Updating the PPP puzzle: should we use nonlinear models?

Zanetti Chini, Emilio

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

15 November 2010

Updating the PPP Puzzle: Should We Use Nonlinear Models?

Emilio Zanetti Chini*

"Tor Vergata" University of Rome

Abstract

We investigate the empirical support to the Purchasing Power Parity hypothesis in sixteen real exchange rates for the decade 1999-2009 by implementing Cointegrated VAR analysis, panel cointegration and nonlinear models. The theory is rejected and both the puzzles remain unsolved if considering linear models, while a nonlinear scenario seems to allow for a partial solution to the puzzle if adopting a modified General-to-Specific modelling strategy. The parameters restrictions commonly used in literature and the automatic use of symmetric transitions between different regimes when estimating the conditional mean are criticized and shown being two plausible candidates for explaining the puzzle.

Key words: PPP, real exchange rates, dynamically symmetric models,STAR models, model specification.
JEL Classification: [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) proposition states that the price of a basket of goods expressed in a common currency should be constantly equal to one (absolute version) or

* *E-mail*: Emilio.Zanetti.Chini@uniroma2.it. - Department in Economics and Institutions. This paper is based on my MSc. dissertation in Macroeconometrics at University of Rome "Tor Vergata" entitled *"Does the Parity of Purchasing Power Hypothesis hold? Evidence from the last decade"*. A special thank to my first supervisor Tommaso Proietti and to Fabrizio Mattesini who was the second supervisor. I am particularly grateful to Giovanni Trovato and Barbara Annicchiarico for their useful comments and suggestions. Any error is my own responsibility.

constant (relative version). Rogoff (1996) highlights that there is consensus on the facts that real exchange rates tend toward PPP in the very long run while the speed of convergence towards is extremely slow and that short run deviations from PPP are large and volatile since the half-live is measured in the range of 3-5 years, hence the "PPP puzzle"[1].

Since PPP is assumed is assumed in many wildly used macroeconomic models, there is a huge literature which uses three main methodologies: linear cointegration analysis, panel methods and univariate nonlinear autoregressive models. Only recently some positive results have been achieved: Taylor *et al.* (2001) (TPS) applies the family of smooth transition autoregressive (STAR) models (see Section 2) to four rates and solves the two PPP puzzles[2] by analyzing the standard post-Bretton-Wood sample.

Panel unit root and cointegration techniques has been developed since last 90's (Maddala and Wu, 1999; Pedroni, 2004), until Banerjee *et al.* (2005) (BMO) noticed 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 leads to a severe critique to the commonly used empirical methodology in macroeconomics in the measure of which panel methods are performed automatically.

The issue of the unobserved heterogeneity seemed to be a plausible candidate to go ahead the BMO critique: Imbs *et al.* (2005) explicitly takes in account the heterogeneous dynamics in a panel of sectoral indexes and shows how this heterogeneity is consistent, from a theoretical point of view, with the high persistence of real exchange rates in aggregates indexes and the faster adjustment in sectoral ones because of an upward bias in the traditional estimates

[1] "The purchasing power parity puzzle then is this: 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.

[2] TPS considers the two empirical facts above mentioned (that is long-run mean reversion of real exchange rates with respect their theoretical values and the volatility) separately and calls them "First" and "Second PPP Puzzles" respectively. In this paper we will use the same notation.

of such persistence. Gadea and Mayoral (2009) replies that the heterogeneity is not a valid answer because the GIRF (see next Section 2.4) used by Imbs *et al.* was seriously biased and, consequently, this bias is the cause of the differences between sectoral and aggregate persistence, so that the games are still open.

Finally Johansen *et al.* (2010) solves the PPP puzzles for the DKR/$ rate using an alternative, fully empirically-based approach in a standard sample by implementing a Cointegrated VAR model under an I(2) scenario, justified by a non-conventional economic theory[3].

The main purpose of this article is to investigate the empirical support to the PPP hypothesis for the last 11 years. In order to do this we compare all the three main methodologies previously mentioned. 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; consequently, none of such studies (also the most recent ones) mention the euro nor, *a fortiori* the effects of the 2008 crisis. Second, the model specification in almost all the literature is highly driven by theoretical reasons (in particular in nonlinear models, see Section 3). In the next sections we will try to bridge this gap in empirical literature and to compare the three methodologies 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 model specification by introducing some modifications to the available strategies 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. We anticipate the two main results: first, the dynamics of real exchange rate seems to be asymmetric contrarily to what suggested by standard literature; second, the peculiarity of the estimates suggests that only a small part of the mean reversion can be captured by standard nonlinear models because of the highly restric-

[3]See Sarno and Taylor (2001) for a survey of standard applied literature prior to 2001

tive definition of (a)symmetry implicitly used in applied literature. The paper is organized as follows: Section 2 states the relations of interest and the new definition of (a)symmetry used in the paper 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 models

2.1 Economic relations and definitions

Following Juselius (2009) notation and working with aggregate terms in logarithmic transformation, we define the PPP as:

$$p_t = p_t^* + s_t + v_t \tag{1}$$

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":

$$v_t \equiv y = p_t - p_t^* - s_t, \tag{2}$$

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

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

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:*

$$CI = (1 \;\; -1 \;\; -1) \tag{4}$$

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

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

The previous postulates and definitions assume a linear model. In practice, it is well known that economic variables behaves nonlinearly in the short/medium-run, as we will proove in the next Sec. 5. Hence the need an appropriate definition of what we call "nonlinearity" from a statistical point of view:

Definition 3 (Nonlinear process)**.** *Let* $\{v_t\}_t^T$ *be a stationary stochastic process with conditional mean* $m_t = E(v_t|v_{t-1}, \dots, y_{t-m})$ *where* v_t *defined as in* (2)*. The process is defined nonlinear if* $m_t = f(\boldsymbol{y_t})$*, where* $f(\boldsymbol{y_t})$ *is any (possibly twice differenciable) function in* $\mathbf{R}^{m+1}$ *and* $\boldsymbol{y_t} = [v_t, \gamma_t]$*,* γ_t *is a velocity parameter.*

The previous Def. 3 only require that the conditional mean is not a constant or a line. In practice, in applied literature, one require some other dynamics properties, such as monotony and symmetry. These two requirements are particularly interesting for our aims because they are assumed in econometric models. Hence the following

Definition 4 (Dynamically (a)symmetric and (a)symmetric models)**.** *Consider a nonlinear model for* m_t *satisfying Def. 3. Suppose that data suggest two different levels for* m_t*, say* m_1 *and* m_2 *respectively. Call the mild-point between* m_1 *and* m_2 *as* $m_a = (m_1+m_2)/2$*. If* $f(\boldsymbol{y_t})$ *is such that: i)* m_t *moves from* m_1 *to* m_2 *with increasing* *velocity until m_a, where it begin to decrease at the same rate of the previous increase; ii) an increase in the velocity produce a monotone increase in m_t; then the model is called* dynamically symmetric. *A model satisfying only ii) is defined* symmetric. *A model not satisfying i) and ii) (not satisfying ii)) is defined* dynamically asymmetric *(*asymmetric*).*

According to our definition, a typical dynamically symmetric process is the Logistic Smooth Transition Auto-Regressive model (see Sec. 2.4).

Remark 1. *Notice that Def 4 does not concern about the permanence of m_t on its new level, but only about the dynamics of its velocity of transition; that is, the fact that m_t does not return to its previous level is not sufficient to define the process as asymmetric.*

Remark 2. *Notice that a dynamically symmetric model is considerably more stringent than a symmetric one from a statistical point of view. Hence it is a testable hypothesis which should be checked any time the econometrician uses a nonlinear model for m_t.*

This definition of (a)symmetry is the main contribution of this paper, since the econometric and applied literature implicitly consider asymmetry only by looking at the level of m_t, or equivalently, uses a dynamically symmetric structure without test for it, see TPS (pag. 1020) and Sec. 2.4.

2.2 Statistical models: CVAR

For what concerns the cointegration analysis of PPP, we use a VECM(p) to model relations (1) and (2)[4]:

$$\boldsymbol{\Delta y_t = \Gamma_1^{(1)} \Delta y_{t-1} + \Gamma_2^{(1)} \Delta y_{t-2} + \ldots + \Gamma_{p-1}^{(1)} \Delta y_{t-p-1} + \alpha \beta' x_{t-1} + \mu_0 + \mu_1 t + \epsilon_t} \tag{5}$$

[4]Remember that the VECM(p) is only a way for re-writing the more conventional VAR(p) when the data are supposed to be I(1), hence the "CVAR" notation remains still valid.

where: $\mathbf{\Gamma_1^{(1)}} = -(\mathbf{\Pi_2} + \mathbf{\Pi_3} + \ldots + \mathbf{\Pi_p})$, $\mathbf{\Gamma_2^{(1)}} = -\mathbf{\Pi_3}$ and $\mathbf{\Pi} = -(\boldsymbol{I} - \mathbf{\Pi_1} - \mathbf{\Pi_2} - \ldots - \mathbf{\Pi_p})$ are the short run matrices and the long run matrix respectively and the integer (1) indicates the lag pleacement of ECM, $\mathbf{\Pi} = \boldsymbol{\alpha\beta'}$ is the reduced rank long run matrix, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $p \times r$ matrices, $r \leq p$, $\boldsymbol{\mu_0} + \boldsymbol{\mu_1} t = \mathbf{\Phi}\boldsymbol{D_t}$ are the unrestricted components (i.e. allowed to enter in cointegrating relation) of deterministic trend. The equation (5) 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:

$$\Delta y_{it} = \rho_i y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{i,L} \Delta y_{i,t-L} + \alpha_{mi} d_{mt} + \epsilon_{it} \;\; m = 1, 2, 3 \tag{6}$$

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 (6) 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 non-stationary 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 used simple panel versions of LM and ADF-based procedure in order to test the two opposite null hypothesis systems. In this paper we implement two on the seven tests developed by Pedroni (2004) because, differently to the previous ones, 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 two statistics are more powerful than Pedroni's ones.

2.4 Statistical models: nonlinear time series

Concerning the nonlinear scenario, we use the standard STAR and its particular case, the self-exciting threshold autoregressive (SETAR) 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 broadly describe the econometric methodology step by step using the notation by Teräsvirta (2006) (since now, Teräsvirta) to which we remind for technicalities. Consider the general additive non-linear model:

$$y_t = \boldsymbol{\phi' z_t} + \boldsymbol{\theta' z_t} G(\gamma, \boldsymbol{c}, s_t) + \epsilon_t \tag{7}$$

where $y_t \equiv v_t$ in equation (2), $\boldsymbol{z_t} = (1, y_1, \ldots, y_{t-p})'$, $\boldsymbol{\phi} = (\phi_0, \phi_1, \ldots, \phi_p)'$, $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_p)'$ are parameter vectors, and $\epsilon_t \sim i.i.d.(0, \sigma^2)$, the transition function $G(\gamma, \boldsymbol{c}, s_t)$ is a continuous function in the transition variable s_t[5] where γ controls the velocity of the transition and $\boldsymbol{c} = (c_1, \ldots, c_K)$ is a vector of transition parameters.

One of the main used functions for $G(\cdot)$ is the (first order) logistic function:

$$G(\gamma, \boldsymbol{c}, s_t) = \left(1 + exp\left\{-\gamma \prod_{k=1}^{K} (s_t - c_k)\right\}\right)^{-1}, \quad \gamma > 0, \tag{8}$$

where $\gamma > 0$ is an identifying restriction. Equations (8) and (7) define the first order Logistic STR (LSTR1) model. The most common choices for K are $K = 1$, in which case the parameters $\boldsymbol{\phi} + \boldsymbol{\theta} G(\gamma, \boldsymbol{c}, s_t)$ change monotonically as a function of s_t from $\boldsymbol{\phi}$ to $\boldsymbol{\phi} + \boldsymbol{\theta}$ and $K = 2$, in which case the parameters $\boldsymbol{\phi} + \boldsymbol{\theta} G(\gamma, \boldsymbol{c}, s_t)$ change *symmetrically*[6] around the mid-point $(c_1 + c_2)/2$ where the logistic

[5] Notice 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.

[6] Notice that here the term "symmetrically" is referred only to the level of the conditional mean, while according to our definition (see Def. 4 on page 5), the LSTR1 is a dynamically symmetric model since γ is implicitly assumed constant. We will show in next Section 5 that this could lead to misleading results.

function attains its minimum, $min_G G(\cdot) \in [0, 1/2]$, and it's such that:

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

If $\gamma = 0$, the transition function $G(\gamma, \boldsymbol{c}, s_t) \equiv 1/2$ so that model (7) nests a linear model. In the latter case, that is when $K = 2$ and $c_1 \neq c_2$ the transition function became a second order Logistic STR (LSTR2). A peculiar form of this latter case is when $K = 2$ and $c_1 = c_2$ and the transition function (8) becames:

$$G(\gamma, \boldsymbol{c}, s_t) = 1 - exp\{-\gamma(s_t - c)^2\}, \;\; \gamma > 0 \tag{9}$$

Equations (7) and (9) define the Exponential STR (ESTR) model. When $\boldsymbol{z_t} \equiv \boldsymbol{y_{t-d}}$ and $s_t \equiv y_{t-d}$, $d > 0$ in (8) and (9), the model becomes an LSTAR1, an LSTAR2 and an ESTAR respectively. Similarly, when $\gamma \to \infty$ and $\boldsymbol{z_t} \equiv \boldsymbol{y_t}$ and $s_t \equiv y_{t-d}$ the model (7) nests a SETAR model:

$$y_t = \sum_{j=1}^{r+1} (\boldsymbol{\phi_j' y_t}) I\big(y_{t-d} \leq c_j\big) + \sum_{j=1}^{r+1} (\boldsymbol{\phi_j' y_t}) I\big(y_{t-d} > c_j\big) + \epsilon_{jt} \tag{10}$$

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

Concerning step (i) (specification), Tsay (1989) 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, deter-

mine 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 (1996) discusses an alternative likelihood-based test. Three statistics are used:

$$\begin{aligned} S_T &= \sup_{\boldsymbol{\gamma} \in \boldsymbol{\Gamma}} S_T(\boldsymbol{\gamma}), \\ aveS_T &= aveS_T(\boldsymbol{\gamma}) = \int_{\boldsymbol{\Gamma}} S_T(\boldsymbol{\gamma}) dW(\boldsymbol{\gamma}), \\ \exp S_T &= \ln\left(\int_{\boldsymbol{\Gamma}} \exp\left\{ \frac{1}{2} S_T(\boldsymbol{\gamma}) \right\} dW(\boldsymbol{\gamma}) \right) \end{aligned} \tag{11}$$

where $\boldsymbol{\Gamma} = \{\boldsymbol{\gamma} : \boldsymbol{\gamma} \in \boldsymbol{\Gamma}\}$, $W(\boldsymbol{\gamma})$ is a weight function such that $\int_{\boldsymbol{\Gamma}} W(\boldsymbol{\gamma}) d\boldsymbol{\gamma} = 1$. Using the likelihood function of the model (7), 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 (8) (or (9)), $T_3(z) = g_1 z + g_3 z^3$ 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, $y_t = \boldsymbol{\phi' z_t} + \boldsymbol{\theta' z_t} T_3(\gamma(y_{t-d} - c)) + \boldsymbol{\epsilon}_t$, leads to the auxiliary regression:

$$\hat{\boldsymbol{\epsilon}}_t = \boldsymbol{\hat{z}'_{1t}} \boldsymbol{\tilde{\beta}_1} + \sum_{j=1}^{p} \beta_{2j} y_{t-j} y_{t-d} + \sum_{j=1}^{p} \beta_{3j} y_{t-j} y_{t-d}^2 + \sum_{j=1}^{p} \beta_{4j} y_{t-j} y_{t-d}^3 + v'_t \tag{12}$$

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

$$LM_2 = (SSR_0 - SSR)/\hat{\sigma}^2 \sim \chi^2(3p) \tag{13}$$

where SSR are the sum of squared residuals form equation (12) or, alternatively, by setting the artificial model $y_t = g_1\gamma_0 + \gamma_1'\boldsymbol{z}_t + \gamma_2'(\boldsymbol{z}_t y_{t-d}) + \gamma_3'(\boldsymbol{z}_t y_{t-d}^2) + \gamma_4'(\boldsymbol{z}_t y_{t-d}^3) + v_t''$ where $v'' \sim nid(0, \sigma_{v''}^2)$, $\gamma_j = (\gamma_{1j}, \cdots, \gamma_{jp})'$, and $j = 1, \cdots, 4$, and $H_0 : \gamma_2 = \gamma_3 = \gamma_4 = 0$. In terms of Taylor approximations we get:

$$\begin{aligned}
\gamma_2 &= g_1\gamma\hat{\boldsymbol{\theta}} + 3g_3\gamma^3 c^2\hat{\boldsymbol{\theta}} - 3g_3\gamma^3 c\theta_0\boldsymbol{e}_d \\
\gamma_3 &= -3g_3\gamma^3 c\hat{\boldsymbol{\theta}} + g_3\gamma^3\theta_0\boldsymbol{e}_d \\
\gamma_4 &= g_3\gamma^3\hat{\boldsymbol{\theta}}
\end{aligned} \tag{14}$$

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

$$\hat{v}_t = \tilde{\boldsymbol{\beta}}_1'\hat{\boldsymbol{z}}_{1t} + \boldsymbol{\beta}_2'\boldsymbol{z}_t y_{t-d} + \boldsymbol{\beta}_3'\boldsymbol{z}y_{t-d}^2 + e_t' \tag{15}$$

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 $\boldsymbol{\beta}_1 = \bar{\boldsymbol{\phi}} - c^2\bar{\theta} + 2c\theta_0\boldsymbol{e}_d$, $\boldsymbol{\beta_2} = 2c\bar{\theta} - \theta_0\boldsymbol{e}_d$. The null of linearity is $H_0' : \boldsymbol{\beta}_2 = \boldsymbol{\beta}_3 = 0$ which is tested by statistic

$$LM_3 = (SSR_0 - SSR)/\hat{\sigma}^2 \sim \chi^2(p) \tag{16}$$

where SSR is the sum of squared residuals from (15). In order to choice the correct transition function, Teräsvirta (1994) proposes the following nested hypotheses test (the so called "Teräsvirta rule"):

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

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) (Dyagnostic 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:

$$\begin{aligned} TIRF(h,\delta,\omega_{t-1}) =& E[y_{t+h}|\epsilon_t=\delta,\epsilon_{t+1}=\cdots=\epsilon_{t+h}=0,\omega_{t-1}]- \\ & -E[y_{t+h}|\epsilon_t=0,\epsilon_{t+1}=\cdots=\epsilon_{t+h}=0,\omega_{t-1}], \end{aligned} \tag{18}$$

for $h=0,1,2,\ldots$. The TIRF is commonly used in linear systems because of its three properties: first, it's symmetric, that is a 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 (18), 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:

$$GIRF(h,\epsilon_t,\Omega_{t-1})=E[y_{t+h}|\epsilon_t,\Omega_{t-1}]-E[y_{t+h}|\Omega_{t-1}]. \tag{19}$$

for $h=0,1,2,\ldots$. 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; in particular the price homogeneity and the order of integration constitute the main block of the whole strategy since all conventional literature makes assumption on them[7]. 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 models[8]. 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 dummies[9]; 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 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

[7]See Juselius (2009) for details and an application to the conventional case of $DMK/\$$ in the conventional sample 1975-1998.

[8]"The 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)

[9]That 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.

forced to use a more traditional theory based approach[10].

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:

i. Select the lag p for the system (5) by using standard information criteria.

ii. Once p is selected, check for normality and residuals first and second-order autocorrelation; in the case, augment p until necessary.

iii. 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.

iv. 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, simulate the critical values when outlier has been found in Step 3.

v. If $rank(\mathbf{\Pi}) \neq 0$, set the rank of $\mathbf{\Pi}$ matrix.

vi. 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.

vii. The so found restricted VECM are analyzed in their cointegrating relations by recursive tests.

[10]However 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.

viii. Check for the presence of I(2)-ness. If found, the whole analysis should be reconsidered in an I(2) scenario.

In the nonlinear framework we follow TPS as benchmark and apply an analysis *à là* Teräsvirta (1994) 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 (7) is restricted for $\boldsymbol{\phi} = -\boldsymbol{\theta}$, because this ensure an economic interpretation similar to Coleman (1995)[11]. 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 level[12].

Since the ESTAR model is a particular case of the more general family of LSTAR, 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[13]. For this reason, we follow a more agnostic policy in modeling our series, allowing some parameters (the constant, in a lot of cases) for no restrictions. The resulting Granger-Teräsvirta modeling procedure has been consequently adapted as here described in detail:

[11] "[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. 1030).

[12] TPS 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. 1021).

[13] This last empirical finding has been checked as preliminary step to the draft of this paper in order to replicate the results by TPS. The results, jointly with more detailed critiques to the TPS methodology, can be sent under request.

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

ii. *Specification of linear AR(p) part of the model (7).* Allow for the possibility for model (7) to have some zero-coefficient in both linear and nonlinear part, starting with the hypothesis of no restrictions.

iii. *Linearity tests.* Apply the Teräsvirta rule (17) for all possible candidates transition variable $s_t = \tilde{\boldsymbol{z}}_t \equiv (y_{t-1}, ..., y_{t-p}, t)'$; 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 ii) 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 i), augment the order p until the maximum p estimated by information criteria (possibly augmented by Hinich' test) and restart the procedure.

iv. *Grid search for starting values of nonlinear parameters.*

v. *Estimation.* Use the selected transition variables, transition function and starting value in order to compute the parameter estimates by NLLS algorithm. Impose the restriction $\boldsymbol{\phi} = -\boldsymbol{\theta}$ 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 all parameters are significantly different from zero), continue with next step; otherwise, restrict y_{t-i}, $i = 1, ..., p$ singularly and continue with next step; otherwise, restrict al 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 ii) and restart the procedure. If also the new procedure defaults, return to Step i), augment the order p until the maximum p estimated by information criteria and restart the whole procedure.

vi. *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 v) 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 iii), set different transition variables and restart the procedure. If this is not an help, return to Step ii) 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 i) 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 indicators[14]. The series of real exchange rates has been built by equation (2).

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 addiction 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. For detailed description of the dataset, see Appendix A.1.

The fact that the dataset is composed by a limited number of advanced economies ensures that the empirical analysis is not seriously biased by Balassa-Samuelson effect when considering very aggregate ("all items") CPIs. Moreover, we stress the fact that the whole applied literature, with the exception constituted by Imbs *et al.* (2005) and Gadea and Mayoral (2009), does not concerns at all of the effect of the various CPI indexation. Finally, we remember that there is no universally accepted price index adjusted for export due to the high measurement error in the computation of the quantity of goods exported.

5 Empirical evidence

A graphical analysis of the series shown in Figure 1 reveals several important features: i) all the real exchange rates follow a positively (negatively) shaped linear broken trend when \$ (€) is the numeraire; ii) the break is located approximately in 2008:04 but continues until

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

first months of 2009, a fact which suggest a regime switching in the levels; iii) this change in regime is weaker when the € is numeraire because of the speculative attack during the financial crisis; iv) all the series show one or more breaks in the middle of the sample, corresponding to the selected dummy variables in Table 3, two of which (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; v) the presence of autocorrelation in rates DN/$, EU/$, DN/€, SD/$, SZ/$ and ARCH-effects is observable in series in first differences (not reported here). While considering these findings, it is not a surprise that strong PPP hypothesis does not seem to hold. Namely we used 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 with a predetermined number of additional lags (namely, $p = 0, \ldots, 3$) by "Marshallian" approach discussed in Section 3. Table 1 shows our results: 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. 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.

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 in order to 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. Table 3 illustrates our results for PPP when using \$ or € as numeraire. In the majority of cases our optimal lag choice is 2. However, because of the presence of some outliers, 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)-ness[15]. Moreover, 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, which is not uncommon in financial variables.

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 dependance across disturbances. Table 4 shows 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 two most powerful statistics on the seven proposed by the author, $\tilde{Z}_\rho$ and $Z_{t_{\hat{\rho}_{NT}}}$. 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, so that, similarly to the CVAR framework, the statistic critical values are biased

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

when allowing for a deterministic kernel to enter in cointegrating relations. In this case, robust critical values can be computed by bootstrap methods. Table 5 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

Concerning the First Puzzle we checked for the presence of ARCH-effects by performing the McLeod and Li (1983) test before to implement the Granger-Teräsvirta procedure. The results are given in Table 2. 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. Table 8 shows the results of the STAR specification procedure above explained for our dataset. We identified seven LSTR1 models when \$ is numeraire (six when € is numeraire, one of which is LSTR2). 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 correction is CAN/\$ (three lag added by procedure described in Sec. 3). The estimates of selected STAR models are reported in Table 9.

Figure 2 plots the transition function $G(\gamma, c, s_t)$ as function of the transition variable s_t. On thirteen estimated nonlinear models, five (DN/\$, CAN/\$, SD/\$, SZ/\$ and €/\$) are at the line with linear models, since their mean reversion is very smooth. The other models 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 models NW/\$ and SD/\$ are the more restricted one, since all lags of $y_t^{NW/\$}$ and $y^{SD/\$}$ enter in the restriction $\boldsymbol{\phi} = -\boldsymbol{\theta}$, so the interpretation is very similar to that given in TPS. On the contrary, model £/€ 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; this is the only estimated LSTR2 model in the dataset. 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, 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 Table 10 does not support the idea of a third regime for the estimated model for any model. Third, and most important, few of the estimated transition function are puzzling if considering the results form linearity tests used in Teräsvirta, unless reconsidering the intrinsically symmetric (in the sense of Def. 4) structure of the same one in STAR family. That is, we suspect that such quasi-linear behavior could be only apparent because data has been forced to be estimated by using a dynamically symmetric rather than a symmetric statistical model. In other world, the fact that the econometric literature does not allow to test for and parametrize an eventual change in the velocity of transition from m_1 to m_a w.r.t. m_a to m_2 seems us able to generate a sort of "neglected dynamic asymmetry" puzzle which could explain some of the difficulties in fitting a statistically significant mean reversion.

Table 9 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 (11) 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 rate £/€as shown in Table 7. 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:

$$
\begin{aligned}
y_t^{£/e} &= \underset{(0.013)}{-0.022} + \underset{(0.055)}{0.849} \cdot y_{t-1}^{£/e} \quad if\ y_{t-1}^{£/e} \leq 0.341, \quad \hat{\sigma}^2 = 0.0007 \\
y_t^{£/e} &= \underset{(0.011)}{0.006} + \underset{(0.026)}{0.984} \cdot y_{t-1}^{£/e} \quad if\ y_{t-1}^{£/e} > 0.341, \quad \hat{\sigma}^2 = 0.0002
\end{aligned}
\tag{20}
$$

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 Table 2 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.

Concerning the Second Puzzle, the previously introduced "neglected dynamic asymmetry puzzle" leads us to not follow TPS pedantically in using the GRT method for GIRF analysis in order to study the shocks persistence of real exchange rates and to use TIRF for the five models previously mentioned. Nevertheless, neither in the other series 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 (neither in the common meaning of the term, although it can be seen as an approximation), hence the economic interpretation of the resulting GIRF could be misleading[16]. Figure 3 plots the TIRF/GIRF of the real exchange rate series for shocks of magnitude $\{1, 2, 3\}$ and an horizon of 12 month. It's easy to notice the different behavior of the two groups: in TIRFs, when the shock is larger (shock=3, green line) the TIRF does converge faster to 0 and tends immediately to go slightly below such threshold; more in general, 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. On the other hand, when considering the GIRFs, the shocks are more problematic, since they does not seem to have a clear path. It's interesting to note that TPS concerns for a larger scaled dataset[17] 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 by using different methodologies in both linear and nonlinear scenario in order to compare and update the available empirical literature.

The general result is ambivalent: the analysis of a dataset of 16 real exchange rates does not support the PPP hypothesis when a linear scenario is used. In particular, the CVAR analysis show the

[16]Moreover, 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 business cycle will tend to weaken or hide the evidence of asymmetry"(*Ibidem*).

[17]TPS used a sample of 288 observations, an horizon of 200 observation and a set of six shocks going from 5 to 50

two cases in which weak PPP holds are found to be strongly I(2). Panel methods for unit root and cointegration confirm the rejection of the theory, although the BMO's critique suggests to take such results very carefully. Things seem different in the nonlinear scenario: 13 rates on 16 are nonlinearly mean reverting, where 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. In particular, two findings are puzzling: first, the qualitative analysis of the nonlinear part of several STAR models shows a quasi-linear transition function, a fact that leads us to suspect a sort of neglected dynamic asymmetry in transition functions due to a gap in the econometric methodology used. Second, data are not able to support the choice of an ESTAR-type of nonlinearity in favor of an LSTAR-type; that is, contrarily to TPS, the speed of adjustment of exchange rates varies according to the over(under)valuation of the numeraire. We conclude that the solution of the two PPP puzzles needs a methodological approach slightly different form the TPS one when considering a nonlinear scenario. In particular, the problem of capturing the effect of the dynamically asymmetry in model parameters is still today unsolved by the econometric literature. Hence standard nonlinear models can be reasonably used only when strongly nonlinear transition functions are estimated. Our modified General-to-Specific modelling strategy seems being an help but only as a very preliminary step. However, for what concerns the Second Puzzle, TIRFs for estimated real exchange rates for an horizon of one year confirm the 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: in period of crisis shocks in real exchange rates tends to no return to their previous level.

Finally, we suggest some lines for future research. Extending our to the multivariate scenario could be the natural extension of this work, but the problem of neglected asymmetry before mentioned could be exacerbated. For this reason we think that relaxing the assumption of symmetry in transition functions could be a preferable strategy. This would require a re-examination of the whole structure of STR family. We remind this extension to further works.

References

Banerjee, A., Marcellino, M. and Osbat, C. (2005). Testing for PPP: Should we use panel methods? *Empirical Economics*, **30**, 77–91, DOI: 10.1007/s00181-004-0222-8.

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. and Juselius, K. (2006). CATS in RATS. Cointegration Analysis of Time Series, Version 2. Estima: Evanston.

Eitrheim, O. and 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. and Stock, J. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, **64**, 813–836.

Gadea, M. and Mayoral, L. (2009). Aggregation is Not the Solution: The PPP Strikes Back. *Journal of Applied Econometrics*, **24**, 875–894, DOI: 10.1002/jae.1078.

Gallant, A., Rossi, P. and Tauchen, G. (1993). Nonlinear Dynamic Structures. *Econometrica*, **61**, 871–907.

Granger, C. and 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 Statistic*, **6**, 205–221, DOI: 10.1080/10485259608832672.

Im, K., Pesaran, M. and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, **115**, 53–74.

Imbs, J., Muntaz, H., Ravn, M. and Rey, H. (2005). PPP Strikes Back: Aggregation and Real Exchange Rate. *Quarterly Journal of Economics*, **120** (1), 1–43.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vectors Autoregressive Models. *Econometrica.*, **59**, 1551–1580.

— (2002). A small sample correction for tests of hypotheses on the cointegrated vectors. *Journal of Econometrics*, **111**, 195–221.

—, Juselius, K., Frydman, R. and 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.

— (2009). The Long Swings Puzzle: What the Data Tell When Allowed to Speak Freely. In K. Patterson and T. Mills (eds.), *The Palgrave Handbook of Empirical Econometrics*, Macmillan, pp. 367–384.

Koop, G., Pesaran, M. and 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. and 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. and 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. and Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika*, **75**, 491–499.

MADDALA, G. and 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.

MCLEOD, A. and 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. and 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.

— (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 Dependance. *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. and TAYLOR, M. (2001). Purchasing Power Parity and the Real Exchange Rates. CEPR Discussion Papers 2913.

TAYLOR, M., PEEL, D. and SARNO, L. (2001). Nonlinear Mean Reversion in Real Exchange Rates: Towards a Solution to the Purchasing Power Parity Puzzles. *International Economic Review*, pp. 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.

— (2006). Forecasting Economic Variables with Nonlinear Models. In G. Elliott, C. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, vol. 1, Elsevier BV., pp. 414–457.

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. Sample: 1999:01-2009:12 (132 observation). Spot rate series with basis USD source: FED of St. Louis. Series names:
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 source: European Central Bank. Dataset name: Exchange Rates; frequency: monthly; currency denominator: Euro; exchange rate type: spot; series variation - EXR context: average or standardized measure for given frequency. Series names:
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 source: OECD. Series names:
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 series has been built.

A.2 Tables and Graphs

Table 1: Univariate Tests on Real Exchange Rates

	US numeraire				EU numeraire			
Lag	0	1	2	3	0	1	2	3
ADF								
CAN	-1.776	-2.320	-2.300	-2.411	$-3.283^{\bullet}$	-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	$-3.269^{\bullet}$
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
DF-GLS								
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
KPSS								
CAN	$1.060^{\bullet\bullet\bullet}$	$0.541^{\bullet\bullet\bullet}$	$0.369^{\bullet\bullet\bullet}$	$0.284^{\bullet\bullet\bullet}$	$0.591^{\bullet\bullet\bullet}$	$0.311^{\bullet\bullet\bullet}$	0.216^{**}	$0.168^{\bullet\bullet}$
DN	$0.799^{\bullet\bullet\bullet}$	$0.414^{\bullet\bullet\bullet}$	$0.285^{\bullet\bullet\bullet}$	$0.221^{\bullet\bullet\bullet}$	$1.310^{\bullet\bullet\bullet}$	$0.675^{\bullet\bullet\bullet}$	$0.460^{\bullet\bullet\bullet}$	$0.352^{\bullet\bullet\bullet}$
JPN	$1.190^{\bullet\bullet\bullet}$	$0.612^{\bullet\bullet\bullet}$	$0.421^{\bullet\bullet\bullet}$	$0.326^{\bullet\bullet\bullet}$	$1.080^{\bullet\bullet\bullet}$	$0.559^{\bullet\bullet\bullet}$	$0.384^{\bullet\bullet\bullet}$	$0.297^{\bullet\bullet\bullet}$
NW	$0.694^{\bullet\bullet\bullet}$	$0.359^{\bullet\bullet\bullet}$	$0.248^{\bullet\bullet\bullet}$	$0.192^{\bullet\bullet}$	$0.885^{\bullet\bullet\bullet}$	$0.463^{\bullet\bullet\bullet}$	$0.324^{\bullet\bullet\bullet}$	$0.254^{\bullet\bullet\bullet}$
SD	$0.837^{\bullet\bullet\bullet}$	$0.430^{\bullet\bullet\bullet}$	$0.294^{\bullet\bullet\bullet}$	$0.226^{\bullet\bullet\bullet}$	$0.755^{\bullet\bullet\bullet}$	$0.394^{\bullet\bullet\bullet}$	$0.273^{\bullet\bullet\bullet}$	$0.212^{\bullet\bullet}$
SZ	$0.686^{\bullet\bullet\bullet}$	$0.359^{\bullet\bullet\bullet}$	$0.349^{\bullet\bullet\bullet}$	$0.194^{\bullet\bullet}$	$1.200^{\bullet\bullet\bullet}$	$0.620^{\bullet\bullet\bullet}$	$0.423^{\bullet\bullet\bullet}$	$0.324^{\bullet\bullet\bullet}$
UK	$1.150^{\bullet\bullet\bullet}$	$0.586^{\bullet\bullet\bullet}$	$0.399^{\bullet\bullet\bullet}$	$0.306^{\bullet\bullet\bullet}$	$1.110^{\bullet\bullet\bullet}$	$0.581^{\bullet\bullet\bullet}$	$0.402^{\bullet\bullet\bullet}$	$0.312^{\bullet\bullet\bullet}$
EU	$0.843^{\bullet\bullet\bullet}$	$0.437^{\bullet\bullet\bullet}$	$0.301^{\bullet\bullet\bullet}$	$0.233^{\bullet\bullet\bullet}$	-	-	-	-
US	-	-	-	-	$0.860^{\bullet\bullet\bullet}$	$0.446^{\bullet\bullet\bullet}$	$0.307^{\bullet\bullet\bullet}$	$0.238^{\bullet\bullet\bullet}$

$\bullet$ 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. Software used: RATS 7.2

Table 2: McLeod-Li test for no ARCH-effects in ΔPPP (p-value)

	US numeraire				EU numeraire			
Lag	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	-	-	-	-

Software used: RATS 7.2

Table 3: Cointegration Analysis for System of Country j

(1) j	(2) SC	(3) HQ	(4) p	(5) 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 4: Panel unit root test for Real Echange Rates (lag=1)

	Lag =1				
	US numeraire		EU numeraire		
Method	Statistics	p-value	Statistics	p-value	Obs
Null: unit root (common unit root is assumed)					
LLC t^*	-1.697	0.045	-1.540	0.062	1040
Null: unit root (individual unit root is assumed)					
IPS w-statistic	-1.683	0.046	-1.290	0.099	1040
CADF Z(t-bar)	-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
Null: no unit root (common unit root is assumed)					
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•••	-	5.192•••		
NH^{**} (nonparametric adjustment of LRV, lag=1)	2.760•••	-	2.730•••		

	Lag =2				
	US numeraire		EU numeraire		
Method	Statistics	p-value	Statistics	p-value	Obs
Null: unit root (common unit root is assumed)					
LLC t^*	-1.519	0.064	-1.233	0.109	1032
Null: unit root (individual unit root is assumed)					
IPS w-statistic	-1.761	0.039	-1.247	0.106	1032
CADF Z(t-bar)	-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
Null: no unit root (common unit root is assumed)					
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•••	-	5.192•••		
NH^{**} (nonparametric adjustment of LRV, lag=2)	1.924•••	-	1.898•••		

	Lag =3				
	US numeraire		EU numeraire		
Method	Statistics	p-value	Statistics	p-value	Obs
Null: unit root (common unit root is assumed)					
LLC t^*	-1.005	0.157	-0.826	0.2043	1024
Null: unit root (individual unit root is assumed)					
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
Null: no unit root (common unit root is assumed)					
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•••	-	5.192•••		
NH^{**} (nonparametric adjustment of LRV, lag=3)	1.499•••		1.477•••		

• 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 5: Pedroni and Westerlund test on panel cointegration*

Case	Pedroni tests				Westerlund test				
Case ($y_t \sim I(0)$)	$\tilde{Z}_\rho$	p-value	$Zt_{\hat{\rho}_{NT}}$	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 6: Linearity testing and model selection: SETAR

(1) Series	(2) BIC	(3) AIC	(4) HQ	(5) Hinich test (H-statistic)	(6) p	(7) Tsay test d=1	(8) Tsay test d=2	(9) Tsay test d=3	(10) d	Model
US numeraire										
DN	2	2	2	0.000	2	0.129	0.098	0.732		
CAN	1	1	1	0.999	4	0.612	0.703	0.347		
JPN	1	2	2	0.000	1	0.964	0.456	0.028	3	SETAR(1; 1, 3)
NW	2	2	2	0.000	2	0.475	0.097	0.406		
SD	2	4	2	0.050	2	0.388	0.227	0.119		
SZ	1	2	2	0.000	3	0.281	0.359	0.568		
UK	2	2	2	0.000	2	0.204	0.273	0.731		
EU	2	2	2	0.000	2	0.059	0.073	0.0564		
EU numeraire										
DN	2	5	4	0.085	3	0.034	0.238	0.376	1	SETAR(1; 3, 1)
CAN	1	1	1	0.000	1	0.095	0.846	0.920		
JPN	2	2	2	0.000	2	0.143	0.110	0.008	3	SETAR(1; 2, 3)
NW	2	3	2	0.000	2	0.023	0.024	0.036	1	SETAR(1; 2, 1)
SD	1	1	1	0.000	1	0.003	0.311	0.194	1	SETAR(1; 1, 1)
SZ	1	1	1	0.000	1	0.425	0.337	0.407		
UK	1	1	1	0.000	1	0.005	0.768	0.759	1	SETAR(1; 1, 1)
US	1	1	2	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) Hinich test statistics in p-values for third order autocorrelation and serial dependence from non gaussian errors; (6): 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; (7)-(9): Tsay test for threshold linearity in p-values using d={1, 2, 3} delay parameters; (10): chosen delay parameter d; (12) selected SETAR(k;p,d), $k = \{1, \ldots, N\}$ regimes. Software used: RATS

Table 7: Hansen' threshold effect test (bootstrapped p-values)

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

Software used: RATS

Table 8: Linearity testing and model selection: STAR

(1) Series	(2) BIC	(3) AIC	(4) HQ	(5) Hinich test H-statistic*	(6) p selected	(7) s_t	(8) F_L F-value	(9) F_4 F-value	(10) F_3 F-value	(11) F_2 F-value	(12) Model
US numeraire											
DN	2	2	2	0.000	2	y_{t-1}	$4.6\ e^{-27}$	$7.2\ e^{-1}$	$9.1\ e^{-2}$	$9.1e^{-30}$	LSTR1
CAN	1	1	1	0.999	4	y_{t-1}	$1.8\ e^{-24}$	$9.2\ e^{-1}$	$6.8\ e^{-1}$	$4.8\ e^{-28}$	LSTR1
JPN	1	2	2	0.000	1	-	-	-	-	-	Linear
NW	2	2	2	0.000	2	t	$1.3\ e^{-2}$	$8.4\ e^{-2}$	$7.4\ e^{-1}$	$4.4\ e^{-3}$	LSTR1
SD	2	4	2	0.050	4	y_{t-1}	$1.5\ e^{-28}$	$1.7\ e^{-1}$	$5.6\ e^{-1}$	$1.7\ e^{-31}$	LSTR1
SZ	1	2	2	0.000	2	y_{t-1}	$2.1\ e^{-23}$	$3.8\ e^{-1}$	$6.1\ e^{-2}$	$1.9\ e^{-25}$	LSTR1
UK	2	2	2	0.000	2	t	$1.1\ e^{-2}$	$5.9\ e^{-2}$	$7.0\ e^{-2}$	$7.6\ e^{-2}$	LSTR1
EU	2	2	2	0.000	2	y_{t-1}	$5.5\ e^{-27}$	$2.3\ e^{-1}$	$9.9\ e^{-2}$	$3.3\ e^{-29}$	LSTR1
EU numeraire											
DN	2	5	4	0.085	3	-	-	-	-	-	Linear
CAN	1	1	1	0.000	1	t	$3.2\ e^{-2}$	$4.0\ e^{-2}$	$2.3\ e^{-1}$	$1.1\ e^{-1}$	LSTR1
JPN	2	2	2	0.000	2	t	$3.4\ e^{-3}$	$5.6\ e^{-2}$	$1.4\ e^{-1}$	$8.3\ e^{3}$	LSTR1
NW	2	3	2	0.000	2	t	$3.1\ e^{2}$	$4.9\ e^{-1}$	$4.6\ e^{-1}$	$4.1\ e^{-3}$	LSTR1
SD	1	1	1	0.000	1	t	$2.1\ e^{-2}$	$3.6\ e^{-2}$	$5.0\ e^{-1}$	$3.2\ e^{-2}$	LSTR1
SZ	1	1	1	0.000	1	t	$3.9\ e^{-3}$	$3.9\ e^{-2}$	$5.9\ e^{-2}$	$2.9\ e^{-2}$	LSTR1
UK	1	1	1	0.000	1	y_{t-2}	$2.5\ e^{-4}$	$1.4\ e^{-1}$	$3.2\ e^{-5}$	$6.0\ e^{-1}$	LSTR2
US	-	-	-	-	-	-	-	-	-	-	-

Legend. (1): Countries; (2): Swartz-Bayesian information criterion; (3): Aikake information criterion; (4): Hannan-Quinn information criterion; (5) Hinich test statistics (p-values) for third order autocorrelation and serial dependence from non gaussian errors; (6): 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; (7): transition variable; (8): Saikkonen-Lukkonen-Teräsvirta linearity test by statistics LM_2 (equation 13 on page 12) or LM_3 (equation 16 on page 12); (9)-(11): results for Teräsvirta rule for the choice of the from of transition variable (see sequence of nested hyphotheses 17 on page 12); (12) Selected model for transition function. Software used: JMulTi 4

Table 9: STAR models estimates

(1) Series	(2) ϕ_0	(3) ϕ_1	(4) ϕ_2	(5) θ_0	(6) θ_1	(7) θ_2	(8) γ	(9) c_1	(10) c_2	(11) Restr	(12) $\bar{R}^2$	(13) J-B	(14) ARCH-LM(p)	(15) σ^2_ϵ	(16) SD_ϵ
US numeraire															
	3.978	-	-0.639	2.679	-	0.639	0.567	-1.997	-			0.922	5.711		
DN	(0.402)	-	(0.187)	(0.300)	-	(0.187)	(0.000)	(0.026)	-	$y^{DN/\$}_{t-2}$	0.982	-	-	0.001	0.023
	[0.000]	-	[0.001]	[0.000]	-	[0.001]	[0.000]	[0.000]	-			[0.631]	[0.004]		
	-1.158	-	-0.436	1.442	-	0.436	0.507	-0.360	-			371.000	0.195		
CAN	(0.000)	-	(0.174)	(0.000)	-	(0.174)	(0.000)	(0.030)	-	$y^{CAN/\$}_{t-2}$	0.979			0.000	0.020
	[0.000]	-	[0.014]	[0.000]	-	[0.014]	[0.000]	[0.000]	-			[0.000]	[0.094]		
	-0.596	1.216	-0.510	0.596	-1.216	0.510	0.991	221.762	-			4.525	30.505		
NW	(0.071)	(0.143)	(0.144)	(0.071)	(0.143)	(0.144)	(0.063)	(6.166)	-	*const*, $y^{NW/\$}_{t-1}$, $y^{NW/\$}_{t-2}$	0.912			0.002	0.042
	[0.000]	[0.000]	[0.006]	[0.000]	[0.000]	[0.006]	[0.000]	[0.000]	-			[0.104]	[0.000]		
	-11.794	-	-1.525	11.794	-	1.525	0.105	-3.435	-			0.160	0.167		
SD	(3.195)	-	(0.400)	(3.195)	-	(0.400)	(0.020)	(0.751)	-	*constant*, $y^{SD/\$}_{t-2}$	0.957			0.001	0.026
	[0.000]	-	[0.000]	[0.000]	-	[0.000]	[0.000]	[0.000]	-			[0.923]	[0.955]		
	-1.127	-	-0.656	1.220	-	0.656	0.615	-0.388	-			4.419	1.078		
SZ	(0.272)	-	(0.303)	(0.330)	-	(0.303)	(0.162)	(0.052)	-	$y^{SZ/\$}_{t-2}$	0.960			0.001	0.023
	[0.000]	-	[0.032]	[0.000]	-	[0.032]	[0.000]	[0.000]	-			[0.110]	[0.343]		
	0.011	1.327	-0.350	0.490	-1.327	0.350	53.692	123.025	-			7.113	3.205		
UK	(0.011)	(0.085)	(0.086)	(0.013)	(0.085)	(0.086)	(30.565)	(0.565)	-	$y^{£/\$}_{t-1}$, $y^{£/\$}_{t-2}$	0.960			0.000	0.020
	[0.308]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.081]	[0.000]	-			[0.029]	[0.044]		
	-0.811	-	-0.745	1.544	-	0.745	0.556	-0.027	-			4.400	12.057		
EU	(0.000)	-	(0.228)	(0.000)	-	(0.228)	(0.000)	(0.000)	-	$y^{e/\$}_{t-2}$	0.979			0.001	0.024
	[0.023]	-	[0.014]	[0.008]	-	[0.014]	[0.105]	[0.528]	-			[0.111]	[0.149]		
EU numeraire															
	0.067	1.135	-	0.067	-1.135	-	14.865	135.977	-			0.474	5.593		
CAN	(0.016)	(0.039)	-	(0.016)	(0.039)	-	(3.794)	(1.574)	-	*constant*, $y^{CAN/e}_{t-1}$	0.800			0.001	0.030
	[0.001]	[0.000]	-	[0.001]	[0.000]	-	[0.001]	[0.000]	-			[0.789]	[0.693]		
	-0.656	3.202	-2.335	0.953	-3.202	3.395	0.189	132.43	-			6.526	14.715		
JPN	(0.607)	(0.000)	(0.000)	(1.227)	(0.000)	(0.000)	(0.079)	(0.000)	-	$y^{JPN/e}_{t-1}$	0.985			0.001	0.027
	[0.282]	[0.000]	[0.007]	[0.438]	[0.000]	[0.002]	[0.018]	[0.041]	-			[0.038]	[0.065]		
	-15.142	-	-6.470	15.142	-	6.470	19.286	86.444	-			12.905	118.97		
NW	(6.473)	-	(3.128)	(6.473)	-	(3.128)	(0.000)	(3.643)	-	$y^{NW/e}_{t-1}$	-			1.656	1.287
	[0.029]	-	[0.041]	[0.029]	-	[0.041]	[0.261]	[0.000]	-			[0.002]	[0.000]		
	0.055	1.026	-	-2.428	-1.026	-	11.233	123.205	-			45.300	19.380		
SD	(0.052)	(0.024)	-	(0.500)	(0.024)	-	(3.396)	(1.702)	-	$y^{SD/e}_{t-1}$	0.974			0.000	0.013
	[0.288]	[0.000]	-	[0.000]	[0.000]	-	[0.001]	[0.000]	-			[0.000]	[0.013]		
	0.015	1.041	-	-0.015	-1.041	-	6.849	147.861	-			172.224	9.957		
SZ	(0.007)	(0.017)	-	(0.007)	(0.017)	-	(1.567)	(4.099)	-	*constant*, $y^{SZ/e}_{t-1}$	0.967			0.000	0.010
	[0.037]	[0.000]	-	[0.037]	[0.000]	-	[0.000]	[0.000]	-			[0.000]	[0.268]		
	-0.019	1.287	-0.245	0.069	-1.424	1.299	6.585	-0.141	-0.541	-		48.886	10.177		
UK	(0.007)	(0.090)	(0.095)	(0.016)	(0.395)	(0.363)	(4.357)	(0.004)	0.006		0.983			0.000	0.015
	[0.011]	[0.000]	[0.011]	[0.000]	[0.000]	[0.001]	[0.133]	[0.000]	0.000			[0.000]	[0.000]		

Legend. (1): Countries; (2)-(4): parameters $\boldsymbol{\phi}$ of linear AR-part; (5)-(7): parameters $\boldsymbol{\theta}$ of nonlinear AR-part; (8)-(10): parameters of transition function G(·); (11): parameters under restriction $\boldsymbol{\theta} = \boldsymbol{\phi}$; (12) Adjusted R^2; (13) Jarque-Brera statistic; (14) Engle's LM test for no ARCH effects for p lags; (15) variance of ϵ; (16) standard deviation of ϵ. Line 1: parameters' estimates; Line 2 (brakets): standard deviations; Line 3 (square brakets): p-values. Software used: JMulTi 4

Table 10: Diagnostic tests for estimated STAR models

	Serial Correlation								No remaining nonlinearity			Parameter Constancy		
	F-values								F-values	DF1	DF2	H_1	H_2	H_3
Lag	1	2	3	4	5	6	7	8						
U.S. numeraire														
DN	0.658	0.871	0.582	0.510	0.747	0.687	0.721	0.762	0.322	6	118	2.636	1.545	1.883
CAN	0.312	0.293	0.229	1.661	1.291	1.282	1.389	1.309	0.801	12	110	1.704	1.186	1.420
NW	0.295	0.644	1.846	1.718	1.622	1.797	1.559	1.675	0.165	6	118	19.031	10.739	10.408
SD	1.639	2.373	1.775	1.708	2.437	2.075	1.864	1.650	0.782	12	110	2.374	1.644	1.717
SZ	0.174	0.319	0.412	0.400	0.340	0.361	0.605	0.570	0.408	6	118	2.418	1.391	1.711
UK*	0.276	1.135	1.103	1.491	1.367	1.314	1.251	1.172	0.698	6	118	1.835	1.801	NaN
EU	0.669	0.769	0.510	0.417	0.836	0.761	0.765	0.760	0.420	6	118	1.474	0.777	0.614
E.U. numeraire														
CAN	5.094	2.742	1.563	1.013	0.775	0.898	0.875	1.179	0.032	3	123	32.968	18.208	12.649
JPN	0.418	0.737	0.615	1.029	0.876	0.791	0.585	0.457	0.001	6	118	0.676	0.005	0.005
NW	2.997	3.541	2.500	1.965	1.576	1.928	1.935	1.744	$3.82\ e^{-34}$	6	120	72,934	36,508	29,320
SD	3.090	1.658	3.886	3.008	2.586	2.316	1.974	1.852	0.000	3	123	0.030	0.024	0.004
SZ	1.547	0.749	0.650	0.642	0.496	0.414	0.759	0.676	0.118	3	123	0.082	0.104	0.658
UK	0.344	0.174	0.142	0.096	0.118	0.326	0.575	0.537	0.962	3	118	1.272	1.333	1.753
	P-values								P-values	DF1	DF2	H_1	H_2	H_3
Lag	1	2	3	4	5	6	7	8						
U.S. numeraire														
DN	0.419	0.421	0.628	0.728	0.590	0.660	0.654	0.637	0.924	"	"	0.037	0.150	0.044
CAN	0.578	0.747	0.876	0.164	0.273	0.272	0.217	0.247	0.648	"	"	0.154	0.314	0.168
NW	0.588	0.527	0.142	0.151	0.160	0.106	0.155	0.113	0.986	"	"	0.000	0.000	0.000
SD	0.203	0.098	0.156	0.153	0.034	0.062	0.082	0.119	0.667	"	"	0.056	0.120	0.073
SZ	0.677	0.728	0.745	0.809	0.888	0.902	0.751	0.119	0.873	"	"	0.052	0.208	0.074
UK*	0.601	0.325	0.351	0.209	0.242	0.257	0.282	0.323	0.651	"	"	0.098	0.057	NaN
EU	0.545	0.466	0.676	0.796	0.527	0.603	0.618	0.639	0.865	"	"	0.233	0.543	0.719
E.U. numeraire														
CAN	0.026	0.068	0.202	0.404	0.569	0.499	0.529	0.318	0.992	"	"	0.000	0.000	0.000
JPN	0.519	0.481	0.607	0.395	0.500	0.579	0.767	0.884	1.000	"	"	0.001	NaN	NaN
NW	0.858	0.032	0.063	0.104	0.172	0.082	0.071	0.096	1.000	"	"	0.000	0.000	0.000
SD	0.081	0.195	0.011	0.021	0.030	0.038	0.065	0.075	1.000	"	"	0.000	0.000	0.000
SZ	0.216	0.475	0.584	0.633	0.779	0.868	0.623	0.712	0.950	"	"	0.000	0.000	0.000
UK	0.559	0.840	0.935	0.983	0.988	0.922	0.775	0.827	0.413	"	"	0.276	0.211	0.042

*Statistic for No remaining nonlinearity refers to the case $s_t = y_{t-1}$. Software used: JMulTi 4

Figure 1: PPP series (in logs)

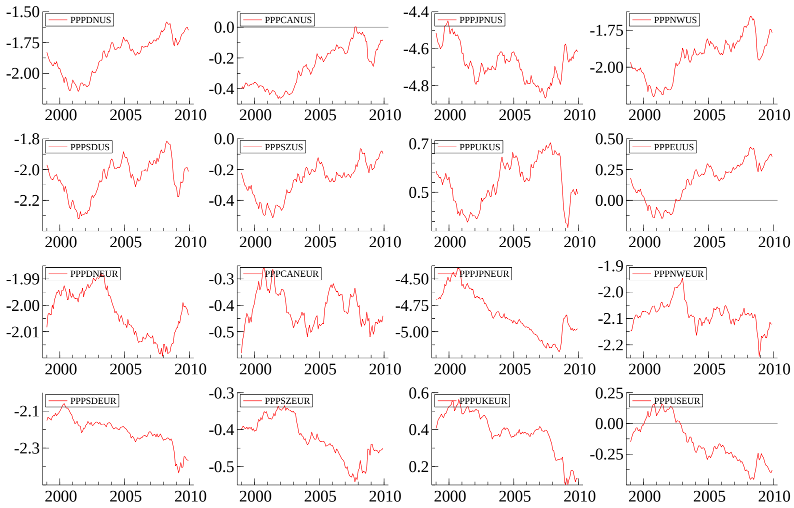
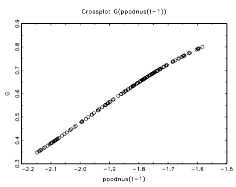
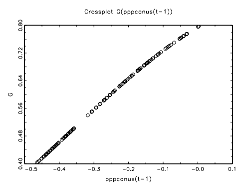


Figure 2: Plots of transition functions for estimated STAR models

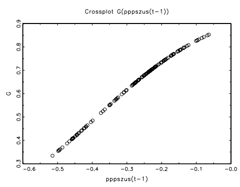
(a) *PPPDNUS*



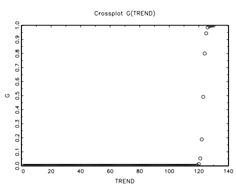
(b) *PPPCANUS*



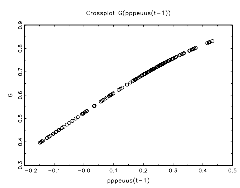
(c) *PPPSZUS*



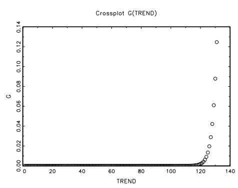
(d) *PPPUKUS*



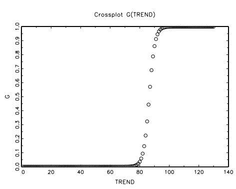
(e) *PPPEUUS*



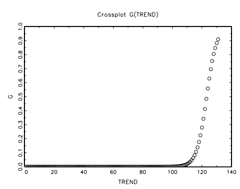
(f) *PPPCANEUR*



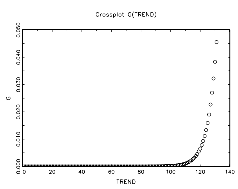
(g) *PPPNWEUR*



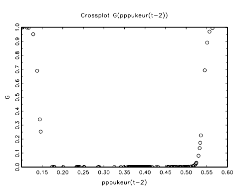
(h) *PPPSDEUR*



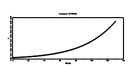
(i) *PPPSZEUR*



(j) *PPPUKEUR*



(k) *PPPNWUS*



(l) *PPPSDUS*

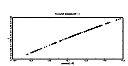
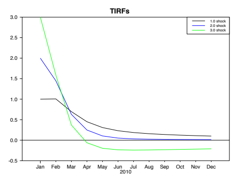
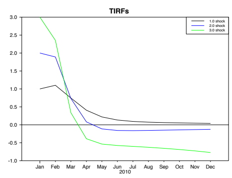


Figure 3: Impulse Response Functions

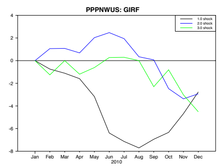
(a) *DN/\$: TIRF*



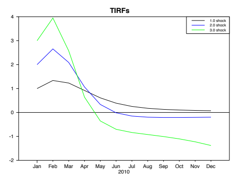
(b) *CAN/\$: TIRF*



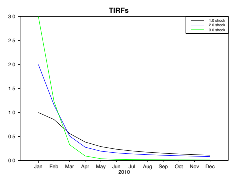
(c) *NW/\$: GIRF*



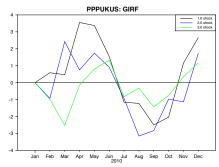
(d) *SD/\$: TIRF*



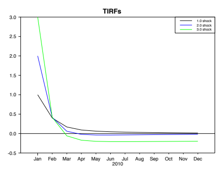
(e) *SZ/\$: TIRF*



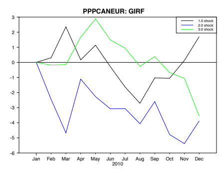
(f) *£/\$: GIRF*



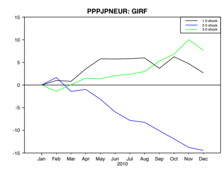
(g) *€/\$: TIRF*



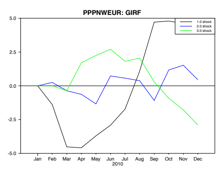
(h) *CAN/€: GIRF*



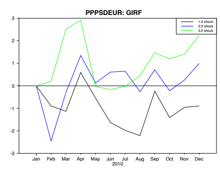
(i) *JPN/€: GIRF*



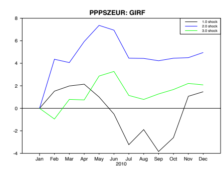
(j) *NW/€: GIRF*



(k) *SD/€: GIRF*



(l) *SZ/€: GIRF*



(m) *£/€: GIRF*