Open economies review **17**: 235–251, 2006

Forecasting Performance of Exponential Smooth Transition Autoregressive Exchange Rate Models

AHMAD ZUBAIDI BAHARUMSHAH* zubaidi@putra.upm.edu.my
Department of Economics, Faculty of Economics and Management, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

VENUS KHIM-SEN LIEW
Labuan School of International Business and Finance, Labuan International Campus, Universiti Malaysia, Sabah P.O. Box 80594-87015 Federal Territory Labuan, Malaysia

Key words: autoregressive, smooth transition autoregressive, nonlinear time series, forecasting accuracy

JEL Classification Numbers: F31, F47

Abstract

This paper compares the forecasting performance of the Smooth Transition Autoregressive (STAR) model with the conventional linear Autoregressive (AR) and Simple Random Walk (SRW) models. The empirical analysis was conducted using quarterly data for the yen-based currencies of six major East Asian countries. We discovered strong evidence on nonlinear mean reversion in deviation from purchasing power parity (PPP). The results suggest that both the STAR and AR models outperform or at least match the performance of the SRW model. The results also show that the STAR model outperforms the AR model, its linear competitor in a 14-quarter forecast horizon. This finding is consistent with the emerging line of research that emphasizes the importance of allowing nonlinearity in the adjustment of exchange rate.

Since the establishment of the free float regime in March 1973, exchange rate forecasting has been an important research issue in exchange rate study. However, previous findings generally cannot negate over the fact that exchange rate models forecast no better than the random walk, the so-called "model of no change" (Meese and Rogoff, 1983a,b; Diebold and Nason, 1990; Meese and Rose, 1991; Lin and Chen, 1998; Kilian and Taylor, 2001). Several authors have argued that this forecast failure is due to the fact that exchange rate models ignore nonlinearity adjustments to their equilibrium values (e.g. Micheal, Nobay and Peel, 1997; Taylor and Peel, 1997; Sarno 2000; Coakley and Fuertes, 2001). In fact, the vast majority of studies on the behaviour of exchange rate rely on the assumption of linearity. The advancement in time

*To whom correspondence should be addressed.

series econometrics in the past two decades or so has made it possible to capture nonlinearities in macroeconomic variables such as gross national product (GNP) and the exchange rate. The Smooth Transition Autoregressive (STAR) model, for example, is well suited to capture the nonlinearities in the exchange rate process. The STAR model allows the variable under investigation to move within two different state spaces with a smooth transition process.[1] This nonlinear time series model offers an alternative approach to the modeling of economic variables that exhibit nonlinearities.

The STAR model is not widely used, and its forecasting performance is yet to be compared with alternative models. For example, Taylor and Peel (2000), Sarno (2000) and Baum, Barkoulas and Caglayan (2001) are among the first to demonstrate the usefulness of the STAR modeling in areas of exchange rate dynamics, but they did not carry out an evaluation of relative forecasting performance.[2] This paper intends to determine the applicability of the STAR model to the yen-based currencies of the two advanced Asian economies of Singapore and South Korea with per capita income of US$10,000 or more and the four emerging economies of the Association of the Southeast Asian Nations (ASEAN), namely Malaysia, Thailand, the Philippines and Indonesia, with a per capita income of less than US$ 4,000. All these ASEAN countries have not been included in the studies cited above. In addition, the present article extends the analysis to include the forecasting performance of the STAR models. To this end, we utilise the linear Autoregressive (AR) and the Simple Random Walk (SRW) models as the yardsticks of comparison.[3],[4] Thus, it is of great interest to investigate whether the STAR model has the potential to beat the naive random walk model in forecasting the exchange rates of South Korea and the founding members of the ASEAN countries.

This paper extends previous studies by evaluating the forecasting performance of the studied models using the commonly used Root Mean Squared Error (*RMSE*). This study differs from those reported in the earlier papers (see e.g., Sarno (2000) for 11 Middle East countries, Taylor and Peel (2000) for UK and Germany, and Baum et al. (2001) for 17 US trading partners) in two important ways. First, we examined the experience of the ASEAN economies and included the recent financial crisis in our forecasting horizon. This allowed us to examine whether the exchange rate follows a nonlinear adjustment process. Second and more importantly, we utilized formal statistics—the w-test due to Granger and Newbold (1986, pp. 278–280)—in the analysis to show the differences in the performance between the linear and nonlinear models.[5]

This study models the adjustment process of the deviations of each yen-based currency's movement from its fundamental equilibrium as determined by the purchasing power parity (PPP) hypothesis. Simply, the PPP hypothesis postulates that the nominal exchange rate is given by the ratio of the domestic and foreign price levels. It states that exchange rates should tend to equalize prices for identical goods in different countries. Recent studies based on careful application of time series econometrics methods are more supportive of the mean reverting behaviour of exchange rates (Azali, Habibullah and Baharumshah, 2001; Baum, Barkoulas and Caglayan, 2001).[6] Moreover, the stylized fact that emerges from recent literature is that exchange rate adjusts nonlinearly to its long-run PPP equilibrium (Baum, Barkoulas and Caglayan, 2001; Coakley and Fuertes, 2001; Sarno, 2000; Mahajan and Wagner, 1999, among others).

The present study, then, offers two distinctive results. The first is a confirmation of nonlinear adjustments of exchange rate to the long-run PPP values. The second is that the nonlinear STAR model outperformed the linear model in the medium-term horizon, particular the SRW. The rest of the paper is organized as follows. Sections 1 and 2 offer a brief review on the development of STAR models and a discussion on the data used in the analysis. In Section 3, we describe the linearity test and present the test results. The estimated models and the results of the model forecasts are presented in Section 4. Finally, the last section offers some concluding remarks.

1. The STAR models

The earlier version of the nonlinear Smooth Transition Threshold Autoregressive or just Smooth Transition Autoregressive (STAR) model could be traced back to the Threshold Autoregressive (TAR) model initially introduced by Tong in 1977 (see Tong and Lim, 1980). The TAR model assumes that a variable has different behaviours within different regimes.[7] An example of the TAR model is the Self-Excited TAR or SETAR model, which assumes that a variable (say, exchange rate), y_t is a linear autoregression within a regime, but it may move between regimes depending on the value taken by a lag of y_t, say y_{t-d}, where d is the delay parameter. For a two-regime case $(q = 2)$ where y_t follows an AR (p_1) process in one regime and AR (p_2) process in the other, the SETAR (2;p_1, p_2) representation of y_t can be written compactly as (see Clement and Smith, 1997):

$$y_t = \beta_0 + \sum_{i=1}^{p_1} \beta_i y_{t-i} + \varepsilon_{1t} + I_{t-d}(r)\left[\beta_0^* + \sum_{i=1}^{p_2} \beta_i^* y_{t-i} + \varepsilon_{2t} - \varepsilon_{1t}\right] \tag{1}$$

where $I_t(r) = 1$ if $y_t > r$ and 0 otherwise is the threshold. For $h = 1, 2$, $\varepsilon_{ht} \sim G$ (0, σ_h^2) where G(·) may be a Gaussian distribution but this is not necessarily the case. β_i for $i = 1, \ldots, p_1$ and β_i^* for $i = 1, \ldots, p_2$ are parameters to be estimated.

The introduction of nonlinear time series models such as the SETAR model is motivated by the fact that the linear time series model should give place to a much wider class of models if we are to gain more understanding into more complicated phenomena such as limit cycles, time irreversibility, amplitude-frequency dependency and jump resonance (see for example, Tong and Lim, 1980). Since its introduction, few attempts have been made in the applying and validating of the SETAR model, and hence the usefulness of the model in empirical work has yet to be determined. For instance, the article by Diebold and Nason (1990) point out that there is no guarantee that the SETAR model will perform better than the linear AR model. A similar view is expressed in Clements and Smith (1997), where they note that the rejection of a null of linearity in favour of nonlinearity does not guarantee that the prediction based on the SETAR model will outperform the AR models.

The deficiency in the SETAR is deemed to be due to the unrealistic fixed threshold in the model. The fixed threshold of the SETAR model was later replaced with a smooth function and thus led to the formation of the STAR model in the early 1990s. The STAR model allows the variable under study to alternate between two different regimes with a smooth transition function between these regimes so that there can be a continuum of states between extreme regimes. The STAR methodology is generally preferred because regime change is smooth rather than discrete (as in TAR) because of heterogeneity behavior of economic agents (see De Grauwe, Dewachter and Embrechts, 1993; Sarantis, 1999). In addition, movement of exchange rates are based on price indices involving a range of goods with different costs of arbitrages (Taylor and Peel, 2000). The STAR representation is given by Teräsvirta and Anderson (1993) as follows:

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + \left[\beta_0^* + \sum_{i=1}^{p} \beta_i^* y_{t-i} \right] F(y_{t-d}) + \varepsilon_t \tag{2}$$

where y_t is mean-corrected, β_0, and β_0^* are constants, β_i and β_i^*, $i = 1, \ldots, p$ are autoregressive parameters, $F(\cdot)$ is the transition function depending on the lagged level, y_{t-d} where d is known as the delay length or delay parameter, and ε_t is a white noise with zero mean and constant variance σ_ε^2.

For the application of the STAR model, Granger and Teräsvirta (1993) have proposed the exponential function as one of the plausible transition functions, thus yielding the exponential STAR or ESTAR model. The exponential function is defined as

$$F(y_{t-d}) = 1 - exp\left[-\gamma^2 (y_{t-d} - u)^2 / \hat{\sigma}_{y_t}^2 \right] \tag{3}$$

where the transition parameter γ^2 is standardized by $\hat{\sigma}_{y_t}^2$, the estimated variance of y_t. u is the equilibrium or threshold value of the mean corrected y_t series and hence E(u) = 0.

Note that the speed of transition between the two regimes is positively related to the value of the transition parameter γ^2. In other words, higher values of γ^2 imply a much faster speed of transition. Taylor and Peel (2000) used a version of the transition function, $F(\cdot)$ with $\hat{\sigma}_{y_t}^2 = 1$. Nevertheless, Granger and Teräsvirta (1993, p. 124) have argued that scaling the exponential term by the sample variance speeds up the convergence and improves the stability of the nonlinear least squares estimation algorithm. It also makes it possible to compare estimates of the transition parameter across equations.

The exponential transition function is bounded between zero and one. Judging from Eq. (3), when y_{t-d} equals its equilibrium value u or when γ^2/

$\hat{\sigma}^2_{y_t}$ goes to zero, $F(\cdot) = 0$ and Eq. (2) reverts to a standard linear AR(p) representation:

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + \varepsilon_t \tag{4}$$

In such a case, the conventional restriction of $\sum_{i=1}^{p} \beta_i < 1$ applies so that y_t is mean-reverting. For extreme deviations from the fundamental equilibrium, $F(\cdot) = 1$ (when $\gamma^2/\hat{\sigma}^2_{y_t}$ approaches infinity), and Eq. (2) becomes a two-regime AR (p) model:

$$y_t = (\beta_0 + \beta_0^*) + \sum_{i=1}^{p} (\beta_i + \beta_i^*) y_{t-i} + \varepsilon_t \tag{5}$$

If Eq. (5) is the correct specification, it is expected that $|\sum_{i=1}^{p} \beta_i| \leq 1$ such that y_t may exhibit local unit root behaviour but the requirement for global stability is that $|\sum_{i=1}^{p} (\beta_i + \beta_i^*)| < 1$ must be met.

The exponential function $F(\cdot)$ allows a smooth transition between regimes and symmetric adjustment for deviations above and below the fundamental or equilibrium value. This function is appropriate for the modeling of exchange rate as it has a number of attractive properties. For instance, it can capture the symmetrical response to positive and negative deviations from its fundamental equilibrium (see for example, Baum, Barkoulas and Caglayan 2001) by its inverse-bell shaped distribution around zero. Another important feature of the model worth mentioning is that it nests the linear regression [AR(p)] model. Thus it allows us to test for linearity prior to applying nonlinear models.

2. Preliminary data analysis

The data used in this paper are end-of-quarter nominal bilateral exchange rates for the yen-based six Asian currencies, namely the Indonesia rupiah (INR/JPY), the Korea won (KRW/JPY), the Malaysian ringgit (MYR/JPY), the Philippines peso (PHP/JPY), the Singapore dollar (SGD/JPY) and the Thai baht (THB/JPY) as well as relative prices (P_ts), which are constructed as the ratio of the consumer price indices (CPIs) of the six Asian countries to CPI of Japan. The data are mainly from the International Monetary Fund's *International Financial Statistics* (IMF/IFS), comprising seasonally unadjusted observations. Our data consists of quarterly time-series observations from 1980: Q1 through 2003: Q4. The full sample period is divided into two portions. The first sub-period, which starts from 1980: Q1 and ends in 1997: Q2 is used for the model estimation purpose while the remaining observations are kept for assessing the out-of-sample forecast performance of the studied models. We note that long spans of high frequency data (e.g. CPI) do not

Table 1. Unit root tests results.

	Intercept Without Trend				Intercept With Trend			
Country	X	△ X	P	△ P	X	△ X	P	△ P
			Augmented Dickey–Fuller Test					
Indonesia	–1.467	–5.600*	–0.702	–4.464*	–2.491	–5.617*	–1.077	–5.167*
Korea	–0.317	–6.730*	1.347	–5.102*	–2.724	–6.733*	–0.802	–5.588*
Malaysia	–0.362	–4.958*	1.989	–5.537*	–2.696	–4.953*	–0.040	–6.260*
Philippines	0.448	–3.988*	–2.042	–3.680*	–1.527	–4.059*	–2.503	–4.112*
Singapore	–0.211	–3.712*	0.054	–5.551*	–0.575	–3.989*	–1.172	–5.744*
Thailand	0.179	–3.712*	1.237	–3.541*	–2.489	–6.482*	–0.985	–3.920*
			Phillip–Perron Test					
Indonesia	–1.631	–8.261*	1.684	–4.758*	–2.768	–8.265*	–0.371	–5.180*
Korea	–0.484	–14.39*	1.124	–10.38*	–4.445*	–14.43*	–0.710	–10.59*
Malaysia	–0.528	–8.653*	3.726	–11.67*	–2.987	–8.653*	0.233	–12.86*
Philippines	0.344	–8.320*	–2.504	–5.276*	–1.949	–8.373*	–1.701	–5.887*
Singapore	0.240	–5.811*	0.991	–11.18*	–0.832	–5.811*	–0.842	–11.43*
Thailand	–0.576	–11.10*	–1.360	–9.811*	–3.107	–11.10*	–0.757	–10.13*

Notes: X and P denote exchange rate and relative price respectively. Variable with △ in front means its first difference. Test-statistics with * imply reject null hypothesis of unit-root at 1% significance level.

exist for most of the countries under investigation. Shiller and Perron (1985) argued that the span of the data set is far more important than the number of observation per se. Our choice of quarterly data should make our analysis of wider interest since quarterly data is available for a wider range of countries than monthly or daily data.

To examine whether each of these exchange rates exhibits mean reverting behaviour to its long-run PPP equilibrium, we tested for the cointegrating relationship between the pairwise exchange rate and relative price data series using the Johansen and Juselius (1990) procedures. However, prior to any cointegration tests, the series involved should be tested for stationarity and order of integration. This is important as only variables of the same order of integration may provide a meaningful relationship in the Johansen framework. The commonly used Augmented Dickey–Fuller (ADF) and non-parametric Phillips–Perron (PP) unit root tests were employed for this purpose. The results of the unit root tests as summarized in Table 1 convincingly suggest that both exchange rates and relative prices are first difference stationary, which implies they are all integrated of the same order, that is, I(1). These results generally hold whether it is with trend or without trend.

Next, we proceeded to investigate whether or not the long-run PPP holds for the currencies selected for this study. To this end, we applied the Johansen and Juselius (1990) multivariate cointegration techniques to test for cointegration between exchange rates and relative prices.[8] Results of the

Table 2. Johansen and Juselius cointegration test results.

Country	Optimal Lag	Likelihood Ratio of Eigenvalue $r=0$	Likelihood Ratio of Eigenvalue $r\leq 1$
Indonesia	12	20.080*	6.685
Korea	12	25.629*	9.072
Malaysia	10	24.369*	5.061
Philippines	12	21.678*	0.039
Singapore	12	24.559*	2.817
Thailand	11	23.884*	9.080
		Critical Values	
	5%	19.90	9.24
	1%	24.60	12.97

Notes: r denotes the hypothesized number of cointegrating relationships. Optimum lag-length is determined by the Akaike Information Criterion (AIC). *Denotes rejection of hypothesis at 5% significance level or better.

Johansen-Juselius trace test are depicted in Table 2. The test results provide strong evidence that all exchange rates and their corresponding relative prices are cointegrated at standard significance levels. We also performed the maximum eigenvalue test for all the currencies. A similar conclusion can be derived using the maximum eigenvalue test (but not reported). These findings reveal that the bilateral rates exhibit mean reverting behaviour to their long-run PPP equilibria. In general, the results obtained so far are consistent with the evidence reported in Baharumshah and Ariff (1997) and Azali, Habibullah and Baharumshah (2001)[9].

This finding enables us to estimate the equilibrium values of the six Asian exchange rates based on the PPP hypothesis. Deviations of each rate from its equilibrium (z_t) can then be deduced by subtracting its observed values from the estimated equilibrium values. Note that the nature of the adjustment process of these deviations towards the equilibrium position is not known yet. To determine the linearity (or nonlinearity) of this adjustment process, we employed the linearity tests against the STAR models. The empirics of this issue are taken up in the next section.

3. Linearity tests

The minimum requirement for the estimation of STAR models is to reject the linearity of the variable under study (Tong and Lim, 1980). Various linearity tests have been developed based on the idea of testing the null hypothesis that all β^* s in Eq. (2) are simultaneously zero, against the alternative hypothesis that at least one β^* is not zero. Notice that failing to reject the null hypothesis, Eq. (2) would simply reduce to the linear AR (p) model. By the same token, rejection of the null hypothesis implies the presence of nonlinearity in favour of STAR (p) model. As the properties of the transition parameter (γ^2), the coefficients of nonlinear terms (β^*s) and the mean value (u) of the variable under estimation are not identified under the null hypothesis, linearity is tested in the context of the auxiliary model instead of the original STAR specification as in Eq. (2). Theoretical issues on linearity tests against STAR models are found in Luukkonen, Saikkonen and Teräsvirta (1998), Saikkonen and Luukkonen (1998), Teräsvirta and Anderson (1993), Teräsvirta (1994) and Eirtheim and Teräsvirta (1996). Interested readers may refer to these articles for more detailed discussion on the tests.

The present study only highlights a specification of the linearity test with alternative hypothesis in favour of the ESTAR model, a variant of STAR model relevant to the exchange rate model. This specification as proposed by Teräsvirta (1994), is based on the following auxiliary regression:

$$z_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i z_{t-i} + \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_{ij}^* z_{t-i} z_{t-j} + \sum_{i=1}^{p} \sum_{j=1}^{p} \tau_{ij}^* z_{t-i} z_{t-j}^2 + \omega_t \tag{6}$$

where z_t is the deviation of exchange rate from its PPP value.

The null hypothesis to be tested is that:

$$H_0 : \alpha_{ij}^* = \tau_{ij}^* = 0; i, j = 1, \ldots, p \tag{7}$$

In practice, Teräsvirta's Lagrange Multiplier (LM) linearity tests can be performed by following these steps:

(1) Regress z_t on $\{1, z_{t-j};\ j = 1, \ldots, p\}$. Obtain the estimated residuals $\hat{\varepsilon}_t$ and compute the residual sum of squares, $SSR_0 = \sum_{t=1}^{T} \hat{\varepsilon}_t{}^2$, where T is the sample size;
(2) Regress $\hat{\varepsilon}_t$ on $\{1, z_{t-i}, z_{t-i}z_{t-j}, z_{t-i}z_{t-j}^2\}$. Obtain the estimated residuals $\hat{\omega}_t$ and compute the residual sum of squares $SSR = \sum_{t=1}^{T} \hat{\omega}_t^2$;
(3) Compute the test statistic

$$\mathrm{LM} = \frac{(SSR_0 - SSR)}{\hat{\sigma}_\varepsilon^2} \text{where } \hat{\sigma}_\varepsilon^2 \text{ is the estimated variance of } \hat{\varepsilon}_t \tag{8}$$

Under the null hypothesis the LM statistic is asymptotically distributed as a chi-square (χ^2) with $2p$ degrees of freedom, given that the delay parameter d is known. If d were unknown, the degrees of freedom would be as large as $0.5p(p+1) + 2p^2$. Thus, prior knowledge about d is very useful in testing linearity against ESTAR models.

Note that the optimum lag length p of the above auxiliary regression is unknown and it has to be determined from the data. To this end, model selection criteria such as Final Prediction Error (FPE), Schwarz Information

Criterion (SIC) and Akaike Information Criterion (AIC) are commonly used for this purpose. However, these criteria are of course not without any shortcomings; see, for instance, Teräsvirta and Anderson (1993) for more on this issue. Briefly, Teräsvirta and Anderson (1993) pointed out that neglecting the autocorrelation structure of the residuals may lead to false rejection of the linearity hypothesis in favour of the nonlinearities alternative[10]. In this study, the optimal lag length p of linear AR (p) model is selected based on AIC and the lag length selected is sufficient to eliminate serial correlation.

Having selected p, d needs to be determined. In order to specify d, a linearity test is carried out for the range of values considered appropriate, in this case, $1 \leq d \leq 12$. If the linearity is rejected for more than one value of d, then d is determined such that $\hat{d} = \text{argmin } p(\text{LM})$ for $1 \leq d \leq 12$ where p(LM) is the marginal significance value of the LM test. The argument behind this rule of minimising the marginal significance value is that the test has maximum power if d is chosen correctly, whereas an incorrect choice of d weakens the power of the test. Additionally, the Ljung-Box portmanteau Q test is employed to confirm the absence of serial correlation up to 20 lags.

Results of the linearity tests are summarized in Table 3. It is clear from the table that linearity is rejected at 5% significance level or better for all six currencies, and hence, in favour of the ESTAR models. The Q statistic suggests that the combination of p and d selected for the models yield residuals that are free from autocorrelation problems up to 20 lags. The results in Table 3 for the ASEAN currencies are representative of the kind of results one finds in the recent literature on exchange rate; see Coakley and Fuertes (2001), Sarno (2000), Micheal, Nobay and Peel (1997) and Taylor and Peel (1997), among others. The finding is appealing as it reveals that exchange rate adjusts differently to positive and negative shocks, and to large or small deviations from its long-run equilibrium value.

4. Forecasting performance of estimated models

In this section we examine the concern about exchange rates being too volatile to be explained by structural models using quarterly data. Meese and Rogoff (1983a) showed that a random walk model of the exchange rate performed better in forecasting than any structural model. To date, this result still holds true, and as noted in Devereux (1997, p. 774), "this result has not been substantially overturned". The ESTAR models are estimated for each of the exchange rate deviations, z_t. The results obtained from these models are reported in Table 4.

Several features for the estimated unrestricted model are noteworthy here: first, the nonlinear parameters (β_1^* and γ) of the unrestricted ESTAR (2) model are statistically significant at 10% or higher level. Second, the residual variance ratios of all ESTAR models to their corresponding linear AR models are smaller than one, indicating that the former have much smaller variances.

This implies that the nonlinear models have the ability to produce smaller forecast errors than the linear models. Third, these models passed a battery of diagnostic tests at conventional significance levels and there is no indication that the model is misspecified. Fourth, the high adjusted R^2 indicates that the explanatory power of these nonlinear models on the adjustment of deviations is fairly high. Fifth, the absolute sum of linear parameters, $|\sum_{i=1}^{p} \beta_i| > 1$ for all ESTAR models (excluding INR/JPY and KRW/JPY), suggesting that with the exception of INR/JPY and KRW/JPY, z_t exhibits unit root behaviour and therefore the linear AR (p) model itself is an inadequate representation of z_t. On the other hand, $|\sum_{i=1}^{p} (\beta_i + \beta_i^*)| < 1$ in all cases implies that the requirement for global stability is met. This confirms that z_t is mean-reverting in the nonlinear specification (see Baum, Barkoulas and Caglayan 2001; Taylor and Peel 2000) in general. Taken together these findings support our contention that the nonlinear model is an appropriate representation of z_t, the deviations of six major East Asian exchange rates. The estimated exponential transition functions, F(•) are depicted in Figure 1.

The in-sample and out-sample forecasting performances of these estimated forecasting models are evaluated using AR and SRW models as benchmarks, based on the commonly used *RMSE*. The in-sample forecasts are evaluated for the period 1981: Q2 to 1997: Q2 with a total of 65 quarterly observations. The estimated forecasting models are utilized to generate, first, a total of 14 out-sample quarterly forecasts over the period 1997: Q3 to 2000: Q4, and second, to generate 26 quarterly ($6\frac{1}{2}$ years) forecast over the period 1997: Q3 to 2004: Q4. The overall forecasting performances are summarized in Table 5. The *RMSE* ratio reported in Table 5 is obtained by dividing the *RMSE* of the forecasting model by the *RMSE* of the benchmark model. The statistical w test by Granger and Newbold (1986, pp. 278–280) is also included

Table 3. Linearity tests results based on LM test

Exchange Rates	Lag Length, p^a	Delay Parameter, d^b	p-value of LM Test^c	p-value of Q(20) Statistic^d
INR/JPY	4	1	0.047	0.232
KRW/JPY	5	1	0.003	0.738
MYR/JPY	2	2	0.025	0.445
PHP/JPY	5	11	0.000	0.725
SGD/JPY	4	12	0.000	0.826
THB/JPY	3	1	0.001	0.995

Notes: [a] The optimal lag length p of linear AR (p) model is determined by the AIC.

[b,c] Optimal delay parameter d is determined by $\hat{d} = \arg\min_{1 \le d \le 12} p$ (LM) where $p(\cdot)$ denotes the p-value of the implied test statistic for the null hypothesis of H_0: Linear model is correct. Rejection of H_0 implies nonlinearity in favour of the ESTAR model.

[d] Ljung-Box portmanteau Q test, denoted as Q(20), is employed to confirm the absence of serial correlation up to 20 lags.

Table 4. Estimated ESTAR models.

Parameters	INR/JPY	KRW/JPY	MYR/JPY	PHP/JPY	SGD/JPY	THB/JPY
			Estimated Values (Standard Errors)			
β_0	–11.981 (12.05)	–0.006 (0.05)	–0.031 (0.19)	0.002 (0.00)	–0.001 (0.01)	–0.001 (0.00)
β_1	–0.977 (1.10)	2.137 (0.89)*	1.799 (0.33)*	0.902 (0.23)*	–4007.748 (3395.86)	2.155 (0.23)*
β_2	0.342 (0.77)	–2.139 (1.24)	–0.340 (0.23)	0.171 (0.34)	–15233.223 (12891.62)	0.010 (0.01)
β_3	0.091 (0.86)	0.948 (0.66)	–	–0.404 (0.40)	–2578.732 (2177.45)	0.010 (0.01)
β_4	0.297 (0.63)	–0.026 (1.66)	–	0.590 (0.43)	–388.558 (334.36)	–
β_5	–	–0.205 (0.48)	–	0.130 (0.49)	–	–
β_1^*	1.832 (1.10)*	–1.307 (0.90)	–1.170(0.40)*	0.003 (0.38)	4008.870 (3395.89)	–1.527 (0.76)
β_2^*	–0.276 (0.82)	2.244 (1.25)*	–0.349(0.50)	–0.801 (0.54)*	15232.908 (12891.63)	–0.156 (0.69)
β_3^*	–0.030 (0.92)	–0.803 (0.70)	–	1.283 (0.58)	2579.160 (2177.37)	–0.004 (0.05)
β_4^*	–0.314 (0.67)	0.239 (1.66)	–	–0.561 (0.55)	388.265 (334.41)	–
β_5^*	–	–0.265 (0.53)	–	–0.621 (0.52)	–	–
γ	3.535 (1.08)*	18.260 (2.48)*	1.364(0.25)*	1.693 (0.68)*	29.218 (4.07)*	15.549 (1.11)*
$\hat{\sigma}_y^2$	42871.850	0.673	0.171	0.001	0.062	0.009
			Diagnostic Tests [Marginal Significance Values]			
$\hat{\sigma}_{ESTAR}^2$	7982.464	0.129	0.002	0.002	0.004	0.001
$\hat{\sigma}_{ESTAR}^2/\hat{\sigma}_{AR}^2$	0.908	0.992	0.769	0.974	0.962	0.779
Q (20)	14.023 [0.83]	6.014 [0.99]	17.522[0.62]	16.099 [0.71]	12.538 [0.90]	9.132 [0.98]
WHITE	1.599 [0.97]	4.044 [0.91]	5.276[0.81]	12.479 [0.19]	6.330 [0.71]	9.382 [0.40]
ARCH (4)	0.572 [0.97]	1.359 [0.85]	1.563[0.82]	5.732 [0.22]	1.861 [0.76]	3.142 [0.53]
GARCH (1, 1)	0.070 [0.97]	0.565 [0.75]	0.546[0.76]	3.280 [0.19]	0.460 [0.79]	0.543 [0.76]
Adjusted R^2	0.817	0.822	0.882	0.751	0.923	0.879

Notes: Asterisk (*) stands for significant at 10% level or better.

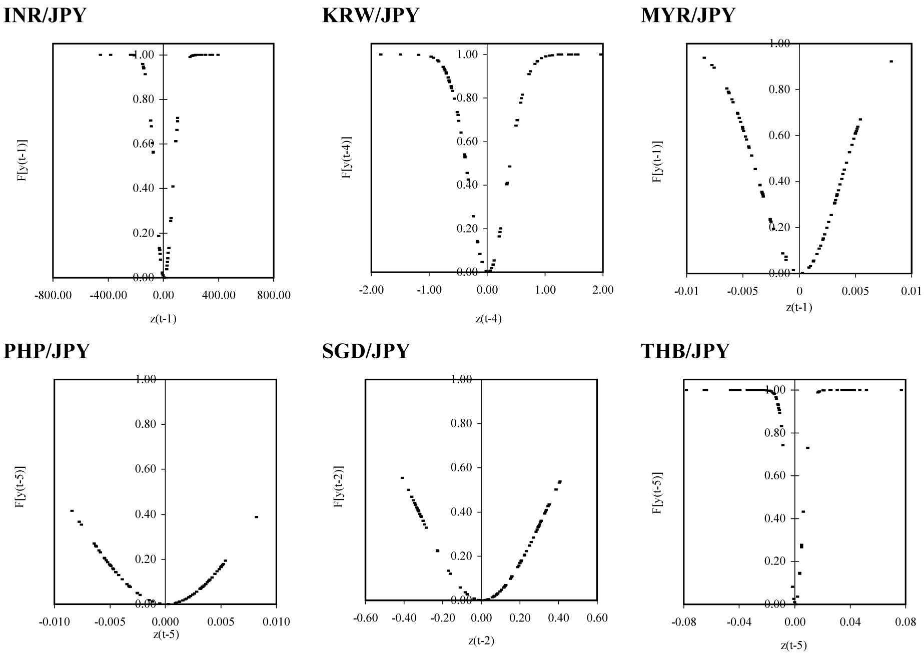

Figure 1. Estimated exponential transition functions.

to show the statistical significance of the *RMSE* ratio. Briefly, the w test is based on the correlation coefficient (r) of the sums and differences of the two models' forecast errors, where $r = 0$ under the null hypothesis of "equal accuracy" for the two competing models. The test is carried out assuming the one-sided alternative that $r > 0$, which implies one model has significantly outpredicted the other.[11]

Table 5 shows that in the case of in-sample forecasts, the ESTAR (p) models perform better than the AR (p) models for all exchange rates except INR/JPY. This conclusion is drawn from the root mean square error (*RMSE*). We find that the former is smaller than the latter, thereby yielding *RMSE* ratios of values less than one. The statistical significance of this finding is reinforced by the results of the w test, which demonstrates that the *RMSE* ratios are significant at 1% in all cases. Compared with random walk forecasts, the results are somewhat mixed. In particular, the ESTAR models are shown to be superior to the random walk for KRW/JPY, PHP/JPY and SGD/JPY by the *RMSE* ratio, whereas the opposite is true for INR/JPY, MYR/JPY and THB/JPY rate.

Turning to the out-of-sample forecasts, the ESTAR models are found to have significantly outperformed both the AR and the SRW models neatly on the basis of the *RMSE* ratio and the w test for the 14-quarter horizon. These results are reported in Table 5. Next, we extend the forecasting horizon to the period ended in 2003: Q4. By and large, we observe that the long-

Table 5. Overall performances of forecasting models.

	In-sample		Out-sample			
	(n = 65 quarters) 1980Q1 to 1997Q2		(n = 14 quarters) 1997Q3 to 2000Q4		(n = quarters) 1997Q3 to 2003Q4	
Accuracy criteria	*RMSE* Ratio	w-Test	*RMSE* Ratio	w-Test	*RMSE* Ratio	w-Test
			Linear AR model as benchmark			
INR/JPY	1.669	0.553^{c}	0.951	–0.599^{c}	0.998	–0.090^{a}
KRW/JPY	0.951	0.043^{c}	0.629	–1.856^{c}	0.869	–0.127^{c}
MYR/JPY	0.824	–1.054^{c}	0.880	–0.627^{c}	0.909	–0.367^{c}
PHP/JPY	0.402	–1.170^{c}	0.653	–0.539^{c}	0.808	0.153^{b}
SGD/JPY	0.956	–0.361^{c}	0.713	–0.677^{c}	0.939	0.089^{a}
THB/JPY	0.723	–0.586^{c}	0.739	–0.727^{c}	0.792	0.293^{c}
			Random walk model as benchmark			
INR/JPY	1.968	0.237^{c}	0.927	–0.242^{c}	1.013	–0.481^{c}
KRW/JPY	0.823	0.012	0.702	–1.880^{c}	0.968	0.041
MYR/JPY	1.205	0.942^{c}	0.858	–0.478^{c}	1.256	0.281^{c}
PHP/JPY	0.951	–0.140^{c}	0.971	–0.731^{c}	1.134	1.104^{c}
SGD/JPY	0.917	–0.648^{c}	0.974	–0.060	1.062	–0.295^{c}
THB/JPY	1.630	1.945^{c}	0.869	–0.386^{c}	0.993	0.028
			Critical Values of *w*-Test			
			1%		5%	10%
In-sample	$w \sim$ N(0, 1/62)		0.033		0.021	0.014
Out-sample	$w \sim$ N(0, 1/23)		0.101		0.072	0.056
	$w \sim$ N(0, 1/11)		0.187		0.117	0.077

Notes: w tests the null hypothesis of "equal accuracy" against one-sided alternative of "one model has significantly outpredicted the other" on the basis *RMSE* ratio. Superscripts ^{a,b} and c denote significant at 10, 5 and 1% level or better, respectively.

term forecast performance of the model deteriorates (see Table 5, column 6). Specifically, the ESTAR model still maintains its superiority over the AR model in all cases. However, the SRW model predicts better than the ESTAR in five out of six cases. We note that the ESTAR model yields superior forecasts over the simple random walk model for the case of South Korea and Thailand but they are not statistically significant based on the w-test. To sum up, this article has provided empirical evidence that the ESTAR models predict better than the linear AR and random walk models on the basis of *RMSE* ratio as well as the statistical significance test in terms of medium out-of-sample forecasts (14-quarter).

5. Conclusion

The empirical performance of exchange rate models has been frequently criticized in recent years. These critiques come from prior studies that have found that exchange rate models poorly predict in-sample forecasts, not to mention the out-of-sample periods. In the case of the East Asian countries, central banks often intervene in the foreign exchange market in an effort either to attenuate or to amplify variations in the exchange rate. This factor explained partly the poor out-of-sample predictions found in past studies.

In this article, the Teräsvirta's (1994) LM-type linearity test was used and the statistical results indicate that a nonlinear model can characterize all the currencies in the sample countries. Importantly, we demonstrate formally that the adjustment of the six major East Asian currencies to their long-run equilibria follows a nonlinearity path. We also show that the adjustment of these exchange rates is in fact predictable based on the nonlinear STAR model. Specifically, the out-of-sample forecasting performance of the STAR model significantly outperforms both the linear autoregressive (AR) and the random walk models in the medium-term. We note that the existing literature on the major currencies have pointed out that exchange rate is not predictable when forecast horizon is short based on structural models. As expected, the result fails to hold when the forecasting horizon is extended to 26 quarters. This finding is in sharp contrast to the earlier works of Sarantis (1999) and Kilian and Taylor (2001) that concluded that there is no clear gain on the ESTAR exchange rate forecasts (for industrialized countries) over the random walk forecasts in terms of forecast accuracy. This suggests that there is a systematic predictable component in the movement of the six major East Asian nominal exchange rates at least in the medium-term. One major implication of this finding is that exchange rates forecasters who are more inclined to the chartists' methods instead of the economic fundamental models (Taylor and Allen 1992) may resort to the nonlinear fundamental models as complementary tools.

Acknowledgment

We are grateful for the financial support for MOSTEC [Project no: S4284]. We acknowledge the helpful suggestions and comments of the anonymous referees and the editor of this journal. The authors would like to thank the participants of Asia Pacific Economics & Business (APEB) Conference 2002 in Sarawak, Malaysia for their constructive comments. The standard disclaimer applies.

Notes

1. Bacon and Watts (1971) were the first to introduce the STAR model but it was Chan and Tong (1986) and Granger and Teräsvirta (1993) who popularized the model.
2. With the exceptions of Sarantis (1999) and Kilian and Peel (2001).
3. These benchmarks are chosen for two reasons: First, the linear AR model is easily nested in the nonlinear STAR model and it is interesting to find out whether the latter yields more accurate forecasts than the former. In view of the fact that the answer to this issue is important, this study contrasts the forecast performance between these two competing nonlinear STAR and linear AR models directly. Second, the outcome of past studies indicate

that exchange rate models rarely beat the SRW model; see, *inter alia*, Meese and Rogoff (1983a, 1983b), Diebold and Nason (1990), Meese and Rose (1991) and Lin and Chen (1998).

4. Sarantis (1999) finds that there is not much to choose between the STAR and linear models based on out-sample forecasting performance. Nonetheless, STAR models do outperform the Markov regime-switching model, its major nonlinear competitor. It is noteworthy that the current study differs from Sarantis (1999) in two ways. First, this study deals with nominal exchange rates of major East Asian countries, whereas the latter forecasts the real effective exchange rates of the major industrial countries. Second, we employ an additional statistical test to show the robustness of our comparison based on *RMSE*.
5. Mizrach (1995) argues that heuristic approaches such as *RMSE* in forecast comparison may yield misleading inference. Thus, it is important to carry out statistical tests complementary to any ratio analysis.
6. The relatively small number of empirical investigations on long run PPP for the East Asian countries has produced to different results.
7. The STAR family of models was originally developed by Teräsvirta and Anderson (1993) for modeling nonlinearities over the business cycle. For their statistical properties and estimation, refer to Granger and Teräsvirta (1993) and Teräsvirta (1994) and for the application of the STAR to exchange rates see, Micheal, Nobay and Peel (1997), Sarantis (1999) and Sarno (2000), among others.
8. To implement the Johansen-Juselius procedure, one needs to determine the optimal lag length in the vector autoregressive (VAR) system. Our procedure for choosing the optimal lag was based on the Akaike information criteria (AIC). In addition, the residuals from the VAR were checked for white noise.
9. It is generally known that the power of the conventional unit root test is fairly poor in the present context, especially if the true DGP is in fact nonlinear mean reverting; see Taylor 2001 and Sarno, Taylor and Chowdhury (2004) on this issue.
10. It is generally known that the test also has low power against serially correlated errors. For detail discussion, see (Teräsvirta and Anderson 1993).
11. The w-test statistic is based on the approximation that $w = 0.5 \ln [(1+r)/(1-r)] \sim N(0, 1/(n-3))$.

References

Bacon, D.W. and D.G. Watts (1971) “Estimating the Transition between Two Intersecting Lines.” *Biometrica* 58:525–534.

Baharumshah, Ahmad Zubaidi and Mohamed Ariff (1997) “Purchasing Power Parity in South East Asian Countries Economies: A Cointegration Approach.” *Asian Economic Journal* 11:141–153.

Baum, Christopher F, John T. Barkoulas, and Mustafa Caglayan (2001) “Non-linear Adjustment to Purchasing Power Parity in the Post-Bretton Woods Era.” *Journal of International Money and Finance* 20:379–399.

Chan, K.S., and Howell Tong (1986) “On Estimation Thresholds in Autoregressive Models.” *Journal of Time Series Analysis* 7:179–190.

Clements, Micheal P and Jeremy Smith (1997) “A Monte Carlo Study of the Forecasting Performance of Empirical SETAR Models.” *International Journal of Forecasting* 13:463–475.

Coakley Jerry and Anna-Maria Fuertes (2001) “Nonparametric Cointegration Analysis of Real Exchange Rates.” *Applied Financial Economics* 11:1–8.

Devereux, Michael B (1997) “Real Exchange Rates and Macroeconomics: Evidence and Theory.” *Canadian Journal of Economics* XXX (4a):773–808.

De Grauwe Paul, Hance Dewachter and Mark Embrechts (1993) *Exchange Rate Theory; Chaotic Models of Foreign Exchange Markets. Oxford: Blackwell.*

Diebold, Francis X. and James A. Nason (1990) “Non-parametric Exchange Rate Predictions.” *Journal of International Economics* 28:315–332.

Eirtheim, Oyvind and Timo Teräsvirta (1996) “Testing the Adequacy of Smooth Transition Autoregressive Model.” *Journal of Econometrics* 74:59–75.

Granger, Clive W.J. and Paul Newbold (1986) *Forecasting Economic Time Series*. San Diego: Academic Press.

Granger, Clive W.J. and Timo Teräsvirta (1993) *Modelling Non-linear Economic Relationships*. London: Oxford University Press.

Johansen, Søren and Katarina Juselius (1990) "Maximum Likelihood Estimation and Inference on Cointegration–with Application to the Demand for Money." *Oxford Bulletin of Economics and Statistics* 52:169–210.

Kilian, Lutz and Taylor Mark P (2001) "Why Is It so difficult to Beat the Random Walk Forecast of Exchange Rates." Working Paper, Centre for Economic Policy Research, US.

Lin, Winston T., and Yueh H. Chen (1998) "Forecasting Foreign Exchange Rates with an Intrinsically Nonlinear Dynamic Speed of Adjustment Model." *Applied Economics* 30:295–312.

Luukkonen, Ritva (1990) "Estimating Smooth Transition Autoregressive Models by Conditional Least Squares." In R. Luukkonen (ed.), *On Linearity Testing and Model Estimation in Non-linear Time Series Analysis*. Helsinki: Finnish Statistical Society.

Luukkonen, Ritva, Pentti Saikkonen and Timo Teräsvirta (1988) "Testing Linearity against Smooth Transition Autoregressive Models." *Biometrika* 75:491–499.

Mahajan, Arvind and Andrew Wagner J. (1999) "Non-linear Dynamics in Foreign Exchange Rates." *Global Finance Journal* 10(1):1–23.

MacKinnon, James G. (1996) "Numerical Distribution Functions for Unit Root and Cointegration Tests." *Journal of Applied Econometrics* 1:601–618.

Meese, Richard A. and Andrew K. Rose (1991) "An Empirical Assessment of Non-linearities in Models of Exchange Rate Determination." *Review of Economics Studies* 58:603–619.

Meese, Richard A. and Kenneth Rogoff (1983a) "Empirical Exchange Rates Models in the Seventies: Do They Fit Out-of-Sample?" *Journal of International Economics* 14:3–74.

Meese, Richard A. and Kenneth Rogoff (1983b) "The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?" In J.A. Frenkel (ed.), *Exchange Rates and International Macroeconomics*, Chicago: Chicago University Press and National Bureau of Economic Research.

Micheal, Panos, A.R. Nobay and David A. Peel (1997) "Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation." Journal of Political Economy 105(4):933–948.

Mizrach, Bruce (1995) "Forecast Comparison in L_2," Working Paper, Department of Economics, Rutgers University.

Mohamed Azali, Habibullah Muzafar Shah, Baharumshah Ahmad Zubaidi (2001) "Does PPP Hold between Asian and Japanese Economies? Evidence using Panel Unit Root and Panel Cointegration." *Japan and the World Economy* 13:35–50.

Shiller, R, and P. Perron (1985) "Testing the Random Walk Hypothesis: Power Versus Frequency of Observation." *Economics Letters* 18:381–386.

Saikkonen, Pentti and Ritva Luukkonen (1988) "Lagrange Multiplier Tests for Testing Nonlinearities in Time Series Models." *Scandinavian Journal of Statistics* 15:55–68.

Sarantis, Nicholas (1999) "Modeling Nonlinearities in Real Effective Exchange Rates." *Journal of International Money and Finance* 18:27–45.

Sarno, Lucio (2000) "Real Exchange Rate Behaviour in the Middle East: A Re-examination." *Economics Letters* 66:127–136.

Sarno, Lucio, Mark P. Taylor and Chowdury Ibrahim (2004) "Non Linear Dynamics in Deviations from the Law of One Price: A Broad-Based Empirical Study." *Journal of International Money and Finance* 23:1–25.

Taylor, Allan M. (2001) "Potential Pitfalls for the Purchasing Power Parity Puzzle?" Sampling and Specification Biases in Mean Reversion of The Law of One Price." *Econometrica* 69:473–498.

Taylor, Mark P. and Allen Helen (1992) "The Use of Technical analysis in the Foreign Exchange Market." *Journal of International Money and Finance* 11:304–314.

Taylor, Mark P. and David A. Peel (1997) "Nonlinearities in Real Exchange Rate Adjustment during the Recent Float: Empirical Evidence and Monte Carlo Analysis." Discussion Paper, Institute of Economics and Statistics, University of Oxford.

Taylor, Mark P. and David A. Peel (2000) "Non-linear Adjustment, Long-run Equilibrium and Exchange Rate Fundamentals." *Journal of International Money and Finance* 19:33–53.

Tong, Howell and K.S. Lim (1980) "Threshold Autoregression, Limit Cycles and Cyclical Data. " *Journal of Royal Statistics Soceity* 42(3):245–292.

Teräsvirta, Timo (1994) "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models." *Journal of American* Statistics Association 89(425):208–218.

Teräsvirta, Timo and Anderson Heather M. (1993) "Characterizing Non-linearities in Business Cycles using Smooth Transition Autoregressive Models." In M. Hashem and Pesaran and Simon M. Potter (eds.), *Non-linear Dynamics, Choas and Econometrics*. US: John Wiley.