Forecasting the Real Exchange Rates Behavior:

An Investigation of Nonlinear Competing Models

Yu Liu and Ruxandra Prodan

Abstract: There is a large amount of literature which finds that real exchange rates appear to be characterized by several non-linear specifications. While each of these nonlinear models "fits" some particular real exchange rates series especially well, leading to good in-sample properties, the recent studies have not come to any consensus whether the nonlinear models provide a better specification than the linear model and/or the random walk model according to their out of sample forecasting performance. Our goal is to examine two important nonlinear methods (Band-TAR and ESTAR models) concerning their ability to generate out of sample forecasts, when estimating real exchange rate series for 20 OECD countries. We find strong evidence that the ESTAR model outperforms the random walk model and that neither the linear model nor the Band-TAR model significantly outperforms the random walk model, when forecasting out of sample. On the other hand, a comparison between the nonlinear models and the linear models would not be conclusive due to the low power of tests for predictive ability when bootstrapping critical values.

Correspondence to:

Ruxandra Prodan, tel. (713) 743-3836, email: rprodan@mail.uh.edu, Department of Economics, University of Houston, Houston, TX 77204.
Yu Liu, tel. (205) 348-0255, email: yuliu@cba.ua.edu, Department of Economics, University of Alabama, Tuscaloosa, AL 35487.

1. Introduction

There is a huge literature on whether Purchasing Power Parity holds in the post-Bretton Woods period. Numerous studies have found that the real exchange rates among industrialized countries appear to be characterized by a non-stationary behavior, implying the absence of any long-run tendency towards PPP.

This view has been lately challenged by a growing amount of nonlinear literature. Numerous empirical studies have found that the real exchange rates among industrialized countries appear to be characterized by several nonlinear specifications. There are reasons to believe that the linear models, which assume that large deviations from PPP are corrected in the same manner as small deviations, are outdated; however some nonlinear specifications are well explained by recent theoretical models. For instance, transaction costs could give rise to a band of inactivity where arbitrage is not profitable, so that the real exchange rate deviations from purchasing power parity are not corrected inside the band. If the real exchange rate moves outside of the band, arbitrage works to bring the real exchange rate back to the edge of the band.[1] To capture this behavior, Obstfeld & Taylor (1997) estimate band-threshold autoregressive (*Band-TAR*) models and find significant evidence towards nonlinearity for the real exchange rates of a large number of industrialized countries. On the other hand, it is possible that aggregation and non-synchronous adjustment by heterogeneous agents will cause regime changes to be smooth rather than discrete even if they individually make dichotomous decisions. Under this assumption, Taylor, Peel, & Sarno (2001) conclude that the real exchange rates that

[1] Transaction costs can be broadly defined to include transportation costs, tariffs and non-tariff barriers, as well as any other costs that agents incur in international trade (Obstfeld & Rogoff, 2000). Dumas (1992), Uppal (1993), Sercu et al. (1995) and Coleman (1995) develop equilibrium models of real exchange rate determination which take into account transaction costs and show that adjustment of real exchange rates toward PPP is necessarily a nonlinear process.

they consider (UK, Germany, France and Japan real exchange rates against the dollar) are very well characterized by an exponential smooth threshold AR (*ESTAR*) model. They furthermore argue that the speed of adjustment of the real exchange rate should not vary according to whether the dollar is undervalued or overvalued.

As described above, each of the two nonlinear models "fits" some real exchange rates series especially well in sample. However, some questions remain un-answered: Would these non-linear models also provide an out of sample forecast which is superior to the random walk model and/or the linear model? Which one of these models provides a superior out of sample forecasting performance?

Our aim is to move a step forward from the in-sample estimations and compare the relative performance of the two nonlinear models concerning their ability to generate out of sample forecasts, when estimating the real exchange rate series.[2] We therefore estimate a Band-TAR and an ESTAR model for 20 OECD countries real exchange rate series and then compute 4 years out-of-sample forecasts. We are aware of the size distortions of the Diebold-Mariano-West statistics as our competing models are nested and we calculate bootstrap critical values to correct this bias. Our main findings are twofold: First, we find strong evidence that the ESTAR model outperforms the random walk model and that neither the linear model nor the Band-TAR model significantly outperforms the random walk model, when forecasting out of sample. Hence we conclude that using the "out-of-sample" criteria, Purchasing Power Parity hypothesis is verified if estimating the real exchange rate as a nonlinear process, specifically an ESTAR model. Second, we find that bootstrapping critical values leads not only to important gains in

[2] This paper is only concerned with analyzing the possible nonlinear processes caused by transaction costs when estimating and forecasting the real exchange rate series. We thus do not investigate other potentially important characteristics of their data generating process as heteroscedasticity and/or structural breaks.

terms of the size properties but also to a significant loss in power to almost no power of the DMW test when testing the predictive accuracy of the nonlinear models versus the linear model; we are therefore not able to draw any conclusion on which models, the nonlinear or the linear, provide a better forecasting performance, when using a correctly sized DMW test. [3]

The rest of the paper is organized as follows. Part 2 presents in detail the data and the model estimation method. Part 3 describes the out of sample forecast method and part 4 illustrates the econometric tests we use to evaluate the models' predictive accuracy. Part 5 presents the empirical results and findings and part 6 concludes this study.

2. Nonlinear Models

2.1 Data

We use monthly nominal exchange rates and CPI data to calculate the real exchange rate series. The data was obtained from the *International Financial Statistics* database. It covers a set of 20 OECD countries with US as the base country. These countries are: UK, France, Italy, Spain, Japan, Korea, Austria, Belgium, Denmark, Finland, Greece, Luxembourg, Netherlands, Norway, Portugal Sweden, Switzerland, Turkey, Canada and Mexico. The monthly real exchange rate data starts from 1973:1 and ends by 2006:6 for all the countries.[4]

---

[3] Rapach and Wohar (2006) have previously analyzed the out-of-sample forecasting performance of the two nonlinear models (Band-TAR and ESTAR) of U.S. dollar real exchange rate behavior for four countries for the post- Bretton Woods. They find that the nonlinear models do not outperform the simple linear autoregressive models in terms of their out-of-sample forecasting performance.

[4] For the European Union countries we use the fixed exchange rate between the currency of each country against euro times the euro-dollar exchange rate to calculate the nominal exchange rate between each Euro country and US since 1998.

Under the hypothesis of Purchasing Power Parity (PPP), the real exchange rate displays long-run mean reversion. The real dollar exchange rate is calculated as follows:

$$q = e + p^* - p, \tag{1}$$

where q is the logarithm of the real exchange rate, e is the logarithm of the nominal exchange rate (the dollar price of the foreign currency) and p and p^* are the logarithms of the US and the foreign price levels, respectively.

2.2 Band-TAR model

Although there is a large array of regime switching models, we will consider specifically the band threshold autoregressive model (Band-TAR) as it has been extensively used when modeling the real exchange rate. Most recent nonlinear research assumes that "iceberg" transportation costs create a band for the real exchange rate within which the marginal cost of arbitrage exceeds the marginal benefit. The Band-TAR model used by Obstfeld & Taylor (1997) is characterized by unit-root behavior in an inner regime and reversion to the edge of the unit-root band in an outer regime. Their Band-TAR model takes the form:

$$\Delta y_t = \begin{cases} \alpha_t^{out}\,(y_{t-1} - \tau) + \varepsilon_t^{out} & if \quad y_{t-1} > \tau; \\ \varepsilon_t^{in} & if \quad \tau \geq y_{t-1} \geq -\tau \\ \alpha_t^{out}(y_{t-1} + \tau) + \varepsilon_t^{out} & if \quad -\tau > y_{t-1}; \end{cases} \tag{2}$$

where ε^{out} is N(0, σ^{out^2}), ε^{in} is N(0, σ^{in^2}), τ is the value of the threshold. We have previously assessed that almost all series can fit pretty well an AR (1) series using Schwarz criteria, therefore we use only one lag.

We implement the maximum likelihood estimation through a grid search over possible threshold values and delay parameters. Chan (1993) shows that a grid search over all potential values of the thresholds yields a superconsistent estimate of the unknown threshold parameter, τ. To use the method, we order the absolute value of observations from smallest to largest such that:

$$y^1 < y^2 < y^3 .... < y^T \tag{3}$$

Each value of y^j is then allowed to serve as an estimate of the threshold τ. For each of these values, the Heaviside indicator is set and estimated using equation 2. Since we have decided to use only one lag for our model estimations, we assume that the delay parameter is also not larger than 1.[5] The regression equation with the smallest residual sum of squares contains the consistent estimate of the threshold τ. We follow the conventional practice of excluding the highest and lowest 15% of the absolute value of y^j values to ensure an adequate number of observations in each regime.

The Band-TAR model allows us to estimate the value of the threshold without imposing a priori line of demarcation between the regimes. The key feature of these models is that a sufficiently large shock can cause the system to switch between regimes. The dates at which the series crosses the threshold are not specified beforehand by the researcher.

2.3. ESTAR model

In contrast to the discrete regime switching that characterizes the Band-TAR model, the exponential STAR (ESTAR) model proposed by Granger and Terasvirta

[5] We also estimate our model by choosing the delay parameter d from 1 to 3 and we find that the selection of d will not cause substantial changes in our results.

(1993) allows for smooth adjustments, so that the speed of adjustment varies with the extent of the deviation from parity. Taylor, Peel, and Sarno (2001) argue that time aggregation and non-synchronous adjustment by heterogeneous agents leads to smooth regime switching.

We use the following parsimonious ESTAR model:

$$y_t = y_{t-1} - \left\{1 - \exp\left[-\alpha \cdot (y_{t-1} - \tau)^2\right]\right\} \cdot (y_{t-1} - \tau) + \varepsilon_t \qquad (4)$$

where ε is N$(0, \sigma^2)$, and τ is the long-run equilibrium level of the series. The real exchange rate behaves as a random walk in the inner regime when $y_{t-1} = \tau$ and there is not too much incentive for arbitrage in the market. The speed of the mean reversion increases gradually as the real exchange rate moves away from the long run equilibrium. We have previously determined, using the Schwarz criteria, that the optimal lag length is 1 for all countries, therefore we will consider only one lag of the real exchange rate. We implement the nonlinear least squares estimation by setting the delay parameter equal to 1.[6] Our experiments using different starting values for the parameters yield similar results, indicating the location of a global optimum.

3. Out-of-sample forecasts

We estimate each series of real exchange rates as a linear autoregressive model (AR(1)), a random walk model (RW) and nonlinear models (Band-TAR and ESTAR), as described in the previous section. Following, we proceed to compute the 48 step forecasts

[6] We also grid search both, the threshold and the delay parameter, when estimating the ESTAR model and the results are similar.

from the linear, the random walk, and the non-linear models.[7] For the AR(1) or RW models, the multi-period forecasting is straightforward because they are linear. On the other hand, forecasting the nonlinear models (Band-TAR and ESTAR), because of their conditional expectations of future innovations, was a nontrivial task. As analyzed in Koop, Pesaran, and Potter (1996) and later in Enders (2004), the forecasts from a nonlinear model are state-dependent. For a model with one lag, we select a particular history of y_{t-1}. Since there is a possibility of regime switching, the multi-step-ahead forecasts from Band-TAR and ESTAR models are more difficult to calculate. To employ Koop, Pesaran, and Potter's methodology, we select 48 randomly drawn realizations of the residuals of the estimated non-linear model. Because the residuals may not have a normal distribution, they are selected using standard ''bootstrapping'' procedures. In particular, the residuals are drawn with replacement using a uniform distribution. Each residual drawn here are multiplied by a random number drawn from a standard normal distribution. We call these residual products $\varepsilon^*_{t+1}, \varepsilon^*_{t+2}, ..., \varepsilon^*_{t+48}$. We then generate y^*_{t+1} through y^*_{t+48} by substituting these ''bootstrapped'' residuals into equation 2 or 4. For this particular history, we repeat the process 1000 times. Under very weak conditions, the Law of Large Numbers (see Koop, Pesaran, and Potter 1996) guarantees that the sample average of the 1000 values of y^*_{t+1} converges to the conditional mean of y_{t+1} denoted by $E_t y_{t+1}$. Similarly, the Law of Large Numbers guarantees that the sample means of the various y^*_{t+i} converge to the true conditional i-step ahead forecasts,

[7] Following Rogoff 's (1994) argument that the real exchange rates can take from 3 to 5 years to converge to a constant mean (Rogoff's puzzle), we use 48 step monthly forecasts to determine if either the nonlinear or the linear model outperform the random walk model when forecasting the real exchange rates series.

$$\lim_{N\to\infty}\left[\sum_{k=1}^{N} y_{t+i}^{*}(k)/N\right]=E_t y_{t+i} \qquad (5)$$

The essential point is that the sample averages of y_{t+1}^{*} through y_{t+48}^{*} yield the one-step through 48-step ahead conditional forecasts of the real exchange series.

4. Comparative performance of the recursive forecasts

In this section we consider expanding-window regressions to obtain multi-step-ahead forecasts from every estimated model. We estimate the parameters of each model using all observations from the start of the series through 1982:12. We repeat the estimation process by adding successive observations through 2002:6. Next, we compute 48-step forecasts from all the forecasting origins, from 1983:1 to 2002:7. At the end of this exercise there are 235 out-of-sample 1-step through 48-step forecasts for each series.

The forecasts are used to obtain the mean square prediction error (MSPE) of the nonlinear, linear and random walk models, for each series at each forecasting horizon. We assess the significance of our evaluation by employing a test statistics for forecasting accuracy.

There is a large amount of research focused on model selection and estimation. Diebold, Mariano (1995) and West (1996) (henceforth DMW) proposed a statistics to test the null hypothesis of equal predictive ability against the one-sided alternative hypothesis. This test allows doe different variants of the loss function and for non-Gaussian, nonzero mean, serially correlated and contemporaneously correlated forecast errors. While Harvey, Leyboure and Newbold (1997) noticed that the DMW statistic is oversized for multi-step ahead forecast, McCracken (2004) and Clark and McCracken (2001, 2003) have shown that DMW statistic is severely undersized when the models are nested. Since our models are nested and we use multi-step ahead forecasts, these arguments are potential concerns for us. Several papers have proposed different adjustments to the DMW statistics to correct these size problems.[8] However, there is no evidence that any of these adjustments would properly correct the size problems when one of the nested models is nonlinear.

To evaluate the performance of the DMW statistics when one of the competing models is nonlinear (in our case Band-TAR and ESTAR) we run two Monte Carlo simulation experiments. We generate the following data process:

$$y_t = \alpha y_{t-1} + \varepsilon_t \quad (6)$$

The level of persistence is measured by the autoregressive coefficient α =1 and 0.98 and the residuals are drawn randomly from a normal distribution with a standard deviation of 0.03.[9] The sample size is 402 (the same length as the actual data) and the process is repeated 1000 times. First, when α =1, the data generated being a random walk process, we asses how well the test finds that the two non-linear models (Band-TAR and ESTAR) and the linear model have a superior predictive ability against the random walk model. Second, when α =0.98, the data generated being a linear process, we asses how well the test finds that two non-linear models (Band-TAR and ESTAR) have a superior

[8] Harvey, Leyboure and Newbold (1997) correct the oversizing problem of DMW statistic for multi-step ahead forecast by multiplying an adjustment term to DMW statistics (MDM statistics). Clark and West (2005) argue that the undersizing of DMW statistic, when two competing models are nested, is mainly caused by noises introduced by the alternative model. They try to fix this problem by proposing a method to adjust the MSPE in order to remove these noises. Their corrections work for models which are both nested and 'smooth' (i.e. twice continuously differentiable conditional mean) covariance stationary.

[9] The persistence level and the standard errors of the residuals are based on an average of the estimated coefficients when estimating the real exchange rates as a linear (or random walk) process for all the countries in our sample.

predictive ability against the linear model.[10] We report results for a 10% nominal size and we investigate the relationship between the empirical and the nominal test size.

As shown in Table 1, we find that DMW statistics is generally undersized at a short forecasting horizon and oversized at a long forecasting horizon. The degree of over sizing in the long run varies from model to model. For instance, for Band-TAR versus RW, Band-TAR versus Linear and ESTAR versus Linear, the DMW statistics is severely oversized by the end of the forecasting period, while for ESTAR versus RW and Linear versus RW the DMW statistics is almost correctly sized by the end of the forecasting period. [11]

We correct this size distortion by calculating parametric bootstrap critical values. The data generating process is constructed as follows. (1) construct the pseudo data by using the AR(1) model with the estimated coefficient; (2) add the artificial residuals which are randomly selected from a normal distribution with the estimated standard error. [12] The random walk generating process is constructed similarly. Following we estimate the above described models and compute the DMW statistics. We repeat this process 1000 times. The critical values for all the models are presented in Table 2.

We have shown that the size distortions of DMW test can be corrected by the use of appropriately adjusted critical values. The issue becomes the loss in power implied by

[10] In other words, we analyze the size of the test, specifically how well the test finds that that the nonlinear models and the random walk/linear model have an equal predictive ability when the data generating processes is a random walk/linear process.

[11] Clark & McCracken (2005) argue that standard normal critical values provide reliable inference when the forecasting horizon is relatively short, and the proportion between the sample size and the forecasting horizon size is quite small. Once the forecasting horizon increases beyond a few periods, neither a standard normal approximation nor the asymptotic distribution yields reliable inference in finite samples, the bootstrap methods being much more reliable.

[12] Instead of generating different critical value for every country, we use one set of coefficients to generate the pseudo linear or random walk data. We estimate the true data first as a random walk/linear model and then save the average standard error of residuals and the coefficient to form our basic assumptions of the relative coefficients. This simplicity does not cause substantial changes in the results and it significantly shortens the computation time.

such gains in the size performance of the test. In order to investigate the size adjusted power of the DMW test we build experiments with artificial data under a true alternative hypothesis where the data process follows an ESTAR or a Band-TAR process. We next examine whether these nonlinear models have a superior predictive ability against the random walk or the linear model. [13]

Within Monte Carlo experiments we consider the data generating processes from equation 2 (Band-TAR model) and equation 4 (ESTAR model). As previously, the generated data is based on the specifications found in our real exchange rate series: we estimate the true exchange rate data for each country as an ESTAR and a Band-TAR model and we use the average coefficient and standard deviation estimated for all countries to generate an ESTAR and a Band-TAR data generating process in our Monte Carlo experiment. For the above ESTAR model we use the coefficients: τ = -0.05, α = -0.42 and the standard error ε_t =0.03 and for the Band-TAR model we use the coefficient: τ=0.15, α_t^{out} =-0.12, the standard errors ε_t^{out} =0.04 and ε_t^{in} = 0.03.

The power of a test is normally analyzed by tabulating how often the null is rejected when it is false. In the table 3 we present the results for the power of the test by tabulating how often the test selects the true nonlinear model. The sample size is 402 (the same length as the actual real exchange rate data we are using) and the process is repeated 1000 times. We use the bootstrap critical values shown in Table 2 and report results for a 10% nominal size. The two nonlinear models are estimated as shown in equation 2 (Band-TAR model) and equation 4 (ESTAR model).

[13] Liu and Prodan (2007) conduct a more extensive study on the power of tests for comparative predictive ability when bootstrapping critical values.

Our Monte Carlo experiments show that the size adjusted power of the DMW test to distinguish between nonlinear (both ESTAR and Band-TAR) and random walk models is moderate: at almost all forecasting horizons the size adjusted power is about 30%. The size adjusted power of the DMW test to distinguish between nonlinear models (both ESTAR and Band-TAR) and linear models is very low: in the case of the ESTAR model the power varies between 5%-8% at longer forecasting horizons and 10% at shorter forecasting horizons; in the case of the Band-TAR model the power varies from 9% at longer forecasting horizons to 15-20% at shorter forecasting horizons.

Since the size adjusted power of the DMW test is very low when testing the nonlinear models versus the linear model, we will not consider further testing the comparative predictive ability of these models. On the other hand, since the DMW's size adjusted power is moderate when distinguishing between the nonlinear and the random walk models, we will next consider only testing the predictive ability between the nonlinear models and the random walk model.

5. Out-of-sample performance of the nonlinear models

The appendix presents our detail forecasting and evaluating results of all countries. For each country, the DMW statistic is computed for Band-TAR vs RW, ESTAR vs RW, and Linear vs RW from step 1 to step 48 forecast. For each pair of the models we report the DMW statistics and the rejections at the 10% significance level.

To describe our methodology of choosing the better fitting model we first examine in detail France's case and subsequently present the results for all countries.

For the case of France we first compare first the out of sample forecasting behavior of the nonlinear model versus the random walk model. For the Band-TAR against the RW model, the DMW statistics is not significant at any step at 10% significance level which implies that the Band-TAR model does not outperform the RW model. For ESTAR against the RW model, the DMW statistics is significant at long horizon (last two years), which provides strong evidence that the ESTAR model outperforms the RW model in long run. We then compare the out of sample forecast behavior of the linear model versus the random walk model and we find that the DMW statistics is not significant at any forecasting horizon. Overall, we can conclude that for France there is strong evidence that in long run the ESTAR model provides a better forecasting performance than the RW model. On the other hand the Band-TAR model and the linear models do not outperform the random walk model. Following a similar process, we have analyzed all the other countries. The results are shown in the Table 4.

We first investigate whether the nonlinear models provide a better forecasting performance than the RW model. Among 20 analyzed countries we find strong evidence that for 14 countries the ESTAR model outperforms the RW model: for 13 countries the ESTAR model outperform the RW model only in the long horizon and for one country the ESTAR model outperform the RW model at all steps. The countries where we did not find any evidence towards ESTAR model are Japan, Korea, Sweden, Turkey, Canada and Mexico[14]. On the other hand we find very little evidence towards the Band-TAR model: only in 5 out of 20 cases, Band-TAR provides a better out of sample forecasting performance than the random walk model at a few forecasting steps (only in the middle).

[14] For some of these countries we find that the mean square prediction error of the ESTAR model is smaller than the mean square prediction error of the RW models in the long horizon. According to the bootstrapping critical value of the DMW statistic, the difference is not significant at 10% level.

The above findings provide some evidence that real exchange rate is characterized by a nonlinear behavior, and the transitions between regimes are smooth rather than discrete.

Second, we investigate whether the linear models provide a better forecasting performance than the RW model: among 20 analyzed countries we find that for only 4 countries the linear model outperforms the RW model, at very long horizons.

As we have shown in this paper, the low power of tests for predictive ability when bootstrapping critical values makes it difficult to compare the performance of the nonlinear models versus the linear models.[15] Even though previous literature finds that generally the nonlinear models do not outperform the linear model when forecasting out of sample, it also provides several explanations for this: One reason is that the 'non-linearity' might fail to persist into the future (e.g., Granger and Terasvirta, 1993) so the lack of forecast gain of non-linear models over linear models might be due to the fact that a more complicated model (the nonlinear model) may be hard to identify and estimate with precision. Second, it is not obvious that features of nonlinear time series such as heteroskedasticity, structural break or outliers will result in improved forecasts compared to ones from linear models. As Clements and Hendry (2001) argue, an incorrect but simple model may outperform a correct model in forecasting. Third, Diebold and Nason (1990) argue that the nonlinearities may be present in even-ordered conditional moments, and therefore are not useful for point prediction and very slight conditional-mean nonlinearities might be truly present and be detectable with large datasets, while nevertheless yielding negligible ex ante forecast improvement. Finally, Liu and Prodan

[15] We test the comparative predictability of the nonlinear versus the linear models, but we are not able to reject the equal predictive ability of these models for any country. We do not report those results as the lack of rejections might be due to the lack of power of the DMW test and we do not find those results as being conclusive.

(2007) perform a series of size-adjusted power simulations of bootstrap tests for comparative predictability and argue that these tests have very little power when the null is a highly persistent nonlinear series.

As a result, there are good reasons to believe that the real exchange rates show evidence of a nonlinear, smooth adjustment to Purchasing Power Parity in the long horizon.

6. Conclusion

There is a large amount of research that focused on assessing the validity of Purchasing Power Parity. Several studies failed to find any evidence of PPP when analyzing the post Bretton Woods era industrialized countries' real exchange rates. Assuming that the continuous and constant assumption of the linear models does not apply to real exchange rates, the focus has moved toward estimating them as nonlinear processes. The conclusion was that several nonlinear models "fit" some particular real exchange rates series especially well, leading to good in-sample properties.

We examine two nonlinear methods (Band-TAR and ESTAR models) concerning their ability to generate out of sample forecasts, when estimating real exchange rate series. We find a significant amount of evidence towards the ESTAR model.

Our results can be summarized as follows: 1) Generally, the ESTAR model provides a better forecasting performance than the RW model 2) The linear model and the Band-TAR model does not outperform the RW model. 3) We are not able to assess the forecasting performance of the nonlinear models versus the linear model as the size adjusted power of the DMW test, in this specific case, is very low to almost no power.

We therefore conclude that using the “out-of-sample” criteria, Purchasing Power Parity hypothesis is verified if estimating real exchange as a nonlinear process, specifically an ESTAR model.

References

Balke, Nathan S & Fomby, Thomas B, 1997. Threshold Cointegration. International Economic Review, vol. 38(3), pages 627-45, August.

Chan, K. S, 1993. Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model. The Annals of Statistics, 21(1), 520-533.

Clark, Todd E., 2004. Can Out-of-Sample Forecast Comparisons Help Prevent Overfitting? Journal of Forecasting. Volume 23.2, 115-139

Clark, Todd E. & West, Kenneth D., 2005. Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. Journal of Econometrics, forthcoming

Clark, Todd E. & West, Kenneth D., 2005. Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis. Journal of Econometrics, forthcoming

Clark, Todd E. & McCracken, Michael W.,2001. Evaluating Long-horizon Forecasts. Research Working Paper RWP 01-14, Federal Reserve Bank of Kansas City.

Clark, Todd E. & McCracken, Michael W., 2003. The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence. Computing in Economics and Finance 2003 183, Society for Computational Economics.

Clark, Todd E. & McCracken, Michael W., 2005. Evaluating direct multi-step forecasts. Econometric Reviews, Taylor and Francis Journals, vol. 24(4), pages 369-404, October.

Clements, M.P. & D.F. Hendry, 2001. Explaining the Results of the M3 forecasting Competition. International Journal of Forecasting, 17, 550-554.

Coleman A, 1995. Arbitrage, Storage and the Law of One Price: New Theory for the Time Series Analysis of an Old Problem. Ph.D Dissertation, Princeton University.

Dick van Dijk &Philip Hans Franses, 2003. Selecting a Nonlinear Time Series Model using Weighted Tests of Equal Forecast Accuracy. Oxford Bulletin of Economics and Statistics, 65, Supplement 0305-9049

Diebold, Francis X. & Roberto S. Mariano, 1995. Comparing Predictive Accuracy. Journal of Business & Economic Statistics, Vol. 13, No. 3. (Jul., 1995), pp. 253-263.

Diebold, Francis X. & James Nason, 1990. Nonparametric exchange rate prediction? Journal of International Economics 28, 315-332.

Dumas B.,1992. Dynamic Equilibrium and the Real Exchange-Rate in a Spatially Separated World. Review of Financial Studies 5 (2): 153-180 1992

Dumas B.,1994. Some Models of the International Capital-Market European Economic Review 38 (3-4): 923-931 APR 1994

Enders, Walter, 2004, Applied Econometric Time Series. 2nd ed. John Wiley and Sons: Hoboken, N.J.

Granger, C. & T. Terasvirta,1993. Modeling Nonlinear Economic Relationships. Oxford University Press, Oxford

Harvey, David; Stephen Leybourne and Paul Newbold, 1997. Testing the equality of prediction mean squared errors. International Journal of Forecasting 13 (1997) 281-291

Inoue, A. & L. Kilian, 2004. In-sample or Out-of-sample Tests of Predictability: Which One Should We Use? Forthcoming Econometric Reviews.

Koop, Gary, M. Hashem Pesaran & Simon Potter,1996. Impulse Response Analysis Nonlinear Multivariate Models. Journal of Econometrics ,74, 119–147.

Liu, Yamei and Walter Enders, 2003. Out-of -Sample Forecasts and Nonlinear Model Selection with an Example of the Term Structure of the Interest Rates. Southern Economic Journal 2003, 69(3), 520-540.

Liu, Yu and Ruxandra Prodan, 2007. The performance of bootstrap tests for equal predictive accuracy, working paper.

McCracken, Michael W., 2004. Parameter Estimation and Tests of Equal Forecast Accuracy Between Non-nested Models. International Journal of Forecasting, Elsevier, vol. 20(3), pages 503-514.

Obstfeld, Maurice and Alan M. Taylor, 1997. Nonlinear Aspects of Goods Market Arbitrage and Adjustment. Journal of the Japanese and International Economies 11, 441–479

Obstfeld, Maurice & Kenneth Rogoff, 2001. The Six Major Puzzles in International Macroeconomics: Is There a Common Cause? International Trade 0012003, EconWPA

Rapach, David E. & Mark E. Wohar, 2006. The Out-of-Sample Forecasting Performance of Nonlinear Models of Real Exchange Rate Behavior. International Journal of Forecasting, Vol. 22, No. 2 (April-June 2006), pp. 341–361

Sercu, P., Uppal, R. & Van Hulle, C., 1995. The Exchange Rate in the Presence of Transaction Costs: Implications for Tests of Purchasing Power Parity. Journal of Finance, 50, 1309-1319.

Taylor, Mark, David A. Peel and Lucia Sarno, 2001. Nonlinear Mean Reversion in Real Exchange Rates: Toward a Solution to the Purchasing Power Parity Puzzles. International Economic Review 42, No 4.

Teräsvirta, T., 1994. Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models. Journal of the American Statistical Association 89, 208-218.

Uppal R., 1993. A General Equilibrium Model of International Portfolio Choice. Journal of Finance, 48.2, 529-553.

West, Kenneth D. 1996. Asymptotic Inference about Predictive Ability. Econometrica, Vol. 64, No. 5. Sep. pp. 1067-1084.

Table 1: The Size Performance of the DMW Statistics (Nominal Size = 10%)

Step	Band-TAR vs Lin	Band-TAR vs RW	Lin vs RW	ESTAR vs RW	ESTAR vs Lin
1	19%	9%	2%	3%	5%
2	30%	16%	3%	5%	8%
3	37%	19%	4%	6%	11%
6	55%	27%	6%	8%	18%
9	64%	33%	7%	8%	23%
12	71%	36%	8%	10%	26%
18	76%	37%	8%	11%	31%
24	78%	39%	10%	11%	31%
30	78%	40%	11%	11%	30%
36	79%	39%	11%	12%	30%
42	79%	41%	12%	13%	30%
48	79%	41%	13%	13%	30%

Note:

1. Steps 1 to 48 indicate the forecast horizon (months-ahead).
2. The DGP is either a random walk or a linear process and the estimated models are Band-TAR, ESTAR and Linear, depending on the case.

Table 2: Bootstrapped Critical Values of DMW Statistics at 10% significance level

Step	Band-TAR vs RW	ESTAR vs RW	Linear vs RW	Band-TAR vs Linear	ESTAR vs Linear
1	1.24	0.42	0.42	1.63	0.87
2	1.59	0.55	0.55	1.98	1.13
3	1.77	0.69	0.61	2.33	1.32
6	2.15	1.02	0.69	2.96	1.79
9	2.38	1.09	0.77	3.44	2.12
12	2.50	1.23	0.89	3.84	2.35
18	2.71	1.34	1.12	4.20	2.48
24	2.73	1.44	1.24	4.32	2.48
30	2.72	1.36	1.40	4.37	2.40
36	2.66	1.57	1.51	4.01	2.40
42	2.76	1.65	1.51	3.61	2.29
48	2.67	1.65	1.65	3.41	2.27

Note:

1. Steps 1 to 48 indicate the forecast horizon (months-ahead).

Table 3: Size Adjusted Power of DMW Statistics at 10% significance level

Step	Band-TAR vs RW	ESTAR vs RW	Linear vs RW	Band-TAR vs Linear	ESTAR vs Linear
1	0.23	0.23	0.24	0.21	0.11
2	0.24	0.25	0.23	0.21	0.10
3	0.27	0.24	0.24	0.19	0.11
6	0.31	0.24	0.27	0.20	0.11
9	0.27	0.26	0.29	0.16	0.09
12	0.28	0.27	0.30	0.15	0.08
18	0.27	0.29	0.29	0.15	0.08
24	0.28	0.32	0.30	0.11	0.08
30	0.28	0.35	0.30	0.11	0.07
36	0.31	0.32	0.29	0.09	0.06
42	0.29	0.32	0.29	0.09	0.08
48	0.29	0.32	0.30	0.10	0.08

Note:

1. See the notes to Table 1.
2. The power is calculated based on to the critical values presented in Table 2.

Table 4:	DMW Statistic Results of France								
	ESTAR vs RW			Band-TAR vs RW			Linear vs RW		
step	**DMW**	**cv-10%**	**sig-10%**	**DMW**	**cv-10%**	**sig-10%**	**DMW**	**cv-10%**	**sig-10%**
1	-1.62	0.42		-1.75	1.24		-0.78	0.42	
2	-1.31	0.55		-1.32	1.59		-0.79	0.55	
3	-1.36	0.69		-0.79	1.77		-0.93	0.61	
6	-1.29	1.02		-0.87	2.15		-0.89	0.69	
9	-0.48	1.09		-0.59	2.38		-0.96	0.77	
12	0.01	1.23		-0.01	2.50		-0.88	0.89	
18	0.65	1.34		1.08	2.71		-0.83	1.12	
24	1.44	1.44	*	1.76	2.73		-0.94	1.24	
30	1.73	1.36	*	5.09	2.72	*	-0.79	1.40	
36	1.92	1.57	*	5.69	2.66	*	-0.43	1.51	
42	2.03	1.65	*	2.36	2.76		-0.32	1.51	
48	2.31	1.65	*	2.54	2.67		-0.19	1.65	

Note:

1. 1. Steps 1 to 48 indicate the forecast horizon (months-ahead).
2. DMW statistic for the null hypothesis that the null model MSPE equals the alternative model MSPE against the alternative hypothesis that the null model MSPE is greater than the alternative model MSPE. "*" indicates that the null hypothesis is rejected at 10% significant level.

Table 5:	**Comparative Performance**		
	Estar vs RW	**Bandtar vs RW**	**Linear vs RW**
UK	Y-all	N	Y-end
France	Y-end	Y-mid	N
Italy	Y-end	N	Y-end
Spain	Y-mid,end	N	N
Japan	N	Y-mid	N
Korea	N	N	N
Austria	Y-end	N	Y-end
Belgium	Y-end	N	N
Denmark	Y-end	N	N
Finland	Y-end	N	N
Greece	Y-end	N	N
Luxembourg	Y-end	Y-mid	N
Netherlands	Y-end	N	N
Norway	Y-end	N	N
Portugal	Y-end	N	N
Sweden	N	Y-mid	N
Switzerland	Y-end	N	Y-end
Turkey	N	Y-mid	N
Canada	N	N	N
Mexico	N	N	N

Note:

1. "Y-beg" - DMW statistics is significant at 10% level in the beginning of the forecasting horizon, "Y - mid" - DMW statistics is significant at 10% level in the middle of the forecasting horizon, "Y-end" - DMW statistics is significant at 10% level in the end of the forecasting horizon, "Y-all" - DMW statistics is significant at 10% level at all forecasting horizons.

Appendix

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
UK	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	0.50	0.42	*	(0.91)	1.24		0.16	0.42	
2	0.91	0.55	*	(1.36)	1.59		0.26	0.55	
3	0.99	0.69	*	(1.44)	1.77		0.23	0.61	
6	1.17	1.02	*	(1.47)	2.15		0.14	0.69	
9	1.14	1.09	*	(1.40)	2.38		0.27	0.77	
12	2.05	1.23	*	(1.31)	2.50		0.77	0.89	
18	1.30	1.34		(1.01)	2.71		1.55	1.12	*
24	1.70	1.44	*	(1.30)	2.73		1.50	1.24	*
30	1.99	1.36	*	(1.06)	2.72		2.81	1.40	*
36	2.05	1.57	*	(1.08)	2.66		4.84	1.51	*
42	2.08	1.65	*	(1.10)	2.76		3.05	1.51	*
48	2.11	1.65	*	(1.11)	2.67		3.60	1.65	*
FRANCE	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.62)	0.42		(1.75)	1.24		(0.78)	0.42	
2	(1.31)	0.55		(1.32)	1.59		(0.79)	0.55	
3	(1.36)	0.69		(0.79)	1.77		(0.93)	0.61	
6	(1.29)	1.02		(0.87)	2.15		(0.89)	0.69	
9	(0.48)	1.09		(0.59)	2.38		(0.96)	0.77	
12	0.01	1.23		(0.01)	2.50		(0.88)	0.89	
18	0.65	1.34		1.08	2.71		(0.83)	1.12	
24	1.44	1.44	*	1.76	2.73		(0.94)	1.24	
30	1.73	1.36	*	5.09	2.72	*	(0.79)	1.40	
36	1.92	1.57	*	5.69	2.66	*	(0.43)	1.51	
42	2.03	1.65	*	2.36	2.76		(0.32)	1.51	
48	2.31	1.65	*	2.54	2.67		(0.19)	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
ITALY	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(2.05)	0.42		(2.02)	1.24		(1.44)	0.42	
2	(1.72)	0.55		(1.34)	1.59		(1.38)	0.55	
3	(1.66)	0.69		(1.03)	1.77		(1.33)	0.61	
6	(0.88)	1.02		(0.73)	2.15		(1.35)	0.69	
9	(0.32)	1.09		(0.99)	2.38		(1.61)	0.77	
12	0.24	1.23		(0.65)	2.50		(1.09)	0.89	
18	0.95	1.34		0.54	2.71		0.25	1.12	
24	1.49	1.44	*	1.37	2.73		0.60	1.24	
30	1.59	1.36	*	2.27	2.72		2.92	1.40	*
36	2.05	1.57	*	1.50	2.66		1.36	1.51	
42	1.86	1.65	*	1.65	2.76		1.56	1.51	*
48	1.97	1.65	*	1.18	2.67		1.83	1.65	*
SPAIN	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	0.03	0.42		(0.22)	1.24		(0.96)	0.42	
2	0.24	0.55		(0.40)	1.59		(1.10)	0.55	
3	0.32	0.69		(0.05)	1.77		(1.09)	0.61	
6	0.04	1.02		(0.35)	2.15		(1.57)	0.69	
9	0.56	1.09		(0.28)	2.38		(1.21)	0.77	
12	2.00	1.23	*	0.29	2.50		(0.32)	0.89	
18	2.65	1.34	*	0.38	2.71		0.11	1.12	
24	1.72	1.44	*	0.42	2.73		0.23	1.24	
30	1.79	1.36	*	0.63	2.72		0.83	1.40	
36	1.94	1.57	*	0.49	2.66		0.52	1.51	
42	2.21	1.65	*	0.67	2.76		0.80	1.51	
48	2.49	1.65	*	0.72	2.67		1.06	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
JAPAN	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(2.46)	0.42		(0.67)	1.24		(1.41)	0.42	
2	(2.53)	0.55		(0.28)	1.59		(1.42)	0.55	
3	(2.51)	0.69		(0.06)	1.77		(1.40)	0.61	
6	(3.37)	1.02		0.54	2.15		(0.92)	0.69	
9	(3.61)	1.09		0.80	2.38		(0.29)	0.77	
12	(3.25)	1.23		1.29	2.50		(0.01)	0.89	
18	(1.12)	1.34		3.05	2.71	*	0.14	1.12	
24	(0.73)	1.44		3.99	2.73	*	0.37	1.24	
30	(0.46)	1.36		1.96	2.72		0.68	1.40	
36	(0.22)	1.57		1.88	2.66		0.70	1.51	
42	(0.14)	1.65		1.94	2.76		0.68	1.51	
48	(0.25)	1.65		1.82	2.67		1.61	1.65	
KOREA	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(0.15)	0.42		0.22	1.24		(1.55)	0.42	
2	0.52	0.55		1.13	1.59		(1.26)	0.55	
3	(0.17)	0.69		1.02	1.77		(1.66)	0.61	
6	(0.33)	1.02		1.20	2.15		(1.12)	0.69	
9	(0.19)	1.09		1.22	2.38		(0.91)	0.77	
12	(0.01)	1.23		1.71	2.50		(0.98)	0.89	
18	0.21	1.34		1.97	2.71		(1.24)	1.12	
24	0.74	1.44		2.07	2.73		(0.44)	1.24	
30	1.05	1.36		1.89	2.72		(0.14)	1.40	
36	1.06	1.57		2.64	2.66		0.44	1.51	
42	1.59	1.65		2.24	2.76		0.76	1.51	
48	2.21	1.65	*	2.71	2.67	*	1.25	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
AUSTRIA	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.17)	0.42		(2.31)	1.24		(0.84)	0.42	
2	(0.87)	0.55		(1.52)	1.59		(0.69)	0.55	
3	(0.77)	0.69		(1.27)	1.77		(0.65)	0.61	
6	(0.32)	1.02		(1.59)	2.15		(1.04)	0.69	
9	0.02	1.09		(1.70)	2.38		(0.99)	0.77	
12	0.34	1.23		(1.34)	2.50		(0.59)	0.89	
18	0.83	1.34		(0.42)	2.71		0.29	1.12	
24	2.04	1.44	*	0.04	2.73		0.73	1.24	
30	1.60	1.36	*	0.55	2.72		2.83	1.40	*
36	1.73	1.57	*	0.62	2.66		1.75	1.51	*
42	1.87	1.65	*	0.78	2.76		1.99	1.51	*
48	1.88	1.65	*	0.53	2.67		2.17	1.65	*
BELGIUM	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.63)	0.42		(1.61)	1.24		(0.43)	0.42	
2	(1.37)	0.55		(1.44)	1.59		(0.40)	0.55	
3	(1.24)	0.69		(1.30)	1.77		(0.46)	0.61	
6	(0.74)	1.02		(0.52)	2.15		(0.59)	0.69	
9	(0.36)	1.09		(0.02)	2.38		(0.75)	0.77	
12	0.13	1.23		0.75	2.50		(0.77)	0.89	
18	0.85	1.34		1.24	2.71		(0.95)	1.12	
24	1.20	1.44		2.72	2.73		(2.01)	1.24	
30	1.71	1.36	*	1.77	2.72		(0.87)	1.40	
36	1.98	1.57	*	1.94	2.66		(0.82)	1.51	
42	2.06	1.65	*	2.59	2.76		(0.80)	1.51	
48	2.23	1.65	*	2.32	2.67		(0.73)	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
DENMARK	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.16)	0.42		(1.83)	1.24		(0.50)	0.42	
2	(1.07)	0.55		(1.38)	1.59		(0.30)	0.55	
3	(0.86)	0.69		(1.08)	1.77		(0.25)	0.61	
6	(0.53)	1.02		(0.89)	2.15		(0.55)	0.69	
9	0.03	1.09		(0.70)	2.38		(0.43)	0.77	
12	0.41	1.23		(0.56)	2.50		(0.40)	0.89	
18	0.88	1.34		(0.15)	2.71		(0.25)	1.12	
24	1.62	1.44	*	0.65	2.73		(0.32)	1.24	
30	1.73	1.36	*	1.72	2.72		(0.09)	1.40	
36	1.86	1.57	*	2.25	2.66		0.06	1.51	
42	1.94	1.65	*	2.35	2.76		0.22	1.51	
48	1.98	1.65	*	2.39	2.67		0.43	1.65	
FINLAND	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(2.29)	0.42		(1.38)	1.24		(1.37)	0.42	
2	(2.19)	0.55		(1.30)	1.59		(1.31)	0.55	
3	(2.27)	0.69		(1.19)	1.77		(1.25)	0.61	
6	(2.22)	1.02		(1.25)	2.15		(1.53)	0.69	
9	(1.80)	1.09		(1.05)	2.38		(0.80)	0.77	
12	(0.45)	1.23		0.04	2.50		(0.87)	0.89	
18	0.34	1.34		2.13	2.71		0.11	1.12	
24	1.05	1.44		1.64	2.73		0.96	1.24	
30	1.32	1.36		1.82	2.72		2.00	1.40	*
36	1.47	1.57		2.06	2.66		1.03	1.51	
42	1.66	1.65	*	2.45	2.76		1.08	1.51	
48	1.91	1.65	*	2.86	2.67	*	1.34	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
GREECE	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.46)	0.42		(0.45)	1.24		(0.94)	0.42	
2	(1.46)	0.55		(0.26)	1.59		(0.64)	0.55	
3	(1.37)	0.69		(0.48)	1.77		(0.53)	0.61	
6	(1.22)	1.02		(0.61)	2.15		(1.03)	0.69	
9	(0.74)	1.09		(0.42)	2.38		(0.74)	0.77	
12	(0.36)	1.23		(0.21)	2.50		(0.53)	0.89	
18	0.10	1.34		0.08	2.71		(0.64)	1.12	
24	0.67	1.44		0.55	2.73		(0.14)	1.24	
30	1.40	1.36	*	1.16	2.72		0.30	1.40	
36	2.25	1.57	*	3.25	2.66	*	0.76	1.51	
42	2.30	1.65	*	2.10	2.76		1.11	1.51	
48	2.32	1.65	*	2.33	2.67		1.37	1.65	
LUXEMBOURG	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.47)	0.42		(1.60)	1.24		(0.43)	0.42	
2	(1.33)	0.55		(1.30)	1.59		(0.38)	0.55	
3	(1.29)	0.69		(1.31)	1.77		(0.41)	0.61	
6	(0.95)	1.02		(0.46)	2.15		(0.55)	0.69	
9	(0.35)	1.09		0.23	2.38		(0.67)	0.77	
12	0.07	1.23		0.64	2.50		(0.80)	0.89	
18	0.59	1.34		1.36	2.71		(0.97)	1.12	
24	1.09	1.44		2.90	2.73	*	(2.04)	1.24	
30	1.68	1.36	*	4.97	2.72	*	(2.28)	1.40	
36	1.94	1.57	*	2.14	2.66		(0.86)	1.51	
42	2.21	1.65	*	2.24	2.76		(0.84)	1.51	
48	2.34	1.65	*	2.58	2.67		(0.78)	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
NETHERLANDS	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.30)	0.42		(1.54)	1.24		(0.61)	0.42	
2	(1.35)	0.55		(1.21)	1.59		(0.44)	0.55	
3	(1.18)	0.69		(1.26)	1.77		(0.51)	0.61	
6	(0.66)	1.02		(1.34)	2.15		(0.71)	0.69	
9	(0.34)	1.09		(0.72)	2.38		(0.87)	0.77	
12	(0.03)	1.23		0.14	2.50		(1.06)	0.89	
18	0.49	1.34		1.13	2.71		(0.61)	1.12	
24	1.25	1.44		1.83	2.73		(0.99)	1.24	
30	1.78	1.36	*	2.17	2.72		(0.33)	1.40	
36	5.35	1.57	*	2.33	2.66		(0.08)	1.51	
42	2.20	1.65	*	2.47	2.76		0.14	1.51	
48	2.37	1.65	*	2.56	2.67		0.33	1.65	
NORWAY	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(2.30)	0.42		(1.50)	1.24		(0.48)	0.42	
2	(1.80)	0.55		(1.20)	1.59		(0.51)	0.55	
3	(1.62)	0.69		(1.10)	1.77		(0.52)	0.61	
6	(0.75)	1.02		(1.23)	2.15		(0.81)	0.69	
9	(0.32)	1.09		(1.04)	2.38		(1.29)	0.77	
12	0.07	1.23		(0.65)	2.50		(1.21)	0.89	
18	0.53	1.34		0.61	2.71		(0.55)	1.12	
24	1.09	1.44		1.00	2.73		(0.32)	1.24	
30	1.65	1.36	*	0.97	2.72		0.01	1.40	
36	1.88	1.57	*	0.53	2.66		0.24	1.51	
42	2.03	1.65	*	0.43	2.76		0.42	1.51	
48	2.18	1.65	*	(0.09)	2.67		0.70	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
PORTUGAL	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(1.55)	0.42		0.19	1.24		(1.12)	0.42	
2	(1.54)	0.55		(0.25)	1.59		(1.15)	0.55	
3	(1.70)	0.69		(0.41)	1.77		(1.31)	0.61	
6	(1.53)	1.02		(0.95)	2.15		(1.12)	0.69	
9	(0.54)	1.09		(0.65)	2.38		(1.40)	0.77	
12	(0.23)	1.23		(0.07)	2.50		(0.54)	0.89	
18	0.36	1.34		(0.06)	2.71		(0.15)	1.12	
24	1.21	1.44		(0.28)	2.73		0.05	1.24	
30	1.51	1.36	*	(0.42)	2.72		0.42	1.40	
36	1.73	1.57	*	(0.45)	2.66		0.86	1.51	
42	1.69	1.65	*	(0.49)	2.76		1.00	1.51	
48	2.51	1.65	*	(0.53)	2.67		1.17	1.65	
SWEDEN	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(2.25)	0.42		(0.60)	1.24		(1.20)	0.42	
2	(2.17)	0.55		(0.04)	1.59		(1.32)	0.55	
3	(2.06)	0.69		0.19	1.77		(1.58)	0.61	
6	(1.22)	1.02		0.28	2.15		(1.35)	0.69	
9	(0.93)	1.09		0.31	2.38		(1.49)	0.77	
12	(0.73)	1.23		1.06	2.50		(1.50)	0.89	
18	0.32	1.34		1.37	2.71		(1.58)	1.12	
24	0.97	1.44		4.10	2.73	*	(2.08)	1.24	
30	1.30	1.36		2.76	2.72	*	(1.28)	1.40	
36	2.16	1.57	*	6.36	2.66	*	(1.35)	1.51	
42	1.43	1.65		2.40	2.76		(1.25)	1.51	
48	1.45	1.65		2.55	2.67		(1.21)	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
SWITZERLAND	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(0.20)	0.42		(1.57)	1.24		(0.58)	0.42	
2	(0.17)	0.55		(1.60)	1.59		(0.42)	0.55	
3	(0.10)	0.69		(1.46)	1.77		(0.28)	0.61	
6	0.13	1.02		(1.70)	2.15		(0.32)	0.69	
9	0.32	1.09		(1.40)	2.38		(0.05)	0.77	
12	0.63	1.23		(0.74)	2.50		0.27	0.89	
18	1.40	1.34	*	0.72	2.71		0.83	1.12	
24	1.60	1.44	*	1.49	2.73		1.45	1.24	*
30	1.50	1.36	*	1.76	2.72		2.15	1.40	*
36	1.50	1.57		1.69	2.66		2.05	1.51	*
42	1.53	1.65		1.50	2.76		2.23	1.51	*
48	1.66	1.65	*	0.68	2.67		2.00	1.65	*
TURKEY	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(4.46)	0.42		0.97	1.24		(1.66)	0.42	
2	(3.53)	0.55		1.17	1.59		(1.38)	0.55	
3	(3.44)	0.69		1.40	1.77		(1.41)	0.61	
6	(3.23)	1.02		1.93	2.15		(1.31)	0.69	
9	(1.62)	1.09		1.94	2.38		(0.66)	0.77	
12	(2.03)	1.23		2.59	2.50	*	(0.61)	0.89	
18	(2.54)	1.34		2.07	2.71		(0.33)	1.12	
24	(1.08)	1.44		1.86	2.73		(0.15)	1.24	
30	(0.78)	1.36		3.55	2.72	*	0.07	1.40	
36	(0.44)	1.57		4.28	2.66	*	0.22	1.51	
42	(0.17)	1.65		2.00	2.76		0.50	1.51	
48	0.01	1.65		2.08	2.67		0.71	1.65	

	ESTAR vs RW			BANDTAR vs RW			LIN vs RW		
CANADA	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(2.48)	0.42		(1.46)	1.24		(1.70)	0.42	
2	(3.05)	0.55		(1.45)	1.59		(2.00)	0.55	
3	(3.29)	0.69		(1.60)	1.77		(2.20)	0.61	
6	(4.22)	1.02		(1.56)	2.15		(2.49)	0.69	
9	(3.01)	1.09		(1.20)	2.38		(1.44)	0.77	
12	(2.20)	1.23		(0.33)	2.50		(1.67)	0.89	
18	(2.71)	1.34		0.91	2.71		(1.46)	1.12	
24	(1.14)	1.44		1.14	2.73		(0.78)	1.24	
30	(0.36)	1.36		1.10	2.72		(0.58)	1.40	
36	(0.07)	1.57		1.11	2.66		(0.49)	1.51	
42	0.23	1.65		1.28	2.76		(0.43)	1.51	
48	0.34	1.65		1.52	2.67		(0.46)	1.65	
MEXICO	dm1	%fract90	0.10	dm1	%fract90	0.10	dm1	%fract90	0.10
1	(2.73)	0.42		(1.75)	1.24		(1.56)	0.42	
2	(2.99)	0.55		(1.69)	1.59		(1.72)	0.55	
3	(3.04)	0.69		(1.54)	1.77		(1.80)	0.61	
6	(1.47)	1.02		(1.08)	2.15		(1.62)	0.69	
9	(1.25)	1.09		(0.41)	2.38		(1.43)	0.77	
12	(0.98)	1.23		0.41	2.50		(0.76)	0.89	
18	0.32	1.34		1.20	2.71		0.16	1.12	
24	0.45	1.44		1.47	2.73		0.37	1.24	
30	0.32	1.36		1.18	2.72		0.31	1.40	
36	0.24	1.57		0.96	2.66		0.28	1.51	
42	0.33	1.65		0.98	2.76		0.38	1.51	
48	0.69	1.65		1.29	2.67		0.75	1.65	

Note:

1. See the notes to Table 1.
2. "*" indicates that the alternative model is better than the null model at 10% significant level.
3. DMW statistic for the null hypothesis that the null model MSFE equals the alternative model MSFE against the alternative hypothesis that the null model MSFE is greater than the alternative model MSFE.