Nonlinear Exchange Rate Predictability

Carlos Felipe Lopez-Suarez and Jose Antonio Rodriguez-Lopez[*]

First version: May 2007
Revised: September 2010

Abstract

We study whether the nonlinear behavior of the real exchange rate can help us account for the lack of predictability of the nominal exchange rate. We construct a smooth nonlinear error-correction model that allows us to test the hypotheses of nonlinear predictability of the nominal exchange rate and nonlinear behavior on the real exchange rate in the context of a fully specified cointegrated system. Using a panel of 19 countries and three numeraires, we find evidence of nonlinear predictability of the nominal exchange rate and of nonlinear mean reversion of the real exchange rate. Out-of-sample Theil's U-statistics show a higher forecast precision of the nonlinear model than the one obtained with a random walk specification. Although the robustness of the out-of-sample results over different forecast windows is somewhat limited, we are able to obtain significant predictability gains—from a parsimonious structural model with PPP fundamentals—even at short-run horizons.

Keywords: exchange rates, predictability, nonlinearities, purchasing power parity.
JEL codes: C53, F31, F47.

---

1   Introduction

During the last three decades a growing amount of literature has shown a poor empirical relation between economic fundamentals and the exchange rate. In their seminal work, Meese and Rogoff (1983) show that the relation between economic fundamentals and the nominal exchange rate is so weak that the short-term forecast of the future behavior of the exchange rate is usually outperformed by naive random walk specifications. As documented in the work of Cheung, Chinn, and Garcia Pascual (2005), Diebold and Nason (1990) and Meese and Rose (1991), this puzzling behavior—commonly denoted as lack of exchange rate predictability—has remained through most of the recent floating period.

The finding of Mark (1995) that an empirical significant relation between economic fundamentals and the exchange rate exists at long horizons, from 1 to 4 years, has also been brought into question. Berben and van Dijk (1998), Kilian (1999), and Berkowitz and Giorgianni (2001) show that the results of high and significant predictability in long horizons can be explained as lack of cointegration between fundamentals and the exchange rate. Furthermore, according to Faust, Rogers, and Wright (2003), Mark's original results have largely disappeared as a result of data revisions and data accumulation through time.

There are several interpretations one can give to these findings. The most common one is that of a dismal performance of our economic models. A second and more recent interpretation comes from Engel and West (2005), who show that under certain conditions the standard linear models of exchange rates and fundamentals behave as (near) random walks. A third interpretation—and more related to this paper—is that linear models provide a poor approximation to the behavior of economic time series and that richer and more general statistical models are needed. Indeed, several theoretical models that take into account the existence of fixed and variable transaction costs across countries conclude that the mean reversion of the real exchange rate should be nonlinear.

Using the purchasing power parity (PPP) model, in this paper we study whether smooth nonlinearities can account for the lack of exchange rate predictability. Building on the work of Granger and Swanson (1996), we construct a generalized cointegrated system that nests the possibilities of having predictability or unpredictability of the nominal exchange rate, as well as linear or nonlinear behavior on the real exchange rate.

For comparison purposes, we follow Mark and Sul (2001) and estimate the model for three

different numeraires—U.S., Japan and Switzerland—using a panel dataset of 19 countries in the post-Bretton Woods period. We perform t-tests of the predictability and nonlinear mean reversion hypotheses and construct Theil's U-statistics that measure the relative out-of-sample accuracy of the nonlinear forecast of the nominal exchange rate versus the random walk. Across numeraires, the t-tests reveal significant nonlinear behavior of the real exchange rate and nonlinear predictability of the nominal exchange rate, while the U-statistics show evidence of an improved forecast precision of the nonlinear error-correction model relative to the random walk. Contrary to recent results by Engel, Mark, and West (2007) and others, our out-of-sample results are significant for many countries at short-term horizons.

Following Rogoff and Stavrakeva (2008), we verify the robustness of the out-of-sample results over different forecast windows. Using a joint test based in the average U-statistic, as Mark and Sul (2001), we find that our nonlinear model significantly outperforms the random walk over different forecast windows for one-quarter-ahead forecasts when the U.S. is the numeraire, and for four- and eight-quarter-ahead forecasts when either Japan or Switzerland is the numeraire. Nevertheless, for individual exchange rates we find that the nonlinear model consistently (and significantly) dominates the random walk over different forecast windows for only four countries (out of 18) in the one-quarter-ahead U.S. case, for nine countries in the four-quarter-ahead Japan case, and for four countries in the four-quarter-ahead Switzerland case. It is also important to mention that in general, the predictability of the nominal exchange rate breaks down in forecast windows starting by the end of 2008 and during 2009. This result is, however, not surprising as it coincides with the most tense period of the recent (unprecedented) financial crisis. In that period of such worldwide economic instability, it would be naive to expect that exchange rates follow traditional economic fundamentals.

In our out-of-sample analysis, we choose the driftless random walk as our benchmark model. Hence, in order to verify that we are comparing our nonlinear model against the best of the naive specifications, we provide an extensive out-of-sample comparison of the random walk with drift versus the random walk without drift. We show that for each of the three numeraires, the driftless random walk is generally a better predictor than the random walk with drift.

The paper is organized as follows. In Section 2 we discuss the theoretical and empirical arguments in favor of a nonlinear specification of the real exchange rate. In Section 3 we show how to construct the smooth nonlinear error-correction model based on a nonlinear specification

of the real exchange rate and the existence of a cointegrated system. We will start with a simple standard linear model and then build up step by step into our more general nonlinear model. In Section 4 we estimate the smooth error-correction model and then compare its forecast accuracy against the random walk. In Section 5 we perform the robustness analysis of the out-of sample results. Finally, Section 6 concludes.

2 Nonlinearities of the Real Exchange Rate

The idea that the mean reversion of the real exchange rate may be nonlinear has both theoretical and empirical motivations. Theoretical support can be found in the work of Benninga and Protopapadakis (1988), Dumas (1992), Sercu, Uppal, and Van Hulle (1995), O'Connell and Wei (2002) and Ohanian and Stockman (1997). The standard argument is that the existence of fixed and/or variable transactions costs across countries generates a positive relation between the magnitude of PPP deviations and the degree of mean reversion. When the real exchange rate is close to its equilibrium level, the difference in effective prices across countries is close to zero, and the real exchange can freely move in any direction according to random shocks in the economy. However, when the real exchange rate deviation from its equilibrium level gets larger, the difference in effective prices across countries gets also larger, and there will be a higher degree of mean reversion due to the existence of arbitrage trade among countries.

For example, let us consider the simple continuous time model with two countries and one good of O'Connell and Wei (2002). In the presence of only iceberg transaction costs, they show the existence of an equilibrium where the real exchange rate is confined between two reflecting barriers that delimit a range of no-arbitrage. Whenever the real exchange rate gets outside the reflecting barriers, there is a minimal amount of trade that rebounds the real exchange rate into the closest reflecting barrier. If there are fixed costs rather than iceberg costs, the authors find that instead of two reflecting barriers there would be two resetting barriers in equilibrium. Whenever the real exchange rate hits one of the resetting barriers, trade is conducted in an amount sufficiently large to get the difference in effective prices across countries equal to zero. Having together both iceberg and fixed costs in the model generates an equilibrium with four barriers. Whenever the real exchange rate hits the outer barriers, it is instantly reset by trade to the closest inner barrier. The bottom line of these models is that we should expect a higher degree of mean reversion in the real exchange rate when the difference in effective prices across

3

countries is larger.

There are two branches of empirical literature studying the predictions of these models. The first branch focuses on cross-country differences in effective prices of single goods or commodities. These studies—e.g., Obstfeld and Taylor (1997) and Imbs, Mumtaz, Ravn, and Rey (2003)—use highly nonlinear statistical specifications and find strong evidence of nonlinearities. A common specification in this kind of papers is the Threshold Autoregressive (TAR) model. Letting y_t denote the difference in effective prices of a single commodity across two countries, the TAR specification is given by

$$\Delta y_t = \begin{cases} \lambda^{\text{out}}(y_{t-1} + c) + \varepsilon_t & \text{if } y_{t-1} < -c \\ \lambda^{\text{in}} y_{t-1} + \varepsilon_t & \text{if } -c \leq y_{t-1} \leq c \\ \lambda^{\text{out}}(y_{t-1} - c) + \varepsilon_t & \text{if } y_{t-1} > c, \end{cases}$$

where $c > 0$.

Note that this statistical specification closely resembles the predictions of the model of O'Connell and Wei (2002) with iceberg transaction costs. When the difference in effective prices is small (i.e., less than c), there is no mean reversion in the price differential, as represented with the a priori selection of a random walk model with $\lambda^{\text{in}} = 0$. When instead the difference in prices gets large (i.e., larger than c), we should expect a large degree of mean reversion towards the reflecting barriers according to the estimated parameter $\lambda^{\text{out}} < 0$. In general, these studies conclude that highly nonlinear models such as TAR provide a good description of the behavior of single goods price differentials and find estimates of mean reversion, λ^{out}, larger in absolute value than the ones obtained using standard linear models.

The second branch of the empirical literature focuses on the study of cross-country differences in effective prices of large bundles of commodities, e.g., deviations from PPP—measured with the Consumer Price Index (CPI). In these papers it is argued that aggregation in goods and time generates a smooth nonlinear mean reversion rather than a one point threshold autoregression specification (see Dumas (1994) and Taylor and Peel (2000)). The statistical model that has been used in this literature—e.g., Taylor, Peel, and Sarno (2001) and Kilian and Taylor (2003)— is the Smooth Transition Autoregressive (STAR) model of Granger and Teräsvirta (1993) and Teräsvirta (1994). Here we consider the parsimonious representation proposed by Kilian and

Taylor (2003) in which

$$z_t = \exp(\gamma z_{t-1}^2)\left[\rho_1 z_{t-1} + \rho_2 z_{t-2}\right] + u_{1t}, \tag{1}$$

where z_t denotes the logarithm of the real exchange rate measured in terms of the CPI and u_{1t} is white noise. As Kilian and Taylor, let us impose the restriction that $\rho = \rho_1 = 1 - \rho_2$. The statistical properties of this specification are very intuitive. Provided $\gamma < 0$, the degree of mean reversion of the real exchange rate is a smooth function on the level of the log real exchange rate. In order to see this, note that we can interpret equation (1) as a smooth transition between two extreme statistical models. On the one hand, whenever $z_{t-1} \to 0$, we have a statistical model that tends to $z_t = \rho z_{t-1} + (1-\rho) z_{t-2} + u_{1t}$ so that the real exchange rate behaves as a random walk with no mean reversion. On the other hand, whenever $|z_{t-1}| \to \infty$, we have a log real exchange rate that tends in the limit to white noise (given by $z_t = u_{1t}$) and complete mean reversion is present. In general, we will have a level of mean reversion between these two extreme points. As $|z_{t-1}|$ goes from zero to infinity, we will have a degree of mean reversion that smoothly increases from no mean reversion to complete mean reversion.

Taylor, Peel, and Sarno (2001) and Kilian and Taylor (2003) show that the STAR model provides a good description of the behavior of the real exchange rate. They show that the linearity hypothesis, $H_0 : \gamma = 0$, is rejected for most industrialized countries and suggest that the model can account for several of the PPP puzzles in the literature, including the lack of power in standard unit root tests and the high persistence of small shocks in the real exchange rate around equilibrium (see Rogoff (1996)).

In this paper we study whether the Smooth Transition model can also account for the lack of predictability of the nominal exchange rate observed in standard linear models. We will do this by including equation (1) in a generalized cointegrated system following the procedure suggested by Granger and Swanson (1996). The resulting model will be a Smooth Transition Error-Correction model where the mean reversion of the nominal exchange rate toward its economic fundamentals will be a smooth function of the effective price differential across countries.

3 The Smooth Transition Error-Correction Model

It is a common approach in the empirical literature to directly specify a cointegrated system either in terms of the error-correction model or in the common trend representation of Stock

and Watson (1988). Granger and Swanson (1996) argue that we can generalize a cointegrated system if we instead specify the model in terms of the cointegrated variables and the stochastic trends using the following specification

$$z_t = \alpha' x_t \tag{2}$$

$$w_t = \lambda' x_t, \tag{3}$$

where x_t is a $n \times 1$ vector of nonstationary variables, z_t is a $r \times 1$ vector of cointegrated variables, with $r < n$, and w_t is a $(n - r) \times 1$ I(1) vector, such as

$$w_t = w_{t-1} + \eta_t. \tag{4}$$

The triangular representation of Phillips is a special case of this specification. Consider for example the case at hand with $x_t = (s_t \ f_t)'$, where s_t is the logarithm of the nominal exchange rate—defined as the local price of one unit of the numeraire's currency—and f_t represents the economic fundamentals. In the PPP model, economic fundamentals are defined as

$$f_t = p_t - p_t^*, \tag{5}$$

where p_t and p_t^* are the log prices at the local and the numeraire country, respectively.

Assuming that the cointegrated variable z_t—the logarithm of the real exchange rate—follows an AR(2) process, the triangular representation is given either by

$$z_t = s_t - f_t, \qquad z_t = \rho_1 z_{t-1} + \rho_2 z_{t-2} + u_{1t} \tag{6}$$

$$\Delta s_t = \eta_t \tag{7}$$

or

$$z_t = s_t - f_t, \qquad z_t = \rho_1 z_{t-1} + \rho_2 z_{t-2} + u_{1t} \tag{8}$$

$$\Delta f_t = \eta_t, \tag{9}$$

where u_{1t} and η_t are possibly correlated white noise processes. Under the assumption that w_t is given by equation (4), both representations are consistent with the specification of equations (2)

and (3) since $\alpha = (1 \ -1)'$ and $\lambda = (1 \ 0)'$ or $\lambda = (0 \ 1)'$ generate equations (6)-(7) and (8)-(9), respectively.

A drawback of the triangular representation is that it assumes a priori whether the nominal exchange rate is predictable or not. In equation (7) we state that the nominal exchange rate is a random walk, while in equations (8) and (9) we state that the exchange rate is predictable. The exact form of the predictability implied by equations (8) and (9) is derived by computing the corresponding error-correction model. Taking first differences in both equations and solving the system in terms of Δs_t we obtain that

$$\Delta s_t = (\rho_1 - 1)z_{t-1} + \rho_2 z_{t-2} + (u_{1t} + \eta_t), \tag{10}$$

where the future movement in the nominal exchange rate depends on previous realizations of the real exchange rate.

In order to circumvent an a priori assumption about the predictability of the nominal exchange rate, we follow Granger and Swanson (1996) and instead of assuming a λ equal to either $(0 \ 1)'$ or $(1 \ 0)'$, we assume a λ equal to $(\lambda_1 \ \lambda_2)'$, where $\lambda_1, \lambda_2 \in \mathbb{R}$. It can be shown by taking first differences to equations (2) and (3) that when $\lambda = (\lambda_1 \ \lambda_2)'$ and $\alpha = (1 \ -1)'$, the cointegrated system has a vector error-correction representation given by

$$\Delta s_t = -r \left(1 - \rho_1 - \rho_2\right) z_{t-1} - r\rho_2 \Delta z_{t-1} + u_{2t}, \tag{11}$$

where $r = \frac{\lambda_2}{\lambda_1 + \lambda_2}$ and $u_{2t} = \frac{\lambda_2 u_{1t} + \eta_t}{\lambda_1 + \lambda_2}$. The predictability of the nominal exchange rate can now be tested using $H_0 : r = 0$ versus $H_A : r > 0$. As long as $r > 0$, we have a system in which the mean reversion of the real exchange rate towards its equilibrium relies, at least partially, in movements of the nominal exchange rate. On the other side, when $r = 0$ the mean reversion of the real exchange rate relies exclusively on future movements in the difference in prices.

It is now straightforward to generalize the linear error-correction representation into a non-linear error-correction representation. Let the cointegrated system be determined by equations (2) and (3) with $\alpha = (1 \ -1)'$, $\lambda = (\lambda_1 \ \lambda_2)'$, and the real exchange rate be given by the STAR specification of equation (1). Taking first differences of equations (2) and (3) and solving the

system for Δs_t, we obtain the Smooth Transition Error-Correction (STEC) model

$$\Delta s_t = -r \left[1 - (\rho_1 + \rho_2) \exp(\gamma z_{t-1}^2)\right] z_{t-1} - r\rho_2 \exp(\gamma z_{t-1}^2)\Delta z_{t-1} + u_{2t}. \tag{12}$$

This equation is a generalization of the linear error-correction equation (11) in the sense that the linear model is just the STEC model restricted to $\gamma = 0$. Using the restriction $\rho = \rho_1 = 1 - \rho_2$ of the STAR specification of Kilian and Taylor (2003), the STEC model is reduced to the parsimonious specification

$$\Delta s_t = -r \left[1 - \exp(\gamma z_{t-1}^2)\right] z_{t-1} - r(1 - \rho) \exp(\gamma z_{t-1}^2)\Delta z_{t-1} + u_{2t}. \tag{13}$$

This equation is the building block of our work.[1] The parameter r maintains the same role as in the linear model. An $r = 0$ implies that mean reversion in the real exchange rate relies exclusively in the difference in prices, while an $r > 0$ implies that mean reversion in the real exchange rate relies at least partially in the nominal exchange rate. Then, we can test the hypothesis of predictability of the nominal exchange rate using $H_0 : r = 0$ versus $H_A : r > 0$. The degree of nonlinearity in the mean reversion of the real exchange rate depends on γ. Given $\gamma < 0$, the mean reversion of the nominal exchange rate towards its fundamentals will be increasing in the absolute value of the difference in effective prices across countries—i.e., the error-correction parameter of equation (13) is an increasing function of $|z_{t-1}|$. Also, the larger the γ (in absolute value), the higher the degree of nonlinearities. Hence, we can test for nonlinearities using $H_0 : \gamma = 0$ versus $H_A : \gamma < 0$.

4 Estimation of the STEC Model

4.1 The Data

We estimate the STEC model using the same panel of countries as Mark and Sul (2001) but with an updated sample. Hence, we obtain quarterly time series for nominal exchange rates and price levels from 1973.Q1 through 2009.Q4 for 19 countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Great Britain, Greece, Italy, Japan, Korea, the

---

[1] Note that even though the STEC model in equation (12) nests a linear model, our parsimonious STEC model in equation (13) does not. This result follows because an error-correction representation cannot exist when $\gamma = 0$ and $\rho = \rho_1 = 1 - \rho_2$ since the real exchange rate would not be stationary and the system would not be cointegrated.

Netherlands, Norway, Spain, Sweden, Switzerland and the United States.[2] All series are from the IMF's International Financial Statistics and are measured at the end of the quarter. Prices correspond to CPI levels and are seasonally adjusted with a one-side moving average of the present observation and three lagged variables. Nominal exchange rates are reported as the price of one U.S. dollar.[3]

4.2 Econometric Specification and Estimation Results

Our econometric specification is given by

$$z_{i,t} = \exp(\gamma z_{i,t-1}^2)\left[\rho z_{i,t-1} + (1-\rho)z_{i,t-2}\right] + u_{1t}^i \tag{14}$$

$$\Delta s_{i,t} = -r\left[1 - \exp(\gamma z_{i,t-1}^2)\right]z_{i,t-1} - r(1-\rho)\exp(\gamma z_{i,t-1}^2)\Delta z_{i,t-1} + u_{2t}^i, \tag{15}$$

where $i = 1, 2, ..., L$ indexes countries and $(u_{1t}^i \ u_{2t}^i)'$ are independent and identically distributed processes with variance-covariance matrix given by $\Sigma_{2L \times 2L}$. As in Mark and Sul (2001), we estimate the model for three numeraires: the United States, Japan and Switzerland. For each numeraire, we define $s_{i,t}$ as the log of the country i's currency price of one unit of the currency of the numeraire, and $z_{i,t}$ as the demeaned log of the real exchange rate between country i and the numeraire. We use a demeaned measure of the log of the real exchange rate to account for the different factors that cause relative purchasing power parity deviations in equilibrium—e.g., Harrod-Balassa-Samuelson effects. We abstract from cointegration tests.[4] Since the model is nonlinear, there are not closed form solutions for the estimators and numerical methods are needed. We estimate the model by constrained maximum likelihood with $\gamma < 0$.

Note that we restrict γ, ρ, and r to be equal across country pairs. These restrictions in the parameter space will provide us with two significant gains. First, they will increase the precision of our estimates by increasing the size of the effective sample used for the estimation of the parameters. Second, they will allow us to substantially reduce the time needed for the nonlinear estimation of the model, making the out-of-sample bootstraps of the following sections

---

[2]Mark and Sul (2001) estimate their linear model of nominal exchange rates for both monetary and PPP fundamentals. As we focus on nonlinearities of the real exchange rate, our estimation is based exclusively on PPP fundamentals.

[3]As in Engel, Mark, and West (2007) and Rogoff and Stavrakeva (2008), starting in 1999 the exchange rates for the Euro area countries differ only by a constant of proportionality.

[4]Mark and Sul (2001) show that nominal exchange rates and PPP fundamentals are cointegrated in their version of this dataset (see their Table 1). See also the references cited therein regarding the stationarity of the real exchange rate.

Table 1: STEC Model Estimation Results

Numeraire		Coefficient	t-stat	p-value
	γ	-0.456	-11.250	0.000
US	ρ	1.014	128.700	0.000
	r	0.833	51.530	0.000
	γ	-0.381	-18.640	0.000
Japan	ρ	0.934	58.570	0.000
	r	0.950	155.900	0.000
	γ	-1.409	-18.530	0.000
Switzerland	ρ	0.926	56.890	0.000
	r	0.957	228.900	0.000

For $\hat{\gamma}$, the p-value is the proportion of the bootstrap distribution (for the t-statistic) to the left of the calculated t-statistic. For $\hat{\rho}$ and $\hat{r}$, the p-value is the proportion of the bootstrap distribution (for the t-statistic) to the right of the calculated t-statistic.

computationally feasible.[5] These a priori restrictions in the parameters do not seem to come at a high cost. Allowing a different γ for each country pair generates similar results to those presented in this paper.[6] Moreover, unrestricted estimation of the model reveals that the estimates of r and ρ are remarkably similar across country pairs for each of the three numeraires.

Table 1 presents the estimation results. Note that under the null $\gamma = 0$, all the variables in the model are nonstationary and the distribution of $\hat{\gamma}$ is not standard. Therefore, the p-value we present for the estimate of γ for each numeraire corresponds to the bootstrapped p-value of the t-statistic under the null that γ equals zero.[7] The asymptotic distributions of $\hat{\rho}$ and $\hat{r}$ and their t-statistics are standard (so that we can interpret these statistics in the usual way). Nevertheless, in order to control for a possible small sample bias, we also obtain bootstrapped p-values for these estimators' t-statistics.

We find that across numeraires the hypotheses of linearity of the real exchange rate ($H_0 : \gamma = 0$) and unpredictability of the nominal exchange rate ($H_0 : r = 0$) are rejected even at a 1% level.

---

[5]The estimation of the unrestricted model requires up to 3 hours, while the estimation of the restricted model requires only a few seconds.

[6]A previous version of this paper included the estimation of the STEC model allowing for a different γ for each country pair. These results are available upon request.

[7]We perform one bootstrap for each numeraire. Each bootstrap is computed using the following procedure. First, we estimate the data generating process (DGP) under the null (given by equations (14) and (15), with the restriction $\gamma = 0$). Second, we construct 1,000 artificial datasets with $100+T$ observations—where T equals 148 and represents the number of quarterly observations in our sample period. We use zeros as initial values and build up the datasets using the recursive procedure of the DGP and independent draws from a multivariate normal distribution with variance-covariance matrix given by $\hat{\Sigma}_{2L \times 2L}$ (estimated with sample moments of the residuals of the DGP). Then, for each dataset we discard the first 100 observations, compute the t-statistic (using MLE) and count the number of times in which this statistic is lower than the real sample t-statistic. The bootstrapped p-value is just the result of this count divided by 1000.

All the estimates of r are relatively close to one, indicating that—no matter the numeraire—most of the mean reversion of the real exchange rate is conducted through future movements in the nominal exchange rate. In other words, we find that whenever the difference between the nominal exchange rate and its fundamentals is low, we should expect to see a nominal exchange rate that behaves much like a driftless random walk, while in those cases where the difference is large we should expect to see strong future adjustments on the nominal exchange rate toward its fundamentals. On the other hand, the estimates of γ are always negative—with values of about -0.38 and lower—and imply the existence of substantial nonlinear mean reversion on the real exchange rate.

4.3   Out-of-Sample Predictability

We now perform out-of-sample tests of the hypothesis that the nominal exchange rate is unpredictable, that is, $H_0 : r = 0$ versus $H_A : r > 0$. We consider bootstrapped out-of-sample tests based on Theil's U-statistic, where the U-statistic is computed as the ratio of the Root Mean Square Error (RMSE) of the out-of-sample forecast of the nonlinear model to the RMSE of the out-of-sample forecast of the driftless random walk.[8] As in Mark and Sul (2001), we also present a joint test of predictability by taking the mean value—across country pairs—of the individual U-statistics.

The root mean square errors of the STEC and random walk models used to compute the U-statistics are based on out-of-sample forecasts at short (one- and four-quarter-ahead) and long (eight- and sixteen-quarter-ahead) horizons for the sample period 1985.Q1-2009.Q4. The procedure to compute the out-of-sample forecasts is as follows. Let k represent the number of periods—i.e., quarters—ahead for the forecast. When $k = 1$, we estimate the models with data through 1984.Q4 and compute the forecasts for 1985.Q1. Then we estimate the models with data through 1985.Q1, compute the forecasts for 1985.Q2 and so on. For $k > 1$ we estimate the models with data through 1985.Q1$-k$ and then compute the forecasts for 1985.Q1. Once the forecasts for 1985.Q1 are constructed, we move up one quarter and estimate the models with data through 1985.Q2$-k$ in order to compute the forecasts for 1985.Q2 and so on up to 2009.Q4.

Let $\hat{s}_{i,t+k}$ denote the STEC model forecast for the nominal exchange rate k periods ahead.

---

[8]According to Rogoff and Stavrakeva (2008), bootstrapped out-of-sample tests such as the Theil's U and the Diebold-Mariano/West tests are "more powerful and better sized" than new asymptotic out-of-sample tests such as the Clark-West and Clark-McCracken tests.

From equation (15) we can see that the computation of the one-period-ahead STEC forecast is straightforward and given by

$$\hat{s}_{i,t+1} = s_{i,t} - \hat{r}\left[1 - \exp(\hat{\gamma}z_{i,t}^2)\right]z_{i,t} - \hat{r}(1 - \hat{\rho})\exp(\hat{\gamma}z_{i,t}^2)\Delta z_{i,t}. \tag{16}$$

However, the computation of STEC forecasts at longer horizons is more problematic since it requires the knowledge of the distribution function of the error u_1^i, denoted by $\Phi(u_1^i)$. For example, consider the case of the two-period-ahead forecast $\hat{s}_{i,t+2} = s_{i,t} + \widehat{\Delta s}_{i,t+1} + \widehat{\Delta s}_{i,t+2}$, where

$$\widehat{\Delta s}_{i,t+1} = -\hat{r}\left[1 - \exp(\hat{\gamma}z_{i,t}^2)\right]z_{i,t} - \hat{r}(1 - \hat{\rho})\exp(\hat{\gamma}z_{i,t}^2)\Delta z_{i,t} \tag{17}$$

$$\widehat{\Delta s}_{i,t+2} = \int \left(-\hat{r}\left[1 - \exp(\hat{\gamma}z_{i,t+1}^2)\right]z_{i,t+1} - \hat{r}(1 - \hat{\rho})\exp(\hat{\gamma}z_{i,t+1}^2)\Delta z_{i,t+1}\right)d\Phi(u_1^i), \tag{18}$$

given that

$$z_{i,t+1} = \exp(\gamma z_{i,t}^2)\left[\rho z_{i,t} + (1 - \rho)z_{i,t-1}\right] + u_{1t+1}^i.$$

Then, in order to get $\hat{s}_{i,t+2}$ we must first know $\Phi(u_1^i)$ so that we can compute the integral in equation (18).

There are several approaches to estimate this integral. Here we consider the parametric and nonparametric techniques presented in Granger and Teräsvirta (1993). The parametric approach estimates the integral by assuming normality in the distribution of the errors. For example, in our two-period-ahead forecast we have that

$$\widehat{\Delta s}_{i,t+2} = \frac{1}{J}\sum_{j=1}^{J}\left(-\hat{r}\left[1 - \exp(\hat{\gamma}\hat{z}_{i,t+1,j}^2)\right]\hat{z}_{i,t+1,j} - \hat{r}(1 - \hat{\rho})\exp(\hat{\gamma}\hat{z}_{i,t+1,j}^2)\Delta\hat{z}_{i,t+1,j}\right), \tag{19}$$

where

$$\hat{z}_{i,t+1,j} = \exp(\hat{\gamma}z_{i,t}^2)\left[\hat{\rho}z_{i,t} + (1 - \hat{\rho})z_{i,t-1}\right] + \hat{u}_{1t+1,j}^i,$$

and $\hat{u}_{1t+1,j}^i$ is one of J independent draws from a multivariate normal distribution with variance-covariance $\hat{\Sigma}_{2L\times 2L}$. The nonparametric approach follows a similar procedure but instead of assuming normality in the errors it resamples them from the residuals of the STEC model.

In this paper we use a J equal to 1000 and follow both the parametric and nonparametric approaches. In all cases the results are remarkably similar and our conclusions do not change.

In order to preserve space we present throughout the paper only the results based on the non-parametric approach.[9]

The data generating process (DGP) implied by the STEC model under the null that the nominal exchange rate is unpredictable (i.e., under $r = 0$) is given by

$$z_{i,t} = \exp(\gamma z_{i,t-1}^2)\left[\rho z_{i,t-1} + (1 - \rho)z_{i,t-2}\right] + u_{1t}^i \tag{20}$$

$$\Delta s_{i,t} = u_{2t}^i. \tag{21}$$

If the null is true, the driftless random walk model should be more accurate than the STEC model—the random walk should have a lower RMSE than the STEC model—and therefore the U-statistics should be larger than one. On the other hand, whenever the exchange rate is predictable (i.e., $r > 0$) we should have more forecast accuracy with the STEC model so that the U-statistics should be lower that one. The p-value of each U-statistic—including the average U-statistic—is calculated by bootstrapping its distribution under the DGP in equations (20) and (21).[10]

Tables 2 and 3 present the out-of-sample predictability results in the short and long run, respectively. The results are remarkable. Looking at the number of U-statistics smaller than one in both tables, we see that with a single exception (the one-quarter-ahead forecast with Switzerland as numeraire) the STEC model outperforms the driftless random walk for 15 or more exchange rates (out of 18). No matter the horizon, country pair, or numeraire considered, the STEC model is generally better.

As shown in Table 2, the statistical significance of the results is good in short-run horizons (one- and four-quarters ahead). In all cases, the joint test provides significance levels of 2.1% or lower for the better forecast accuracy of the STEC model. For individual exchange rates, the best out-of-sample predictability results of the paper come from four-quarter-ahead forecasts,

---

[9]Results based on the parametric approach are available upon request.

[10]We perform one bootstrap for each numeraire. The procedure for each bootstrap is as follows. First, we estimate the DGP under the null (given by equations (20) and (21), with the restriction $\gamma \leq 0$). Second, we construct 1,000 artificial datasets with $100 + T$ observations—where T equals 148 and represents the number of quarterly observations in our sample period. We use zeros as initial values and build up the datasets using the recursive procedure of the DGP and independent draws from a multivariate normal distribution with variance-covariance matrix given by $\hat{\Sigma}_{2L \times 2L}$ (estimated with sample moments of the residuals of the DGP). Then, for each dataset we discard the first 100 observations, compute the out-of-sample forecasts of the two competing models, their RMSE, and then the individual and average U-statistics. Finally, in order to compute each p-value we just count the number of times in which the U-statistic from the artificial samples is lower than the real sample U-statistic and then divide this number by 1000.

Table 2: Out-of-Sample Statistics of the STEC Model: Short-Run

Numeraire	United States				Japan				Switzerland			
Quarters ahead	1		4		1		4		1		4	
	U-stat	p-value	U-stat	p-value	U-stat	p-value	U-stat	p-value	U-stat	p-value	U-stat	p-value
Australia	0.988	0.107	**0.963**	0.093	0.999	0.362	0.963	0.120	**0.953**	0.000	**0.887**	0.000
Austria	**0.984**	0.039	**0.925**	0.007	0.995	0.213	**0.951**	0.051	1.003	0.631	**0.969**	0.058
Belgium	**0.981**	0.021	**0.923**	0.008	0.990	0.109	**0.951**	0.044	1.002	0.581	**0.970**	0.060
Canada	**0.978**	0.011	**0.956**	0.029	0.993	0.191	**0.931**	0.023	0.992	0.153	**0.894**	0.001
Denmark	**0.982**	0.023	**0.927**	0.009	**0.981**	0.024	**0.955**	0.059	1.005	0.746	0.981	0.155
Finland	**0.968**	0.004	**0.943**	0.024	**0.984**	0.037	**0.944**	0.033	**0.990**	0.085	**0.916**	0.000
France	**0.987**	0.065	**0.933**	0.013	**0.988**	0.077	**0.954**	0.055	1.001	0.552	**0.973**	0.074
Germany	**0.985**	0.051	**0.918**	0.005	0.996	0.271	**0.952**	0.057	1.006	0.768	0.976	0.109
Great Britain	**0.983**	0.029	**0.936**	0.016	0.993	0.207	**0.953**	0.064	0.995	0.199	**0.930**	0.004
Greece	1.024	0.880	1.034	0.768	1.011	0.775	1.044	0.849	1.053	0.983	1.085	0.956
Italy	**0.976**	0.011	**0.935**	0.017	**0.963**	0.001	**0.953**	0.050	0.991	0.103	**0.933**	0.003
Japan	0.995	0.247	**0.956**	0.077					**0.979**	0.008	**0.927**	0.002
Korea	0.989	0.123	**0.949**	0.061	1.006	0.496	**0.941**	0.053	**0.957**	0.000	**0.908**	0.006
Netherlands	**0.986**	0.050	**0.927**	0.011	0.995	0.247	**0.954**	0.061	1.003	0.642	0.983	0.175
Norway	0.990	0.103	**0.950**	0.046	1.002	0.499	0.967	0.145	**0.989**	0.064	**0.957**	0.042
Spain	**0.977**	0.011	**0.940**	0.025	**0.971**	0.003	**0.957**	0.073	**0.986**	0.044	**0.963**	0.059
Sweden	**0.965**	0.001	**0.918**	0.005	**0.966**	0.005	0.973	0.177	**0.984**	0.030	**0.945**	0.022
Switzerland	0.996	0.279	**0.927**	0.011	0.992	0.172	**0.956**	0.067				
United States					0.999	0.389	**0.913**	0.003	**0.982**	0.027	**0.875**	0.000
Average	**0.985**	0.004	**0.942**	0.001	**0.990**	0.016	**0.956**	0.013	**0.993**	0.021	**0.948**	0.000
Countries<1	17		17		15		17		11		17	

The p-value is the proportion of the bootstrap distribution to the left of the calculated U-statistic.
Bolded coefficients are statistically significant at a 10% level.

Table 3: Out-of-Sample Statistics of the STEC Model: Long-Run

Numeraire	United States				Japan				Switzerland			
Quarters ahead	8		16		8		16		8		16	
	U-stat	p-value	U-stat	p-value	U-stat	p-value	U-stat	p-value	U-stat	p-value	U-stat	p-value
Australia	**0.916**	0.077	**0.855**	0.099	**0.906**	0.072	**0.827**	0.057	**0.829**	0.003	**0.796**	0.005
Austria	**0.892**	0.039	**0.845**	0.091	0.939	0.170	0.940	0.391	**0.931**	0.032	0.949	0.348
Belgium	**0.881**	0.023	**0.833**	0.073	0.948	0.235	0.947	0.424	0.964	0.207	0.971	0.471
Canada	**0.926**	0.045	**0.881**	0.096	**0.873**	0.026	**0.812**	0.045	**0.810**	0.001	**0.745**	0.001
Denmark	**0.897**	0.053	0.864	0.150	0.935	0.143	0.923	0.327	0.963	0.180	1.025	0.785
Finland	**0.905**	0.048	**0.851**	0.078	**0.898**	0.035	**0.826**	0.054	**0.838**	0.001	**0.707**	0.000
France	**0.903**	0.053	0.876	0.159	0.931	0.127	0.895	0.205	**0.939**	0.054	0.909	0.137
Germany	**0.874**	0.018	**0.815**	0.053	0.937	0.171	0.928	0.340	**0.943**	0.068	0.917	0.165
Great Britain	**0.895**	0.030	0.855	0.111	**0.903**	0.061	0.855	0.114	**0.842**	0.001	**0.756**	0.002
Greece	1.055	0.819	1.074	0.842	1.072	0.878	1.048	0.803	1.096	0.943	1.066	0.863
Italy	**0.916**	0.089	0.927	0.335	0.941	0.186	0.936	0.379	**0.903**	0.017	**0.880**	0.080
Japan	0.919	0.108	0.882	0.176					**0.886**	0.010	**0.869**	0.070
Korea	0.925	0.119	0.884	0.159	**0.909**	0.085	0.884	0.179	**0.844**	0.005	**0.870**	0.059
Netherlands	**0.892**	0.035	**0.846**	0.095	0.949	0.229	0.951	0.445	0.980	0.361	1.060	0.888
Norway	**0.918**	0.083	0.884	0.183	0.925	0.107	0.881	0.170	**0.915**	0.036	**0.859**	0.063
Spain	0.924	0.108	0.906	0.271	0.969	0.345	0.940	0.384	0.945	0.121	0.918	0.188
Sweden	**0.854**	0.008	**0.805**	0.031	0.939	0.170	0.870	0.131	**0.907**	0.030	**0.860**	0.060
Switzerland	**0.879**	0.018	**0.821**	0.055	**0.922**	0.093	0.886	0.179				
United States					**0.867**	0.013	**0.794**	0.023	**0.796**	0.000	**0.738**	0.001
Average	**0.909**	0.003	**0.873**	0.033	**0.931**	0.040	**0.897**	0.097	**0.907**	0.000	**0.883**	0.004
Countries<1	17		17		17		17		17		15	

The p-value is the proportion of the bootstrap distribution to the left of the calculated U-statistic.
Bolded coefficients are statistically significant at a 10% level.

where the STEC model is significantly better at a 10% level for 17 (out of 18) countries when the U.S. is the numeraire, and for 14 countries when Japan or Switzerland are the numeraires. For one-quarter-ahead forecasts, the STEC model is significantly better than the driftless random walk (at a 10% level) for 12 exchange rates when the U.S. is the numeraire, and for 6 and 8 countries when Japan and Switzerland (respectively) are the numeraires.

As seen in Table 3, for long-run horizons (eight- and sixteen-quarters ahead) the U-statistics are on average lower than in the short run—implying more forecast accuracy of the STEC model over the random walk in the long run than in the short run. Moreover, the joint test reveals statistical significance at a 4% level for 5 of the 6 cases (the remaining case is statistically significant at a 10% level). For individual exchange rates, the results are very good for eight-quarter-ahead forecasts when the U.S. or Switzerland are the numeraires (the STEC model is significantly better at a 10% level for 14 and 13 exchange rates, respectively). For sixteen-quarter-ahead forecasts, the STEC model is significantly better (at a 10% level) for 9 exchange rates when the U.S. is the numeraire, and for 10 exchange rates when Switzerland is the numeraire. When Japan is the numeraire, the significance of the long-run results for individual exchange rates are not as good (the STEC model is significantly better only for 7 and 4 exchange rates for eight- and sixteen-quarter-ahead forecasts).

In comparison with the out-of-sample one- and sixteen-quarter-ahead results of Mark and Sul (2001), the STEC model has a better performance than their linear PPP estimation—in terms of U-statistics values and their significance—for one-quarter-ahead forecasts when the U.S. and Japan are the numeraires. The opposite is true when Switzerland is the numeraire. With respect to sixteen-quarter-ahead forecasts, the U-statistics of Mark and Sul are very low and highly significant when the U.S. and Switzerland are the numeraires. There is, however, an important caveat in Mark and Sul's results. As noted also by Engel, Mark, and West (2007), Mark and Sul base their out-of-sample results in a comparison against a random walk with drift, which turns out to perform worse than a driftless random walk. We look further into the comparison of the driftless random walk and the random walk with drift in the following section.

5 Out-of-Sample Predictability Robustness

The out-of-sample results of the previous section show that no matter the horizon or the numeraire, the STEC model is generally a better forecaster than the driftless random walk. In this

section we verify the robustness of these results in two dimensions. First, we follow Rogoff and Stavrakeva (2008) and analyze the statistical significance of our results over different forecast windows. Second, we compare the driftless random walk against the random walk with drift to verify that we are comparing our STEC model against the best of the naive specifications.

5.1 Predictability over Different Forecast Windows

In their evaluation of recent models of exchange rate forecasting, Rogoff and Stavrakeva (2008) propose a simple, yet strict, way to test "how consistently reliable" a forecast is. The test consists in looking at how the estimated model compares to the driftless random walk over different forecast windows. In our Tables 2 and 3, for example, we present out-of-sample results based on a forecast window given by 1985.Q1-2009.Q4. The Rogoff-Stavrakeva proposal consists in obtaining the information of Tables 2 and 3 for each of the remaining possible forecast windows: from 1985.Q2-2009.Q4 to 2009.Q3-2009.Q4.[11]

5.1.1 Joint Test

We begin by looking into the joint test based on the average U-statistic. For the initial forecast window, we obtained in the tables above that the STEC model was on average (significantly) better than the driftless random walk for each of the three numeraires over every horizon. In Figure 1 we present graphs for the bootstrapped p-values of the joint test for the different forecast windows (for each numeraire over the different forecast horizons). The x-axis represents the first quarter of the forecast window (so that the first plotted point for each line corresponds to a p-value—for the average U-statistic—in Table 2 or 3).

As we can observe in Figure 1a, when the U.S. is the numeraire, the p-values for the average U-statistic over the different forecast horizons are below 0.1 (with a couple of exceptions where the p-values are just above the 0.1 line) for the forecast windows with starting quarter between 1985 and 2003. From around 2003, the STEC model is not a reliable forecaster for four- and eight-quarter-ahead horizons. The same happens with sixteen-quarter-ahead forecasts starting in 2005. Nevertheless, the one-quarter-ahead forecast is robust to the different forecast windows up to the last quarter of 2008. At that time, it jumps above the 0.1 line and remains there for the last quarters. The fact that the reliability of the one-quarter-ahead forecast breaks down
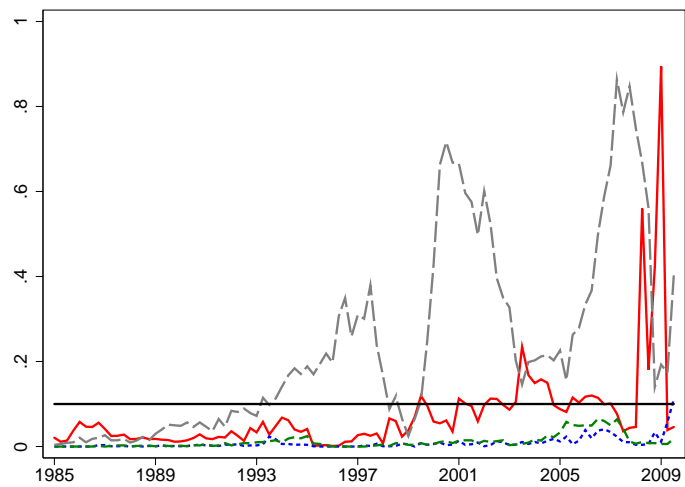
---

[11]In the worst case scenario, the estimated model could be significantly better than the random walk in the initial window, but not significantly better for the rest of the forecast windows.
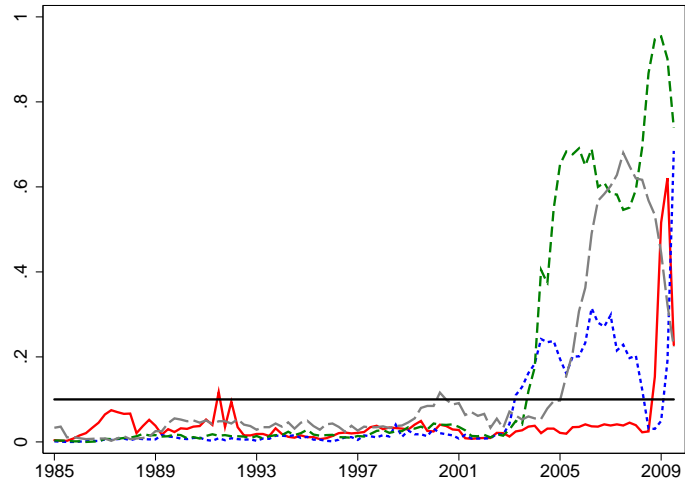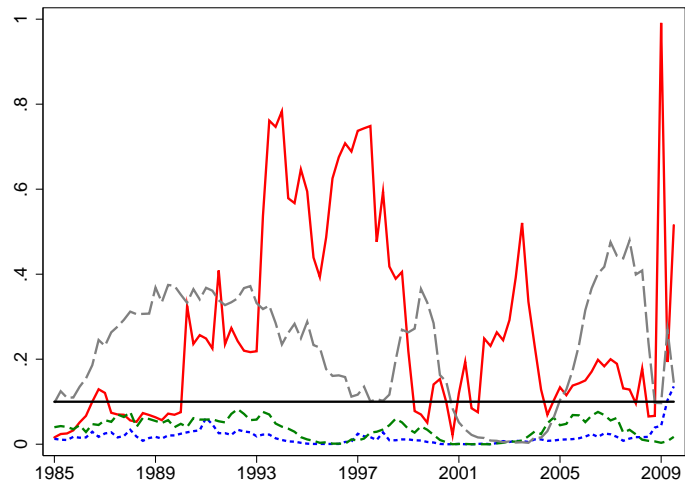
Figure 1: Joint test *p*-values: one quarter ahead *(solid)*, four quarters ahead *(dot)*, eight quarters ahead *(dash)*, sixteen quarters ahead *(long dash)*

by the end of 2008 and during 2009 should not be surprising, given the unprecedented world economic turmoil unleashed by the events in the U.S. financial system during September 2008.

In Figure 1b we see the evolution of the p-values when Japan is the numeraire. We obtain contrasting results. On the one hand, one- and sixteen-quarter-ahead forecasts perform poorly over most of the forecast windows. On the other hand, the four- and eight-quarter-ahead forecasts are remarkably reliable: the p-value is always below 0.1 for the eight-quarter-ahead forecast and only one time just above 0.1 (for the last forecast window) for the four-quarter-ahead forecast. In Figure 1c we observe that the same is true for four- and eight-quarter-ahead forecasts when Switzerland is the numeraire. Also, for Switzerland, the one-quarter-ahead forecast is mostly reliable (with the exception of the forecast windows between 2003 and 2005, and between 2008 and 2009), while the sixteen-quarter-ahead forecast performs poorly for forecast windows starting in 1994.

To sum up, on average the STEC model is consistently and significantly better than the driftless random walk for one-quarter-ahead horizons when the U.S. is the numeraire, and for four- and eight-quarter-ahead horizons when either Japan or Switzerland is the numeraire.

5.1.2 Individual Exchange Rates

Let us now look into the bootstrapped p-values for individual exchange rates. To preserve space, we focus on the short-run cases that are robust over different forecast windows for the joint test: one-quarter-ahead forecasts when the U.S. is the numeraire, and four-quarter-ahead forecasts when Japan and Switzerland are the numeraires.

Figure 2 shows the one-quarter-ahead U.S. case. In Figure 2a we show those countries that are robust throughout most of the forecast windows, while Figure 2b shows some other countries that perform well for some forecast windows but do poorly for some others.[12]

The one-quarter-ahead predictability for the exchange rates (against the U.S. dollar) of Australia, Canada and Sweden is quite robust (and for Finland to a lesser degree). As in the joint test, the predictability breaks down by the end of 2008 and during 2009, but again, it is not surprising that the exchange did not follow PPP fundamentals during the most unstable economic period in the U.S. since the Great Depression. It is interesting to note that Australia, Canada, and Sweden are the same countries for which Rogoff and Stavrakeva (2008) find robust

---

[12]The forecastability results for the countries that are not in Figure 2 (Greece, Japan, Korea, Spain, and Switzerland) are in general poor.
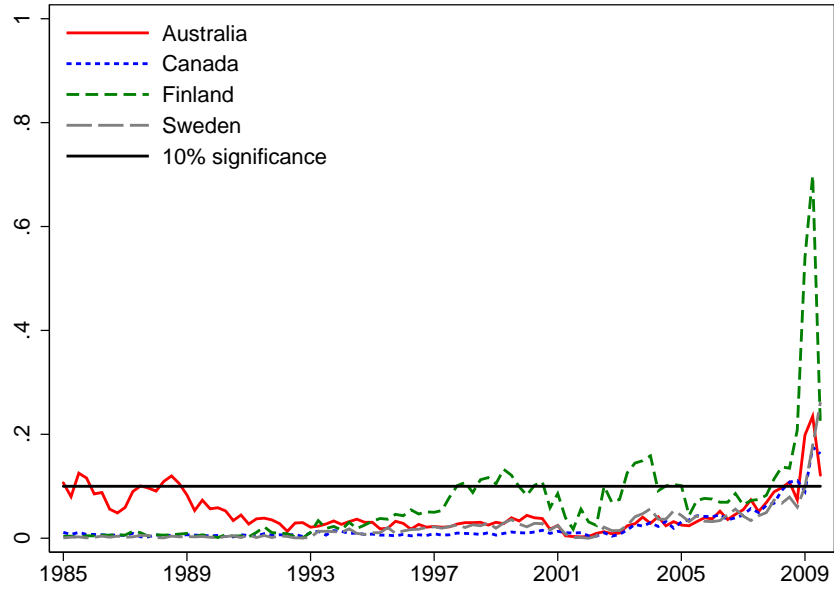
(a) Robust group
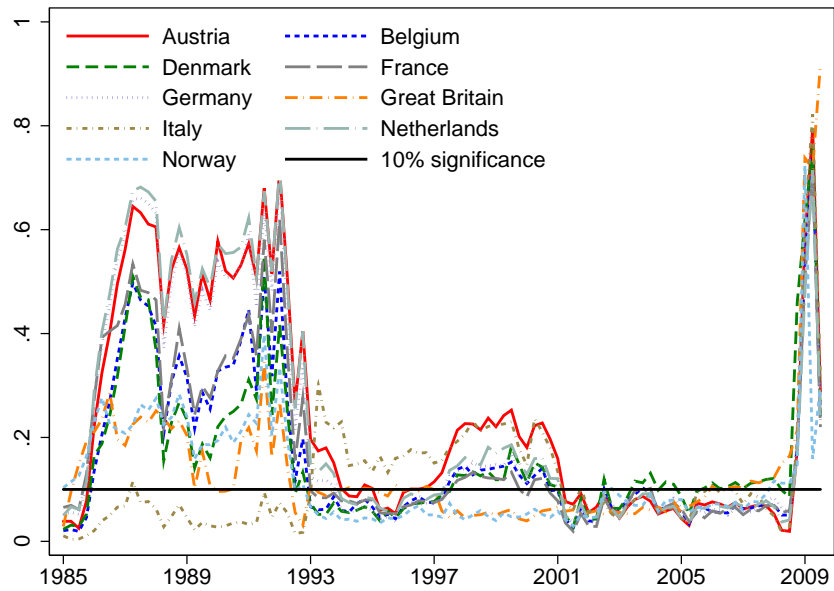
(b) Other countries

Figure 2: Individual *p*-values with the U.S. as numeraire (one quarter ahead)

predictability with their method of cross-country shocks with forecast "pooling." The difference is that Rogoff and Stavrakeva's method relies on *ad hoc* weighting procedures, while the STEC model is based exclusively on fundamentals under a nonlinear structural model.

Figure 2b includes Euro area countries and Denmark, Great Britain, and Norway. The figure shows a typical example on why the Rogoff and Stavrakeva proposal of checking forecastability over different windows is highly relevant. Note that for our original forecast window (the first point in every line), all countries (but Norway) are below the 10% significance level line. However, for windows starting half a year later and up to 1993, only Italy's forecastability remains robust. From 1993, the forecastability for Great Britain and Norway becomes significant (up to the 2008 crisis). Note also that from 2001 and up to 2008, the euro forecastability is statistically significant for all the Euro area countries in the Figure, suggesting (as we should expect) a close relationship between these countries' price levels.

Figure 3 shows the countries with better four-quarter-ahead forecastability when Japan and Switzerland are the numeraires. As we can see in Figure 3a, when Japan is the numeraire, the STEC model robustly outperforms the driftless random walk for half the countries in our sample (again, the predictability breaks down towards the end of our sample period). These are all European countries, seven from the Euro area plus Denmark and Great Britain. In Figure 3b we observe that the results are not as good when Switzerland is the numeraire, with only four countries with significant forecastability throughout most of the sample period.

Hence, although our parsimonious model based in economic fundamentals is on average better than the random rank even in short-run horizons, the number of individual exchange rates for which the significance of the predictability is robust (to different forecast windows) remains low.

5.2 Driftless Random Walk versus Random Walk with Drift

There has been a debate on whether or not we should include a drift in the random walk specification for out-of-sample forecast accuracy comparisons. As first shown by Meese and Rogoff (1983), the most important argument in favor of the driftless random walk specification is that it is a far more accurate out-of-sample predictor of the nominal exchange rate—in U.S. dollars—than the random walk with drift. On the other hand, Engel and Hamilton (1990) find that conventional in-sample statistics of the price of the dollar against some major currencies
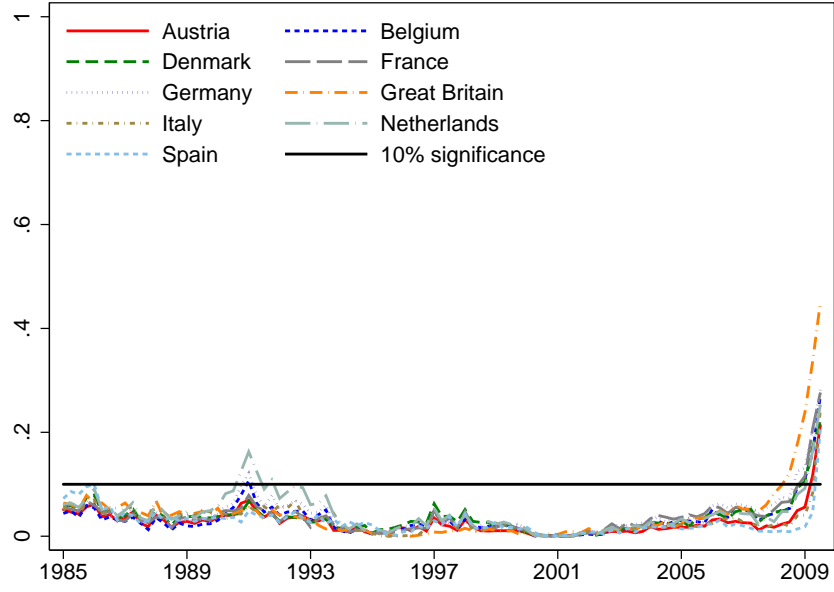
(a) Numeraire: Japan
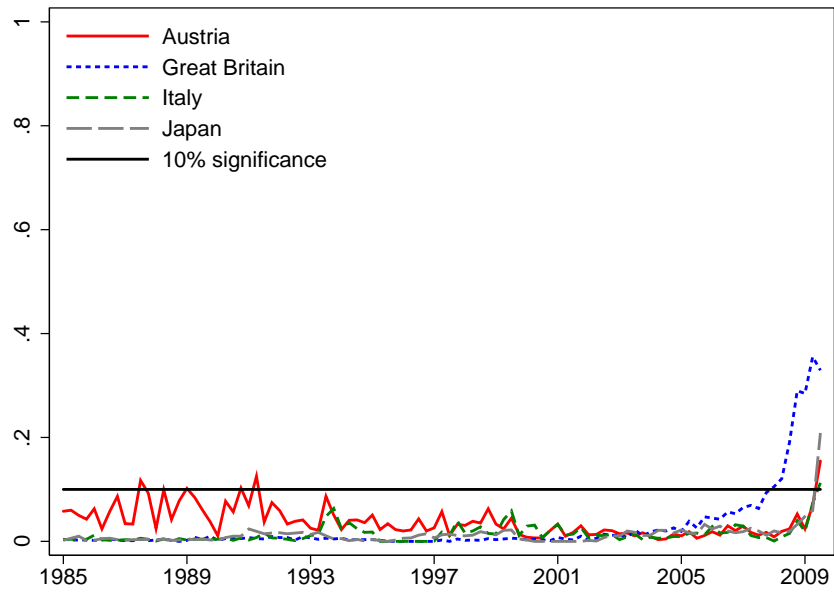
(b) Numeraire: Switzerland

Figure 3: Individual p-values with Japan and Switzerland as numeraires (four quarters ahead)

reject the hypothesis that the drift has been constant across pre-selected periods of time, so that it is arbitrary to set the drift equal to zero. In this paper, we focus exclusively on out-of-sample forecast accuracy and therefore we worry only about comparing our STEC model against the best of the random walk specifications. Given that in the previous sections we have used the driftless random walk as our competing model, we devote this section to verify if indeed the driftless random walk is better than the random walk with drift.

As mentioned above, Engel, Mark, and West (2007) revise the results of Mark and Sul (2001) for an updated sample (up to 2005.Q4) and show—with the U.S. as numeraire—that the linear PPP model outperforms the random walk with drift for one-quarter-ahead forecasts for 15 out of 18 countries, but only for 7 out of 18 when the driftless random walk is used. Results are better for sixteen-quarter-ahead forecasts, with the PPP model outperforming the random walk with drift in 17 out of 18 countries, and for 14 out of 18 when compared against the driftless random walk. Instead of comparing the STEC model against both random walk specifications, we provide an extensive out-of-sample forecast accuracy comparison of the random walk without drift versus the random walk with drift. We do this for our updated version of Mark and Sul's data (up to 2009.Q4) and for the three numeraires of their original work.

Based on our initial forecast window (1985.Q1-2009.Q4), Table 4 presents out-of-sample U-statistics across different horizons for our three numeraires. A U-statistic less than one implies a higher forecast precision of the driftless random walk. As we can observe, the driftless random walk is generally a better predictor for each of the three numeraires: it is better than the random walk with drift for 15 or more countries (out of 18) across different horizons when the U.S. and Switzerland are the numeraires, and for 14 or more countries when Japan is the numeraire. The average U-statistics show that the relative accuracy of the driftless random walk improves the longer the forecast horizon. Hence, if we were to compare the STEC model with the random walk with drift, we would find lower U-statistics for the vast majority of countries. This would appear to generate large predictability gains coming from economic fundamentals, when it is actually a lower forecasting performance of the benchmark model.

The better forecast accuracy of the driftless random walk remains over most of the forecast windows. Figure 4 shows the number of countries for which the driftless random walk is better— with a U-statistic less than one—than the random walk with drift over the different windows. As before, the x-axis represents the first quarter of the forecast window (so that the first point

Table 4: Out-of-Sample U-Statistics for Comparison of Random Walks: No–Drift versus Drift

Numeraire	United States				Japan				Switzerland			
Quarters ahead	1	4	8	16	1	4	8	16	1	4	8	16
Australia	0.986	0.955	0.906	0.858	1.000	1.001	0.986	0.978	0.997	0.985	0.948	0.898
Austria	1.000	0.994	0.988	0.978	0.994	0.976	0.956	0.907	0.978	0.901	0.773	0.497
Belgium	0.990	0.964	0.936	0.877	0.989	0.961	0.925	0.879	0.941	0.744	0.551	0.425
Canada	0.988	0.956	0.916	0.855	1.000	0.997	0.985	0.956	0.998	0.984	0.957	0.838
Denmark	0.982	0.941	0.898	0.823	0.986	0.955	0.917	0.865	0.921	0.681	0.504	0.336
Finland	0.983	0.949	0.910	0.833	0.993	0.976	0.950	0.928	0.979	0.910	0.812	0.687
France	0.975	0.923	0.864	0.779	0.981	0.938	0.883	0.836	0.901	0.671	0.495	0.390
Germany	0.999	0.991	0.982	0.969	0.993	0.975	0.954	0.908	0.973	0.878	0.734	0.455
Great Britain	0.977	0.926	0.857	0.763	0.995	0.987	0.974	0.951	0.988	0.956	0.901	0.776
Greece	0.964	0.906	0.898	0.967	1.013	1.030	1.056	1.103	1.019	1.019	1.035	1.102
Italy	0.952	0.864	0.777	0.653	0.977	0.933	0.884	0.829	0.933	0.790	0.660	0.545
Japan	1.004	1.008	1.003	0.995					0.994	0.974	0.937	0.820
Korea	0.991	0.969	0.929	0.829	1.006	1.027	1.034	1.048	1.001	1.003	0.988	0.865
Netherlands	0.997	0.984	0.971	0.949	0.992	0.973	0.949	0.903	0.965	0.839	0.673	0.404
Norway	0.987	0.955	0.914	0.855	0.996	0.990	0.980	0.970	0.990	0.963	0.901	0.810
Spain	0.952	0.864	0.771	0.696	0.976	0.928	0.868	0.839	0.923	0.750	0.628	0.562
Sweden	0.982	0.946	0.890	0.816	0.998	0.991	0.975	0.981	0.992	0.964	0.883	0.814
Switzerland	1.003	1.002	0.994	0.950	0.994	0.974	0.937	0.820				
United States					1.004	1.008	1.003	0.995	1.003	1.002	0.994	0.950
Average	0.984	0.950	0.911	0.858	0.994	0.979	0.956	0.927	0.972	0.890	0.799	0.676
Countries<1	15	16	17	18	14	14	15	16	15	15	17	17

The U-statistics are computed as the ratio of the RMSE of the driftless random walk to the RMSE of the random walk with drift. A value less than one implies that the driftless random walk performs better.

in each line is a number in the last row of Table 4). Figure 4a shows that when the U.S. is the numeraire, the driftless random walk outperforms the random walk with drift for 10 or more countries (at any forecast horizon) for forecast windows starting before 2007. After that year, it only goes below 9 for eight-quarter-ahead forecasts and it reaches 9 for one-quarter-ahead forecasts only in the last forecast window. When Japan is the numeraire, the driftless random walk is better over forecast windows with an initial quarter between 1985 and 2006. After that, the random walk with drift becomes better. Something similar happens when Switzerland is the numeraire, with the random walk with drift becoming better only after 2007. Thus, the driftless random walk dominates the random walk with drift throughout the vast majority of our windows.

With respect to the individual exchange rates for which the STEC model consistently and significantly outperforms the driftless random walk (in Section 5.1.2), the driftless random walk is better than the random walk with drift in almost all forecast windows for almost all cases.[13] For the one-quarter-ahead U.S. case, the driftless random walk is better than the random walk with drift in every forecast window for Australia, Canada, and Finland. For Sweden, the random walk with drift becomes slightly better in forecast windows starting between 2007.Q4 and 2008.Q3. For the robust countries when Japan is the numeraire (in Figure 3a), with the exception of Great Britain, the random walk with drift only becomes better than the driftless random walk in forecast windows starting after 2006. Finally, for the countries in Figure 3b with Switzerland as numeraire, the driftless random walk always dominates for Japan and it dominates up to windows starting in 2007 for Austria and Italy.[14]

6 Conclusions

We present a Smooth Transition Error-Correction Model—in the spirit of the generalized cointegrated system of Granger and Swanson (1996)—for the relationship between PPP fundamentals and the nominal exchange rate. Using a panel dataset of 19 countries and estimating the model for three numeraires, we find strong evidence of nonlinear mean reversion of the real exchange rate and of nonlinear predictability of the nominal exchange rate.

Out-of-sample statistics show a better forecast accuracy of the STEC model than the driftless

---

[13]Individual U-statistics for the comparison of the two random walk specifications are available upon request.
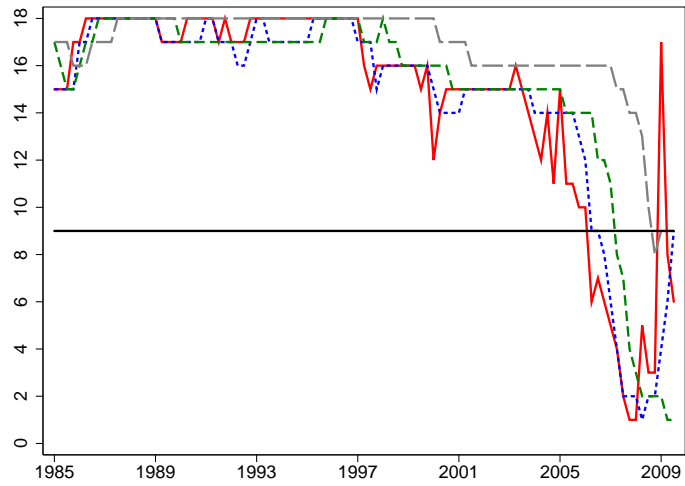
[14]For Great Britain, the random walk with drift dominates for forecast windows starting around 2003 when Japan is the numeraire, and for forecast windows starting around 1998 when Switzerland is the numeraire.
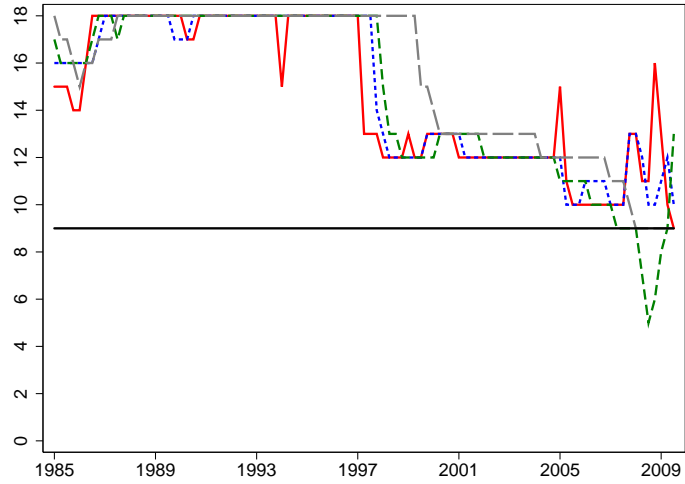
(a) Numeraire: United States
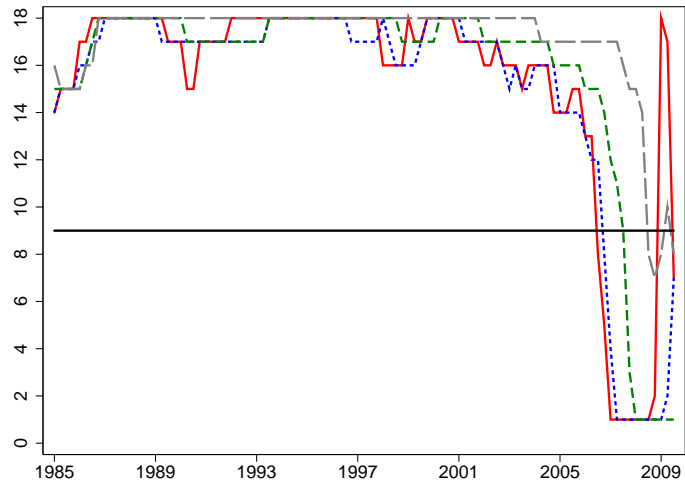
(b) Numeraire: Japan

(c) Numeraire: Switzerland

Figure 4: Number of countries for which the driftless random walk is better than the random walk with drift: one quarter ahead *(solid)*, four quarters ahead *(dot)*, eight quarters ahead *(dash)*, sixteen quarters ahead *(long dash)*

random walk specification across numeraires and horizons. We follow Rogoff and Stavrakeva (2008) and verify the robustness of the out-of-sample results over different forecast windows. We find that on average, the STEC model significantly dominates the driftless random walk over different forecast windows for one-quarter-ahead horizons when the U.S. is the numeraire and for four- and eight-quarter-ahead horizons when Japan and Switzerland are the numeraires.

These results differ with recent contributions from Engel, Mark, and West (2007) and others, who find weak (nominal exchange rate) predictability results for short-run horizons—U-statistics very close to one (about 0.99 or more) and not significant. A possible explanation of the difference in results might be that nonlinear models with panel data perform better than traditional linear models in the short run. Moreover, as Rogoff and Stavrakeva (2008), we find robust one-quarter-ahead predictability when the U.S. is the numeraire for Australia, Canada, and Sweden. The difference is that in contrast to the forecast pooling method with cross-country shocks of Rogoff and Stavrakeva, which requires specific weighting procedures, our results come from the direct estimation of an structural nonlinear model.

References

BENNINGA, S., AND A. PROTOPAPADAKIS (1988): "The Equilibrium Pricing of Exchange Rates and Assets when Trade Takes Time," *Journal of International Money and Finance*, 7(2), 129–149.

BERBEN, R., AND D. VAN DIJK (1998): "Does the Absence of Cointegration Explain the Typical Findings in Long Horizon Regression?," Working Paper 9814, Erasmus University of Rotterdam - Econometric Institute.

BERKOWITZ, J., AND L. GIORGIANNI (2001): "Long-Horizon Exchange Rate Predictability?," *Review of Economics and Statistics*, 83(1), 81–91.

CHEUNG, Y.-W., M. D. CHINN, AND A. GARCIA PASCUAL (2005): "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?," *Journal of International Money and Finance*, 24(7), 1150–1175.

DIEBOLD, F. X., AND J. A. NASON (1990): "Nonparametric Exchange Rate Prediction?," *Journal of International Economics*, 28(3-4), 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.

——— (1994): "Partial Equilibrium versus General Equilibrium Models of International Capital Markets," in *Handbook of International Macroeconomics*, ed. by F. Van Der Ploeg, chap. 10. Blackwell, Oxford.

ENGEL, C., AND J. D. HAMILTON (1990): "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?," *American Economic Review*, 80(4), 689–713.

ENGEL, C., N. C. MARK, AND K. D. WEST (2007): "Exchange Rate Models Are Not as Bad as You Think," *NBER Macroeconomics Annual*, 22, 381–441.

ENGEL, C., AND K. D. WEST (2005): "Exchange Rates and Fundamentals," *Journal of Political Economy*, 113(3), 485–517.

FAUST, J., J. H. ROGERS, AND J. H. WRIGHT (2003): "Exchange Rate Forecasting: the Errors We've Really Made," *Journal of International Economics*, 60(1), 35–59.

GRANGER, C. W. J., AND N. SWANSON (1996): "Future Developments in the Study of Cointegrated Variables," *Oxford Bulletin of Economics and Statistics*, 58(3), 537–53.

GRANGER, C. W. J., AND T. TERÄSVIRTA (1993): *Modelling Nonlinear Economic Relationships*. Oxford University Press, Oxford.

IMBS, J., H. MUMTAZ, M. O. RAVN, AND H. REY (2003): "Nonlinearities and Real Exchange Rate Dynamics," *Journal of the European Economic Association*, 1(2), 639–649.

KILIAN, L. (1999): "Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?," *Journal of Applied Econometrics*, 14(5), 491–510.

KILIAN, L., AND M. P. TAYLOR (2003): "Why is it so Difficult to Beat the Random Walk Forecast of Exchange Rates?," *Journal of International Economics*, 60(1), 85–107.

MARK, N. C. (1995): "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85(1), 201–218.

MARK, N. C., AND D. SUL (2001): "Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Amall Post-Bretton Woods Panel," *Journal of International Economics*, 53(1), 29–52.

MEESE, R. A., AND K. ROGOFF (1983): "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics*, 14(1), 3–24.

MEESE, R. A., AND A. K. ROSE (1991): "An Empirical Assessment of Non-linearities in Models of Exchange Rate Determination," *Review of Economic Studies*, 58(3), 603–619.

OBSTFELD, M., AND A. M. TAYLOR (1997): "Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited," *Journal of the Japanese and International Economies*, 11(4), 441–479.

O'CONNELL, P. G. J., AND S.-J. WEI (2002): "The Bigger They Are, the Harder They Fall: Retail Price Differences Across U.S. Cities," *Journal of International Economics*, 56(1), 21–53.

OHANIAN, L. E., AND A. C. STOCKMAN (1997): "Arbitrage Costs and Exchange Rates," University of Rochester, unpublished.

ROGOFF, K. (1996): "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 34(2), 647–668.

ROGOFF, K. S., AND V. STAVRAKEVA (2008): "The Continuing Puzzle of Short Horizon Exchange Rate Forecasting," NBER Working Papers 14071, National Bureau of Economic Research, Inc.

SERCU, P., R. UPPAL, AND C. VAN HULLE (1995): "The Exchange Rate in the Presence of Transaction Costs: Implications for Tests of Purchasing Power Parity," *Journal of Finance*, 50(4), 1309–1319.

STOCK, J. H., AND M. W. WATSON (1988): "Variable Trends in Economic Time Series," *Journal of Economic Perspectives*, 2(3), 147–174.

TAYLOR, M. P., AND D. A. PEEL (2000): "Nonlinear Adjustment, Long-run Equilibrium and Exchange Rate Fundamentals," *Journal of International Money and Finance*, 19(1), 33–53.

TAYLOR, M. P., D. A. PEEL, AND L. SARNO (2001): "Nonlinear Mean-Reversion in Real Exchange Rates: Toward a Solution to the Purchasing Power Parity Puzzles," *International Economic Review*, 42(4), 1015–42.

TERÄSVIRTA, T. (1994): "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89(425), 208–218.