on copula theories, we will analysis dependence structures between these two countries' exchange rates. If these two countries' exchange rates dependence measures are high, it means these two countries' currencies show high mutual influence. If the dependence measures are low, it means these two countries' foreign exchange markets relatively independent. This step not only help international trader solve the settlement problem and help investors make a reasonable portfolio in the future, but also help government notice currency crisis and ensure their own currency keep health growth.

3.1.1 Data processing and testing

For data processing, we will switch our original data into return in percentage. Then use ADF and PP test to check our data stationary or not. Last, we use two-sample Kolmogorov-Smirnov test to check our data obey the uniform distribution or not, it means we can use parametric copula or not.

(1) Rate of Return

Adjust information of Chinese exchange rates and Thai Baht exchange rates sample observations in form of return terms follow as:

$$X_C = \ln(\frac{P_{c,t}}{P_{c,t-1}}) \quad (3.1)$$

$$X_B = \ln(\frac{P_{B,t}}{P_{B,t-1}}) \quad (3.2)$$

where, X_C is the return on China's exchange rates, X_B is the return on Thailand's exchange rates. $P_{C,t}$ is the China exchange rates at time t, $P_{C,t-1}$ is the China's exchange rates at time t-1, $P_{B,t}$ is the Thailand's exchange rates at time t, $P_{B,t-1}$ is the Thailand's exchange rates at time t-1.

(2) Unit root test by using ADF test and PP test

1. ADF test, test the time series data for stationary follow as:

$$\Delta G_{C,t} = \alpha_1 + \beta_1 t + \delta_1 G_{C,t-1} + \sum_{i=1}^{m} s_i \Delta G_{C,t-1} + \varepsilon_{1t} \quad (3.3)$$

$$\Delta G_{B,t} = \alpha_2 + \beta_2 t + \delta_2 G_{B,t-1} + \sum_{i=1}^{m} s_i \Delta G_{B,t-1} + \varepsilon_{2t} \quad (3.4)$$

where, $G_{C,t}$ is the China's exchange rates return in percentage at time t, $G_{C,t-1}$ is the China's exchange rates return in percentage at time t-1, $G_{B,t}$ is the Thailand's exchange rates return in percentage at time t, $G_{B,t-1}$ is the Thailand's exchange rates return in percentage at time t-1. $\alpha_1, \beta_1, \delta_1, s\ \alpha_2, \beta_2, \delta_2$ are the parameters, $\varepsilon_{1t}\ \varepsilon_{2t}$ are the error term, t is the trend

The hypotheses for test are following as:

H_0: δ_i =0 (non-stationary)

H_1: δ_i <0 (stationary) when i=1, 2

If accept H_0 means the exchange rate has unit root and non-stationary but if accept H_1 means that the exchange rate has no unit root and stationary.

2. PP test, this test was developed by Phillips and Perron (1988). The model as follows:

$$\tilde{t} = t_p (\frac{w_0}{B})^{1/2} - \frac{T(B - w_0)(s_p(\hat{\varphi}))}{2B^{1/2}e} \quad (3.5)$$

where $\tilde{t}$ is the ratio of p, B is the residual of estimate, w_0 is the consistent estimate of error variance, e is the standard error from the regression test. The hypotheses for test same as ADF test.

(3) Two sample Kolmogorov-Smirnov test

K-S two-sample test is a distribution test of fitness; the purpose of this test is that test two observations come from the same distribution. It compares the cumulative distribution function with a specific distribution.

The K-S two-sample test statistic explained as follows:

$$D = \left| E_1(d) - E_2(d) \right| \tag{3.6}$$

where $E_1(d)$ and $E_2(d)$ are observations of People's Republic of China's exchange rates return in percentage and that of Thailand's empirical distribution functions.

H_0: The two samples (one of them is random uniform distribution) come from a common distribution.

H_1: The two samples (one of them is random uniform distribution) do not come from a common distribution.

The K-S two-sample test provides the information that we can use parametric copulas or not.

3.1.2 Linear and Nonlinear modeling and selection

This step will base on AIC, BIC and MAPE from linear and nonlinear models we find the appropriate exchange rates forecasting models for People's Republic of China and Thailand.

(1) Autoregressive-linear model (AR-linear Model)

The basic linear model AR-linear model explained as follows:

$$y_{ct+s} = \phi + \phi_0 y_{ct} + \phi_1 y_{1-d} + \ldots + \phi_m y_{ct-(m-1)d} + \varepsilon_{ct+s} \tag{3.7}$$

$$y_{bt+s} = \phi + \phi_0 y_{bt} + \phi_1 y_{1-d} + \ldots + \phi_m y_{bt-(m-1)d} + \varepsilon_{bt+s} \tag{3.8}$$

where y_{ct} is China's exchange rates at time t, y_{bt} is Thailand's exchange rates at time t ϕ is parameter and coefficient of y_t, ε is error term of this equation.

(2) Self-Exciting Threshold Autoregressive Model (SETAR Model)

The general Self-Exciting Threshold Autoregressive Model explained as follows:

$$y_{ct+s} = \begin{cases} \phi_1 + \phi_{10} y_{ct} + \phi_{11} y_{ct-d} + \ldots + \phi_{1L} y_{ct-(L-1)d} + \varepsilon_{ct+s} & Z_{ct} \le th \\ \phi_2 + \phi_{20} y_{ct} + \phi_{21} y_{ct-d} + \ldots + \phi_{2H} y_{ct-(H-1)d} + \varepsilon_{ct+s} & Z_{ct} > th \end{cases} \quad (3.9)$$

$$y_{bt+s} = \begin{cases} \phi_1 + \phi_{10} y_{bt} + \phi_{11} y_{bt-d} + \ldots + \phi_{1L} y_{bt-(L-1)d} + \varepsilon_{bt+s} & Z_{bt} \le th \\ \phi_2 + \phi_{20} y_{bt} + \phi_{21} y_{t-d} + \ldots + \phi_{2H} y_{bt-(H-1)d} + \varepsilon_{bt+s} & Z_{bt} > th \end{cases} \quad (3.10)$$

where y_{ct} is China's exchange rates at time t, y_{bt} is Thailand's exchange rates at time t, ϕ is the parameter and coefficient of equation, ε is error term of this equation and Z_{ct} and Z_{bt} are threshold variables in the model. The "L" is represented lower regime of model and "H" is represented the higher regime of the model.

(3) Logistic Smooth Transition Autoregressive Model (LSTAR)

The general Logistic Smooth Transition Autoregressive Model explained as follows:

$$\begin{aligned} y_{ct+s} &= (\phi_1 + \phi_{10} y_{ct} + \phi_{11} y_{ct-d} + \ldots + \phi_{1L} y_{ct-(L-1)d})(1 - G(z_{ct}, \gamma, th)) + \\ &(\phi_2 + \phi_{20} y_{ct} + \phi_{21} y_{ct-d} + \ldots + \phi_{2H} y_{ct-(H-1)d}) G(z_{ct}, \gamma, th) + \varepsilon_{ct+s} \end{aligned} \quad (3.11)$$

$$\begin{aligned} y_{bt+s} &= (\phi_1 + \phi_{10} y_{bt} + \phi_{11} y_{bt-d} + \ldots + \phi_{1L} y_{bt-(L-1)d})(1 - G(z_{bt}, \gamma, th)) + \\ &(\phi_2 + \phi_{20} y_{bt} + \phi_{21} y_{bt-d} + \ldots + \phi_{2H} y_{bt-(H-1)d}) G(z_{bt}, \gamma, th) + \varepsilon_{bt+s} \end{aligned} \quad (3.12)$$

where y_{ct} is the China's exchange rates at time t, y_{bt} is the Thailand's exchange rates at time t, ϕ is the parameter and coefficient of equation, ε is error term of this equation and Z_{ct} and Z_{bt} are threshold variables in the model. The "L" is represented lower regime of model and "H" is represented the higher regime of the model. Moreover, "G" is the logistic function and ϕ, γ , th are the parameters to be computed.

(4) Neural Network Models (NNT Model)

The Neural Network Model explained as follows:

$$y_{ct+s} = \beta_0 + \sum_{j=1}^{D} \beta_j g(\gamma_{0j} + \sum_{i=1}^{m} \gamma_{ij} y_{ct-(i-1)d}) \quad (3.13)$$

$$y_{bt+s} = \beta_0 + \sum_{j=1}^{D} \beta_j g(\gamma_{0j} + \sum_{i=1}^{m} \gamma_{ij} y_{bt-(i-1)d}) \quad (3.14)$$

where y_{ct} is China's exchange rates at time t, y_{bt} is Thailand's exchange rates at time t, the β_0 is parameter of equation. In a hidden units and activation function g.

(5) Additive Autoregressive Model (AAR Model)

The generalized non-parametric additive model (Generalized Additive Model) explained as follows:

$$y_{ct+s} = \mu + \sum_{i=1}^{m} s_i(y_{ct-(i-1)d}) \tag{3.15}$$

$$y_{bt+s} = \mu + \sum_{i=1}^{m} s_i(y_{bt-(i-1)d}) \tag{3.16}$$

where y_{ct} is China's exchange rates at time t, y_{bt} is Thailand's exchange rates at time t. S_i are smooth functions represented by penalized cubic regression.

(6) Information Criteria: Akaike Information Criteria (AIC), Schwartz Information Criteria (SIC or BIC)

The Akaike (1974, 1976) and Schwarz (1978) information criteria for selecting the most parsimonious correct model are respectively.

$$\text{Akaike:}\quad c_n(h) = \frac{-2 \cdot \ln(L_n(h))}{n} + \frac{2h}{n} \tag{3.17}$$

$$\text{Schwarz:}\quad c_n(h) = \frac{-2 \cdot \ln(L_n(h))}{n} + \frac{h \cdot \ln(n)}{n} \tag{3.18}$$

where $L_n(h)$ is the maximized value of the likelihood function for the estimated model, h is the number of parameters used, n is the each country's exchange rates sample size.

(7) The Mean Absolute Percentage Error (MAPE)

The formula of MAPE follows as:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{S_t - R_t}{S_t}\right| \tag{3.19}$$

where S_t is the actual value for each country's exchange rates, R_t is the each country's exchange rates forecast value, n is the data of forecast.

If the MAPE value is less than 10%, it is highly accurate forecast. If the MAPE value is between 10% and 20%, it is good forecast. If the MAPE value is between 20% and 50%, it is reasonable forecast. If the MAPE value is greater than 50%, it is inaccurate forecast.

3.1.3 Copulas Theory and modeling

Copula connects a joint distribution to its marginal distributions. It will solve the high dimensional (equal or more than two dimensions) process, in this paper we will focus on bivariate case or two-dimensional process. Step 3 will base on copula theories to analysis these two countries' exchange rates dependence measures.

3.1.3.1 Rank correlation

Rank correlation reflects the monotonic dependence structure between observations. The widely typical rank correlation coefficients we used are Kendall.tau and Spearman.rho.

(1) Kendall.tau

Suppose $(x_1, y_1),(x_2, y_2)$ are i.i.d vector, $x_1, x_2 \in x$, $y_1, y_2 \in y$.

τ is between [-1, 1], suppose the copula function of (x_1,y_1) is C (u, v), that the function τ explained as follows:

$$\tau = 4\iint {}_0^1 C(u,v)dC(u,v) - 1 \tag{3.20}$$

(2) Spearmam.rho

H(x, y) is the joint distribution of (x, y).

The copula function C (u, v) is given, where u=F(x), v=G (y), that the function ρ explained as follows:

$$\rho = 12\iint {}_0^1 C(u,v)dC(u,v) - 3 \tag{3.21}$$

u and v stand for uniform distribution of People's Republic of China's exchange rates return in percentage and that of Thailand respectively.

3.1.3.2 The dependence of upper tail and lower tail

Joe in 1997 describes the dependence structure of the bivariate tails distribution between left lower quadrant and right upper quadrant which we called tail dependence.

F (x) and L (y) are the marginal distribution functions of continuous random variables X and Y respectively. The correlation coefficient of the distribution in upper tail and lower tail explained as follows:

$$\lambda^{up} = \lim_{u\to 1} P\{Y > L^{-1}(u) \mid X > F^{-1}(u)\} = \lim_{u\to 1} \frac{1-2u+C(u,v)}{1-u} \qquad (3.22)$$

If $\lambda^{up} \in (0,1)$, the upper tail is dependent; if $\lambda^{up} = 0$, the upper tail is independent.

$$\lambda^{low} = \lim_{u\to 1} P\{Y < L^{-1}(u) \mid X < F^{-1}(u)\} = \lim_{u\to 1} \frac{C(u,v)}{1-u} \qquad (3.23)$$

If $\lambda^{low} \in (0,1)$, the low tail is dependent; if $\lambda^{low} = 0$, the lower tail is independent.

u and v stand for uniform distribution of People's Republic of China's exchange rates return in percentage and that of Thailand respectively.

3.1.3.3 Maximum Likelihood estimation method

We suppose f(x, y) as the joint distribution F(x, y) density function; we do the partial derivative on both sides of equation

$$H(u_1,...u_n) = C(F_1(u_1),...F_n(u_n))$$

We get:

$$f(x, y) = C_\theta(F_x(x;\alpha), F_Y(y;\beta)) \cdot f_x(x;\alpha) \cdot f_Y(y;\beta) \qquad (3.24)$$

where f(x; α), f(y; β) is the marginal density function of f(x, y), α, β are the parameters, θ is the copula parameter.

The copula density function explained as follows:

$$C_\theta(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v} \qquad (3.25)$$

u and v stand for uniform distribution of People's Republic of China's exchange rates return in percentage and that of Thailand respectively.

Using the maximum likelihood method, the log-likelihood function of equations (3.24) explained as follows:

$$l(v) = \sum_{t=1}^{T} \ln c(F_X(x_t;\alpha), F_Y(y_t;\beta);\theta) + \sum_{t=1}^{T} \ln f_X(x_t;\alpha) + \sum_{t=1}^{T} \ln f_Y(y_t;\beta) \quad (3.26)$$

where α,β,θ, are the parameter vectors of marginal distribution and copula function, respectively.

3.1.3.4 Generalized Pareto Distribution (GPD) models

Since the distribution of excesses can be estimated by a GPD, this study using the GPD functions to estimate the tail of the original distribution F. We use GPD converse the exceed threshold sequence into the [0, 1] uniform distribution.

3.1.3.5 Copula modeling

As we know if the marginal distribution is not continuous, the best empirical copula will not unique. For some parametric copulas the density function will not easy to calculate. Therefore we put empirical copula and an appropriate parametric copula which we selected together analysis the dependence measures.

(1) The empirical copula

Deheuvels in 1979 introduced Empirical copula theory. Suppose that the marginal distribution F is continuous, therefore the copula associated to F is unique.

Copula explained as follows:

$$C = \{(\frac{t_1}{T},...,\frac{t_N}{T});1 \le n \le N, t_n = 0,...,T\} \tag{3.27}$$

Through

$$\hat{C}(\frac{t_1}{T},...,\frac{t_N}{T}) = \frac{1}{T}\sum_{t=1}^{T}\prod_{n=1}^{N} 1_{[\gamma_n^t \le t_n]} \tag{3.28}$$

We will get the empirical copula.

Suppose $\{x_k, y_k\}_{k=1}^{n}$ stand for the bivariate time series observation distribution and length of the observation is n, the Empirical copula C_n explained as follows:

$$C_n(\frac{i}{n},\frac{j}{n}) = \{\sum \text{ The sample which satisfied } x \le x_i \cap y \le y_j\}/\text{n} \tag{3.29}$$

x, y stand for the People's Republic of China's exchange rates return in percentage and that of Thailand, respectively.

(2) Gumbel Copula

$$C(u,v) = \exp(-[(-\ln u)\theta + (-\ln v)\theta]^{\frac{1}{\theta}} \quad (3.30)$$

when $\theta = 1$, the random variables u and v are independent; when $\theta = +\infty$, the random variables u and v are completely related. The density distribution shows "J" form which means that the upper tail dependence is high, and the copula reflects on the change of upper tail sensitively.

u and v stand for uniform distribution of People's Republic of China's exchange rates return in percentage and that of Thailand respectively.

(3) Clayton Copula

$$C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \quad (3.31)$$

where the parameters $\theta \in (0,+\infty)$, when $\theta \to 0$, the random variables u and v are independent, when $\theta = +\infty$, the random variables u and v tend to be perfect correlation. The density distribution shows "L" form which means that the lower tail dependence is high. This copula reflects the change of lower tail sensitively.

u and v stand for uniform distribution of People's Republic of China's exchange rates return in percentage and that of Thailand respectively.

(4) Frank Copula

$$C(u,v) = -\frac{1}{\theta}\ln(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}) \tag{3.32}$$

Where the $\theta \neq 0$, if $\theta > 0$, the random variables u and v are positive correlation; if $\theta \to 0$ the random variables u and v are independent; if $\theta < 0$, the random variables u and v are negative correlated. The density distribution shows symmetric situation.

u and v stand for uniform distribution of People's Republic of China's exchange rates return in percentage and that of Thailand respectively.

(5) Plackett Copula

The properties of Plackett copula are: first, for general marginal processes the density function is high elastic; second, a single parameter will explain the copula's two margins association.

Plackett copula (with $\theta > 0$) explained as follows:

$$C(u,v|\theta) = \{\frac{1}{2}(\theta-1)L - \sqrt{L^2 - 4uv(\theta-1)}\} \quad \text{if } \theta \neq 1 \qquad (3.33)$$

$$= uv \qquad \text{if } \theta = 1$$

where $L = 1 + (\theta-1)(u+v)$. There is zero tail dependence.

u and v stand for uniform distribution of People's Republic of China's exchange rates and Thailand exchange rates respectively.

(6) t Copula

Mashal & Zeevi (2002) and Breymann et al. (2003) in their paper showed the empirical t copula fit results. The t copula explains the dependence structure which the margins follow multivariate t distribution and we also call it Gaussian copula.

The t copula explained as follows:

$$C^t_{\nu,P}(u) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \cdots \int_{-\infty}^{t_\nu^{-1}(u_d)} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\pi\nu)^d |P|}} (1 + \frac{x'P^{-1}x}{\nu})^{-\frac{\nu+d}{2}} dx \qquad (3.34)$$

where t_ν^{-1} is the standard univariate t_ν distribution quantile function.

The density function of t copula explained as follows:

$$C^{t}_{\nu,P}(u)=\frac{f_{\nu,P}(t_{\nu}^{-1}(u_1),\dots,t_{\nu}^{-1}(u_d))}{\Pi_{i=1}^{d} f_{\nu}(t_{\nu}^{-1}(u_i))}, u\in(0,1)^d \tag{3.35}$$

where $f_{\nu,p}$ is the joint density function and f_{ν} is the density function of the univariate standard t distribution which freedom degree is ν .

u and v stand for uniform distribution of People's Republic of China's exchange rates return in percentage and that of Thailand respectively.