



---

**Centre for International Capital Markets**

**Discussion Papers**

***ISSN 1749-3412***

---

**Non-normality and Recursive Unit Root Tests for PPP:  
Solving the PPP Puzzle?**

Guglielmo Maria Caporale and Andros Gregoriou

No 2007-6

**NON-NORMALITY  
AND RECURSIVE UNIT ROOT TESTS FOR PPP:  
SOLVING THE PPP PUZZLE?**

**Guglielmo Maria Caporale  
Andros Gregoriou**

Brunel University, London

April 2007

*Corresponding author: Professor Guglielmo Maria Caporale, Brunel Business School, Brunel University, Uxbridge, Middlesex UB8 3PH, UK. Tel.: +44 (0)1895 266713. Fax: +44 (0)1895 269770. E-mail: [Guglielmo-Maria.Caporale@brunel.ac.uk](mailto:Guglielmo-Maria.Caporale@brunel.ac.uk)*

## Abstract

In this paper we carry out unit root tests on real exchange rates recursively as in Caporale et al (2003), but, following Arghyrou and Gregoriou (2007), we adjust the residuals for non-normality using a wild bootstrap method. The results are striking: the correction for non-normality dramatically increases the rejection percentages of the unit root null, and attenuates the erratic behaviour of the t-statistic, thus providing strong evidence in favour of PPP, and suggesting that such a correction might at least go some way towards solving the “PPP puzzle”.

**Keywords:** Purchasing Power Parity (PPP), Real Exchange Rate, Unit Roots, Non-Normality, Wild Bootstrap

**JEL Classification:** C12, C22, F31

## 1 INTRODUCTION

Most economists are in agreement that some form of Purchasing Power Parity (PPP) determines the long-run equilibrium level of the exchange rate (see, e.g., Taylor and Taylor, 2004 for a critical review of the PPP debate). However, the available empirical evidence is not always consistent with the PPP condition. This apparent contradiction is known as the “PPP puzzle” (see Rogoff, 1996), and has frequently been attributed to the limitations of standard unit root and cointegration tests (see Froot and Rogoff, 1995). Subsequent studies using panel methods (see Caporale and Cerrato, 2006 for a survey), or examining non-linearities (see, e.g., Taylor et al, 2001) still reach mixed conclusions.

Recently, Caporale et al (2003) have provided evidence of erratic behaviour of standard unit root tests, which might arguably reflect some form of non-stationarity in the underlying process more complex than the unit root one usually assumed (see Caporale and Pittis, 2002).<sup>1</sup> In another recent study, Arghyrou and Gregoriou (2007) focus on non-normality of the estimated residuals of the PPP equation, and conclude that PPP is still rejected when correcting for non-normality in US dollar PPP regressions for the G7 bilateral exchange rates using the wild bootstrap technique.

In the present paper we carry out the same type of recursive analysis as in Caporale et al (2003), but, following Arghyrou and Gregoriou (2007), we also adjust the residuals for non-normality using a wild bootstrap method. The results are striking: the correction for non-normality increases dramatically the rejection percentages of the unit root null, and attenuates the erratic behaviour of the t-statistic, thus providing strong evidence in favour of PPP, and suggesting that such a correction might solve the “PPP puzzle”.

The layout of the paper is as follows. Section 2 describes the data, and the recursive unit root tests. In Section 3 the wild-bootstrap correction for non-normality is introduced. Section 4 concludes.

## 2 RECURSIVE UNIT ROOT TESTS

Our dataset consists of the bilateral US dollar exchange rates for the G7 countries and the Euro/US dollar rate (ECU/USD prior to 1999), and the consumer price index

---

<sup>1</sup> Similar results are reported in the case of trivariate cointegration tests by Caporale and Hanck (2006).

(CPI). The frequency is monthly and covers the post-Bretton Woods period (1973:1-2005:12). For the French Frank, German Mark and Italian Lira the tests are carried out for the period 1973:1-1998:12, as these currencies were subsequently replaced by the Euro. For the Euro/Dollar exchange rate, the sample is 1978:12-2005:12. The data are taken from the IMF's International Financial Statistics Databank, except for the Euro/USD exchange rate and EMU CPI price index, for which the source is the OECD Historical Data databank.<sup>2</sup>

As a first step, we carried out Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) tests (see Dickey and Fuller, 1979, and MacKinnon, 1991 for the critical values) on the log of the real exchange rate. The general regression model is the following:

$$\Delta y_t = \gamma_0 + \gamma_1 y_{t-1} + \sum_{i=1}^n \phi_i \Delta y_{t-i} + \varepsilon_t \quad (1)$$

where  $y_t$  is the log of the real exchange rate, the  $\gamma$ 's and  $\phi$ 's are constant parameters and  $\varepsilon_t$  is a random disturbance term. Rejection of the null of non-stationarity requires  $\gamma_1$  to be negative and significantly different from zero. A lag length of one was chosen for the countries under investigation on the basis of the Akaike Information and Schwarz Information Criteria (AIC, SIC).

Next, following Caporale et al (2003), we created a new time series "t-stat", defined as  $\gamma_{1t}/s.e.(\gamma_{1t})$ , which is the computed t-statistic from the recursive estimation of the coefficients of the AR(1) model selected using AIC and SIC. The plot of the recursive t-stat can then be inspected for evidence of erratic behaviour: frequent jumps from the rejection to the acceptance region or vice versa being an indication of endemic instability in the underlying Data Generation Process (DGP). Table 1 reports in each case the minimum and maximum value of the t-statistic, as well as the rejection and acceptance percentages (see columns 2 to 5). As can be seen, the rejection percentages are rather low, suggesting in all cases that PPP does not hold. Also, graphs of the recursive t-stat (not included for reasons of space) provide more evidence of the type of erratic behaviour previously detected by Caporale et al (2003). However, the Jarque-Bera test implies that the null of normality is strongly rejected

---

<sup>2</sup> The analysis was also carried out using PPI data; these results, qualitatively similar, are available from the authors upon request.

for all estimated residuals' series (see Table 1, column 6 and 7, reporting the min and max value respectively for each country). Whether correcting for non-normality affects the inference drawn from the reported ADF statistics is examined next.

### 3 CORRECTING FOR NON-NORMALITY

The distribution of the standard DF test is computed under the assumption that the series  $y_t$  is normally distributed. In case of non-normality, the standard parametric correction uses the White (1980) heteroscedastic-consistent covariance matrix. However, a parametric form might not be able to capture certain forms of conditional heteroscedasticity, especially in the presence of instabilities in the data. Consequently, as stressed by Arghyrou and Gregoriou (2007), it is preferable to adopt a non-parametric specification for the conditional variance of the residuals of equation (1) (see Goncalves and Kilian, 2004).

The wild bootstrap technique (see, e.g., Davidson and Flachaire, 2001 and Ioannidis and Peel, 2005) is a suitable method to obtain appropriate critical values for the DF test in this context. It entails estimating equation (1) by OLS, obtaining the estimated  $\varepsilon_t$  series and generating a new series of residuals given by

$$\varepsilon_t^* = \varepsilon_t u_t \tag{2}$$

where  $u_t$  is drawn from a two-point distribution

$$u_t \begin{pmatrix} -(5^{0.5}-1)/2 \text{ with probability } p = \frac{(1+5^{0.5})}{2(5^{0.5})} \\ (5^{0.5}+1)/2 \text{ with probability } (1-p) \end{pmatrix}$$

The  $u_t$  terms are mutually independent drawings from a distribution which is independent of the original data and has the properties  $E(u_t) = 0$ ,  $E(u_t^2) = 1$  and  $E(u_t^3) = 1$ . As a result, any non-normality and heteroscedasticity in the estimated residuals  $\varepsilon_t$  of equation (1) is preserved in the created residuals  $\varepsilon_t^*$ . We generate 10,000 sets of the  $\varepsilon_t^*$  residuals. Subsequently, for each bootstrap iteration a series of DF tests are constructed under the null hypothesis  $\gamma_1 = 0$ , so that

$$\Delta y_t^* = \gamma_0^* + \sum_{i=1}^n \phi_i \Delta y_{t-i}^* \quad t = 1 \dots 10,000 \quad (3)$$

As can be seen from (3), the generated sequence of artificial data has a true  $\gamma_1$  coefficient of zero. However, when one regresses the artificial DF test for a given bootstrap sample  $0t$  estimated values of  $\gamma_1$  that differ from zero will result. This procedure provides an empirical distribution for  $\gamma_1$  and their associated standard errors based exclusively on the resampling of the residuals from the original regression (1). Therefore appropriate critical values are obtained for the null hypothesis of non-stationarity  $\hat{\gamma}_1 = 0$ . in equation (1).

We repeat the recursive analysis of Section 2 using the wild bootstrap correction. Table 1 (column 8 and 9) presents the minimum and maximum critical values obtained from a 95% confidence interval when the test statistic is less than zero. The lower limits of these intervals provide the critical values to be used for the estimated t-statistics. The table also shows the corresponding rejection and acceptance percentages (column 10 and 11). As can be seen, the results are striking: in all cases there is a dramatic increase in the rejection percentages, the evidence becoming strongly supportive of PPP. In other words, using the bootstrap critical values appears to reverse the earlier findings, implying that tests of the null hypothesis of non-stationary residuals in DF/ADF regressions are not robust to correcting for non-normality. Moreover, plots of the recursive t-stat indicate that the erratic behaviour has become much less evident (see Figure 1 for the case of the UK).<sup>3</sup> This suggests that failure to correct for non-normality might be responsible for earlier findings being inconsistent with PPP, and also that non-normality, rather than some other complex dynamic structure or endemic instability in the underlying process, might account for the erratic behaviour of standard unit root tests.

#### 4 CONCLUSIONS

This paper aims to shed light on the PPP puzzle by investigating the implications of violations of the normality assumption for the residuals of standard unit root tests of PPP - an issue which the earlier literature, with the notable exception of Arghyrou and Gregoriou (2007), has overlooked. We analyse the US dollar exchange rate against

---

<sup>3</sup> A very similar pattern is observed for the other countries. The corresponding figures are available upon request.

the currencies of the G7 countries and the Euro during the period of flexible exchange rates, and show that correcting the critical values of the standard DF statistic for non-normality using a wild bootstrap method has very important consequences: unit root tests are not any longer wildly erratic, and the evidence becomes strongly supportive of the PPP hypothesis, This finding is in contrast to Arghyrou and Gregoriou (2007), and can be explained by the fact that in the present study we correct for non-normality recursively, thereby obtaining more robust results in the presence of possibly volatile unit root tests. It would appear that taking into account non-normality in a recursive framework might be the way to solve the “PPP puzzle”.

## REFERENCES

Arghyrou, M. and A. Gregoriou (2007), "Testing for purchasing power parity correcting for non-normality using the wild bootstrap", *Economics Letters*, forthcoming.

Caporale, G.M. and M. Cerrato (2006), "Panel data tests of PPP: a critical overview", *Applied Financial Economics*, 16, 1-2, 73-91.

Caporale, G.M. and N. Pittis (2002), "Unit roots versus other types of time-heterogeneity, parameter time dependence and superexogeneity", *Journal of Forecasting*, 21,3, 207-223.

Caporale, G.M., N. Pittis and P. Sakellis (2003), "Testing for PPP: the erratic behaviour of unit root tests", *Economics Letters*, 80, 2, 277-284.

Caporale, G.M. and C. Hanck (2006), "Cointegration tests of PPP: do they also exhibit erratic behaviour?", *CESifo Working Paper no. 1811*, Munich, Germany.

Davidson, R. and E. Flachaire (2001), "The Wild Bootstrap, Tamed at Last," *mimeo*, Department of Economics, Queens University, Kingston, Ontario.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of the estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association*, 74, 427-431.

Froot, K.A. and K. Rogoff (1995), "Perspectives on PPP and long-run real exchange rates", in G. Grossman and K. Rogoff (eds.), *The Handbook of International Economics*, vol. 3, Elsevier Press, Amsterdam.

Goncalves, S. and L. Kilian (2004), "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form", *Journal of Econometrics*, 123, pp. 89-120

Ioannidis, C. and D. Peel (2005), "Testing for market efficiency in gambling markets when the errors are non normal and heteroskedastic: an application of the wild bootstrap", *Economics Letters*, 87, pp. 221-226.

MacKinnon, J.G. (1991), "Critical values for cointegration tests", in R.F. Engle and C.W.J. Granger (eds.), *Long-Run Economic Relationships*, 267-276, Oxford University Press, Oxford.

Rogoff, K. (1996), "The purchasing power parity puzzle", *Journal of Economic Literature*, 34, 647-668.

Taylor, M.P., Peel, D. and L. Sarno (2001), "Non-linear mean reversion in real exchange rates: towards a solution to the purchasing power parity puzzles", *International Economic Review*, 42, 1015-1042.

Taylor, A.M. and M.P Taylor (2004), “The purchasing power debate”, *Journal of Economic Perspectives*, 18, 4, 135-158.

White, H. (1980), “A heteroskedasticity-consistent covariance matrix estimator and a direct test of heteroskedasticity”, *Econometrica*, 48, pp. 817-838.

**Table 1**  
**Unit Root Tests**

Country	Min	Max	Reject	Accept	NORM MIN	NORM MAX	CV Min	CV Max	Reject 2	Accept 2
Canada	-3.93	-1.84	30	70	17.02 (0)	19.36 (0)	-3.22	-2.20	69	31
Emu	-4.02	-1.98	13	87	16.83 (0)	17.04 (0)	-3.75	-2.20	71	29
France	-3.80	-1.79	3	97	18.04 (0)	20.08 (0)	-3.18	-2.15	60	40
Germany	-3.90	-1.83	30	70	16.08 (0)	19.36 (0)	-3.12	-2.27	63	37
Italy	-3.90	-1.92	32	68	18.03 (0)	18.98 (0)	-3.20	-2.20	73	27
Japan	-3.96	-1.80	12	88	15.02 (0)	18.34 (0)	-3.01	-2.30	68	32
UK	-4.02	-1.99	0	100	15.84 (0)	16.94 (0)	-3.10	-2.26	73	27

Notes: p values are in brackets. Min and Max represent the minimum and maximum value of the ADF t- statistic constructed under the null hypothesis of a unit root. Reject and Accept show the percentage of the rejection – acceptance of the null hypothesis of a unit root. NORM MIN and NORM MAX are the minimum and maximum values of the Jacque-Bera test for normality. CV Min and CV Max are the minimum and maximum values of the ADF statistic constructed under the null hypothesis of a unit root, using the critical values obtained from the lower limits of a 95% confidence interval. The confidence interval was derived from a wild bootstrap simulation with replacement in 10,000 replications. Reject 2 and Accept 2 show the percentage of the rejection – acceptance of the null hypothesis of a unit root, using the lower limit of a 95% confidence interval obtained from the wild bootstrap simulation.

**Figure 1: UK ADF t-statistics**

