Confidence intervals for half-life deviations from Purchasing Power Parity

Barbara Rossi*

This draft: April 2004

Abstract

According to Rogoff (1995), Purchasing Power Parity (PPP) fits well one empirical feature of the data, namely the high short-run volatility of real exchange rates, but also implies that shocks should die away in one to two years (the time interval compatible with price and wage stickiness). However, existing point estimates of half-life deviations from PPP are in the order of 3 to 5 years. This paper assesses how much uncertainty there is around these point estimates by using local to unity asymptotic theory to construct confidence intervals that are robust to high persistence in the presence of small sample sizes. The empirical evidence suggests that the lower bound of the confidence interval is around 4 to 8 quarters for most currencies, which is not inconsistent with PPP. However, the upper bounds are infinity for all currencies so we cannot provide conclusive evidence in favor of PPP either.

Keywords: Purchasing power parity; half-life; persistence; roots close to unity.

JEL classification: F300; F400.

*Department of Economics, Duke University, Durham NC27708 USA.

1. Introduction

What determines nominal exchange rates in the long-run? According to Purchasing Power Parity (PPP), since the (bilateral) nominal exchange rate (E_t) is the relative price of two currencies,[1] in equilibrium it should reflect their relative purchasing powers. So, if P_t is the price level in the home country and P_t^* is the price level in the foreign country, then PPP requires:

$$E_t = \frac{P_t}{P_t^*} \tag{1}$$

Thus, the logarithm of the real exchange rate, defined as $y_t = \ln\left(\frac{E_t P_t^*}{P_t}\right)$, should be constant if PPP holds at every point in time. A weaker version of the PPP, which is followed in this paper and in most of the literature, requires only that (1) holds in the long run.

The empirical evidence on PPP is mixed. Although casual evidence suggests that the two series, E_t and P_t/P_t^*, tend to revert towards each other over time, there are protracted periods in which the nominal exchange rate deviates from its PPP level. How persistent are these deviations? A measure of persistence is the half-life of PPP deviations. To motivate this measure, suppose that the deviations of the logarithm of the real exchange rate y_t from its long run value y_0, which is constant under PPP, follow an autoregressive process of order one:

$$y_t - y_0 = \rho \left(y_{t-1} - y_0\right) + \epsilon_t \tag{2}$$

where ϵ_t is a white-noise. Then, at horizon h, the percentage deviation from equilibrium is ρ^h. The half-life deviation from PPP is defined as the smallest value of h such that $\mathcal{E}\left(y_{t+h} - y_0 \mid y_{t-s} - y_0, s \leq 0\right) \leq \frac{1}{2}\left(y_t - y_0\right)$, where $\mathcal{E}$ is the expectation operator. That is:[2]

$$\rho^h = \frac{1}{2} \quad \Rightarrow \quad h = \frac{\ln(1/2)}{\ln(\rho)} \tag{3}$$

Using data under floating exchange rate regimes, estimates of h range between 2 to 5 years for most countries, with an average of 3.7 years (see table 7.2 in Mark (2001)).[3]

The existing point estimates of half-life deviations from PPP are difficult to reconcile with PPP. According to Rogoff (1995), deviations from PPP can be attributed to transitory disturbances, like financial and monetary shocks, which buffet the nominal exchange rate and translate into real exchange rate variability because of nominal price stickiness. Thus,

while PPP is compatible with the enormous short-term volatility of real exchange rates, it also implies that deviations should be short-lived, as they can only occur during a time frame in which nominal wages and prices are sticky (that is one to two years). The existing point estimates imply instead that deviations are much more persistent than that.

The contributions of this paper are two. First, it introduces a measure of the half-life for a general AR(p) process that allows better asymptotic approximations in the presence of a root close to unity. Although the methods for deriving the half-life are quite standard, there is no such result in the literature. Abuaf and Jorion (1990) discuss half-lives in the context of an AR(1) process only. Mark (2001) discusses measures of half-lives for general AR(p) processes, but for *stationary* processes only. Andrews (1993) proposed a measure of half-life for an AR(1) process which is robust to the presence of high persistence. Andrews and Chen (1994) generalized the method to obtain an approximate median unbiased estimate of AR(p) coefficients in the presence of high persistence. They showed how to construct an approximately median unbiased estimator of the impulse-response function, but did not provide an analytic measure of the half-life for AR(p) processes. Second, it uses this measure to provide a simple method for constructing confidence intervals for the half-life. The issue is complicated by both the high persistence in real exchange rates and the small samples usually available. For these reasons, this paper considers an alternative asymptotic theory based on local-to-unity asymptotics and a half-life that grows to infinity at the rate of the sample size, as in Stock (1996) and Phillips (1998). How good this approximation is relative to the conventional (normal sampling) asymptotic theory is discussed in a Monte Carlo experiment.

This is not the first paper on inference about half-life deviations from PPP. Four recent works that address this issue are Cheung and Lai (2000), Murray and Papell (2002), Lopez, Murray and Papell (2003), Gospodinov (2002) and Kilian and Zha (2002). All these papers calculate confidence intervals by estimating impulse responses with various methods: Cheung and Lai rely on stationary, normal sampling distributions, Murray and Papell rely on Andrews and Chen (1994), Gospodinov (2002) relies on Hansen (1999) and Kilian and Zha (2002) use Bayesian methods. However, estimating the whole impulse-response function may quickly become computationally intensive.

Overall, the results of this paper are not inconsistent with the PPP, although they do not

solve the PPP puzzle either. The existing point estimates, although too high to be reconciled with the PPP, also have huge variability. As a result, confidence intervals with 95% coverage for most currencies include 4 to 8 quarters as their lower bound, a time interval in which deviations from PPP are compatible with nominal price and wage stickiness. However, since we cannot rule out the possibility of an infinite half-life, we interpret the evidence as being simply not informative enough.

The paper is organized as follows. The next section introduces the data generating process considered in this paper and derives the measure of half-life used in this paper. Section 3 describes the methods used to construct the confidence intervals for h. The fourth section discusses a small Monte Carlo experiment that compares the coverage of the various confidence intervals discussed in section 3. The fifth section discusses the empirical results and section 6 concludes.

2. Measuring the half-life

Let the data generating process (DGP) for the log of the real exchange rate, y_t, be:

$$
\begin{aligned}
y_t &= d_t + u_t & t = 1, 2..., T \\
u_t &= \rho u_{t-1} + v_t
\end{aligned}
\tag{4}
$$

where $d_t = \mu_0$ is a deterministic component,[4] v_t is a zero mean, stationary and ergodic process, with finite autocovariances $\gamma(k) = \mathcal{E} v_t v_{t-k}$, $\omega^2 = \sum_{k=-\infty}^{+\infty} \gamma(k)$ is finite and non-zero and $v_t = b(L)^{-1} \epsilon_t$, where ϵ_t is a martingale difference sequence with finite fourth moments and constant variance σ_ϵ^2 and $b(L)$ is finite order and has $p < \infty$ (stable) roots.

In order to provide better asymptotic approximations to the statistics of interest in small samples when variables are highly persistent, we use local to unity asymptotic theory (see Stock (1991) among others):

$$
\rho = e^{c/T} \simeq 1 + \frac{c}{T},
\tag{5}
$$

where c is a constant (negative, if the process is highly persistent but mean reverting) and T is the sample size. In order to provide better small sample approximations in situations where the true half-life, h, can be "big" relative to the sample size, we derive the asymptotic

distributions by letting h increase as the sample size T increases in such a way that their ratio remains a fixed number δ. That is:

$$\frac{h}{T} \underset{T \to \infty}{\to} \delta \tag{6}$$

We will refer to δ as the *"half-life as a fraction of the sample size"*.

The persistence[5] of the process in small samples, measured by c, is relevant for our purposes, since we are trying to estimate at which horizon the deviations from PPP are back to one-half after a shock. As we will show later, the speed at which the effect of a shock dies away depends on a function of the largest root of the process, ρ^h, and, under assumption (6):

$$\rho^h \underset{T \to \infty}{\to} e^{c\delta} \tag{7}$$

In order to derive an expression for the half-life in this general $AR(p)$ process, we need to derive an expression for the effect of the shock ϵ_t on y_t after h periods. We will derive it in terms of the eigenvalues of the process.[6] We factorize (4) as:

$$(1 - \lambda_1 L)(1 - \lambda_2 L) \dots (1 - \lambda_p L)(y_t - d_t) = \epsilon_t \tag{8}$$

where, for convenience, $\lambda_1 = \rho$ is the root close to unity and λ_2, λ_3, ...,λ_p are the (stable) roots, the inverse of the roots of the polynomial $b(L)$. We also define $\boldsymbol{\lambda}$ to be a $(p \times 1)$ vector containing all the eigenvalues of the data generating process, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ...,\lambda_p]'$. We assume that the eigenvalues are distinct. Suppose we start at time $t - 1$ in the long run equilibrium μ_0, and at time t there is a shock ϵ_t. No other shocks hit the economy subsequently. The shock ϵ_t measures the initial deviation from equilibrium, which we denote by $\widetilde{y}_t \equiv y_t - \mu_0 = \epsilon_t$. It follows that the deviation from equilibrium after h periods will be $\widetilde{y}_{t+h} = \mathbf{c}'\boldsymbol{\lambda}^h \epsilon_t$, where $\mathbf{c}$ is a $(p \times 1)$ vector with generic element:

$$c_i = \frac{\lambda_i^{p-1}}{\prod\limits_{k=1, k \neq i}^{p} (\lambda_i - \lambda_k)} \tag{9}$$

and $\boldsymbol{\lambda}^h$ is the $(p \times 1)$ vector containing all the eigenvalues to the h power (see Hamilton, p. 12). After h periods, the percentage deviation from equilibrium relative to the initial percentage deviation from equilibrium is:[7]

$$\frac{(y_{t+h} - \mu_0)}{(y_t - \mu_0)} = \frac{\partial \widetilde{y}_{t+h}}{\partial \epsilon_t} = \mathbf{c}'\boldsymbol{\lambda}^h \tag{10}$$

We call $\partial \widetilde{y}_{t+h}/\partial \epsilon_t$ (which is the usual definition of an impulse-response) "the effect of a shock ϵ_t after h periods". By combining (9) and (10) and isolating the largest root $\lambda_1 (= \rho)$:

$$\frac{\partial \widetilde{y}_{t+h}}{\partial \epsilon_t} = \frac{\lambda_1^{h+p-1}}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \dots (\lambda_1 - \lambda_p)} + \sum_{i=2}^{p} \frac{\lambda_i^{h+p-1}}{\prod\limits_{k=1, k \neq i}^{p} (\lambda_i - \lambda_k)} \tag{11}$$

As all eigenvalues except the first one are in modulus less than one then, as $h \rightarrow \infty$, $\lambda_i^{h+p-1} \rightarrow 0 \; \forall i \neq 1$ so the second component in (11) disappears. Also, since by assumption p is finite, by combining (6) and (7), as $T \rightarrow \infty$ we have that $\lambda_1^{h+p-1} = \rho^{h+p-1} = \left(1 + \frac{c}{T}\right)^{h+p-1} \rightarrow e^{c\delta}$ (since $\rho^{p-1} \rightarrow 1$). Also asymptotically, $\lambda_1 - \lambda_i \simeq 1 - \lambda_i \; \forall i \neq 1$. Notice finally that $(1 - \lambda_2)(1 - \lambda_3) \dots (1 - \lambda_p) = b(1)$. Thus, the effect of the shock after h periods is:[8]

$$\frac{\partial \widetilde{y}_{t+h}}{\partial \epsilon_t} \underset{T \rightarrow \infty}{\rightarrow} e^{c\delta} b(1)^{-1} \tag{12}$$

The half-life is defined as the horizon h at which the effect of the shock is one-half. Hence, from (12), we obtain that the half-life as a fraction of the sample size is:

$$\delta^* \equiv \max \left\{ \frac{\ln\left(\frac{1}{2} b(1)\right)}{c}, 0 \right\} \tag{13}$$

where we call $b(1)$ the "correction factor". We let $\delta^* \equiv +\infty$ if $c \geq 0$.[9] It follows that the half-life is:

$$h^* \equiv \max \left\{ \frac{\ln\left(\frac{1}{2} b(1)\right)}{\ln \rho}, 0 \right\} \tag{14}$$

Again, $h^* \equiv +\infty$ if $\rho \geq 1$. Note the monotonicity of the relationship between δ^* and c in (13). The monotonicity arises because, in the long-run, it is the root close to unity that is relevant. The monotonicity is not assured if the autoregressive process is not persistent.

Let us compare our measure of the half-life with those one would get by running the regression in the form of an AR(1), thus ignoring short run dynamics (as in Abuaf and Jorion (1990), Frankel and Rose (1996) and Lothian and Taylor (1996)), or by running the regression in ADF form and calculating the half-life on the basis of the coefficient on the lagged level variable only (see Mark (2001), par. 2.4, and references therein). The measure of the half-life in the former approach, let's call it $h_{AR(1)}$, would ignore the correction factor $b(1)$:[10]

$$h_{AR(1)} \equiv \max \left\{ \frac{\ln\left(\frac{1}{2}\right)}{\ln \rho}, 0 \right\} \tag{15}$$

By comparing (14) with (15), it is clear that they will differ unless the true DGP is an AR(1) – in which case $b(1) = 1$. In the second approach, the researcher relies on estimates from the ADF regression. As in Stock (1991), the data generating process (4), can be rearranged to yield the following ADF regression:

$$y_t = \widetilde{\mu}_0 + \alpha(1) y_{t-1} + \sum_{j=1}^{k} \alpha_{j-1}^* \Delta y_{t-j} + \epsilon_t \tag{16}$$

where $\alpha(1) = 1 + \frac{c}{T} b(1)$, $\widetilde{\mu}_0 = -\frac{c}{T} \mu_0 b(1)$, $k = p - 1$ and $\alpha_j^* = -\sum_{i=j+1}^{k} \alpha_j$. The researcher then calculates the half-life based on the AR(1) model (3), let's call it h_a, as:

$$h_a \equiv \max\left\{ \frac{\ln\left(\frac{1}{2}\right)}{\ln \alpha(1)}, 0 \right\} \tag{17}$$

which corresponds to an approximate half-life as a fraction of the sample size equal to:

$$\delta_a \equiv \max\left\{ \frac{\ln\left(\frac{1}{2}\right)}{cb(1)}, 0 \right\} \tag{18}$$

Expression (17) has also been proposed by Andrews (1993) for the simple AR(1) case. Although $\alpha(1)$ is estimated from the (correct) AR(p) process, the half-life is calculated as if $\alpha(1)$ were equal to the largest autoregressive root (ρ) in (2). By comparing (13) with (18), it is clear that, again, they will differ unless the true data generating process is an AR(1). The intuition behind this result is that (16) is not the Sims, Stock and Watson (1990) canonical representation, and this will matter at long horizons (see the appendix for additional technical explanations and clarifications).

3. Econometric methods

After discussing different measures of half-lives, we address the issue of how to construct confidence intervals with the correct coverage. First, it is well known that $b(1)$ can be consistently estimated by using the estimates from the ADF regression (see Stock (1991)), so the correction factor can be consistently estimated by:

$$\widehat{b(1)} = \left(1 - \sum_{j=1}^{k} \widehat{\alpha}_{j-1}^* \right) \tag{19}$$

Second, there are a variety of methods for constructing confidence intervals for the root close to unity, ρ. In the remainder of this section, we will discuss them and show how

to construct confidence intervals for the half-life. The discussion that follows focuses on two-sided confidence intervals; the construction of one-sided confidence intervals follows in a straightforward way.

(i) Confidence intervals based on normal sampling distributions

Starting from (3) and a usual Augmented Dickey-Fuller (ADF) regression (16), where y_t is the real exchange rate, some empirical papers estimates the half-life to be:

$$\widehat{h}_a = \frac{\ln(0.5)}{\ln(\widehat{\alpha}(1))} \tag{20}$$

where we use "hats" above a parameter to denote its estimated value. Hence, by using a delta method approximation, a conventional two-sided 95% confidence interval for h_a, (h_l^a, h_u^a), is:

$$\widehat{h}_a \pm 1.96\widehat{\sigma}_{\widehat{\alpha}(1)} \left(\frac{\ln(0.5)}{\widehat{\alpha}(1)} \left[\ln\left(\widehat{\alpha}(1)\right)\right]^{-2} \right) \tag{21}$$

where $\widehat{\sigma}_{\widehat{\alpha}(1)}$ is an estimate of the standard deviation of $\alpha(1)$.

(ii) Confidence intervals for persistent time series based on Stock (1991) method

Stock (1991) proposed a method for constructing confidence intervals based on median unbiased estimates of the largest autoregressive root of the process (4). Note that a given sample size T and a given c identify the length of the half-life deviation, h, and that the true half-life as a fraction of the sample size, δ, is a monotone decreasing function of c. It is then possible to construct a confidence interval for the half-life by using Stock (1991) method to construct a two-sided confidence interval for c, $(\widehat{c}_l, \widehat{c}_u)$, and then, by monotonicity, obtain a confidence interval for the half-life as a fraction of the sample size, $(\widehat{\delta}_l, \widehat{\delta}_u)$ by applying equations (13) and (18). These can be directly transformed to confidence intervals for the half-life, $(\widehat{h}_l^a, \widehat{h}_u^a)$, by multiplying by the sample size T.

(iii) Confidence intervals based on Elliott and Stock (2001) method

In empirical applications, Stock's method may deliver wide confidence intervals. The reason is that it is based on inverting the ADF test statistic, which has poor power properties (see Elliott, Rothemberg and Stock, 1996). Another drawback of Stock's approach is that it

allows the construction of a confidence interval by inverting a statistic for testing whether $c = 0$; in general, it might be interesting to test other null hypotheses. Elliott and Stock (2001) discuss how to build confidence intervals based on the point optimal test proposed by Elliott, Rothemberg and Stock (1996) and for more general null hypotheses. This section builds on their results.

To construct the confidence interval, we follow Elliott and Stock (2001) and invert a sequence of test statistics, each test statistic being the point-optimal one for testing $H_0 : c = c_0$. We chose the coverage of the confidence interval constructed in this way to be at least 95%. The lag length in the test statistic is chosen according to the MAIC criterion based on GLS detrended data, as suggested by Ng and Perron (2001).

(iv) Hansen's grid-bootstrap method

An alternative method for constructing confidence intervals is by applying bootstrap methods. However, as Hansen (1999) showed, the problem with conventional bootstrap methods is that, even in large samples, their coverage probability is quite poor if the true value of the highest autoregressive root is close to unity and that root is the parameter of interest.

A bootstrap method that is valid in the presence of highly persistent variables is Hansen's (1999) grid-α bootstrap method. Since the half-life is a monotone transformation of the ADF parameters, by the transformation-respecting property of the percentile method we can construct a confidence interval for the half-life by taking the monotone transformation of the corresponding confidence interval for $\widehat{\alpha}(1)$. Furthermore, by the range-preserving property of the grid-α method, the constraint that the half-life cannot be negative directly translates into a constraint on $\alpha(1)$ (namely $\alpha(1) \leq 1$) so it will be automatically satisfied.[11]

As emphasized in Stock (1991), the first order asymptotic theory described in (i) does not provide a suitable framework for the construction of confidence intervals because it is discontinuous at $\rho = 1$. This motivates the construction of confidence intervals for h based on confidence intervals for c obtained by using the procedures (ii) to (iv) described above. (ii) is expected to provide better approximations than the usual normal sampling theory when ρ is close to one. This is based on the conjecture that the inversion of the ADF test statistic to form confidence intervals will work better than normal sampling for $|c|$ small.

Thus, the performance of the proposed method will depend upon ρ (it will work better for ρ close to one) and on the magnitude of h relative to T (the larger h relative to T, the better the procedure). In particular, the theory does not imply that the confidence intervals will have good coverage for ever-growing sample sizes for a fixed value of h. However, for some intermediate combinations of parameter values and sample sizes the proposed procedure may work well. Only Monte Carlo evidence will establish how good these approximations are, and the next section provides such evidence. It is important to emphasize that not only the absolute performance is important, but also the relative performance (relative that of other available methods).

4. A small Monte Carlo experiment

To evaluate the performance of the methods used in this paper, we perform some small Monte Carlo experiments. The experiments are useful to address two questions: "how big does the true half-life have to be in order for the $\frac{h}{T} \to \delta$ asymptotics to be better than the usual asymptotics?" and "how does serial correlation in $b(L)$ affect the performance of different methods?".

In the first experiment, data are generated as $y_t = \rho y_{t-1} + \epsilon_t$, $\epsilon_t \sim N(0,1)$, $\rho = \left(1 + \frac{c}{T}\right)$ and we compare the coverage of the confidence intervals for different values of the true half-life as a fraction of the sample size, δ. The number of Monte Carlo replications is 5000, the sample size is $T = 100$. Results are reported in figure 1.

Insert Figure 1

The figure compares confidence intervals based on four tests:

- the normal sampling test for testing $H_0 : \delta = \delta_0$ (where δ_0 is the true value of δ) versus $H_A : \delta \neq \delta_0$, labeled "*normal sampling h*" test;

- the normal sampling test for testing $H_0 : \rho = \rho_0$ (where ρ_0 is the true value of ρ) versus $H_A : \rho \neq \rho_0$, labeled "*normal sampling rho*";

- the ADF-t unit root test for testing $H_0 : \rho = 1$ versus $H_A : \rho < 1$, labeled "*unit root*";

- Stock's test as described in the previous section, labeled "*Stock*".

As expected, the coverage of the normal sampling confidence intervals is fine when the half-life is small relative to the sample size (δ is smaller than 0.02, say); however, when the half-life is big then the normal sampling test rejects too often and, as a consequence, coverage is lower than nominal. Note that the same problems would arise if one would construct the confidence intervals based on impulse-response functions without taking into account the persistent nature of the process. The problem with unit root tests, instead, is that they lack power. Finally, notice that confidence intervals based on Stock method have coverage close to nominal for most true half-lives (unless they are considerably short).

We also compared the performance of the various methods that are robust to high persistence in another small Monte Carlo experiment. We generate the data as above for four different true half-lives: 2, 6, 20 and infinity.[12] Table 1(a) reports the results. Notice that the coverage of the confidence interval based on normal sampling asymptotic theory (denoted by *Normal*) tends to zero as h increases and is pretty poor relative to, say, that based on Stock (1991) method *even for very small half-lives*. Similar results hold for the grid-bootstrap method. Elliott and Stock (2001) method seems to be the most sensitive to the presence of stationarity.

Finally, Table 1(b) compares confidence intervals based on different methods in simple AR(p) processes (p=2,3,4,5) described by (8) for sample sizes usually available for quarterly real exchange rates (T=100). We compare confidence intervals based on the following methods: Stock applied to (14) (labeled *"Stock"*) and Stock applied to (17), (labeled *"ADF"*). The table reports the actual confidence interval type I error, which ideally should be 0.10. Stock's method performs well as long as the roots other than ρ are not too close to unity (as otherwise the second component in (11) becomes asymptotically non-negligible). The performance of the ADF method worsens as the amounts of serial correlation increases. Some of the DGPs are calibrated on actual estimates: AR(3) with $\lambda_2 = -0.4, \lambda_3 = 0.4$ corresponds to Denmark (results are similar for Belgium); AR(5) with $\lambda_2 = -0.68, \lambda_3 = \lambda_4 = 0.15$, $\lambda_5 = 0.08$ corresponds to Sweden; AR(5) with $\lambda_2, \lambda_3 = 0.3, \lambda_4, \lambda_5 = -0.5$ corresponds to Finland (results are similar for Greece and Japan). Table 1(b) shows that, for the latter case, using h_a rather than h_e may lead to slightly higher confidence interval type I error. Most countries with AR(2) processes have a very small second root, usually smaller than 0.2 in absolute value, for which the two methods provide similar results. Thus, looking ahead

to the empirical section, for most currencies the serial correlation is so small that the two methods are expected to give similar results (except for Finland, Greece and Japan).

Evidence on the coverage properties of our method is not sufficient to conclude that the method is reliable, as we would also like the confidence intervals to exclude an infinite half-life with sufficient probability. Table 1(b) also reports evidence on the empirical proportion of times in which the confidence interval based on Stock (1991) does not include an infinite half-life (labeled "h_∞"). The table shows that our method can potentially rule out an infinite half-life, and, for given values of the short run serial correlation, these rejections are more likely the more stationary the true DGP is. For example, in the AR(2) case (Norway, Switzerland and the U.K.), rejection frequencies corresponding to values of c around the lower bound of the confidence interval (between -15 and -20) can be as high as 0.40. Finally, Table 1(c) shows that, for a given value of ρ, rejection rates become closer to one as the sample size increases.

Insert Tables 1(a), 1(b) and 1(c)

5. Empirical results

The data used in this paper are from Datastream (IMF Database). Data on the nominal exchange rate are end-of-period and data on prices are seasonally unadjusted, to avoid temporal aggregation issues (as discussed in Taylor, 2001a). The nominal exchange rate is expressed as national currency units in terms of 1 U.S. dollar (that is the price of dollars in terms of other national currency units). Data are quarterly from 1973:3 to 2002:2 for non-EMU countries and from 1973:3 to 1998:2, for a total of 100 observations, for EMU countries. The price indices are consumer price indexes (CPI), so they don't distinguish between tradeables and non-tradeables. The log of the real exchange rate is constructed as the log of the bilateral nominal exchange rate plus the log of the CPI in the U.S. minus the log of CPI in the reference country. The series used are "XXI64...F" (CPI) and "XXI..AE." (bilateral nominal exchange rate), where "XX" is the mnemonic used by the IMF to denote the country (e.g. "US" for the U.S.).

Table 2 reports confidence intervals based on standard asymptotic theory. The lag length of the ADF test statistic is chosen by the MAIC criterion based on OLS demeaned data. Since the half-life cannot be negative, we imposed a lower bound of zero (which implies immediate adjustment). According to the table, point estimates of h^*, are around 8 to 12 quarters (2 to 3 years) for most currencies. As discussed previously, researchers used to report h_a. From table 2, notice that, due to the absence of the correction factor, this procedure generally underestimates the true half-life.[13]

Based on normal sampling asymptotic theory, the 95% normal confidence intervals include zero to twenty (or more) quarters for most currencies. However, the real exchange rate is a highly persistent series. In fact, table 2 also reports the conventional Augmented-Dickey Fuller tests for the real exchange rate series. As the 5% critical value is -2.89, we cannot reject the null hypothesis that there is a unit root in any of the currencies considered, and thus that the PPP does not hold. The table also reports the DF-GLS efficient test statistic of Elliott, Rothemberg and Stock (1996), which is more powerful to reject the null hypothesis of a unit root. However, even this test does not reject the hypothesis that the real exchange rate is not mean reverting. But, although one cannot reject that the half-life can be infinity, these results do not determine how low the lower bound for the estimated half-life can be. To answer this question, the remainder of the paper focuses on the construction of confidence intervals that are robust to highly persistent variables.

Insert Table 2

Table 3 shows the estimated confidence intervals for h^* and h_a based on Stock's method. We reported only the lower bounds of the confidence intervals because the upper bounds were infinity for all currencies. For a few currencies, also the median half-life is infinity. However, the uncertainty over the estimated half-life is so big that a half-life of 6 to 8 quarters is compatible with the observed data for almost all currencies as well.[14] Thus, on the one hand, the upper bounds can be as high as infinity, invalidating the PPP. On the other hand, the lower bounds are compatible with a time horizon in which prices may be sticky and, thus, horizons in which deviations from PPP might be explained in the light

of monetary and financial shocks. The confidence intervals are thus too wide to provide conclusive support in favor of the PPP.

Insert Tables 3 and 4

Table 4 reports the confidence intervals based on the Elliott and Stock (2001) and Hansen (1999) methods. According to our estimates, the lower bounds for the half-life deviations from PPP are between 4 to 7 quarters for most currencies, except for a few outliers (Canada and Australia in particular). However, overall, this additional evidence delivers confidence intervals of roughly the same magnitude as Stock's method, thus confirming the previous results. It is disappointing that the confidence intervals are not much shrunk by the inversion of more powerful tests, but this might be due to the fact that we imposed an upper bound on the half-life. Also, given the non-linearity of the transformation between c and δ, it is not clear that good confidence intervals for c imply good confidence intervals for δ as well.

6. Conclusions

The objective of this paper was to propose a better approximation to the half-life for highly persistent processes in the presence of small samples and to use it to build confidence intervals for half-life deviations from PPP. We first showed that normal sampling methods for constructing confidence intervals might be unreliable when variables are highly persistent and the sample size is small. As the empirical evidence suggests that the real exchange rate is such a variable, we then consider methods that are robust to persistence. These methods (Stock, Elliott and Stock, and Hansen) imply confidence intervals that cover half-lives as low as 5 to 8 quarters for most real exchange rates considered in this paper (although some countries are outliers).

Overall, the results of this paper suggest that the real exchange rate is a highly persistent variable and that the confidence intervals for the half-life are extremely wide. On the one hand, they are compatible with processes that can halve in 4 to 8 quarters, which are compatible with PPP. On the other hand, they do not rule out the possibility of an infinite half-life. The empirical results thus show that the data are not informative enough to

support any alternative hypothesis regarding the half-life based on the tests used in this paper. This is likely due to the limited power of the unit root tests and the short data we use. The conclusion is similar to that in Murray and Papell (2002) and Kilian and Zha (2001). In Kilian and Zha (2001), for example, the Bayes factors, which provide useful results beyond what a confidence or posterior probability interval gives, similarly imply that the data are not informative enough. More powerful tests might be able to provide more empirical support for the PPP. For example, Elliott and Pesavento (2001) apply higher power tests for mean reversion against close alternatives by exploiting information on other economic variables.

Some researchers have recently addressed the question of size distortions of tests in the presence of persistent real exchange rates and their implications for the PPP debate (see Engel, 2000 and Caner and Kilian, 2001). However, Caner and Kilian focus on size problems of tests of the null hypothesis of stationarity versus the alternative of a unit root. Engel (2000) discusses both tests on a unit root and tests on stationarity, but he focuses on simulations (calibrated on real data) to show the existence of possible size distortions in the former and low power in the latter.

This paper assumes a linear data generating process. Recent research points to non-linearities as a candidate for explaining the apparent high persistence and excess volatility of exchange rates (e.g. Taylor (2001), Shintani (2002)). Also, this paper does not address panel issues, nor it considers longer datasets that merge floating and fixed exchange rate samples. We would expect panel tests and tests based on longer data series to have more power to reject a unit root and, hence, confidence intervals obtained by inverting these tests to strengthen the empirical evidence in favor of PPP. However, this would require additional assumptions on the joint distribution and a careful investigation of cross sectional dependence (O'Connell (1998a)). Finally, it might be worth investigating potential aggregation biases (as in Imbs, Mumtaz, Ravn and Rey (2002)) and the reasons why some countries' half-lives are much higher than those of other countries. However, all these questions are left for future research.

References

Abuaf, Niso and Jorion, Philippe, "Purchasing power parity in the long run", *Journal of Finance*, 45(1), March 1990.

Andrews, D. W. K., "Exactly median-unbiased estimation of autoregressive/unit root models", *Econometrica*, 61, 1993, pp. 139-166.

Andrews, D. W. K., and Chen, H.-Y., "Approximately median-unbiased estimation of autoregressive models", *Journal of Business and Economic Statistics*, 12, 1994, pp. 187-204.

Caner, Mehmet, and Kilian, Lutz, "Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate", *Journal of International Money and Finance* 20(5), October 2001.

Cheung, Yin-Wong and Kon Lai, "On the purchasing power parity puzzle", *Journal of International Economics* 52, 2000, pp. 321-330.

Elliott, Graham, Rothemberg, Thomas, and Stock, James, "Efficient tests for an autoregressive unit root", *Econometrica,* Vol. 64(4), July 1996, pp. 813-36.

Elliott, Graham, and Stock, James, "Confidence intervals for autoregressive coefficients near one", *Journal of Econometrics* 103, 2001, pp. 155-181.

Elliott, Graham and Elena Pesavento, "Higher power tests for bilateral failure of PPP after 1973", *mimeo*, 2001.

Engel, Charles, "Long-run PPP may not hold after all", *Journal of International Economics* 51, August 2000.

Frankel, Jeffrey, and Rose, Andrew, "A panel project on purchasing power parity: mean reversion within and between countries", *Journal of International Economics* 40, 1996.

Gospodinov, Nikolay, "Asymptotic confidence intervals for impulse responses for near-integrated processes: an application to Purchasing Power Parity", *mimeo*, 2002.

Hamilton, James, *Time series analysis*, Princeton: Princeton University Press, 1994.

Hansen, Bruce, "Bootstrapping the autoregressive model", *Review of Economics and Statistics 81(4)*, November 1999, pp.594-607.

Inoue, Atsushi, and Kilian, Lutz, "Bootstrapping Autoregressions with Possible Unit Roots", *Econometrica* 70(1), January 2002, pp. 377-91.

Kilian, Lutz, and Zha, Tao, "Quantifying the Uncertainty about the Half-Life of Deviations from PPP", *Journal of Applied Econometrics* 17(2), March-April 2002.

Imbs, Jean, Mumtaz, Haaron, Ravn, Morten and Rey, Helene, "PPP strikes back: aggregation and real exchange rate dynamics", *mimeo,* 2002.

Lopez, Claude, Murray, Christian and David Papell, "State of the Art Unit Root Tests and the PPP Puzzle", *mimeo,* August 2003.

Lothian, James, and Taylor, Mark, "Real exchange rate behavior: the recent float from the perspective of the past two centuries", *Journal of Political Economy* 104(3), 1996.

Mark, Nelson M., *International Macroeconomics and Finance: Theory and Empirical Methods*, Blackwell Publishers, 2001.

Murray, Christian J., and Papell, David H., "The Purchasing Power Parity Persistence Paradigm", *Journal of International Economics* 56, 2002, pp. 1-19.

Ng, Serena, and Perron, Pierre, "Lag length selection and the construction of unit root tests with good size and power", *Econometrica* 69(6), November 2001, pp. 1519-54.

O' Connell, P.G.J., "The overvaluation of purchasing power parity", *Journal of International Economics* 44, 1998(a), pp. 1-19.

Phillips, Peter C.B., "Impulse response and forecast error variance asymptotics in nonstationary VARs", *Journal of Econometrics* 83, 1998, pp. 21-56.

Rogoff, Kenneth, "The purchasing power parity puzzle", *Journal of Economic Literature,* June 1996, pp. 647-68.

Shintani, Mototsugu, "A non-parametric measure of convergence toward PPP", *mimeo*, 2002.

Sims, Christopher, James Stock and Mark Watson, "Inference in linear time series models with some unit roots", *Econometrica* 58, 1990, pp. 113-144.

Stock, James, "Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series", *Journal of Monetary Economics* 28, 1991, pp. 435-59.

Stock, James, "VAR, error correction and pretest forecasts at long horizons", *Oxford Bullettin of Economics and Statistics,* 58(4), 1996.

Taylor, Alan, "Potential pitfalls for the purchasing-power-parity puzzle? Sampling and specification biases in mean-reversion tests of the law of one price", *Econometrica* 69(2), March 2001(a), pp. 473-98.

Taylor, Alan, "A century of purchasing power parity", *Review of Economics and Statistics*, forthcoming, 2001(b).

Figure 1
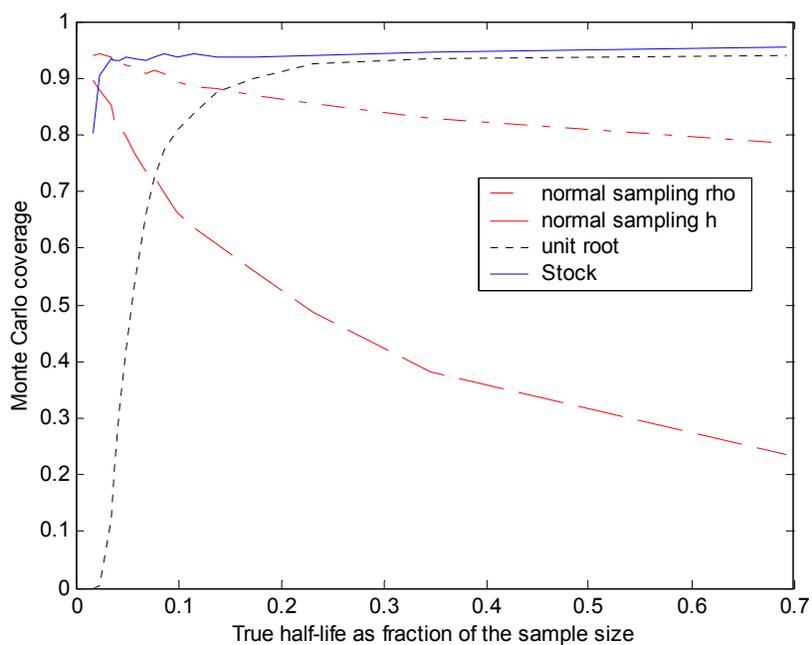
Figure 1. Comparison of coverage of various confidence intervals

as a function of the true half-life (nominal coverage=0.95).

Tables

Table 1(a). Comparison of coverage rates

h	c	Normal	Unit root	Stock	Elliott-Stock	Hansen
2	-30	0.718	0.00	0.90	0.34	0.96
6	-11.5	0.700	0.57	0.93	0.62	0.92
20	-3.5	0.468	0.91	0.94	0.84	0.96
$+\infty$	-.01	0.020	0.95	0.96	0.97	0.95

Note: The table reports, for each true half-life (and its corresponding measure of persistence in the data, c) the coverage rate of the confidence intervals based on: approximate normal sampling distribution ("Normal"); unit root test ("Unit root"); Stock (1991) method ("Stock"); Elliott and Stock (2001) method ("Elliott-Stock"); Hansen (1999) grid-α bootstrap method ("Hansen"). Ideally, these coverage rates should be close to 0.95. In the case of multiple intersections between the estimated test statistic and the critical values, we convexified the confidence interval. In the case of Stock's method, we simulated the critical values of the ADF test statistic outside the range considered by Stock (1991). The number of Monte Carlo simulations was 5000 for all methods except for Hansen, which is computationally intensive, so we chose 100 Monte Carlo replications only. T=100.

Table 1(b). Comparison of actual confidence interval type I error

c	λ_2	λ_3	λ_4	λ_5	h	$Stock$	ADF	h_∞
-5	0.2	–	–	–	19	0.11	0.11	0.14
-5	0.6	–	–	–	33	0.12	0.12	0.13
-5	0.9	–	–	–	71	0.14	0.22	0.11
-8	0.2	–	–	–	12	0.10	0.10	0.22
-8	0.6	–	–	–	21	0.13	0.12	0.21
-8	0.9	–	–	–	50	0.13	0.22	0.13
-11.5	0.2	–	–	–	8	0.11	0.10	0.40
-11.5	0.6	–	–	–	15	0.12	0.12	0.33
-11.5	0.9	–	–	–	39	0.13	0.21	0.17
-5	0.6	0.7	–	–	60	0.12	0.26	0.11
-8	0.6	0.7	–	–	39	0.13	0.33	0.16
-11.5	0.6	0.7	–	–	28	0.14	0.41	0.22
-5	0.4	-0.4	–	–	18	0.10	0.10	0.12
-8	0.4	-0.4	–	–	11	0.10	0.10	0.22
-11.5	0.4	-0.4	–	–	8	0.10	0.10	0.37
-5	-0.68	0.15	0.15	0.09	12	0.10	0.11	0.11
-5	0.3	0.3	-0.5	-0.5	12	0.10	0.10	0.11
-8	-0.68	0.15	0.15	0.09	7	0.13	0.15	0.18
-8	0.3	0.3	-0.5	-0.5	8	0.08	0.10	0.18
-11.5	-0.68	0.15	0.15	0.09	5	0.16	0.20	0.27
-11.5	0.3	0.3	-0.5	-0.5	4	0.37	0.43	0.28

Table 1(c). Comparison of actual confidence interval type I error

(upper bound of the confidence interval) for various values of T

	$T =$	250			$T =$	500	
ρ	λ_2	λ_3	h_∞	ρ	λ_2	λ_3	h_∞
0.95	0.2	–	0.46	0.95	0.2	–	0.95
0.92	0.2	–	0.84	0.92	0.2	–	1
0.885	0.2	–	0.98	0.885	0.2	–	1
0.95	0.6	0.7	0.35	0.95	0.6	0.7	0.85
0.92	0.6	0.7	0.58	0.92	0.6	0.7	0.98
0.885	0.6	0.7	0.79	0.885	0.6	0.7	0.99
0.95	0.4	-0.4	0.44	0.95	0.4	-0.4	0.94
0.92	0.4	-0.4	0.79	0.92	0.4	-0.4	0.99
0.885	0.4	-0.4	0.96	0.885	0.4	-0.4	1

Notes to Tables 1(b) and 1(c): Monte Carlo simulations (5,000 replications) are based on the AR(p) process (p=2,3,4,5) described by (8). T=100 in Table 1(b), whereas the value of T is specified in Table 1(c). The table reports the actual confidence interval type I error based on: Stock (1991) method for h* (labeled "Stock") and for h_a (labeled "ADF"). The table also reports the true half-life, h, and the empirical fraction of times in which the upper bound of the confidence interval based on Stock (1991) method for h^* rejects an infinite half-life – i.e. it rejects a unit root (labeled h_∞). The lower bound was almost always below infinity. When $\lambda_i = 0 \ \forall i$, coverage is close to nominal for all methods. The lag length is treated as known in all methods. Ideally, these percentages should be close to 0.10. In Table 1(c), we report only a representative subset of the parameters examined in Table 1(b).

Table 2 Confidence intervals based on standard asymptotics and ADF tests

	$\widehat{\alpha}\left(1\right)$	ADF	ADF_{GLS}	$\widehat{h}_a$	$\left(\widehat{h}_a^l,\widehat{h}_a^u\right)$	$\widehat{h}^*$	$\widehat{\gamma}$
Austria	0.936	-1.8	-1.94	10.4	(0; 22.1)	10.4	- -
Australia	0.962	-1.4	-.264	18	(0; 43)	18	- -
Belgium	0.948	-1.55	-2.1	13	(0; 30)	13.3	.35
Canada	0.99	-.588	-0.37	66.6	(0; 134)	66.6	- -
Denmark	0.942	-1.71	-2.39	11.6	(0; 25.3)	11.5	.4
Finland	0.89	-3	-2.1	5.96	(1.83; 10.1)	5.16	.5
France	0.932	-1.83	-1.97	9.92	(0; 20.9)	9.92	- -
Germany	0.936	-1.81	-1.9	10.6	(0; 26.2)	10.6	- -
Greece	0.919	-2.32	-2.04	8.17	(0.97; 15.4)	8.47	.5
Italy	0.918	-2.18	-2.03	8.1	(0.48; 15.7)	8.54	.17
Japan	0.93	-2.48	-1.2	9.56	(1.72; 17.4)	9.96	.4
Netherl.	0.922	-2.03	-1.68	8.55	(0; 17.1)	8.9	.1
Norway	0.931	-1.99	-2.12	9.72	(0; 19.6)	9.72	- -
Spain	0.956	-1.51	-1.64	15.3	(0; 35.6)	15.3	- -
Sweden	0.945	-1.99	-1.24	12.2	(1.14; 17.1)	12.5	.7
Switzerl.	0.905	-2.48	-1.87	6.97	(0.28; 15.6)	7.13	.06
U.K.	0.897	-2.72	-1.65	6.39	(1.53; 11.2)	6.73	.16

Note: For each bilateral real exchange rate we report: the estimated test statistic of the demeaned Augmented Dickey-Fuller regression (ADF), the estimated coefficient of the lagged regressor ($\widehat{\alpha}\left(1\right)$ as defined in (16)) and the DF-GLS test proposed by Elliott, Rothemberg and Stock, 1996 (ADF_{GLS}). The lag lengths are selected by the MAIC criterion based, respectively, on the OLS and on the GLS detrending methods proposed by Ng and Perron (2001). $\widehat{h}_a$, $\widehat{h}^*$ are the estimates of the half-life from (17) and (14); $(\widehat{h}_a^l,\widehat{h}_a^u)$ is based on (21). The 5% critical value of both the ADF and ADF_{GLS} test statistics is -2.89. $\widehat{\gamma}$ is the absolute value of the estimated second largest root (significantly less than one for all countries).

Table 3 Confidence intervals based on Stock (1991)

	$lags$	$(\widehat{c}_l, \widehat{c}_u)$	$\widehat{c}_l^{05}$	$\widehat{h}_{.05}^a$	$\widehat{h}_{median}^a$	$\widehat{b}(1)$	$\widehat{h}_{.05}^*$	$\widehat{h}_{median}^*$
Austria	1	(-12.5; 3.67)	-10.93	6.34	29.96	1.00	6.34	29.96
Australia	1	(-8.87; 4.13)	-7.53	10.67	$+\infty$	1.00	10.67	$+\infty$
Belgium	3	(-9.91; 3.99)	-8.53	8.59	$+\infty$	0.95	8.77	$+\infty$
Canada	1	(-3.06; 4.74)	-1.75	45.97	$+\infty$	1.00	45.97	$+\infty$
Denmark	3	(-11.4; 3.79)	-9.98	7.91	64.31	1.02	7.84	63.74
Finland	5	(-27.9; 0.799)	-25.71	8.00	13.80	0.34	6.93	11.95
France	1	(-12.8; 3.62)	-11.23	6.17	26.48	1.00	6.17	26.48
Germany	1	(-12.5; 3.66)	-10.97	6.32	29.50	1.00	6.32	29.50
Greece	5	(-18.4; 2.7)	-16.67	7.14	15.80	0.58	7.40	16.39
Italy	2	(-16.7; 2.95)	-14.97	5.64	13.84	0.82	5.95	14.60
Japan	5	(-20.4; 2.32)	-18.60	7.22	14.73	0.60	7.52	15.35
Netherl.	2	(-15; 3.22)	-13.35	5.88	16.72	0.88	6.12	17.41
Norway	1	(-14.6; 3.3)	-12.93	6.22	18.66	1.00	6.22	18.66
Spain	1	(-9.52; 4.05)	-8.15	8.50	$+\infty$	1.00	8.50	$+\infty$
Sweden	5	(-14.5; 3.3)	-12.89	11.31	34.10	0.55	11.59	34.95
Switzerl.	2	(-20.4; 2.32)	-18.59	4.59	9.36	0.94	4.69	9.57
U.K.	2	(-23.8; 1.65)	-21.86	4.44	8.27	0.83	4.68	8.71

Note. For each bilateral real exchange rate we run a demeaned ADF regression. The median-unbiased two-sided and one-sided confidence intervals for c, denoted respectively by $(\widehat{c}_l, \widehat{c}_u)$ and $\widehat{c}_l^{05}$, are obtained directly by inverting Stock's table A1 (with a linear interpolation from its grid values). We report one-sided lower bounds for the median unbiased confidence intervals for the half-life (h) with coverage 0.95, denoted by subscript .05. Superscripts $*$ and a denote measures of the half-life obtained by multiplying (13) and (18) (based on $\widehat{c}_l^{05}$), by T respectively, where T is the sample size. Upper bounds were $+\infty$ for all currencies so they were not reported. h_{median} is the median unbiased estimate of the half-life (based on the median unbiased estimate of c).

Table 4 Confidence intervals based on Elliott and Stock (2001) and Hansen (1999)

	Elliott	& Stock	Hansen	
	$\widehat{h}^a_{.05}$	$\widehat{h}^*_{.05}$	$\widehat{h}^a_{.05}$	$\widehat{h}^*_{.05}$
Austria	5.87	5.87	6.07	6.79
Australia	61.28	61.28	9.88	11.54
Belgium	4.13	4.22	7.59	8.55
Canada	60.63	55.59	17.60	23.25
Denmark	3.21	3.18	6.72	7.62
Finland	10.23	8.86	3.53	3.80
France	5.10	5.10	5.83	6.61
Germany	4.50	4.64	6.48	7.41
Greece	6.88	7.13	5.13	5.74
Italy	5.61	5.92	5.30	6.11
Japan	21.93	22.85	6.27	6.84
Netherl.	7.40	7.70	5.34	6.15
Norway	4.48	4.48	5.93	6.60
Spain	7.54	7.54	8.42	9.98
Sweden	26.83	11.02	5.97	6.41
Switzerl.	4.04	4.53	4.89	5.48
U.K.	8.09	8.52	4.40	4.81

Note. For each bilateral real exchange rate we run a demeaned ADF regression, where the lag length is chosen according to MAIC. We report the 1-sided confidence intervals for the half-life (h) with coverage 0.95, denoted by subscript .05. Superscripts $*$ and a denote measures of the half-life in equations (14) and (17) respectively. Upper bounds were infinity for all currencies so they were not reported.

Notes

[1]The bilateral nominal exchange rate is defined here as the price of the foreign country's currency in terms of the home country's currency.

[2]We generalized the process to be an AR(p) in the empirical estimation and in the discussion in the text. However, for simplicity, we introduce the concept here by using an AR(1).

[3]These point estimates are introduced to motivate the paper but no longer represent the current status of our knowledge. For example, Murray and Papell (2002) note that, after accounting for serial correlation and small sample bias, these point estimates become very difficult to believe.

[4]We do not allow the presence of a deterministic time trend in the theoretical DGP, nor in the empirical estimation. The reason is that if a deterministic time trend is present, PPP in levels won't hold. If a deterministic trend is present, so that $d_t = \mu_0 + \mu_1 t$, then the calculations that follow continue to hold provided that we define a time-varying long-run equilibrium, i.e. such that the long-run equilibrium at time τ is defined as $\overline{y}_\tau = \mu_0 + \mu_1 \tau$. This is the equilibrium path that would have prevailed in the absence of the shock. The empirical results for detrended real exchange rates are similar to those reported in this paper and are available upon request.

[5]We will provide detailed empirical evidence on the degree of persistence in the bilateral exchange rates considered in this paper in the empirical section.

[6]We follow Hamilton (1994) in referring to the inverse of the roots of the polynomial $b(L)(1 - \rho L)$ as the eigenvalues (or the roots) of the DGP .

[7]Recall that, in this paper, y_t is the *logarithm* of the real exchange rate, so $y_{t+h} - \mu_0$ measures a percentage deviation.

[8]Note that this approximation assumes that there is only one root close to unity. It is possible to extend (12) to processes integrated of order higher than one. In that case, the

second component in (11) becomes asymptotically relevant as well. However, the empirical evidence (see table 2) suggests that there is only one root close to unity in real exchange rates so we specialize the result to this case. In the Monte Carlo section we provide some sensitivity results to processes with a high second largest root.

[9]In fact, when $c \geq 0$, the process is (mildly) explosive so the half-life will be infinite with our long-horizon approximation. On the other hand, when $c < 0$, $\frac{\ln\left(\frac{1}{2}b(1)\right)}{c}$ might be negative when mean reversion is considerably fast. In this case, we let $\delta^* = 0$. In this case, our method, which relies on (6), may not provide a good approximation.

[10]Note that, even if the true DGP is an AR(p), the estimated coefficient in the AR(1) regression will be a consistent estimate of the largest root because $T\left(\widehat{\rho} - \rho\right) = \left(\frac{1}{T^2}\Sigma_{t=1}^{T}y_{t-1}^2\right)^{-1}$ $\left(\frac{1}{T}\Sigma_{t=1}^{T}y_{t-1}u_t\right) = O_p(1)$.

[11]Hansen (1999) also suggested a grid$-t$ bootstrap method. However, the half-life is a non-linear transformation of the parameter $\alpha(1)$. In practice, the grid-t method requires an estimate of the variance which, if obtained by the delta-method approximation, makes the coverage quite poor (results of a Monte Carlo experiment are available upon request). Also, constructing confidence intervals by minimizing their length gave similar empirical results.

[12]In practice, the infinite half-life is 6900, corresponding to a value of $c = -0.01$.

[13]To save space, confidence intervals for h^* are not reported. They generally comprise higher values for the half-life. We chose to report confidence intervals based on h_a in order for the results to be comparable with those existing in the literature.

[14]The most notable exception is the Canadian real exchange rate. However, the data clearly show that there is a time trend in that case. Other countries for which the lower bound of the half-life is quite high are Australia and Sweden. But for most countries, the lower bound is less than 8 quarters anyway.

Appendix

Comparison of exact and approximate half-lives with an application to an AR(2) process

Let us compare three candidate measures of the long-run (l.r.) effects of a unitary shock: (i) the "*AR(1) l.r. effect*", that depends only on the largest unit root of the process and not on short-run dynamics, equal to ρ^h; (ii) the "*approximate l.r. effect*", frequently used by applied researchers, equal to $\alpha(1)^h$; (iii) the "*exact l.r. effect*", proposed in this paper, equal to $\rho^h \phi_b$, where $\phi_b \equiv b(1)^{-1}$.

To highlight the relationship between the approximate and the exact l.r. effects, we rearrange the *ADF regression* (16) as:

$$\widetilde{y}_t = \sum_{j=1}^{k} \alpha_{j-1}^* \Delta \widetilde{y}_{t-j} + \alpha(1)\,\widetilde{y}_{t-1} + \epsilon_t \tag{22}$$

where $\widetilde{y}_t \equiv y_t - dt = u_t$ measures deviations of y_t from its long-run equilibrium value. We can rewrite (22) as a VAR(1):

$$Y_t = \mathbf{A} Y_{t-1} + e_t \tag{23}$$

$$\mathbf{A} \equiv \left(\begin{array}{c|cccc} \alpha(1) & \alpha_0^* & \alpha_1^* & ... & \alpha_{k-1}^* \\ \hline \alpha(1)-1 & \alpha_0^* & \alpha_1^* & ... & \alpha_{k-1}^* \\ 0 & 1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & 0 & 1 & 0 \end{array} \right) = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right) \tag{24}$$

(where $A_{11} = \alpha(1) = 1 + \frac{c}{T} b(1)$ is a scalar and the rest of the matrix $\mathbf{A}$ is partitioned accordingly), $Y_t \equiv [\widetilde{y}_t, \Delta\widetilde{y}_t, \Delta\widetilde{y}_{t-1}, ... \Delta\widetilde{y}_{t-k+1}]'$ is a $(k+1) \times 1$ vector, $\Delta\widetilde{y}_{t-j} \equiv \widetilde{y}_{t-j} - \widetilde{y}_{t-j-1}$, $e_t = [\epsilon_t, \epsilon_t, \mathbf{0}_{1\times(k-1)}]'$ is a $(k+1) \times 1$ vector and $\mathbf{0}_{1\times(k-1)}$ is a $1 \times (k-1)$ vector of zeros.

An alternative approach is to follow Stock (1991) and Phillips (1998) in rewriting the *ADF* regression in terms of the canonical regressors, let's call it the *ADF canonical regression*:[15]

$$\widetilde{y}_t = \rho\widetilde{y}_{t-1} - \sum_{j=1}^{k} b_j \widetilde{\Delta}\widetilde{y}_{t-j} + \epsilon_t \tag{25}$$

which can be rewritten in a VAR(1) format:

$$\widetilde{Y}_t = \mathbf{E}\widetilde{Y}_{t-1} + e_t \tag{26}$$

$$\mathbf{E} \equiv \left(\begin{array}{c|cccc} \rho & -b_1 & -b_2 & ... & -b_k \\ \hline \rho - 1 & -b_1 & -b_2 & ... & -b_k \\ 0 & 1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & 0 & 1 & 0 \end{array} \right) = \left(\begin{array}{c|c} E_{11} & E_{12} \\ \hline E_{21} & E_{22} \end{array} \right) \tag{27}$$

(where $E_{11} = \rho$ and the rest of the matrix $\mathbf{E}$ is partitioned accordingly), $\widetilde{Y}_t \equiv [\widetilde{y}_t, \widetilde{\Delta}\widetilde{y}_t, \widetilde{\Delta}\widetilde{y}_{t-1},$ $...\widetilde{\Delta}\widetilde{y}_{t-k+1}]'$ is a $(k+1) \times 1$ vector and $\widetilde{\Delta}\widetilde{y}_{t-j} \equiv \widetilde{y}_{t-j} - \rho\widetilde{y}_{t-j-1}$. We also find it convenient to define $\underline{\mathbf{1}}_k$ to be the first column of the $(k \times 1)$ identity matrix and I_k to be the identity matrix with k elements.

Suppose we start at time t-1 in the long run equilibrium $\overline{y}_{t-1}^{LR}$ and at time t there is a shock ϵ_t. The initial deviation from equilibrium is thus:

$$\widetilde{y}_t = \underline{\mathbf{1}}_k' e_t = \epsilon_t \tag{28}$$

By using (26), the effect of the shock in the subsequent periods becomes:

$$\widetilde{y}_{t+1} = \underline{\mathbf{1}}_k' \mathbf{E}[1, 1, \mathbf{0}_{1\times(k-1)}]' \epsilon_t \tag{29}$$

$$...$$

$$\widetilde{y}_{t+h} = \underline{\mathbf{1}}_k' \mathbf{E}^h [1, 1, \mathbf{0}_{1\times(k-1)}]' \epsilon_t$$

Hence, after h periods, the percentage deviation from equilibrium relative to the initial percentage deviation from equilibrium is:

$$\frac{\widetilde{y}_{t+h}}{\widetilde{y}_t} = \underline{\mathbf{1}}_k' \mathbf{E}^h [1, 1, \mathbf{0}_{1\times(k-1)}]' \tag{30}$$

This measures the effect of a shock ϵ_t after h periods. The usefulness of the VAR(1) representation above is that, since $\rho - 1 = c/T$, it follows that $E_{21} \simeq \mathbf{0}_{k\times 1}$, so that $\mathbf{E}$ (as partitioned) is asymptotically an upper diagonal matrix. As a consequence:

$$\mathbf{E}^h = \left(\begin{array}{c|c} E_{11}^h & E_{11}^{h-1}E_{12}E_{22}^0 + E_{11}^{h-2}E_{12}E_{22} + ... + E_{11}^0 E_{12}E_{22}^{h-1} \\ \hline \mathbf{0}_{k\times 1} & E_{22}^h \end{array} \right) \tag{31}$$

$$= \left(\begin{array}{c|c} E_{11}^h & E_{11}^{h-1}E_{12}\left(I_k + E_{22}/E_{11} + ... + (E_{22}/E_{11})^{h-1} \right) \\ \hline \mathbf{0}_{k\times 1} & E_{22}^h \end{array} \right)$$

$$= \left(\begin{array}{c|c} E_{11}^h & E_{11}^{h-1}E_{12}\left(I_k - E_{22}/E_{11} \right)^{-1}\left(I_k - (E_{22}/E_{11})^h \right) \\ \hline \mathbf{0}_{k\times 1} & E_{22}^h \end{array} \right)$$

As $h \to \infty$, $E_{22}^h \to \mathbf{0}_{k\times k}$ (because all of its roots are in absolute value less than one) and $E_{11}^h = \rho^h \to e^{c\delta}$ (so it is bounded asymptotically), which imply that $(E_{22}/E_{11})^h$ also vanishes asymptotically. Hence, the effect of the shock on the first component of Y_t after h periods, $\mathbf{E}^h[1, 1, \mathbf{0}_{1\times(k-1)}]'$, will be equal to:

$$E_{11}^h + E_{11}^{h-1}E_{12}\left(I_k - E_{22}/E_{11} \right)^{-1} \tag{32}$$

$$= E_{11}^h \left(1 + E_{12}\left(E_{11}I_k - E_{22} \right)^{-1}\mathbf{1}_k \right)$$

$$\simeq E_{11}^h \left(1 + E_{12}\left(I_k - E_{22} \right)^{-1}\mathbf{1}_k \right) = E_{11}^h \phi_b$$

where in the last line we use the approximation that $E_{11} = \rho = 1 + \frac{c}{T} \simeq 1$. Thus:

$$\left(\frac{\partial \widetilde{y}_{t+h}}{\partial \epsilon_t} \right)_{\text{canonical}} = E_{11}^h \phi_b \tag{33}$$

where ϕ_b is the correction factor:[16]

$$\phi_b \equiv 1 + E_{12}\left(I_k - E_{22} \right)^{-1}\mathbf{1}_k \tag{34}$$

We could repeat the same reasoning for the matrix $\mathbf{A}$. By doing the same calculations, we find that, for the ADF regression (22), the l.r. effect is:

$$\left(\frac{\partial \widetilde{y}_{t+h}}{\partial \epsilon_t} \right)_{ADF} = A_{11}^h \phi_{\alpha^*} \tag{35}$$

where the correction factor is:

$$\phi_{\alpha^*} \equiv 1 + A_{12}\left(I_k - A_{22} \right)^{-1}\mathbf{1}_k \tag{36}$$

To highlight the differences between the two results, we introduce a simple example. We consider an AR(2) process without deterministic components:

$$a\left(L\right)y_t = \epsilon_t \tag{37}$$

where $a\left(L\right) = \left(1 - \rho L\right)\left(1 - \gamma L\right)$. Notice that ρ and γ are the roots and, by assumption, $\left|\gamma\right| << 1$ and $\rho = 1 + \frac{c}{T}$. Note that $a\left(L\right) = 1 - \left(\gamma + \rho\right)L + \rho\gamma L^2$ so the process can be rewritten in the familiar form:

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \epsilon_t \tag{38}$$

where:

$$\rho_1 = \gamma + \rho \tag{39}$$

$$\rho_2 = -\rho\gamma$$

The canonical ADF representation is:

$$y_t = \rho y_{t-1} + \gamma\left(y_{t-1} - \rho y_{t-2}\right) + \epsilon_t \tag{40}$$

Finally, the ADF representation is:

$$y_t = \alpha\left(1\right)y_{t-1} - \rho_2\left(y_{t-1} - y_{t-2}\right) + \epsilon_t \tag{41}$$

where $\alpha\left(1\right) = \rho_1 + \rho_2 = \gamma + \rho - \rho\gamma$.

Since $b\left(1\right) = 1 - \gamma$, the l.r. effect of the shock ϵ_t after h periods derived from the DGP representation in section 2 is:

$$\left(\frac{\partial y_{t+h}}{\partial \epsilon_t}\right)_{\text{exact}} = \rho^h\left(1 - \gamma\right)^{-1} \rightarrow e^{c\delta}\left(1 - \gamma\right)^{-1} \tag{42}$$

From (27) and (31), note that $1 + E_{12}\left(I_k - E_{22}\right)^{-1}\mathbf{1}_k = 1 + \frac{\gamma}{1-\gamma} = \left(1 - \gamma\right)^{-1}$ so:

$$\left(\frac{\partial \widetilde{y}_{t+h}}{\partial \epsilon_t}\right)_{\text{canonical}} = \rho^h\left(1 - \gamma\right)^{-1} \rightarrow e^{c\delta}\left(1 - \gamma\right)^{-1} \tag{43}$$

corresponds to the exact long-run effect. Instead, from (24) and (36), $\phi_{\alpha^*} = \left(1 + \rho_2\right)^{-1} \simeq \left(1 - \gamma\right)^{-1}$

$$\left(\frac{\partial \widetilde{y}_{t+h}}{\partial \epsilon_t}\right)_{ADF} = \alpha\left(1\right)^h\left(1 - \gamma\right)^{-1} \rightarrow e^{c\delta b(1)}\left(1 - \gamma\right)^{-1} \tag{44}$$

The reason why we think that the ADF representation and the canonical ADF representation give different answers is that in the ADF representation the regressors are not the Sims, Stock and Watson (1990) canonical regressors; thus the regressors $y_t - y_{t-1}$ will be over-differenced and, cumulated over time, this will matter asymptotically. In other words, one can rewrite (25) as:

$$\widetilde{y}_t = \rho \widetilde{y}_{t-1} - \sum_{j=1}^{k} b_j \Delta \widetilde{y}_{t-j} + \xi_t + \epsilon_t, \quad \xi_t \equiv (\rho - 1) \sum_{j=1}^{k} b_j \widetilde{y}_{t-j} \tag{45}$$

where ξ_t can be interpreted as an omitted variable in regression (22), whose effect is non-negligible asymptotically, when added over time.

Let's compare the exact long-run effect with the other two measures. What we called the "AR(1) long-run effect" is the effect of a one-time unitary shock to v_t rather than to ϵ_t. In fact, from the data generating process $y_t = \rho y_{t-1} + v_t$ we have that:

$$\frac{\partial y_{t+h}}{\partial v_t} = \rho^h \to e^{c\delta} \tag{46}$$

Since $v_t = b\left(L\right)^{-1} \epsilon_t$, the long-run effect of a shock to ϵ_t will be $\rho^h b\left(1\right)^{-1}$, which is our exact measure. What we called "approximate long-run effect", which corresponds to what most empirical researchers report, is instead $\alpha\left(1\right)^h \to e^{c\delta b(1)}$ so it will be different from the exact long-run effect unless $c = 0$ or the true process is an AR(1), for which $b\left(1\right) = 1$. Hence it seems that the use of this approximation is not justified asymptotically (under the assumptions of this paper). It will also be different from the "approximate long-run effect" calculated by using the ADF representation, which will be (see (44)):

$$\alpha\left(1\right)^h \to \frac{e^{c\delta b(1)}}{b\left(1\right)} \tag{47}$$