CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Robust Estimation of Real Exchange Rate Process Half-life

Michele Bergamelli

CEA@Cass Working Paper Series

WP–CEA–01-2015

Robust Estimation of Real Exchange Rate Process Half-life*

Michele Bergamelli†

THIS VERSION: 19th January 2015

Abstract

In this paper, we show that the data generating process of the real exchange rate is likely to include outliers that, if not accounted for, lead to unreliable half-lives estimates. In order to obtain robust estimates of the half-life, we propose to identify outlying observations by means of a dummy saturation type algorithm designed for ARMA processes which enables to detect additional and innovative outliers as well as level shifts in the real exchange rate process. In a Monte Carlo study, we show that the proposed estimation procedure delivers the correct retention rates of the dummy variables included to capture the outliers. An empirical application involving US dollar real exchange rates shows that the estimated half-lives are consistently shorter when outlying observations are correctly modelled, thus shedding some light on the PPP puzzle.

Keywords: Exchange Rates, Purchasing Power Parity, Half-life, Outliers, Dummy Saturation, Robust Estimation.
JEL Classification: C13, C51, F31.

*I wish to thank participants in the 15th OxMetrics' User Conference" (Cass Business School, 4-5 September 2014) in particular Neil Ericsson, Sir David Hendry, and Bent Nielsen for very useful comments. The usual disclaimer applies. Financial support from the "PhD Scholarship in Memory of Ana Timberlake" and the Centre for Econometric Analysis is gratefully acknowledged.

†Centre for Econometric Analysis, Cass Business School, City University London, 106 Bunhill Row EC1Y 8TZ London (UK). E-mail: Michele.Bergamelli.2@cass.city.ac.uk.

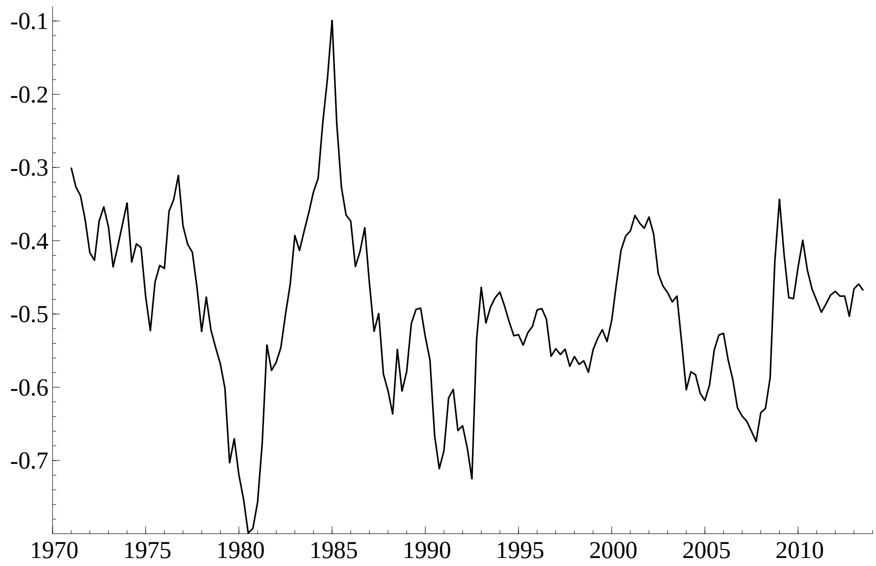
1 Introduction

The purchasing power parity (PPP) condition holds if the real exchange rate process is stationary and it reverts to its mean in about one to two years, the necessary time to absorb financial and monetary shocks which are thought to affect the exchange rate. Over the last thirty years, the validity of the PPP condition has been subject to several empirical tests by means of alternative econometric techniques. Sarno and Taylor (2002) provide an extensive literature review from the early seventies to recent years (see also James et al., 2012).

The main empirical evidence is that the PPP is valid in the long-run while substantial deviations are usually observed in the short-run. In other terms, the process followed by the real exchange rate is found to be stationary but with high persistence. To measure the degree of persistence of the real exchange rate process, the concept of *half-life* is usually employed (see Mark, 2001; Rossi, 2005; Chortareas and Kapetanios, 2013). This is defined as the period of time necessary for the real exchange rate process to dissipate by half a unitary shock (or its cumulative effect) and it is commonly used as a measure to quantify to which extent the purchasing power parity condition holds. Empirical findings seem to confirm that the half-life of the real exchange rate ranges from about three to five years (see Frankel and Rose, 1996, for an extended study involving over a hundred of countries), giving rise to the so called "PPP puzzle" as defined in Rogoff (1996). Taylor (2001) explores possible sources of bias which might explain the puzzling half-lives measures obtained in the literature. The author suggests that the puzzle might be mitigated by taking into account data aggregation issues and by allowing for non-linear dynamics in the real exchange rate process. Notwithstanding, together with the correct specification of the statistical framework used to model the real exchange rate process, correct inference for the model parameters is crucial for obtaining reliable half-life measures.

In this paper, we argue that the data generating process of the real exchange rate is likely to include outliers that, if not accounted for, distort the estimated half-lives since they alter the autocorrelation structure of the observed time series (see for instance Tsay, 1986). From a visual inspection of the USD/GBP real exchange rate time series in Figure 1,

Figure 1: USD/GBP Real Exchange Rate (log).



it is evident that our conjecture is not completely at odds with empirical evidence. This paper contributes to the literature on the PPP puzzle in three directions.

First, we propose a new framework to model the real exchange rate process. In particular, we allow the real exchange rate to follow an ARMA process contaminated with additive and innovative outliers as well as level shifts. Further, we devise a fast and accurate procedure to estimate the half-life of the ARMA process in this framework.

Second, in order to estimate the outlier-contaminated model, we consider an extension of the Dummy Saturation approach introduced by Hendry (1999) and Hendry et al. (2008) which considers saturating in turn with additive outliers, innovative outliers and level shifts in a maximum likelihood framework. The performance of the procedure, in terms of retention rate of the insignificant outliers, is explored using a Monte Carlo simulation.

Finally, in order to show the severity of the effect of unaccounted outliers on the estimation of the half-life, we carry out empirical application involving US dollar real exchange rates for a group of developed countries. The results we obtain are consistent with our claim and we find that half-life estimates can indeed change dramatically when outlying observations are accounted for. In particular, we find that estimated half-lives are considerably shorter for the exchange rates which exhibit the most puzzling behaviour than when computing the half-life disregarding the effect of outliers.

The reminder of the paper is organised as follows: Section 2 briefly reviews the existing literature focusing on the existing approaches to model the real exchange rate process and computing the half-life; Section 3 presents our model and describes the estimation approach; Section 4 reports the empirical application and Section 5 concludes.

2 Review of the Literature

The real exchange rate process is usually regarded as a measure of deviation from the PPP relationship. In particular, if we denote with S_t the spot nominal exchange rate while with $P_{i,t}^h$ and $P_{i,t}^f$ the domestic and foreign price of the i^{th} good respectively, the PPP condition can be formulated as a generalisation of the law of one price to a basket of goods, such that

$$S_t = \frac{\sum_{i=1}^N w_i P_{i,t}^h}{\sum_{i=1}^N w_i P_{i,t}^f} = \frac{\bar{P}_t^h}{\bar{P}_t^f} \qquad t = 1, \dots, T \tag{1}$$

where N denotes the number of goods in the basket and w_i, such that $\sum_{i=1}^N w_i = 1$, denotes the weight assigned to the i^{th} good. The logarithmic form of the PPP is then given by

$$s_t = \bar{p}_t^h - \bar{p}_t^f, \tag{2}$$

and from (2), the real exchange rate is defined as

$$q_t \equiv s_t - \bar{p}_t^h + \bar{p}_t^f \tag{3}$$

which should then be equal to 0 for all $t = 1, \dots, T$ if PPP holds exactly. In general, however, $\{q_t\}$ follows a stochastic process which properties can be used to study the deviations from PPP. Empirically, as reported in Rogoff (1996), we usually proxy the price of the goods in one country by the corresponding CPI (of traded goods), where $CPI_t^h = \bar{P}_t^h / \bar{P}_0^h$ and $CPI_t^f = \bar{P}_t^f / \bar{P}_0^f$ with the subscript "0" denoting some chosen base year [1]. In practise, what we are actually computing when using the CPI (or every other

[1] We are assuming to be in the ideal situation where the price index is computed from the same basket of goods in the two countries. In practise, this is almost never satisfied and together with the way in which

price index) is then

$$s_t - (\bar{p}_t^h - \bar{p}_0^h) + (\bar{p}_t^f - \bar{p}_0^f) = q_t + \bar{p}_0^h - \bar{p}_0^f = q_t + const \tag{4}$$

which implies that the real exchange rate obtained in this way has an expectation different from zero by construction.

In the economic literature, we can find two different ways to model the process followed by $\{q_t\}$. The vast majority of the empirical studies models the real exchange rate process according to linear dynamic models assuming either an AR(1) (Abuaf and Jorion, 1990) or an AR(p) (Rossi, 2005; Chortareas and Kapetanios, 2013) structure. As noted in Chortareas and Kapetanios (2013), assuming a simplistic AR(1) is however suboptimal whenever the dynamics followed by the real exchange rate process can be captured by higher order models although the computation of the half-life complicates. Further, we think that the use of ARMA processes has been largely overlooked in favour of simple AR processes. The reason is not completely clear thus one explanation might lie in the fact that the computation of the half-life is computationally more cumbersome. To the best of our knowledge, the only exceptions are Diebold et al. (1991) and Cheung and Lai (2000) where the authors allow $\{q_t\}$ to follow a long-memory process of the ARFIMA class and a stationary ARMA process, respectively. All these studies, with the exception of Chortareas and Kapetanios (2013) who adopt a different half-life measure, find that PPP holds in the long-run whereas substantial deviations occurs in the short-run. Another stream of literature, building on the "bands of inaction" argument raised in Taylor (2001), considers instead non-linear dynamic models like self-exciting threshold autoregression (SETAR) and smooth transition autoregression (STAR). According to Taylor (2001), "bands of inaction" due to transaction costs, where the real exchange rate behaves like a random walk, would bias upward the autoregressive coefficient giving rise to half-lives bigger than those that would be observed if measuring the half-life outside these bands.

As mentioned in the introduction, testing the PPP condition involves the computation of a measure of persistence of the process followed by $\{q_t\}$. Typically, this is done by

the weights are formed is one of the main source of the index problem.

computing the *half-life* of the process[2]. According to Mark (2001) and Rossi (2005), the half-life is formally defined as the smallest h such that $\mathbb{E}(q_{t+h} - q_0 | q_{t-s} - q_0, s \leq 0) \leq \frac{1}{2}(q_t - q_0)$, i.e. the time necessary for q_t to revert back at half its initial post shock value. Denoting with $\psi(t)$ for $t \geq 0$ the impulse response function (IRF) of $\{q_t\}$ and considering an initial unitary shock to give $\psi(0) = 1$, the above definition of half-life corresponds to the instant h such that

$$\psi(h) = \frac{1}{2}. \tag{5}$$

More in detail, Mark (2001) provides a formula to compute half-lives for stationary AR(p) processes while Rossi (2005) derives an asymptotic approximation and confidence intervals for the half-life of AR(p) processes when one root of the characteristic polynomial is close to unity. The results are consistent with the puzzling behaviour of the PPP condition.

Recently, Chortareas and Kapetanios (2013) propose an alternative definition of half-life based on the decline of the cumulative effect of a shock rather than its point value. In particular, half-life is defined as the instant h such that half of this cumulative effect has dissipated. In terms of the IRF $\psi(t)$, it can be expressed as

$$\int_0^h f[\psi(i)]\mathrm{d}i = \int_h^\infty f[\psi(i)]\mathrm{d}i \tag{6}$$

with $f[\psi(i)] = |\psi(i)|$, to accommodate negative values of the IRF, or $f[\psi(i)] = \psi(i)^2$ to take also into account the possibility of long-memory processes. The authors apply this new definition of half-life using US dollar exchange rates for a set of developed countries for the period 1957:Q1-1998:Q4 and finding a reduction in the half-lives with respect to previous studies. Notwithstanding, in this paper, we adopt the standard definition of half-life as in (5), both for ease of comparison with the previous results and for the fact that the PPP puzzle has been built around this definition. In general, however, it is clear that independently of which of the two definitions we decide to adopt, the computation of the half-life depends on the estimated parameters of the underlying model and thus, outliers robust methods are equally advised.

[2] Alternative measures are given in Andrews and Chen (1994).

In what follows, we extend the current literature not restricting $\{q_t\}$ to follow only AR(p) processes but considering more general stationary ARMA(p,q) process contaminated by outliers and level shifts (see Section 3). Further, we describe a fast method to compute exact half-lives for ARMA(p,q) processes according to (5) (see Section 3.3).

3 Real Exchange Rate Process: an Outliers-based Approach

Assume that the real exchange rate process $\{q_t\}$ is well represented by a stationary ARMA(p,q) process contaminated by outliers and level shifts. In particular, let the process followed by $\{q_t\}$ be described by

$$q_t = q_0 + \sum_{i=1}^{k} \delta_i V_i(L) \mathbf{1}(t = T_i) + v_t \tag{7}$$

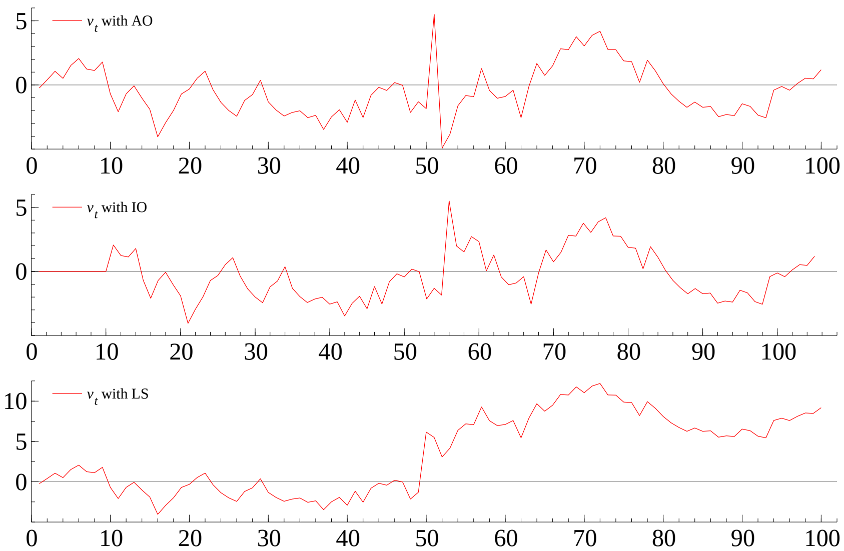
$$\phi(L) v_t = \theta(L) \varepsilon_t \qquad t = 1, \ldots, T \tag{8}$$

where q_0 can be thought as the long-run value of the real exchange rate, k denotes the number of outlying events, δ_i is the outlier or level shift size, $V_i(L)$ (with L denoting the lag operator) defines the outlier type, $\mathbf{1}(t = T_i)$ is an impulse indicator assuming value 1 for $t = T_i$ and 0 otherwise, $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$ are lag polynomials with roots outside the unit circle, and $\varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$. The specification of $V_i(L)$ allows to characterise different kinds of outlying observations (see Tsay, 1988). Three specifications are particularly relevant to our analysis:

$V_i(L) = 1$	Additive Outlier (AO)
$V_i(L) = \phi^{-1}(L)\theta(L)$	Innovative Outlier (IO)
$V_i(L) = (1 - L)^{-1}$	Level Shift (LS) (Since $(1 - L)^{-1}\mathbf{1}(t = T_i) = \mathbf{1}(t \geq T_i)$).

The difference between AO and IO concerns with the way they affect the time series. To better understand their impact, assume to be able to distinguish k^A AOs, k^I IOs and k^L

Figure 2: Plot of AO, IO and LS at $T_1 = 50$ for $T = 100$ and $v_t = 0.8v_{t-1} + \varepsilon_t$.



LSs, and rewrite (7)-(8) in terms of ε_t to give

$$q_t = q_0 + \sum_{i=1}^{k^A} \delta_i^A \mathbf{1}(t = T_i) + \frac{\theta(L)}{\phi(L)} \left(\sum_{l=1}^{k^I} \delta_l^I \mathbf{1}(t = T_l) + \varepsilon_t \right) + \sum_{j=1}^{k^L} \delta_j^L \frac{1}{(1-L)} \mathbf{1}(t = T_k). \quad (9)$$

From equation (9), we can see that an IO affects directly the innovation process and thus it propagates to future observations through the multiplier $\phi(L)^{-1}\theta(L)$ while an AO just affect the single observation at the time of the shock. Finally, a LS produces a shift in the long-run value of the real exchange rate from the time of the shock onwards. Figure 2 gives an idea of how the impact of the different outliers looks like using a simulated process.

As largely documented in the statistical and econometric literature (see Chang et al., 1988; Tsay, 1988; Chen and Liu, 1993; Sánchez and Peña, 2003; Cavaliere and Georgiev, 2009, and references therein), outlying observations distort the autocorrelation structure of the time series under exam. This has an impact on the identification of the appropriate ARMA order and most importantly it causes biases in the estimated ARMA coefficients. The implications for forecasting follow from the just mentioned problems.

Two major consequences for the PPP relationship can be identified. First, strongly upward-biased autoregressive coefficients might lead to the conclusions that $\{q_t\}$ is nonstationary and thus rejecting the existence of PPP in a first place[3]. Second, biased ARMA coefficients such that $\{q_t\} \sim I(0)$ can still lead to biased half-life measures either upwards or downwards, being the half-life computed from these coefficients (see Section 3.3). As noted in Tsay (1986), the exact effect (bias direction) of multiple outliers is difficult to quantify and it depends also on the interaction between the outliers themselves (e.g. two outliers with magnitudes of opposite sign can cancel out) and their location in the sample. Sparse results are only available in the literature for a maximum of two outliers and low order ARMA processes which are of scarce interest for our purposes.

The conclusion is that to which extent the bias induced by unaccounted outliers exacerbates or solves the PPP puzzle, it is worth to be investigated.

3.1 Estimation Strategy and Outliers Detection

In this section, we deal with the estimation of equation (9). Using a compact matrix notation, the real exchange rate process contaminated with outliers is given by

$$q_t = q_0 + \mathbf{x}_t^\top \boldsymbol{\delta}^A + \mathbf{y}_t^\top \boldsymbol{\delta}^L + \phi^{-1}(L)\theta(L)(\varepsilon_t + \mathbf{z}_t^\top \boldsymbol{\delta}^I) \qquad t = 1, \ldots, T \tag{10}$$

$$= q_0 + \mathbf{x}_t^\top \boldsymbol{\delta}^A + \mathbf{y}_t^\top \boldsymbol{\delta}^L + \tilde{\mathbf{z}}_t^\top \boldsymbol{\delta}^I + v_t \qquad v_t = \phi^{-1}(L)\theta(L)\varepsilon_t, \tag{11}$$

where $\mathbf{x}_t$, $\mathbf{y}_t$ and $\mathbf{z}_t$ are vectors containing impulse or step dummies of dimension k^A, k^I and k^L, respectively and $\tilde{\mathbf{z}}_t$ denote the filtered IOs vector. Finally, we can write

$$\mathbf{q} = \mathbf{W}\boldsymbol{\delta} + \mathbf{v} \qquad \mathbf{v} = \phi^{-1}(L)\theta(L)\boldsymbol{\varepsilon}, \tag{12}$$

where $\mathbf{W}$ is a matrix of size $(1 + k^A + k^L + k^I) \times T$ mostly made up of 0-1 entries with the exception of the last k^I columns (those corresponding to the innovational outliers). Further, under the assumption that $\{\varepsilon_t\}$ is a sequence of IID normally distributed random variables, it follows that

$$\mathbf{v} \sim \mathcal{N}_T(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}) \tag{13}$$

[3]Consider for instance the low power of unit root tests when applied to a stationary processes with level shifts.

where $\mathbf{\Omega}$ is a symmetric $T \times T$ Toeplitz matrix containing the autocovariances of the ARMA process followed by $\{v_t/\sigma_\varepsilon\}$. The autocovariances are functions of the ARMA parameters ($\boldsymbol{\phi} = [\phi_1, \ldots, \phi_p]^\top$ and $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_q]^\top$) and need to be estimated along with the regression coefficients.

Given the form of (12), we are in the framework of a regression with ARMA errors where the exogenous regressors are the outlying observations. Typical estimation methods for such regressions are GLS (in particular see Galbraith and Zinde-Walsh, 1992) or ML based on (13). If the time-series parameters ($\boldsymbol{\phi}$, $\boldsymbol{\theta}$, σ_ε^2) are known and under normality, GLS and ML yield the same estimator of the regression coefficients, which take on the following form

$$\hat{\boldsymbol{\delta}}_{GLS} = \hat{\boldsymbol{\delta}}_{ML} = (\mathbf{W}^\top \mathbf{\Omega}^{-1} \mathbf{W})^{-1} \mathbf{W}^\top \mathbf{\Omega}^{-1} \mathbf{q}. \tag{14}$$

Assuming a single outlier at $T_1 < T$, the estimators of the impact of the different kind of outliers can be derived from (14) and are given by

$$\text{AO} : \hat{\delta}^A(T_1) = (\Omega^{-1})_{T_1,T_1} (\mathbf{\Omega}^{-1})_{T_1,\cdot} \mathbf{q} \tag{15}$$

$$\text{LS} : \hat{\delta}^L(T_1) = \left(\sum_{i=T_1}^{T} \sum_{j=T_1}^{T} (\Omega^{-1})_{i,j} \right)^{-1} \left(\sum_{i=T_1}^{T} (\mathbf{\Omega}^{-1})_{i,\cdot} \right) \mathbf{q} \tag{16}$$

$$\text{IO} : \hat{\delta}^I(T_1) = \phi(L)\theta^{-1}(L) q_{T_1} \tag{17}$$

where $\mathbf{\Omega}_{i,\cdot}$ denotes the i^{th} row of $\mathbf{\Omega}$. Note that the impact of an IO equals the residual at that time (Chang et al., 1988).

In general, however, $\boldsymbol{\phi}$, $\boldsymbol{\theta}$ and σ_ε^2 are unknown and hence, $\mathbf{\Omega}$ needs to be replaced by a consistent estimator $\hat{\mathbf{\Omega}}$. In this situation, (*iterated feasible*) GLS and ML provide different estimators of $\mathbf{\Omega}$, thus yielding different estimates of $\boldsymbol{\delta}$. Hereafter, we will focus exclusively on ML estimation which has the advantage to allow simultaneous estimation of both the ARMA parameters and the regression coefficients. In particular, the log-likelihood function of model (12) under the normality assumption (13) is given by

$$\ell(\boldsymbol{\delta}, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma_\varepsilon^2) = -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log\sigma_\varepsilon^2 - \frac{1}{2}\log|\mathbf{\Omega}| - \frac{1}{2\sigma_\varepsilon^2}(\mathbf{q} - \mathbf{W}\boldsymbol{\delta})^\top \mathbf{\Omega}^{-1} (\mathbf{q} - \mathbf{W}\boldsymbol{\delta}).$$

In our setting, the main challenge is that the regressors matrix ($\mathbf{W}$) is not known as well. Determining $\mathbf{W}$ amounts to the problem of selecting the outlying observations. In the next section, we describe in detail the procedure we propose to detect outliers in an ARMA modelling framework.

3.2 Dummy Saturation in the Presence of AOs, IOs and LSs

In order to locate outlying observations, we employ the Dummy Saturation principle proposed by Hendry (1999) and theoretically explored by Hendry et al. (2008) and Johansen and Nielsen (2009). The original contributions involve saturation with AOs (referred as IIS, *impulse indicator saturation*) while in a recent paper Doornik et al. (2013) consider saturation with LSs (referred as SIS, *step indicator saturation*).

In this paper, we consider a ML based procedure which looks for outliers in the real exchange rate process by saturating in turn with AOs, IOs and LSs. The steps of the procedure can be summarised as follows:

Step 1. Start by selecting the ARMA process that best fits the real exchange rate under the assumption of no outliers. To this extent, we estimate via ML processes of the form

$$\phi(L)(q_t - q_0) = \theta(L)\varepsilon_t, \tag{18}$$

increasing consecutively the order of the lag polynomials. Finally, we select the model according to the Akaike information criteiron (AIC) and denote the corresponding order with $(\tilde{p}, \tilde{q})$. The ARMA order is kept fixed until Step 3.

Step 2. This step involves the first search for outliers. We search sequentially for AOs, IOs and LSs and we store the selected outliers after each saturation. The significance level adopted in selecting each outlying observations is denoted with α. We can identify three sub-steps:

1. Saturate (18) with **AOs**[4]. Following Hendry et al. (2008)[5]:

[4]The choice of starting the saturation with AOs is purely casual as there is no difference in starting with either IOs or LSs in spite of AOs.

[5]Castle et al. (2012) and Bergamelli and Urga (2013) study by simulations the performance of the

(a) Add the first half of AOs, say $x_{j,t}$, $j = 1, \dots, \lfloor T/2 \rfloor$, and estimate by ML the following regression

$$q_t = q_0 + \sum_{j=1}^{\lfloor T/2 \rfloor} \delta_j^A x_{j,t} + \phi^{-1}(L)\theta(L)\varepsilon_t. \tag{19}$$

(b) Store all $\mathbf{x}_j$ such that $|\mathsf{t}_{\hat{\delta}_j^A}| > c_{\alpha/2}$. Denote the matrix of retained AOs with $\ddot{\mathbf{X}}_{(1)}$.

(c) Repeat by saturating with the second half of AOs, i.e. estimating (19) with $x_{j,t}$, $j = \lfloor T/2 \rfloor + 1, \dots, T$, and again define $\ddot{\mathbf{X}}_{(2)}$ the matrix of the outliers for which $|\mathsf{t}_{\hat{\delta}_j^A}| > c_{\alpha/2}$.

(d) Estimate (19) including only the AOs selected at the two previous stages and denote $\ddot{\mathbf{X}}$ the matrix with the statistically significant outliers.

2. Saturate (18) with **LS**s. Repeat all the procedure described in steps (a)-(d) in order to get $\ddot{\mathbf{Y}}$, the matrix containing the retained LSs.

3. Saturate (18) with **IO**s. Repeat all the procedure described in steps (a)-(d) in order to get $\ddot{\mathbf{Z}}$, the matrix containing the retained IOs.

Step 3. In the last step, we select the final model whose time series parameters are then used to obtain a robust half-life estimate. In practice, we start by estimating via ML the following regression with ARMA errors[6]

$$q_t = q_0 + \ddot{\mathbf{x}}_t^\top \boldsymbol{\delta}^A + \ddot{\mathbf{y}}_t^\top \boldsymbol{\delta}^L + \phi^{-1}(L)\theta(L)(\varepsilon_t + \ddot{\mathbf{z}}_t^\top \boldsymbol{\delta}^I) \tag{20}$$

and drop the insignificant outliers. Estimation of (20) is then iterated until the remaining outliers are all statistically significant. Once the selection of the outliers is terminated, we check whether the coefficients of the lag polynomials are all statistically significant as well and we modify the ARMA order accordingly.

Dummy Saturation when applied through the model selection algorithm *Autometrics* (see Doornik, 2009). However, here we resort on a more traditional split algorithm (see for instance Hendry et al., 2008) as estimation of ARMA processes is not currently available through *Autometrics*.

[6]The implicit assumption is that the number of outliers at this point is such that there are enough degrees of freedom. If this is not the case an easy solution is to strengthen the significance level used in the saturation.

The procedure outlined above extends the existent applications of the Dummy Saturation principle in two directions. First, it considers outliers detection in ARMA models which puts the Dummy Saturation in the context of maximum likelihood estimation. Second, it generalises the search of outlying observations considering also outliers of the innovative form. As noted above, innovative outliers have a slowly decaying effect in contrast with the instantaneous effect of additional outliers. This leads to the intuition that saturating with innovative outliers might help to capture more parsimoniously what otherwise would be captured by a series of additive outliers with decreasing coefficients.

As long as the steps of the procedure are structured, they match rather closely those of other procedures for outliers detection in time series model, in particular the widely used iterative approach as presented in Chen and Liu (1993). The main difference lies in the way the outliers are identified, i.e. saturation with dummies instead of the inclusion of the outliers one-by-one. In particular, we can identified three separate saturations: a first one involving AOs, a second one involving IOs and a third involving LSs. The retained outliers from each single saturation are then combined for the selection of the final model to take place. Additionally, note that the first and the last step are those where the ARMA order is determined.

Finally, when saturating a regression with dummies, an important aspect to keep under control is the retention rate of the outliers. We know from Hendry et al. (2008), Castle et al. (2012) and Doornik et al. (2013) that, given a significance level α and under normality, αT outliers are retained on average under the null of no outliers. Since in our case we are working with three separate saturations, we must set in each single saturation a significance level equal to approximately one third the desired retention rate before the final selection in Step 3. We also stress the fact that, though we present the procedure splitting the dummies into two halves, splitting the set of dummies in more than two parts does not alter the finite-sample properties of the procedure (see Table 1).

To verify that the outlined procedure delivers controlled retention rates, we carry out a small simulation study involving an ARMA$(1,1)$ process, under the null of no outliers.

Table 1: Retention Rates of Not Significant Outliers

$T = 100$			$T = 200$			$T = 300$		
AO	IO	LS	AO	IO	LS	AO	IO	LS
				$\alpha = 0.01$				
				$n = 2$				
1.242	6.472	0.901	1.215	6.488	0.983	1.092	6.651	0.989
				$n = 5$				
1.189	2.125	0.989	1.079	2.088	0.973	1.062	2.105	1.000
				$n = 10$				
1.116	1.530	0.993	1.096	1.493	0.972	1.045	1.479	1.004
				$n = 20$				
1.116	1.462	1.041	1.054	1.273	1.005	1.040	1.233	1.002
				$\alpha = 0.05$				
				$n = 2$				
5.260	15.880	4.655	5.223	16.085	4.818	5.103	16.265	4.899
				$n = 5$				
5.245	7.756	4.567	5.096	7.822	4.951	5.092	7.875	4.981
				$n = 10$				
5.206	6.346	4.894	5.188	6.235	4.965	5.038	6.271	5.010
				$n = 20$				
5.332	6.140	4.995	5.096	5.680	4.994	5.132	5.669	5.022

Notes: The reported values are the retention frequencies of the dummies computed across $M = 1,000$ simulations as $\frac{1}{T \times M} \sum_{m=1}^{M} \sum_{j=1}^{T} \mathbf{1}(|\mathrm{t}_{\hat{\delta}_j}| > c_{\alpha/2})$ where $\hat{\delta}_j$, $j = 1, \ldots, T$ denotes the estimated dummy coefficient associated with the j^{th} dummy and $T = \{100, 200, 300\}$ the sample size.

In particular, we simulate 1,000 paths from the following process

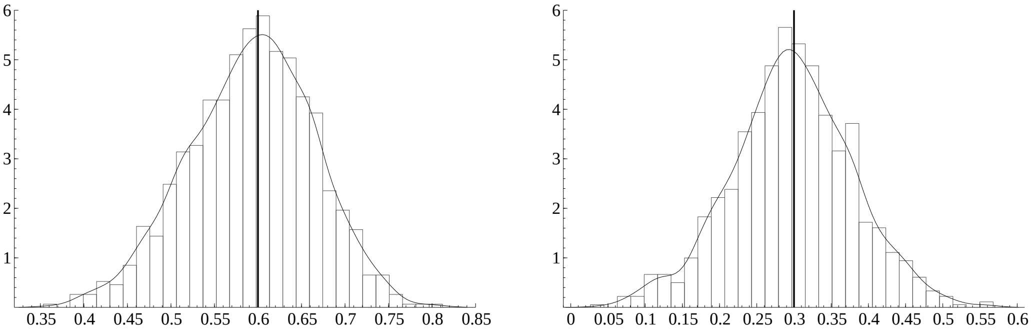
$$y_t = 0.6y_{t-1} + \eta_t + 0.3\eta_{t-1} \qquad t = 1, \ldots, T$$

for $T = \{100, 200, 300\}$ and $\eta_t \sim \mathcal{N}(0, 1)$ for all t. We compute the retention rates[7] of the three sets of dummies setting $\alpha = \{0.01, 0.05\}$. Table 1 reports the results using in each experiment a different number of splits of the dummies, i.e. $n = \{2, 5, 10, 20\}$. The results

[7] All the computations are carried out using OxMetrics 6.30. The ARFIMA (Doornik and Ooms, 2012) package has been used for estimation purposes.

Figure 3: Distribution of the ARMA Coefficients Estimates after Dummy Saturation

(a) $\phi = 0.6$ **(b)** $\theta = 0.3$



allow us to conclude that the retention rates under the null of no outliers are very close to the nominal level α with different block splits for AOs and LSs. For IOs, we observe instead deviations of the retention rates from α when the number of splits is small. In particular, the procedure retains too many IOs with the retention rates converging to the nominal level as the number of splits increases. In particular, the simulations show that the minimum number of splits that allows to reach the nominal level is around 20 which implies that the ideal number of dummies to be considered in each split is about $T/20$. In the empirical application, we will take into account this fact by setting the number of dummies splits to 20. Further, as commonly observed, the results show that the retention rates get closer to the nominal levels as the sample size T increases. As far as the time-series parameters are concerned, Figure 3 shows that the distributions of the ARMA coefficients estimates are correctly centred around the coefficients true values.

3.3 Half-life Computation for ARMA(p, q) Models

After controlling for outlying observations as described in the previous section, we use the robust ARMA parameters estimates in order to compute the half-life. Starting from (10), the outlier-free series is given by

$$\tilde{q}_t \equiv q_t - q_0 - \mathbf{x}_t^\top \boldsymbol{\delta}^A - \mathbf{y}_t^\top \boldsymbol{\delta}^L - \phi(L)^{-1}\theta(L)\mathbf{z}_t^\top \boldsymbol{\delta}^I = \phi^{-1}(L)\theta(L)\varepsilon_t = v_t. \tag{21}$$

Further, we can define $\psi(L) = \phi^{-1}(L)\theta(L)$ to give $\tilde{q}_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j}$, with $\sum_{j=0}^{+\infty} \psi_j^2 < \infty$ (under the assumption that the roots of $\phi(L)$ all lie outside the unit circle), such that $\lim_{j\to\infty} \psi_j = 0$. As already mentioned, the value assumed by the coefficients ψ_j as a function of time is regarded as the impulse response function (IRF) and denoted $\psi(j) = \psi_j$ with $\psi(0) = 1$. Following the definition of half-life given in (5), we are interested in finding the first instant h such that $\psi(h) = \frac{1}{2}$.

To compute the IRF of a general univariate ARMA(p, q) process, it is convenient to arrange its components in a VAR(1) form. Assuming $\{\tilde{q}_t\}_{t=1}^{T} \sim ARMA(p, q)$,

$$\underbrace{\begin{bmatrix} \tilde{q}_t \\ \tilde{q}_{t-1} \\ \vdots \\ \tilde{q}_{t-p+1} \\ \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q+1} \end{bmatrix}}_{\boldsymbol{\xi}_t} = \underbrace{\begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_p & \theta_1 & \dots & \theta_q \\ 1 & 0 & \dots & & & & \\ 0 & \ddots & 0 & \dots & & & \\ 0 & \dots & 1 & 0 & \dots & & \\ 0 & \dots & & & & & \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix}}_{\mathbf{F}} \underbrace{\begin{bmatrix} \tilde{q}_{t-1} \\ \tilde{q}_{t-2} \\ \vdots \\ \tilde{q}_{t-p} \\ \varepsilon_{t-1} \\ \varepsilon_{t-2} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix}}_{\boldsymbol{\xi}_{t-1}} + \underbrace{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{G}} \varepsilon_t$$

The IRF can be obtained as

$$\psi(j) = \mathbf{e}(\mathbf{F}^j\mathbf{G}) \tag{22}$$

where $\mathbf{e} = [1\ 0\ \dots\ 0]^\top$ is a selection vector to pick up the IRF for $\tilde{q}_t$. Alternatively, it is possible to compute $\psi(j)$ recursively. The half-life is thus given by $\psi(h) = \mathbf{e}(\mathbf{F}^h\mathbf{G}) = 0.5$. For AR(1) processes, the solution is trivially given by $h = \log(0.5)/\log(\rho)$ while for higher order auto-regression we can use the eigenvalues based formula given in Hamilton (1994, p. 12). However, if we want to compute the exact half-life of a general ARMA(p, q) process, we need to use some numerical procedure since no closed form solution exists. Also, the fact that h might no be an integer rules out standard Newton-Raphson methods.

In order to compute the half-life, we propose the following numerical procedure that exploits spline function interpolation and it runs as follows:

(a) Compute ψ_j for $j = 1, \ldots, J$ (typically $50 \leq J \leq 100$), using (22).

(b) Interpolate the above values using a spline function to get

$$\boldsymbol{\psi} = [\psi_0, \psi_{0+1/\Delta}, \psi_{0+(2/\Delta)}, \ldots, \psi_1, \psi_{1+(1/\Delta)}, \psi_{1+(2/\Delta)}, \ldots, \psi_J]$$

where $\Delta \in \mathbb{N}$ denotes the number of values in between two impulse responses. The size of $\boldsymbol{\psi}$ is then $\Delta(J-1)+J$. Define the associated index vector $\boldsymbol{\lambda}$ of size $\Delta(J-1)+J$ such that $\lambda_j = j$.

(c) Compute a new vector, $\boldsymbol{\iota}$, of the same size as $\boldsymbol{\psi}$ such that its j^{th} element is defined as

$$\iota_j = \begin{cases} j & \psi_j \geq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

and denote $\tilde{\boldsymbol{\iota}}$ of size $N < \Delta(J-1)+J$ a vector containing the same elements as $\boldsymbol{\iota}$ but with all null entries removed. Note that $\tilde{\boldsymbol{\iota}}$ is the vector of the time periods (x-axis coordinates) for which the IRF is above or equal 0.5.

(d) Define another index vector $\tilde{\boldsymbol{\lambda}}$ such that $\tilde{\lambda}_i = i$ for $i = 1, \ldots, N$ and compute $\boldsymbol{\delta}$ where $\delta_i = I(\tilde{\lambda}_i < \tilde{\iota}_i)$, $I(\cdot)$ denoting the indicator function. The estimated half-life h is given by

$$h = \begin{cases} \lambda_{\max\{\tilde{\iota}\}} & \max\{\boldsymbol{\delta}\} < 1 \\ \lambda_{\min\{\tilde{\iota}\}} | \delta_{\min\{\tilde{\iota}\}} = 1 & \text{otherwise} \end{cases}$$

where the condition is needed to ensure that if the IRF is equal to 0.5 for different time periods, we want the half-life to be the first of these periods.

The procedure described above is computationally very fast and it allows to obtain precise half-life estimates.

4 Empirical Application

In order to illustrate the empirical relevance of the impact of unaccounted outliers on the half-life estimation, we analyse US dollar bilateral exchange rates. The data used

Table 2: Half-life ML Estimates Without Outliers Detection

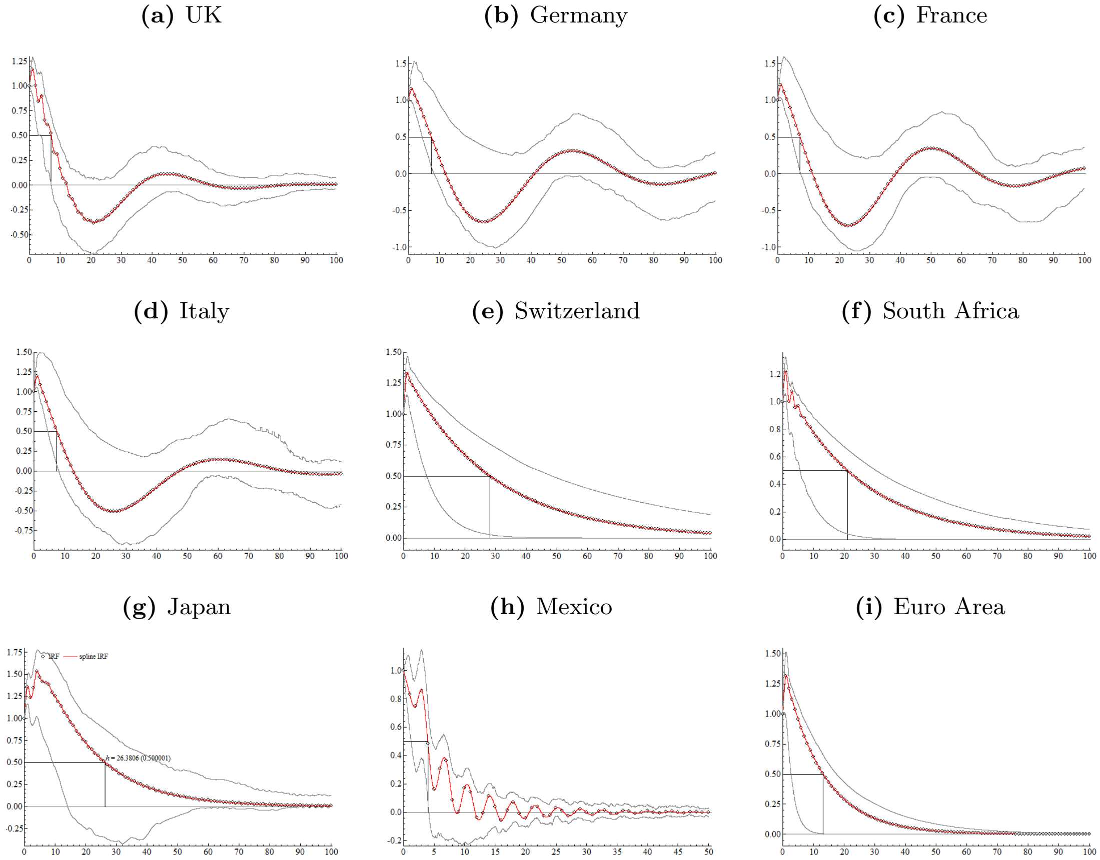
	$\hat{h}$	$\hat{c}_{low}$	$\hat{c}_{upp}$	$(p,\, q)$	AIC	J-B
UK	1.78	0.85	2.26	4,3	-584.61	[0.0050]**
Germany	1.86	1.11	4.82	2,2	-400.49	[0.3885]
France	1.82	1.14	4.27	2,2	-419.27	[0.0456]*
Italy	1.87	1.13	4.74	2,2	-414.70	[0.0306]*
Switzerland	7.05	1.89	12.12	1,1	-541.48	[0.9039]
South Africa	5.27	1.40	7.62	2,1	-471.59	[0.0001]**
Japan	6.60	2.31	11.02	5,1	-545.81	[0.0503]
Mexico	0.99	0.31	1.13	3,3	-211.69	[0.0000]**
Euro Area	3.30	0.67	4.38	1,1	-208.00	[0.2370]

Notes: $\hat{h}$ denotes the annualised half-life estimate, $\hat{c}_{low}$ and $\hat{c}_{upp}$ are the lower and upper endpoint of the bootstrapped confidence interval, (p, q) denotes the ARMA order, *AIC* the Akaike Information Criterion and J-B the p-value of the Jarque-Bera test with '**' and '*' denoting rejection of the null of Normality at 1% and 5% significance level respectively.

are taken from the Federal Reserve Bank of Saint Louis database (FRED). Following the literature, the nominal exchange rate is expressed as national currency units in terms of one US dollar (daily averages) while the price indexes are consumer price indexes (CPI) not seasonally adjusted. We consider quarterly data for the following countries: United Kingdom, Germany, France, Italy, Switzerland, Japan, South Africa, Mexico and the Euro Area. The data span the period 1971:1 to 2013:3 though the number of observations varies from a maximum of 171 to a minimum of 59 for the Euro Area. The log real exchange rate for the i^{th} country is computed as $q_{i,t} = s_{i,t} - p_{i,t} + p_{US,t}$ where $s_{i,t}$ is the logarithm of the nominal exchange rate, $p_{i,t}$ the logarithm of the CPI for that country and $p_{US,t}$ the logarithm of the CPI of the United States.

We start by computing the half-life for the above countries disregarding the possibility of outlying observations. Then, we fit an ARMA model to each real exchange rate series and use the estimated coefficients to compute the IRF response function and the corresponding half-life measure according to the procedure described in Section 3.3. To estimate the ARMA models, we use maximum likelihood methods based on (13) which also avoids problems related with the finite sample bias of the least squares estimator. Table 2 reports the annualised estimated half-life together with 95% bootstrap confidence intervals, obtained by non-parametric bootstrapping with 299 replications, as well as the ARMA order, the AIC and the Jarque-Bera normality test p-value.

Figure 4: Estimated IRFs Without Outliers Detection



Notes: The black circles denotes point estimates of the IRFs, the red line is an interpolating spline and the grey lines are bootstrapped confidence bands (lower band corresponds to the 2.5^{th} percentile while the upper band to the 97.5^{th} percentile). Half-lives are the values (quarterly frequency) on the x-axis in correspondence to the black vertical lines.

Analysing carefully the results, we can conclude that, over the timespan considered, the PPP puzzle is more evident for Switzerland, South Africa, Japan and Euro Area with half-lives between three and seven years whereas is decisively less pronounced for UK, Germany, France, Italy and Mexico. In Figure 4, we report the estimated IRFs with bootstrapped confidence bands.

Next, we re-fit an ARMA model to each real exchange rate time-series following the outlier detection procedure described in Section 3.2. In order to select the significant dummies, we set a significance level of $\alpha = 0.01$ in each individual saturation and, based on the simulations results reported in Table 1, we use 20 blocks of dummies throughout. This allows to control the retention rate of AOs, IOs and LSs at $0.01 \times T$ individually.

Table 3: Half-life ML Estimates With Outliers Detection

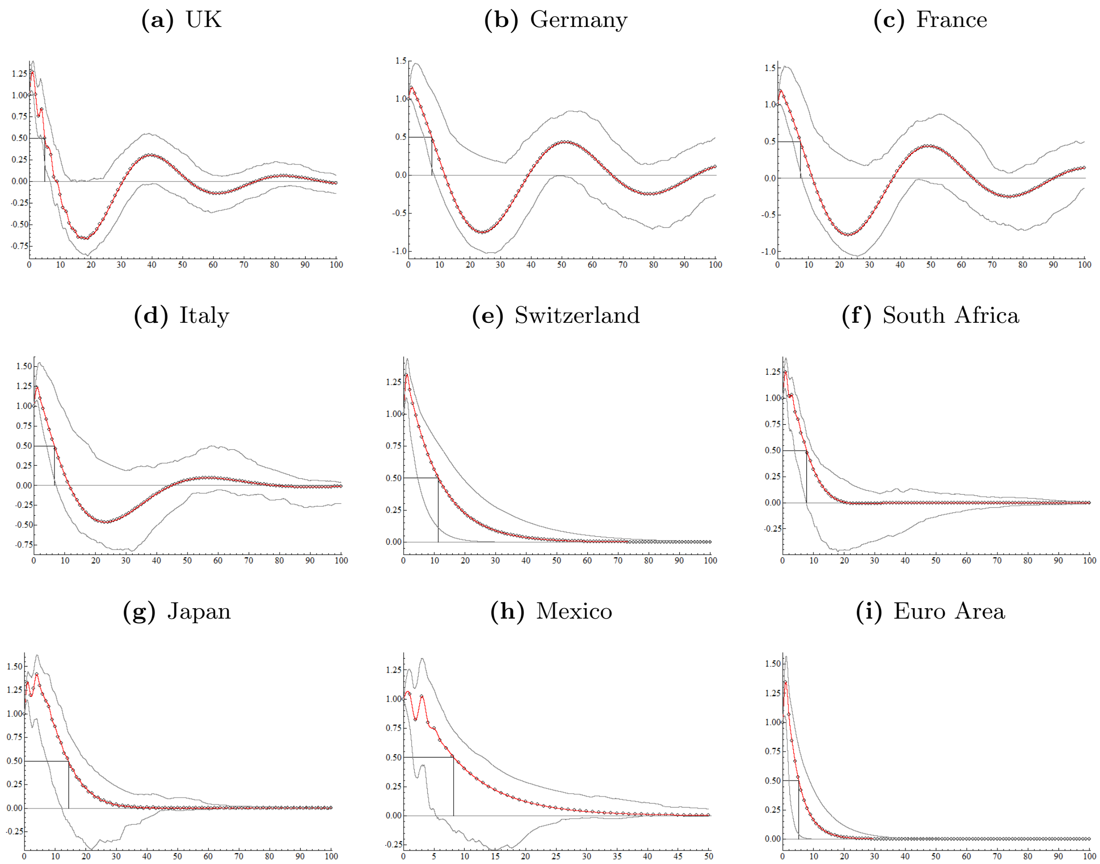
	$\hat{h}$	$\hat{c}_{low}$	$\hat{c}_{upp}$	$(p,\, q)$	AOs	IOs	LSs	AIC	J-B
UK	1.25	0.86	1.93	4,3	1981(3) 1985(1) 1988(2)	2008(4)	1990(3) 1992(4)	-658.67	[0.4640]
Germany	1.89	1.27	3.72	2,2			1974(3) 1975(3) 1984(3) 1988(3)	-412.26	[0.7323]
France	1.85	1.24	3.88	2,2	1985(1)	1991(2)		-425.33	[0.1589]
Italy	1.66	1.03	4.45	2,2	2000(4)	1976(1) 1984(3) 1992(3)		-425.99	[0.0117]*
Switzerland	2.85	1.15	4.82	1,1	1985(1)	1971(1) 1978(1)		-558.30	[0.7162]
South Africa	1.95	0.96	2.42	3,2		2001(4) 2008(4)	1975(4) 1998(3)	-516.03	[0.0000]**
Japan	3.59	1.85	5.07	5,1	1979(4) 1995(2)	1971(1) 1998(4) 2008(4)	1978(3) 2013(1)	-593.94	[0.8416]
Mexico	2.05	0.38	3.26	2,3		1995(1) 2008(4)	1995(2)	-306.87	[0.3196]
Euro Area	1.31	0.51	2.19	1,1	2000(4)		2003(4) 2004(1)	-218.31	[0.2711]

Notes: See below Table 2. Further, under the columns tagged "AOs", "IOs" and "LSs", we report the dates of the additive outliers, innovative outliers and level shifts, respectively.

Hence, under the null of no outliers we should expect $0.03 \times T$ not significant outliers on average before the combined selection to take place in Step 3 where a significance level of $\alpha = 0.05$ is used. The estimated ARMA coefficients after dummy saturation are then used to obtain robust half-lives estimates. Table 3 reports the results while Figure 5 the IRFs.

The results after outliers detection by dummy saturation are very interesting. The main evidence is that for the countries where the real exchange rate shows the most puzzling behaviour, modelling outliers seems to drastically reduce the extent of this puzzle. In particular, for Switzerland, South Africa, Japan and the Euro Area, we observe that their half-lives are reduced by a factor of three or two, the most striking result being South Africa which half-life passes from a point value of 5.27 to 1.95. For the countries where instead the PPP puzzle is less evident or even absent, including outliers seem not to change the half-lives estimates. However, the benefit in accounting for outlying observations appears in the tighter confidence intervals and in the restored normality. The only country with increasing half-life after including outliers is Mexico with a point value of 2.05 after saturation compared to 0.99 before saturation. As far as the outliers

Figure 5: Estimated IRFs With Outliers Detection



Notes: See below Figure 4.

are concerned, we see from Table 3 that the average number of outliers retained for each country is around four while their location is pretty widespread along all the time period considered. The most recurring outlier is the innovative outlier in the fourth quarter of 2008 ("IO:2008(4)") that is in the final model for UK, South Africa, Japan and Mexico. It is straightforward to notice that the mentioned outlier is capturing the effect on the exchange rates of the recent financial crisis which apparently died out – at least for these series – following a decaying effect.

A further investigation of the dynamics followed by the exchange rates is reported in the next section, where we also link our analysis with the "bands of inaction" literature (Taylor, 2001).

4.1 Testing for Non-linear Effects

In this section, we test whether the real exchange rate time series considered in the empirical application exhibit non-linear effects or not. This is of interest for at least two reasons. First, as mentioned above, part of the literature uses non-linear models to describe the real exchange rate process based on the "bands of inaction" argument (Taylor, 2001). Thus, testing for non-linearities can be seen as an empirical test of this phenomenon. Second, by testing for non-linear effects with and without outliers removal, we can also explore the implications of removing outlying observations on non-linearities. This is of particular interest as we want to explore whether non-linearities might be due to unaccounted outliers or, conversely, outliers inclusion is masking non-linearities.

To this purpose, we consider the Brock-Dechert-Scheinkman test statistic (hereafter, BDS) introduced by Brock, Dechert, and Scheinkman (1987) and theoretically established in Brock, Dechert, Scheinkman, and Le Baron (1996). The statistic is based on the concept of *correlation integral* which aims at measuring the frequency with which temporal patterns are repeated in the data. Briefly, consider the time series $\{x_t\}_{t=1,\dots,T}$ which is embedded in the m-space by forming m-histories $x_t^m = (x_t, x_{t-1}, \dots, x_{t-m+1})$. Brock et al. (1996)

shows that the U-statistic

$$C(\eta, m, T) = \frac{2}{(T-m+1)(T-m)} \sum_{m \le s <} \sum_{t \le T} \mathbf{1}(|x_{t-i} - x_{s-i}| < \eta : i = 1, \ldots, m-1)$$

is a consistent estimator of the correlation integral under fairly general assumptions on $\{x_t\}$. If the process is i.i.d. then it is possible to show that $C(\eta, m, T) = C(\eta, 1, T)^m$. Using this fact with the properties of the U-statistics, Brock et al. (1996) show that

$$BDS(\eta, m, T) = \sqrt{T} \frac{C(\eta, m, T) - C(\eta, 1, T)^m}{\sigma(\eta, m, T)} \overset{A}{\sim} \mathcal{N}(0, 1)$$

where $\sigma(\eta, m, T)$ is the standard deviation of $\sqrt{T}C(\eta, m, T) - C(\eta, 1, T)^m$ and $\overset{A}{\sim}$ indicates *"asymptotically distributed as"*.

In particular, we apply the BDS test to the residuals of the fitted ARMA models before and after outliers removal[8]. Rejection of the null indicates that unaccounted dynamics is present in the residuals which can be interpreted as unaccounted non-linearities since Brock et al. (1996) show by simulations that the test has high power against non-linear alternatives. Table 4 reports the p-values of the BDS test statistic at different embedding dimensions. For each country, the first row indicates the p-value of the BDS statistic before outliers removal while the second row the p-value after robust estimation. Empirical evidence of the presence of non-linear effects is not uniform across our sample even before outliers detection. In fact, for only half of the countries considered we fail to accept the null and the results are also dependent on the embedding dimension adopted. Accounting for outliers allows to reduce the number of rejections and in particular, for UK and EMU there is no evidence of non-linear effects using dummy saturation while the null hypothesis is rejected at all embedding dimensions without modelling outliers. However, for countries like South Africa and Italy, where the evidence of misspecified linear models is quite strong, the inclusion of outliers has no effect on the test outcome.

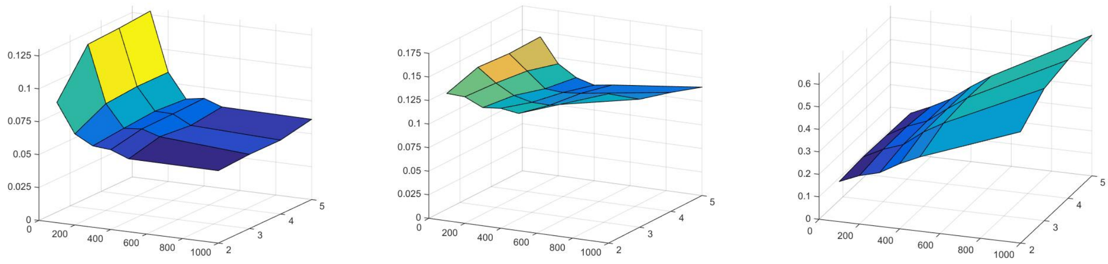
To summarise, empirical evidence suggests that non-linear effects caused by "bands of inaction" is not uniform across the countries considered, and even before removing

[8]The value of the test statistic is obtained using the C code made available by Le Baron at `http://people.brandeis.edu/~blebaron/software/bds/bdsccode/`.

Table 4: BDS Test p-values with embedding dimensions $m = \{2, 3, 4, 5\}$.

$\eta = 0.5$	2	3	4	5
UK	0.003**	0.005**	0.002**	0.000**
	0.484	0.313	0.313	0.004**
Germany	0.116	0.025*	0.960	0.097
	0.002**	0.285	0.689	0.136
France	0.447	0.992	0.603	0.711
	0.116	0.313	0.741	0.857
Italy	0.001**	0.000**	0.000**	0.000**
	0.037*	0.006**	0.007**	0.000**
Switzerland	0.749	0.535	0.126	0.022*
	0.772	0.294	0.562	0.352
South Africa	0.003**	0.001**	0.000**	0.000**
	0.001**	0.000**	0.000**	0.000**
Japan	0.689	0.298	0.230	0.478
	0.757	0.407	0.711	0.332
Mexico	0.000**	0.002**	0.052	0.099
	0.000**	0.004**	0.067	0.072
EMU	0.215	0.002**	0.000**	0.000**
	0.555	0.522	0.119	0.555

Notes: ** and * denote presence of non-linear effects at 1% and 5% significance level. For each country, the first row reports p-values without outliers detection while the second row p-values with robust estimation.



(a) No outliers **(b)** AOs **(c)** LSs

Figure 6: Empirical rejection frequencies of the BDS test. The plots show how often the BDS test rejects the null hypothesis of no non-linear effects (Z axis) for different sample sizes (X axis) and different embedding dimensions (Y axis). The DGP is an ARMA(1,1) process with no outliers **(a)**, an ARMA(1,1) contaminated by AOs **(b)**, and an ARMA(1,1) contaminated by LSs **(c)**.

outliers this is observed only in half of the sample. Moreover, the number of rejections of the null hypothesis drops further when modelling outlying observations suggesting that model misspecification – at least in some cases – can be reasonably thought to be due to unaccounted outliers. The intuition behind this claim is further confirmed by Figure 6 where we compute by simulations the empirical rejection frequency of the BDS test when applied to the residuals of an ARMA model contaminated by outliers and level shifts. Clearly, the test over-rejects the null hypothesis if outlying observations are not taken into account during the estimation process. Nevertheless, in our empirical application there are cases where we fail to reject the null hypothesis disregarding the fact that we take into account the presence of outliers or not, thus suggesting that non-linearities might indeed play a role in such situations.

5 Conclusions

In this paper, we studied to which extent the half-life estimates and the related PPP puzzle are affected by unaccounted outlying observations. In particular, we modelled the observed real exchange rate process as a linear ARMA process contaminated by additional and innovative outliers as well as level shifts. In order to estimate such models, we proposed a sequential Dummy Saturation approach combined with ML estimation of the relevant parameters. In a Monte Carlo simulation exercise, we showed that the proposed estimation procedure delivers the correct retention rates of the dummy variables included to capture the outliers.

We illustrate the impact of removing outliers on the half-life estimates through an empirical application involving a panel of US dollar denominated real exchange rates over a period spanning the last four decades. Our findings can be summarised as follows. First, for countries where the real exchange rate shows the most puzzling behaviour, the inclusion of outliers drastically reduces the extent of the PPP puzzle. Second, for countries where half-life estimates are contained even without including outliers, the benefit of a robust estimation approach is reflected in tighter confidence intervals for the half-life as well as in the restoration of normality. Therefore, in light of the evidence emerging from the empirical study, the main conclusion is that robust estimation methods allow to reduce the impact of the PPP puzzle and are anyhow beneficial for statistical inference even when the half-life estimates are not affected when accounting for outlying observations. Additionally, motivated by the "bands of inaction" argument of Taylor (2001), we tested for the presence of non-linear effects in the real exchange rate process with and without the inclusion of outliers. We found mixed evidence of the presence of non-linear effects across our sample even before removing outlying observations. The evidence stabilises against the presence of non-linearities when accounting for outliers. This is symptomatic of the fact that unaccounted outliers may indeed lead to spurious non-linear effects.

The findings in this paper suggest a number of stimulating avenues for future research. First, from a methodological point of view, it would be of interest to extend the results in Johansen and Nielsen (2009), which relate the Dummy Saturation to robust M-estimation, to our ML-based setting. Second, in light of the empirical application, it will be of some interest to investigate the role of outlying observations in the presence of time series which exhibit genuine non-linear dynamics (see Franses et al., 1996). Third, given the generality of the proposed procedure for outliers detection in a time-series setting, it will be interesting to extend its application to other areas of financial economics such as for instance stock return predictability. This is part of an ongoing research agenda.

References

Abuaf, N. and Jorion, P. (1990). Purchasing Power Parity in the Long-Run. *Journal of Finance 45*, 157–174.

Andrews, D. K. W. and Chen, H. Y. (1994). Approximately Median Unbiased Estimation of Autoregressive Models. *Journal of Business & Economic Statistics 12*, 187–204.

Bergamelli, M. and Urga, G. (2013). Detecting Multiple Structural Breaks: Dummy Saturation vs Sequential Bootstrapping. *Discussion Paper, Cass Business School, London*.

Brock, W.A. and Dechert, W.D. and Scheinkman, J.A. (1987). A Test of Independence Based on the Correlation Dimension. *Working Paper. Department of Economics, University of Winsconsin at Madison, University of Houston, and University of Chicago*.

Brock, W.A. and Dechert, W.D. and Scheinkman, J.A. and Le Baron, B. (1996). A Test of Independence Based on the Correlation Dimension. *Econometrics Reviews 15*, 197–235.

Castle, J. and Doornik, J.A. and Hendry, D.F. (2012). Model Selection when there are Multiple Breaks. *Journal of Econometrics 169*, 239–246.

Cavaliere, G. and Georgiev, I. (2009). Robust Inference in Autoregressions with Multiple Outliers. *Econometric Theory 25*, 1625–1661.

Chang, I. and Tiao, C. G. and Chen, C. (1988). Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics 30*, 193–204.

Chen, C. and Liu, L.-M. (1993). Joint Estimation of Model Parameters ans Outlier Effects in Time Series. *Journal of the American Statistical Association 88*, 284–297.

Cheung, Y. W. and Lai, K. S. (2000). On the Purchasing Power Parity Puzzle. *Journal of International Economics 52*, 321–330.

Chortareas, G. and Kapetanios, G. (2013). How Puzzling is the PPP Puzzle? An Alternative Half-Life Measure of Convergence to PPP. *Journal of Applied Econometrics 28*, 435–457.

Diebold, F. X. and Husted, S. and Rush, M. (1991). Real Exchange Rates under the Gold Standard. *Journal of Political Economy 99*, 1252–1271.

Doornik, J. A. (2009). Autometrics. In J. Castle and N. Shepard (Eds.), *The Methodology and Practice of Econometrics*, Chapter 4, pp. 88–121. Oxford: Oxford University Press.

Doornik, J. A. and Hendry, D. F. and Pretis, F. (2013). Step Indicator Saturation. *Discussion Paper 658,* Department of Economics, University of Oxford.

Doornik, J. A. and Ooms, M. (2012). A Package for Estimating, Forecasting and Simulating Arfima Models: ARFIMA package 1.06 for Ox. *User Manual*.

Frankel, J. A. and Rose, A. K. (1996). A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries. *Journal of International Economics 40*, 209–224.

Franses, P.H and van Dijk, D. J.C. and Lucas, A. (1996). Testing for Smooth Transition Nonlinearity in the Presence of Outliers. *Journal of Business & Economic Statistics 17*, 217–235.

Galbraith, J. W. and Zinde-Walsh, V. (1992). The GLS Transformation Matrix and a Semi-Recursive Estimator for the Linear Regression Model with Arma Errors. *Econometric Theory 8*, 95–111.

Hamilton, J.D. (1994). *Time Series Analysis.* Princeton: Princeton University Press.

Hendry, D.F. and Johansen, S. and Santos, C. (2008). Automatic Selection of Indicators in a Fully Saturated Regression. *Computational Statistics 23*, 317–335 and Erratum 337–339.

Hendry, D. F. (1999). An Econometric Analysis of US Food Expenditure, 1931-1989. In J. R. Magnus and M. S. Morgan (Eds.), *Methodology and Tacit Knowledge: Two Experiments in Econometrics*, Chapter 17, pp. 341–361. Chicester: John Wiley & Sons.

James, J., I. Marsh, and L. Sarno (Eds.) (2012). *Handbook of Exchange Rates.* Hoboken, New Jersey: John Wiley & Sons.

Johansen, S. and Nielsen, B. (2009). An Analisys of the Indicator Saturation Estimator as a Robust Regression Estimator. In J. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics*, Chapter 1, pp. 1–36. Oxford: Oxford University Press.

Mark, N. C. (2001). *International Macroeconomics and Finance: Theory and Empirical Methods*. Blackwell.

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

Rossi, B. (2005). Confidence Intervals for Half-Life Deviations from Purchasing Power Parity. *Journal of Business & Economic Statistics 23*, 432–442.

Sánchez, M. J. and Peña. (2003). The Identification of Multiple Outliers in ARIMA Models. *Communications in Statistics 32*, 1265–1287.

Sarno, L. and Taylor, M. (2002). *The Economics of Exchange Rates*. Cambridge, UK: Cambridge University Press.

Taylor, A. M. (2001). Potential Pitfalls for Purchasing Power Parity Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the Law of One Price. *Econometrica 69*, 473–498.

Tsay, R. S. (1986). Time Series Model Specification in the Presence of Outliers. *Journal of the American Statistical Association 81*, 132–141.

——— (1988). Outliers, Level Shifts, and Variance Changes in Time Series. *Journal of Forecasting 7*, 1–20.