

# Confidence intervals for half-life deviations from Purchasing Power Parity

Barbara Rossi\*

First draft: June 6, 2001

This draft: December 2, 2001

## Abstract

According to the Purchasing Power Parity (PPP) theory, real exchange rate fluctuations are mainly caused by transitory shocks. The theory 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). Existing point estimates of half-life deviations from PPP are in the order of 3 to 5 years, too big to be reconciled with the PPP. The scope of this paper is to assess how much uncertainty there is around these point estimates. We 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 6 quarters for most currencies. With a few exceptions, the results show that the data are not inconsistent with the PPP theory, although we cannot provide conclusive evidence in favor or PPP either.

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

*JEL classification:* F300; F400.

*Acknowledgements.* The idea of the paper came out during discussions with Mark Watson, whom I thank for many valuable suggestions and continued support. I would also like to thank Bo Honore', Tim Bollerslev, Charles Engel and Christian Murray for comments. Financial support provided by the International Finance Section of Princeton University is gratefully acknowledged. All mistakes are mine.

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

E-mail: brossi@econ.duke.edu. Tel.: (919) 660 1801. Fax: (919) 684 8974.

## 1. Introduction

What determines nominal exchange rates in the long-run? According to the theory of Purchasing Power Parity (PPP), since the (bilateral) nominal exchange rate ( $E_t$ ) is the relative price of two currencies,<sup>1</sup> 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:

$$E_t = \frac{P_t}{P_t^*} \quad (1)$$

This theory, then, predicts that the relative price level should drive the dynamics of the nominal exchange rate. In other words, the logarithm of the real exchange rate, defined as  $y_t = \ln\left(\frac{E_t P_t^*}{P_t}\right)$ , should be constant. Notice that (1) requires PPP to hold 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:

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

where  $\epsilon_t$  is a white-noise. Then, at horizon  $h$ , the percentage deviation from equilibrium is  $\rho^h$ . The half-life deviation from PPP is defined as the horizon at which the percentage deviation from equilibrium is one half, that is:<sup>2</sup>

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

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)). Actually, the deviations from PPP are so persistent that researchers have started to worry whether there is reversion at all. Thus, a related strand of the literature has focused on whether

the real exchange rate behaves like a stationary series around a mean<sup>3</sup> or whether it is non-stationary and behaves like a random walk. Early tests for a unit root have been presented by Meese and Rogoff (1988), who could not reject the hypothesis that the real exchange rate was a random walk. However, as Froot and Rogoff (1995) argue, this might be an artifact of small samples and of the low power of unit-root tests against persistent alternatives. In fact, two recent papers, Taylor (2001b) and Cheung and Lai (1998) use Elliott, Rothenberg and Stock (1995) point optimal test for unit root and find that it is possible to reject the unit root hypothesis for some currencies over the modern floating period. Additional empirical evidence comes from analysis based on long data sets, where consensus estimates of  $h$  range between three to five years (see the references in Froot and Rogoff (1995) and in Rogoff (1995)). However, these data sets merge fixed and floating exchange rate data and the implications are unclear (Rogoff, 1995). In fact, during periods of nominal exchange rate stabilization, the real exchange rate is stabilized as well, so it might be easier to reject the random walk model and we would not know whether the tests are not rejecting because of the bigger sample size or because of the addition of the earlier part of the sample, when the exchange rate was stabilized.

The existing point estimates of half-life deviations from PPP are difficult to reconcile with the PPP theory. According to that theory, deviations from PPP are 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. More persistent deviations can be rationalized only if the predominant sources of shocks were real, as in the Balassa-Samuelson model. However, while the Balassa-Samuelson model can explain persistent deviations from PPP, it cannot explain the huge short-term variability in the real exchange rate, as real shocks should, theoretically, be rare events. This is the so-called “Purchasing Power Parity Puzzle”, as Rogoff (1995) called it. To discriminate among these two competing models, it is then important to correctly measure how persistent the deviations from PPP really are.

The contribution of this paper is to assess how much uncertainty there is around estimates of half-life deviations from PPP. That is, this paper aims to construct confidence intervals for  $h$ . The issue is hard to address because of the complications due to the high persistence in real exchange rates. In fact, conventional methods for constructing confidence intervals may not be reliable because the true half-life could be large relative to the available sample size, in which case the confidence intervals would be affected. The recently developed local-to-unity asymptotic framework suggests that the method of constructing confidence intervals based on classic asymptotic theory is only reliable when  $\rho \ll 1$  for any fixed  $h$ , but will break down if  $\rho$  is “close” to unity and  $h$  is “big” relative to the sample size. For these reasons, this paper considers alternative confidence intervals, whose asymptotic theory is based on letting the half-life grow to infinity at the rate of the sample size, as in Stock (1996) and Phillips (1998). How “big”  $h$  has to be in order for the conventional (normal sampling) asymptotic theory to break down is discussed in a Monte Carlo experiment.

Two recent works that address the issue of evaluating the degree of imprecision in estimated half-lives are Cheung and Lai (2000) and Murray and Papell (2001). The approach of Murray and Papell (2001) is similar in spirit to the contents of this paper, and the empirical results in the present paper are similar to theirs. However, as they discuss, measures of “...the half-life calculated on the lagged level of the real exchange rate in standard Dickey-Fuller and Augmented Dickey-Fuller regressions are no longer appropriate measures of persistence”. We will call these measures *approximate*, because they are valid only if the true data generating process is an AR(1). This paper, instead, provides a formal expression of the half-life that is valid for a general AR( $p$ ) process in the presence of a root close to unity. We call these measures *exact* (meaning that they are accurate asymptotic approximations under the assumptions of this paper). Although the methods for deriving the half-life are quite standard, there is no such result in the literature. Abuaf and Jorion (1990) provide a formal definition of half-life in the context of an AR(1) process only. Mark (2001) discusses exact 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, but did not provide a measure of the half-life for AR( $p$ ) processes. In fact, Murray

and Papell (2001) relied on confidence intervals based on the impulse-response functions obtained by using this computationally intensive method. This paper, instead, discusses *exact* measures of half-lives for  $AR(p)$  processes in the presence of high persistence. The same criticisms, i.e. approximate versus exact measures, applies to analysis of the half-life based on impulse-response functions when the lead time of the impulse-response is proportional to the sample size (see Phillips (1998) and Wright (2000)). While empirical analyses based on impulse-response functions have interesting insights regarding the nature of the adjustment process (see Cheung and Lai (2000)), these analyses usually treat the lead time of the impulse-response as fixed, so they are not robust to departures from stationarity in small samples.

Overall, the results of this paper are not inconsistent with the PPP. 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 6 quarters as their lower bound, a time interval in which deviations from PPP are compatible with nominal price and wage stickiness.<sup>4</sup> However, since we cannot rule out the possibility of an infinite median half-life, it is possible also to interpret the evidence as being simply non-informative, as Murray and Papell (2001) do.

The paper is organized as follows. The next section introduces the data generating process considered in this paper and derives the exact measure of half-life in the presence of persistence. Section 3 describes the methods used to construct the confidence intervals for  $h$ . We first describe the construction of confidence intervals based on conventional, normal sampling asymptotic theory. As the usual asymptotic theory provides inadequate approximations in small samples when the variable is highly persistent, we consider next confidence intervals based on the local-to-unity asymptotic theory, as developed by Stock (1991), which provide better approximations in small samples. The drawback of these confidence intervals is that they are large, because they invert inefficient tests for unit roots, and their coverage is less than the nominal level if the process is sufficiently stationary. In order to investigate the robustness of Stock's results, the remainder of the section explores the alternative methods proposed by Elliott and Stock (2001) and Hansen (1999). The former method delivers tighter confidence intervals, based on point optimal test statistics, whereas the latter has the advantage of having the correct coverage independently of the degree of persistence in

the data. 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) be:

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

where  $d_t = \mu_0$  is a deterministic component,<sup>5</sup>  $v_t$  is a zero mean, stationary and ergodic process, with finite autocovariances  $\gamma(k) = \mathcal{E}v_t v_{t-k}$  (where  $\mathcal{E}$  denotes the unconditional expectation operator),  $\omega^2 = \sum_{k=-\infty}^{+\infty} \gamma(k)$  is finite and non-zero and  $v_t = b(L)^{-1} \epsilon_t$ , where  $\epsilon_t$  is a martingale difference sequence with finite fourth moments and constant variance  $\sigma_\epsilon^2$  and  $b(L)$  is finite order and has  $p < \infty$  (stable) roots.

Under the local to unity asymptotic theory proposed by Bobkoski (1983), Cavanagh (1985), Phillips (1987), Chan and Wei (1987) and Stock (1991),  $\rho$  is a function of the sample size such that it stays in a decreasing neighborhood of one:

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

where  $c$  is a constant 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  $\delta$ . That is:

$$\frac{h}{T} \xrightarrow{T \rightarrow \infty} \delta \tag{6}$$

We will refer to  $\delta$  as the “*half-life as a fraction of the sample size*”.

The problem of persistence is relevant for our purposes, since we are trying to estimate the half-life of PPP deviations. That is, we are trying to estimate at which horizon the deviations from PPP are back to one-half after a shock. Under assumption (5), the process is highly persistent,<sup>6</sup> so won’t be able to estimate consistently how close to unity the true

parameter is in a neighborhood  $1/T$  of the null value. At first sight, this might seem an irrelevant problem, or at least a problem with irrelevant practical consequences, as the object of interest is an estimate of the persistence, and not how close that is to unity in a neighborhood  $1/T$  of the null value. In other words, one might object that, by looking in a neighborhood  $1/T$ , we are using a sophisticated microscope when there is no need. But the half-life, by looking  $h$ -steps ahead – where  $h$  can be “big” relative to the sample size – operates exactly like a telescope and blows up the process in that neighborhood. In fact, 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,  $\rho^h$ , and, under assumption (6):

$$\rho^h \xrightarrow{T \rightarrow \infty} e^{c\delta} \quad (7)$$

Thus, how close the process is to unity, measured by  $c$ , is asymptotically relevant. Furthermore,  $c$  cannot be consistently estimated, nor can  $e^{c\delta}$ . That means that even small deviations from unity can have a huge impact on the estimated confidence intervals and that standard methods for constructing confidence intervals can be very unreliable.

In order to derive an exact expression for the half-life in this general AR( $p$ ) process, we need to derive an expression for the effect of the shock  $\epsilon_t$  on  $y_t$  after  $h$  periods. One way is to derive it in terms of the eigenvalues of the process. We factorize (4) as:

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

where, for convenience,  $\lambda_1 = \rho$  is the root close to unity and  $\lambda_2, \lambda_3, \dots, \lambda_p$  are the (stable) 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,  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_p]'$ . We assume that the eigenvalues are distinct and real. Suppose we start at time  $t - 1$  in the long run equilibrium  $\mu_0$ , and at time  $t$  there is a shock  $\epsilon_t$ . No other shocks hit the economy subsequently. The shock  $\epsilon_t$  measures the initial deviation from equilibrium, which we denote by  $\tilde{y}_t \equiv y_t - \mu_0 = \epsilon_t$ . It follows that the deviation from equilibrium after  $h$  periods will be  $\tilde{y}_{t+h} = \mathbf{c}' \boldsymbol{\lambda}^h \epsilon_t$ , where  $\mathbf{c}$  is a  $(p \times 1)$  vector with generic element:

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

and  $\boldsymbol{\lambda}^h$  is the  $(p \times 1)$  vector containing all the eigenvalues to the  $h$  power. After  $h$  periods, the percentage deviation from equilibrium relative to the initial percentage deviation from equilibrium is:<sup>7</sup>

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

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

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

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 (7) and (6), as  $T \rightarrow \infty$  we have that  $\lambda_1^{h+p-1} = \rho^{h+p-1} = (1 + \frac{c}{T})^{h+p-1} \rightarrow e^{c\delta}$  (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:

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

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:

$$\delta_e \equiv \frac{\ln(\frac{1}{2}\phi_b^{-1})}{c}, \quad \phi_b \equiv b(1)^{-1} \quad (13)$$

where we call  $\phi_b$  the “correction factor”. It follows that an asymptotically correct (which we will call *exact*) measure of the half-life is:

$$h_e \equiv \frac{\ln(\frac{1}{2}\phi_b^{-1})}{\ln \rho} \quad (14)$$

Note the monotonicity of the relationship between  $\delta_e$  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 existing in the literature. It seems an established empirical practice to either run 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 to run the regression in *ADF* form and calculate 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 former approach still gives a consistent estimate of the parameter  $\rho$  but the estimate of the half-life will be biased because it ignores short-run dynamics. In fact, in this case the estimate, let's call it  $h_{AR(1)}$ , would ignore the correction factor:

$$h_{AR(1)} \equiv \frac{\ln\left(\frac{1}{2}\right)}{\ln \rho} \quad (15)$$

so it will be wrong unless  $b(1) = 1$ . In other words, (15) is not the population value of interest. Furthermore, the parameter  $\rho$  cannot be consistently estimated in a neighborhood  $1/T$  of the true value so the estimate of the half-life as a fraction of the sample size will be *inconsistent*. 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:<sup>8</sup>

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

where  $\alpha(1) = 1 + \frac{c}{T}b(1)$ ,  $\tilde{\mu}_0 = -\frac{c}{T}\mu_0 b(1)$  and  $\alpha_j^* = -\sum_{i=j+1}^k \alpha_i$ . The researcher then calculates the half-life based on the AR(1) model (3) as:

$$h_a = \frac{\ln\left(\frac{1}{2}\right)}{\ln \alpha(1)} \quad (17)$$

which we call *approximate* half-life, and which corresponds to an approximate half-life as a fraction of the sample size equal to:

$$\delta_a = \frac{\ln\left(\frac{1}{2}\right)}{cb(1)} \quad (18)$$

Expression (17) has also been proposed by Andrews (1993) for the simple AR(1) case. We call this measure approximate because, 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 ( $\rho$ ) in (2). By comparing (18) with (13), it is clear that this measure is not correct under the assumptions of this paper, unless the true data generating process is an AR(1) – in which case  $b(1) = 1$ . That is, it is not the population value of interest. We discuss the issue more thoroughly in the appendix. However, in the empirical section we also report estimates of the approximated half-life, as it is commonly used by empirical researchers.<sup>9</sup>

### 3. Econometric methods

So far, we discussed why and when we think that the existing measures of half-lives can be misleading and why they may not be consistently estimated. Let us now discuss 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, so the correction factor can be consistently estimated by:

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

Second, as argued above,  $c$  cannot be consistently estimated, nor can  $\rho$  in a neighborhood  $1/T$  of the true value. However, there are a variety of methods for constructing confidence intervals for the root close to unity,  $\rho$ . 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, the empirical literature estimates the (approximated) half-life to be:

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

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:

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

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). That is, he focused

on the parameter  $\rho$ , because he was interested in tests for unit roots. In our case, we are interested in constructing confidence intervals for the half-life. 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  $\delta$  is a monotone decreasing function of  $c$ . Hence, it is 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 for persistent time series based on Elliott and Stock (2001) method

In empirical applications, Stock's method delivers 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  $H_0 : c = c^*$ . The *lower* quantile (whose critical value is the *upper* line in figure 1) corresponds to the point optimal test of  $H_0 : c = c^*$  versus  $H_A : c = c^* + \bar{c}_L < c^*$ . The upper quantile (whose critical value is the lower line in figure 1) corresponds to the point optimal test of  $H_0 : c = c^*$  versus  $H_A : c = c^* + \bar{c}_U > c^*$ . It is a sequence of test statistics because for each  $c^*$  we choose  $\bar{c}_L(c^*)$  and  $\bar{c}_U(c^*)$  in order to have good power properties (in practice, we chose them such that the point optimal test is tangent to the power envelope when the power was equal to 0.5 – which ensured that the power of the test statistics were indistinguishable from the power envelope). We chose the coverage of the confidence interval constructed in this way to be at least 95%.

Figure 1 shows how the confidence interval is constructed for the Yen/\$ (demeaned) exchange rate data. Figure 1 shows both the point optimal test statistic critical values for testing  $c = c^*$  versus  $c < c^*$  (called PTcritical low) and the point optimal test statistic critical values for testing  $c = c^*$  versus  $c > c^*$  (called PTcritical up). The figures also report their estimated values (called, respectively, PThat low and PThat up). 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). The confidence interval is the set of  $c$ 's such that both the lower and the upper test statistic do not reject the null hypothesis, that is where the estimated test statistic is bigger than its critical value. The confidence interval associated with the lower point optimal test statistic is  $c \in [2, -\infty)$  and that associated with the upper is  $c \leq -15.69$ . Hence, the joint confidence interval is  $c \in [2, -15.69]$ , which, by monotonicity of (13), translates into a confidence interval for  $\delta$ :  $\delta \in [.042, +\infty)$ . By (14), it follows that  $h \in [4.2, +\infty)$ .

Insert Figure 1

#### (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. While it is true that the bootstrap is valid for inference about the non-unit root parameters, as Inoue and Kilian (2001) show, it is also true that it fails when the parameter of interest is the unit root parameter, and that is the parameter relevant for our purposes, as discussed in section 1.

A bootstrap method that is valid in the presence of highly persistent variables is Hansen's (1999) grid- $\alpha$  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  $\hat{\alpha}(1)$ . Furthermore, by the range-preserving

property of the grid- $\alpha$  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.<sup>10</sup>

We have now described all the methods that we will use in the empirical application.<sup>11</sup> Before turning to the empirical results (presented in section 5), in the next section we consider a small Monte Carlo experiment to compare their performance in small sample sizes (i.e. for the sample size that we have available, which is 100 observations).

## 4. A small Monte Carlo experiment

To evaluate the performance of the methods used to construct confidence intervals in the paper, we perform a small Monte Carlo experiment. The experiment is useful to address the question: “how big does the true half-life have to be in order for the  $\frac{h}{T} \rightarrow \delta$  asymptotics to be better than the usual asymptotics?”. We will suggest that the normal sampling confidence intervals have good coverage as long as  $\delta < 0.02$ ; for bigger values there will be size distortions in the tests, which will be increasing with  $\delta$ , and thus the confidence intervals will have coverage different from the nominal one. Hence, with a sample of, say, 100 quarters, a horizon of 3 quarters has to be considered “big”.

Data are generated as  $y_t = \rho y_{t-1} + \epsilon_t$ ,  $\epsilon_t \sim N(0, 1)$ ,  $\rho = (1 + \frac{c}{T})$  and we compare the coverage of the confidence intervals used in the paper for different values of the true half-life as a fraction of the sample size,  $\delta$ .<sup>12</sup> The number of Monte Carlo replications is 5000, the sample size is  $T = 100$ . Results are reported in figure 2.

Insert Figure 2

The figure compares confidence intervals based on four tests:

- the normal sampling test for testing  $H_0 : \delta = \delta^*$  (where  $\delta^*$  is the true value of  $\delta$ ) versus  $H_A : \delta \neq \delta^*$ , labeled “normal sampling h” test;
- the normal sampling test for testing  $H_0 : \rho = \rho^*$  (where  $\rho^*$  is the true value of  $\rho$ ) versus  $H_A : \rho \neq \rho^*$ , labeled “normal sampling rho”;
- the *ADF-t* unit root test for testing  $H_0 : \rho = 1$  versus  $H_A : \rho < 1$ , labeled “unit root”.

Unit root tests are equivalent to testing the null hypothesis that the half-life is infinity. That

is, the figure shows how many times (as a percentage of the replications) we do not reject the hypothesis that  $h$  is infinity.

- Stock's test, labeled "Stock".<sup>13</sup>

As expected, the coverage of the normal sampling confidence intervals is fine when the half-life is small relative to the sample size ( $\delta$  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. In fact, for  $\delta = 0.5$  then the coverage of confidence intervals based on "normal sampling  $h$ " asymptotic theory is around 0.3 and that of the "normal sampling  $\rho$ " test is around 0.8. 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. In this example, the data generating process has a unit root only when  $c = 0$ , which corresponds to  $\delta = +\infty$ . Hence, the test has the correct coverage of 0.95 for  $\delta = +\infty$ , but too high acceptance probability even when  $\delta < +\infty$ . In fact, coverage remains 0.95 even for  $\delta = 0.5$  and decreases to zero only for very small  $\delta$ .

Finally, notice that confidence intervals based on Stock method have a 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.<sup>14</sup> 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. Even more interestingly, the coverage is pretty poor relative to, say, that based on Stock (1991) method *even for very small half-lives* (e.g., when  $h = 6$ , the coverage of the confidence interval based on normal sampling asymptotics is 0.70 whereas the coverage of that based on Stock (1991) is 0.93). Similar results hold for the grid-bootstrap method, whose performance is not surprising as, by construction, it is valid *globally* in the parameter space. Elliott and Stock (2001) method seems to be the most sensitive to the presence of stationarity.

Finally, Table 1(b) shows a comparison of the exact (14) and approximate (17) measures of the half-life in a simple AR(2) setting. For simplicity, both confidence intervals are

constructed by using Stock (1991) method. The two measures are numerically identical when  $\gamma = 0.5$ . As expected, the table shows that the performance of the approximate measure worsens relative to the exact one as  $\gamma$  deviates more from 0.5.

In conclusion, this Monte Carlo example shows that empirical results based on tests other than the normal sampling asymptotic can be more reliable even for *very small* half-lives. Also, we documented the poor coverage of confidence intervals based on (17) as the true data generating process departs from an AR(1).

Insert Tables 1(a) and 1(b)

## 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 1998:2, for a total of 100 observations. 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 "XXX64..F" (CPI) and "XXX..AE." (bilateral nominal exchange rate), where "XXX" is the mnemonic used by the IMF to denote the country (e.g. "USI" for the U.S.).

Table 2 reports confidence intervals based on standard asymptotic theory. The lag length ( $k$ ) 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 the exact half-life,  $h_e$ , are around 8 to 12 quarters (2 to 3 years) for most currencies. As discussed previously, most researchers report only an approximation to the true half-life,  $h_a$ . Unless the underlying data generating process is an AR(1) the approximation is not correct, even asymptotically.

From table 2, notice that, due to absence of the correction factor, this procedure generally underestimates the true half-life.<sup>15</sup>

Based on normal sampling asymptotic theory, the 95% normal confidence intervals include zero to twenty (or more) quarters for most currencies. However, it is well known that when the time series process under consideration is highly persistent, normal sampling methods for constructing confidence intervals will not be correct. The real exchange rate is such a series. In fact, table 2 also reports the conventional Augmented-Dickey Fuller tests for the real exchange rates considered in this paper. 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. If a unit root is present in the data then the real exchange rate will not be mean reverting, thus implying that the PPP theory does not hold in the data. The table also reports the DF-GLS efficient test statistic of Elliott, Rothenberg 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. Hence, the data are considerably persistent. 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 exact and approximate confidence intervals based on Stock's method. We reported only the lower bounds of the confidence intervals because the upper bounds were infinity for all currencies. That is, based on these results, one cannot reject that there is a unit root in the data, that is one cannot reject that the half-life can be infinity. In fact, for a few currencies, the median half-life is infinity. However, the uncertainty over the estimated half-life is so big that a half-life of 6 or 7 quarters is compatible with the observed data for almost all currencies as well.<sup>16</sup> The fact that confidence intervals can be so big is due to the high persistence in the series. The results do not change much when considering exact half-lives. The exact half-lives (and the implied lower bounds for the confidence intervals) are in general higher, especially if one considers one-sided confidence

intervals,<sup>17</sup> but still for most currencies remain around 6 to 8 quarters. Although it is true that a median-unbiased estimate of the half-life can be as high as infinity, these results do not contradict the PPP, as the lower bounds are compatible with a time horizon in which it is possible that prices are sticky and, thus, that deviations from PPP can be explained in the light of monetary factors.

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 approximate half-life deviations from PPP are between 4 to 6 quarters for most currencies, except for a few outliers (Canada and Australia in particular). Estimates of the exact half-life are slightly higher and present more outliers (Finland, Greece and Sweden in addition to the cases already mentioned). However, overall, this additional evidence delivers confidence intervals of roughly the same magnitude as Stock's method, thus confirming the previous results. Note that it is disappointing that the confidence intervals are not much shrunked 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.

## 6. Conclusions

This paper offers an empirical assessment of the Purchasing Power Parity. The theory suggests that fluctuations in the nominal exchange rate between two countries are driven by their price differentials. One implication of the theory is that we should observe mean reversion of the real exchange rate, and such reversion should happen in a time frame compatible with price adjustment lags (usually in the order of 1 to 2 years). Existing point estimates of half-life deviations, in the order of 3 to 5 years, are difficult to reconcile with the theory. However, when the variability of the estimates is sufficiently high, even a point estimate of 5 years could be reconciled with theory. This could be the case especially when variables are highly persistent, as the variance becomes particularly high and the corresponding confidence intervals large. The objective of this paper was to build confidence intervals for half-life deviations from Purchasing Power Parity.

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. We first consider Stock's method based on inverting the Augmented Dickey-Fuller test, whose confidence intervals cover half-lives as low as 5 to 6 quarters for most real exchange rates considered in this paper (although some countries are outliers). These results are encouraging, as nominal prices and wages can be sticky in this time frame, and are robust to the use of other methods (Elliott and Stock's lower bound is around 4 to 6 quarters for most currencies and Hansen's is around 5 to 6 quarters).

Overall, the results of this paper suggest that, although the real exchange rate is a highly persistent variable, whose median half-life can even be infinity for some currencies, the deviations from its long run equilibrium are compatible with processes that can halve in 4 to 6 quarters. Thus, deviations from PPP can reasonably be justified by price stickiness. This result reconciles the observed enormous short term volatility in the real exchange rate with an admissible lower bound for the PPP half-life as low as 1 to 2 years for most currencies. The admissibility of such low bounds is justified by the interaction between the high persistence in the real exchange rate and the small sample sizes usually available. However, since we cannot rule out the possibility of an infinite median half-life, it is also possible to interpret the evidence as simply being non informative, as Murray and Papell (2001) do.

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. Neither paper constructs confidence intervals.

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 (see Taylor (2001) and Panos, Nobay and Peel (1997) among others).

Also, this paper deals only with time-series data over the floating-period. That is, this paper does not address panel issues, nor it considers longer datasets that merge floating and fixed exchange rate samples. It would be interesting to see whether the results continue to hold in panels and with longer data series. However, 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. It would also be interesting to pool data across countries to get more accurate inference, although this would require additional assumptions on the joint distribution and a careful investigation of cross sectional dependence (see O'Connell (1998a) for a discussion of the latter). Finally, it might be worth investigating the reason 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.
- Blough, Stephen, "The relationship between power and level for generic unit root tests in finite samples", *Journal of Applied Econometrics* 7, 1992, pp. 295-308.
- Bobkoski, M. J., "Hypothesis testing in nonstationary time series", Unpublished Ph.D. thesis (Department of Statistics, University of Wisconsin, Madison, WI), 1983.
- Caner, Mehmet, and Kilian, Lutz, "Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate", *mimeo*, 2001.
- Cavanagh, C., "Roots local to unity", Manuscript (Department of Economics, Harvard University, Cambridge, MA), 1985.
- Chan, N. H., and C. Z. Wei, "Asymptotic inference for nearly nonstationary AR(1) processes", *Annals of Statistics* 15, 1987, pp. 1050-63.
- Cheung, Yin-Wong and Kon Lai, "Parity reversion in real exchange rates during the Post-Bretton Woods period", *Journal of International Money and Finance* 17(4), August 1998, pp. 597-614.
- Cheung, Yin-Wong and Kon Lai, "On the purchasing power parity puzzle", *Journal of International Economics* 52, 2000, pp. 321-330.
- Efron, Bradley, and Tibshirani, Robert J., *An Introduction to the Bootstrap*, New York: Chapman and Hall/CRC, 1998.

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.

Froot, K. and Rogoff, Kenneth, "Perspectives on PPP and Long-Run Real Exchange Rates", in Grossman, Gene, and Rogoff, Kenneth, *Handbook of International Economics*, vol. 4, North Holland, 1995.

Hall, Peter, *The Bootstrap and Edgeworth Expansion*, New York: Springer-Verlag, 1992.

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*, forthcoming.

Kilian, Lutz, and Zha, Tao, "Quantifying the Uncertainty about the Half-Life of Deviations from PPP", *Journal of Applied Econometrics*, forthcoming.

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, forthcoming, 2001.

Meese, R. and Rogoff, Kenneth, "Was it real? The exchange rate interest differential relation over the modern floating exchange rate period", *Journal of Finance* 43, 1988, pp. 933-48.

Murray, Christian J., and Papell, David H., "The Purchasing Power Parity Persistence Paradigm", June 2001, forthcoming in the *Journal of International Economics*.

Ng, Serena, and Perron, Pierre, "Lag length selection and the construction of unit root tests with good size and power", forthcoming *Econometrica*.

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

O'Connell, P.G.J, "Market frictions and real exchange rates", *Journal of International Money and Finance* 17, 1998(b), pp. 71-95.

Panos, Michael, Nobay, Robert and Peel, David, "Transaction Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation", *Journal of Political Economy* (105), 1997, pp. 862-879.

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

Phillips. Peter C.B., "Towards a unified asymptotic theory for autoregression", *Biometrika* 74, 1987, pp. 535-47.

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

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 Bulletin 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).

Wright, Jonathan, “Confidence Intervals for Impulse Responses with a Near Unit Root”, *Journal of Business and Economic Statistics* (18), 200, pp. 368-373.

## Figures

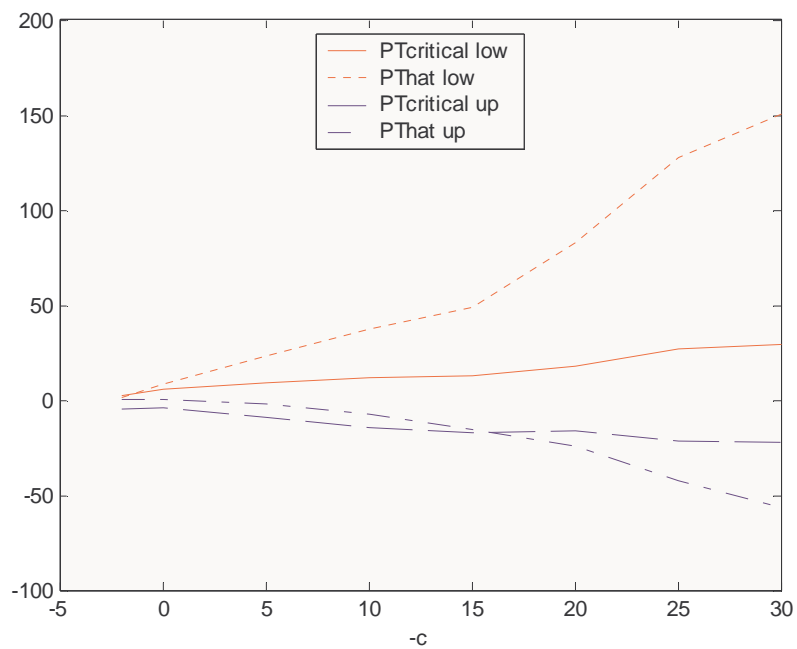


Figure 1. Two-sided Elliott and Stock (2001) test for demeaned Yen/\$ data.

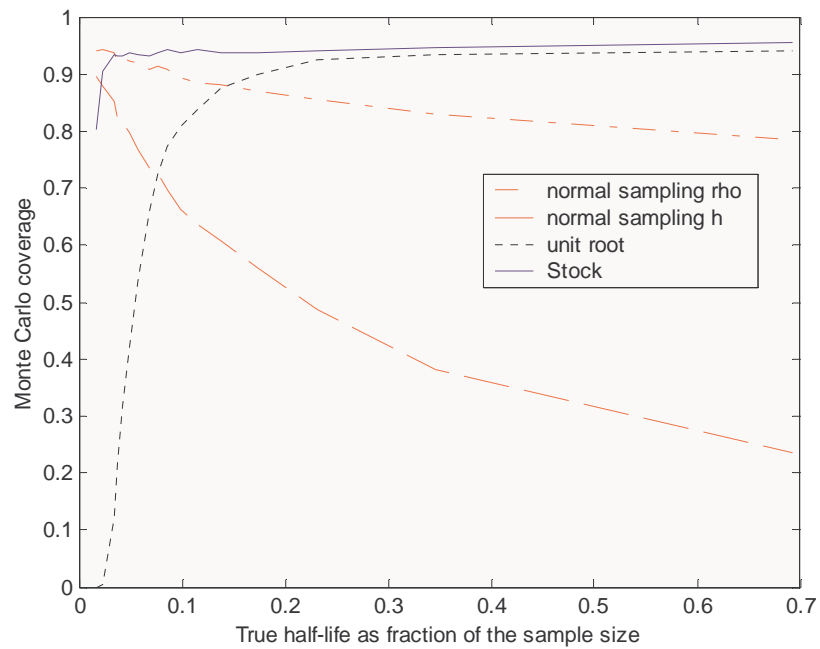


Figure 2. Comparison of coverage of various confidence intervals as a function of the true half-life (nominal coverage=0.95).

## Tables

**Table 1(a).** Comparison of coverage rates

$h$	$c$	<i>Normal</i>	<i>Unit root</i>	<i>Stock</i>	<i>Elliott-Stock</i>	<i>Hansen</i>
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
$+\infty$	-0.1	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- $\alpha$  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.

**Table 1(b).** Comparison of coverage rates

$c$	$\gamma$	$h_e$	$h_a$	<i>Exact</i>	<i>Approx.</i>	$c$	$\gamma$	$h_e$	$h_a$	<i>Exact</i>	<i>Approx.</i>
-5	0.6	32	35	0.93	0.93	-8	0.8	29	43	0.89	0.82
-5	0.7	38	46	0.92	0.91	-8	0.9	37	87	0.84	0.60
-5	0.8	46	69	0.91	0.88	-11.5	0.6	14	15	0.90	0.91
-5	0.9	59	139	0.89	0.76	-11.5	0.7	16	20	0.88	0.86
-8	0.6	20	22	0.91	0.91	-11.5	0.8	20	30	0.83	0.73
-8	0.7	24	29	0.91	0.89	-11.5	0.9	26	60	0.74	0.46

Note: Monte Carlo simulations (5,000 replications) are based on the following AR(2) process:

$$y_t = \left(1 + (1 - \gamma) \frac{c}{T}\right) y_{t-1} + \gamma \left(1 + \frac{c}{T}\right) (y_{t-1} - y_{t-2}) + \epsilon_t.$$

The table reports, for each true half-life the coverage rate of the confidence intervals based on Stock (1991) method using both (14), labeled “Exact”, and (17), labeled “Approx.”. Ideally, these coverage rates should be close to 0.95.

**Table 2** Confidence intervals based on standard asymptotics and *ADF* tests

	$\hat{\alpha}(1)$	<i>ADF</i>	<i>ADF</i> <sub>GLS</sub>	$\hat{h}_a$	$(\hat{h}_l^a, \hat{h}_u^a)$	$\hat{h}_e$
Austria	0.936	-1.8	-1.94	10.4	(0; 22.1)	10.4
Australia	0.943	-1.57	-0.396	11.8	(0; 26.9)	11.8
Belgium	0.948	-1.55	-2.1	13	(0; 30)	13.3
Canada	0.981	-0.765	-0.381	35.3	(0; 127)	36.7
Denmark	0.941	-1.64	-2.2	11.4	(0; 25.4)	11.4
Finland	0.89	-3	-2.1	5.96	(1.83; 10.1)	5.16
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
Italy	0.918	-2.18	-2.03	8.1	(0.48; 15.7)	8.54
Japan	0.935	-1.8	-1.85	10.3	(0; 21.8)	10.3
Netherl.	0.922	-2.03	-1.68	8.55	(0; 17.1)	8.9
Norway	0.909	-2.11	-1.77	7.28	(0.17; 14.4)	7.49
Spain	0.956	-1.51	-1.64	15.3	(0; 35.6)	15.3
Sweden	0.928	-2.08	-1.46	9.26	(1.17; 13)	9.62
Switzerl.	0.916	-2.15	-1.7	7.9	(0.37; 15.4)	7.9
U.K.	0.902	-2.43	-1.47	6.69	(0.99; 12.4)	7.06

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 ( $\hat{\alpha}(1)$  as defined in (16)) and the DF-GLS test proposed by Elliott, Rothenberg and Stock, 1996 (*ADF*<sub>GLS</sub>). 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).  $\hat{h}_a, \hat{h}_e$  are the estimates of the half-life from (17) and (14);  $(\hat{h}_l^a, \hat{h}_u^a)$  is based on (21). The 5% critical value of both the *ADF* and *ADF*<sub>GLS</sub> test statistics is -2.89.

**Table 3** Approximated and exact confidence intervals based on Stock (1991)

	$(\hat{c}_l, \hat{c}_u)$	$\hat{h}_{.025}^a$	$\hat{h}_{.05}^a$	$\hat{h}_{median}^a$	$\hat{b}(1)$	$\hat{h}_{.025}^e$	$\hat{h}_{.05}^e$	$\hat{h}_{median}^e$
Austria	(-12.5; 3.67)	5.57	6.34	29.96	1.00	5.57	6.34	29.96
Australia	(-10.1; 3.97)	6.87	7.97	720.04	1.00	6.87	7.97	720.04
Belgium	(-9.91; 3.99)	7.39	8.59	$+\infty$	0.95	7.55	8.77	$+\infty$
Canada	(-4.03; 4.63)	19.33	28.72	$+\infty$	0.89	20.09	29.85	$+\infty$
Denmark	(-10.8; 3.87)	6.49	7.47	109.71	0.99	6.52	7.51	110.24
Finland	(-27.9; 0.799)	7.38	8.00	13.80	0.34	6.39	6.93	11.95
France	(-12.8; 3.62)	5.43	6.17	26.48	1.00	5.43	6.17	26.48
Germany	(-12.5; 3.66)	5.55	6.32	29.50	1.00	5.55	6.32	29.50
Greece	(-18.4; 2.7)	6.45	7.14	15.80	0.58	6.69	7.40	16.39
Italy	(-16.7; 2.95)	5.07	5.64	13.84	0.82	5.34	5.95	14.60
Japan	(-12.5; 3.67)	5.57	6.34	29.97	1.00	5.57	6.34	29.97
Netherl.	(-15; 3.22)	5.23	5.88	16.72	0.88	5.44	6.12	17.41
Norway	(-15.8; 3.07)	4.72	5.29	13.81	0.93	4.86	5.44	14.20
Spain	(-9.52; 4.05)	7.28	8.50	$+\infty$	1.00	7.28	8.50	$+\infty$
Sweden	(-15.6; 3.11)	7.54	8.45	22.59	0.59	7.83	8.78	23.48
Switzerl.	(-16.4; 2.99)	4.24	4.73	11.86	1.00	4.24	4.73	11.86
U.K.	(-19.8; 2.47)	4.27	4.70	9.84	0.82	4.50	4.96	10.37

Note. For each bilateral real exchange rate we run a demeaned *ADF* regression. The median-unbiased two-sided confidence interval for  $c$ ,  $(\hat{c}_l, \hat{c}_u)$ , is obtained directly by inverting Stock's table A1 (with a linear interpolation from its grid values). We report both two-sided and one-sided lower bounds for the median unbiased confidence intervals for the half-life ( $h$ ) with coverage 0.95, denoted by subscripts .025 and .05 respectively. Superscripts  $e$  and  $a$  denote exact and approximated measures, as described in equations (14) and (17) respectively, where  $T = 100$ . 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** Approximated and exact confidence intervals  
based on Elliott and Stock (2001) and Hansen (1999)

	<i>Elliott &amp; Stock</i>				<i>Hansen</i>			
	$\widehat{h}_{.025}^a$	$\widehat{h}_{.025}^e$	$\widehat{h}_{.05}^a$	$\widehat{h}_{.05}^e$	$\widehat{h}_{.025}^a$	$\widehat{h}_{.025}^e$	$\widehat{h}_{.05}^a$	$\widehat{h}_{.05}^e$
Austria	4.42	4.42	5.87	5.87	5.83	6.19	6.65	7.00
Australia	26.17	26.17	53.67	53.67	6.61	6.96	7.34	7.69
Belgium	3.54	3.61	4.13	4.22	6.80	7.31	8.09	8.63
Canada	10.13	10.53	32.29	33.56	14.58	15.52	19.75	20.89
Denmark	3.32	3.33	3.84	3.86	5.89	6.27	7.45	7.84
Finland	8.74	7.57	10.23	8.86	3.82	3.62	4.06	3.82
France	4.01	4.01	5.10	5.10	5.46	5.81	6.09	6.45
Germany	3.70	3.82	4.50	4.64	6.20	6.55	7.37	7.72
Greece	5.85	6.07	6.88	7.13	4.71	5.25	5.30	5.87
Italy	4.68	4.94	5.61	5.92	4.71	5.35	5.21	5.87
Japan	4.42	4.42	5.89	5.89	5.65	6.00	6.63	6.98
Netherl.	5.35	5.57	7.40	7.70	4.90	5.47	5.47	6.07
Norway	4.87	5.01	6.55	6.74	4.25	4.74	4.64	5.14
Spain	5.16	5.16	7.54	7.54	8.28	8.63	9.14	9.49
Sweden	16.13	5.82	19.39	6.99	4.08	4.61	4.41	4.96
Switzerl.	4.49	4.49	7.41	7.41	4.57	4.92	5.09	5.45
U.K.	5.70	6.01	10.11	10.66	3.98	4.57	4.47	5.09

Note. For each bilateral real exchange rate we run a demeaned *ADF* regression, where the lag length is chosen according to MAIC. We report both 2-sided and 1-sided confidence intervals for the half-life ( $h$ ) with coverage 0.95, denoted by subscripts .025 and .05 respectively. Superscripts  $e$  and  $a$  denote exact and approximate bounds. Upper bounds were infinity for all currencies so they were not reported.

## 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  $\rho^h$ ; (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  $\rho^h \phi_b$ .

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

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

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

$$Y_t = \mathbf{A}Y_{t-1} + e_t \quad (23)$$

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

(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),  $Y_t \equiv [\tilde{y}_t, \Delta \tilde{y}_t, \Delta \tilde{y}_{t-1}, \dots, \Delta \tilde{y}_{t-k+1}]'$  is a  $(k+1) \times 1$  vector,  $\Delta \tilde{y}_{t-j} \equiv \tilde{y}_{t-j} - \tilde{y}_{t-j-1}$ ,  $e_t = [\epsilon_t, \epsilon_t, \mathbf{0}_{1 \times (k-1)}]'$  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*:<sup>18</sup>

$$\tilde{y}_t = \rho \tilde{y}_{t-1} - \sum_{j=1}^k b_j \tilde{\Delta} \tilde{y}_{t-j} + \epsilon_t \quad (25)$$

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

$$\tilde{Y}_t = \mathbf{E}\tilde{Y}_{t-1} + e_t \quad (26)$$

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

(where  $E_{11} = \rho$  and the rest of the matrix  $\mathbf{E}$  is partitioned accordingly),  $\tilde{Y}_t \equiv [\tilde{y}_t, \tilde{\Delta}\tilde{y}_t, \tilde{\Delta}\tilde{y}_{t-1}, \dots, \tilde{\Delta}\tilde{y}_{t-k+1}]'$  is a  $(k+1) \times 1$  vector and  $\tilde{\Delta}\tilde{y}_{t-j} \equiv \tilde{y}_{t-j} - \rho\tilde{y}_{t-j-1}$ . We also find it convenient to define  $\mathbf{1}_k$  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  $\bar{y}_{t-1}^{LR}$  and at time  $t$  there is a shock  $\epsilon_t$ . The initial deviation from equilibrium is thus:

$$\tilde{y}_t = \mathbf{1}'_k e_t = \epsilon_t \quad (28)$$

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

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

...

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

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

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

This measures the effect of a shock  $\epsilon_t$  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:

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

As  $h \rightarrow \infty$ ,  $E_{22}^h \rightarrow \mathbf{0}_{k \times k}$  (because all of its roots are in absolute value less than one) and  $E_{11}^h = \rho^h \rightarrow e^{c\delta}$  (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:

$$\begin{aligned}
&E_{11}^h + E_{11}^{h-1} E_{12} (I_k - E_{22}/E_{11})^{-1} \\
&= E_{11}^h \left( 1 + E_{12} (E_{11} I_k - E_{22})^{-1} \mathbf{1}_k \right) \\
&\simeq E_{11}^h \left( 1 + E_{12} (I_k - E_{22})^{-1} \mathbf{1}_k \right) = E_{11}^h \phi_b
\end{aligned} \tag{32}$$

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

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

where  $\phi_b$  is the correction factor:<sup>19</sup>

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

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:

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

where the correction factor is:

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

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

$$a(L)y_t = \epsilon_t \quad (37)$$

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

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

where:

$$\rho_1 = \gamma + \rho \quad (39)$$

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

The canonical *ADF* representation is:

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

Finally, the *ADF* representation is:

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

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

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

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

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

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

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

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

The reason why 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:

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

where  $\xi_t$  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  $\epsilon_t$ . In fact, from the data generating process  $y_t = \rho y_{t-1} + v_t$  we have that:

$$\frac{\partial y_{t+h}}{\partial v_t} = \rho^h \rightarrow e^{c\delta} \quad (46)$$

Since  $v_t = b(L)^{-1} \epsilon_t$ , the long-run effect of a shock to  $\epsilon_t$  will be  $\rho^h b(1)^{-1}$ , which is our exact measure. What we called "approximate long-run effect", which corresponds to what most empirical researchers report, is instead  $\alpha(1)^h \rightarrow e^{c\delta b(1)}$  so it will be different from the exact long-run effect unless  $c = 0$  or the true process is an AR(1), for which  $b(1) = 1$ . Hence 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)):

$$\alpha(1)^h \rightarrow \frac{e^{c\delta b(1)}}{b(1)} \quad (47)$$

## Notes

<sup>1</sup>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.

<sup>2</sup>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). Under more general AR(p) processes, this result will not hold. However, empirical researchers use (3) as an approximation to the true half-life, so we will use a subscript "a" to denote this measure and to highlight this fact in the next sections. Later, we will discuss the assumptions under which it is correct.

<sup>3</sup>The mean is needed because of the base year normalization of price indexes.

<sup>4</sup>Similar results, showing some empirical support for the PPP theory in small samples, are obtained by Elliott and Pesavento (2001). They apply higher power tests for mean reversion against close alternatives by exploiting information on other economic variables.

<sup>5</sup>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  $\tau$  is defined as  $\bar{y}_\tau = \mu_0 + \mu_1 \tau$ . 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.

<sup>6</sup>We will provide detailed empirical evidence on the degree of persistence in the bilateral exchange rates considered in this paper in the empirical section.

<sup>7</sup>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.

<sup>8</sup>This equation is the one that we will use in the empirical estimation.

<sup>9</sup>A comparison of the coverage of the confidence intervals based on approximate and exact measures of the half-life is addressed in section 4.

<sup>10</sup>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).

<sup>11</sup>We also explored the construction of confidence intervals by minimizing the length of the confidence interval. The empirical results were similar.

<sup>12</sup>For this simple process, the exact and the approximated measures of the half-life coincide.

<sup>13</sup>The small sample performance of the Elliott and Stock (2001) and the Hansen (1999) methods will be considered later.

<sup>14</sup>In practice, the infinite half-life is 6900, corresponding to a value of  $c = -0.01$ .

<sup>15</sup>To save space – and because confidence intervals constructed with usual asymptotics are not correct anyways – confidence intervals for the exact half-life are not reported. They comprise higher values for the half-life. We chose to report confidence intervals based on the approximated – rather than exact – half-life in order to compare our results with those existing in the literature.

<sup>16</sup>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 bigger than 6 quarters are Finland, Spain and Sweden. But for most of them, the lower bound is less than 8 quarters anyway.

<sup>17</sup>The reason is that lower bounds of one-sided confidence intervals correspond to lower bounds of two-sided confidence intervals with twice the nominal coverage, that is 10%.

<sup>18</sup>The canonical regression approach was introduced by Sims, Stock and Watson (1990). It corresponds to a transformed DGP where all the regressors (in the representation for  $\tilde{y}_t$ ) except one have mean zero and are stationary.

<sup>19</sup>It can be shown that this is the same correction factor derived in section 2. This is the reason why we call it in the same way. See Hamilton, chps. 1 and 2, for a discussion on the equivalence of the eigenvalue factorization and the VAR representation. We won't prove this, but only show that this is true in the AR(2) example below.