Chapter 2

Related Theories and Literature Reviews

In this chapter this study will introduce the exchange rate determination theoretical theories and some econometric theories which this study use in our study and help us understand the foreign exchange market mechanism. The theoretical theories have two parts: the asset market view of exchange rate determination and monetary models of exchange rate determination, respectively. The theoretical model this study will not use them to modeling our study. For the econometric theories, this study lists some of the models this study will use in our study later, the actual modeling process will in chapter 3.

2.1 The Theoretical model theories

2.1.1 The Asset Market View of Exchange Rate Determination

This simple model assumes that the logarithm of the exchange rate at time t, e (t); is determined by

$$e(t) = X(t) + \alpha \cdot E\{[e(t+1) - e(t)]; t\} \tag{2.1}$$

where $E\{[e(t+1)-e(t)];t\}$ denotes the expected percentage rate of change of the exchange rate between t and t + 1, conditional on information available at t , and where X (t) represents the basic conditions of supply and demand that affect the foreign exchange market at time t.

For $j \geq 0$, depends on a weighted average of expected future X ' s , starting at t + j and extending farther into the future; specifically,

$$E(e(t+j);t) = (1/(1+\alpha)) \cdot \sum_{i=0}^{\infty} (\alpha/(1+\alpha))^{i} \cdot E(X(t+j+i);t) \tag{2.2}$$

Setting j = 0, we obtain the expression for the current exchange rate as a weighted average of present and expected future X's.

Using equation (2.2), we may obtain a convenient decomposition of the actual change in the exchange rate, $D[e(t)] = e(t+1) - e(t)$ into its expected change component, $D^{e}[e(t)] = E\{D[e(t)];t\} = E[e(t+1);t] - e(t)$, and its unexpected change component $D^{u}[e(t)] = e(t+1) - E[e(t+1);t]$. Specifically, applying the expected change operator $D^{e}(\ )$ to (2) with j = 0, we may conclude that

$$D^{e}[e(t)] = [1/(1+\alpha)] \cdot \sum_{i=0}^{\infty} [\alpha/(1+\alpha)]^{i} \cdot E\{D[X(t+j);t]\} \tag{2.3}$$

Thus, the expected change in the exchange rate is a weighted average of all expected future changes in the X's. This result may also be written in the alternative form,

$$D^{e}[e(t)] = [1/(1+\alpha)] \cdot \{E[e(t+1);t] - X(t)\} \qquad (2.4)$$

Which expresses the expected change in the exchange rate as proportional to the difference between the weighted average of all expected future X's that determines E[e(t + 1); t] and the current X (t) , with a factor of proportionality of [1/(1 + a)]. The unexpected change in the exchange rate is determined by applying the unexpected change operator D^{e} ( ) to (2.2) with j = 0;

$$D^{u}[e(t)] = [1/(1+\alpha)] \cdot \sum_{i=0}^{\infty}[\alpha/(1+\alpha)]^{i} \cdot \{E[X(t+j+1);t+1] - E[X(t+j+1);t]\} \qquad (2.5)$$

Thus, the unexpected component of the change in the exchange rate is a weighted average of the change in expectations about future X's, based on new information that is received between t and t + 1.

From these results, it is possible to argue that expected changes in exchange rates are unlikely to be very large.

2.1.2 Monetary Models of Exchange Rate Determination

The first class of monetary models, which have been widely applied in empirical studies of exchange rate behavior, expresses the current exchange rate as a function of the current stocks of domestic and foreign money and the current determinants of the demands for these monies, including domestic and foreign income and interest rates. The second class of monetary models, which has been more widely used in theoretical work, focuses on the influence on the current exchange rate of the expected future path of money supplies and of factors affecting money demands. The essential content of the first class of monetary models may be summarized in an equation of the form where e is the logarithm of the price of foreign money in terms of domestic money, m is the logarithm of the domestic money supply, I is the logarithm of demand for domestic money (a function of domestic income, y, the domestic interest rate, i, and other factors k), and an asterisk (*) indicates variables for the foreign country.

$$e = m - m^{*} - (l[y, i, k] - l^{*}[y^{*}, i^{*}, k^{*}]) \tag{2.6}$$

The critical condition determining the exchange rate for this country is the requirement of money market equilibrium; where m is the logarithm of the domestic money supply, e is the logarithm of the price of foreign money in terms of domestic money, k summarizes all exogenous factors affecting the logarithm of the demand for domestic money, and $D^e(e)=E(e(t+1);t)-e(t)$ is the expected rate of change of the exchange rate. Equation (2.7) should be thought of as a reduced-form equilibrium condition derived from a more basic model of goods and asset market equilibrium. In this reduced form, the parameter ζ captures all of the mechanisms through which an increase in the price of foreign money increases the demand for domestic money, and the parameter η captures all of the mechanisms through which an increase in the expected rate of change of the price of foreign money affects the demand for domestic money.

$$m = k + \zeta \cdot e - \eta \cdot De(e), \quad \zeta, \eta > 0 \tag{2.7}$$

Since the reduced-form demand for domestic money depends on the expected rate of change of the exchange rate, it follows that the current equilibrium exchange rate depends not only on the current values of m and k , but also on the expectation of next period's exchange rate;

$$e(t) = [1/(\zeta + \eta)] \cdot [m(t) - k(t)] + [\eta/(\zeta + \eta)] \cdot E(e(t+1); t) \tag{2.8}$$

Forward iteration of (2.8), justified by the assumption of rational expectations, leads to the conclusion that the exchange rate expected at any future date is an exponentially weighted sum of expected future differences between m and k;

$$E(e(s);t) = [\zeta/(\zeta+\eta)] \cdot \sum_{j=0}^{\infty} \eta/(\zeta+\eta)]^{j} E(w(s+j);t), \qquad (2.9)$$

where $w(u) = (1/\zeta) \cdot [m(u) - k(u)]$. The current exchange rate $e(t) = E(e(t);t)$ is found by setting s = f in (2.9). This result reveals the fundamental principle that the current exchange rate depends on the entire future expected path of differences between (the logarithms of) the money supply and the exogenous component of money demand.

The function (2.9) may be used to decompose the change in the exchange rate into its expected and unexpected components. The expected change in the exchange rate is given by

$$D^{e}[e(t)] = (\zeta/(\zeta+\eta)) \cdot [E(e(t+1);t) - E(w(t);t)] \qquad (2.10)$$

The unexpected change in the exchange rate is given by

$$D^{u}[e(t)] = (\zeta/(\zeta+\eta)) \cdot \sum_{j=0}^{\infty} (\eta/(\zeta+\eta))^{j} \cdot [E(w(t+j+1);t+1) - E(w(t+j+1);t)] \quad (2.11)$$

2.2 The Econometric model theories

2.2.1 CLS (Conditional Least Squares) method

The approach which this study develop an estimation procedure for dependent observations based on the minimization of a sum of squared deviations about conditional expectations, we call "Conditional Least Squares" (CLS). This approach provides a unified treatment of estimation problems for widely used classes of stochastic models.

2.2.2 Nonlinear method models

In time series literature reviews, linear methods often run better than nonlinear methods and they often provide an adequate approximation for the object well. However, for the real economic situation linear methods could not capture the behaviors of data or observations accurately. Therefore using non-linear models will provide a potentially promising. And this study also see linear model as one of the special nonlinear models.

Nonlinear autoregressive time series models consider the discrete-time univariate stochastic process $\{X_t\}_{t\in T}$. Suppose X_t is generated by the map:

$$X_{t+s} = f(X_t, X_{t-d}, ...X_{t-(m-1)d}; \theta) + \varepsilon_{t+s}$$

With $\{X_t\}_{t\in T}$ white noise, ε_{t+s} independent w.r.t. X_{t+s}, and with f a generic function from R^m to R. This class of models is frequently referenced in the literature with the acronym NLAR (m), which stands for Nonlinear Auto Regressive of order m.

In this function, this study has implicitly defined the embedding dimension m, the time delay d and the forecasting steps s. The vector θ indicates a generic vector of parameters governing the shape of f, which this study would estimate on the basis of some empirical evidence (i.e., an observation time series $\{ x_1, x_2, ..., x_N \}$).

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

The basic linear model is AR (m) model and also this model can be written in equation:

$$y_{t+s} = \phi + \phi_0 y_t + \phi_1 y_{1-d} + ... + \phi_m y_{t-(m-1)d} + \varepsilon_{t+s} \quad (2.12)$$

The equation (2.12) is represented the AR (m) model and y_t is time series data at time t, ϕ is parameter and coefficient of y_t in the model. In addition, ε is error term of this equation.

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

The general Self-Exciting Threshold Autoregressive Model or SETAR model can be written in equation:

$$y_{t+s} = \begin{cases} \phi_1 + \phi_{10} y_t + \phi_{11} y_{t-d} + \ldots + \phi_{1L} y_{t-(L-1)d} + \varepsilon_{t+s} & Zt \leq th \\ \phi_2 + \phi_{20} y_t + \phi_{21} y_{t-d} + \ldots + \phi_{2H} y_{t-(H-1)d} + \varepsilon_{t+s} & Zt > th \end{cases} \quad (2.13)$$

The equation (2.13) is represented the SETAR models and y_t is time series data at time t, ϕ is the parameter and coefficient of equation. In addition, ε is error term of this equation and Z_t is a threshold variable 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 model)

The general Logistic Smooth Transition Autoregressive Model or LSTAR model can be written in equation:

$$\begin{aligned} y_{t+s} &= (\phi_1 + \phi_{10} y_t + \phi_{11} y_{t-d} + \ldots + \phi_{1L} y_{t-(L-1)d})(1 - G(z_t, \gamma, th)) + \\ &(\phi_2 + \phi_{20} y_t + \phi_{21} y_{t-d} + \ldots + \phi_{2H} y_{t-(H-1)d}) G(z_t, \gamma, th) + \varepsilon_{t+s} \end{aligned} \quad (2.14)$$

The equation (2.14) is represented the LSTAR model and y_t is the time series data at time t, ϕ is the parameter and coefficient of equation. In addition, ε

is error term of this equation and Z_t is a threshold variable 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)

From the neural network model was used for estimation in this research can be explained by equation:

$$y_{t+s} = \beta_0 + \sum_{j=1}^{D} \beta_j g(\gamma_{0j} + \sum_{i=1}^{m} \gamma_{ij} y_{t-(i-1)d}) \quad (2.15)$$

The equation (2.15) is represented the NNT model and y_t is time series data 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) or AAR model can be written in equation:

$$y_{t+s} = \mu + \sum_{i=1}^{m} s_i(y_{t-(i-1)d}) \quad (2.16)$$

The equation (2.16) is represented the generalized non-parametric additive model and y_t is time series data at time t. S_i are smooth functions represented by penalized cubic regression.

2.2.3 Copulas

Copula is a function that connects multi-dimensional distribution function with one-dimensional marginal distribution function, it's emergence not only bring risk analysis and multivariate time series analysis to a new stage, but also promote the financial risk measurement method to a new breakthrough which measuring the risk of non-ellipsoidal shape distribution. At meanwhile, it can accurately describe the correlation of multi-variable distribution.

The name "copula" was chosen to emphasize how a copula "couples" a joint distribution to its marginal distributions. The concept applies to high dimensional processes, but for simplicity this study focus on the bivariate case in this paper.

The use of copulas therefore splits a complicated problem (finding a multivariate distribution) into two simpler tasks. The first task is to model the univariate marginal distributions and the second task is finding a copula that summarizes the dependence structure between them.

(1) The Sklar's theorem

Sklar's theorem implies that a joint distribution can be factored into two marginal distributions of the components and a copula describes the dependence between the components. With Sklar's theorem, one can first estimate suitable marginal distributions of the components of a multivariate system by any possible method, then link them together through an appropriate copula to form a joint distribution.

Suppose the marginal distribution function of a multi-dimensional distribution function H is that $F_1(x_1),...,F_n(x_n)$, then there exists a Copula function that satisfies:

$$H(x_1,...x_n) = C(F_1(x_1),...F_n(x_n)) \tag{2.17}$$

If $F_1(x_1),...,F_n(x_n)$ is continuous, then the Copula function is uniquely determined, and vice versa. From this theorem, it can be inference that this study can easily calculate the joint distribution of financial time series when we identified the marginal distribution a number of financial time series and selected a suitable Copula function, which is also the advantage of Copula functions in practical application.

(2) Rank correlation

Rank correlation reflects the monotonic dependence between variables, so it remains unchanged under the non-linear monotone transformation that has good statistical properties, and should be superior to a traditional linear correlation. The most representative rank correlation coefficients are Kendall.tau and Spearman.rho.

(i) Kendall.tau

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

$$\tau = p\{(x_1 - x_2)(y_1 - y_2) > 0\} - p\{(x_1 - x_2)(y_1 - y_2) < 0\} \quad (2.18)$$

Then, τ measures the degree of consist change in x and y. It can be proved:

$$\tau = 2P\{(x_1 - x_2)(y_1 - y_2) > 0\} - 1 \quad (2.19)$$

It can be seen that the τ is between [-1, 1], suppose the copula function of (x_1, y_1) is C (u, v), then Schwettzer and Wolff (1998) proved that the function τ is given by copula:

$$\tau = 4\iint_0^1 C(u,v)\,dC(u,v) - 1 \quad (2.20)$$

(ii) Spearmam.rho

Let (x, y) have the joint distribution H(x, y). Their corresponding marginal distribution is F_x and F_y, $x_0 \in x$, $y_0 \in y$ and $(x, y) \sim F(x)G(y)$, that is, x_0, y_0 are independent. Assume (x,y) and (x_0, y_0) are also independent, let

$$\rho = 3[P\{(x - x_0)(y - y_0) > 0\} - p\{(x - x_0)(y - y_0) < 0\}] \qquad (2.21)$$

When the Copula function C (u, v) is given, where u=F(x), v=G (y), Schwettzer and Wolff (1998) proved that ρ is given by the corresponding Copula function:

$$\rho = 12\iint_0^1 C(u,v)dC(u,v) - 3 \qquad (2.22)$$

(3) The dependence of upper tail and lower tail

Tail dependence (Joe 1997) describes the dependence of the tails of the bivariate distribution between left lower quadrant and right upper quadrant. It is the Copula of the two variables.

The dependence of tail is a kind of correlation of tail data in multi-dimensional distribution, which is a concept of link with extreme value theory. The dependence of tail between variables is the probability when the random variable

X increase or decrease significantly, the random variable Y has also undergone a substantial increase or decrease significantly.

For continuous random variables X, Y, the marginal distribution function is F (x) and G (y) respectively, then the correlation coefficient of the distribution curve in the upper tail and lower tail can be expressed as:

$$\lambda^{up} = \lim_{u \to 1} P\{Y > L^{-1}(u) \mid X > F^{-1}(u)\} = \lim_{u \to 1} \frac{1-2u+C(u,v)}{1-u} \quad (2.23)$$

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

$$\lambda^{low} = \lim_{u \to 1} P\{Y < L^{-1}(u) \mid X < F^{-1}(u)\} = \lim_{u \to 1} \frac{C(u,v)}{1-u} \quad (2.24)$$

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

(4) Pearson linear correlation

Pearson's correlation coefficient between two variables is defined as the covariance of the two variables divided by the product of their standard deviations

For a population:

Pearson's correlation coefficient when applied to a population is commonly represented by the Greek letter ρ (rho) and may be referred to as the population correlation coefficient or the population Pearson correlation coefficient. The formula for ρ is:

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu x)(Y - \mu y)]}{\sigma_X \sigma_Y} \tag{2.24}$$

For a sample:

Pearson's correlation coefficient when applied to a sample is commonly represented by the letter r and may be referred to as the sample correlation coefficient or the sample Pearson correlation coefficient. We can obtain a formula for r by substituting estimates of the covariances and variances based on a sample into the formula above. That formula for r is:

$$\gamma = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{2.25}$$

An equivalent expression gives the correlation coefficient as the mean of the products of the standard scores. Based on a sample of paired data (Xi, Yi), the sample Pearson correlation coefficient is

$$\gamma = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{s_X}\right)\left(\frac{Y_i - \overline{Y}}{s_Y}\right) \tag{2.26}$$

where $\frac{X_i - \overline{X}}{s_X}$, $\overline{X}$,and s_X are the standard score, sample mean, and sample standard deviation, respectively.

The absolute value of both the sample and population Pearson correlation coefficients are less than or equal to 1. Correlations equal to 1 or -1 correspond to data points lying exactly on a line (in the case of the sample correlation), or to a bivariate distribution entirely supported on a line (in the case of the population correlation). The Pearson correlation coefficient is symmetric: corr(X,Y) = corr(Y,X).

A key mathematical property of the Pearson correlation coefficient is that it is invariant (up to a sign) to separate changes in location and scale in the two variables. That is, this study may transform X to a + bX and transform Y to c + dY, where a, b, c, and d are constants, without changing the correlation coefficient (this fact holds for both the population and sample Pearson correlation coefficients). Note that more general linear transformations do change the correlation: see a later section for an application of this.

The Pearson correlation can be expressed in terms of uncentered moments. Since $\mu_X = E(X), \sigma_X^2 = E[(X - E(X))^2] = E(X^2) - E^2(X)$ and likewise for Y, and since

$$E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y) \tag{2.27}$$

the correlation can also be written as

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - (E(X))^2}\sqrt{E(Y^2) - (E(Y))^2}} \tag{2.28}$$

Alternative formulae for the sample Pearson correlation coefficient are also available:

$$\gamma_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{(n-1)s_x s} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}} \tag{2.29}$$

The above formula suggests a convenient single-pass algorithm for calculating sample correlations, but, depending on the numbers involved, it can sometimes be numerically unstable.

(5) Dependence Measures properties

The general properties of dependence measures can be explained by the 4 items properties shown below (Embrechts, Lindskog and McNeil (2003))

1. $\delta(X,Y)=\delta(Y,X)$

2. $-1\le\delta(X,Y)\le 1$

3. $\delta(X,Y)=1$ If X and Y are comonotonic; as well as $\delta(X,Y)=-1$ if X and Y are comontonic.

4. If T is exactly monotonic, then $\delta(T(X),Y)=\{\delta(X,Y), T = increa\sin g$ Or $-\delta(X,Y), T = decrea\sin g\}$

Normally, the Pearson linear correlation fits only the first two properties but the rank correlation measures Spearman's rho and Kendall's tau fits all of the 4 properties. Therefore, the copulas uses the Spearman's rho and Kendall's tau to calculate the dependence measures between X and Y which are random variables.

2.3 Literature reviews

In most empirical studies, which refer to the post-Bretton Woods floating exchange-rate regime, it is shown that the nature of the dependency can be adequately represented by the autoregressive conditional heteroskedastic (ARCH) model proposed by Engle (1982), or by its generalization represented by the GARCH model, suggested by Bollerslev (1986). This class of models is particularly suitable to describe the typical behaviour of financial time series, namely the fact that large (small) price changes tend to be followed by large (small) price changes of either sign.

However, this kind of dependency can be exploited only to improve interval or density forecasts, but not point forecasts. An improvement in point forecasts can be achieved by the GARCH in Mean (GARCH-M) model, where the conditional variance estimate enters as a regressor in the mean equation of the series.

Recently, many authors have also stressed the empirical relevance of non-linearity in mean for the exchange rate returns; Meese and Rose (1991), Krager and Kugler (1993), Peel and Speight (1994), Chappell et al. (1996) and Brooks (1997). However, the significant presence of mean non-linearities for the in-sample period only rarely has provided better out-of-sample forecasts compared with those obtained from a simple linear or a random walk model. Furthermore, the results are often sensitive to the length of the forecast horizon and to the metric adopted to measure the forecasting accuracy.

Diebold and Nason (1990) suggest four different reasons why non-linear models cannot provide better out-of-sample forecasts than the simpler linear model even when linearity is significantly rejected.

The result of this study is there are (1) non-linearities concern the even-ordered conditional moments and therefore are not useful for improving forecasts, (2) in-sample non-linearities are due to structural breaks or outliers which cannot be exploited to improve out-of-sample forecasts, (3) conditional means non-linearities are a feature of the DGP but are not large enough to offer better forecasts, and (4) non-linearities are present but they are captured by the wrong type of non-linear model.

Dumas (1992, 1994), Sercu et al. (1995) and McMillan (2005) have indicated that when conducting arbitrage in international commodity markets, the presence of transaction costs induces nominal exchange rates that deviate from the equilibrium to adjust in such a way that the rate process exhibits nonlinear properties.

Krager and Kugler (1993) estimate threshold autoregressive models for the returns of the French franc, Italian lire, Japanese yen, German mark and Swiss franc, all quoted against the US dollar (weekly observations for the period 1980.6–1990.1). Krager and Kugler find evidence of three different regimes, with the outer regimes exhibiting much higher estimated standard deviations than the inner regime, and argue that this finding is probably due to the central bank interventions aimed at avoiding excessive appreciation (first regime) or depreciation (third regime). The theoretical background of the empirical analysis presented in Krager and Kugler is the rational expectations monetary model with stochastic intervention rules proposed by Hsieh (1989). Thus, a three-regime autoregressive model is considered a good candidate to approximate Hsieh's model, which, according to Krager and Kugler, provides a better understanding of the managed floating exchange-rate regime than the target zone model (Krugman, 1991). In Hsieh's model central bank intervention is triggered by large exchange-rate changes, while in Krugman's model the intervention takes place when the level of the variable is in the vicinity of the bounds. In order to evaluate the relative importance of mean and variance non-linearities, GARCH models are also estimated for the variables listed above.

The result of this study is Krager and Kugler conclude that neither the threshold models nor the GARCH models prove successful in describing adequately the non-linearity present in the series.

Peel and Speight (1994) analyze the changes of the British pound exchange rate against the US dollar, the French franc and the Reichsmark for the interwar period (weekly data). Having found strong evidence of a generic form of non-linearity in all the series, the authors precede by estimating alternative non-linear models: GARCH, bilinear and threshold autoregressive models.

The result of this study is the forecasting performance of the models is evaluated only for a one-step-ahead horizon: the linear-ARCH models exhibit a lower MSFE compared to the bilinear models for all the series, but in the case of the pound/US dollar exchange rate the most accurate forecasts are provided by the threshold models.

Frankel and Rose(1995), Taylor(1995) and Sarno and Taylor(2002) shows that the persistent divergence from equilibrium causes linear PPP based fundamentals exchange rates models not to perform well in predicting or explaining future or past exchange rate movements. **Taylor et al. (2001) and Kilian and Taylor (2003)** argued that allowing for nonlinearities in real exchange rate adjustment is key both to detect mean reversion in the real exchange rate and to solve the PPP puzzle. **Haidar (2011b)** showed that certain measurements of currency valuation are misleading for economies whose markets are structurally different from the benchmark currency countries. MacDonald (1999, 2004) believes that some form of PPP does in fact hold at least as a long run relationship.

Taylor (1995) show a survey that panel unit root and long-run studies have reported evidence favorable to parity reversion, however, Rogoff (1996) and Obstfeld and Rogoff (2001) pointed out it is impossible to reconcile the high short-term volatility of real exchange rates with the slow rate at which shocks in the real exchange rate appear to die out in those studies.

Chappell et al. (1996) differs from those presented in Krager and Kugler (1993) and in Peel and Speight (1994) since it is focused on the forecasting performance of non-linear models fitted to the levels, rather than the changes, of some bilateral ERM exchange rates considered at daily frequency. It is important to stress that if the forecast assessment (for more than one step ahead) is carried out on the basis of criteria such as the MSFE, the choice of data transformations is not neutral as shown by Clements and Hendry (1993, 1995): evaluation in differences is penalizing relative to evaluation in levels. The issue of whether to evaluate the forecasting performance for the differences or the levels of the series is distinct from the issue of whether to estimate a model in the differences rather than the levels. According to Chappell et al. (1996), the inherent design of the ERM, based on the existence of a band in which the exchange rate is allowed to fluctuate without intervention by the central banks, could be the rationale for the presence of at least one threshold. Thus, the exchange rate follows a random walk process within the band but stationary autoregressive processes in the proximity of the ceiling or the floor such that the whole process exhibits mean-reverting features.

The result of this study is that the process is globally, but not locally, stationary. In contrast with most of the studies in which it is documented that the forecasting superiority of the non-linear models is often confined to the one-step-ahead horizon, the SETAR models estimated by Chappell et al. (1996) yield noticeable gains outperforming the random walk and the linear model at horizons as long as five and ten steps ahead.

Brooks (1996, 1997) analyses the daily British pound/US dollar exchange rate returns for the period 1974.1–1994.7. The main findings are that the non-linear models adopted, namely GARCH, SETAR and bilinear, produce forecasts only marginally more accurate than the ones obtained from a random walk model for all the horizons considered (up to 20 steps ahead). Moreover, on the basis of the Pesaran–Timmermann (1992) test, Brooks shows that the estimated models do not feature any market timing ability.

Dacco and Satchell (1999) and Clements and Smith (2001) argue that the alleged poor forecasting performance of non-linear models can also be due to the evaluation and measurement method adopted. On the basis of an extensive Monte Carlo study, Clements and Smith (2001), using the SETAR specifications of Krager and Kugler (1993) discussed above, show that whether the nonlinearities present in the data can be exploited to forecast better than a random walk depends both on how forecast accuracy is measured and on the state of nature.

The results of these studies are the evaluation of the whole forecast density may reveal gains to the non-linear models that are systematically masked if the comparison is carried out only in terms of MSFE.

Dacco and Satchell (1999) point out the predominance of the random walk model in forecasting an exchange rate is mostly based on MSFE measures. Therefore, they suggest that the method of evaluation has to be chosen according to the nature of the problem examined. Methods based on the profitability criterion should turn out to be more adequate in the case of financial variables. Tests for the percentage of correct sign predictions, such as that proposed by Pesaran and Timmermann(1992), are expected to be more informative in deciding whether to buy or sell foreign currencies.

Sarantis (1999) used the smooth transition autoregressive (STAR) family of models to test the existence of two distinct regimes of real exchange rates for 10 major industrial countries; evidence from this study shows that the rate process is cyclical in both regimes for most of the countries and that the STAR models outperform Hamilton's Markov regime-switching model in an out-of-sample forecasting context.

Arnaud Costinot and Thierry Roncalli (2000) consider the problem of modeling the dependence between financial markets. Show that this coefficient does not give a precise information on the dependence structure. Instead, they propose a conceptual framework based on copulas. **V. Durrleman, A. Nikeghbali & T. Roncalli(2000)** gave a few methods for the choice of copulas in financial modeling. **Wu Zhenxiang, Ye Wuyi and Miao Baiqi (2004)** using Archimedean Copula to analysis two assets portfolio. The least VaR portfolio of two assets portfolio can be found by selecting proper Copula. In the practice of foreign exchange markets, the least VaR portfolio of European dollar and Japanese yen is gotten. Also the sensitivity of VaR to the combination coefficients is given.

Liew et al. (2002) showed that the adjustment behavior of nominal rates more accurately fits the exponential smooth transition autoregressive (ESTAR) model that assumes a symmetric distribution of exchange rates. They also demonstrated that the re-parameterization of the logistic smooth transition autoregressive (LSTAR) model greatly improves the efficiency of coefficient estimation. **Liew et al. (2003)** have obtained using the nonlinear testing methodology developed by Luukkonen et al.(1988) indicated that the behaviors of real rates for 11 Asian countries fit the nonlinear STAR models rather than the linear AR models. **Baharumshah et al. (2003)** conducted a test for nonlinear unit roots, which indicated that the real rates for five ASEAN countries exhibit nonlinear mean reversion.

Imbs et al. (2003) found that mean reversion speed increases using TAR models and sectoral disaggregated price data. Following their argument, the further away the real exchange rate is from its long-run equilibrium, the stronger will be the forces driving it back towards equilibrium.

Schnatz et al. (2004) found that although the productivity differential was an important determinant of the real dollar–euro exchange rate, its ability to explain the real depreciation of the euro in the late nineties could be considered very limited. **Camarero, Ordóñez and Tamarit (2005)** estimated a long-run model for a synthetic preeuro– dollar exchange rate, finding that the main factor explaining the dynamic adjustment in the error correction model was the productivity differential. **Lothian**

and Taylor (2008) investigated the influence of productivity differentials on the equilibrium level of the pound–dollar and pound–franc real exchange rates. Although these authors found statistically significant evidence of the HBS effect for the pound–dollar real exchange rate, they failed to find any significant evidence of the HBS effect for the pound–franc real exchange rates.

Peel and Venetis (2005) found evidence that the ESTAR model can capture the smoothly symmetric adjustment process of nominal rates when real rates converge toward their long-run equilibrium levels. Marko (2005) found that a specific relationship between nominal rates and macroeconomic fundamentals exists in terms of adjusted consumer and import price indices. **Zhao Zhenquan (2011)** use the KPSS test based on ESTAR model, found that Chinese exchange rate fit the PPP theory.

Liu Wei (2008) according to analyzing real exchange rate theories and properties of RMB real exchange rate, a linear autoregressive model and two nonlinear regime switch model are selected as tool for the research. Estimating these models and fitting sample data lead to a conclusion that the smooth transition autoregressive model is the best one which describes the RMB real exchange rate behavior well.

The result of this study is that all outcomes indicate that RMB real exchange rate bears the characteristic of non-linear dynamic behavior and asymmetry.

Zhang Yuqin, Lin Guijun, Wang Shouyang (2011) focus on testing for nonlinearity of the daily exchange rate. Specifically, they use a discrete parametric modeling approach to compute an efficient test statistics for ASEAN-5 daily dollar exchange rates. To validate the test statistics, they use surrogate data testing method. The empirical results show that there exists nonlinear in ASEAN-5 daily dollar exchange rates.

Liu Qiongfang and Zhang Zongyi (2011) apply the Copulas theory to investigate the dependence structure between real estate and finance industries. Based on AIC and BIC minimum theories，the Gumbel Copula function shows that a correlation between these two markets exists in only upper tail for single parameter Copulas. However，the BB3 Copula has a higher correlation in the lower tail than the upper tail for a variety of parameters used in the Copula function. The GPD model needs to estimate the threshold values in order to exactly fit the margin distribution of Copula function. In addition，creating an investment portfolio of two different stocks won't necessary help reduce investment risk.

Pisit Leeahtam, Chukiat Chaiboonsri, Prasert Chaitip, Kanchana Chokethaworn, Songsak Sriboonchitta (2011) aim to forecast a single market bridging the development gap among members of ASEAN that find itself facing important financial opportunities.

The result of this study confirmed that the Autoregressive-linear model was suggested as appropriate model to forecast for Thailand's exchange rate and for Malaysia's exchange rate during the period of 2008-2011. And the relationship between these two countries' exchange rate return in percentage is not strong.

Table 2.1 Summary of literature reviews for Non-linear models

Authors	Method/Variable	Results
Diebold and Nason (1990)		Suggest four different reasons why non-linear models cannot provide better out-of-sample forecasts than the simpler linear model even when linearity is significantly rejected.
Krager and Kugler (1993)	French franc, Italian lire, Japanese yen, German mark and Swiss franc	Neither the threshold models nor the GARCH models prove successful in describing adequately the non-linearity present in the series.
Peel and Speight (1994)	British pound the French franc and the Reichsmark	The linear-ARCH models exhibit a lower MSFE compared to the bilinear models for all the series, in the case of the pound/US dollar exchange rate the most accurate forecasts are provided by the threshold models.
Frankel and Rose(1995), Taylor(1995) and Sarno and Taylor(2002		The persistent divergence from equilibrium causes linear PPP based fundamentals exchange rates models not to perform well in predicting or explaining future or past exchange rate movements.
Taylor (1995)	Panel unit root and long-run studies	Evidence favorable to parity reversion.
Chappell et al. (1996)	Forecasting performance of non-linear models	Process is globally, but not locally, stationary.

Source: From Summarized

Table 2.1 Summary of literature reviews for Non-linear models (Continue)

Authors	Method/Variable	Results
Brooks (1996, 1997)	Daily British pound/US dollar exchange rate returns for the period 1974.1–1994.7	Non-linear models adopted, namely GARCH, SETAR and bilinear, produce forecasts only marginally more accurate than the ones obtained from a random walk model for all the horizons considered.
Dacco and Satchell (1999) and Clements and Smith (2001)	Poor forecasting performance of non-linear models due to the evaluation and measurement method adopted.	Evaluation of the whole forecast density may reveal gains to the non-linear models that are systematically masked if the comparison is carried out only in terms of MSFE.
Dacco and Satchell (1999)		Predominance of the random walk model in forecasting an exchange rate is mostly based on MSFE measures.
Sarantis (1999)	Smooth transition autoregressive(STAR) family	STAR models outperform Hamilton's Markov regime-switching model in an out-of-sample forecasting context.
Taylor et al. (2001) and Kilian and Taylor (2003)		Allowing for nonlinearities in real exchange rate adjustment is the key both to detect mean reversion in the real exchange rate and to solve the PPP puzzle.
Liew et al. (2002)		The adjustment behavior of nominal rates more accurately fits the exponential smooth transition autoregressive (ESTAR) model
Imbs et al. (2003)		Mean reversion speed increases using TAR models and sectoral disaggregated price data.

Source: From Summarized

Table 2.1 Summary of literature reviews for Non-linear models (Continue)

Authors	Method/Variable	Results
Liew et al. (2003)	Nonlinear testing	Behaviors of real rates for 11 Asian countries fit the nonlinear STAR models rather than the linear AR models.
Baharumshah et al. (2003)	Nonlinear unit roots	Real rates for five ASEAN countries exhibit nonlinear mean reversion.
Peel and Venetis (2005)		ESTAR model can capture the smoothly symmetric adjustment process of nominal rates when real rates converge toward their long-run equilibrium levels.
Liu Wei (2008)	RMB exchange rate	Fitting sample data lead to a conclusion that the smooth transition autoregressive model is the best one which describes the RMB real exchange rate behavior well.
Zhao Zhenquan (2011)	KPSS test based on ESTAR model	Chinese exchange rate fit the PPP theory.
Zhang Yuqin, Lin Guijun, Wang Shouyang (2011)	ASEAN-5 daily exchange rate	There exists nonlinear in ASEAN-5 daily dollar exchange rates.
Pisit Leeahtam, Chukiat Chaiboonsri, Prasert ChaitipTIP, Kanchana Chokethaworn, Songsak Sriboonchitta (2011)	Linear, Non-linear and Copulas Approach	Autoregressive-linear model was suggested as appropriate model to forecast for Thailand’s exchange rate and for Malaysia’s exchange rate during the period of 2008-2011.And the relationship between these two countries’ exchange rate return in percentage is not strong.

Source: From Summarized

Table 2.2 Summary of some of the literature reviews for Copulas

Authors	Method/Variable	Results
Arnaud Costinot and Thierry Roncalli (2000)	Copula dependence structure	Show that this coefficient does not give precise information on the dependence structure. Instead, they propose a conceptual framework based on copulas.
Wu Zhenxiang, Ye Wuyi and Miao Baiqi(2004)	Archimedean Copula	In the practice of foreign exchange markets, the least VaR portfolio of European dollar and Japanese yen is gotten. Also the sensitivity of VaR to the combination coefficients is given.
Liu Qiongfang and Zhang Zongyi (2011)	Copula dependence structure	Gumbel Copula function shows that a correlation between these two markets exists in only upper tail. BB3 Copula has a higher correlation in the lower tail than the upper tail. The GPD model needs to estimate the threshold values in order to exactly fit the margin distribution of Copula function.Creating an investment portfolio of two different stocks won't necessary help reduce investment risk.

Source: From Summarized