Empir Econ (2015) 48:407–426
DOI 10.1007/s00181-014-0888-5

Forecasting major Asian exchange rates using a new semiparametric STAR model

Nan Cai · Zongwu Cai · Ying Fang · Qiuhua Xu

Received: 4 June 2013 / Accepted: 11 October 2014 / Published online: 3 December 2014

Abstract To forecast exchange rates, this paper proposes a new semiparametric smooth transition autoregressive model by allowing state variables to enter into the transition function in a nonparametric way. We propose a three-stage estimation procedure to estimate both the parametric and nonparametric parts in the new model, and a simulation study is conducted to demonstrate satisfactory finite sample performance. The empirical results, based on the proposed model applied to forecasting five major

We thank the valuable comments from two anonymous referees and the editors Subal Kumbhakar and Christopher Parameter. Their comments improved the quality of our paper immensely. Zongwu Cai acknowledges the financial support from the NSFC #71131008. Ying Fang acknowledges the financial supports from NSFC #70971113, #71271179 and NCET-12-0328.

N. Cai
School of Public Finance and Taxation, Northeastern University of Finance and Economics,
Dalian 116025, China
e-mail: cainan2006@yahoo.com.cn

Z. Cai
Department of Economics, University of Kansas, Lawrence, KS 66045, USA
e-mail: zongwucai@gmail.com

Z. Cai · Y. Fang (✉)
The Wang Yanan Institute for Studies in Economics, MOE Key Laboratory of Econometrics
(Xiamen University), Fujian Provincial Key Laboratory of Statistical Science,
Xiamen University, Xiamen 361005, Fujian, China
e-mail: yifst1@gmail.com

Q. Xu
The Wang Yanan Institute for Studies in Economics, Xiamen University,
Xiamen 361005, Fujian, China
e-mail: qiuhua814@gmail.com

Asian exchange rates, show that the new model has some advantages in out-of-sample forecasting performance.

Keywords Nonlinearity · Out-of-sample forecasting · Semiparametric estimation · STAR model · Time-varying

JEL Classification C53 · C14 · C21

1 Introduction

The exchange rate, reflecting relative values between different currencies, is one of the most important financial and macroeconomic indicators in an economy. The movement of exchange rates produces influential impacts on international trade, capital flows and asset portfolio management. Although forecasting exchange rates has received much attention and extensive research recently, it is still a very challenging and active research area due to the nonlinearity and time-varying features existing in exchange rate data, see the papers by Hong and Lee (2003) and Cai et al. (2012) for details.

Many parametric nonlinear models have been developed to deal with the nonlinearity and time-varying features in exchange rates. To name just a few, for example, Quant (1958) proposed a linear regression system by allowing two separate regimes. Many authors studied the variants of the regime switching models and discussed the choice of the transition function which determines how to switch from one regime to another, see Bacon and Watts (1971), Goldfeld and Quandt (1972), Maddala (1977), Haggan and Ozaki (1981), Chan and Tong (1986), among others. The smooth transition autoregressive (STAR) model (Granger and Teräsvirta 1993; Teräsvirta 1994; Eitrheim and Teräsvirta 1996) is most widely employed in the class of regime switching models. The STAR model can be simply regarded as a time-varying weighted average of two linear autoregressive models, and the smooth change between two regimes is determined by a so-called transition function, which commonly adopts either a logistic function or an exponential function. Dijk et al. (2002) provided a comprehensive survey about STAR models and their applications in economics. However, empirical studies provide mixed evidence in support of the STAR model. For example, using monthly data on real effective exchange rates of ten major industrial countries, Sarantis (1999) found that the STAR model is not significantly better than linear models in terms of predictive ability. Stock and Watson (1996) employed the STAR model to forecast various monthly US macroeconomic data, but in most cases, the forecasting performance of the STAR model is inferior to linear models. Boero and Marrocu (2002) obtained similar results using several exchange rate data. A recent study by Rapach and Wohar (2006) also found that, in terms of short term prediction, the STAR model and the autoregressive (AR) model were similar in stock return data. One of the possible explanations to the aforementioned problems is that parametric STAR models may not be flexible enough to capture the nonlinear dynamics in exchange rates.

An alternative approach to model the nonlinearity is to use nonparametric techniques. For example, Diebold and Nason (1990) applied kernel estimation to ten major dollar spot rates in the post-1973 period. Kuan and Liu (1995) modeled the exchange rates using a feed-forward and recurrent network. Mizrach (1992) predicted the European monetary system currencies via a multivariate nearest neighbor estimation. Moreover, Hong and Lee (2003) and Cai et al. (2012) adopted the functional coefficient model to predict exchange rates. On the other hand, some authors provided empirical evidence that a complicated nonparametric model cannot even beat a simple random walk model, particularly in terms of out-of-sample forecasting. For example, Diebold and Nason (1990) and Meese and Rose (1991) found that nonparametric models provided little improvement in out-of-sample prediction for many exchange rates. Another concern is that the nonparametric techniques usually provide very little economic explanation, and the estimation and prediction process are criticized as a black box.

In this paper, we propose a new semiparametric STAR model to forecast the major five Asian exchange rates. The new semiparametric model adopts the basic framework of the STAR model but allows the state variables to enter into the transition function in a nonparametric way. Therefore, the new semiparametric STAR model not only inherits some merits of the STAR model, such as a two-regime structure and a clear economic explanation,[1] but also provides more flexibility to fit the data and alleviate the risk of misspecification. We propose a three-stage estimation procedure to estimate both the parametric and nonparametric parts in the new model. Our simulation results show that the proposed estimation method works very well even in a finite sample setting. Furthermore, we apply the new model to forecast five major Asian exchange rates: the Indian rupee (INR), Japanese yen (JPY), Singapore dollar (SGD), Korean won (KRW) and Thai baht (THB). The empirical results show that the semiparametric model has some advantages over other alternative models in terms of out-of-sample forecasting performance.

The rest of the paper is organized as follows. Section 2 introduces the new model and discusses the three-stage estimation method. Monte Carlo simulations are conducted in Sect. 3 to demonstrate the finite sample performance of the estimation method. In Sect. 4, to forecast five major Asian exchange rates, we compare the semiparametric STAR model with other models in terms of the out-of-sample forecasting ability. Section 5 concludes.

2 Econometric models and their modeling procedures

2.1 Review of STAR models

For a time series $\{y_t\}$, the STAR model of Teräsvirta (1994) can be expressed as follows:

$$y_t = \phi_{1,0} + \sum_{j=1}^{p} \phi_{1,j} y_{t-j} + \left(\phi_{2,0} + \sum_{j=1}^{p} \phi_{2,j} y_{t-j} \right) G(s_t; \gamma, c) + \varepsilon_t,$$

[1] Similar to the TAR model, both the STAR model and the semiparametric STAR model proposed in this paper can be extended to a three-regime structure, which would be of inherent interests in some economic and financial applications such as modeling financial returns. We thank one referee for pointing out this.

or

$$y_t = \phi_1' x_t + \phi_2' x_t G(s_t; \gamma, c) + \varepsilon_t,$$

where $x_t = (1, y_{t-1}, \ldots, y_{t-p})'$ is a vector of explanatory (lagged) variables, $\phi_i = (\phi_{i,0}, \phi_{i,1}, \ldots, \phi_{i,p})'$, $i = 1, 2$ is a vector of the corresponding coefficients of the explanatory variables, ε_t is a sequence of i.i.d. random variables, the transition function $G(s_t; \gamma, c)$ is a continuous function in the range of 0 and 1, s_t is a transition variable which can be lagged dependent variables, that is, $s_t = y_{t-d}$ for certain integer $d > 0$, c is the so-called threshold value, and γ is a smooth parameter, which determines the speed of transition between two regimes.

The STAR model can be regarded as a time-varying weighted average of two linear autoregressive models. When $G(s_t; \gamma, c)$ equals zero or one, y_t obeys an AR process. However, when $0 < G(s_t; \gamma, c) < 1$, y_t switches from one regime to another smoothly. Two types of transition functions are commonly used in the literature. One is a logistic function given by

$$G(s_t; \gamma, c) = \frac{1}{1 + \exp\{-\gamma(s_t - c)\}}, \quad \gamma > 0,$$

and the corresponding model is called the logistic STAR (LSTAR) model. The other is an exponential function

$$G(s_t; \gamma, c) = 1 - \exp\{-\gamma(s_t - c)^2\}, \quad \gamma > 0,$$

and the corresponding model is called an exponential STAR (ESTAR) model. If the threshold value c is an $n \times 1$ vector, then such a model is called an nth-order LSTAR or ESTAR model, which allows multiple switches between two regimes. For example, the transition function of an nth-order LSTAR model is given by

$$G(s_t; \gamma, c) = \frac{1}{1 + \exp\left\{-\gamma \prod_{i=1}^{n} (s_t - c_i)\right\}}, \quad \gamma > 0;$$

see Dijk et al. (2002) for more discussions.

2.2 Semiparametric STAR model and its estimation procedure

The STAR model can well characterize dynamic characteristics of changes between two regimes and possess good economic implications. The logistic and exponential transition functions clearly define two regimes which can be used to model different states in real economy, for example, the expansion and the recession. This paper further improves the STAR model by adopting a semiparametric transition function by allowing the transition variables to enter into the transition function in a nonparametric way.

The semiparametric STAR model, termed as SSTAR model, is given by

$$y_t = \phi_{1,0} + \sum_{j=1}^{p} \phi_{1,j} y_{t-j} + \left(\phi_{2,0} + \sum_{j=1}^{p} \phi_{2,j} y_{t-j} \right) G^*(f(s_t)) + \varepsilon_t,$$

or in a compact vector form

$$y_t = \phi_1' x_t + \phi_2' x_t G^*(f(s_t)) + \varepsilon_t, \quad (1)$$

where $f(\cdot)$ is an unknown function and for simplicity, $G^*(u)$[2] is taken to be either logistic

$$G^*(u) = \frac{1}{1 + \exp\{-u\}},$$

or exponential

$$G^*(u) = 1 - \exp(-u^2/2).$$

Indeed, the model given in (1) can be re-expressed as

$$y_t = a(s_t, \phi)' x_t + \varepsilon_t,$$

where $a(s_t, \phi) = \phi_1 + \phi_2 G^*(f(s_t))$, which is a semiparametric form. Therefore, the proposed model given in (1) can be regarded as a generalization of the classical functional coefficient regression model proposed in Cai et al. (2000) and is flexible enough to capture the nonlinear dynamics.

One of our main interests in this paper is to estimate the nonparametric functional $f(\cdot)$ in (1). To this end, we propose a three-stage estimation procedure to estimate the above semiparametric STAR model in (1), described as follows. At the first stage, the first-order Taylor expansion is applied to the unknown function $f(s_t)$, when s_t is in a neighborhood of the grid point s_0 from the domain of s_t,

$$f(s_t) \approx a + b(s_t - s_0),$$

where $a = f(s_0)$ and $b = f'(s_0)$, and then, the semiparametric STAR model in (1) is approximated by

$$y_t \approx \phi'_1 x_t + \phi'_2 x_t \frac{1}{1 + \exp\{a + b(s_t - s_0)\}} + \varepsilon_t.$$

[2] If $G^*(\cdot)$ is an unknown function but not either logistic or exponential, some identification conditions are needed. Therefore, for simplicity, we take it to be either logistic or exponential as defined above.

The first-stage estimates are obtained by minimizing the following localized objective function

$$L(\phi_1, \phi_2, a, b) = \sum_{t=1}^{T} \left(y_t - \phi'_1 x_t - \phi'_2 x_t \frac{1}{1 + \exp\{a + b(s_t - s_0)\}} \right)^2 K_{h_1}(s_t - s_0),$$

where $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function, and h_1 is the bandwidth at the first step. For given s_0 in the domain of s_t, we can obtain the nonparametric estimates $\{\hat{\phi}_1(s_0), \hat{\phi}_2(s_0), \hat{a}(s_0), \hat{b}(s_0)\}$. Based on the nonparametric theory, one may show easily under some regularity conditions that the nonparametric estimates $\{\hat{\phi}_1(s_0), \hat{\phi}_2(s_0), \hat{a}(s_0), \hat{b}(s_0)\}$ are $(T h_1)^{1/2}$ consistent and the asymptotic bias is of the order $O(h_1^2)$, see Cai et al. (2000).

Since ϕ_1 and ϕ_2 are constant parameters, the above nonparametric estimation for them is inefficient due to using local information. At the second stage, we obtain root-T consistent estimators of ϕ_1 and ϕ_2, denoted by $\tilde{\phi}_1$ and $\tilde{\phi}_2$, using the following average method. That is,

$$\tilde{\phi}_1 = \frac{1}{T} \sum_{t=1}^{T} \hat{\phi}_1(s_t), \quad \text{and} \quad \tilde{\phi}_2 = \frac{1}{T} \sum_{t=1}^{T} \hat{\phi}_2(s_t).$$

As expected, both $\tilde{\phi}_1$ and $\tilde{\phi}_2$ are the root-T consistent estimators of ϕ_1 and ϕ_2. However, it is common that at the first step, the bandwidth h_1 is under-smoothed to get rid of the asymptotic bias term. In other words, h_1 is taken to satisfy $T h_1^5 \to 0$, see Cai and Xiao (2012).

It is worth to point out that the well-known Robinson (1988) type or profile least squares type of Speckman (1988) estimation approach for classical semiparametric regression models might not be suitable to model (1) due to its nonlinearity. For example, for a profile least squares method, to estimate the parameters in the linear component under the least squares framework, one usually multiplies a projection matrix to remove the nonparametric component and then fit a linear model. But this approach is not applicable to the current model setting due to lack of explicit normal equations. Also, note that to gain the efficiency of $\tilde{\phi}_1$ and $\tilde{\phi}_2$, one may follow Cai and Xiao (2012) to use the weighted average method by choosing the efficient weights, see Cai and Xiao (2012) for more details.

Finally, at the last stage, we plug the estimates $\tilde{\phi}_1$ and $\tilde{\phi}_2$ obtained at the second stage into the objective function and estimate the nonparametric part again,

$$L(f) = \sum_{t=1}^{T} \left(y_t - \tilde{\phi}'_1 x_t - \tilde{\phi}'_2 x_t \frac{1}{1 + \exp\{a + b(s_t - s_0)\}} \right)^2 K_{h_3}(s_t - s_0),$$

where h_3 is the bandwidth at the third step. By moving s_0 over the whole the domain of s_t, we obtain the whole nonparametrically estimated curve $\tilde{f}(s_t)$.

Note that the Epanechnikov kernel function, $K(u) = 0.75(1 - u^2)1(|u| \leq 1)$, is employed at the first and last stages. At the first stage, as mentioned earlier, we need under-smoothing by selecting a small value of bandwidth h_1 to alleviate the impact of bias on latter stages. However, at the third stage, we can adopt some popular procedures such as the cross validation method to choose an optimal bandwidth for h_3.

3 Monte Carlo simulation

In this section, Monte Carlo experiments are conducted to evaluate the finite sample performance of the proposed three-stage estimation method. Consider the following data-generating process (DGP),

$$y_t = \phi_1 + \phi_2 y_{t-1} + (\phi_3 + \phi_4 y_{t-1})G^*(f(s_t)) + e_t, \tag{2}$$

where the transition variable s_t is generated from a uniform distribution in the range of $(-3, 3)$, and e_t is generated from a normal $N(0, 0.2^2)$. To investigate whether, for different forms of transition functions, our method can well estimate both the parametric coefficients and the unknown functions, given a finite sample size, we set the parameters as the following, corresponding to ESTAR, LSTAR1 and LSTAR2, respectively.

DGP1 [LSTAR1]:

$$\begin{cases} \phi_1 = 0.3, \phi_2 = -0.4, \phi_3 = -0.7, \phi_4 = 0.6; \\ f(s) = -2s \end{cases}$$

DGP2 [LSTAR2]:

$$\begin{cases} \phi_1 = 0.5, \phi_2 = -0.65, \phi_3 = -0.55, \phi_4 = 0.75; \\ f(s) = -0.5(s-1)(s+1) \end{cases}$$

DGP3 [ESTAR]:

$$\begin{cases} \phi_1 = 0.25, \phi_2 = -0.45, \phi_3 = -0.65, \phi_4 = 0.55; \\ f(s) = -0.5s^2 \end{cases}$$

The sample size T is taken to be 200, 500 and 800, respectively. For each sample size, we repeat 1,000 times in the Monte Carlo simulations.

Table 1 reports the simulation results of parametric coefficients under three DGPs, LSTAR1, LSTAR2 and ESTAR, respectively. In Table 1, ϕ_1, ϕ_2, ϕ_3 and ϕ_4, respectively, represent the parametric coefficients in model (2). We report the median of the absolute deviations between the true values and their estimates in 1,000 replications, and the standard deviations are reported in parentheses. For all coefficients under various transition functions, both the medians of absolute deviation and the standard deviations decrease as the sample size increases.

Table 2 shows the simulation results of nonparametric parts. The estimation performance is measured by the mean absolute deviation error (MADE), which is the deviation between the true values and their estimated values. The MADE is given by

Table 1 The medians and standard deviations (in parentheses) of 1,000 absolute deviations for parametric estimation

Model	Sample size	ϕ_1	ϕ_2	ϕ_3	ϕ_4
LSTAR1	200	0.0045 (0.0040)	0.0113 (0.0105)	0.0125 (0.0106)	0.0135 (0.0122)
	500	0.0040 (0.0036)	0.0075 (0.0064)	0.0111 (0.0098)	0.0106 (0.0088)
	800	0.0038 (0.0033)	0.0060 (0.0054)	0.0108 (0.0090)	0.0092 (0.0081)
LSTAR2	200	0.0047 (0.0043)	0.0118 (0.0112)	0.0123 (0.0106)	0.0147 (0.0132)
	500	0.0042 (0.0036)	0.0076 (0.0071)	0.0113 (0.0096)	0.0108 (0.0089)
	800	0.0037 (0.0036)	0.0066 (0.0057)	0.0105 (0.0092)	0.0098 (0.0081)
ESTAR	200	0.0049 (0.0043)	0.0106 (0.0101)	0.0092 (0.0081)	0.0124 (0.0113)
	500	0.0045 (0.0040)	0.0077 (0.0070)	0.0085 (0.0071)	0.0091 (0.0086)
	800	0.0044 (0.0036)	0.0062 (0.0055)	0.0079 (0.0069)	0.0085 (0.0078)

Table 2 The medians and standard deviations (in parentheses) of 1,000 MADEs for nonparametric estimation

Model	MADE		
	$T = 200$	$T = 500$	$T = 800$
LSTAR1	0.0184 (0.0051)	0.0129 (0.0031)	0.0106 (0.0025)
LSTAR2	0.0228 (0.0066)	0.0205 (0.0051)	0.0172 (0.0046)
ESTAR	0.0258 (0.0056)	0.0175 (0.0035)	0.0144 (0.0028)

$$\text{MADE} = m^{-1} \sum_{j=1}^{m} \left| \hat{f}(z_j) - f(z_j) \right|,$$

where $\{z_j\}_{j=1}^m$ are grid points. From Table 2, we can observe that the median and the standard deviation of the 1,000 MADE values shrink in a reasonable speed as the sample size is enlarged. This implies that the nonparametric estimation works very well even in a small sample. Figure 1 compares the estimated curves of $f(s_t)$ in solid lines to the corresponding true curves, which are given in dotted lines. All three cases show that the nonparametric estimation fits the true curves reasonably well.

4 Empirical results

4.1 The data

We use monthly data of five Asian exchange rates, which are the Indian rupee (INR/USD), the Japanese yen (JPY/USD), the Singapore dollar (SGD/USD), the Korean won (KRW/USD) and the Thai baht (THB/USD), over the period 1994: 1–2013:2 (230 observations). We reserve the last 50 observations, from 2009:11 to 2013:12, for checking the out-of-sample forecasting performance. We only look at the one-step ahead forecasting. All data for the exchange rates and the returns of the exchange rates are plotted in Figs. 2 for JPY/USD, SGD/USD and INR/USD, and 3 for KRW/USD and THB/USD. From Figs. 2 and 3, we observe that the original

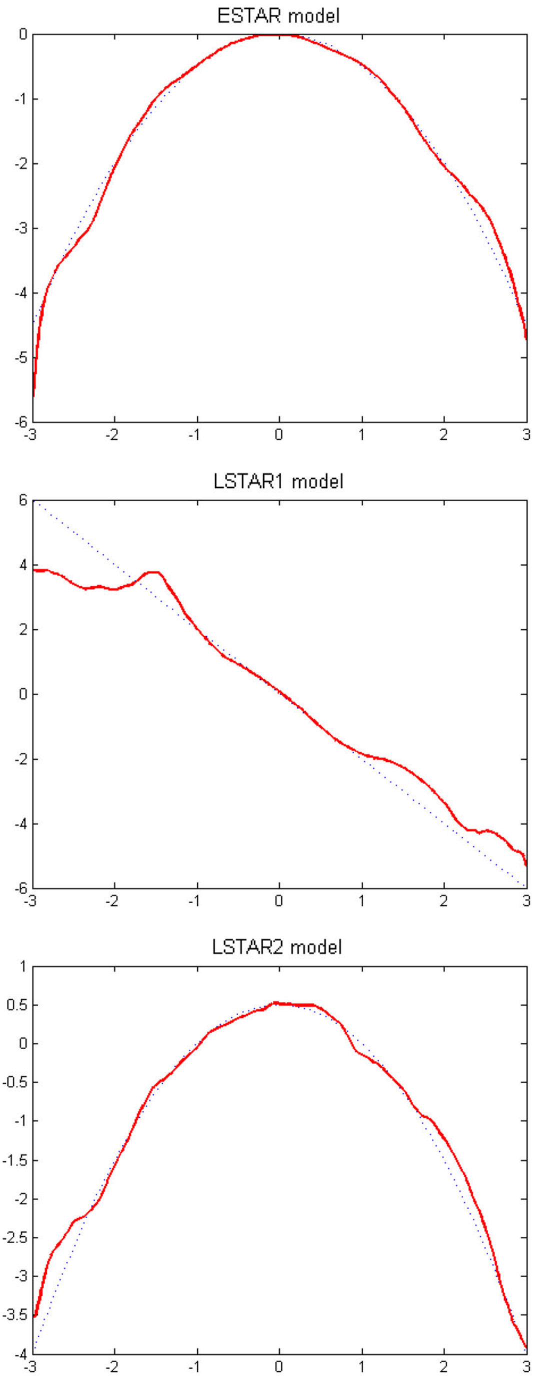

Fig. 1 The estimated curves of $f(s_t)$ (*bold solid lines*) versus their corresponding true curves (*dotted lines*). From the *top* to the *bottom*: ESTAR model, LSTAR1 model and LSTAR2 model

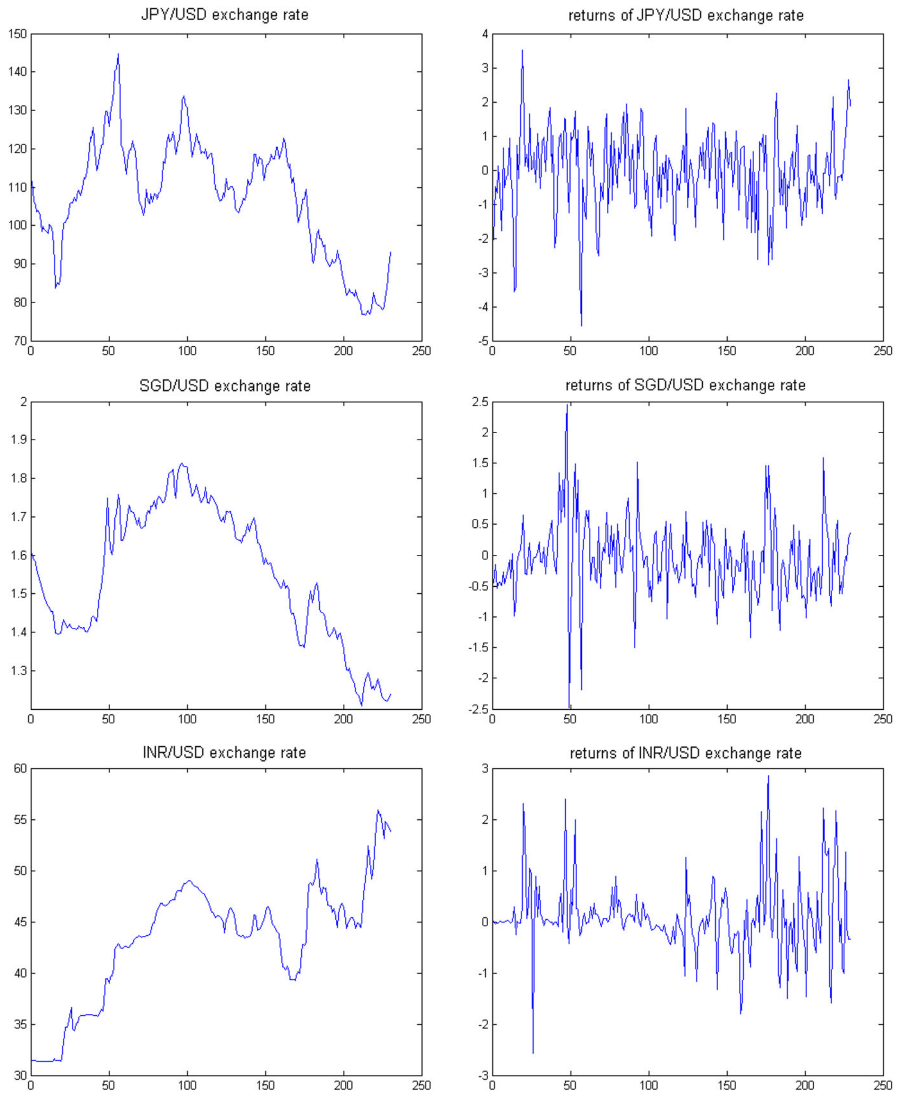

Fig. 2 The time series plots for exchange rates on the *left panel* and for the returns of the exchange rates on the *right panel*. From the *top* to the *bottom*: JPY/USD, SGD/USD and INR/USD

exchange rates data have an obvious time trend, but the return data are relatively stable. Table 3 reports the augmented Dickey–Fuller (ADF) unit root tests for the level and first difference of the exchange rates, both measured in logarithms. These results indicate that all the level time series are clearly integrated of order 1. Hence, the variable, denoted as Z_t, used in all estimations is the log return, the first logarithmic difference of exchange rates. Next, we apply the linearity test allowing for conditional heteroskedasticity proposed by Becker and Hurn (2009) to Z_t. Table 4 shows that all these transformed exchange rates are nonlinear.

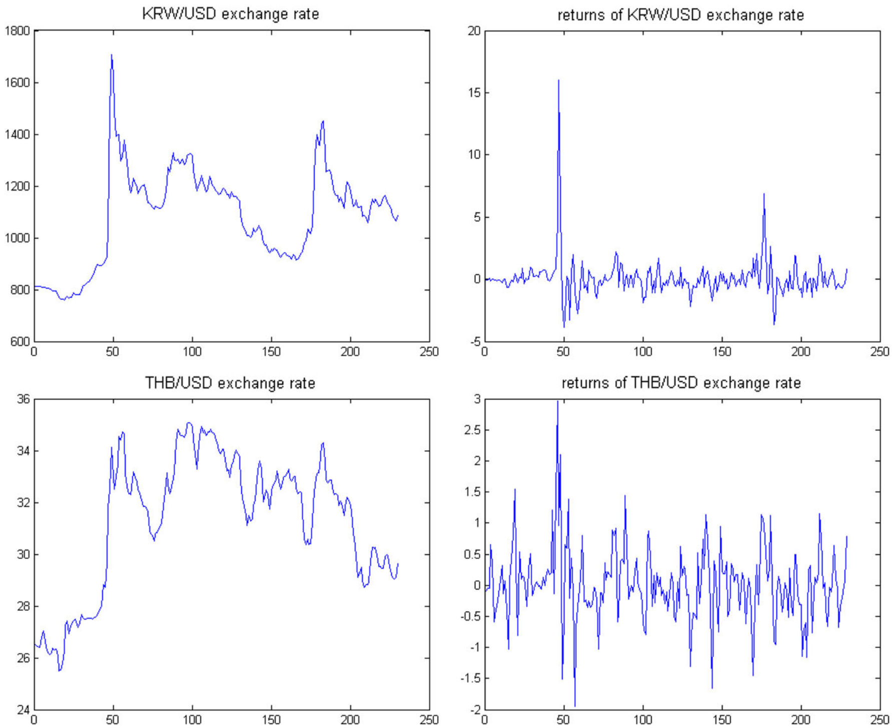

Fig. 3 The time series plots for exchange rates on the *left panel* and for the returns of the exchange rates on the *right panel*. From the *top* to the *bottom*: KRW/USD and THB/USD

Table 3 *P* values of ADF unit root tests for the Asian exchange rates

	JPY/USD	SGD/USD	INR/USD	KRW/USD	THB/USD
Level	0.3679	0.8752	0.4321	0.1842	0.1633
First difference	0.0000	0.0000	0.0000	0.0000	0.0000

Table 4 *P* values of the linearity test

	JPY/USD	SGD/USD	INR/USD	KRW/USD	THB/USD
P value	0.000	0.0608	0.0246	0.0774	0.0662

Next, we also do some diagnostic checks on the use of the semiparametric STAR model. Figures 4 and 5 provide the time plots of the in-sample fitting errors and out-of-sample forecasting errors, respectively, based on the semiparametric STAR model. Table 5 reports the p-values of Q-tests for both the in-sample fitting errors and the out-of-sample forecasting errors. For most cases except the KRW/USD exchange rate, we cannot reject the null hypothesis that both the in-sample errors and the out-of-sample errors are not serially correlated. For the KRW/USD exchange rate, the $Q(10)$ rejects

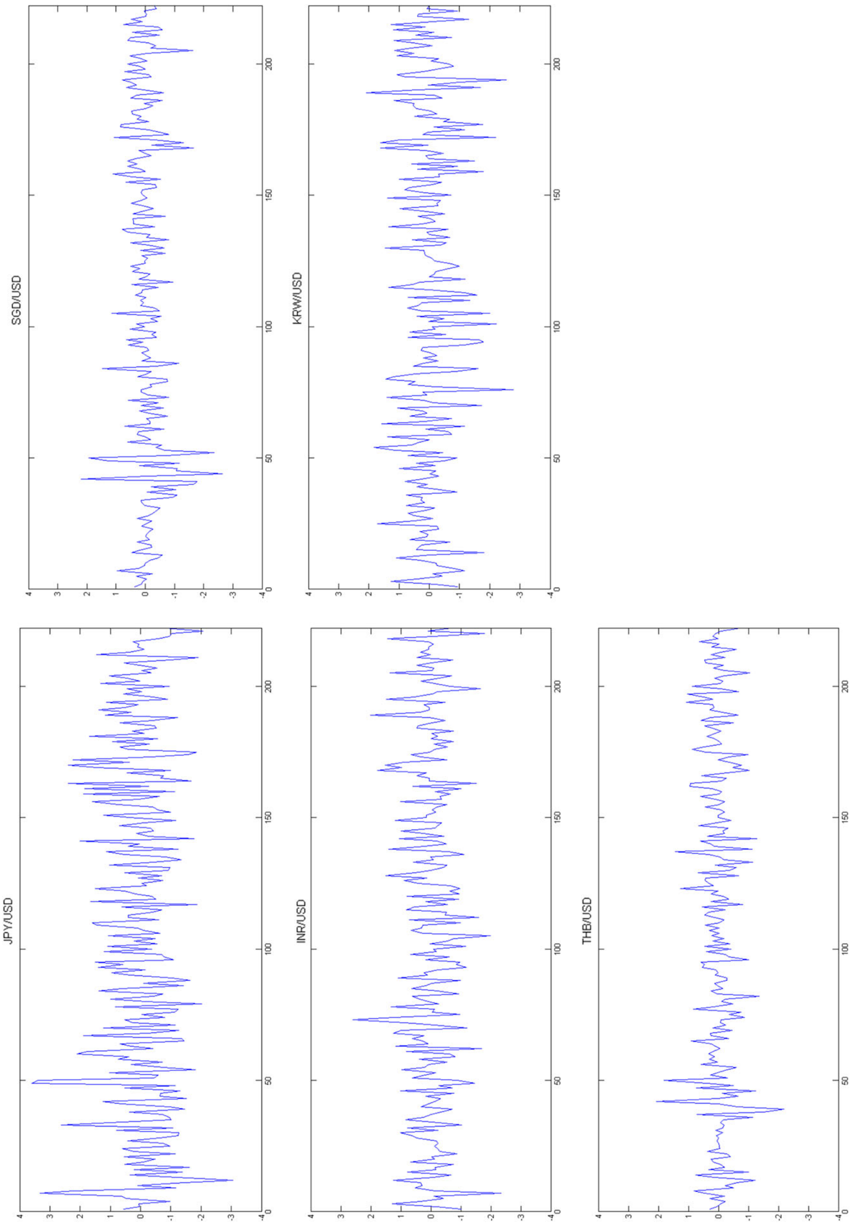

Fig. 4 The time plots of in-sample fitting errors computed from the SSTAR model

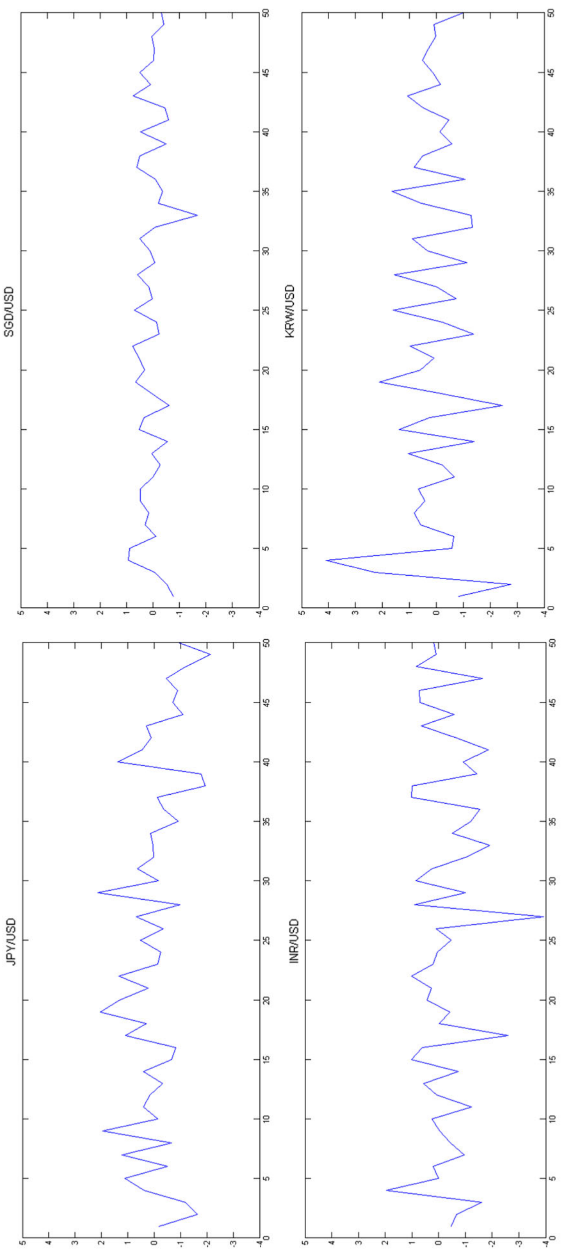

Fig. 5 The time plots of out-of-sample forecasting errors computed from the SSTAR model

Table 5 *P* values of *Q* tests for in-sample and out-of-sample errors

	SSTAR					
	In-sample			Out-of-sample		
	Q(5)	Q(10)	Q(20)	Q(5)	Q(10)	Q(20)
JPY/USD	0.102	0.152	0.101	0.677	0.836	0.947
SGD/USD	0.351	0.336	0.728	0.328	0.172	0.365
INR/USD	0.035	0.193	0.799	0.491	0.423	0.547
KRW/USD	0.371	0.037	0.201	0.007	0.004	0.000
THB/USD	0.593	0.944	0.766	0.183	0.258	0.113

the null hypothesis for the in-sample errors with a p value of 0.037, and all three Q tests reject the null for the out-of-sample errors with p values 0.007, 0.004 and 0.000, respectively.

4.2 Alternative models

To evaluate the performance of the semiparametric STAR model, we compare its in-sample fit and out-of-sample forecasts to other popular models including the random walk model, the autoregressive model, the threshold autoregressive model, the STAR model and the artificial neural network model. All models are estimated by using the first logarithmic difference of exchange rates.

The random walk model (RW), first proposed by Bachelier (1990), is given by

$$y_t = y_{t-1} + e_t, \quad t = 1, \ldots, T,$$

where y_t is the value in time period t, y_{t-1} is the value in time period $t-1$, and e_t is the random error term in time period t.

The autoregressive (AR) model describes a dynamic relationship between the current value and the historical values:

$$y_t = c + \sum_{i=1}^{p} \varphi_i y_{t-i} + e_t,$$

where $\{\varphi_i\}$ are the autoregression coefficients, y_t is the time series under investigation, and p is the order of the filter, which is generally much less than the length of the series, and e_t is assumed to be white noise. An autoregressive model can also be viewed as the output of an all-pole infinite impulse response filter whose input is white noise.

The threshold autoregressive (TAR) model was first proposed by Tong (1983, 1990). A widely used TAR model, the two-regime threshold autoregressive model, is

Table 6 Ratios of mean square errors for in-sample forecasting

	MSE				
	RW	AR	TAR	STAR	ANN
JPY/USD	1.7065	1.0544	1.0599	1.0568	0.5466
SGD/USD	1.6168	1.0276	0.9673	0.9713	0.5259
INR/USD	1.5230	0.9780	0.9144	0.9164	0.5199
KRW/USD	1.8607	1.2386	1.0468	1.0506	0.3801
THB/USD	1.7592	1.1838	1.1435	0.9948	0.7026

$$\Delta y_t = \left(\mu_1 + \sum_{i=1}^{k} \varphi_{1,i} \Delta y_{t-i}\right) I(s_{t-d}; c) + \left(\mu_2 + \sum_{i=1}^{k} \varphi_{2,i} \Delta y_{t-i}\right)(1 - I(s_{t-d}; c)) + e_t,$$

where s_{t-d} is the state determining variable, c is the threshold parameter, and d is the delay parameter which determines how many lags the state determining variable influences the regime in time t, and $I(x_t; c)$ is an indicator function defined by $I(x_t; c) = 1$ if $x_t < c$ and 0 otherwise. When $s_{t-d} = \Delta y_{t-d}$, it becomes the so-called self-exciting TAR (SETAR) model, see Tong (1990) for more details.

Finally, the last alternative model, the ANN model, was proposed by McCullogh and Pitts (1943). The artificial neural network (ANN) model is a data processing system based on the topological structure of the human brain, or the mathematical model of simulating organization and function of biological neurons. The basic unit of the ANN model is neurons, composed of three layers: input layer, hidden layer and output layer. To apply the model, one first needs to define an evaluation function to measure the gap between the network output and the excepted output, and then assign a set of random starting values of weights which describe the relationship between input and output variables. The optimal weights can be obtained by minimizing the gap through a process known as "training." Some widely used training methods include the back-propagation algorithm, collapsible neural network and genetic algorithm. Moreover, the nonlinear adaptive information processing ability owned by the neural network makes itself have the self-learning function, the associative memory function and a high-speed ability to find the optimal solution.

4.3 In-sample and out-of-sample performance comparison

We first evaluate the in-sample goodness of fit by calculating in-sample mean square errors (MSE). Table 6 reports ratios of the in-sample mean square errors with respect to the SSTAR model. Compared to the random walk model, the semiparametric STAR model has uniformly smaller mean square errors for all five exchange rates. The results are mixed when we compare the semiparametric STAR model to the AR, TAR and STAR models, but most ratios are either larger than one or very close to one. The ANN model produces the smallest in-sample mean square errors for all five exchange rates.

Table 7 Forecasting performance comparison in terms of MSFE and MAFE

	RW	AR	TAR	STAR	ANN
JPY/USD					
MSFE	1.5012	1.6746	1.0973	1.0782	1.3506
MAFE	1.2717	1.2521	1.0335	1.0464	1.1556
SGD/USD					
MSFE	1.5359	1.0521	1.0919	1.1542	1.1287
MAFE	1.2560	1.0413	1.0259	1.0908	1.0263
INR/USD					
MSFE	1.5196	0.9573	1.0255	1.1602	0.9503
MAFE	1.2209	0.9767	0.9904	1.0265	0.9395
KRW/USD					
MSFE	1.4090	1.1186	0.9257	1.2599	1.2020
MAFE	1.1738	1.0670	0.9550	1.0428	1.1190
THB/USD					
MSFE	1.3452	1.5361	1.1049	1.0453	1.8230
MAFE	1.1732	1.2898	1.0780	1.0304	1.3798

However, a pure nonparametric model is inclined to overfit in-sample, see Diebold and Nason (1990) and Meese and Rose (1991). We therefore move to compare the out-of-sample forecasting performance.

The out-of-sample forecast performance[3] is evaluated by the mean squared forecasting error (MSFE) and the mean absolute forecasting error (MAFE), given by

$$\text{MSFE} = m^{-1}\sum_{i=1}^{m}(Y_{T+i} - \hat{Y}_{T+i})^2 \quad \text{and} \quad \text{MAFE} = m^{-1}\sum_{i=1}^{m}\left|Y_{T+i} - \hat{Y}_{T+i}\right|,$$

where $m = 50$ is the forecasting period. Table 7 presents computed ratios of out-of-sample MSFE and MAFE values with respect to the SSTAR model. In most cases, we can observe that the semiparametric STAR model has the small MSFE and MAFE values relative to other models. However, for the INR/USD, both the AR and the ANN models are better than the SSTAR model in terms of MSFE and MAFE, but the difference is small. For KRW/USD, the TAR model is also slightly better than the SSTAR model in both the MSFE and MAFE.

From Table 7, we can see that semiparametric STAR model has an advantage in out-of-sample forecasting, but we need to make sure whether the advantage is significant. We check this through the so-called superior predictive ability (SPA) tests. The null hypothesis is that the proposed model is not inferior to all alternative models. The SPA test was first proposed by White (2002), also called the reality check (RC) test. Hansen (2005) proposed an improved version of the RC test and suggested a new

[3] Inoue and Kilian (2004, 2006) discussed cases, especially for nested models, when predictability tests and the selection of forecasting models based on out-of-sample error comparisons can lead misleading results.

Table 8 Comparison for superior predictive ability

	RC	SPA
JPY/USD		
SFE	0.1776	0.5124
AFE	0.2544	0.7922
SGD/USD		
SFE	0.1634	0.7600
AFE	0.4524	0.7322
INR/USD		
SFE	0.1166	0.2172
AFE	0.3078	0.5550
KRW/USD		
SFE	0.3196	0.4828
AFE	0.2408	0.4506
THB/USD		
SFE	0.3896	0.6652
AFE	0.4040	0.6206

Table 9 Comparison for superior predictive ability II

	RW		STAR	
	RC	SPA	RC	SPA
JPY/USD				
SFE	0.0050	0.0074	0.3152	0.8938
AFE	0.0020	0.0030	0.1666	0.5538
SGD/USD				
SFE	0.0000	0.0000	0.0794	0.2530
AFE	0.0004	0.0010	0.1672	0.3632
INR/USD				
SFE	0.0004	0.0006	0.5020	0.933
AFE	0.0088	0.0196	0.1712	0.377
KRW/USD				
SFE	0.0092	0.0174	0.0834	0.2260
AFE	0.0052	0.0178	0.0894	0.1974
THB/USD				
SFE	0.0002	0.0002	0.0136	0.0176
AFE	0.0002	0.0004	0.1188	0.2258

testing procedure known as the SPA test, which is more powerful and less sensitive to poor and irrelevant alternatives compared to the reality check test. In Table 8, we choose the random walk model, the autoregressive model, the threshold autoregressive model, the STAR model and the ANN model as the group of alternative models. The squared forecasting errors (SFE) and the absolute forecasting errors (AFE) are employed as the

Table 10 Inference results using out-of-sample forecast errors

	RW	STAR
	JPY/USD	
r_{MSE}	0.7583	1.0559
$Q_{0.50}$	0.0175	0.7645
$[Q_{0.25}, Q_{0.75}]$	[0.0040, 0.0373]	[0.7208, 0.7949]
	SGD/USD	
r_{MSE}	0.6765	0.9274
$Q_{0.50}$	0.0000	0.1575
$[Q_{0.25}, Q_{0.75}]$	[0.0000, 0.0000]	[0.1175, 0.1989]
	INR/USD	
r_{MSE}	0.6544	0.9798
$Q_{0.50}$	0.0000	0.4093
$[Q_{0.25}, Q_{0.75}]$	[0.0000, 0.0000]	[0.3763, 0.4416]
	KRW/USD	
r_{MSE}	0.8287	1.1435
$Q_{0.50}$	0.0000	0.8980
$[Q_{0.25}, Q_{0.75}]$	[0.0000, 0.0015]	[0.8759, 0.9241]
	THB/USD	
r_{MSE}	0.7214	0.8539
$Q_{0.50}$	0.0000	0.0678
$[Q_{0.25}, Q_{0.75}]$	[0.0000, 0.0000]	[0.0505, 0.0945]

loss function to evaluate model performance. The p values of all tests are presented in Table 8. All tests fail to reject the null hypothesis that the SSTAR model is at least as good as the alternatives, and for the SPA test, which are considered to be more powerful than the RC test, the p values are larger than 50 % in most cases. However, for the INR/USD exchange rates, particularly for the SFE, the p values of two tests are just 0.11 and 0.22, which implies the SSTAR just passes the SPA tests marginally.

Since it is difficult to reach some informative conclusions for model comparison when one fails to reject the null hypothesis in the RC and SPA tests, to obtain robust testing results, we set the random walk model and the STAR model as the null model, respectively, and put the semiparametric STAR model and the remaining models into the alternative group. Table 9 shows that the null hypothesis of the random walk model is strongly rejected for all exchange rates. However, for the STAR model, the test results are mixed. In most cases, we cannot reject the null hypothesis that the STAR model is not inferior to alternative models. To be concluded, in terms of the out-of-sample forecasting performance, the semiparametric STAR model is significantly better than the random walk model. Compared to the STAR model, the semiparametric

STAR model has smaller out-of-sample MSFE and MAFE, but the improvement is not significant.

All the above inference procedures are justified only for large samples. In order to consider small sample uncertainty, as suggested by the editor, we apply the bootstrapping-based postsample inference procedure proposed by Ashley (1998). For two postsample forecasting errors produced from different forecasting models, we first generate 100 starting samples based on the estimation of a VAR model with a maximum lag order of two. Given any starting sample, say the 24th starting sample, we then generate 2000 bootstrap samples using the same method. For each bootstrap sample, we calculate the ratio $\frac{\hat{r}_i}{r_{24}}$ where $\hat{r}_i$ is the ratio of two mean square errors of the two postsample forecasting errors in the i^{th} bootstrap sample, and r_{24} is the ratio of mean square errors from a single sample of 5000 observations generated from the 24th starting sample. Using these 2000 bootstrap samples, we can easily compute the fraction of the event of $\frac{\hat{r}_i}{r_{24}} \geqslant \hat{r}_{\text{orig}}$, where $\hat{r}_{\text{orig}}$ is the ratio of mean square errors computed from the original sample. Finally, the 100 stating sample can provide a distribution of these fractions, and then, we can report the median and the corresponding empirical 50 % confidence interval.

Table 10 summarizes all results of the bootstrap-based postsample inference procedure. We report the sample ratios of mean square errors $\hat{r}_{\text{orig}}$, the median fraction $Q_{0.50}$ and the corresponding empirical confidence interval $[Q_{0.25}, Q_{0.75}]$. When comparing SSTAR to the random walk model, we find that the mean square errors can be improved by at least 25 % for all five exchange rates, and the improvement is significant even considering the small sample uncertainty. However, the results for the comparison between SSTAR and STAR are again mixed. In most cases, the mean square errors of the SSTAR model are smaller than those of the STAR model, where the only exception is the exchange rate of JPY/USD, but most of them are not significant which reconfirm the results in Table 9.

5 Conclusion

This paper proposes a new semiparametric STAR model to forecast five Asian exchange rates. Compared to the traditional STAR models, the new model can provide more flexibility to characterize the nonlinearity and time-varying features of the exchange rate data by allowing the state variable to enter into the transition function in a nonparametric way, while at the same time, the new model can still inherit some advantages of the classical STAR model, such as the two-regime structure and good economic intuition. In terms of the out-of-sample forecasting performance, the semiparametric STAR model is superior to other models including the random walk model, the autoregressive model, the threshold autoregressive model and the ANN model in various superior predictive ability tests. However, although the semiparametric STAR model has smaller out-of-sample MSFE and MAFE than the STAR model, the improvement is not significant.

References

Ashley R (1998) A new technique for postsample model selection and validation. J Econ Dyn Control 22:647–665
Bachelier L (1990) Théorie de la Spéculation. Annales Scientifiques de l'E.N.S. 17:21–86
Bacon DW, Watts DG (1971) Estimating the transition between two interesting straight lines. Biometrika 58:525–534
Becker R, Hurn AS (2009) Testing for nonlinearity in mean in the presence of heteroskedasticity. Econ Anal Policy 39:311–326
Boero G, Marrocu E (2002) The performance of non-linear exchange rate models: a forecasting comparison. J Forecast 21:513–542
Cai Z, Chen L, Fang Y (2012) A new forecasting model for USD/CNY exchange rate. Stud Nonlinear Dyn Econom 16:1558–3708
Cai Z, Fan J, Yao Q (2000) Functional-coefficient regression models for nonlinear time series. J Am Stat Assoc 95:941–956
Cai Z, Xiao Z (2012) Semiparametric quantile regression estimation in dynamic models with partially varying coefficients. J Econ 167:413–425
Chan WS, Tong H (1986) On tests for non-linearity in time series analysis. J Forecas 5:217–228
Diebold FX, Nason JA (1990) Nonparametric exchange rate prediction? J Int Econ 28:315–332
Eitrheim Ø, Teräsvirta T (1996) Testing the adequacy of smooth transition autoregressive models. J Econom 74:59–76
Goldfeld SM, Quandt R (1972) Nonlinear Methods Econom. North Holland, Amsterdam
Granger CWJ, Teräsvirta T (1993) Model Nonlinear Econ Relat. Oxford University Press, Oxford
Haggan V, Ozaki T (1981) Modeling nonlinear vibrations using an amplitude-dependent autoregressive time series model. Biometrika 68:189–196
Hansen PR (2005) A test for superior predictive ability. J Bus Econ Stat 23:365–380
Hong Y, Lee T (2003) Inference on via generalized spectrum and nonlinear time series models. Rev Econ Stat 85:1048–1062
Inoue A, Kilian L (2004) In-sample or out-of-sample tests of predictability: which one should we use? Econom Rev 23:371–402
Inoue A, Kilian L (2006) On the selection of forecasting models. J Econom 130:273–306
Kuan CM, Liu T (1995) Forecasting exchange rates using feed-forward and recurrent neural networks. J Appl Econom 10:347–364
Maddala DS (1977) Econometrics. McGraw-Hill, New York
McCullogh WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5:115–133
Meese RA, Rose AK (1991) An empirical assessment of non-linearities in models of exchange rate determination. Rev Econ Stud 58:603–619
Mizrach B (1992) Multivariate nearest-neighbor forecasts of EMS exchange rate. J Appl Econom 7:151–163
Quant RE (1958) The estimation of the parameters of a linear regression system obeying two separate regimes. J Am Stat Assoc 53:873–880
Rapach DE, Wohar ME (2006) In-sample versus out-of-sample tests of stock return predictability in the context of data mining. J Empir Financ 13:231–247
Robinson PM (1988) Root-N-consistent semiparametric regression. Econometrica 56:931–954
Sarantis N (1999) Modeling non-linearities in real effective exchange rate. J Int Money Financ 18:27–45
Speckman P (1988) Kernel smoothing partial linear models. J R Stat Soc Ser B 50:413–426
Stock JH, Watson MW (1996) Evidence on structural instability in macroeconomic time series relations. J Bus Econ Stat 14:11–30
Teräsvirta T (1994) Specification, estimation, and evaluation of smooth transition autoregressive models. J Am Stat Assoc 89:208–218
Tong H (1983) Threshold models in non-linear time series analysis. In: Lecture Notes in Statistics, No. 21, Springer, Heidelberg
Tong H (1990) Non-linear time series: a dynamical system approach. Clarendon Press, Oxford
van Dijk D, Teräsvirta T, Franses PH (2002) Smooth transition autoregressive models—a survey of recent developments. Econom Rev 21:1–47
White H (2002) A reality check for data snooping. Econometrica 68:1097–1126