



Munich Personal RePEc Archive

Does the purchasing power parity hypothesis hold after 1998?

Zanetti Chini, Emilio

University of Rome "Tor Vergata" - Department in Economics and Institutions

15 November 2010

Online at <https://mpra.ub.uni-muenchen.de/27225/>
MPRA Paper No. 27225, posted 06 Dec 2010 01:02 UTC

Does the Purchasing Power Parity hypothesis hold after 1998?*

Emilio Zanetti Chini[†]

15/11/2010

Abstract

We investigate the empirical support to the Purchasing Power Parity hypothesis by using sixteen real exchange rates for the decade 1999-2009. The literature has recently arrived to a solution to the two PPP puzzles if considering the post-Bretton Woods period from 1975 to 1998. Time series-based studies consider few cases, while panel-based studies have been recently criticized. Multivariate and panel cointegration, and nonlinear models are here implemented. The theory is rejected and both the puzzles remain unsolved if considering a linear structure, while a nonlinear scenario seems to allow for a partial solution to the first puzzle.

Key words: PPP, unit roots, cointegration, nonlinear models, IRF.

JEL: C32, C33, C50, F31.

1 Introduction

Real exchange rates are source of one of the six main puzzles in macroeconomics: the Purchasing Power Parity (PPP) puzzle, explicitly defined in Rogoff (1996)¹. The main

*This paper is based on my MSc Dissertation in Macroeconometrics at University of Rome "Tor Vergata". I thank my supervisor Tommaso Proietti, Giovanni Trovato and Barbara Annichiarico for their useful comments. A special thanks to Luigi Benfratello for his support.

[†]University of Rome II "Tor Vergata", Department of Economics and Institutions, Via Columbia 2, 00133, Rome, Italy - *E-mail*: Emilio.Zanetti.Chini@uniroma2.it - *Phone*: (+39) 06 72595715 - *Mobile*: (+39) 340 9641055 - *Fax*: (+39) 06 2020500.

¹[There is] a surprising degree of consensus on a couple of basic facts: first, [...] real exchange rates [...] tend toward PPP in the very long run. Consensus suggests, however, the speed of convergence towards PPP is extremely slow [...]. Second, short run deviation from PPP are large and volatile. [...] The question is: how can one reconcile the enormous short-term volatility with the extremely slow rate at which shocks appear to damp out ?"(Rogoff, 1996, pag. 647).

purpose of this article is to investigate the empirical support to the PPP hypothesis for the last 11 years. This implies, from an econometrical point of view, to search for two main objects: mean reversion in real exchange rates and cointegrating relations in their components (nominal exchange rates, domestic and foreign prices). In order to do this we compare three main methodologies: linear cointegration analysis, panel unit root and cointegration and univariate nonlinear autoregressive models.

Since PPP is assumed in many widely used macroeconomic models, the literature is huge and only recently some positive results have been achieved (Sarno and Taylor, 2001). A whole research field in applied econometrics has grown in order to find persistencies in this kind of data. The workhorses are: the family of smooth transition regression (STR) models (Bacon and Watts, 1971), in which the more specific smooth transition autoregressive (STAR) models (Chan and Tong, 1986) and (self-exciting) threshold autoregressive ((SE)TAR) models (Tong, 1983) (see Section 2) are nested; and panel cointegration techniques Pedroni (2001, 2004). Taylor et al. (2001) (TPS) applies these nonlinear techniques to four rates and solve the two PPP puzzles by analyzing the standard post-Bretton-Wood sample. Panels was also successful since last '90s literature, until Banerjee et al. (2005) (BMO) notice that commonly used panel unit root test critical values, if not allowing for cross-countries cointegrating relationships, are severely biased towards rejecting the null hypothesis of a unit root; this implies a critique to the automatic use of panel data in macro. Finally, Juselius (2006) shows an alternative, fully empirically-based approach to cointegration analysis which offers a new perspective on these (and, at least potentially, on many other) puzzles. Johansen et al. (2010) solves the PPP puzzles for the DKR/\$ rate using the standard sample by implementing a CVAR model under I(2) scenario.

This work originates from three findings: first, almost all the most influential studies - and, *in primis*, the ones supporting the theory - are based on a very peculiar sample (1975:04-1998:12 at the best) and on few currencies. Second and consequently, none of such studies (also the most recent ones) mention the euro; this seemed to us a serious lack in empirical literature. Third, the model specification in almost all the literature is highly driven by theoretical reasons (in particular in nonlinear models, see Section 3). Moreover, the recent financial crisis gives us an opportunity to catch an important shock to take in account when holding with evaluation. In the next sections we will try to bridge this gap

in empirical literature and to compare the "traditional" econometric models for last 11 years data and will check whether the conclusions of the previously mentioned studies are still valid or not. Secondly, we will extensively discuss the issue of models' specification introducing modifications whether necessary that allow us to estimate a higher number of models for real exchange rates than the ones we would estimate whether considering a more theory-based specification.

The paper is organized as following: Section 2 states the relations of interest for testing PPP hypothesis and briefly describes the statistical models; Section 3 points out the empirical strategy; Section 4 describes the data set; Section 5 shows the empirical evidence of both weak and strong PPP hypotheses for our dataset for each methodology used; Section 6 concludes.

2 The economic theory and the models

2.1 The economic theory of PPP

Following Juselius (2006) notation and working with aggregate terms in logarithmic transformation, we define the PPP as:

$$p_t = p_t^* + s_t + v_t \quad (1)$$

where p_t and p_t^* are the domestic and foreign consumer price indexes, s_t are defined as above and v_t is the time t error term. Hence the model can be written in deviation from PPP, which correspond to what literature calls "*strong* PPP hypothesis":

$$v_t \equiv y = p_t - p_t^* - s_t, \quad (2)$$

where y corresponds to the real exchange rate. The "*weak* PPP hypothesis", is a generalization of model (2) and is defined as:

$$\hat{v}_t \equiv \hat{y}_t = p_t - \alpha p_t^* - \beta s_t \quad (3)$$

where α and β represent measurement errors as transaction and transport costs and the hat is only for notation.

In term of cointegrating relations we can state two postulates:

Postulate 1. *If strong PPP holds, the corresponding cointegrating relation is:*

$$CI = (1 \quad -1 \quad -1) \quad (4)$$

Postulate 2. *If weak PPP holds, $\exists CI$ s.t. $\hat{y}_t \sim I(0)$*

where CI indicates the cointegrating relation and $I(0)$ indicates integrated of order zero process. Testing for strong PPP means testing for unit root of real exchange rates, while testing for weak PPP means testing for cointegration.

We define the two PPP puzzles directly from Postulates 1 and 2:

Definition 1 (1st Puzzle). *Neither Postulate 1 nor 2 holds. That is, the real exchange rates deviate systematically from their theoretical (PPP) values.*

Definition 2 (2nd Puzzle). *These deviation are permanent in the long run, contrary to what the economic theory suggests.*

2.2 Statistical models: CVAR

For what concerns the cointegration analysis of PPP, we use a VAR(p) to model relations (1) and (2):

$$\mathbf{y}_t = \Pi_1 \mathbf{y}_{t-1} + \Pi_2 \mathbf{y}_{t-2} + \dots + \Pi_p \mathbf{y}_{t-p} + \Phi D_t + \epsilon_t \quad (5)$$

where $t = 1, \dots, T$ and $T = 132$, $\mathbf{y}_t = \begin{bmatrix} p_t & p_t^* & s_t \end{bmatrix}'$, $\epsilon_t \sim IN_p(\mathbf{0}, \Omega)$ and IN means identically and p-normally distributed. Rewriting model (5) in VECM form we get:

$$\Delta \mathbf{y}_t = \Gamma_1^{(1)} \Delta \mathbf{y}_{t-1} + \Gamma_2^{(1)} \Delta \mathbf{y}_{t-2} + \dots + \Gamma_{p-1}^{(1)} \Delta \mathbf{y}_{t-p-1} + \alpha \beta' x_{t-1} + \mu_0 + \mu_1 t + \epsilon_t \quad (6)$$

where: $\Gamma_1^{(1)} = -(\Pi_2 + \Pi_3 + \dots + \Pi_p)$, $\Gamma_2^{(1)} = -\Pi_3$ and $\Pi = -(I - \Pi_1 - \Pi_2 - \dots - \Pi_p)$ are the short run matrices and the long run matrix respectively and the integer (1) indicates the lag placement of ECM, $\Pi = \alpha \beta'$ is the reduced rank long run matrix, α and β are $p \times r$ matrices, $r \leq p$, $\mu_0 + \mu_1 t = \Phi D_t$ are the unrestricted components (i.e.

allowed to enter in cointegrating relation) of deterministic trend. The equation (6) is the cointegrated VAR (CVAR) model under I(1) hypothesis (see Johansen (1991) for further details and estimation).

2.3 Statistical models: panel methods

For what concerns panel data methods, the general model can be formulated as the following regression:

$$\Delta y_{it} = \rho_i y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{i,L} \Delta y_{i,t-L} + \alpha_{mi} d_{mt} + \epsilon_{it} \quad m = 1, 2, 3 \quad (7)$$

where: $y_{it} = [p_{it}, p_{it}^*, s_{it}]'$, $\epsilon_{it} \sim IID(0, \sigma^2)$, $E(\epsilon_{it}\epsilon_{jt}) = 0$, $i \neq j \forall t$, d_{mt} indicates the vector of deterministic terms and α_{mi} the corresponding vector of coefficients for model $m = 1, 2, 3$ and p_i is unknown. In particular, $d_{1,t} = \emptyset$, $d_{2,t} = \{1\}$ and $d_{3t} = \{1, t\}$.

By starting from model (7) we can test for unit root (that is, for strong PPP) the panel of exchange rates using a battery of tests allowing for slightly more general assumptions and making the investigator able to answer to three different questions: (i) is panel supporting strong PPP? (ii) Conversely, is panel rejecting strong PPP? (iii) Finally, are there cointegrating cross-sections (that is, is panel supporting weak PPP)? Levin et al. (2002) (LLC), Im et al. (2003) (IPS), Pesaran (2007) (CADF), Maddala and Wu (1999) (MW) are used to answer to question (i). Hadri (2000) and Nyblom and Harvey (2000) (NH) answer to question (ii). Pedroni (2004) and Westerlund (2007) answer to question (iii).

We refer to the original papers for technicalities. We just underline that these different tests are today used to analyze the nonstationary behavior of data from slightly different perspectives; that is, since a test which is robust to all possible features in the panel does not exist, a battery of partial tests can be build in order to cover particular lacks which remain unsolved by other tests (see BMO as an example). In particular the LLC test has the strongest hypothesis system: each series is unit root against each series is stationary. For this reason the LLC test is one of the more frequently used and criticized. The IPS test solves this problem but the cost is that it can be applied only to balanced panels; moreover, both LLC and IPS are built under cross-sectional independence hypothesis. This

last peculiarity is treated by CADF test while MW test is in turn the solution to IPS lack of adequacy in unbalanced panels and by construction can be used for other unit root test. Again, the problem is in that p -values needed to perform it have to be computed by Monte Carlo simulation. Concerning the tests for the opposite null of stationarity, the Hadri test is the the panel analogue of univariate Kwiatkowski et al. (1992) (KPSS) test. Differently, the NH test is its multivariate version which allows to test the presence of an additive random walk in the data generating process. Concerning panel cointegration, the first tests was build up by McCoskey and Kao (1998) and Kao (1999) who used LM and ADF-based procedure in order to test two opposite null hypothesis systems (no cointegration and cointegration respectively). However we choose to implement two on the seven tests by Pedroni (2004), differently to the other two, it allows for individual heterogeneity, fixed effects and trends terms. Westerlund (2007) uses a different kind of test in order to test the same null hypothesis of no cointegration, but its statistics are more powerful than Pedroni's ones.

2.4 Statistical models: univariate nonlinear models

Concerning univariate nonlinear models, we use the standard STAR/(SE)TAR models in order to replicate the analysis by TPS. Granger and Teräsvirta (1993) recommends a specific-to-general modelling procedure based on the following steps: (i) select an appropriate linear AR(p) model for the series under investigation; (ii) test the null hypothesis of linearity against the alternative of STAR/SETAR-type nonlinearity and select the appropriate transition variable(s); (iii) estimate the parameters; (iv) evaluate the model using diagnostic tests; (v) if necessary, modify the model; (vi) use the model for descriptive or forecasting objectives. We now broadly describe the econometric methodology step by step; for technicalitis, see Tsay (1989), Hansen (1996) ((SE)TAR models) and Teräsvirta (1994) (STAR models).

Consider the general additive non-linear model:

$$y_t = \phi' z_t + \theta' z_t G(\gamma, \mathbf{c}, s_t) + \epsilon_t \quad (8)$$

where $y_t \equiv v_t$ in equation (2), $\mathbf{z}_t = (1, y_1, \dots, y_{t-p})'$, $\boldsymbol{\phi} = (\phi_0, \phi_1, \dots, \phi_p)'$, $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_p)'$ are parameter vectors, and $\epsilon_t \sim i.i.d.(0, \sigma^2)$, the transition function $G(\gamma, \mathbf{c}, s_t)$ is a continuous function in the transition variable s_t ² where γ is the slope parameter and controls the velocity of the transition, $\mathbf{c} = (c_1, \dots, c_K)$ is a vector of (transition') location parameters. One of the main used functions for $G(\cdot)$ is the (first order) logistic function:

$$G(\gamma, \mathbf{c}, s_t) = \left(1 + \exp \left\{ -\gamma \prod_{k=1}^K (s_t - c_k) \right\} \right)^{-1}, \quad \gamma > 0, \quad (9)$$

where $\gamma > 0$ is an identifying restriction. Equations (9) and (8) define the Logistic STR (LSTR) model. The most common choices for K are $K = 1$, in which case the parameters $\boldsymbol{\phi} + \boldsymbol{\theta}G(\gamma, \mathbf{c}, s_t)$ change monotonically as a function of s_t from $\boldsymbol{\phi}$ to $\boldsymbol{\phi} + \boldsymbol{\theta}$ and $K = 2$, in which case the parameters $\boldsymbol{\phi} + \boldsymbol{\theta}G(\gamma, \mathbf{c}, s_t)$ change symmetrically around the mid-point $(c_1 + c_2)/2$ where the logistic function attains its minimum, $\min_G G(\cdot) \in [0, 1/2]$, and it's such that:

$$\min_G G(\cdot) = \begin{cases} 0 & \text{if } \gamma \rightarrow \infty \\ 1/2 & \text{if } c_1 = c_2 \text{ and } \gamma < \infty \end{cases}$$

For notational convenience, we will consider only the first case ($K = 1$). If $\gamma = 0$, the transition function $G(\gamma, \mathbf{c}, s_t) \equiv 1/2$ so that model (8) nests a linear model. When $K = 2$ and $c_1 = c_2$ the transition function (9) becomes:

$$G(\gamma, \mathbf{c}, s_t) = 1 - \exp\{-\gamma(s_t - c)^2\}, \quad \gamma > 0 \quad (10)$$

Equations (8) and (10) define the Exponential STR (ESTR) model.

When $\mathbf{z}_t \equiv \mathbf{y}_{t-d}$ and $s_t \equiv y_{t-d}$, $d > 0$ in (9) and (10), the model becomes an LSTAR and an ESTAR respectively. Similarly, when $\gamma \rightarrow \infty$ and $\mathbf{z}_t \equiv \mathbf{y}_t$ and $s_t \equiv y_{t-d}$ the model (8) nests a SETAR model:

$$y_t = \sum_{j=1}^{r+1} (\boldsymbol{\phi}'_j \mathbf{y}_t) I(y_{t-d} \leq c_j) + \sum_{j=1}^{r+1} (\boldsymbol{\phi}'_j \mathbf{y}_t) I(y_{t-d} > c_j) + \epsilon_{jt} \quad (11)$$

²Notice that here s_t is a generic transition variable which can coincide (but not necessarily) with $y_t \neq s_t$. This change in notation is only for convenience when comparing the literature in STR models.

where ϕ, \mathbf{y}_t are defined as before, s_t is a continuous switching r.v., c_0, c_1, \dots, c_{r+1} are threshold parameters, $c_0 = -\infty, c_{r+1} = +\infty, \epsilon_{jt} \sim i.i.d.(0, \sigma_j^2), j = 1, \dots, r$.

Concerning step (i) (specification), Tsay proposes a four-step specification procedure for SETAR model: select the AR order p and the set of possible threshold lags S , fit arranged autoregressions for a given p and every element d of S and perform threshold nonlinearity test $\hat{F}(p, d)$; if some nonlinearity is detected, select the delay parameter d_p such that $\hat{F}(p, d_p) = \max_{v \in S} \{\hat{F}(p, v)\}$; for given p, d , locate the threshold variables by using scatterplot of predictive residuals derived by the arranged autoregression against y_{t-d} ; finally, refine the order and threshold values by linear techniques. Teräsvirta proposes a similar procedure for STAR models: specify a linear AR(p) model; test linearity for different values of d and, if rejected, determine the d parameter following the same criterion above mentioned.

Concerning step (iii), the estimation is done by OLS in (SE)TAR models while in STAR models the NLLS algorithm is required.

The step (ii) (Linearity testing) for (SE)TAR models is discussed in Tsay. The idea is to perform an arranged autoregression and the resulting parameter are estimated by recursive least squares. The resulting predictive and standardized predictive residuals are used to build the F-type test from a least square regression. Hansen discusses an alternative likelihood-based test. Three statistics are used:

$$\begin{aligned} S_T &= \sup_{\gamma \in \Gamma} S_T(\gamma), \\ ave S_T &= ave S_T(\gamma) = \int_{\Gamma} S_T(\gamma) dW(\gamma), \\ \exp S_T &= \ln \left(\int_{\Gamma} \exp \left\{ \frac{1}{2} S_T(\gamma) \right\} dW(\gamma) \right) \end{aligned} \quad (12)$$

where $\Gamma = \{\gamma : \gamma \in \Gamma\}$, $W(\gamma)$ is a weight function such that $\int_{\Gamma} W(\gamma) d\gamma = 1$. Using the likelihood function of the model (8), Hansen derives the the score function for Wald and LM test; the empirical distribution of the last one is computed by bootstrap simulation and can be used to show whether the null hypothesis has to be rejected or not. The analogue test for STAR models is discussed in Luukkonen et al. (1988). It is based on a Taylor expansion of the transition function (9) (or (10)), $T_3(z) = g_1 z + g_3 z^3$ where

$g_1 = \partial G / \partial z|_{z=0}$ and $g_3 = (1/6)\partial^3 G / \partial z^3|_{z=0}$, so that the approximation:

$$y_t = \boldsymbol{\phi}' \mathbf{z}_t + \boldsymbol{\theta}' \mathbf{z}_t T_3(\gamma(y_{t-d} - c)) + \boldsymbol{\epsilon}_t \quad (13)$$

leads to the auxiliary regression:

$$\hat{\boldsymbol{\epsilon}}_t = \hat{\mathbf{z}}'_{1t} \tilde{\boldsymbol{\beta}}_1 + \sum_{j=1}^p \beta_{2j} y_{t-j} y_{t-d} + \sum_{j=1}^p \beta_{3j} y_{t-j} y_{t-d}^2 + \sum_{j=1}^p \beta_{4j} y_{t-j} y_{t-d}^3 + v'_t \quad (14)$$

The null hypothesis for linearity against LSTAR is $H_0 : \beta_{2j} = \beta_{3j} = \beta_{4j} = 0, j = 1, \dots, p$, which, under the conditions that a linear autoregressive model holds and $E\epsilon_t < \infty$, is tested by statistic:

$$LM_2 = (SSR_0 - SSR) / \hat{\sigma}^2 \sim \chi^2(3p) \quad (15)$$

where SSR are the sum of squared residuals from equation (14) or, alternatively, by setting the artificial model:

$$y_t = g_1 \gamma_0 + \gamma'_1 \mathbf{z}_t + \gamma'_2 (\mathbf{z}_t y_{t-d}) + \gamma'_3 (\mathbf{z}_t y_{t-d}^2) + \gamma'_4 (\mathbf{z}_t y_{t-d}^3) + v''_t \quad (16)$$

where: $v'' \sim nid(0, \sigma_{v''}^2)$, $\gamma_j = (\gamma_{1j}, \dots, \gamma_{jp})'$, and $j = 1, \dots, 4$, and $H_0 : \gamma_2 = \gamma_3 = \gamma_4 = 0$. In terms of Taylor approximations we get:

$$\begin{aligned} \gamma_2 &= g_1 \gamma \hat{\boldsymbol{\theta}} + 3g_3 \gamma^3 c^2 \hat{\boldsymbol{\theta}} - 3g_3 \gamma^3 c \theta_0 \mathbf{e}_d \\ \gamma_3 &= -3g_3 \gamma^3 c \hat{\boldsymbol{\theta}} + g_3 \gamma^3 \theta_0 \mathbf{e}_d \\ \gamma_4 &= g_3 \gamma^3 \hat{\boldsymbol{\theta}} \end{aligned} \quad (17)$$

where $\hat{\boldsymbol{\theta}}$ and \mathbf{c} and d are previously defined. Similarly, if the model is an ESTAR(p) model, $\hat{\mathbf{z}}_{1t} = -\mathbf{z}_t$ and $\hat{\mathbf{z}}_{2t}(\boldsymbol{\pi}) = -(y_{t-d} - c)^2 (\hat{\boldsymbol{\theta}}' \mathbf{z}_t) = -(\boldsymbol{\theta}' \mathbf{z}_t y_{t-d}^2 + \theta_0 y_{t-d}^2 - 2c \bar{\boldsymbol{\theta}}' \mathbf{z}_t y_{t-d} + c^2 \bar{\boldsymbol{\theta}}' \mathbf{z}_t - 2c \theta_0 y_{t-d} + c^2 \theta_0)$. This yields to the following auxiliary regression:

$$\hat{v}_t = \tilde{\boldsymbol{\beta}}'_1 \hat{\mathbf{z}}_{1t} + \boldsymbol{\beta}'_2 \mathbf{z}_t y_{t-d} + \boldsymbol{\beta}'_3 \mathbf{z}_t y_{t-d}^2 + e'_t \quad (18)$$

where \hat{v}_t is the analogue of $\hat{\boldsymbol{\epsilon}}_t$, e'_t is an error term and $\tilde{\boldsymbol{\beta}}_1 = (\beta_{10}, \boldsymbol{\beta}'_1)'$ with $\beta_{10} = \phi_0 - c^2 \theta_0$ and $\boldsymbol{\beta}_1 = \bar{\boldsymbol{\phi}} - c^2 \bar{\boldsymbol{\theta}} + 2c \theta_0 \mathbf{e}_d$, $\boldsymbol{\beta}_2 = 2c \bar{\boldsymbol{\theta}} - \theta_0 \mathbf{e}_d$. The null of linearity is $H'_0 : \boldsymbol{\beta}_2 = \boldsymbol{\beta}_3 = 0$

which is tested by statistic

$$LM_3 = (SSR_0 - SSR)/\hat{\sigma}^2 \sim \chi^2(p) \quad (19)$$

where SSR is the sum of squared residuals from (18). In order to choose the correct transition function, Teräsvirta proposes the following nested hypotheses test (the so called "Teräsvirta rule"):

$$\begin{aligned} H_{04} : \gamma_4 = 0 \text{ against } H_{14} : \gamma_4 \neq 0 \text{ in (17).} \\ H_{03} : \gamma_3 = 0 \mid \gamma_4 = 0 \text{ against } H_{13} : \gamma_3 \neq 0 \mid \gamma_4 = 0 \text{ in (17).} \\ H_{02} : \gamma_2 = 0 \mid \gamma_3 = \gamma_4 = 0 \text{ against } H_{12} : \gamma_2 \neq 0 \mid \gamma_3 = \gamma_4 = 0 \text{ in (17).} \end{aligned} \quad (20)$$

if the p -value of H_{03} is the smallest of the three, select an ESTAR model; otherwise, select an LSTAR model.

Concerning the step (iv) (Diagnostic tests), Eitrheim and Teräsvirta (1996) provides three LM tests for serial auto-correlation, remaining nonlinearity and parameter constancy.

Finally, the last step (Evaluation and/or forecasting) can be performed by using the impulse response functions (IRF). Formally the "traditional" impulse response function (TIRF) is defined as:

$$\begin{aligned} TIRF(h, \delta, \omega_{t-1}) = & E[y_{t+h} \mid \epsilon_t = \delta, \epsilon_{t+1} = \dots = \epsilon_{t+h} = 0, \omega_{t-1}] - \\ & - E[y_{t+h} \mid \epsilon_t = 0, \epsilon_{t+1} = \dots = \epsilon_{t+h} = 0, \omega_{t-1}], \end{aligned} \quad (21)$$

for $h = 0, 1, 2, \dots$. The TIRF is commonly used in linear systems because of its three properties: first, it's symmetric, that is a shock of size $-\delta$ has an effect exactly opposite to that of shock of size δ ; second, it's proportional to the size of the size of the shock; third, it's history independent, that is its shape does not depend on the particular history ω_{t-1} . These properties does not hold in nonlinear models.

In order to solve this problem, Koop et al. (1996) proposes a generalization of (21), called Generalized Impulse Response Function (GIRF). The GIRF for a shock $\epsilon_t = \delta$ and history ω_{t-1} , for both δ and ω_{t-1} function of the random variable ϵ_t and Ω_{t-1} (the set of all possible

histories $\{\omega_{t-1}\}$, is defined as:

$$GIRF(h, \epsilon_t, \Omega_{t-1}) = E[y_{t+h}|\epsilon_t, \Omega_{t-1}] - E[y_{t+h}|\Omega_{t-1}]. \quad (22)$$

for $h = 0, 1, 2, \dots$. In linear models TIRF and GIRF coincide. The applied econometric literature uses the IRF analysis in order to study the second PPP puzzle, namely to measure the half-life of the deviation of real exchange rates from their theoretical PPP value.

3 Empirical strategy

We follow the Juselius' "Marshallian" approach to cointegration analysis when testing weak PPP hypothesis because of its completeness and its agnosticism. It can be summarized in three main points: first, much more importance is put on the specification rather than the prior role of a theoretical economic model. Second, and consequently, the theoretical model is re-parametrized in such a way that all possible testable hypotheses can be analyzed. Third, the econometrician should minimize the restrictions that could be needed during the specification in order to let the data speak freely.

The Marshallian strategy briefly described above is in contraposition to the theory-based "Walrasian" approach, represented by DSGE family of models³. However, Juselius' philosophy clearly presents some problems, first of all the probability of rejection of investigator' searched relation, relatively higher than in any other theory-based econometric model. A second more important problem is the treatment of extraordinary events, modeled by using shift and blip dummies⁴; namely, the problem is in that such dummies derive from the search of large errors in distribution of the series, for which not always an economic explanation is available and in that parameters estimates are strongly sensitive to such dummies and linear trends entering in cointegrating vectors. A third problem is that this strategy is currently not available for other methodologies. In that case we' re forced to

³The VAR procedure is less pretentious about the prior role of a theoretical economic model, but it avoids the lack of empirical relevance the theory-based approach has often been criticized for" (Juselius, 2006, Ch. 1, pag. 9)

⁴That is, a series could present several changes in terms of mean, trends or transitory shift in levels and these can easily observed by fitting a (preliminary) model assuming gaussian errors and then the differences between this and the original series: if such differences is grater than a pre-specified threshold, they can be seen as extraordinary events to take in account in phase of specification.

use a more traditional theory based approach⁵.

The Marshallian approach to cointegration analysis allows us to choose the appropriate specification for model (3) for the j -th system of country, that is to check if $(p_j - p_j^* - s_j) \sim I(0)$ for $j = CAN, DN, JPN, NW, SD, SZ, UK, EU$ or US . It is implemented by the following step procedure (see Juselius (2006) for details):

- Step 1: Select the lag p for the system (6) by using standard information criteria.
- Step 2: Once p is selected, check for normality and residuals first-order and second-order autocorrelation; if autocorrelation and non-normality is found, augment p until necessary.
- Step 3: Check for the presence of transitory or permanent shocks in the series. This can be done by looking at residuals, imposing a threshold and checking for the presence of errors exceeding it; if outliers exist, impose an appropriate dummy variable in the month corresponding to the outlier and repeat the procedure from Step 2.
- Step 4: Perform the Johansen' Rank test in order to check for the presence of cointegrating relations in the system. Since the test is not invariant to changes in deterministic kernel, if outlier are found in Step 3, simulate the critical values.
- Step 5: If $rank(\mathbf{\Pi}) \neq 0$, set the rank of $\mathbf{\Pi}$ matrix.
- Step 6: Set the identifying restriction corresponding to the searched relation (4); the (possibly, more than one) resulting restricted (cointegrated) VECM model(s) are selected if the p -value is sufficiently high.
- Step 7: The so found restricted VECM are analyzed in their cointegrating relations by recursive tests.
- Step 8: Check for the presence of I(2)-ness. If found, the whole analysis should be reconsidered in an I(2) scenario.

⁵However one should consider that the empirical problem in this article is relatively simple, and the data set used relatively small so the two approaches, in principle, should coincide.

In the nonlinear framework we follow TPS as benchmark and apply Teräsvirta' analysis on our dataset. TPS finds that the exchange rates under investigation are nonlinear mean reverting. In particular, the delay parameter is *a priori* considered as small (near 1) although a grid search by NLLS is performed. Then the model (8) is restricted for $\phi = -\theta$, because this ensure an economic interpretation similar to Coleman (1995)⁶. Using the Monte Carlo method by Gallant et al. (1993) (GRT) for GIRFs, TPS finds that large shocks are faster mean reverting than smaller ones. This means that also the second PPP puzzle is solved. However, this result has strongly driven from the *a priori* choice for exponential smooth transition for STAR models, which is justified it by its property of symmetric adjustment of transition variable around an equilibrium level⁷.

Since the ESTAR model is a particular case of the more general family of LSTR, we find such *a priori* unjustified, specially for a problematic dataset as our one. Moreover, we found that the above mentioned restriction caused an artificial reduction of estimated parameters' p-values. For this reason, we follow a more agnostic policy in modelling our series, allowing some parameters (the constant, in a lot of cases) for no restrictions. The resulting Granger-Tersvirta modelling procedure has been consequently adapted as here described in detail:

- Step 1: *Selection of p-order*. Use AIC, BIC, and HQ criteria, with particular attention to the second one because is known to be the more conservative. Then, in order to take in account the possibility of the presence of non gaussian residuals due to high instability of the series, and secondly to explore its ability to testing for third-order residual correlation, implement an Hinich (1996) test and corrected the choice of lag order when necessary. In particular, this is the criterion in adjusting the AR order: when the one of both Hinich' statistics *p*-values are less the 0.10, add one or more lag in function of its nearness to 0; when slightly higher than such threshold, the test is able to reject the null of no third-order autocorrelation, so the order is not increased.

⁶"[This restriction] implies an equilibrium log-level of real exchange rates [called μ and found being zero] in the neighborhood of which, real exchange is close to a random walk, beginning increasingly mean reverting as going far away from it" (TPS, pag. 14).

⁷TPS consider the logistic smooth transition inappropriate because "[...] It's hard to think economic reasons why the speed of adjustment of the real exchange rate should vary according to whether the dollar is over-valued or undervalued, specially if one is thinking of goods arbitrage as ultimately diving the impetus towards the long run equilibrium and one is dealing with major dollar exchange rates against the currencies of other developed countries" (TPS, pag. 7).

However, this test remain as boldly indicative and does not constrain us to a limit in adding lags; this postulate that a limit of three-four added lags is a reasonable choice. Consider such "auxiliary" lags as potential: this means, start the specification procedure by giving priority to orders selected by traditional criteria.

- Step 2: *Specification of linear AR(p) part of the model (8)*. Allow for the possibility for model (8) to have some zero-coefficient in both linear and nonlinear part, starting with the hypothesis of no restrictions.
- Step 3: *Linearity tests*. Apply the Teräsvirta rule (20) for all possible candidates transition variable $s_t = \tilde{z}_t \equiv (y_{t-1}, \dots, y_{t-p}, t)'$; that is, consider as candidates all lags of $\{y_t\}$ and a linear trend t . If the model is linear for all possible candidate, return to Step 2 and start to restrict the model until some nonlinearity is detected. In particular, use a progressive criterion in putting restriction: start with one zero coefficient, then augment their number until having one constant and a non-zero coefficient. If the model is linear again, return at Step 1, augment the order p until the maximum p estimated by information criteria (possibly augmented by Hinich' test) and restart the procedure.
- Step 4: *Grid search for starting values of nonlinear parameters*.
- Step 5: *Estimation*. Use the selected transition variables, transition function and starting value in order to compute the parameter estimates by NLLS algorithm. Impose the restriction $\phi = -\theta$ only when the unrestricted model is not able to produce reasonable estimates. In that case, perform a progressive criterion in imposing such restriction: start with restricting the constant; if the restriction produce reasonable estimates (γ and c are not high and p-values are lower than 0.05), continue with next step; otherwise, restrict y_{t-i} , $i = 1, \dots, p$ singularly and continue with next step; otherwise, restrict all possible combinations of y_{t-i} , until all lagged y_{t-i} are restricted and continue with next step; otherwise, add also the constant to the y_{t-i} previously restricted and continue with next step; otherwise, return at Step 2 and restart the procedure. If also the new procedure defaults, return to Step 1, augment the order p until the maximum p estimated by information criteria and restart the whole

procedure.

- Step 6: *Diagnostic tests*. Apply the three tests described in Section 2.4. If the resulting p-value are high, accept the selected model. Otherwise, check for different restrictions in Step 5 and accept estimates with slightly higher p-value; then, perform the diagnostic tests for the new (sub-optimal) model; if the resulting statistics are highly significant, accept the model. Otherwise, return to Step 3, set different transition variables and restart the procedure. If this is not an help, return to Step 2 and set new linear AR(p) specification and restart the procedure; if the estimates are significant, accept the model. Otherwise, as last possibility, return to Step 1 and augment the order p until the maximum in the information criteria (possibly augmented by Hinich' test) and restart the whole procedure. If the result is negative, the series is not mean-reverting, hence the (first) PPP puzzle cannot be solved.

4 The data

We consider 9 countries, namely: Denmark (DN), Canada (CAN), Japan (JPN), Norway (NW), Sweden (SD), Switzerland (SZ), U.K., U.S., and E.U (euro area); the numeraires are U.S. and E.U. Hence, we have 16 series for nominal exchange rates and 8 series for seasonally adjusted consumer price indicators⁸. The series of real exchange rates has been built by equation (2). In this article we label "DNUS", "CANUS", etc ("DNEUR", "CANEUR", etc.) the series of nominal exchange rates of various currencies vs. US dollar (vs. euro) Similarly, "PPPDNUS", "PPPCANUS", etc. ("PPPDNEUR", "PPPCANEUR", etc.) is the label of real exchange rates of national currencies vs. \$ (vs. €).

The considered sample is 1999:01 - 2009:12, so we have to hold with the presence of a structural break in 2008:04 due to the financial crisis and, in addition for the euro, with the "Greece Effect" in the last observations. The sources of these series are: FED of St. Louis for spot rates with basis \$, ECB for spot rates with basis € and OECD website for CPIs (see Section A.1 for series details).

⁸The original CPI series was not adjusted for seasonality. We did it by using the X-12 ARIMA procedure

5 Empirical evidence

Figures 1 and 2 plot the series of real exchange rates in the data set. It can be noticed that almost all spot rate follows a positively shaped linear broken trend, where the break corresponds to the beginning of the financial crisis (approximately in 2008:04). Moreover the linear trend is weaker and negatively shaped when E.U. is numeraire; this is the effect of the progressive appreciation of the € against the other currencies and the subsequent depreciation due to the speculative attack during the financial crisis. All the series show one or more breaks in the middle of the sample, corresponding to the selected dummy variables in Table 1. Two of them (2003:01 and 2003:03) seem to be consistent with the turbulence of oil market immediately before and after the Iraqi political crisis in that months. Figures 3 and 4 plot the real exchange rates in first differences, which allows for checking the presence of irregularities in mean reversion such as autocorrelation (clearly observable in $(\Delta\text{PPPDNUS}, \Delta\text{PPPEUUS}, \Delta\text{PPPDNEU}, \Delta\text{PPPSDUS}, \Delta\text{PPPSZUS})$) and ARCH-effects. We can see that all series show a break in mean reversion after 2008:05 approximately; moreover, in the first half of the sample the mean reversion seems to be weaker than in the second half. These irregularities are the source of all difficulties in finding good specification for linear and nonlinear models for all series.

5.1 CVAR

For what concerns the CVAR approach for weak PPP hypothesis, the procedure described in Section 3 is performed in a semi-automatic way by CATS package (Dennis et al., 2006). Concerning for Step 2, we use the Shenton-Bowman test for normality and the Ljung-Box LM test for autocorrelation. For simulation of critical values of Johansen' Trace test we apply the Johansen (2002) bootstrap procedure with 2,500 draws for each possible rank. For the choice of the rank, we consider both standard and Bartlett-corrected for small samples p -values; in a standard scenario, they are really similar. The plausible restricted models are selected by using the automatic procedure "CATS Mining", which show all potential cointegrating relation between covariates and select them as option. Clearly, since cointegrating relations are simply linear combinations, the number of candidates is often

so high that we have selected them by using the following criteria: first, p -value of the candidate should be at least 0.20 (this ensure some stability to the cointegrating relation which is essential for being economically meaningful, see Juselius (2006)); second, the sign of α and β in (3) should be at least similar to what theory suggests; finally, their absolute values should not be extravagant. The presence of I(2)-ness is checked by looking at: (i) the graphs of the cointegrating relations in their two specifications: if not strictly similar, this is a sign of I(2) behavior; (ii) the characteristic roots of the model for a reasonable choice of cointegration rank: if there's no difference between couples corresponding the candidate rank and ones immediately after (that is, they are all near to unit), there's I(2)-ness; (iii) rank test statistic p -values: considerable differences between Bartlett and non-Bartlett corrected p -values. Since the statistical theory of cointegration analysis in an I(2) scenario is not complete and does not necessarily add economically meaningful results to the empirical analysis, we stopped when I(2)-ness was found. Tab.1 illustrates our results for PPP when using \$ or € as numeraire. In the majority of cases our optimal lag choice is 2. However, since some outliers are present, we will specify the next test for $p = 3$. Almost all systems do not reject the null of no cointegration, hence, in practice, we stopped to Step 4. Two exceptions are constituted by Norway and U.K. However, since also the other ranks hypotheses has a small p -values respect on the other systems, the found relations for these countries are affected by I(2)-ness⁹. Moreover, it's interesting to note that the large number of shift dummies used corresponding to an equivalently large number of outliers in residuals implies that gaussianity assumption of the statistical model could be seriously suspect to not hold. This is not uncommon in financial variables.

The strong PPP hypothesis is tested by using three tests: the Augmented Dikey-Fuller (ADF) test, the GLS-robust Dikey-Fuller (DF-GLS) (Elliott et al., 1996) for the null of unit root in the real exchange rates and the KPSS test for the null of stationarity. The Marshallian approach allows us to select a predetermined number of additional lags (namely, $p = 0, \dots, 3$) in order to analyze the series. Clearly, a failure to reject the null hypothesis of unit root (a rejection of the null of stationarity) implies an irregularity in the real exchange rate mean reversion and so a lack of empirical support for strong PPP hypothesis.

⁹We do not report all the data which confirm this finding for space motivation. They can be provided under request

The results for the ADF, KPSS and DF-GLS tests for lags the first three lags are shown in Tables 2, 3 and 4. When \$ is numeraire we cannot reject the null of unit root for any country and any lag, while some exception is observable when € is numeraire but the result is the same; the only relevant rejection is the case of Norway when testing for lag 1. The hypothesis of stationarity is almost always rejected at 1%, regardless to the numeraire, coherently with the above results. The results for null unit root under GLS estimation confirm and, possibly, enforce the ADF ones.

5.2 Panel methods

Concerning panel methodology for strong PPP hypothesis, we have by 8 cross-sections. Since the results country-by country shows that it's reasonable to model until $p = 3$, we test for the first three lags, so that the number of observations varies between 1,024 and 1,040. The strong PPP hypothesis is investigated by performing the six tests previously described (see Section 2.3). For MW method we use both ADF and Phillips-Perron tests. Concerning Hadri test, the statistic is robust to heteroskedasticity and serial dependence across disturbances. Tables 5, 6 and 7 show that for both numeraires, panel unit root tests are not able to reject the null hypotheses of unit root with few exceptions and, coherently with this finding, reject the null of no unit root. Hence the data do not provide empirical evidence for strong PPP hypothesis.

The weak PPP hypothesis is investigated by performing the Pedroni and Westerlund tests on all possible triples of variables. Concerning Pedroni test, we show only the \tilde{Z}_ρ and $Z_{t_{\hat{\rho}_{NT}}}$ statistics on the seven proposed by the author, since they are the most powerful. For the same reason, concerning Westerlund tests, we show only the P_γ statistic. Both of the tests are based on the null hypothesis of no cointegrating relation, hence a failure to reject the null hypothesis implies the failure in finding empirical evidence for weak PPP hypothesis. Notice that Westerlund test is based on an error correction model; this implies that, similarly to the multivariate framework, when allowing for a deterministic kernel to enter in cointegrating relations, the statistic critical values are biased. Again, robust critical values can be computed by bootstrap methods. The whole analysis can be easily performed using *ad hoc* STATA procedures. Tab. 8 shows results for each triple, for which

has been provided the statistics, the corresponding z -value, standard p -value and, for the Westerlund test, bootstrapped p -value and the automatic selection of lags and leads. The two tests show that spot rates, domestic and foreign prices are strongly not cointegrated in two cases on three, regardless to the numeraire, and in the one where cointegration cannot be rejected the variables are positioned differently from what theory suggest. This finding leads us to reject the weak PPP hypotheses, on the contrary of Pedroni (2001).

5.3 Nonlinear models

For what concerns the nonlinear framework, the procedure described in Section 3 is performed by JMulTi free package and can be extended to SETAR models. However, for the last ones we used the analogue RATS procedures, in which case we provide the result for Tsay' test for different transition variables (that is for different delay parameters).

Before to implement the Granger-Teräsvirta procedure, we checked for the presence of ARCH-effects by performing the McLeod and Li (1983) test. The results are given in Tab. 9. We can see that the test fails to reject the null of no ARCH-effect for almost all the series, but if considering the lag corresponding to the p order of selected model, they became less problematic. Tab. 10 shows the results of the STAR specification procedure above explained for our dataset in levels. Data show some differences between numeraires: the detected nonlinear series are five when U.S. is the numeraire, one when E.U. is the numeraire, corresponding to the same national currency of one of the selected models in US numeraire case. This is the effect of the financial crisis: a more (nonlinearly) volatile behavior of real exchange rates towards euro which contrasts a good specification. It's interesting to notice that the errors of the considered series are not third-order correlated, as suggested by Hinich test: the only case of relevant correction is the series for real exchange rate between Canada and US (three lag added by procedure described in Sec. 3). The estimates of selected STAR models are reported in Appendix A.2. Fig. 5 shows the grid search for starting values used to arrive to the above estimates using negative SSR for contour plot, while fig. 6 the transition function $G(\gamma, c, s_t)$ as function of the transition variable s_t . These graphs shows a peculiarity in our dataset: on six estimated nonlinear models, four (PPPDNUS, PPPCANUS, PPPSDUS and PPPSZUS) are seen to be at the

line with linear models, since their mean reversion is very smooth. The other two models (PPPUKUS and PPPUKEUR) are more clearly nonlinear and, consequently, more interesting from an economic point of view because they correspond to very different situations: in the set of models under investigation, the model (27) is the more restricted one, since all lags of $y_t^{\mathcal{L}/e}$ enter in the restriction $\phi = -\theta$, hence the interpretation is very similar to that given in TPS, with the exception for the constant, which does not enter in the restriction. On the contrary, model (28) has no restrictions, so there's no equilibrium around which the model is a random walk, but more meaningfully a simple (quasi-exponential) mean reversion. Notice that this is the only non-LSTR1 model in the 6 estimates, namely a second-order logistic STAR. The only model estimated with restriction for all parameters is (25). These findings leads to several implications: first, the methodological choice to not use the restriction for all parameters jointly and the exponential smooth transition function as *a priori* was really critical. Second, real exchange rates have really an asymmetric behavior respect to their numeraires; this is an effect of financial instability in the last observations in our sample. Third, the nonlinear asymmetric mean reversion of exchange rates suggests a change in long run, if considering the TPS' s implicit observation that goods arbitrage are driving the market towards it. In this sense, the diagnostic tests in Tables 13, 14 and 15 do not support the idea of a third regime for the estimated model for any model; on the contrary, the rejected models are characterized by high parameter variability and serial correlation.

Tab. 11 shows the results for SETAR(k; p, d) specification procedures above explained for our dataset in levels. The Tsay 's test allows five currency to be nonlinearly mean-reverting. These results should not be taken as definitive, since a key role is put on the Hansen's test for threshold effect: if a series allows for a SETAR-type nonlinearity but the threshold effect is weak, the nonlinearity should be interpreted as spurious. This is exactly what happens to our data. We used all three statistics (12) for testing threshold effect in both homoskedasticity and heteroskedasticity cases. The p-values are bootstrapped using 1,000 draws. On six plausible one threshold SETAR models, the threshold effect holds in only one of them, namely in the real exchange rates between U.K. and E.U, as shown in Tab. 12. Moreover, this is a limit-case, since only the $supS_T$ -statistic is in rejection region, when the test is robust to heteroskedasticity. These are the resulting estimates where the

values in brackets are robust standard errors:

$$\left\{ \begin{array}{ll} y_t^{\mathcal{L}/e} = -0.022 + 0.849 \cdot y_{t-1}^{UK/EUR} & \text{if } y_{t-1}^{\mathcal{L}/e} \leq 0.341, \quad \hat{\sigma}^2 = 0.0007 \\ (0.013) \quad (0.055) & \\ y_t^{\mathcal{L}/e} = 0.006 + 0.984 \cdot y_{t-1}^{UK/EUR} & \text{if } y_{t-1}^{\mathcal{L}/e} > 0.341, \quad \hat{\sigma}^2 = 0.0002 \\ (0.011) \quad (0.026) & \end{array} \right.$$

An interesting feature is that the estimated SETAR model (as all plausible threshold models, too) are not so sensitive to heteroskedasticity, and the above model is the only exception. That is, the relevant ARCH-effects showed in Tab. 9 does not involve any relevant difference in (bootstrapped) p-values when performing an heteroskedasticity robust test for threshold effect respect on the non-robust one.

The above empirical analysis suggests that only a restricted number of series in our dataset are nonlinearly mean-reverting in a way that such mean reversion is sufficient in order to specify reasonable models; moreover, the majority of them show a behavior really near to the line with linear models. This suggests to us to not follow TSP pedantically in using the GRT method for GIRF analysis in order to study the shocks persistence of real exchange rates. On the contrary, by looking at the plots of transition functions we think that the TIRF are still a correct methodology for all models, except for (27) and (28) which presents the most nonlinear behavior. Nevertheless, neither in this case we agree in performing the GRT method, since our modeling strategy and the resulting models was different from TPS: their solution to PPP puzzle was based on ESTAR model, which has been shown in Section 2.4 to be a very peculiar case of the more general family of LSTR. The only model which does not follow an LSTR is an LSTR2, since $c_1 \neq c_2$; this means that the transition function is not perfectly symmetric (although it can be seen as an approximation), hence the economic interpretation of the resulting GIRF could be misleading¹⁰.

Fig. 7 plots the TIRF of the real exchange rate series for shocks of magnitude $\{1, 2, 3\}$ and an horizon of 12 month. We can see the almost regular behavior in the majority of cases, but also a couple of peculiarities: first, when the shock is larger (shock=3, green

¹⁰Moreover, the GRT method for GIRF is based on several strong assumption: parametric distribution; the mean as statistic measure of baseline forecasts, "[...] which under stationarity is the unconditional mean" (Koop et al. (1996, pag.130)); the GIRF is zero if the initial shock is zero. We note that the second assumption is the most problematic one because when shocks have asymmetric effect, "[...] then averaging across phases of the bussiness cycle will tend to weaken or hide the evidence of asymmetry" (*Ibidem*).

line) the TIRF does converge faster to 0 but tends immediately to go below such threshold; secondly, the series for £/\$ rate, which we have seen to be very nonlinear, seems to be the most "conventional" one for TIRF behavior in the long run; moreover its transition function has been seen to be logistic, which implies that the speed of adjustment should vary asymmetrically, and so a non regular behavior in TIRF. Instead, this happens in the series for £/€ rate, which was seen to have a quasi-symmetric transition function. However, a common feature of all the TIRFs is that small shocks tend to disappear slower than big shocks (6-8 months against 2-4) month. This confirms in some way the TPS result, showing the nonlinearity of the real exchange rate adjustment toward theoretical PPP equilibrium. It's interesting to note that TPS concerns for a larger scaled dataset¹¹ so that our result can be seen, nevertheless all the mentioned empirical problems, as a reasonable approximation to it.

6 Conclusions

In this article we studied the empirical support for the PPP theory after 1998 using different methodologies and two different scenarios, linear and nonlinear, in order to compare and update the empirical literature.

The general result is ambivalent: the analysis of a dataset of 16 real exchange rates does not support the PPP hypothesis for two of the three methodologies used. In particular, the CVAR analysis show the data are found to be strongly I(2). Panel methods for unit root and cointegration confirm the rejection of the theory, while the BMO's critique suggests to take such results very carefully. Things seem different in the nonlinear scenario: 6 rates on 16 are nonlinearly mean reverting, but the change in regime is located in correspondence of the crisis. This implies that the financial crisis in 2008 has been a source of nonlinear behavior which allowed to explain the movements for some rates better than using a linear framework, although this is not the case of €/\$, that it the most important case. Hence, despite the number of estimated models is higher than using other approaches (that is, weak PPP holds in more cases than TPS), we are careful to consider the PPP

¹¹TPS used a sample of 288 observations, an horizon of 200 observation and a set of shock going from 5 to 50

puzzles solved yet. In particular, two findings are critical: first, the qualitative analysis of the nonlinear part of STAR models shows a quasi-linear transition function. Second, data are not able to support the choice of an ESTAR-type of nonlinearity in favor of an LSTAR-type. This implies that, contrarily to TPS, the speed of adjustment of exchange rates varies according to the over(under)valuation of the numeraire. For these reasons, the solution of the second puzzle needs a methodological approach slightly different from the TPS one. However, TIRFs for estimated real exchange rates for an horizon of one year confirm nonlinear adjustment of the real exchange rates towards their theoretical PPP value and, in an approximate way, the TPS result. This is consistent with the economic intuition underlying the use of LSTAR as transition function: shocks in real exchange rates are not symmetric in period of crisis.

Finally, we suggest some lines for future research. Johansen et al.'s CVAR model under I(2) scenario seems the more advanced strategy today available, although it's not the most immediate to understand for practitioners. Other nonlinear cointegration techniques (nonlinear, nonparametric and a combination of both) are currently at first stage of development. Multivariate nonlinear models seems to ensure the correct mid-point in the trade-off between tractability and heaviness of structural assumptions (parametric framework). Camacho (2004) and González et al. (2005) generalize STR model to VAR and panel respectively; applying our dataset to such models could be interesting. A second strategy to adopt in this kind of models could be relaxing the assumption of symmetry in transition functions. We remind these extensions to further works.

References

- Bacon D, Watts D. 1971. Estimating the transition between two intersecting straight lines. *Biometrika* **58**: 525–534.
- Banerjee A, Marcellino M, Osbat C. 2005. Testing for PPP: Should we use panel methods? *Empirical Economics* **30**: 77–91. DOI: 10.1007/s00181-004-0222-8.
- Camacho M. 2004. Vector Smooth Transition Regression models for US GDP and the composite index of leading vectors. *Journal of Forecasting* **23**: 173–196. DOI: 10.1002/for.912.
- Chan K, Tong H. 1986. On estimating thresholds in autoregressive models. *Journal of Time Series Analysis* **7**: 178–190. DOI: 10.1111/j.1467-9892.1986.tb00501.x.
- Coleman A. 1995. Arbitrage, Storage and the Law of One Price: New Theory for the Time Series Analysis of an Old Problem. Discussion Paper. Department of Economics. Princeton University.
- Dennis J, Hansen H, Johansen S, Juselius K. 2006. CATS in RATS. Cointegration Analysis of Time Series, Version 2. Estima: Evanston.
- Eitrheim O, Teräsvirta T. 1996. Testing the adequacy of smooth transition autoregressive models. *Journal of Econometrics* **74**: 59–75. DOI: 10.1016/0304-4076(95)01751-8.
- Elliott G, Rothemberg T, Stock J. 1996. Efficient Tests for an Autoregressive Unit Root. *Econometrica* **64**: 813–836.
- Gallant A, Rossi P, Tauchen G. 1993. Nonlinear Dynamic Structures. *Econometrica* **61**: 871–907.
- González A, Teräsvirta T, van Dijk D. 2005. Panel Smooth Transition Regression Models. Quantitative Finance Research Centre. Research Paper.
- Granger C, Teräsvirta T. 1993. *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

- Hadri K. 2000. Testing for stationarity in heterogeneous panel data. *Econometrics Journal* **3**: 148–161. DOI: 10.1111/1368-423X.00043.
- Hansen B. 1996. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* **64**: 413–30.
- Hinich M. 1996. Testing for dependence in the input to a linear time series model. *Journal of Nonparametric Statistics* **6**: 205–221. DOI: 10.1080/10485259608832672.
- Im K, Pesaran M, Shin Y. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* **115**: 53–74.
- Johansen S. 1991. Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vectors Autoregressive Models. *Econometrica*. **59**: 1551–1580.
- Johansen S. 2002. A small sample correction for tests of hypotheses on the cointegrated vectors. *Journal of Econometrics* **111**: 195–221.
- Johansen S, Juselius K, Frydman R, Goldberg M. 2010. Testing hypotheses in an I(2) model with piecewise linear trends. An analysis of the persistent long swings in the Dmk/\$ rate. *Journal of Econometrics* **158**: 117–129. DOI: 10.1016/j.jeconom.2010.03.018.
- Juselius K. 2006. *The Cointegrated VAR Model: Methodology and Applications*. Oxford: Oxford University Press.
- Kao C. 1999. Spurious regressions and residual-based tests for cointegration in panel data. *Journal of Econometrics* **90**: 1–44.
- Koop G, Pesaran M, Potter S. 1996. Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics* **74**: 119–147. DOI: 10.1016/0304-4076(95)01753-4.
- Kwiatkowski D, Phillips P, Shmidt P, Shin Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* **54**: 159–178. DOI: 10.1016/0304-4076(92)90104-Y.
- Levin A, Lin C, Chu C. 2002. Unit root test in panel data: Asymptotic and finite sample properties. *Journal of Econometrics* **87**: 207–237. DOI: 10.1016/S0304-4076(01)00098-7.

- Luukkonen R, Saikkonen P, Teräsvirta T. 1988. Testing linearity against smooth transition autoregressive models. *Biometrika* **75**: 491–499.
- Maddala G, Wu S. 1999. A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* **61**: 631–652. DOI: 10.1111/1468-0084.0610s1631.
- McCoskey S, Kao C. 1998. A residual-based test of the null of cointegration in panel data techniques. *Econometric Reviews* **17**: 57–84. DOI: 10.1080/07474939808800403.
- McLeod A, Li W. 1983. Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis* **4**: 269–273. DOI: 10.1111/j.1467-9892.1983.tb00373.x.
- Nyblom J, Harvey A. 2000. Tests of Common Stochastic Trends. *Econometric Theory* **16**: 176–199.
- Pedroni P. 2001. Purchasing power parity tests in cointegrated panels. *The Review of Economics and Statistics* **83**: 727–731. doi:10.1162/003465301753237803.
- Pedroni P. 2004. Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* **20**: 597–625. DOI: 10.1017/S0266466604203073.
- Pesaran M. 2007. A Simple Panel Unit Root Test for Cross-Section Dependence. *Journal of Applied Econometrics* **22**: 265–312. DOI: 10.1002/jae.951.
- Rogoff K. 1996. The Purchasing Power Parity Puzzle. *Journal of Economic Literature* **34**: 647–668.
- Sarno L, Taylor M. 2001. Purchasing Power Parity and the Real Exchange Rates. CEPR Discussion Papers 2913.
- Taylor M, Peel D, Sarno L. 2001. Nonlinear Mean Reversion in Real Exchange Rates: Towards a Solution to the Purchasing Power Parity Puzzles. *International Economic Review* : 1015–1042 DOI: 10.1111/1468-2354.0014.

- Teräsvirta T. 1994. Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* **89**: 208–218.
- Tong H. 1983. *Threshold Models in Non-Linear Time Series Analysis*. Number 21 in Lecture Notes in Statistics. New York: Springer-Verlag.
- Tsay R. 1989. Testing and modeling threshold autoregressive processes. *Journal of the American Statistical Association* **84**: 231–240.
- Westerlund J. 2007. Testing for Error Correction in Panel Data. *Oxford Bulletin of Economics and Statistics* **69**: 709–748. DOI: 10.1111/j.1468-0084.2007.00477.x.

A Appendix

A.1 Data

Our original dataset is constituted of monthly series of spot rates (currency basis United States Dollar and Euro) and consumers' price indices. The data sample goes from 1999:01 to 2009:12 (132 observation).

Spot rate series with basis USD has been downloaded largely from FED of St. Louis and these are the sources:

Canada: EXCAUS, Board of Governors of Federal Reserve System;
 Denmark: EXDNUS, Board of Governors of Federal Reserve System;
 Japan: EXJPUS, Board of Governors of Federal Reserve System;
 Norway: EXNOUS, Board of Governors of Federal Reserve System;
 Sweden: EXSDUS, Board of Governors of Federal Reserve System;
 Switzerland: EXSZUS, Board of Governors of Federal Reserve System;
 U.K.: United Kingdom, Exchange Rates, OECD;
 EU: EU-12-Extra EU, Exchange Rates, OECD.

Spot rate time series with basis EUR are downloaded from European Central Bank.

Dataset description: Dataset name: Exchange Rates ; Frequency: Monthly ; Currency denominator: Euro ; Exchange rate type: Spot ; Series variation - EXR context: Average or standardized measure for given frequency.

These are the sources for Country:

Canadian dollar: EXR.M.CAD.EUR.SP00.A;
 Danish krone: EXR.M.DKK.EUR.SP00.A;
 Japanese yen: EXR.M.JPY.EUR.SP00.A;
 Norwegian krone: EXR.M.NOK.EUR.SP00.A;
 Swedish krona: EXR.M.SEK.EUR.SP00.A;
 Swiss franc: EXR.M.CHF.EUR.SP00.A;
 U.K. pound sterling: EXR.M.GBP.EUR.SP00.A;
 U.S. dollar: EXR.M.USD.EUR.SP00.A.

CPI series has been downloaded from OECD. These are the name of the series for Country:

Canada: CAN CPI - All items - Index publication base - units: 2005=100;

Denmark: DNK CPI - All items - Index publication base - units: 2005=100;
Japan: JPN CPI - All items Tokyo - Index publication base - units: 2005=100;
Norway: NOR CPI - All items - Index publication base - units: 2005=100;
Sweden: SWE CPI - All items net - Index publication base - units: 2005=100;
Switzerland: CHE CPI - All items - Index publication base - units: 2005=100;
U.K.: GBR CPI - All items - Index publication base - units: 2005=100;
U.S: USA CPI - All items SA - Index publication base - units: 2005=100;
E.U.: EMU CPI HICP - All items - Index publication base - units: 2005=100.

All CPI series (except USA CPI which is seasonally adjusted) has been de-seasonalised by X-12 ARIMA procedure. Then, these preliminary data has been transformed in logarithms, from which PPP (or real exchange rates) series has been built.

A.2 STAR model estimates

$$y_t^{DN/US} = -3.978 - 0.639 \cdot y_{t-2}^{DN/US} + (2.679 + 0.639 \cdot y_{t-2}^{DN/US}) \cdot \{1 - \exp[0.567 \cdot (y_{t-1}^{DN/US} - (-1.997))]\}^{-1} \quad (23)$$

(0.402)	(0.187)	(0.300)	(0.187)	(0.000)	(0.026)
[0.000]	[0.001]	[0.000]	[0.001]	[0.000]	[0.000]

AR - part : constant, y_{t-2} *T* = 130 *Niter* = 1 $\phi = -\theta : y_{t-2}^{DN/US}$ $R^2 = 0.9780$ $\bar{R}^2 = 0.9820$
 $\hat{\sigma}_{s_t}^2 = 0.0235$ $SD_{s_t} = 0.1533$ $\hat{\sigma}_{\epsilon_t}^2 = 0.0005$ $SD_{\epsilon_t} = 0.0233$
Sk = 0.1918 *Ek* = 3.1518 *JB* = 0.9218(0.6307) *ARCH - LM(p)* = 5.7114(0.0042)

$$y_t^{CAN/US} = -1.158 - 0.436 \cdot y_{t-2}^{CAN/US} + (1.442 + 0.436 \cdot y_{t-2}^{CAN/US}) \cdot \{1 - \exp[0.507 \cdot (y_{t-1}^{CAN/US} - (-0.360))]\}^{-1} \quad (24)$$

(0.000)	(0.174)	(0.000)	(0.174)	(0.000)	(0.030)
[0.000]	[0.014]	[0.000]	[0.014]	[0.000]	[0.000]

AR - part : constant, y_{t-2} *T* = 128 *Niter* = 1 $\phi = -\theta : y_{t-2}^{CAN/US}$ $R^2 = 0.9797$ $\bar{R}^2 = 0.9799$
 $\hat{\sigma}_{s_t}^2 = 0.0181$ $SD_{s_t} = 0.1344$ $\hat{\sigma}_{\epsilon_t}^2 = 0.0004$ $SD_{\epsilon_t} = 0.0195$
Sk = -1.1166 *Ek* = 11.0359 *JB* = 371.0014(0.0000) *ARCH - LM(p)* = 0.1947(0.9407)

$$y_t^{SD/US} = -11.794 - 1.525 \cdot y_{t-2}^{SD/US} + (11.794 + 1.525 \cdot y_{t-2}^{SD/US}) \cdot \{1 - \exp[0.105 \cdot (y_{t-1}^{SD/US} - (-3.435))]\}^{-1} \quad (25)$$

(3.195)	(0.400)	(3.195)	(0.400)	(0.020)	(0.751)
---------	---------	---------	---------	---------	---------

[0.000] [0.000] [0.000] [0.000] [0.000] [0.000]

AR - part : constant, y_{t-2} $T = 128$ $Niter = 75$ $\phi = -\theta : constant, y_{t-2}^{SD/US}$ $R^2 = 0.9568$ $\bar{R}^2 = 0.9572$
 $\hat{\sigma}_{s_t}^2 = 0.0154$ $SD_{s_t} = 0.1240$ $\hat{\sigma}_{\epsilon_t}^2 = 0.0007$ $SD_{\epsilon_t} = 0.0261$
 $Sk = -0.0234$ $Ek = 2.8335$ $JB = 0.1596(0.9233)$ $ARCH - LM(p) = 0.1671(0.9547)$

$$y_t^{SZ/US} = -1.127 - 0.656 \cdot y_{t-2}^{SZ/US} + (1.220 + 0.656 \cdot y_{t-2}^{SZ/US}) \cdot \{1 - \exp[0.615 \cdot (y_{t-1}^{SZ/US} - (-0.388))]\}^{-1} \quad (26)$$

(0.272) (0.303) (0.330) (0.303) (0.162) (0.052)

[0.000] [0.032] [0.000] [0.032] [0.000] [0.000]

AR - part : constant, y_{t-2} $T = 130$ $Niter = 5$ $\phi = -\theta : y_{t-2}^{SZ/US}$ $R^2 = 0.9595$ $\bar{R}^2 = 0.9598$
 $\hat{\sigma}_{s_t}^2 = 0.0129$ $SD_{s_t} = 0.1134$ $\hat{\sigma}_{\epsilon_t}^2 = 0.0005$ $SD_{\epsilon_t} = 0.0234$
 $Sk = 0.4484$ $Ek = 3.1078$ $JB = 4.4190(0.1098)$ $ARCH - LM(p) = 1.0784(0.3433)$

$$y_t^{\pounds/\$} = 0.011 + 1.327y_{t-1}^{\pounds/\$} - 0.350y_{t-2}^{\pounds/\$} + (0.490 - 1.327y_{t-1}^{\pounds/\$} + 0.350y_{t-2}^{\pounds/\$}) \cdot \{1 - \exp[53.692 \cdot (t - (123.025))]\}^{-1} \quad (27)$$

(0.011) (0.085) (0.086) (0.013) (0.085) (0.086) (30.565) (0.565)

[0.308] [0.000] [0.000] [0.000] [0.000] [0.000] [0.081] [0.000]

AR - part : constant, y_{t-1}, y_{t-2} $T = 130$ $Niter = 12$ $\phi = -\theta : y_{t-1}^{\pounds/\$}, y_{t-2}^{UK/US}$ $R^2 = 0.9585$ $\bar{R}^2 = 0.9589$
 $\hat{\sigma}_{s_t}^2 = 1,419.1667$ $SD_{s_t} = 37.6718$ $\hat{\sigma}_{\epsilon_t}^2 = 0.0004$ $SD_{\epsilon_t} = 0.0199$
 $Sk = -0.3798$ $Ek = 3.858$ $JB = 7.1131(0.0285)$ $ARCH - LM(p) = 3.2050(0.0439)$

$$y_t^{\pounds/e} = -0.019 + 1.287y_{t-1}^{\pounds/e} - 0.245y_{t-2}^{\pounds/e} + (0.069 - 1.4242y_{t-1}^{\pounds/e} + 1.299y_{t-2}^{\pounds/e}) \cdot \{1 - \exp[6.585(y_{t-2}^{\pounds/e} - 0.141)(y_{t-2}^{\pounds/e} - 0.541)]\}^{-1} \quad (28)$$

(0.007) (0.090) (0.095) (0.016) (0.395) (0.363) (4.357) (0.004) (0.006)

[0.011] [0.000] [0.011] [0.000] [0.000] [0.001] [0.133] [0.000] [0.000]

$$\begin{aligned}
AR - part : constant, y_{t-1}, y_{t-2} \quad T = 130 \quad Niter = 13 \quad \phi = -\theta : - \quad R^2 = 0.9830 \quad \bar{R}^2 = 0.9831 \\
\hat{\sigma}_{s_t}^2 = 0.0121 \quad SD_{s_t} = 0.1079 \quad \hat{\sigma}_{\epsilon_t}^2 = 0.0002 \quad SD_{\epsilon_t} = 0.0154 \\
Sk = -0.9135 \quad Ek = 5.3848 \quad JB = 48.8857(0.0000) \quad ARCH - LM(p) = 10.1769(0.0001)
\end{aligned}$$

where: $y_t \equiv v_t \equiv ppp_t$ as in (2), the values in brackets in the estimated model are standard deviations and the one in square brackets is the p-values, $Niter$ the number of iterations needed for the NLLS algorithm, $\phi = -\theta$ is the restriction in the STAR model (specifying the particular restriction we opted for), \bar{R}^2 the adjusted R^2 , $\hat{\sigma}_{s_t}$, $\hat{\sigma}_{\epsilon_t}^2$, SD_{s_t} and SD_{ϵ_t} the estimated variance and standard deviation of transition variable and sample errors respectively, Sk is the skewness, Ek the excess of kurtosis and JB the Jarque-Brera statistics (with p-value in brackets), $ARCH - LM(p)$ is the Engle test for the estimated model (in levels) for p lags.

A.3 Tables and Figures

Table 1: Cointegration Analysis for System of Country j

(1) j	(2) SC	(3) HQ	(4) p	(5) Shift Dummies	(6) Normality (p-value)	(7) Autocorr. (LM 1)	(8) Autocorr. (LM 2)	(9) Rank Test (p-value)	(10) Rank Test (Bartlett-corr.)	(11) Sim. Rank Test (p-value)	(12) Sim. Rank Test (Bartlett-corr.)
U.S. numeraire											
CAN	2	2	3	2005:08	0.000	0.320	0.865	0.107	0.164	0.265	0.361
DN	2	2	3	2005:09, 2007:11	0.001	0.242	0.116	0.231	0.325	0.228	0.329
JPN	1	2	1	2005:09, 2008:11	0.128	0.102	0.174	0.369	0.451	0.079	0.121
NW	2	2	3	2003:01, 2005:09, 2008:10	0.001	0.174	0.770	0.000	0.000	0.003	0.008
SD	1	2	1	-	0.000	0.151	0.120	0.012	0.030		
SZ	1	2	2	2003:03, 2005:09, 2008:10, 2008:11	0.352	0.083	0.064	0.000	0.000	0.120	0.173
UK	2	2	1	2005:09, 2008:11	0.171	0.535	0.216	0.076	0.117	0.004	0.009
EU	2	2	2	2008:11	0.003	0.056	0.152	0.142	0.197	0.080	0.119
US	-	-	-	-	-	-	-	-	-	-	-
E.U. numeraire											
CAN	1	1	3	-	0.036	0.109	0.556	0.019	0.022	0.095	0.104
DN	1	1	2	2007:11	0.006	0.977	0.617	0.247	0.301	0.315	0.377
JPN	1	1	1	2008:10	0.067	0.240	0.250	0.000	0.000	0.000	0.000
NW	2	1	3	2003:01, 2008:12	0.018	0.697	0.496	0.000	0.000	0.000	0.000
SD	1	1	1	2008:12, 2009:08	0.000	0.056	0.634	1.000	1.000	0.398	0.418
SZ	1	1	2	2008:10	0.002	0.476	0.188	0.024	0.037	0.013	0.020
UK	2	2	2	2008:12	0.008	0.451	0.836	0.000	0.000	0.000	0.000
EU	-	-	-	-	-	-	-	-	-	-	-
US	1	2	2	2005:09, 2008:12	0.035	0.088	0.271	0.080	0.110	0.060	0.093

Legend: Column (1): Countries for which the PPP is tested; columns (2)-(3): results of Shwartz and Hannan and Quinn Information criteria for lag selection; column (4): selected lag; column (5) shift dummies introduced in order to take in account of shocks in the sample; column (6): Shenton-Bowman test for normality in p -values; columns (7)-(8): Ljung-Box tests for first-order and second-order residual autocorrelation; column (9) Johansen' Trace test statistic in p -value; column (10): Bartlett nonparametric corrected Trace test; column (11): simulated Johansen' Trace test; column (12): Bartlett-corrected simulated Trace test.

Simulation technique: Bootstrap;

no. of draws: 2500

Software used: CATS

Table 2: Univariate ADF Test on Real Exchange Rates

Lag	US numeraire				EU numeraire			
	0	1	2	3	0	1	2	3
CAN	-1.776	-2.320	-2.300	-2.411	-3.283 [•]	-2.990	-3.037	-2.743
DN	-2.900	-2.940	-2.687	-2.579	-2.174	-1.468	-1.412	-1.858
JPN	-1.314	-1.873	-2.085	-1.731	-1.620	-2.265	-2.302	-2.318
NW	-2.358	-2.911	-2.742	-2.828	-2.825	-3.480 [*]	-2.951	-2.951
SD	-2.058	-2.531	-2.189	2.420	-2.323	-2.455	-2.455	-3.269 [•]
SZ	-3.067	-3.106	-2.887	-2.650	-1.635	-1.656	-1.601	1.798
UK	-1.433	-1.982	-2.039	-2.091	-2.458	-2.580	-2.235	-2.239
EU	-3.001	-2.919	-2.701	-2.613	-	-	-	-
US	-	-	-	-	-2.990	-2.958	-2.709	-2.679

• Rejection at 10% of the null hypothesis; •• rejection at 5% of the null hypothesis; ••• rejection at 1% of the null hypothesis.

Table 3: Univariate KPSS Test on Real Exchange Rates

Lag	US numeraire				EU numeraire			
	0	1	2	3	0	1	2	3
CAN	1.060 ^{**}	0.541 ^{**}	0.369 ^{**}	0.284 ^{**}	0.591 ^{**}	0.311 ^{**}	0.216 ^{**}	0.168 [*]
DN	0.799 ^{**}	0.414 ^{**}	0.285 ^{**}	0.221 ^{**}	1.310 ^{**}	0.675 ^{**}	0.460 ^{**}	0.352 ^{**}
JPN	1.190 ^{**}	0.612 ^{**}	0.421 ^{**}	0.326 ^{**}	1.080 ^{**}	0.559 ^{**}	0.384 ^{**}	0.297 ^{**}
NW	0.694 ^{**}	0.359 ^{**}	0.248 ^{**}	0.192 [*]	0.885 ^{**}	0.463 ^{**}	0.324 ^{**}	0.254 ^{**}
SD	0.837 ^{**}	0.430 ^{**}	0.294 ^{**}	0.226 ^{**}	0.755 ^{**}	0.394 ^{**}	0.273 ^{**}	0.212 [*]
SZ	0.686 ^{**}	0.359 ^{**}	0.349 ^{**}	0.194 [*]	1.200 ^{**}	0.620 ^{**}	0.423 ^{**}	0.324 ^{**}
UK	1.150 ^{**}	0.586 ^{**}	0.399 ^{**}	0.306 ^{**}	1.110 ^{**}	0.581 ^{**}	0.402 ^{**}	0.312 ^{**}
EU	0.843 ^{**}	0.437 ^{**}	0.301 ^{**}	0.233 ^{**}	-	-	-	-
US	-	-	-	-	0.860 ^{**}	0.446 ^{**}	0.307 ^{**}	0.238 ^{**}

• Rejection at 10% of the null hypothesis; •• rejection at 5% of the null hypothesis; ••• rejection at 1% of the null hypothesis

Table 4: Univariate DF-GLS Test on Real Exchange Rates (ERS critical values)

Lag	US numeraire				EU numeraire			
	0	1	2	3	0	1	2	3
CAN		-2.108	-2.143	-2.143		-1.485	-1.514	-1.504
DN		-1.495	-1.296	-1.308		-0.844	-0.959	-1.232
JPN		-1.792	-2.047	-1.732		-1.770	-1.779	-1.837
NW		-2.248	-2.019	-2.067		-2.134	-1.819	-1.950
SD		-1.831	-1.617	1.898		-1.925	-1.842	-2.560
SZ		-1.495	-1.381	-1.291		-1.418	-1.377	1.481
UK		-1.807	-1.828	-1.922		-1.299	-1.156	-1.178
EU		-1.412	-1.243	-1.272		-	-	-
US		-	-	-		-1.475	-1.293	-1.330

• Rejection at 10% of the null hypothesis; •• rejection at 5% of the null hypothesis; ••• rejection at 1% of the null hypothesis.

Table 5: Panel unit root test for Real Exchange Rates (lag=1)

Method	US numeraire		EU numeraire		Obs
	Statistics	<i>p</i> -value	Statistics	<i>p</i> -value	
<i>Null: unit root (common unit root is assumed)</i>					
LLC t^*	-1.697	0.045	-1.540	0.062	1040
<i>Null: unit root (individual unit root is assumed)</i>					
IPS w-statistic	-1.683	0.046	-1.290	0.099	1040
CADF $Z(\bar{t})$	-0.818	0.207	0.472	0.682	1040
ADF Fisher χ^2	21.543	0.159	21.272	0.168	1040
PP Fisher χ^2	18.615	0.289	20.993	0.179	1040
<i>Null: no unit root (common unit root is assumed)</i>					
Hadri Z-statistic (assuming heterosk. across disturbances)	55.907	0.000	60.480	0.000	
Hadri Z-statistic (assuming serial dependance across disturbances)	5.956	0.000	6.653	0.000	
NH^* (i.i.d. RW errors)	5.210 $\bullet\bullet\bullet$	-	5.192 $\bullet\bullet\bullet$		
NH^{**} (nonparametric adjustment of LRV, lag=1)	2.760 $\bullet\bullet\bullet$	-	2.730 $\bullet\bullet\bullet$		

• Rejection at 10% of the null hypothesis; •• rejection at 5% of the null hypothesis; ••• rejection at 1% of the null hypothesis; * No lag specified for LRV; ** With lag 1, 2 or 3 for LRV; Software used: STATA 10.

Table 6: Panel unit root test for Real Exchange Rates (lag=2)

Method	US numeraire		EU numeraire		Obs
	Statistics	<i>p</i> -value	Statistics	<i>p</i> -value	
<i>Null: unit root (common unit root is assumed)</i>					
LLC t^*	-1.519	0.064	-1.233	0.109	1032
<i>Null: unit root (individual unit root is assumed)</i>					
IPS w-statistic	-1.761	0.039	-1.247	0.106	1032
CADF $Z(\bar{t})$	-1.132	0.129	0.347	0.636	1032
ADF Fisher χ^2	17.691	0.342	17.026	0.394	1032
PP Fisher χ^2	19.825	0.228	21.870	0.147	1032
<i>Null: no unit root (common unit root is assumed)</i>					
Hadri Z-statistic (assuming heterosk. across disturbances)	55.907	0.000	60.480	0.000	
Hadri Z-statistic (assuming serial dependance across disturbances)	5.956	0.000	6.653	0.000	
NH^* (i.i.d. RW errors)	5.208 $\bullet\bullet\bullet$	-	5.192 $\bullet\bullet\bullet$		
NH^{**} (nonparametric adjustment of LRV, lag=2)	1.924 $\bullet\bullet\bullet$	-	1.898 $\bullet\bullet\bullet$		

• Rejection at 10% of the null hypothesis; •• rejection at 5% of the null hypothesis; ••• rejection at 1% of the null hypothesis; * No lag specified for LRV; ** With lag 1, 2 or 3 for LRV; Software used: STATA 10.

Table 7: Panel unit root test for Real Exchange Rates (lag=3)

Method	US numeraire		EU numeraire		Obs
	Statistics	p-value	Statistics	p-value	
<i>Null: unit root (common unit root is assumed)</i>					
LLC t^*	-1.005	0.157	-0.826	0.2043	1024
<i>Null: unit root (individual unit root is assumed)</i>					
IPS w-statistic	-1.538	0.062	-1.074	0.141	1024
CADF Z(t-bar)	-1.308	0.096	0.319	0.375	1024
ADF Fisher χ^2	16.899	0.392	19.690	0.235	1024
PP Fisher χ^2	20.446	0.201	22.824	0.112	1024
<i>Null: no unit root (common unit root is assumed)</i>					
Hadri Z-statistic (assuming heterosk. across disturbances)	55.907	0.000	60.480	0.000	
Hadri Z-statistic (assuming serial dependance across disturbances)	5.956	0.000	6.653	0.000	
NH^* (i.i.d. RW errors)	5.210 $\bullet\bullet\bullet$	-	5.192 $\bullet\bullet\bullet$		
NH^{**} (nonparametric adjustment of LRV, lag=3)	1.499 $\bullet\bullet\bullet$		1.477 $\bullet\bullet\bullet$		

\bullet Rejection at 10% of the null hypothesis; $\bullet\bullet$ rejection at 5% of the null hypothesis; $\bullet\bullet\bullet$ rejection at 1% of the null hypothesis; * No lag specified for LRV; ** With lag 1, 2 or 3 for LRV; Software used: STATA 10.

Table 8: Pedroni and Westerlund test on panel cointegration*

Case ($y_t \sim I(0)$)	Pedroni tests				Westerlund test				
	\tilde{Z}_ρ	p-value	$Z_{t_{\hat{\rho}_{NT}}}$	p-value	P_γ	Z-value	p-value	p-value (Robust)	Lag (AIC)
$(p - \alpha s^{US} - \beta p^{US})$	0.888	0.375	0.485	0.628	-5.383	2.136	0.984	0.952	1
$(s^{US} - \alpha p - \beta p^{US})$	-2.517	0.012	-2.977	0.003	-11.012	-0.219	0.413	0.379	1
$(p^{US} - \alpha s^{US} - \beta p)$	0.598	0.550	0.190	0.849	-9.920	-0.237	0.594	0.542	1
$(p - \alpha s^{EU} - \beta p^{EU})$	0.942	0.346	0.540	0.589	-	-	-	-	-
$(s^{EU} - \alpha p - \beta p^{EU})$	-11.355	0.000	-11.963	0.000	-11.736	-0.522	0.301	0.268	1
$(p^{EU} - \alpha s^{EU} - \beta p)$	-0.190	0.849	0.437	0.662	-	-	-	-	-

* Common Features: H_0 : no cointegration; deterministic term: constant + linear trend. Westerlund test features: lag range: (0 - 3); lead range: (0 - 1); width of Bartlett's Kernel window: 3; bootstrap n. of replications: 1000; Software used: RATS (Pedroni test), STATA 10 (Westerlund).

Table 9: McLeod-Li test for no ARCH-effects in ΔPPP (p-value)

Lag	US numeraire				EU numeraire			
	1	2	3	4	1	2	3	4
DN	0.442	0.013	0.034	0.049	0.753	0.839	0.894	0.502
CAN	0.867	0.732	0.847	0.937	0.807	0.969	0.592	0.489
JPN	0.152	0.082	0.171	0.285	0.000	0.000	0.000	0.000
NW	0.000	0.000	0.001	0.002	0.017	0.000	0.000	0.000
SD	0.009	0.032	0.048	0.092	0.097	0.016	0.013	0.014
SZ	0.115	0.247	0.320	0.471	0.380	0.427	0.023	0.047
UK	0.000	0.000	0.000	0.000	0.011	0.015	0.022	0.039
US	-	-	-	-	0.048	0.012	0.031	0.056
EU	0.175	0.004	0.010	0.012	-	-	-	-

Table 10: Linearity testing and model selection: STAR (levels)

(1) Series	(2) BIC	(3) AIC	(4) HQ	(5) Hinich test		(6) p selected	(7) s_t	(8)	(9)	(10)	(11)	(12) Model
				H-statistic*	C-statistic*			F_L F-value	F_4 F-value	F_3 F-value	F_2 F-value	
<i>US numeraire</i>												
DN	2	2	2	0.000	0.000	2	y_{t-1}	4.6170 exp ⁻²⁷	7.1771 exp ⁻¹	9.1470 exp ⁻²	9.1223 exp ⁻³⁰	LSTR1
CAN	1	1	1	0.999	0.000	4	y_{t-1}	1.8133 exp ⁻²⁴	9.1771 exp ⁻¹	6.8183 exp ⁻¹	4.7939 exp ⁻²⁸	LSTR1
JPN	1	2	2	0.000	0.000	1	-	-	-	-	-	Linear
NW	2	2	2	0.000	0.000	2	-	-	-	-	-	Linear
SD	2	4	2	0.050	0.000	4	y_{t-1}	1.4752 exp ⁻²⁸	1.6756 exp ⁻¹	5.6180 exp ⁻¹	1.7128 exp ⁻³¹	LSTR1
SZ	1	2	2	0.000	0.000	2	y_{t-1}	2.0722 exp ⁻²³	3.7734 exp ⁻¹	6.0788 exp ⁻²	1.8551 exp ⁻²⁵	LSTR1
UK	2	2	2	0.000	0.000	2	t	1.0997 exp ⁻²	5.8995 exp ⁻²	6.9881 exp ⁻²	7.5855 exp ⁻²	LSTR1
EU	2	2	2	0.000	0.000	2	-	-	-	-	-	Linear
<i>EU numeraire</i>												
DN	2	5	4	0.085	0.000	3	-	-	-	-	-	Linear
CAN	1	1	1	0.000	0.000	1	-	-	-	-	-	Linear
JPN	2	2	2	0.000	0.000	2	-	-	-	-	-	Linear
NW	2	3	2	0.000	0.000	2	-	-	-	-	-	Linear
SD	1	1	1	0.000	0.000	1	-	-	-	-	-	Linear
SZ	1	1	1	0.000	0.000	1	-	-	-	-	-	Linear
UK	1	1	1	0.000	0.000	1	y_{t-2}	2.5356 exp ⁻⁴	1.4176 exp ⁻¹	3.1568 exp ⁻⁵	5.9932 exp ⁻¹	LSTR2
US	1	1	2	0.000	0.000	1	-	-	-	-	-	Linear

Legend. (1): Countries; (2): Swartz-Bayesian information criterion; (3): Aikake information criterion; (4): Hannan-Quinn information criterion; (5)-(6) Hinich test statistics (p -values) for third order autocorrelation and serial dependence from non gaussian errors; (7): selected order p using the following rule: $p = k + n$, $k = \max\{AIC, BIC, HQ\}$ is the maximum number selected by standard information criteria, $n = \{1, 2, 3\}$ are the eventually additional order suggested by Hinich test; (8): transition variable; (8): Saikkonen-Lukkonen-Teräsvirta linearity test by statistics LM_2 (equation 15 on page 9) or LM_3 (equation 19 on page 10); (9)-(11): results for Teräsvirta rule for the choice of the form of transition variable (see sequence of nested hypotheses 20 on page 10); (12) Selected model for transition function.

Software used: JMulTi 4

Table 11: Linearity testing and model selection: SETAR (levels)

(1) Series	(2) BIC	(3) AIC	(4) HQ	(5) Hinich test		(6) Selected p	(7)	(8)	(9)	(10) d	(11) Model
				H-statistic	C-statistic		Tsay test d=1	Tsay test d=2	Tsay test d=3		
<i>US numeraire</i>											
DN	2	2	2	0.000	0.000	2	0.129	0.098	0.732		
CAN	1	1	1	0.999	0.000	4	0.612	0.703	0.347		
JPN	1	2	2	0.000	0.000	1	0.964	0.456	0.028	3	SETAR(1; 1, 3)
NW	2	2	2	0.000	0.000	2	0.475	0.097	0.406		
SD	2	4	2	0.050	0.000	2	0.388	0.227	0.119		
SZ	1	2	2	0.000	0.000	3	0.281	0.359	0.568		
UK	2	2	2	0.000	0.000	2	0.204	0.273	0.731		
EU	2	2	2	0.000	0.000	2	0.059	0.073	0.0564		
<i>EU numeraire</i>											
DN	2	5	4	0.085	0.000	3	0.034	0.238	0.376	1	SETAR(1; 3, 1)
CAN	1	1	1	0.000	0.000	1	0.095	0.846	0.920		
JPN	2	2	2	0.000	0.000	2	0.143	0.110	0.008	3	SETAR(1; 2, 3)
NW	2	3	2	0.000	0.000	2	0.023	0.024	0.036	1	SETAR(1; 2, 1)
SD	1	1	1	0.000	0.000	1	0.003	0.311	0.194	1	SETAR(1; 1, 1)
SZ	1	1	1	0.000	0.000	1	0.425	0.337	0.407		
UK	1	1	1	0.000	0.000	1	0.005	0.768	0.759	1	SETAR(1; 1, 1)
US	1	1	2	0.000	0.000	1	0.489	0.386	0.664		

Legend. (1): Countries; (2): Swartz-Bayesian information criterion; (3): Aikake information criterion; (4): Hannan-Quinn information criterion; (5)-(6) Hinich test statistics in p -values for third order autocorrelation and serial dependence from non gaussian errors; (7): selected order p using the following rule: $p = k + n$, $k = \max\{AIC, BIC, HQ\}$ is the maximum number selected by standard information criteria, $n = \{1, 2, 3\}$ are the eventually additional order suggested by Hinich test; (8)-(10): Tsay test for threshold linearity in p -values using $d = \{1, 2, 3\}$ delay parameters; (11): chosen delay parameter d ; (12) selected SETAR($k;p,d$), $k = \{1, \dots, N\}$ regimes.

Software used: RATS

Table 12: Hansen' threshold effect test (bootstrapped p-values)

Rate	supLM	expLM	aveLM	$supLM^h$	$expLM^h$	$aveLM^h$
$y_t^{\mathcal{L}/e}$	0.050	0.088	0.391	0.040	0.083	0.407

Software used: RATS

Table 13: Test for serial correlation in STAR models

Lag	F-values							
	1	2	3	4	5	6	7	8
<i>U.S. numeraire</i>								
DN	0.6577	0.8706	0.5816	0.5104	0.7466	0.6872	0.7210	0.7619
CAN	0.3120	0.2931	0.2291	1.6609	1.2908	1.2815	1.3886	1.3085
SD	1.6392	2.3734	1.7753	1.7076	2.4365	2.0747	1.8644	1.6499
SZ	0.1740	0.3187	0.4115	0.3997	0.3395	0.3610	0.6045	0.5696
UK	0.2756	1.1349	1.1029	1.4913	1.3671	1.3136	1.2506	1.1720
<i>E.U. numeraire</i>								
EU	0.3441	0.1742	0.1418	0.0964	0.1177	0.3255	0.5751	0.5365
Lag	P-values							
	1	2	3	4	5	6	7	8
<i>U.S. numeraire</i>								
DN	0.4190	0.4213	0.6282	0.7282	0.5902	0.6604	0.6544	0.6369
CAN	0.5775	0.7465	0.8760	0.1639	0.2729	0.2716	0.2173	0.2470
SD	0.2029	0.0975	0.1556	0.1530	0.0338	0.0617	0.0822	0.1192
SZ	0.6773	0.7277	0.7450	0.8085	0.8880	0.9021	0.7511	0.1192
UK	0.6005	0.3249	0.3509	0.2093	0.2419	0.2569	0.2817	0.3227
<i>E.U. numeraire</i>								
UK	0.5586	0.8403	0.9347	0.9834	0.9882	0.9223	0.7748	0.8266

Software used: JMulTi 4

Table 14: Test for no remaining nonlinearity

	F-value	DF1	DF2	P-value
<i>U.S. numeraire</i>				
DN	0.32178	6	118	0.924
CAN	0.80135	12	110	0.648
SD	0.78239	12	110	0.667
SZ	0.40780	6	118	0.873
UK*	0.69841	6	116	0.651
<i>E.U. numeraire</i>				
UK	0.96193	3	118	0.413

* Statistic refers to the case $s_t = y_{t-1}$

Software used: JMulTi 4

Table 15: Test for parameter constancy

	F-value			P-value		
	H_1	H_2	H_3	H_1	H_2	H_3
<i>U.S. numeraire</i>						
DN	2.6360	1.5448	1.8829	0.0374	0.1495	0.0440
CAN	1.7044	1.1861	1.4203	0.1537	0.3137	0.1677
SD	2.3735	1.6442	1.7166	0.0562	0.1202	0.0729
SZ	2.4180	1.3907	1.7106	0.0524	0.2080	0.0739
UK	1.8348	1.8005	NaN	0.0984	0.0568	NaN
<i>E.U. numeraire</i>						
UK	1.2717	1.3326	1.7530	0.2759	0.2108	0.0419

Software used: JMulTi 4

Figure 1: PPP, US numeraire (in logs)

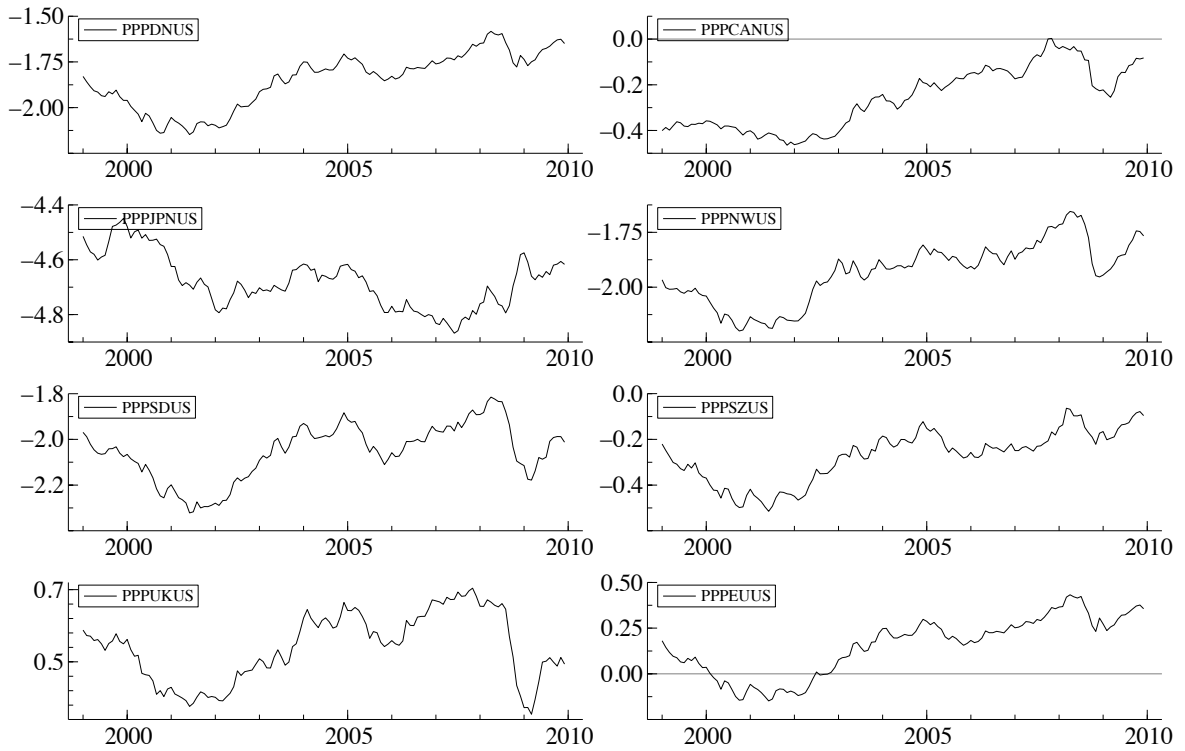


Figure 2: PPP, EU numeraire (in logs)

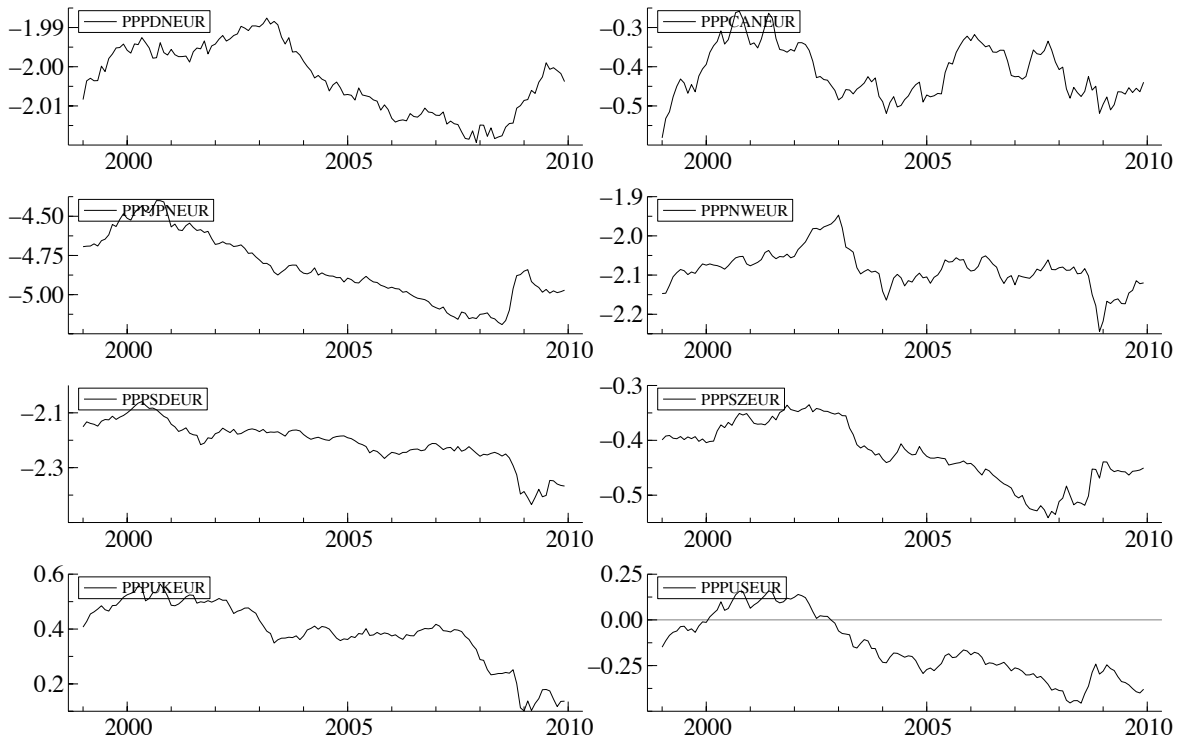


Figure 3: PPP, US numeraire (in logs), first differences

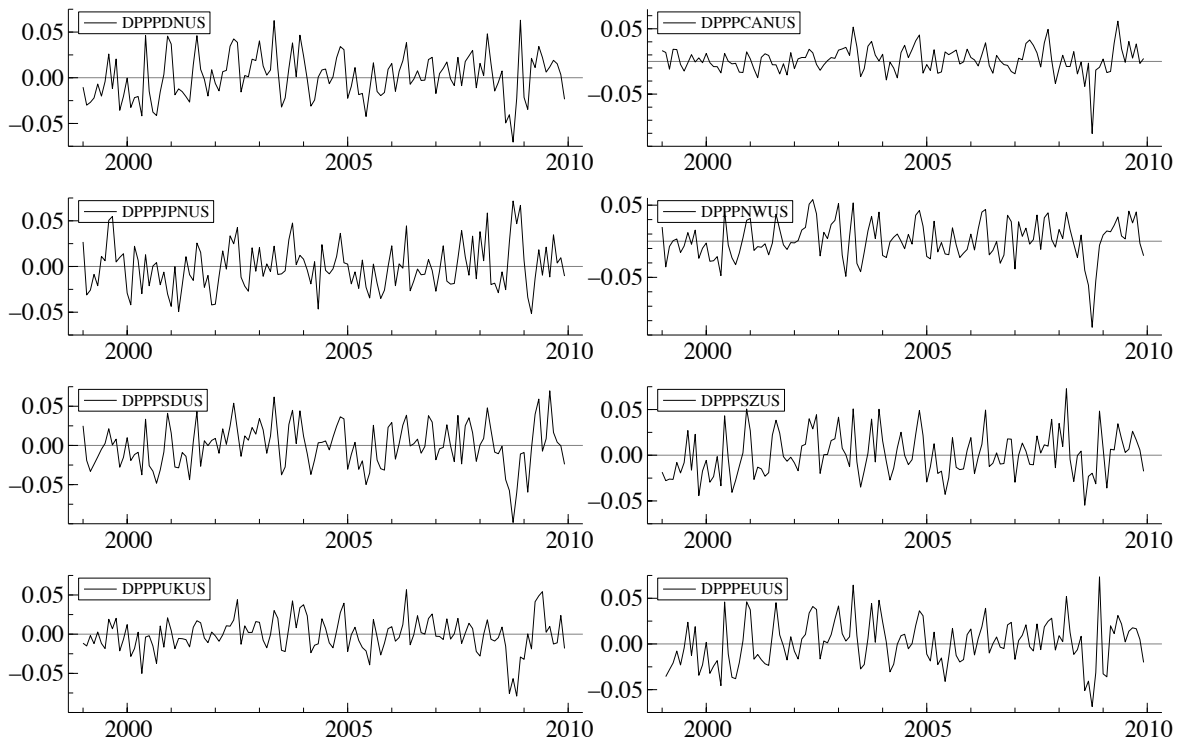


Figure 4: PPP, EU numeraire (in logs), first differences

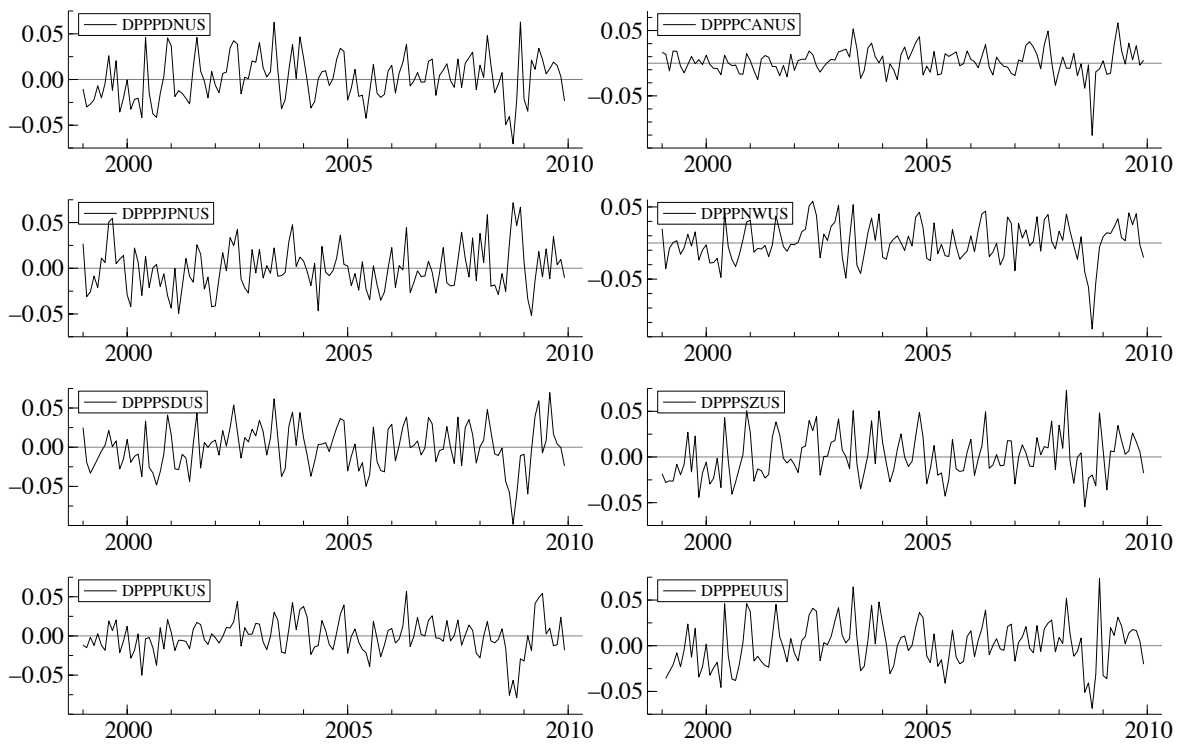
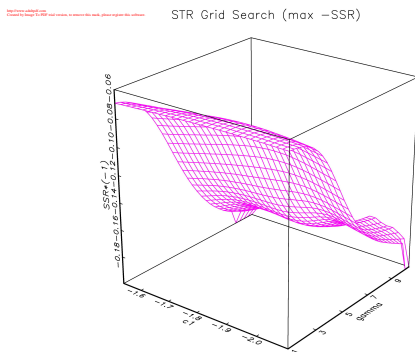
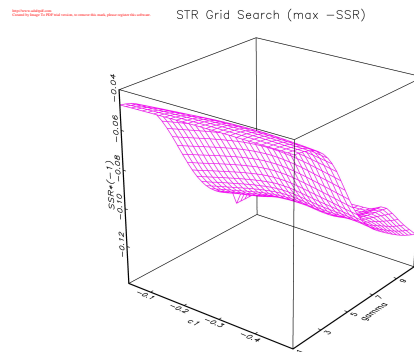


Figure 5: Grid search for starting values of γ and c

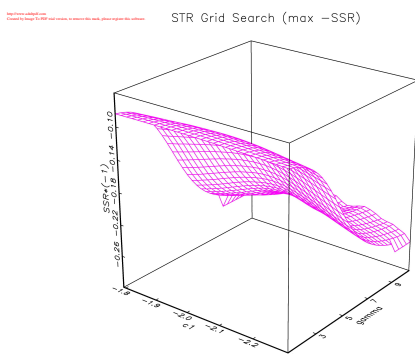
(a) PPPDNUS



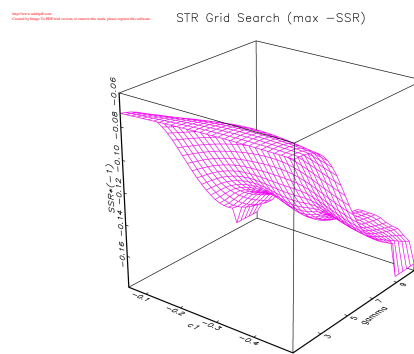
(b) PPPCANUS



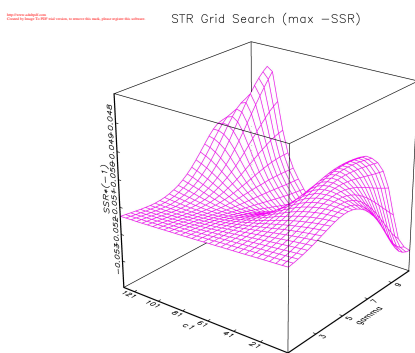
(c) PPPSDUS



(d) PPPSZUS



(e) PPPUKUS



(f) PPPUKEUR

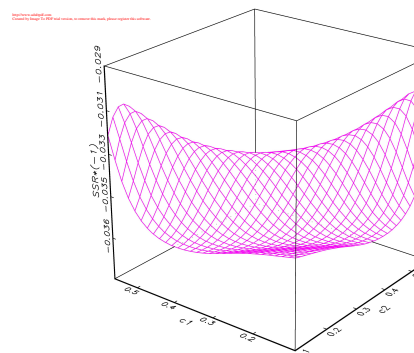
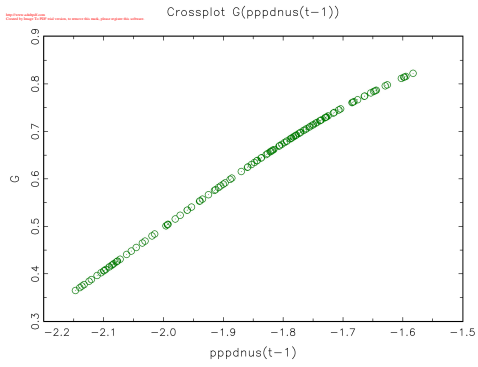
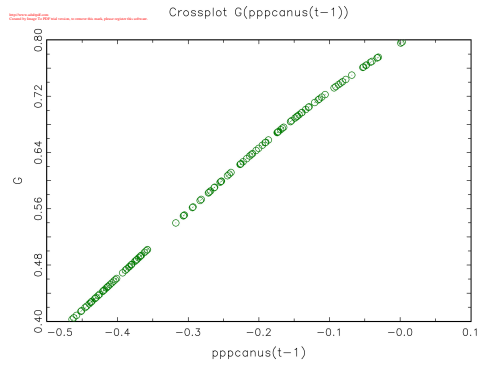


Figure 6: Transition function G for estimated STAR models

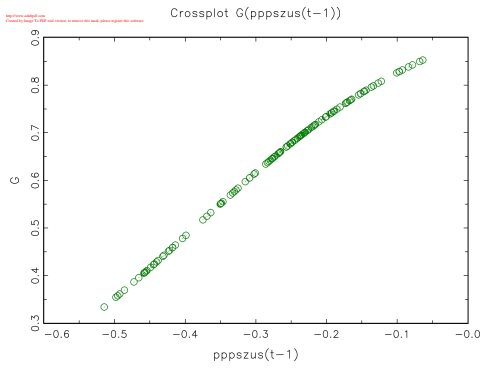
(a) *PPPDNUS*



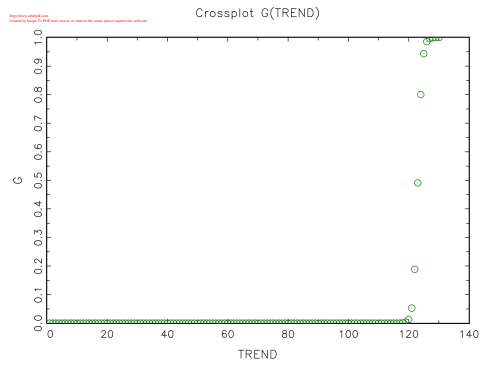
(b) *PPPCANUS*



(c) *PPPSZUS*



(d) *PPPUKUS*



(e) *PPPUKEUR*

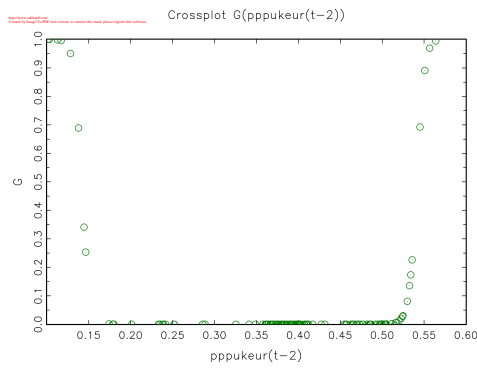
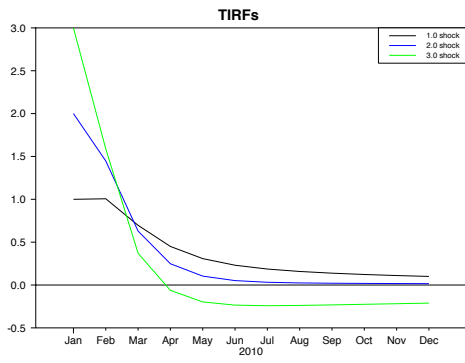
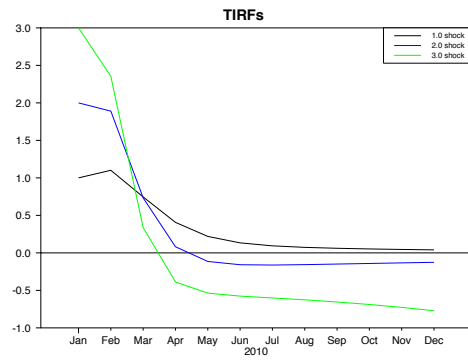


Figure 7: Traditional Impulse Response Functions

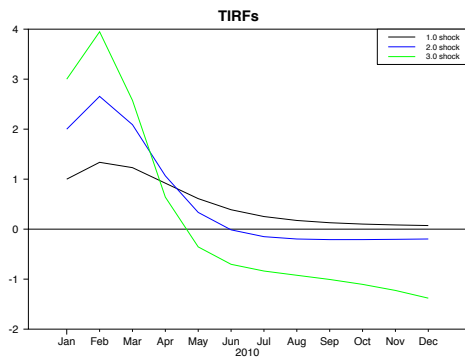
(a) PPPDNUS



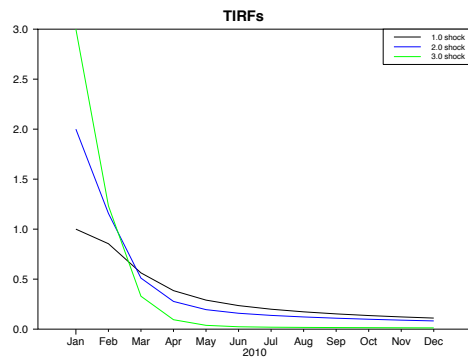
(b) PPPCANUS



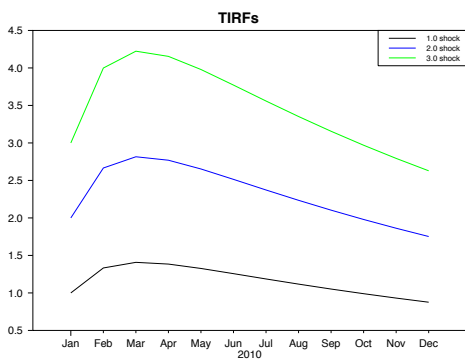
(c) PPPSDUS



(d) PPPSZUS



(e) PPPUKUS



(f) PPPUKEUR

