The dynamics of real exchange rates – A reconsideration*

Hendrik Kaufmann, Florian Heinen, Philipp Sibbertsen[1]

Institute of Statistics, Faculty of Economics and Management,
Leibniz University Hannover, D-30167 Hannover, Germany

Abstract

While it is widely agreed that Purchasing Power Parity (PPP) holds as a long-run concept the specific dynamic driving the process is largely build upon a priori economic belief rather than a thorough statistical specification. The two prevailing time series models, i.e. the exponential smooth transition autoregressive (ESTAR) model and the Markov switching autoregressive (MSAR) model, are both able to support the PPP as a long-run concept. However, the dynamic behavior of real exchange rates implied by these two models is very different and leads to different economic interpretations.
In this paper we approach this problem by offering a bootstrap based testing procedure to discriminate between these two rival models. We further study the small sample performance of the test.
In an application we analyze several major real exchange rates to shed light on the question which model describes these processes best. This allows us to draw conclusions about the driving forces of real exchange rates.

JEL-Numbers: C12, C15, C22, C52, F31

Keywords: Nonlinearities · Markov switching · Smooth transition · Specification testing · Real exchange rates

---

*The authors are grateful to the participants of the 19th Symposium of the Society for Nonlinear Dynamics and Econometrics in Washington, D.C., the ESEM 2011 in Oslo, the NBER-NSF Time Series Conference in Michigan, the Statistische Woche in Nuremberg, the German Econometric Study Group, the Workshop on "Nonlinear and persistent Time Series", the editor and three anonymous referees for inspiring discussion and helpful comments. Financial support by the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

[1] Corresponding author:
Phone: +49-511-762-3783
Fax: +49-511-762-3923
E-Mail: sibbertsen@statistik.uni-hannover.de

1 Introduction

An ongoing debate about the behavior of real exchange rates suggests that Purchasing Power Parity (PPP) holds as a long-run concept (see e.g. Edison and Klovland, 1987, MacDonald, 1998 or Taylor et al., 2001). Econometrically speaking PPP states that real exchange rates fluctuate finitely around an equilibrium, i.e. a time stable mean, and are thus weakly stationary. However, stationarity of real exchange rates does not say anything about the detailed dynamics driving them. Modeling real exchange rates by linear stationary models does not lead to convincing results as standard unit root tests can not reject the null of a random walk in the linear framework and thus can not confirm PPP (see e.g. Adler and Lehmann, 1983, Meese and Rogoff, 1983, Meese and Rogoff, 1988 or Caporale et al., 2003).

Dumas (1992) shows in a theoretical model with two countries and proportional transaction costs that the real exchange rate behaves nonlinear and mean reverting. In this framework, the probability of moving away from the equilibrium is higher than moving towards it. As a consequence the exchange rate spends the most time away from PPP. However, the adjustment speed increases the further the series is away from parity. This can explain the long swings in exchange rates (see Engel and Hamilton, 1990). Therefore, nonlinear models came into the focus of economists. Empirical evidence that a nonlinear adjustment mechanism could solve the PPP puzzle is recently provided by Lo (2008) and Norman (2010).

The two prevailing approaches to model the dynamic of nonlinear adjustment towards PPP are exponential smooth transition autoregressive (ESTAR) models (see e.g. Michael et al., 1997, Taylor et al., 2001 and Kilian and Taylor, 2003) and Markov switching autoregressive (MSAR) models (see e.g. Bergman and Hansson, 2005 and Kanas, 2006).[1] Both approaches imply under certain conditions that PPP may hold as a long-run concept. However, the nonlinear dynamics driving the respective processes are very distinct. The dynamic of an ESTAR process is driven by lagged values of the endogenous, and therefore observable, variable while an unobservable Markov process governs the dynamic in the MSAR case. This implies different economic intuitions, arbitrage and market imperfections on the one hand and structural instability on the other.

The ESTAR approach is based on the idea that international trade only starts if the price differences between countries exceed a certain level which is determined by the costs of trading such as transportation costs, taxes and many others. As long as the price differences are smaller than this level no trading takes place and therefore real

[1] An exception is Lahtinen (2006) who uses a symmetric, second order logistic STAR model. However, second order logistic STAR models are designed to closely resemble ESTAR models and the estimation results imply a dynamic very similar to a switching regression.

exchange rates fluctuate freely and behave like a random walk. As soon as the price differences exceed this level international trading starts and real exchange rates are pulled back to a long-run equilibrium leading to an economic view of real exchange rates which is expressed in the ESTAR model by an inner unit root regime and stabilized by a stationary outer regime. This arbitrage-based adjustment models market behavior without exogenous shocks as the adjustment is driven by an observable endogenous variable. The approach has been applied to real exchange rates by Taylor et al. (2001), Kapetanios et al. (2003) and Rapach and Wohar (2006) amongst others.

In contrast, the adjustment in the MSAR model is driven by an unobservable Markov chain. Therefore, the nonlinear dynamic of the process is exogenous which emphasizes a Peso process or infrequent regime or policy rule changes as a driving force of real exchange rates. Compared to the symmetric ESTAR adjustment based on arbitrage opportunities these effects might cause an asymmetric behavior around the equilibrium. van Norden (1996) shows that MSAR models are able to capture such occasional, sudden and large exchange rate changes. Regarding the MSAR approach there are two possibilities of modeling this economic behavior. One possibility is to implement Markov switching within the autoregressive parameter and thus having a model with two autoregressive regimes, one of them being explosive and the other stabilizing. This model was applied among others by Kanas (2006). A second option is to implement Markov switching in the mean rather than in the autoregressive parameter. This approach was motivated by Engel and Hamilton (1990) and Hamilton (1993) who argue that switches in the mean are more appropriate for modeling the dynamics of financial data. Bergman and Hansson (2005) apply this approach to real exchange rates.

The application of either approach has so far been motivated by an a priori economic belief about the behavior of real exchange rates rather than formal statistical specification. However, arbitrage as well as structural instability are probably present in each real exchange rate. Thus, we do not see the ESTAR and the MSAR approach as mutually exclusive. It is possible and likely that both described effects influence the behavior of real exchange rates. For example Sarno and Valente (2006) estimated an MS-VECM model with an ESTAR function. However, this is not the scope of this paper. The main question here is if one of these influences has only a minor impact so that the explanatory power of the model is sufficient without including additional dynamics. In fact there is no guarantee that one of the applied models is correct. Therefore, our proposed test is not a model selection procedure, but a specification test saying which model approximates the data better and therefore gives an idea which of the proposed dynamics is the dominating one for the considered real exchange rates.

This statistical test procedure is based on a parametric bootstrap version of the Cox test to discriminate between ESTAR and MSAR models. We study the finite sample behavior of the proposed test and analyze several major real exchange rates to shed some light on the question which model and by that which effect influences the dynamic behavior of a particular real exchange rate most.

The rest of the paper is organized as follows: In section 2 we describe the two competing models for real exchange rates. In section 3 we discuss the testing procedure with the corresponding bootstrap in order to distinguish the competing models and investigate the finite sample properties of this test. Section 4 applies the test procedure to empirical data before section 5 concludes.

2 Two Competing models for real exchange rates

The general ESTAR model is given by two autoregressive regimes connected by a smooth exponential transition function $\mathcal{G}(y_{t-d};\gamma,c) : \mathbb{R} \to [0,1]$. This function governs the transition between the two extreme regimes in a smooth way. Alternatively, an ESTAR model can also be interpreted as a continuum of regimes which is passed through by the process.

In general, a univariate ESTAR(p) model with $p \geq 1$ and $d \leq p$ is given by

$$
\begin{aligned}
[y_t - c] &= \left[\sum_{i=1}^{p} \psi_i [y_{t-i} - c]\right] \times [1 - \mathcal{G}(y_{t-d};\gamma,c)] + \left[\sum_{i=1}^{p} \theta_i [y_{t-i} - c]\right] \times \mathcal{G}(y_{t-d};\gamma,c) + \varepsilon_t \quad (1)\\
&= \sum_{i=1}^{p} \psi_i [y_{t-i} - c] + \left[\sum_{i=1}^{p} \phi_i [y_{t-i} - c]\right] \times \mathcal{G}(y_{t-d};\gamma,c) + \varepsilon_t, \quad t \geq 1, \quad (2)
\end{aligned}
$$

with $\varepsilon_t \overset{iid}{\sim} (0,\sigma^2)$.

The exponential transition function $\mathcal{G}(y_{t-d};\gamma,c)$ is given by

$$\mathcal{G}(y_{t-d};\gamma,c) = 1 - \exp\left\{-\gamma (y_{t-d} - c)^2\right\}; \quad \gamma > 0 \,. \quad (3)$$

This transition function provides a symmetric adjustment towards the equilibrium c. Surveys of the broad field of nonlinear time series models in general and STAR models in particular are given by van Dijk et al. (2002) and Teräsvirta (1994).

The most frequently used special case of the general ESTAR model in (2) is the ESTAR(1) model

$$[y_t - c] = \psi [y_{t-1} - c] + \phi [y_{t-1} - c] \left\{1 - \exp\left(-\gamma (y_{t-1} - c)^2\right)\right\} + \varepsilon_t \,. \quad (4)$$

To model real exchange rate behavior, Taylor et al. (2001) and Rapach and Wohar (2006) impose an inner unit root regime, $\psi = 1$. This regime is corrected back by a white noise process for the outer regime, $\phi = -1$, to ensure global stationarity. In this case the process behaves like a random walk if $y_{t-1} = c$. When the deviation $|y_{t-1} - c|$ from the long-run equilibrium c grows, the process is corrected back by the more and more influential stationary outer regime. The parameter γ governs the speed of the mean reversion. In the unrestricted case global stationarity is given as long as $|\psi + \phi| < 1$. Estimation of these models either by nonlinear least squares or maximum likelihood techniques is treated by Klimko and Nelson (1978) and Tjøstheim (1986) respectively.

For the Markov switching framework we use the framework based on Lindgren (1978), and first applied to exchange rates by Engel and Hamilton (1990) and Engel (1994):

$$y_t = \mu_{s_t} + \phi_{1,s_t} y_{t-1} + \ldots + \phi_{p,s_t} y_{t-p} + \varepsilon_t \,. \tag{5}$$

The values of the autoregressive parameters $\phi_{1,s_t}, \ldots, \phi_{p,s_t}$ and the mean μ_{s_t} and thus the regime switching is governed by an unobservable first order Markov chain

$$\mathcal{P}(s_t = j | s_{t-1} = i, s_{t-2} = k, \ldots, y_{t-1}, y_{t-2}, \ldots) = \mathcal{P}(s_t = j | s_{t-1} = i) = p_{ij} \,.$$

The transition probabilities p_{ij} are interior points in the open unit interval to ensure an ergodic Markov chain and $\mu_1 \le \mu_2 \le \cdots \le \mu_k$ for a clear identification of the k regimes. Extensions of this basic framework are possible, see e.g. Hamilton and Raj (2002) and the papers cited therein.

Although it is possible to use such a general model, the models usually found in applied work are more restrictive and have only a few parameters that are regime-dependent. In real exchange rate work the most frequently used model is a first order autoregression with a Markov switching mean and/or autoregressive parameter (see e.g. Engel and Hamilton, 1990, Bergman and Hansson, 2005 or Kanas, 2006). The switching mean model is empirically justified by Hegwood and Papell (1998) and Montañés (1997) who find reversion to an equilibrium which is subject to structural breaks.

In our analysis we use the same two-state model as Bergman and Hansson (2005),

$$y_t = \mu_{s_t} + \phi y_{t-1} + \varepsilon_t \,, \tag{6}$$

where $\varepsilon_t \overset{iid}{\sim} (0, \sigma^2)$ as before. The model in (6) is locally stationary as long as the usual condition $|\phi| < 1$ holds. Global stationarity is given when the following conditions hold (see Francq and Zakoïan, 2001):

$$(p_{11} + p_{22})\phi^2 + (1 - p_{11} - p_{22})\phi^4 < 1 \tag{7}$$

$$(p_{11} + p_{22})\phi^2 < 2 \,. \tag{8}$$

Maximum likelihood estimation by direct, numerical maximization of the likelihood function of such models is treated in detail by Hamilton (1989).

It should be mentioned that the ESTAR and MSAR model are partially non-nested in the sense of Pesaran (1987) as both models nest linear AR-models and thus also the random walk as a special case. As soon as some kind of nonlinear dynamic comes into play, the models are non-nested.

3 Testing non-nested hypotheses

Testing non-nested hypotheses was pioneered by Cox (1961, 1962). He shows that the usual unadjusted likelihood-ratio statistic does not converge to zero if the two models are non-nested. He solves this problem by generalizing the principle and introducing a centered likelihood ratio approach widely known as Cox test. Alternatively Gourieroux et al. (1983) and Mizon and Richard (1986) modify the encompassing principle to allow for non-nested hypothesis testing. This Wald- or Score test based approach has good statistical properties but needs the existence of a Hessian under the null and alternative. As the transition function of the Markov-Switching model is not differentiable the encompassing approach cannot be applied in our context. For this reason we apply a bootstrap version of the Cox test. For a detailed overview about non-nested hypothesis testing see Pesaran and Weeks (2001) or Gourieroux and Monfort (1994).

For the Cox test let the two models for the conditional density be denoted by

$$H_f \quad : \quad \mathcal{F}_\theta = \{f(y_t|\mathfrak{F}_{t-1};\theta),\ \theta \in \Theta \subseteq \mathbb{R}^r\},\ t = 1,\ldots,T \tag{9}$$

$$H_g \quad : \quad \mathcal{F}_\lambda = \{g(y_t|\mathfrak{F}_{t-1};\lambda),\ \lambda \in \Lambda \subseteq \mathbb{R}^q\}. \tag{10}$$

The process is given by y_t and $\mathfrak{F}_{t-1}$ is the sigma algebra generated by $(y_{t-1}, y_{t-2}, \ldots)$. For the sake of brevity we will suppress the dependence on y_t and $\mathfrak{F}_{t-1}$ and simply write $f(\theta)$ and $g(\lambda)$ whenever possible. The likelihood-ratio statistic of Cox (1961, 1962) is given by

$$T_f(\hat{\theta},\hat{\lambda}) = \left[\log f(\hat{\theta}) - \log g(\hat{\lambda})\right] - E_f\left[\log f(\hat{\theta}) - \log g(\hat{\lambda})\right], \tag{11}$$

where $\hat{\theta}$ and $\hat{\lambda}$ denote the maximum likelihood estimates. $E_f[\cdot]$ is the expectation operator evaluated with respect to the 'true' density $f(\theta)$. This serves as a measure of closeness between the two densities and is defined by the Kullback-Leibler information criterion (KLIC) which is defined by

$$E_f[\cdot] = \int_{\mathbb{R}} \log \frac{f(\theta)}{g(\lambda)} f(\theta)\, dy\,. \tag{12}$$

The quantity in (12) is then minimized by choosing a parameter vector λ that is closest to the true density. This can be equivalently reformulated as

$$\max_{\lambda} E_f [\log g(\lambda)] \ . \tag{13}$$

The solution to this problem is called the pseudo-true value of λ given θ.
Regularity conditions for the applicability of Cox's test are given by White (1982b) for the *i.i.d.* case. It is basically required that consistent and asymptotically normally distributed estimators for the parameters θ and λ can be obtained. This is especially important for the estimator of λ as a consistent and asymptotically normally distributed estimator is needed which allows the likelihood function to fail to correspond to the true joint density of the observations. Such an estimator is the quasi-maximum-likelihood estimator (QMLE). Consistency and asymptotic normality results are provided by White (1982a) for the *i.i.d.* case. Gallant and White (1988) treat the dynamic nonlinear case and derive results regarding the QMLE which ensure the applicability of the Cox test in the setting under consideration.

The main problem when using the Cox test statistic is that the measure of closeness can be analytically derived in a closed form only for very specific and simple cases such as the linear versus log-linear model (see Aneuryn-Evans and Deaton, 1980). It is, however, not possible in general. In cases where the exact expression of the KLIC is very complicated or even impossible to obtain. E.g. for nonlinear models, Pesaran and Pesaran (1993) and Lu et al. (2008) propose simulation based methods for approximating the quantity. Coulibaly and Brorsen (1999) report the results of a simulation study in which they compare different ways of computing the Cox test statistic. Their results suggest that simulation of the whole test statistic and the use of Monte Carlo p-values instead of simulating only parts of the test statistic and relying on asymptotic critical values is the more promising approach in finite samples. These results have also been confirmed more recently by Godfrey and Santos Silva (2005) and Kapetanios and Weeks (2003). The latter authors consider non-nested testing in a time series context to distinguish between several non-nested nonlinear time series models for the conditional mean. Different methods and test statistics based on the likelihood ratio principle are explored. Similar to Coulibaly and Brorsen (1999) Kapetanios and Weeks (2003) find that a studentized but not mean-adjusted test statistic with a simple variance estimator performs best over a variety of different settings; see also Lee and Brorsen (1997) for an application to nonlinear models for the conditional variance.
To derive the test statistic used in our set-up we write the average log-likelihood functions

for the models (9) and (10) as

$$\begin{aligned} L_f(\theta) &= \frac{1}{T}\sum_{t=1}^{T} l_{f,t}(\theta) \\ L_g(\lambda) &= \frac{1}{T}\sum_{t=1}^{T} l_{g,t}(\lambda)\,, \end{aligned}$$

where $l_{i,t}$ is the log-likelihood of model i at time t, $i = \{f, g\}$. Let $\hat{\theta}$ and $\hat{\lambda}$ denote the parameter values maximizing these functions. Then the averaged log-likelihood ratio reads $L_f(\hat{\theta}) - L_g(\hat{\lambda})$. In order to studentize this likelihood ratio Coulibaly and Brorsen (1999) and Kapetanios and Weeks (2003) consider different estimators based on the outer-product of the scores of the models under consideration as in Berndt et al. (1974) and based on the information equality. A third alternative calculates

$$\hat{V}^2 = \frac{1}{T-1}\sum_{t=1}^{T}\left(d_t - \bar{d}\right)^2, \tag{14}$$

where $d_t = l_{f,t}(\hat{\theta}) - l_{g,t}(\hat{\lambda})$ is the likelihood ratio for the t-th observation of y_t and $\bar{d}$ is the respective arithmetic mean. These three methods are asymptotically equivalent but Pesaran and Pesaran (1993), Coulibaly and Brorsen (1999) and Kapetanios and Weeks (2003) report superior performance of the simple variance estimator in (14). Therefore, we adopt this approach in our test.

We consider the test statistic

$$S = \frac{\sqrt{T}\left\{L_f(\hat{\theta}) - L_g(\hat{\lambda})\right\}}{\sqrt{\hat{V}^2}}\,. \tag{15}$$

Note that this test statistic is not mean adjusted by an estimate of a measure of closeness of the two distributions in (9) and (10) such as the KLIC. This reduces the computational burden significantly compared to the methods of Pesaran and Pesaran (1993) and Lu et al. (2008) but renders the test statistic asymptotically non pivotal (see e.g. Pesaran and Pesaran, 1993 and Godfrey, 2007). However, as Hall and Titterington (1989) show, non pivotal statistics will have the same asymptotic accuracy regarding size and power as pivotal statistics. Thus we can reduce the computational burden by using the non mean adjusted statistic in (15) as we only need one bootstrap loop instead of two nested loops for computing an estimate of the KLIC (see Lee and Brorsen, 1997 for a related approach).

We use the following parametric bootstrap to resample the likelihood ratio statistic in (15):

(i) Obtain the initial estimates $\hat{\theta}$ and $\hat{\lambda}$ from y_t and compute the test statistic S in (15).

(ii) Generate bootstrap samples by parametric resampling from the fitted model under the null $f(\hat{\theta})$. y_t^b denotes the t-th observation of the b-th bootstrap sample which is dependent on $\hat{\theta}$, i.e. $y_t^b\left(\hat{\theta}\right)$.

(iii) For the b-th bootstrap sample $y_t^b\left(\hat{\theta}\right)$ let $\hat{\theta}^b$ and $\hat{\lambda}^b$ denote the parameter estimates obtained from maximizing $L_f^b = T^{-1}\sum_{t=1}^T l_{f,t}^b(\theta)$ and $L_g^b = T^{-1}\sum_{t=1}^T l_{g,t}^b(\lambda)$. Use $\hat{\theta}^b$ and $\hat{\lambda}^b$ to compute the bootstrap analog to S in (15):

$$S^b = \frac{\sqrt{T}\left\{L_f^b\left(\hat{\theta}^b\right) - L_g^b\left(\hat{\lambda}^b\right)\right\}}{\sqrt{\hat{V}_b^2}}.$$

(iv) Repeat steps (ii) – (iii) B times and save the bootstrap test statistics S^b. This will give a small sample approximation of the distribution of S in (15).

(v) Compare S from (i) with critical values obtained from the distribution of S^b to decide which model captures the data best.

This bootstrap algorithm can be easily implemented using only a single loop design. It is also possible to exchange the standard parametric bootstrap with the wild bootstrap to allow for unknown forms of conditional or unconditional heteroscedasticity in the data. In this case generate the bootstrap samples in (ii) by a recursive design wild bootstrap (see Gonçalves and Kilian, 2004). Using Rademacher variables the bootstrap residuals are generated as follows:

$$\varepsilon_t^b = \hat{\varepsilon}_t * \eta_t \tag{16}$$

$$\eta_t = \begin{cases} -1 \text{ with probability } 1/2, \\ +1 \text{ with probability } 1/2. \end{cases} \tag{17}$$

A special case arises if the MSAR model is under the null. Here the residuals cannot be obtained directly, but estimated via one-step forecast errors (see Krolzig, 1997). Because of the possible difference of the estimated variance and the variance in the residuals we suggest to use a normalization of the estimated residuals $\hat{u}_t$,

$$\hat{\varepsilon}_t = \frac{\hat{u}}{\hat{\sigma}_{\hat{u}}} \cdot \hat{\sigma}, \tag{18}$$

to receive correctly scaled residuals for the bootstrap without affecting a possibly heteroscedastic structure of the variance.

It should be mentioned again that the test merely decides whether the null model, ESTAR or MSAR, is the better approximation to the data compared to the alternative. It does not indicate that one of the models under consideration is the correct one. Therefore, in order to obtain more decisive results the null and alternative should be exchanged after the first test.

3.1 Finite sample properties

We investigate the finite sample properties of the bootstrap based likelihood ratio test by a Monte Carlo analysis. Therefore, we simulate data generating processes (DGPs) with different parameterizations of the two models under consideration as well as AR(1) series to include the nesting case. We use each process to investigate the behavior with the ESTAR model and with the MSAR model under the null hypothesis. Each experiment is computed with 2000 Monte Carlo repetitions and 500 bootstrap replications. The first 1000 generated values of each time series are discarded. All cases are computed with the normal parametric resampling. An application and comparison of the wild bootstrap can be found in section 4.

For the ESTAR DGPs we use the specification in (4) with the parameter combinations $\psi = \{1.5, 1, 0\}$, $\phi = \{0.7, 0.9, -0.5, -1, -1.4\}$, $\gamma = 1$, $c = 0$ and $\varepsilon_t \sim NID(0,1)$. Only representative parameter combinations which ensure global stationarity are considered. These parameter combinations include the very parsimonious ESTAR model used by Rapach and Wohar (2006) and Taylor et al. (2001). The MSAR model for the Monte Carlo

		H_0: ESTAR		H_0: MSAR	
ϕ	ψ	$T = 200$	$T = 400$	$T = 200$	$T = 400$
1	-0.5	3.3	3.7	31.5	55.1
	-1	2.3	2.3	81.2	98.2
1.5	-1	3.1	2.8	80.0	97.3
	-1.4	2.4	2.4	97.3	100.0
0	0.7	3.0	2.1	53.0	79.9
	0.9	2.1	2.2	65.6	90.8

Table 1: Empirical sizes of nominal 5%-level bootstrap-based likelihood ratio test for different ESTAR DGPs with ESTAR or MSAR under the null.

ϕ	μ_1 - μ_2	H_0: ESTAR		H_0: MSAR	
		$T = 200$	$T = 400$	$T = 200$	$T = 400$
0.5	0	5.75	6.80	7.70	8.60
	1	13.25	21.00	8.00	7.74
	2	71.00	95.65	7.00	5.35
0.8	0	5.25	5.45	9.10	7.94
	1	13.85	25.15	8.40	7.10
	2	93.65	99.85	4.90	3.60
0.95	0	3.90	4.70	14.15	13.35
	1	39.35	70.24	7.25	5.80
	2	99.40	100.00	3.55	3.05

Table 2: Empirical sizes of nominal 5%-level bootstrap-based likelihood ratio test for different MSAR DGPs with ESTAR or MSAR under the null.

analysis is given in (6). The parametrizations for the MSAR DGPs are $\phi = \{0.5, 0.8, 0.95\}$, $-\mu_1 = \mu2 = \{0, 0.5, 1\}$, $p_{11} = p_{22} = 0.9$ and $\varepsilon_t \sim NID(0,1)$. This includes three linear AR(1) nesting cases.

The results for the ESTAR DGPs are shown in Table 1. The size of the test with the ESTAR model under the null is below the nominal level in all cases. This behavior is quite stable, only minor deviations in both directions occur if the number of observations increases. The power of the test with the MSAR model under the null is promising even for small sample sizes. The power increases steeply with T and depends on the difference between the inner and outer regime. The higher the difference, the higher the power. In most cases for $T = 400$ the power is close to 1.

The first two DGPs with $\psi = 1$ and $\phi = \{-0.5, -1\}$ are close to the previous empirical work. Especially in the very parsimonious parametrization with $\phi = -1$ the power is above 80% even for small samples.

The results for the MSAR DGPs are shown in Table 2. Three observations can be made. First, the size properties for the linear nesting cases with the ESTAR model under the null are around the nominal 5% level. Second, the power of the test with the ESTAR model under the null increases with T and rapidly with the difference between the states. The power is near 1 when the difference between μ_1 and μ_2 is at least 2. Third, the size of the test with the MSAR model under the null is above the nominal level for the linear DGPs and small switches. This result occurs because of estimation errors under the null. On average the switch is overestimated, so that the bootstrap distribution shifts to the right and the null is rejected too often. For larger regime switches and larger values of T the size level tends to go to the conservative ESTAR results.

The last group in Table 2 with $\psi = 0.95$ is close to a unit root and to the empirically estimated results by Bergman and Hansson (2005). Surprisingly, the power of the test increases rapidly when ψ goes against one. On the other hand the size distortions due to estimation errors are more volatile than in the other groups.

4 Modeling real exchange rates

We apply the testing procedure on real exchange rates. Real exchange rates are often used to proof (or disprove) the existence of PPP in a world of transaction costs and other market imperfections. On the other hand, PPP can also hold in a world of structural instabilities and Peso effects. The two competing models in this paper, ESTAR and MSAR, capture only one economic theory sufficiently. As pointed out before, every country may face transaction costs as well as some structural instability - the question is which effect dominates the time series. By using two models which are designed for one type of PPP disturbance, we should be able to give a suggestion which model is more appropriate for a specific exchange rate and by that be able to conclude which effect on the exchange rate is stronger.

We use monthly (end of period) real exchange rates of 24 countries against the US dollar to apply the proposed test procedure. The data is from the International Monetary Fund (IMF) and starts January 1973 and ends June 2011. All available data is used, so that a maximum of $T = 462$ observations is available. For the countries in the Euro area $T = 312$ and $T = 150$ observations are available for the Euro.

For real exchange rates Taylor et al. (2001) propose the ESTAR formulation in (4) which we use in the test. Even though Taylor et al. (2001) restrict this model for their final results, we test with the unrestricted version.

This model is tested against the MSAR(1) model of Bergman and Hansson (2005) with a switching mean as in (6). We use $B = 500$ bootstrap repetitions for all tests. Because of the high probability of heteroscedasticity in at least some exchange rates we apply the standard parametric bootstrap as well as the recursive wild bootstrap.

To get a feeling of how persistent the time series are the results of four unit root tests are given in Table 3. In the first three columns the ADF test, the KSS test by Kapetanios et al. (2003) and the inf-t test by Park and Shintani (2009) are presented. All tests are performed with demeaned data and a lag length according to the BIC criterium. Although these tests are designed against a linear or an ESTAR alternative, the tests have also power against a MSAR DGP as shown by Choi and Moh (2007). As it has turned out in former work (see Taylor et al., 2001) that the power of univariate unit root tests in real exchange rate settings is rather poor, it is not surprising that the tests can reject the null only in 4, 3 and 4 cases respectively. These results are supported by Figure 1. This exemplary graph shows that the time series for the estimated ESTAR

	ADF	KSS	inf-t	Kruse (2011)
Argentina	-2.931**	-6.477***	-6.442***	42.910***
Brazil	-1.736	-2.544	-2.528	6.637
Canada	-1.448	-1.442	-1.460	3.157
Chili	-1.295	-0.850	-1.318	9.964*
Colombia	-1.326	-1.080	-1.329	1.291
Denmark	-2.136	-1.754	-2.203	9.601*
Euro Area	-1.032	-1.454	-1.451	3.238
Finland	-2.101	-2.369	-2.456	5.826
France	-1.972	-1.472	-2.044	8.727*
Germany	-2.041	-1.491	-2.044	8.558
Italy	-1.959	-2.230	-2.219	6.270
Japan	-2.237	-2.327	-2.325	8.742*
Mexico	-3.342**	-2.506	-3.449**	14.409***
Netherlands	-2.120	-1.660	-2.139	6.438
Norway	-2.641*	-2.137	-2.684	11.764**
Peru	-1.566	-1.574	-1.943	9.204*
Portugal	-1.567	-1.320	-1.616	3.304
Russia	-3.378**	-4.357***	-4.493***	46.726***
South Africa	-2.224	-2.175	-2.451	4.722
Spain	-1.713	-1.532	-1.717	4.162
Sweden	-2.060	-2.408	-2.442	6.952
Switzerland	-2.245	-2.073	-2.405	5.470
Turkey	-1.706	-4.303***	-4.271***	18.476***
United Kingdom	-2.513	-2.596	-2.638	11.240**

Table 3: Results of different unit root tests. Critical values are {-2.57, -2.87, -3.66} for the ADF test, {-2.66, -2.93, -3.48} for the KSS test, {-3.03, -3.30, -3.86} for the inf-t test and {8.60, 10.17, 13.75} for the Kruse (2011) test. All critical values are given for the {10%, 5%, 1%} significance level.

model is highly persistent, even if deviations occur. On the other hand the results of the unit root test of Kruse (2011) in column four show 11 rejections. This test incorporates the location parameter c. As shown in Table 5, the arithmetic means of the real exchange rates, $\bar{y}$, differ slightly from the estimated equilibrium $\hat{c}$ for all ESTAR models. Therefore, the latter test seems to be the most appropriate one.

The results of the test procedure are presented in Table 4. The column $|S|$ shows the test statistic in absolute value for both testing directions. In 17 cases the log likeli-

		Standard Bootstrap			Wild Bootstrap		
		H_0: ESTAR	H_0: MSAR	Suggested	H_0: ESTAR	H_0: MSAR	Suggested
	|S|	p-Value	p-Value	model	p-Value	p-Value	model
Argentina	2.270	0.000	0.032	-	0.082	0.028	-
Brazil	1.850	0.000	0.252	MSAR	0.320	0.104	-
Canada	0.967	0.054	0.394	MSAR	0.532	0.154	-
Chili	0.615	0.180	0.436	-	0.098	0.406	MSAR
Colombia	1.790	0.008	0.112	MSAR	0.318	0.040	ESTAR
Denmark	0.210	0.562	0.032	ESTAR	0.710	0.032	ESTAR
Euro Area	0.102	0.586	0.008	ESTAR	0.846	0.018	ESTAR
Finland	1.552	0.008	0.550	MSAR	0.032	0.612	MSAR
France	0.339	0.480	0.090	ESTAR	0.478	0.088	ESTAR
Germany	0.806	0.338	0.370	-	0.432	0.318	-
Italy	0.447	0.174	0.440	-	0.168	0.512	-
Japan	0.215	0.528	0.184	-	0.242	0.080	ESTAR
Mexico	0.517	0.182	0.454	-	0.060	0.426	MSAR
Netherlands	1.189	0.202	0.752	-	0.240	0.594	-
Norway	0.910	0.298	0.204	-	0.486	0.194	-
Peru	3.771	0.000	0.570	MSAR	0.002	0.716	MSAR
Portugal	0.338	0.132	0.576	-	0.156	0.558	-
Russia	2.690	0.000	0.134	MSAR	0.098	0.106	MSAR
South Africa	2.420	0.000	0.274	MSAR	0.052	0.104	MSAR
Spain	0.151	0.106	0.588	-	0.108	0.556	-
Sweden	1.730	0.728	0.050	ESTAR	0.732	0.032	ESTAR
Switzerland	0.063	0.588	0.672	-	0.680	0.112	-
Turkey	2.887	0.002	0.364	MSAR	0.000	0.334	MSAR
United Kingdom	0.809	0.392	0.062	ESTAR	0.450	0.054	ESTAR

Table 4: Test results for 24 exchange rates against the US dollar. $|S|$ is the test statistic in absolute value. The other columns contain p-values and model suggestions based on the standard bootstrap and the wild bootstrap test procedure.

hood of the MSAR model is higher, in 7 cases the ESTAR log likelihood. In the next block 'Standard Bootstrap' the first two columns provide the p-values of the test with the normal parametric bootstrap. If the null model is ESTAR, the test rejects the null in 9 cases. If the null model is MSAR, 6 rejections occur. Only one rejection on each side is not mutually exclusive (Argentina). The third column contains the suggested model. A suggestion is only made if one test cannot reject the null and the other test can reject at least at the 10% level. Following this procedure the test provides suggestions in 13 of 24 cases. The block 'Wild Bootstrap' contains p-values and model suggestions of the test procedure using the recursive wild bootstrap. In this case the test rejects 9 times if

Country	$\hat{\gamma}$	$\hat{c}$	$\hat{\sigma}$	LR p-value	$\bar{y}$
Colombia	0.035	7.570	0.025	0.127	7.627
Denmark	0.287	1.957	0.033	0.125	1.858
Euro Area	0.491	-0.126	0.031	0.161	-0.169
France	0.356	1.758	0.032	0.265	1.651
Japan	0.193	4.661	0.033	0.989	4.696
Sweden	0.216	1.927	0.033	0.082	1.886
United Kingdom	0.369	-0.409	0.030	0.169	-0.477

Table 5: Estimation and test results for the restricted ESTAR model with $\psi = 1$ and $\phi = -1$. LR p-value is the p-value of the likelihood ratio test on this restriction. $\bar{y}$ is the arithmetic mean of the time series.

the null model is ESTAR and 8 times if the null model is MSAR. Here the test provides suggestions in 14 cases. A comparison between the bootstrap methods shows that the test decision does not change in 18 of 24 cases. The difference to the former procedure is that Brazil and Canada are not identified as MSAR model anymore, Chili and Mexico are now in the MSAR group and for Japan and Columbia an ESTAR model is suggested. Columbia is the only switch between suggestions. This is not completely surprising since the p-value of the standard bootstrap with the MSAR model under the null is only 11.2%.

In the next step we estimate all exchange rates where at least one bootstrap-type of the test suggests a specific model. The estimation results of the ESTAR model are given in Table 5. Even though we test with an unrestricted ESTAR model to allow for high flexibility, i.e. an explosive inner regime, the estimated parameters for the seven ESTAR suggestions are very close to an inner unit root regime and an outer white noise. Similar to previous work we cannot reject the null that $\psi = 1$ and $\phi = -1$ to the 5% level in all cases. The results show that (i) the Euro has the highest adjustment speed, followed by the British Pound and the French Franc, (ii) that the estimated equilibrium level c is very close to the average of the time series $\bar{y}$ and (iii) that the results are quite close to the results of Taylor et al. (2001) and Rapach and Wohar (2006) given that we use an extended data set.

The results of the 10 MSAR suggestions are presented in Table 6. All countries have a highly persistent autoregressive parameters, Colombia and Turkey are extremely close to a unit root. From the pure parameter estimates all time series fulfill the stationary conditions in (7) and (8). The regime probabilities are also persistent for most countries. One crucial exception in the sample is Turkey. All unit root tests reject the null for this series while the estimation result is more a random walk with drift than a nonlinear switching model. In this case both competing models do not seem to be adequate.

	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\phi}$	$\hat{p}_{11}$	$\hat{p}_{22}$	$\hat{\sigma}$
Brazil	0.041	0.087	0.923	0.988	0.977	0.050
Canada	0.002	0.013	0.965	0.989	0.976	0.018
Chili	0.904	0.823	0.857	0.992	0.984	0.054
Colombia	0.047	0.107	0.993	0.978	0.585	0.021
Finland	0.097	0.125	0.927	0.982	0.960	0.029
Mexico	0.114	2.534	0.954	0.999	0.475	0.053
Peru	0.188	0.351	0.829	0.993	0.990	0.077
Russia	0.148	0.399	0.954	0.984	0.528	0.041
South Africa	0.047	0.194	0.971	0.977	0.171	0.034
Turkey	-0.004	0.297	0.998	0.987	0.001	0.039

Table 6: Estimation results for the MSAR model in (6).

From an economic angle the major difference between the two models is the nonlinear switching behavior. In the MSAR model (6) only exogenous effects affect the series through the mean in a nonlinear fashion. Thus, a country where the MSAR model is suggested for the corresponding real exchange rate should be influenced by exogenous shocks like policy changes or economic crisis. In contrast, in the ESTAR model (4) lagged observations have a nonlinear effect on y_t. Therefore, endogeneous market effects should have a crucial impact on real exchange rates identified as ESTAR.

In a last step we discuss some major historical developments in specific parts of the world and show that our empirical findings match to the expected behavior in general, but that there are also a three exceptions (Finland, Canada and Columbia) and some countries where no suggestion is made but expected. We classify two parts of the world, Western Europe and Latin America. All tested real exchange rates can be assigned to one of these areas except Canada, Japan, Russia, South Africa and Turkey.

For five European countries an ESTAR model is suggested and only one MSAR suggestion can be found. During the sample period the western European countries started to cooperate politically as well as in economic policy in particular within the European Economic Community (EEC). This collaboration got even closer with the transition to the European Community (EC). This period of political stability might explain why exogenous effects have only minor influence on these real exchange rates. Another important development was the European Monetary System from 1979-1993 where the member countries of the EC (first without the UK) fixed their exchange rates against each other.This explains the similar behavior of the series over time. Even though the central banks of the western European countries are independent from each other they pursued a similar interest rate policy to control inflation on a rather low level. This co-

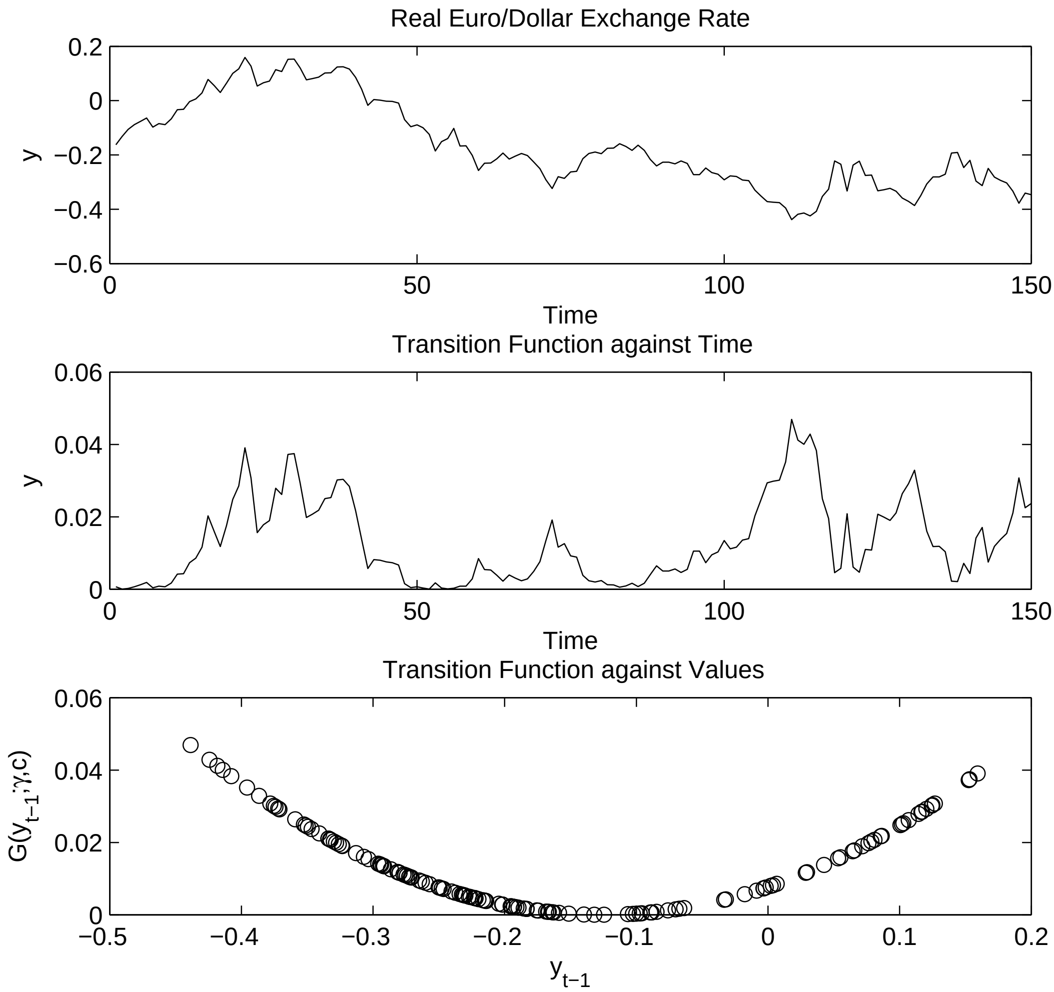

Figure 1: Real exchange rate Euro area with transition function against time and lagged values.

operation led to the adoption of inflation targeting during the 1990's, which is another indicator for stability in this region.

For five Latin American countries at least one type of bootstrap procedure suggests an MSAR model. Columbia is identified as MSAR and ESTAR depending on the bootstrap but shows a complete different behavior than the other Latin American time series. Developments and events which support the MSAR model for this region are exogenous political shocks like the end of the military regime in Brazil in 1985 or the military coup in Peru in 1975. Beside these two major political events many examples of a very volatile and diverse policy making in all Latin American countries during the second half of the twentieth century can be found. Other important external impacts are the oil crises in 1973 and 1979, which led to the Latin American Debt Crisis in 1982 and hyperinflation

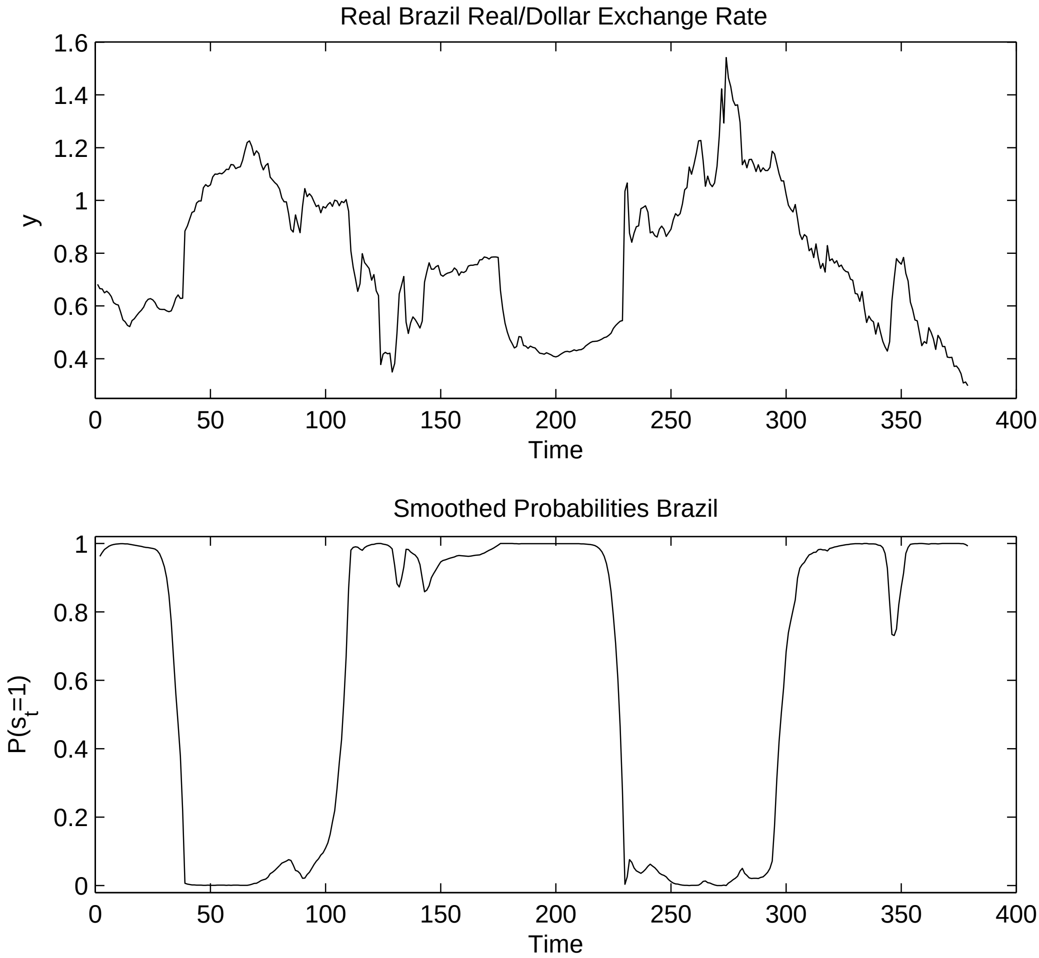

Figure 2: Real exchange rate Brazil with smoothed transition probabilities.

during the 1980s. Another major event was the economic crisis in Mexico in 1994 which had also influence on the other Latin American countries. All these events had an direct effect on the real exchange rate.

Three other countries in the MSAR group show a very similar behavior compared to Latin America: Turkey, Russia and South Africa. Some key events during the sample period are the debt crisis in Russia in 1998, high political and economic instability in Turkey during the 1970s or the end of apartheid in South Africa in 1990. The two remaining countries, Canada and Japan, are only classified for one bootstrap type. Besides neither of the two countries fall in any of the before described clusters. Therefore, no further conclusions for these countries are drawn.

5 Conclusion

Different theories about the existence of the PPP support different economic views of the in-sample dynamic driving real exchange rates. These different views result in different models frequently used in the analysis of PPP, namely ESTAR and MSAR models. As both models are able to support PPP under certain conditions the question which model to use for the analysis is usually answered upon prior economic belief rather than statistical specification. However, as the dynamics of the competing models are rather different the question which model captures the data best is important as it results in different economic theories. In this paper we propose a bootstrap based likelihood ratio test that allows us to discriminate between both classes of nonlinear time series models. The bootstrap approximation of the asymptotic distribution of the test statistic allows us to obtain convincing power results for sample sizes frequently encountered in empirical studies.

In an empirical application we find that the real exchange rates of countries with high inflation rates such as Brazil, Columbia or Peru are modeled best using MSAR models, thus supporting Peso effects and structural instability. Countries not suffering from high inflation like many Western European countries are better described by an ESTAR model. In general continuous adjustment towards a long-run PPP equilibrium can be concluded.

References

Adler, M. and Lehmann, B. (1983). Deviations from the purchasing power parity in the long run. *Journal of Finance*, 38:1471–1487.

Aneuryn-Evans, G. and Deaton, A. (1980). Testing linear versus logarithmic regression models. *The Review of Economic Studies*, 47:275–291.

Bergman, U. M. and Hansson, J. (2005). Real exchange rates and regime switching regimes. *Journal of International Money and Finance*, 24:121–138.

Berndt, E. K., Hall, B. H., Hall, R. E., and Hausman, J. A. (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement*, 3:653–665.

Caporale, G. M., Pittis, N., and Sakellis, P. (2003). Testing for PPP: The erratic behaviour of unit root tests. *Economics Letters*, 80:277–284.

Choi, C.-Y. and Moh, Y.-K. (2007). How useful are tests for unit-root in distinguishing unit-root processes from stationary but non-linear processes? *Econometrics Journal*, 10:82–112.

Coulibaly, N. and Brorsen, B. W. (1999). Monte carlo sampling approach to testing nonnested hypotheses: Monte carlo results. *Econometric Reviews*, 18:195–209.

Cox, D. R. (1961). Tests of seperate families of hypotheses. In *Proceedings of the Fourth Berkley Symposium on Mathematical Statistics and Probability*, volume 1, pages 105–123. University of California Press.

Cox, D. R. (1962). Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society. Series B*, 24:406–424.

Dumas, B. (1992). Dynamic equilibrium and the real exchange rate in a spatially separated world. *The Review of Financial Studies*, 5:153–180.

Edison, H. J. and Klovland, J. T. (1987). A quantitative reassessment of the purchasing power parity hypothesis: Evidence from Norway and the United Kingdom. *Journal of Applied Econometrics*, 2:309–333.

Engel, C. (1994). Can the markov switching model forecast exchange rates? *Journal of International Economics*, 36:151–165.

Engel, C. and Hamilton, J. D. (1990). Long swings in the Dollar: Are they in the data and do markets know it? *The American Economic Review*, 80:689–713.

Francq, C. and Zakoïan, J. M. (2001). Stationarity of mulitvariate Markov-switching ARMA models. *Journal of Econometrics*, 102:339–364.

Gallant, A. R. and White, H. (1988). *A unified theory of estimation and inference for nonlinear dynamic models*. Basil Blackwell.

Godfrey, L. G. (2007). On the asymptotic validity of a bootstrap method for testing nonnested hypotheses. *Economics Letters*, 94:408–413.

Godfrey, L. G. and Santos Silva, J. M. C. (2005). Bootstrap tests of nonnested hypotheses: Some further results. *Econometric Reviews*, 23:325–340.

Gonçalves, S. and Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics*, 123(1):89–120.

Gourieroux, C. and Monfort, A. (1994). Testing non-nested hypotheses. In Engle, R. F. and McFadden, D. L., editors, *Handbook of Econometrics*, volume 4. Elsevier Science Publisher B.V.

Gourieroux, C., Monfort, A., and Trognon, A. (1983). Testing nested or non-nested hypotheses. *Journal of Econometrics*, 21(1):83–115.

Hall, P. and Titterington, D. M. (1989). The effect of simulation order on level accuracy and power of monte carlo tests. *Journal of the Royal Statistical Society. Series B*, 51:459–467.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57:357–384.

Hamilton, J. D. (1993). Estimation, inference and forecasting of time series subject to changes in regime. In Maddala, G., Rao, C., and Vinod, H., editors, *Handbook of Statistics*, volume 11. Elsevier Science Publisher B.V.

Hamilton, J. D. and Raj, B. (2002). New directions in business cycle research and financial analysis. *Empirical Economics*, 27:149–162.

Hegwood, N. D. and Papell, D. H. (1998). Quasi purchasing power parity. *International Journal of Finance and Economics*, 3:279–289.

Kanas, A. (2006). Purchasing power parity and Markov regime switching. *Journal of Money, Credit and Banking*, 38:1669–1687.

Kapetanios, G., Shin, Y., and Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics*, 112:359–379.

Kapetanios, G. and Weeks, M. (2003). Non-nested models and the likelihood ratio statistic: A comparison of simulation and bootstrap based tests. Working Paper Queen Mary, University of London No. 490.

Kilian, L. and Taylor, M. P. (2003). Why is it so difficult to beat the random walk for exchange rates? *Journal of International Economics*, 60:85–107.

Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics*, 6:629–642.

Krolzig, H. M. (1997). *Markov-Switching vector autoregressions: modelling, statistical inference, and application to business cycle analysis*. Springer Berlin.

Kruse, R. (2011). A new unit root test against estar based on a class of modified statistics. *Statistical Papers*, 52(1):71–85.

Lahtinen, M. (2006). The purchasing power parity puzzle: A sudden nonlinear perspective. *Applied Financial Economics*, 16:119–125.

Lee, J. H. and Brorsen, B. W. (1997). A non-nested test of GARCH vs. EGARCH models. *Applied Economics Letters*, 4:765–768.

Lindgren, G. (1978). Markov regime models for mixed distributions and switching regressions. *Scandinavian Journal of Statistics*, 5:81–91.

Lo, M. C. (2008). Nonlinear PPP deviations: A monte carlo investigation of their unconditional half-life. *Studies in Nonlinear Dynamics and Econometrics*, 12:1–29.

Lu, M., Mizon, G. E., and Monfardini, C. (2008). Simulation encompassing: Testing non-nested hypotheses. *Oxford Bulletin of Economics and Statistics*, 70:781–806.

MacDonald, R. (1998). What determines real exchange rates? The long and the short of it. *Journal of International Financial Markets, Institutions and Money*, 8:117–153.

Meese, R. and Rogoff, K. (1983). Emprical exchange rate models of the seventies – Do they fit out of sample? *Journal of International Economics*, 14:3–24.

Meese, R. and Rogoff, K. (1988). Was it real? The exchange rate-interest differential relation over the modern floating-rate period. *Journal of Finance*, 43:933 – 948.

Michael, P., Nobay, A. R., and Peel, D. A. (1997). Transactions costs and nonlinear adjustments in real exchange rates: An empirical investigation. *The Journal of Political Economy*, 105:862–879.

Mizon, G. and Richard, J. (1986). The encompassing principle and its application to testing non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 657–678.

Montañés, A. (1997). Unit roots, level shifts and purchasing power parity. Working Paper University of Zaragoza.

Norman, S. (2010). How well does nonlinear mean reversion solve the PPP puzzle? *Journal of International Money and Finance*, 29:919–937.

Park, J. Y. and Shintani, M. (2009). Testing for a unit root against transitional autoregressive models. *Working Paper*.

Pesaran, M. H. (1987). Global and partial non-nested hypotheses and asymptotic local power. *Econometric Theory*, 3(01):69–97.

Pesaran, M. H. and Pesaran, B. (1993). A simulation approach to the problem of computing Cox's statistic for testing nonnested models. *Journal of Econometrics*, 57:377–392.

Pesaran, M. H. and Weeks, M. (2001). Nonnested hypothesis testing: an overview. In Baltagi, B. H., editor, *Companion to Theoretical Econometrics*. Oxford: Basil Blackwell.

Rapach, D. E. and Wohar, M. E. (2006). The out-of-sample forecasting performance on nonlinear models of real exchange rate behavior. *International Journal of Forecasting*, 22:341–261.

Sarno, L. and Valente, G. (2006). Deviations from purchasing power parity under different exchange rate regimes: Do they revert and, if so, how? *Journal of Banking & Finance*, 30(11):3147–3169.

Taylor, M. P., Peel, D. A., and Sarno, L. (2001). Nonlinear mean-reversion in real exchange rates: Toward a solution to the purchasing power parity puzzles. *International Economic Review*, 42:1015–1042.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89:208–218.

Tjøstheim, D. (1986). Estimation in nonlinear time series models. *Stochastic Processes and their Applications*, 21:251–273.

van Dijk, D., Teräsvirta, T., and Franses, P. H. (2002). Smooth transition autoregressive models - a survey of recent developments. *Econometric Reviews*, 21:1–47.

van Norden, S. (1996). Regime switching as a test for exchange rate bubbles. *Journal of Applied Econometrics*, 11:219–251.

White, H. (1982a). Maximum likelihood estimation of misspecified models. *Econometrica*, 50:1–26.

White, H. (1982b). Regularity conditions for Cox's test of non-nested hypotheses. *Journal of Econometrics*, 19:301–318.