New Multi-Country Evidence on Purchasing Power Parity: Multivariate Unit Root Test Results *

Jan J.J. Groen†
Tinbergen Institute, Erasmus University Rotterdam

Econometric Institute Report EI 2000-9/A

Abstract

In this paper a likelihood-based multivariate unit root testing framework is utilized to test whether the real exchange rates of G10 countries are non-stationary. The framework uses a likelihood ratio statistic which combines the information across all involved countries while retaining heterogeneous rates of mean reversion. This likelihood ratio statistic has an asymptotic distribution which can be typified as a summation of squared, univariate Dickey and Fuller (1979) distributions. Our multivariate unit root tests indicate that bilateral G10 real exchange rates are stationary, irrespective of the *numeraire* country. We also analyze per panel the time necessary to have an adjustment to a shock in the individual real exchange rates. From this analysis it becomes apparent that there are significant cross-country differences in the adjustment of individual real exchange rates within each panel.

Keywords: Multivariate unit root testing, maximum likelihood estimation, PPP, real exchange rates.
JEL classification: C12, C23, F31.

1 Introduction

Purchasing power parity [PPP] is a main building bloc for open-economy macroeconomic models and it implies that real exchange rates are stationary. Testing the validity of PPP has provided an impetus to a whole literature on testing for stationary real exchange

*This paper has benefited from discussions with Frank Kleibergen and from comments by seminar participants at the University of Amsterdam and Maastricht University.

†Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands; E-mail: groen@few.eur.nl.

rates. In general, applying conventional augmented Dickey and Fuller (1979) [ADF] unit root tests on real exchange rates relative to the United States [U.S.] does not result in a rejection of the null of non-stationary real exchange rates. For example, Mark (1990) is not able to reject the null of non-stationarity for monthly real exchange rates relative to the U.S. and the United Kingdom [U.K.] for the 1973-1988 period whereas Papell (1997) has the same result for both monthly and quarterly U.S. real exchange rates over the 1973-1994 period. With respect to Germany-based real exchange rates both Mark (1990) and Papell (1997) provide more positive estimation results, albeit that they still do not significantly reject the hypothesis of non-stationarity for a majority of their real exchange rates.[1]

Since the Monte Carlo analysis in Shiller and Perron (1985) it is well known that the power of ADF unit root tests depend on the time span of the sample utilized in testing. As the time span of the post-Bretton Woods floating rate sample is rather short, 1973 up to the present, one can be doubtful that conventional ADF unit root tests are capable of detecting persistent, but stationary patterns in real exchange rates. One possible remedy for this problem is to look at panel data sets of real exchange rates. One can discern two groups of panel-based unit root tests of real exchange rates. Studies like Frankel and Rose (1996), MacDonald (1996), Oh (1996) and Papell (1997) have conducted panel unit root testing on real exchange rates using a version of the Levin and Lin (1992) panel unit root test. In general these studies find evidence for stationary real exchange rates in panels for 6 to 100 real exchange rates relative to both the U.S. and Germany on post-Bretton Woods samples. However, the evidence within panels of less than 10 countries is weak. Also, Papell (1997) fails to find evidence for stationarity within several samples of quarterly U.S.-based real exchange rates.

A major disadvantage of panel unit root testing based on the Levin and Lin (1992) approach is the assumption of cross-sectional independence between the different real exchange rates within the panel. Monte Carlo experiments in O'Connell (1998) indicate that panel unit root tests that neglect cross-sectional dependence yields severely biased test results on cross-sectionally correlated data. Given the fact that real exchange rates relative to the same base country are contemporaneously correlated, one should be doubtful with respect to test results based on the Levin and Lin (1992) approach. A second group of panel-based studies, most notably Abuaf and Jorion (1990) and O'Connell (1998), utilize panel unit root test regressions where they allow for cross-sectional correlation across the included real exchange rates. On a monthly sample of G10 real exchange rates over the period 1973-1987 Abuaf and Jorion (1990) only rejects the null of non-stationarity marginally at a 10% significance level. O'Connell (1998) in panels of 12 to 64 countries with quarterly data over the period 1973-1995 cannot reject the null of non-stationary real exchange rates at all.

When properly conducted, *i.e.* allowing for cross-sectional dependence, panel unit root tests give mixed results on the issue whether or not real exchange rates are stationary. However, the bulk of panel-based studies are based on the assumption of identical rates

[1] Froot and Rogoff (1995) contains a detailed survey of time series-based empirical tests of PPP.

of mean reversion and the weak panel-based evidence in favor of PPP could very well be caused by inappropriately assuming homogeneous speeds of mean reversion across countries, as suggested by O'Connell (1998, p. 18). For example, bilateral real exchange rates behave differently when monetary shocks in the bilateral relation ship are dominant than when real shocks are dominant. It is known from the literature that deviations from PPP are of short duration for high inflation countries. Also, when the home country has linked its monetary policy to that of the base country, based for example on a target zone regime, PPP deviations do not last long. On the other hand, Balassa (1964) and Samuelson (1964) have argued that in fast growing economies productivity growth in the traded goods sector is higher than in the non-traded goods sector and the relative price of traded/non-traded goods rises quickly. Consequently, if the home country grows faster than the base country the corresponding bilateral real exchange rate will exhibit a sustained appreciation (or a sustained depreciation in the inverse case), implying a low rate of mean reversion. Finally, the mean reversion of real exchange rates can be slowed down by the existence of transportation costs (see Dumas 1992) and when these transportation costs differ across countries they could lead to differing speeds of adjustment. Hence, in order to profit from the extra information in multi-country samples it could be worthwhile to conduct multi-country tests of PPP based on cross-sectional *heterogeneity* of mean reversion parameters.

Multi-country tests of PPP under parameter heterogeneity have up to now not been applied on a frequent basis. Coakley and Fuertes (1997) test the validity of PPP for U.S.-based real exchange rates of G10 countries over the 1973-1995 period within the heterogeneous panel unit root testing framework of Im *et al.* (1997) and they can reject the null of non-stationary real exchange rates. But, the results of Coakley and Fuertes (1997) should be treated with suspicion as the Im *et al.* (1997) framework, like the Levin and Lin (1992) framework, is based on the assumption of cross-sectional independence. Hakkio (1984) does allow for cross-sectional dependence as he estimates a system of four U.S.-based real exchange rates with generalized least squares [GLS], and his estimation results does not provide evidence for PPP. However, the Hakkio (1984) results are not explicitly based on the non-stationarity of real exchange rates under the null and are therefore unreliable. The most reliable results available in case of heterogeneous panels are provided by Engel *et al.* (1997), who use dollarized price levels over the period 1978-1994 for two cities in each of the U.S., Canada, Germany and Switzerland. Engel *et al.* (1997) construct three panel models comprising intra-national real exchange rates, national real exchange rates and continental (North-America versus Europe) real exchange rates, and they simultaneously estimate these three panel models with GLS. Based on parametric bootstrap distributions they test if each of the three panels are composed of non-stationary real exchange rate data and these tests reject the validity of PPP. Yet, Engel *et al.* (1997) only allow for a limited degree of parameter heterogeneity: across the three panels there is heterogeneity and *within* each of the three panel models the mean reversion speeds are homogeneous. This particular specification could very well be the cause of their negative results on the PPP hypothesis.

As an alternative to existing studies, our paper proposes to estimate a system of N ADF test regressions with iterative seemingly unrelated regression estimation [SURE] where the parameters differ for each equation. Likelihood ratio statistics are constructed to test the null hypothesis that all N series are non-stationary versus the alternative hypothesis that all N series are stationary. Compared to the existing literature our framework has several advantageous features. First, the set-up of our multivariate unit root testing framework is such that it allows for different rates of mean reversion under the alternative of stationary series. Next, the estimates and tests within our likelihood-based framework are robust to contemporaneous correlation across the series in our panel. In fact, our likelihood-based framework actually utilizes the presence of contemporaneous correlation to *enhance* the power of the multivariate unit root test. Existing studies of panel unit root tests on contemporaneously correlated data use (parametric) bootstrap distributions, as they claim that "...if there is cross-correlation in the data (...) the distributions of the statistics are not the same as before and are not known." (Maddala and Wu 1996, p. 14). Yet, for our multivariate likelihood ratio unit root test we are able to determine the distribution even if the data are cross-correlated.

The multivariate unit root test is used to test for the validity of PPP under cross-sectional heterogeneity for G10 real exchange rates within the 1973-1997 post-Bretton Woods period. In contrast to the existing literature, we not only use the U.S. as the *numeraire* country. Both within pure time series data (Frenkel 1981, Mark 1990) and within panel data sets (Jorion and Sweeney 1996, Papell 1997) there is more evidence for stationary real exchange rates when instead of the U.S. Germany is used as the base country. Therefore, we use Germany as one of our base countries. Also, like Mark (1990) we use the U.K. as a *numeraire* country. Finally, we use Japan as a base country for our G10 bilateral real exchange rates as this is the second largest non-European country within the set of G10 countries and because the Japanese economy has undergone several structural changes during this period. The multivariate unit root test results indicate that irrespective of the base country G10 bilateral real exchange rates are stationary. We also analyze the mean reversion speeds across the G10 real exchange rates in each panel, and this analysis shows that there is a severe cross-country heterogeneity in the mean reversion speeds within each of our four panels.

The remainder of this paper is organized as follows. In section 2 we provide an overview of existing panel unit root tests, including a Monte Carlo analysis. The likelihood-based multivariate unit root testing framework is described in section 3. Multivariate tests on the stationarity of G10 real exchange rates are reported in section 4. Section 5 concludes the paper.

2 Existing Panel Unit Root Tests

In order to improve upon the negative results of standard time series unit root tests, unit root testing on real exchange rates has recently been conducted within panels of N real exchange rates. Most studies base their analysis on the Levin and Lin (1992) framework which utilizes a test regression like[2]

$$\Delta x_{it} = \delta_i + \alpha x_{i,t-1} + \sum_{j=1}^{p} \gamma_{ij} \Delta x_{i,t-j} + \epsilon_{it}, \;\; i = 1, \ldots, N; \;\; t = 1, \ldots, T, \tag{1}$$

where $\Delta x_{it} = x_{it} - x_{i,t-1}$, δ_i is a constant which can differ across the cross-sections, i is the cross-section index and t is the time series index. Levin and Lin (1992) assume in (1) cross-sectionally unrelated disturbances: $\epsilon_{it} \sim \mathrm{N}(0, \sigma_i^2)$ for $i = 1, \ldots, N$, and p lagged first differences are added to guarantee that the ϵ_{it}'s are not autocorrelated. The non-stationarity of x_{it} for $i = 1, \ldots, N$ can now be tested in (1) through a t-statistic t_α for $H_0\colon \; \alpha = 0$ versus $H_1\colon \; \alpha < 0$. Levin and Lin (1992) derive that for $T \to \infty$, $N \to \infty$ and $\sqrt{N}/T \to 0$ a proper transformation of t_α converges in the limit to a standard normal distribution:[3]

$$\sqrt{1.25} t_\alpha + \sqrt{1.875} N \Rightarrow \mathrm{N}(0, 1). \tag{2}$$

A drawback of panel unit root testing based on (1), is the assumption of a homogeneous adjustment speed under the alternative hypothesis. Such an alternative hypothesis implies two things:

(a) $\alpha_i < 0$ for $i = 1, \ldots, N$;

(b) and conditional on (a): $\alpha_1 = \cdots = \alpha_N$.

When in reality only (a) is valid, assuming a common α in (1) can be too restrictive and could decrease the power to reject the null in favor of a true alternative hypothesis. A possible solution is to base multi-country unit root testing of real exchange rates on the framework of Im *et al.* (1997). This framework is based on the estimation of the ADF test regression for each $x_{1t}, \ldots, x_{Nt}$ separately:

$$\Delta x_{it} = \delta_i + \alpha_i x_{i,t-1} + \sum_{j=1}^{p_i} \gamma_{ij} \Delta x_{i,t-j} + \epsilon_{it}, \tag{3}$$

and constructing N conventional ADF t-statistics $t_{\alpha,i}$ under the null $\alpha_i = 0$ for $i = 1, \ldots, N$. Assuming $\mathrm{Cov}(\epsilon_{it}, \; \epsilon_{jt}) = 0$ for $i, j = 1, \ldots, N$ with $i \neq j$, Im *et al.* (1997) propose to test $H_0\colon \alpha_i = 0$ versus $H_1\colon \alpha_i \leq 0$ through

$$\Pi_{\bar{t}} = \frac{\sqrt{N}(\bar{t} - E(t_{\alpha,i} | \alpha_i = 0))}{\sqrt{\mathrm{Var}(t_{\alpha,i} | \alpha_i = 0)}} \Rightarrow \mathrm{N}(0, \; 1), \tag{4}$$

where $\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{\alpha,i}$ and the asymptotic distribution is valid for $N \to \infty$ and $T \to \infty$. In (4) $E(t_{\alpha,i} | \alpha_i = 0))$ and $\mathrm{Var}(t_{\alpha,i} | \alpha_i = 0)$ are the cross-sectional mean and variance of the

[2] The most appropriate specification for unit root tests on real exchange rates is the specification with a constant included in the test regression.

[3] A symbol "$\Rightarrow$" indicates convergence in distribution.

$t_{\alpha,i}$'s under the null which are calculated through Monte Carlo simulations by Im *et al.* (1997). Im *et al.* combine the individual ADF statistics into a common statistic, as in general the combining of multiple series into one statistic increases the power relative to the case where one bases the test on only one series.

Both the Levin and Lin (1992) and the Im *et al.* (1997) approaches suffer from a number of disadvantages which makes them inappropriate for testing the empirical validity of PPP across N real exchange rates. Firstly, the limiting distributions in both (2) and (4) rely heavily on a large number of cross-section observations N. However, the number of cross-sections for panels of macroeconomic data and in particular real exchange rate data is in most cases limited, especially for samples with quarterly or monthly data. This lack of a significant number of cross-sections could result in a lack of power for both the Levin and Lin (1992) and the Im *et al.* (1997) tests in quarterly or monthly panels of real exchange rates. Also, the panel unit root tests by Levin and Lin (1992) and Im *et al.* (1997) are based on cross-sectional independence between the involved real exchange rates and we argued before that this is a very unlikely assumption. As a consequence the asymptotic distributions in (2) and (4) are invalid.

To investigate the aforementioned problems with power and cross-correlated data we conduct several Monte Carlo experiments for the Im *et al.* and Levin and Lin panel unit root tests (IPS and LL respectively hereafter). We are especially interested in the size and power of the IPS and LL tests in panels of the size typically used in real exchange rate studies. Within the Monte Carlo experiments the data generating process [DGP] of the artificial series y_{it} used in our tests equals:

$$y_{it} = c_i + \rho_i y_{i,t-1} + \mu_{it}, \quad i = 1, \dots, N; \; t = 1, \dots, T. \tag{5}$$

A specification with a constant is chosen in (5) as we use this specification in section 4 and it is the most appropriate one for testing the PPP hypothesis. Also, we set $T = 100$ in (5) which is comparable to the number of quarterly observations within the 1973-1997 sample used in section 4. The cross-section dimension is set at $N = 9$ as we have 9 real exchange rates in the multi-country systems of section 4. We also set $N = 3, \; 6$ so that we can determine how the sizes and power ratios react to increases in the number of cross-sections. The innovations μ_{it} in (5) are generated through

$$\mu_{it} = \lambda_i \mu_{i,t-1} + \epsilon_{it}, \tag{6}$$

where $(\epsilon_{1t} \cdots \epsilon_{Nt})' \sim \mathrm{N}(\mathbf{0}_N, \Sigma)$ with the N-dimensional vector of zeros $\mathbf{0}_N$ and[4]

$$\Sigma = \Lambda'\Lambda, \quad \Lambda \text{ is } N \times N \text{ and } \Lambda \sim \mathrm{U}(0, \; 1). \tag{7}$$

Randomly generating the elements of the Λ matrix in (7) from an uniform distribution $\mathrm{U}(0, \; 1)$ guarantees that the ϵ_{it}'s in (6) are positively cross-correlated, as in the historical samples from section 4. Sizes and power ratios are computed both with and without first order serially correlated μ_{it}'s in (6):

[4]The denomination $\mathrm{U}(k_1, \; k_2)$ indicates that we draw from an uniform distribution on the interval between, but NOT including, k_1 and k_2.

Size without serial correlation: for $i = 1, \ldots, N$ we have in (5) $c_i = 0$ and $\rho_i = 1$, and in (6) $\lambda_i = 0$.

Size with serial correlation: for $i = 1, \ldots, N$ we have in (5) $c_i = 0$ and $\rho_i = 1$, and in (6) $\lambda_i \sim \mathrm{U}(0,\ 0.5)$.

Power without serial correlation: for $i = 1, \ldots, N$ we have in (5) $c_i \sim \mathrm{U}(-1,\ 1)$ and $\rho_i \sim \mathrm{U}(0.9,\ 1)$, and in (6) $\lambda_i = 0$.

Power with serial correlation: for $i = 1, \ldots, N$ we have in (5) $c_i \sim \mathrm{U}(-1,\ 1)$ and $\rho_i \sim \mathrm{U}(0.9,\ 1)$, and in (6) $\lambda_i \sim \mathrm{U}(0,\ 0.5)$.

For the power computations we have chosen to draw the mean reversion parameters from $\mathrm{U}(0.9,\ 1)$ in order to have an ample amount of heterogeneity, comparable with the range of estimated parameters in section 4, combined with a significant degree of persistence. All other parameters were also drawn from uniform distributions for each $i = 1, \ldots, N$ so that we have heterogeneity across the N cross-sections. As a benchmark we also calculate the sizes and power ratios for the univariate ADF unit root test, based on the above mentioned DGP's only now with $N = 1$.

The size and power computations are reported in table 1 and in the case of serially correlated errors we have fitted the test regressions (1) and (3) with a *common* lag order $p = 1$, 2 and 3 to measure the effect of overfitting the lag order. Except for $N = 3$ in case of the IPS test, both panel unit root tests are heavily oversized. This results confirms the fact that in case of cross-correlated data limiting distributions (2) and (4) are incorrect. As we combine in both the LL and the IPS tests a multiple of time series into one statistic the power of these tests should be higher than in case of the univariate ADF test, and this is what we observe in table 1. On the other hand, given the fact that both the LL and the IPS tests are oversized in samples of cross-correlated data the reported power ratios from table 1 are not very impressive. As both panel unit root tests are based on a framework with a large number of cross-sections, the small cross-section dimensions in the Monte Carlo experiments could explain this last observation. Also, as mentioned before, the LL test is based on a homogeneous rate of mean reversion and this could decrease the power of the LL test in our experiments which are based on heterogeneous mean reversion rates. In the next section we propose an alternative framework, which allows for both heterogeneous rates of mean reversion and cross-sectional dependence. Inference in our framework is solely based on large T asymptotics and as such the power of this method does not rely on the presence of a large number of cross-sections.

3 A Multivariate Framework for Unit Root Testing

In this section we propose a likelihood-based framework in which we simultaneously test for non-stationarity across N series. We first discuss in section 3.1 the involved estimation issues. Next, we construct in section 3.2 our multivariate likelihood ratio unit root test statics and discuss the corresponding asymptotic distribution. Results of a Monte Carlo analysis of our test statistics can be found in section 3.3.

3.1 Maximum Likelihood Estimation

In order to conduct a unit root test on an individual variable x_t one can run a ADF test regression

$$\Delta x_t = \delta z_t + \alpha x_{t-1} + \gamma w_{pt} + \epsilon_t;\ t = 1, \ldots, T. \tag{8}$$

In (8) $\Delta x_t = x_t - x_{t-1}$, the $m \times 1$ deterministic components vector z_t either contains a constant: $z_t = 1$, or a constant plus a linear time trend: $z_t = (1 \quad t)'$ with the $1 \times m$ coefficient vector δ, and $w_{pt} = (\Delta x_{t-1} \cdots \Delta x_{t-p})'$ with the $1 \times p$ coefficient vector γ. The unit root test is a test if in (8) $\alpha = 0$.

To conduct unit root testing on a variable x_{it} of the i^{th} cross-section within a panel of N cross-section observations and T time series observations , we can stack N ADF regressions like (8) into one system,

$$\begin{aligned}\Delta X_t &= \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_N \end{pmatrix} z_t + \begin{pmatrix} \alpha_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \alpha_N \end{pmatrix} X_{t-1} + \begin{pmatrix} \gamma_1 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \gamma_N \end{pmatrix} W_{pt} + \varepsilon_t \\ &= \bar{\delta} z_t + \Phi X_{t-1} + \Gamma W_{pt} + \varepsilon_t, \end{aligned} \tag{9}$$

where α_i relates Δx_{it} to $x_{i,t-1}$ and γ_i relates Δx_{it} to $\Delta x_{i,t-1}, \ldots, \Delta x_{i,t-p_i}$.[5] The model in (9) consists of the $N \times 1$ vectors $X_{t-1} = (x_{1,t-1} \cdots x_{N,t-1})'$, $\Delta X_t = X_t - X_{t-1}$ and $\varepsilon_t = (\epsilon'_{1t} \cdots \epsilon'_{Nt})'$, and the $(\sum_{i=1}^N p_i) \times 1$ vector $W_{pt} = (w'_{p_1 t,1} \cdots w'_{p_N t,N})'$ for $t = 1, \ldots, T$ and $i = 1, \ldots, N$. In (9) z_t, x_{it}, $w_{p,i}$, δ_i, α_i and γ_i have an identical definition as in (8) for $i = 1, \ldots, N$, and the coefficient matrices $\bar{\delta}$, Φ and Γ have dimensions equal to $N \times m$, $N \times N$ and $N \times (\sum_{i=1}^N p_i)$ respectively. We assume a multivariate normal distribution for the disturbance vector ε_t: $\varepsilon_t \sim \mathrm{N}(\mathbf{0}_N,\ \Omega)$ with the $N \times N$ covariance matrix structure,

$$\Omega = \begin{pmatrix} \omega_{11} & \cdots & \omega_{1N} \\ \vdots & \ddots & \vdots \\ \omega_{N1} & \cdots & \omega_{NN} \end{pmatrix}. \tag{10}$$

In (10) $\omega_{ij} \equiv \mathrm{Cov}(\epsilon_{it},\ \epsilon_{jt})$ for $i, j = 1, \ldots, N$.

The panel of the N variables $x_{1t}, \ldots, x_{Nt}$ in (9) can be interpreted as a restricted vector autoregressive [VAR] model. This restricted VAR model hinges on the following assumption:

Assumption 3.1 *There is* ***no*** *linear dependence between the variable x_{it} of individual i and lags of the variable x_{jt} of individual j for $i \neq j$.*

[5]Note that the number of lagged first differences can differ across the equations of (9).

Proper estimation of the restricted VAR model (9) involves the usage of feasible GLS (or SURE), see Lütkepohl (1993, Section 5.2). Unit root testing across N cross-sections simultaneously within the restricted VAR model (9) involves testing the parameter restriction $\alpha_1 = \cdots = \alpha_N = 0$. Interpreting the panel as a restricted VAR model allows us to adopt the estimation and testing framework for VAR models to analyze panels with a limited cross-section dimension.

The log-likelihood function for model (9) can be written as,[6]

$$\ell(\bar{\delta},\ \Phi,\ \Gamma,\ \Omega) = -\frac{NT}{2}\ln(2\pi) + \frac{T}{2}\ln|\Omega^{-1}| - \frac{1}{2}tr\left(\Omega^{-1}(\Delta X - Z\bar{\delta}' - X_{-1}\Phi' - W_p\Gamma')'(\Delta X - Z\bar{\delta}' - X_{-1}\Phi' - W_p\Gamma')\right), \quad (11)$$

where $\bar{\delta}$, Φ and Γ are defined in (9) and Ω has an identical structure as (10). The $T \times N$ matrices ΔX, X_{-1} and the $T \times (\sum_{i=1}^{N} p_i)$ matrix W_p in (11) can be defined as:

$$\Delta X = \begin{pmatrix} \Delta X_1' \\ \vdots \\ \Delta X_T' \end{pmatrix},\ X_{-1} = \begin{pmatrix} X_0' \\ \vdots \\ X_{T-1}' \end{pmatrix} \quad \text{and} \quad W_p = \begin{pmatrix} W_{p1}' \\ \vdots \\ W_{pT}' \end{pmatrix},$$

and the $T \times m$ matrix Z equals $Z = \iota_T$ or $Z = (\iota\ \tau)$ with ι_T is a $T \times 1$ vector of ones and the $T \times 1$ vector $\tau = (1 \cdots T)'$.

Maximum likelihood estimates of $\bar{\delta}$, Φ, Γ and the disturbance covariance matrix Ω in (9) can be obtained through iterative SURE (ISURE). Essential for this ISURE procedure is proper estimation of Ω, and based on (11) Ω is estimated with the standard conditional maximum likelihood estimator:[7]

$$\hat{\Omega}(\hat{\bar{\delta}}, \hat{\Phi}, \hat{\Gamma}) = \frac{1}{T}\left(\Delta X - Z\hat{\bar{\delta}}' - X_{-1}\hat{\Phi}' - W_P\hat{\Gamma}'\right)'\left(\Delta X - Z\hat{\bar{\delta}}' - X_{-1}\hat{\Phi}' - W_P\hat{\Gamma}'\right). \quad (12)$$

The ISURE procedure starts off with a consistent *initial* estimate of Ω:

$$\hat{\Omega}(\hat{\Phi}_{\text{OLS}}) = \left(\hat{\Omega}_{ij}\right)_{i,j=1,\ldots,N} \text{with } \hat{\Omega}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{it}\hat{\epsilon}_{jt}'. \quad (13)$$

In (13) $\hat{\epsilon}_{it}$ and $\hat{\epsilon}_{jt}$ are residuals resulting from N OLS regressions of Δx_{it} on z_t, $x_{i,t-1}$ and $\Delta x_{i,t-1}, \ldots, \Delta x_{i,t-p}$ as in (8). The initial estimate (13) is used to estimate $\bar{\delta}$, Φ and Γ through SURE and these SURE estimates in turn can be used to construct a new estimate of Ω based on (12). Next, we can construct *new* SURE estimates of $\bar{\delta}$, Φ and Γ using the estimate of Ω based on the *old* SURE estimates of $\bar{\delta}$, Φ and Γ. Magnus (1978) shows that iterating in this manner until convergence of the estimators yields maximum likelihood estimates of $\bar{\delta}$, Φ, Γ and Ω.

[6] The determinant of Ω^{-1} is indicated with $|\Omega^{-1}|$ and the trace of a matrix is indicated with $tr(\cdots)$.

[7] The number of time series T is identical for each equation as this greatly simplifies the estimation of covariance matrix Ω. Hence, we consider in this paper only systems with balanced times series observations.

3.2 Multivariate Unit Root Testing

For unit root testing across N individuals simultaneously based on a specification like (9), we make use of SURE estimators as outlined in section 3.1. Hence, we can only consider the cases where $T > N$ and the limiting behaviour of our test statistics are based on large T asymptotics while assuming a fixed cross section dimension N.

As the null hypothesis of N unit roots involves a restriction on N parameters simultaneously, we shall use a likelihood ratio test to test for non-stationarity in our SURE system. One can straightforwardly show that the maximized value of (11) conditional on the maximum likelihood estimates $\hat{\hat{\delta}}$, $\hat{\Phi}$ and $\hat{\Gamma}$ in combination with disturbance covariance matrix estimator (12) can be specified as[8]

$$\ell_{\max}[\hat{\hat{\delta}}, \hat{\Phi}, \hat{\Gamma}, \hat{\Omega}(\hat{\hat{\delta}}, \hat{\Phi}, \hat{\Gamma})] = \ell^1_{\max} = -\frac{NT}{2}(1 + \ln(2\pi)) - \frac{T}{2}\ln|\hat{\Omega}(\hat{\hat{\delta}}, \hat{\Phi}, \hat{\Gamma})|. \tag{14}$$

Under the unit root restriction, *i.e.* $\alpha_1 = \cdots = \alpha_N = 0$ in (9), maximum likelihood estimation is identical as in section 3.1 but without $x_{1,t-1}, \ldots, x_{N,t-1}$ included in our restricted VAR model. The corresponding maximized log-likelihood function equals:

$$\ell_{\max}[\hat{\hat{\delta}}, \hat{\Gamma}, \hat{\Omega}(\hat{\hat{\delta}}, \hat{\Gamma})] = \ell^0_{\max} = -\frac{NT}{2}(1 + \ln(2\pi)) - \frac{T}{2}\ln|\hat{\Omega}(\hat{\hat{\delta}}, \hat{\Gamma})|. \tag{15}$$

The likelihood ratio test statistic for $H_0\colon \alpha_1 = \cdots = \alpha_N = 0$ within (9) versus $H_1\colon \alpha_i \neq 0$ for $i = 1, \ldots, N$ is now identical to:

$$\mathrm{LR}_{\Phi=0} = 2(\ell^1_{\max} - \ell^0_{\max}) = T[\ln|\hat{\Omega}(\hat{\hat{\delta}}, \hat{\Gamma})| - \ln|\hat{\Omega}(\hat{\hat{\delta}}, \hat{\Phi}, \hat{\Gamma})|]. \tag{16}$$

The asymptotic behaviour of the multivariate unit root test statistic in (16) can be typified as

Proposition 3.1 *Let,*

(a) *the estimates of $\delta_1, \ldots, \delta_N$, $\alpha_1, \ldots, \alpha_N$, $\gamma_1, \ldots, \gamma_N$ and Ω be fully converged estimates from the iterative estimation schemes of section 3.1 both under the null hypothesis ($\alpha_1 = \cdots = \alpha_N = 0$) and the alternative hypothesis,*

(b) *each of the N series $x_{1t}, \ldots, x_{Nt}$ be I(1),*

(c) *the cross-section dimension N be fixed and the time series dimension $T \to \infty$.*

Then the limiting distribution of $LR_{\Phi=0}$ in (16) equals:

$$LR_{\Phi=0} \Rightarrow \sum_{i=1}^{N} \left[\left(\int \check{B}_i dB_i \right)^2 \left(\int \check{B}_i^2 \right)^{-1} \right]. \tag{17}$$

[8] Note that $\ln|\Omega^{-1}| = -\ln|\Omega|$.

In (17) "$\Rightarrow$" denotes convergence in distribution, $B_i(u)$ is a scalar standard Brownian motion for individual i on the interval $u \in [0,\ 1]$, $\int \check{B}_i dB_i \equiv \int_0^1 \check{B}_i(u) dB_i(u) du$ and $\check{B}_i(u) = B_i(u)$ if in (9) $\delta_1 = \cdots = \delta_N = 0$ or $\check{B}_i(u) = \bar{B}_i(u)$. When appropriate, $\bar{B}_i(u)$ equals for individual i $\bar{B}_i(u) = B_i(u) - \int_0^1 B(u) du$ if in (9) $z_t = 1$ or $\bar{B}_i(u) = B_i(u) - a_i - b_i t$ if in (9) $z_t = (1\ \ t)$ with a_i and b_i resulting from regressing $B_i(u)$ on a constant and a linear time trend.

Proof: *See Appendix A.*

Expression (17) is identical to a summation of N squared Dickey and Fuller (1979) limiting distributions for the univariate ADF unit root test. Appendix B describes how we compute the critical values for test statistic (16) based on the asymptotic distributions from proposition 3.1.

The finite sample properties of test statistic (16) can be improved through a degrees of freedom correction as suggested by Sims (1980). It involves replacing T in (16) by the average degrees of freedom per cross-section under the alternative hypothesis:

$$\mathrm{CLR}_{\Phi=0} = (T-d)[\ln|\hat{\Omega}(\hat{\bar{\delta}}, \hat{\Gamma})| - \ln|\hat{\Omega}(\hat{\bar{\delta}}, \hat{\Phi}, \hat{\Gamma})|], \tag{18}$$

where[9]

$$d = \frac{1}{N}\left(N(m+1) + \sum_{i=1}^{N} p_i \right).$$

Obviously, the corrected likelihood ratio test statistic (18) has smaller values than (16) and in finite samples combined with a large number of parameters $\mathrm{CLR}_{\Phi=0}$ could under a true null very well be much closer to the asymptotic distribution (17) than $\mathrm{LR}_{\Phi=0}$.

3.3 Monte Carlo Evaluation

To study the behaviour of our multivariate unit root test statistics (16) and (18) we conduct a Monte Carlo analysis on artificial samples with comparable dimensions as the multi-country systems used in section 4. The Monte Carlo experiments have the same set-up as in section 2 and, as in section 2, these experiments are based on 10,000 replications, $T = 100$ and $N = 3$, 6 or 9.

The results of the Monte Carlo experiments on our multivariate unit root tests are reported in table 2. When we have no serially correlated innovations we see that both the $\mathrm{LR}_{\Phi=0}$ and the $\mathrm{CLR}_{\Phi=0}$ statistics have a correct size at the 95% quantile from distribution (17). As in section 2 we have fitted our SURE system (9) in case of first order serially correlated innovations with a *common* lag order p equal to 1, 2 and 3. For $p = 1$ we have again in all cases a correct size. When the utilized lag order increases from 1 to 2 and 3

[9] The number of deterministic components per cross-section equals m ($m = 0$, $m = 1$ or $m = 2$), the number of lagged first differences per cross-section equals p_j and we have 1 lagged level $x_{i,t-1}$ per cross-section.

we see in table 2 that at $N = 9$ the $\text{LR}_{\Phi=0}$ statistics has a tendency to slightly overreject the true null hypothesis. The $\text{CLR}_{\Phi=0}$ statistic, however, retains a correct size when at $N = 9$ the lag order increases to 2 and 3. Overall, the $\text{CLR}_{\Phi=0}$ statistic has a better size than the $\text{LR}_{\Phi=0}$ statistic when the number of parameters increase substantially.

When we look at the power ratios in table 2 we see that at $N = 3$ we have for both multivariate unit root test statistics power ratios in the range of 60%-72%. An increase in the number of series from 3 to 6 and 9 results in a substantial increase in the power ratios to levels beyond the 90% value. Next, compare the power performance of our multivariate unit root tests with the power results for the univariate ADF test and the two panel unit root tests in table 1. Such a comparison makes it clear that already at very moderate cross-section dimensions our $\text{LR}_{\Phi=0}$ and $\text{CLR}_{\Phi=0}$ statistics have a superior power performance relative to both univariate and panel unit root tests. This is caused by the fact that within our framework the N test regressions are *jointly* estimated by taken into account the covariances between the N cross-sections, rather than estimating the N regressions *separately* as in the case of the ADF and Im *et al.* (1997) tests. As such our approach yields more efficient estimates of the mean reversion parameters $\alpha_1, \dots, \alpha_N$ resulting in a higher power under a true alternative hypothesis. Hence, our likelihood-based multivariate unit root test statistics are the most appropriate for a multi-country analysis of real exchange rates.

4 A Multi-Country Test of PPP

In this section we apply the multivariate unit root testing framework from section 3 on the real exchange rates of the G10 countries in order to test the validity of PPP for all these countries. Section 4.1 contains an description of the data. Also, we conduct in this subsection univariate unit root tests on bilateral G10 real exchange rates relative to the U.S., Germany Japan and the U.K. Next, we report in section 4.2 multivariate unit root test results for our four sets of G10 bilateral real exchange rates.

4.1 The Data and Univariate Unit Root Test Results

In its logarithmic form the real exchange rate for the home country *versus* a foreign country is defined as

$$q = e + p^* - p, \tag{19}$$

where q, e, p^* and p are the logarithm of the real exchange rate, the nominal exchange rate, the foreign aggregate price level and the home aggregate price level respectively. *Long-run* PPP is valid when the real exchange rate has a constant mean through time, implying an equalized relative competitiveness in the long-run between two countries.

Thus q in (19) must be stationary, *i.e.* one should reject the null hypothesis

$$H_0:\ \Delta q_t = \sum_{j=1}^{p} \gamma_j \Delta q_{t-j} + \epsilon_t,\ \epsilon \sim \text{i.i.d.}(0, \sigma^2);\ t = 1, \ldots, T, \tag{20}$$

in favor of the alternative hypothesis

$$H_1:\ \Delta q_t = \delta + \alpha q_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta q_{t-j} + \epsilon_t,\ \alpha < 0. \tag{21}$$

An intercept δ is included in (21) to correct for measurement errors due to the fact that we use in practice price indices and not actual price levels. Note that (21) allows for *short-run* deviations from PPP.

We consider real exchange rates for 10 of the most important industrialized countries [G10], *i.e.* Canada, France, Germany, Italy, Japan, The Netherlands, Sweden, Switzerland, the U.K. and the U.S. Quarterly observations from 1973.1 through 1997.4 are used in the estimation of our systems of real exchange rates. Logarithms of real exchange rates are constructed as in (19), where we use the consumer price index [CPI] as a proxy of the aggregate price level. Data on the CPI's and exchange rates are obtained from the IMF's *International Financial Statistics* [IFS].[10] G10 real exchange rates are constructed relative to four *numeraire* countries: the U.S., Germany, Japan and the U.K. In constructing real exchange rates relative to the U.S. we use quarterly average U.S. dollar exchange rates as the CPI data are also quarterly averages.[11] In case of real exchange rates relative to Germany, Japan and the U.K., the nominal exchange rates are calculated through cross-rates based on the U.S. dollar exchange rates.

To get a feel of the degree of persistence within bilateral G10 real exchange rates, we conduct univariate ADF unit root tests for G10 real exchange rates relative to our four base countries. We use the ADF unit root test with a constant included in the test regression, that is we conduct a t-test for $\alpha = 0$ in (21). The lag order for the ADF test regressions is selected as follows. First, we determine an optimal lag order through the Schwartz Information Criterion [SIC], based on a comparison of SIC criteria computed for lag orders ranging from 0 to 8 in (21). Next, we used Lagrange-Multiplier [LM] serial correlation tests at 1, 4 and 8 lags to determine whether the residuals of (21) at the optimal SIC lag order are white noise. If that is not the case, we increase the lag order until the LM serial correlation tests indicate that the residuals of (21) are indeed white noise.

From table 3 it becomes clear that irrespective of the base country univariate unit root tests are in general not able to reject the null of non-stationary real exchange rates. The ADF tests for Germany-based real exchange rates provide the most favorable evidence for the PPP hypothesis, as we can reject the null of non-stationarity for France and

[10] The CPI data are from IFS line code 64.
[11] The exchange rate data are from IFS line code "rf".

Switzerland. For the other base countries we are only able to reject the null of non-stationary real exchange rates in case of the real exchange rates of Switzerland relative to Japan and the U.K. The estimated measures of mean reversion seems to have more favorable values when we use Germany and the U.K. as the base countries. All things considered, the results in table 3 indicate that even if real exchange rates are stationary their degree of persistence is such that *univariate* unit root tests are not able reject the null of non-stationarity.

4.2 Multivariate Unit Root Test Results

The failure of univariate unit root tests to reject in section 4.1 the null of non-stationary real exchange rates could be due to slow rates of mean reversion such that one only can find evidence for stationarity within samples of data with a long time span. One possible solution is the usage of panel techniques described in section 2, but these techniques are based on the possibly invalid assumption of homogeneous cross-country rates of mean reversion. As an alternative we apply in this subsection the multivariate framework of section 3.

For G10 bilateral real exchange rates relative to the U.S., Germany, Japan and the U.K. we conduct ISURE estimation on a system like (9) with $z_t = 1$, *i.e.*

$$\begin{array}{ccccccccc} \Delta q_{1t} & = & \delta_1 & + & \alpha_1 q_{1,t-1} & + & \sum_{j=1}^{p_1} \gamma_{1j} \Delta q_{1,t-j} & + & \epsilon_{1t} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \Delta q_{Nt} & = & \delta_N & + & \alpha_N q_{N,t-1} & + & \sum_{j=1}^{p_N} \gamma_{Nj} \Delta q_{N,t-j} & + & \epsilon_{Nt} \end{array} \tag{22}$$

and $N = 9$. Next, after we have estimated (22) under the restriction $\alpha_1 = \cdots = \alpha_N = 0$ we construct our likelihood ratio unit root test statistics (16) and (18). The lag orders for each equation of system (22) are selected through the "bottom-up" approach of Hsiao (1979, 1982) and Lütkepohl (1993, pp.182-183) for restricted VAR models: appropriate lag orders are selected for each equation separately based on the optimal SIC for that particular equation. Given assumption 3.1, we already applied such a strategy in section 4.1 in determining the lag orders for the univariate ADF tests. Hence, we use for each equation of (22) the lag order of the corresponding equation in table 3.

Conditional on the stationarity of our G10 real exchange rates, we are also interested in the speed of mean reversion after the occurrence of a shock in a real exchange rate. In order to achieve this we reinterpret (22), analogous to section 3.1, as a restricted VAR model:

$$Q_t = \Theta_1 Q_{t-1} + \cdots + \Theta_{p_{\max}+1} Q_{t-p_{\max}-1} + \bar{\delta} + \varepsilon_t. \tag{23}$$

In (23) $Q_t = (q_{1t} \cdots q_{Nt})'$, $\bar{\delta} = (\delta_1 \cdots \delta_N)'$ from (22), $p_{\max}$ is the maximum of $p_1, \ldots, p_N$ in (22) and Θ_j is a $N \times N$ diagonal matrix for $j = 1, \ldots, p_{\max} + 1$:

$$\Theta_j = \begin{pmatrix} \theta_{1j} & 0 \cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 \cdots 0 & \theta_{Nj} \end{pmatrix},$$

where if $p_i < p_{\max}$ then $\theta_{i,p_i+2} = \cdots = \theta_{i,p_{\max}+1} = 0$ for $i = 1, \ldots, N$. Define the lag operator L such that $L^j q_{it} = q_{i,t-j}$ and $L^j Q_t = Q_{t-j}$, and define the matrix lag polynomial

$$\begin{aligned}\Theta(L) &= I_N - \Theta_1 L - \cdots - \Theta_{p_{\max}+1} L^{p_{\max}+1} \\ &= (I_N - \bar{\Theta} L) - (\sum_{s=1}^{p_{\max}} \Theta_s^* L^s)(1 - L),\end{aligned} \tag{24}$$

where

$$\bar{\Theta} = \sum_{j=1}^{p_{\max}+1} \Theta_j \quad \text{and} \quad \Theta_s^* = \sum_{l=s+1}^{p_{\max}+1} -\Theta_l.$$

Hence, in (22) $\alpha_1, \ldots, \alpha_N$ and $\gamma_{1j}, \ldots, \gamma_{Nj}$ are equal to the diagonal elements of $(\bar{\Theta} - I_N)$ and Θ_s^* respectively.

The fact that we can read the panel model in (22) as the restricted VAR model in (23) gives us the opportunity to calculate the mean reversion speeds for each cross-section through the corresponding impulse response function. That is, td_i is the number of periods after which in absolute terms $\tau \times 100\%$ of a unit shock in real exchange rate q_{it} has been reversed:[12]

$$td_i = \max(d) \quad \text{for} \quad d = 1, 2, \ldots \quad \text{until} \quad \left| \frac{\partial q_{i,t+d}}{\partial \epsilon_{it}} \right| \leq 1 - \tau, \tag{25}$$

with $0 < \tau < 1$ and $i = 1, \ldots, N$. The estimated mean reversion speeds $td_1, \ldots, td_N$ in (25) depend on estimates of the parameters in (22) and parameter uncertainty thus has an impact on the estimates of $td_1, \ldots, td_N$. We therefore compute 95 % confidence intervals for the estimated mean reversion speeds, based on 10,000 parametric bootstrap simulations. These parametric bootstrap simulations are organized as follows:

- the initial startup values for $q_1, \ldots, q_N$ are taken from the historical data,
- a sequence of $(T - p_{\max} - 1)$ ϵ_{it}'s are drawn for $i = 1, \ldots, N$ from a N-dimensional multivariate normal distribution calibrated to the estimation of (22),
- given the initial values, the artificial $\epsilon_{1t}, \ldots, \epsilon_{Nt}$ and the parameters from (22) estimated on the historical data, we generate for each $i = 1, \ldots, N$ $(T - p_{\max} - 1)$ artificial values of q_{it},
- we re-estimate (22) on these artificial $q_{1t}, \ldots, q_{Nt}$ and calculate for each i the corresponding td_i through (25).

[12] See Ng and Perron (1999). An alternative measure of mean reversion speed equals $\ln(1-\tau)/\ln(1+\alpha_i)$, but this measure only uses the *sum* of autoregressive parameters (see (24)). Our measure of adjustment speed utilizes the moving average representation $Q_t = \bar{\delta} + \Theta(L)^{-1}\varepsilon_t$ and as such makes use of *all* the individual autoregressive parameters.

The resulting parametric bootstrap samples of td_i's are then used to compute the 95% confidence intervals.

The results of the ISURE estimation of panel model (22) for our four panels of G10 bilateral real exchange rates are summarized in table 4. The results for the likelihood-based test statistics $\mathrm{LR}_{\Phi=0}$, reported in the lower part of table 4, indicates a rejection of the null of non-stationary real exchange rates for all four sets of G10 real exchange rates. In case of the base countries Germany and the U.K. the p-values of the test statistics indicate that we easily can reject the null at the 5% significance level. When Japan is used as the *numeraire* country we can reject the null comfortably at the 1% significance level, whereas the results for the U.S. indicates a rejection at the 10% significance level with a p-value very close to 5%. The test results are qualitatively the same when we take into account the number of parameters through the usage of the $\mathrm{CLR}_{\Phi=0}$ statistic. The power analysis in section 3.3 indicates that with cross-correlated data, our multivariate framework has ample power to reject the null when the data are stationary but persistent in nature. On average the cross-correlations of the relative changes in real exchange rates with respect to the U.S., Germany, Japan and the U.K. equals 0.56, 0.33, 0.69 and 0.61 respectively. Consequently, we would expect *a priori* that our multivariate unit root test statistics yield more positive results with respect to the PPP hypothesis than the univariate unit root tests, especially for the base countries the U.S., Japan and the U.K. The multivariate unit root test results in table 4 therefore confirms our prior that the usage of cross-country information in the analysis of real exchange rates results in more positive findings regarding the PPP hypothesis.

In the upper part of table 4 we report the maximum likelihood estimates of the mean reversion coefficients, and the cross-country variability of these estimates within a panel seems to depend on the choice of the *numeraire* country. One notices from table 4 that when the U.S. is used as the base country the estimated α_i's are, with the exception of Canada, very close to each other. In contrast we observe for the base countries Germany and Japan that the estimated mean reversion coefficients per country are quite heterogeneous in nature, *i.e.* they range from -0.036 to -0.197 and -0.033 to -0.145 for Germany and Japan respectively. However, a more fruitful way to determine the degrees of persistence is to look at impulse response functions, as in (25), instead of the α_i's which sum away the information in the individual autoregressive parameters (see footnote 12). The results of such an approach can be found in table 5 where we report the rate in quarters at which 50% and 90% of a shock in an individual real exchange rates has been reverted, that is we compute $td_1, \ldots, td_N$ in (25) for $\tau = 0.5$ and $\tau = 0.9$. We have chosen $\tau = 0.5$ to determine the half life of a shock in a real exchange rate and $\tau = 0.9$ is chosen in order to pin point the period after which a shock does not anymore has an economically significant influence.

From the first column of table 5 we observe that the estimated half life of a shock in the Canada-U.S. real exchange rate equals 6.75 years, whereas the half lives for the remaining U.S.-based rates are more or less identical to a period of 3 years. The cross-country differences across the U.S.-based real exchange rates becomes more pronounced when we look at the 90% absorption rates. In this case France, Germany, Italy, Japan, the Netherlands and Switzerland exhibit a 90% absorption after about 6.5 years, but note that the corresponding confidence interval for Japan is much wider. For Canada, Sweden and the U.K. a shock is for 90% reversed after approximately 8.5 years. In the case of base country Germany we have half lives of 1.5 and 2 years for France and the U.K., which is significantly lower than the half lives of 4.5 and 5.25 years for the U.S. and Canada. In the remaining Germany-based real exchange rates we have a 50% shock reversion after approximately 3.25 years, but the corresponding confidence intervals indicate that the difference with Canada and the U.S. is in reality small. Judging from the fourth column of table 5 the Germany-based rates of France, the Netherlands and the U.K. have a 90% shock reversion after 2, 5.25 and 5.5 years respectively. In all other cases the adjustment speed is significantly slower, especially for Italy and Switzerland. In column 5 a shock in the Switzerland-Japan rate has a half life of 1.25 years, Canada and the U.S. have significantly larger half lives of 5.75 and 4.75 years respectively and all other Japan-based rates have an identical 50% absorption pace of about 2 years. The conclusions for the 90% shock reversion periods are similar: the Switzerland-Japan rate has the highest full adjustment speed with 3.25 years and the North-American rates versus Japan have the slowest full adjustment speeds of 11 to 14 years. Finally, we report the 50% and 90% absorption rates for U.K.-based real exchange rates in columns 7 and 8 of table 5. The North-American rates versus the U.K. have a significantly longer duration of mean reversion than for the other U.K.-based rates with half lives of at least 5 years and full absorption after at least 11.75 years.

The discussion of the results regarding the mean reversion speeds in table 5 indicates that cross-country parameter heterogeneity in the four panels seems to be caused by the behaviour of sub-groups of bilateral relationships within each panel. Most noticeably is the behaviour of the Canadian and, if appropriate, U.S. real exchange rates relative to our four base countries, where shocks can have an influence on these respective rates of up to 10 years on average. In contrast to that we observe for the France-Germany real exchange rate, the U.K.-Germany rate and the European Japan-based rates that shocks in the respective real exchange rates die out relatively fast. We can relate these differences to the dominance of monetary versus real shocks: inflationary/deflationary shocks are of short duration whereas Balassa-Samuelson-type shocks trigger persistent deviations in real exchange rates from their long-run mean, see section 1. Both France and the U.K. has known periods with significantly looser monetary policy relative to Germany and Japan since the early 1990s has experienced a pronounced price deflation. Therefore, the high mean reversion speeds in these real exchange rates seems to be due to the dominance of monetary shocks, which is also reflected by the relatively narrow confidence intervals of these rates at both the 50% and 90% absorption rates. The high persistence in the North American real exchange rates could very well be caused by the Balassa-Samuelson effect. The very large corresponding confidence intervals at the 90% absorption rate seem to confirm this, as a shock in, for example, the Canada-U.S. real exchange rate can potentially influence this rate up to 28 years! We can now also interpret another peculiarity in table 5: for the Italy-Germany, Switzerland-Germany and Sweden-U.K. real exchange rates the reversion of the first 50% of a shock takes place at a faster pace than for the remainder of such a shock. Apparently, these real exchange rates are influenced by both inflation shocks and Balassa-Samuelson-type shocks, resulting in 90% absorption periods which are three times as large as the corresponding half life.

5 Conclusions

The validity of long-run PPP implies that real exchange rates are stationary, *i.e.* in time real exchange rates revert back to a constant mean. This paper proposes and employs a multivariate framework for unit root testing in multi-country panels of real exchange rates, while retaining cross-country differences in mean reversion rates. By treating the panel of data explicitly as a restricted, high dimensional VAR model we are able to derive appropriate estimation and testing methods based on the corresponding log-likelihood function. Utilizing time series-based asymptotics in combination with a fixed cross-section dimension enables us to derive limiting distributions which are also applicable when the series in the panel are contemporaneously correlated. Monte Carlo experiments for systems with empirically sensible dimensions show that our multivariate unit root test statistic behaves well both under a true null of non-stationarity and under a true alternative of stationarity, especially when a degrees-of-freedom correction is employed. The Monte Carlo results indicate that our multivariate unit root test is not only robust to cross-correlations in the data, the usage of cross-correlated data also improves the power of the test significantly.

The empirical tests are conducted on the bilateral real exchange rates of 10 large industrialized [G10] countries. We construct four panels of G10 real exchange rates relative to the U.S., Germany, Japan and the U.K. In all four multi-country panels our multivariate approach is able to reject the null of non-stationary real exchange rates. When we look at the time necessary to have a 50% and 90% completion of an adjustment to a shock, it becomes apparent that there are significant cross-country differences within each panel. As these differences seems to be concentrated within a sub-group of real exchange rates within each panel, we postulate that parameter heterogeneity within our G10 panels is mainly caused by the predominance of Balassa-Samuelson-type shocks in certain countries.

It would be interesting for future research to assess more explicitly the part monetary and real shocks play in the observed asymmetric pattern of mean reversion. A further research topic is to apply the framework of this paper on real exchange rates based on disaggregated price data, *e.g.* city-based price indexes or sector-based prices. Finally, based on an appropriately restricted disturbance covariance matrix our framework could be extended to the case where we have both a large number of cross-sections and time series observations.

Appendix

A Proof of Proposition 3.1

In the following proofs we discard the presence of lagged first differences in (9), and we assume that we have $\gamma_1 = \cdots = \gamma_N = 0$ combined with a vector of disturbances ε_t which does not exhibit serial correlation. From Dickey and Fuller (1979) and Said and Dickey (1984) we know that the inclusion of lagged first differences within ADF test regressions, in order to guarantee white noise innovations, does not influence the asymptotic behaviour of the ADF t-statistic relative to the case of no higher order dynamics. Johansen (1991) has an identical result in the case of likelihood ratio cointegration rank statistics within unrestricted VAR models of non-stationary variables. As (9) can both be considered as a system of N ADF test regressions and as a restricted VAR model of N non-stationary variables, $\mathrm{LR}_{\Phi=0}$ is under the null asymptotically identical whether or not $\gamma_1 = \cdots = \gamma_N = 0$ in (9) as long as we have white noise disturbances. Hence, for notational convenience we base all our proofs on the absence of higher order dynamics in (9). Also, our proofs are at first based on the absence of deterministic components in (9) but we discuss at the end of this Appendix the extension to the case of deterministic components.

In deriving the limiting behaviour of $\mathrm{LR}_{\Phi=0}$ we make use of the following results:

1. We make use of the properties of "vec"-operators and Kronecker-product operators as summarized in Lütkepohl (1993, Appendix A.11 and A.12), we use in particular:

$$\begin{aligned} \mathrm{vec}(ABC) &= (C' \otimes A)\mathrm{vec}(B), \\ (A \otimes B)(C \otimes D) &= (AC \otimes BD), \end{aligned} \tag{A.1}$$

   where A, B and C are appropriate matrices and "vec" denotes vectorization of a matrix by stacking the columns of this matrix.

2. For $T \to \infty$ we have (see Hamilton 1994, chapters 17 and 18):

$$\begin{aligned} \frac{1}{T^2} X_{-1}' X_{-1} &\Rightarrow \Omega^{\frac{1}{2}} \left(\int W_N W_N' \right) \Omega^{\frac{1}{2}'}, \\ \frac{1}{T} X_{-1}' \Delta X &\Rightarrow \Omega^{\frac{1}{2}} \left(\int W_N dW_N' \right) \Omega^{\frac{1}{2}'}. \end{aligned} \tag{A.2}$$

In (A.2) $W_N(u) = (B_1(u) \cdots B_N(u))'$ is a N-dimensional vector Brownian Motion with covariance matrix I_N and $u \in [0,\ 1]$, $B_i(u)$ is a scalar standard Brownian Motion, $\int W_N dW_N' \equiv \int_0^1 W_N(u) dW_N(u) du$, and Ω is the true non-diagonal disturbance covariance matrix as in (10). Note that "$\Rightarrow$" indicates convergence in distribution, whereas in the remainder of this Appendix "$\xrightarrow{p}$" indicates convergence in probability.

The Proof

For $\delta_1 = \cdots = \delta_N = 0$, $\gamma_1 = \cdots = \gamma_N = 0$ and no serial correlation within the innovations vector ε_t in (9), log-likelihood function (11) can also be written as

$$\ell(\Phi,\ \Omega) = -\frac{NT}{2}\ln(2\pi) + \frac{T}{2}\ln|\Omega^{-1}| \\ - \frac{1}{2}\mathrm{vec}(\Delta X - X_{-1}\Phi)'(\Omega^{-1} \otimes I_T)\mathrm{vec}(\Delta X - X_{-1}\Phi), \quad (A.3)$$

with the $T \times T$ identity matrix I_T. We can write within the last part of log-likelihood function (A.3)

$$\begin{aligned} \mathrm{vec}(\Delta X - X_{-1}\Phi) = & \\ & \mathrm{vec}\,(\Delta X) - \mathrm{vec}\,(X_{1,-1}\alpha_1 \ldots X_{N,-1}\alpha_N) \\ & = \mathrm{vec}\,(\Delta X) - (I_N \otimes X_{-1})F_\Phi \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix}, \end{aligned} \quad (A.4)$$

where I_N is a $N \times N$ identity matrix and F_Φ is a $N^2 \times N$ selection matrix,

$$F_\Phi = \begin{pmatrix} e_1 & \mathbf{0}_N & \cdots & \mathbf{0}_N \\ \mathbf{0}_N & e_2 & & \mathbf{0}_N \\ \mathbf{0}_{N(N-3)} & \mathbf{0}_{N(N-3)} & \ddots & \mathbf{0}_{N(N-3)} \\ \mathbf{0}_N & \mathbf{0}_N & & e_N \end{pmatrix} = ((e_1 \otimes e_1) \cdots (e_N \otimes e_N)). \quad (A.5)$$

In (A.5) e_i is the i^{th} column of the identity matrix I_N and $\mathbf{0}_s$ is a s-dimensional column vector of zeros with $s = N$ or $N(N-3)$. Substituting (A.4) in log-likelihood (A.3) and maximizing (A.3) with respect to Φ given Ω, yields the following estimator of $\alpha_1, \ldots, \alpha_N$:

$$\begin{pmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_N \end{pmatrix} = (F'_\Phi(\Omega^{-1} \otimes X'_{-1}X_{-1})F_\Phi)^{-1}F'_\Phi(\Omega^{-1} \otimes I_T)\mathrm{vec}(X'_{-1}\Delta X), \quad (A.6)$$

which is a SURE estimator. The conditional maximum likelihood estimator of Ω given the estimate $\hat{\Phi}$ equals:

$$\hat{\Omega}(\hat{\Phi}) = \frac{1}{T}\left(\Delta X - X_{-1}\hat{\Phi}\right)'\left(\Delta X - X_{-1}\hat{\Phi}\right). \quad (A.7)$$

Using (A.6) and (A.7) in the ISURE procedures from section 3.1 yields maximum likelihood estimates but Magnus (1978) has shown that the estimates after one iteration have the same asymptotic distribution as fully converged estimates. In the following we make use of this property of the one-step SURE estimator.

Following Lütkepohl (1993, pp.123-124), it can be shown that the likelihood ratio statistic for the null hypothesis $\alpha_1 = \cdots = \alpha_N = 0$ versus $\alpha_i \neq 0$ can be written as

$$\text{LR}_{\Phi=0} = 2\left[\ell(\hat{\Phi},\ \Omega) - \ell(\Omega)\right] = \text{vec}(X_{-1}\hat{\Phi})'(\hat{\Omega}^{-1} \otimes I_T)\text{vec}(X_{-1}\hat{\Phi}) + \text{o}_p(1), \tag{A.8}$$

where $\hat{\Phi}$ contains the estimated α_i's from (A.6) and $\hat{\Omega}$ is a consistent estimate of the disturbance covariance matrix Ω. Under $H_0 : \alpha_1 = \cdots = \alpha_N = 0$, $\hat{\Omega}(\hat{\Phi})$ in (A.7) is a consistent estimate of Ω. Hence, given (A.8) and $\hat{\Omega} = \hat{\Omega}(\hat{\Phi}) \xrightarrow{p} \Omega$ we have

$$\begin{aligned} \text{LR}_{\Phi=0} &\backsimeq \text{vec}(X_{-1}\hat{\Phi})'(\hat{\Omega}^{-1} \otimes I_T)\text{vec}(X_{-1}\hat{\Phi}) \\ &= \begin{pmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_N \end{pmatrix}' F_\Phi'(\hat{\Omega}^{-1} \otimes X_{-1}'X_{-1})F_\Phi \begin{pmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_N \end{pmatrix} \\ &= (F_\Phi \text{vec}(X_{-1}'\Delta X\hat{\Omega}^{-1}))'[F_\Phi'(\hat{\Omega}^{-1} \otimes X_{-1}'X_{-1})F_\Phi]^{-1}F_\Phi'\text{vec}(X_{-1}'\Delta X\hat{\Omega}^{-1}), \end{aligned} \tag{A.9}$$

where F_Φ is defined in (A.5) and the third expression results from substituting estimator (A.6).

Based on $\hat{\Omega} = \hat{\Omega}(\hat{\Phi}) \xrightarrow{p} \Omega$, (A.1), (A.2) and the continuous mapping theorem, we have for $T \to \infty$:

$$\begin{aligned} \frac{1}{T^2}[F_\Phi'(\hat{\Omega}^{-1} \otimes X_{-1}'X_{-1})F_\Phi]^{-1} &\Rightarrow [F_\Phi'(\Omega^{-1} \otimes \Omega^{\frac{1}{2}} \int W_N W_N' \Omega^{\frac{1}{2}'})F_\Phi]^{-1} \\ &= [F_\Phi'(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})(I_N \otimes \int W_N W_N')(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})'F_\Phi]^{-1}, \end{aligned} \tag{A.10}$$

and

$$\begin{aligned} \frac{1}{T}F_\Phi'\text{vec}(X_{-1}'\Delta X\hat{\Omega}^{-1}) &\Rightarrow F_\Phi'\text{vec}(\Omega^{\frac{1}{2}} \int W_N dW_N' \Omega^{-\frac{1}{2}'}) \\ &= F_\Phi'(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})\text{vec}(\int W_N dW_N'). \end{aligned} \tag{A.11}$$

In order to manipulate the expressions in (A.10) and (A.11) we define the following:

$$\begin{aligned} \Omega^{-\frac{1}{2}} &= (\Psi_1' \cdots \Psi_N')', \quad \text{with } \Psi_i \text{ is } 1 \times N, \\ \Omega^{\frac{1}{2}} &= (\Upsilon_1' \cdots \Upsilon_N')', \quad \text{with } \Upsilon_i \text{ is } 1 \times N. \end{aligned} \tag{A.12}$$

Utilizing (A.12) we can now write

$$\begin{aligned} F_\Phi'(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})'F_\Phi &= \begin{pmatrix} \Psi_1 \otimes \Upsilon_1 \\ \vdots \\ \Psi_N \otimes \Upsilon_N \end{pmatrix} \begin{pmatrix} \Psi_1 \otimes \Upsilon_1 \\ \vdots \\ \Psi_N \otimes \Upsilon_N \end{pmatrix}' \\ &= \begin{pmatrix} (\Psi_1\Psi_1')(\Upsilon_1\Upsilon_1') & \cdots & (\Psi_1\Psi_N')(\Upsilon_1\Upsilon_N') \\ \vdots & \ddots & \vdots \\ (\Psi_N\Psi_1')(\Upsilon_N\Upsilon_1') & \cdots & (\Psi_N\Psi_N')(\Upsilon_N\Upsilon_N') \end{pmatrix} \\ &= PP', \end{aligned} \tag{A.13}$$

where PP' is the Choleski decomposition of the $N \times N$ matrix in the second right hand side expression in (A.13). Using (A.13), pre-multiplying the expression within square brackets in (A.10) with P^{-1} and post-multiplying with $P^{-1'}$ yields

$$[P^{-1}F'_{\Phi}(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})(I_N \otimes \int W_N W'_N)(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})' F_{\Phi} P^{-1'}]^{-1}$$

$$= \left(P^{-1} \begin{pmatrix} (\Psi_1\Psi'_1)(\Upsilon_1 \int W_N W'_N \Upsilon'_1) & \cdots & (\Psi_1\Psi'_N)(\Upsilon_1 \int W_N W'_N \Upsilon'_N) \\ \vdots & \ddots & \vdots \\ (\Psi_N\Psi'_1)(\Upsilon_N \int W_N W'_N \Upsilon'_1) & \cdots & (\Psi_N\Psi'_N)(\Upsilon_N \int W_N W'_N \Upsilon'_N) \end{pmatrix} P^{-1'} \right)^{-1}$$

$$= \begin{pmatrix} \int B_1^2 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \int B_N^2 \end{pmatrix}^{-1}. \qquad \text{(A.14)}$$

For (A.11) we have, based on (A.1) and (A.13), the following result:

$$P^{-1}F'_{\Phi}(\Omega^{-\frac{1}{2}} \otimes \Omega^{\frac{1}{2}})\text{vec}(\int W_N dW'_N) = P^{-1} \begin{pmatrix} \Upsilon_1 \int W_N dW'_N \Psi'_1 \\ \vdots \\ \Upsilon_N \int W_N dW'_N \Psi'_N \end{pmatrix} = \begin{pmatrix} \int B_1 dB_1 \\ \vdots \\ \int B_N dB_N \end{pmatrix}. \qquad \text{(A.15)}$$

As the P matrix appears in both (A.14) and (A.15), we are able to substitute (A.14) and (A.15) in (A.9) and this results in the following limiting expression for $\text{LR}_{\Phi=0}$:

$$\text{LR}_{\Phi=0} \Rightarrow \begin{pmatrix} \int B_1 dB_1 \\ \vdots \\ \int B_N dB_N \end{pmatrix}' \begin{pmatrix} \int B_1^2 & 0\cdots 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0\cdots 0 & \int B_N^2 \end{pmatrix}^{-1} \begin{pmatrix} \int B_1 dB_1 \\ \vdots \\ \int B_N dB_N \end{pmatrix} = \sum_{i=1}^{N} \left[\left(\int B_i dB_i \right)^2 \left(\int B_i^2 \right)^{-1} \right]. \qquad \blacksquare$$

Deterministic Components

We can concentrate log-likelihood function (11) with respect to the deterministic components through OLS regressions of the elements of ΔX_t, X_{t-1} and W_{pt} on the deterministic component vector z_t, as z_t has an identical content for each equation of (9).[13] Hence, we

[13] See also the Frisch-Waugh-Lovell theorem in Davidson and MacKinnon (1993, pp. 19-24).

have after adjusting for the effect of the deterministic components:

$$\begin{aligned} \Delta\tilde{X}_i &= M_Z \Delta X_i \text{ and } \Delta\tilde{X} = (\Delta\tilde{X}_1 \cdots \Delta\tilde{X}_N), \\ \tilde{X}_{i,-1} &= M_Z X_{i,-1} \text{ and } \tilde{X}_{-1} = (\tilde{X}_{1,t-1} \cdots \tilde{X}_{N,-1}), \\ \tilde{W}_{p,i} &= M_Z W_{p,i} \text{ and } \tilde{W}_p = (\tilde{W}_{p,1} \cdots \tilde{W}_{p,N}), \end{aligned} \tag{A.16}$$

with $M_Z = I_T - Z(Z'Z)^{-1}Z'$. Replacing ΔX, Z, X_{-1} and W_p with the variables of (A.16) in the ISURE procedure from section 3.1 yields therefore identical maximum likelihood estimates of Φ and Γ in (9) as in the original ISURE procedure. Under the null of N non-stationary variables, *i.e.* $\delta_1 = \cdots = \delta_N = \alpha_1 = \cdots = \alpha_N = 0$, we now have $\bar{B}_i(u) = B_i(u) - \int_0^1 B_i(u)du$ or $\bar{B}_i(u) = B_i(u) - a_i - b_i t$, and $dB_i(u) - \int_0^1 dB_i(u)du = dB_i(u)$ or $dB_i(u) - a_i - b_i t = dB_i(u)$.[14] Hence, we replace in all relevant formulae of the previously described proof $W_N(u)$ with $\bar{W}_N(u) = (\bar{B}_1(u) \cdots \bar{B}_N(u))'$ while retaining dW_N.

B Critical Values

The asymptotic distribution of our multivariate likelihood ratio unit root test, as summarized in proposition 3.1, is a functional of Brownian Motions. As these are continuous time variables, one has to rely in practice on approximations to get proper critical values for our multivariate unit root tests. Nielsen (1997) observes that within a single equation model the asymptotic behaviour of a likelihood ratio unit root test is very well approximated by a Gamma-distribution, especially for quantiles $\geq 50\%$. The limiting distribution of a likelihood ratio unit root test within the single equation framework equals a squared Dickey and Fuller (1979) distribution and the limiting distribution in proposition 3.1 equals a summation of N squared Dickey-Fuller distributions. Hence, we can use a Gamma-distribution to approximate the curvature of our asymptotic distributions.

The usage of the Gamma-distribution has several advantageous features. First, we do not have to simulate and report critical values for our multivariate unit root tests at every value of N and the Gamma-distribution can therefore be considered as a "cross-sectional response surface". Next, the usage of the Gamma-distribution allows one to compute the p-values of the multivariate unit root test in a convenient way. The Gamma-distribution can be written as

$$\Gamma(z; r, a) = \int_0^z \frac{a^r}{\Gamma(r)} x^{r-1} \exp(-ax)dx, \quad z > 0,\ r > 0,\ a > 0, \tag{B.1}$$

where $\Gamma(\cdot)$ is the Gamma-function. When we can find proper values for the parameters a and r, we can use (B.1) to approximate the distribution of our test statistic z under the null. Following Doornik (1998), we can calibrate (B.1) through

$$\hat{a} = \frac{m}{v}, \quad \hat{r} = \frac{m^2}{v}, \tag{B.2}$$

[14] Parameters a_i and b_i results from regressing $B_i(u)$ on an intercept and a linear time trend.

where m is the mean of z under the null and v is the variance. Doornik (1998) shows in Monte Carlo experiments that the above described procedures yields very accurate approximations of the asymptotic distributions of multivariate cointegration tests, which basically are squared multivariate Dickey-Fuller distributions.

Proposition 3.1 indicates that the asymptotic distribution of our test statistic is a summation of N independent, squared Dickey-Fuller distributions. Therefore, the mean and variance of these distributions equals N times the mean and N times the variance of a single squared Dickey-Fuller distribution. Thus, we first approximate the mean and variance of a single squared Dickey-Fuller distribution. MacKinnon (1991) provides very accurate approximations of the 10%, 5% and 1% quantiles of the Dickey-Fuller distribution using response surface regressions. Consequently we also utilize response surface regressions to approximate the mean and variance of a single *squared* Dickey-Fuller distribution. In these response surface regressions we use 13 different values of the number of time series observations T: 50, 65, 80, 100, 150, 200, 250, 350, 500, 750, 1,000, 2,000 and 5,000. To control for experimental randomness we performed 40 separate Monte Carlo simulations for every T each with 50,000 replications except for $T > 750$ where we used 25,000 replications in each experiment. In each replication we generate a discrete time random walk with $T + 1$ observations, compute the Dickey-Fuller t-value and take the square of this t-value, and calculate the mean and variance across the 25,000 or 50,000 generated squared t-values. As such the mean and variance of the squared Dickey-Fuller distribution are for each value of T based on either 2 million iterations or in case of $T > 750$ on 1 million replications. These exercises are repeated for specifications with a constant or a constant plus trend added to the test regression, where we use either a demeaned or a detrended random walk.

For each deterministic specification we now have 520 approximations for both the mean and the variance of the single squared Dickey-Fuller distribution at various T and analogous to MacKinnon (1991) we use these approximations to fit a response surface regression for both the mean and variance:

$$C_i^l = \theta_\infty^l + \theta_1^l T_i^{-1} + \theta_2^l T_i^{-2} + e_i^l, \quad i = 1, \dots, 520; \; l = \text{mean or variance.} \tag{B.3}$$

In (B.3) T_i is the number of time series observations in the i^{th} experiment and C_i^l is the estimate of either the mean or variance from the i^{th} experiment. The first parameter θ_∞^l is either the mean or the variance of the asymptotic squared Dickey-Fuller distribution and the other two parameters allows one to determine the mean or variance in finite samples. The error terms e_i^l are heteroskedastic and (B.3) is therefore estimated with a weighted least squares [WLS] procedure. In this procedure we first regress C_i^l on 13 dummy variables, where the first dummy variable is equal to 1 if $T_i = 50$, the second is 1 if $T_i = 65$ and so forth, resulting in the residuals $\check{e}_1^l, \dots, \check{e}_{520}^l$. Next, we regress $(\check{e}_i^l)^2$ on a constant, T_i^{-1} and T_i^{-2}, and the inverses of the square roots of the fitted values of this auxiliary regression are used as weights in WLS estimation of (B.3). We used tests on both the individual and joint significance of parameters plus the Schwartz Information Criterion [SIC] to check the adequacy of the specification in (B.3) relative to other possible specifications such as adding T_i^{-3} as an extra regressor to (B.3) or deleting T_i^{-2} from the equation. The specification tests favored in case of the mean always a version of (B.3) with solely T_i^{-1}. The response surface regression of the variance in the case of no deterministic components also included only T_i^{-1} and for the other cases the response surface for the variance included both T_i^{-1} and T_i^{-2}. See also the expressions within parentheses in table B.1.

Approximations for the asymptotic or exact sample mean and variance of our distributions in proposition 3.1 are now equal to N times the corresponding fitted value of θ^l_∞ or C^l_i in (B.3), see also table B.1. Based on these approximations of m and v we can now determine the values of r and a in (B.1) through (B.2). The resulting calibrated Gamma-distribution can now be used to compute asymptotic or exact sample critical values or p-values for our multivariate unit root test.[15]

Table B.1: Mean and variance of the limiting distributions of Proposition 3.1[a]

Case	m	v
1	$(1.1420 + 0.690T^{-1})\,N$	$(2.2243 + 9.128T^{-1})\,N$
2	$(3.0573 + 1.548T^{-1})\,N$	$(7.0103 + 41.004T^{-1} + 239.48T^{-2})\,N$
3	$(5.3235 + 2.179T^{-1})\,N$	$(11.2478 + 94.101T^{-1} + 504.40T^{-2})\,N$

[a] The values equal N times the mean or variance of a single squared Dickey-Fuller distribution approximated by a response surface regression as in (B.3). The denomination N indicates the cross-section dimension, T is the (balanced) number of time series observations and m, v indicates the approximations of the mean and variance of limiting distribution (17) respectively. Case 1 is the specification without deterministic components, Case 2 is the specification with a constant for each cross-section and Case 3 is the specification with a constant and linear time trend for each cross-section.

References

Abuaf, N. and P. Jorion, 1990, Purchasing Power Parity in the Long Run, *Journal of Finance* **44**, 157–175.

Balassa, B., 1964, The Purchasing Power Parity Doctrine: A Reappraisal, *Journal of Political Economy* **72**, 584–596.

[15] A GAUSS procedure for calculating the 90%, 95% and 99% quantiles or p-values based on the fitted Gamma-distribution is available from the author.

Coakley, J. and A.M. Fuertes, 1997, New Panel Unit Root Tests of PPP, *Economics Letters* **57**, 17–22.

Davidson, R. and J.G. MacKinnon, 1993, *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

Dickey, D.F. 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.

Doornik, J.A., 1998, Approximations to the Asymptotic Distributions of Cointegration Tests, *Journal of Economic Surveys* **12**, 573–593.

Dumas, B., 1992, Dynamic Equilibrium and the Real Exchange Rate in a Spatially Seperated World, *Review of Financial Studies* **5**, 153–180.

Engel, C., M.K. Hendrickson and J.H. Rogers, 1997, Intra-National, Intra-Continental and Intra-Planetary PPP, *Journal of the Japanese and International Economies* **11**, 480–501.

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

Frenkel, J., 1981, Flexible Exchange Rates, Prices and the Role of 'News': Lessons from the 1970s, *Journal of Political Economy* **89**, 665–705.

Froot, K.A. and K. Rogoff, 1995, Perspectives on PPP and Long-Run Real Exchange Rates, *in* G. Grossman and K. Rogoff (editors), *Handbook of International Economics, Vol. III*, North Holland, Amsterdam.

Hakkio, C.S., 1984, A Re-examination of Purchasing Power Parity: A Multi-country and Multi-period Study, *Journal of International Economics* **15**, 265–278.

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

Hsiao, C., 1979, Autoregressive Modelling of Canadian Money and Income Data, *Journal of the American Statistical Association* **74**, 553–560.

Hsiao, C., 1982, Time Series Modelling and Causal Ordering of Canadian Money, Income and Interest Rates, *in* O.D. Anderson (editor), *Time Series Analysis: Theory and Practice*, North-Holland, Amsterdam.

Im, K.S., M.H. Pesaran and Y. Shin, 1997, Testing for Unit Roots in Heterogeneous Panels, *mimeo*, Department of Applied Economics, Cambridge University.

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

Jorion, P. and R. Sweeney, 1996, Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting, *Journal of International Money and Finance* **15**, 535–550.

Levin, A. and Chien-Fu Lin, 1992, Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties, *Discussion Paper 92-23*, Department of Economics, University of California at San Diego.

Lütkepohl, H., 1993, *Introduction to Multiple Time Series Analysis*, 2nd edition, Springer-Verlag, Berlin.

MacDonald, R., 1996, Panel Unit Root Tests and Real Exchange Rates, *Economics Letters* **50**, 7–11.

MacKinnon, J.G., 1991, Critical Values for Cointegration Tests, *in* R.F. Engle and C.W.J. Granger (editors), *Long Run Economic Relationships: Readings in Cointegration*, Oxford University Press, Oxford, pp. 267–276.

Maddala, G.S. and S. Wu, 1996, A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test: Evidence from Simulations and the Bootstrap, *mimeo*, Department of Economics, Ohio State University.

Magnus, J.R., 1978, Maximum Likelihood Estimation of the GLS Model with Unknown Parameters in the Disturbance Covariance Matrix, *Journal of Econometrics* **7**, 281–312.

Mark, N.C., 1990, Real Exchange Rates in the Long-Run: An Empirical Investigation, *Journal of International Economics* **28**, 115–136.

Ng, S. and P. Perron, 1999, PPP May Not Hold Afterall: A Further Investigation, *Working Paper 444*, Department of Economics, Boston College.

Nielsen, B., 1997, Bartlett Correction of the Unit Root Test in Autoregressive Models, *Biometrika* **84**, 500–504.

O'Connell, P., 1998, The Overvaluation of Purchasing Power Parity, *Journal of International Economics* **44**, 1–19.

Oh, K., 1996, Purchasing Power Parity and Unit Root Tests Using Panel Data, *Journal of International Money and Finance* **15**, 405–418.

Papell, D.H., 1997, Searching for Stationarity: Purchasing Power Parity Under the Current Float, *Journal of International Economics* **43**, 313–332.

Said, S. and D.A. Dickey, 1984, Testing for Unit Roots in Autoregressive Moving-Average Models of Unknown Order, *Biometrika* **71**, 599–607.

Samuelson, P.A., 1964, Theoretical Notes on Trade Problems, *Review of Economics and Statistics* **46**, 145–154.

Shiller, R. and P. Perron, 1985, Testing the Random Walk Hypothesis: Power versus Frequency of Observation, *Economics Letters* **18**, 381–386.

Sims, C.A., 1980, Macroeconomics and Reality, *Econometrica* **48**, 1–48.

Table 1: Size and power of the Levin and Lin (1992) and Im *et al.* (1997) panel unit root tests with constant terms in cross-correlated data for a nominal size of 5%.[a]

		No Serial Correlation		Serial Correlation					
				$p=1$		$p=2$		$p=3$	
		Size	Power	Size	Power	Size	Power	Size	Power
Augmented Dickey and Fuller (1979) Unit Root Test									
		0.050	0.170	0.047	0.153	0.048	0.132	0.044	0.129
Panel Unit Root Tests									
$N=3$	LL	0.157	0.344	0.140	0.287	0.138	0.272	0.143	0.257
	IPS	0.072	0.222	0.073	0.185	0.078	0.187	0.078	0.177
$N=6$	LL	0.241	0.449	0.205	0.371	0.207	0.349	0.211	0.338
	IPS	0.136	0.378	0.142	0.330	0.147	0.331	0.146	0.310
$N=9$	LL	0.303	0.519	0.244	0.427	0.253	0.398	0.260	0.370
	IPS	0.182	0.472	0.179	0.429	0.184	0.419	0.189	0.392

[a] The Monte Carlo experiments are based on $T = 100$ and 10,000 simulations both with or without first order serially correlated innovations, see the text. Denomination p indicates the lag order used in the test procedures. Rows with "LL" ("IPS") report the results for the Levin and Lin (1992) (Im *et al.* 1997) panel unit root test. The results for the univariate ADF test are based on the appropriate 5% critical value from MacKinnon (1991).

Table 2: Size and power of the multivariate unit root test with constant terms for a nominal size of 5%.[a]

		No Serial Correlation		Serial Correlation					
				$p=1$		$p=2$		$p=3$	
		Size	Power	Size	Power	Size	Power	Size	Power
$N=3$	$\mathrm{LR}_{\Phi=0}$	0.050	0.729	0.054	0.699	0.057	0.666	0.060	0.637
	$\mathrm{CLR}_{\Phi=0}$	0.044	0.717	0.045	0.677	0.045	0.640	0.045	0.599
$N=6$	$\mathrm{LR}_{\Phi=0}$	0.057	0.965	0.068	0.955	0.076	0.943	0.073	0.930
	$\mathrm{CLR}_{\Phi=0}$	0.049	0.962	0.055	0.948	0.057	0.931	0.054	0.913
$N=9$	$\mathrm{LR}_{\Phi=0}$	0.067	0.997	0.073	0.994	0.090	0.992	0.101	0.988
	$\mathrm{CLR}_{\Phi=0}$	0.057	0.996	0.059	0.993	0.065	0.990	0.067	0.985
		95% Quantiles							
$N=3$		18.112		18.116		18.119		18.123	
$N=6$		30.631		30.636		30.641		30.647	
$N=9$		42.356		42.362		42.369		42.375	

[a] See the notes of table 1. The statistics $\mathrm{LR}_{\Phi=0}$ and $\mathrm{CLR}_{\Phi=0}$ are defined in (16) and (18). Size and power calculations are based on the exact sample 95% quantiles in the lower part of the table, which are computed through the procedures of Appendix B.

Table 3: Univariate unit root test results for G10 real exchange rates, 1973.1-1997.4[a]

Relative to:	U.S.			Germany			Japan			U.K.		
	α	t_α	p	α	t_α	p	α	t_α	p	α	t_α	p
Canada	−0.030	−1.615	3	−0.071	−2.328	3	−0.028	−1.504	1	−0.058	−1.800	1
France	−0.070	−2.280	1	−0.185	−3.080**	6	−0.047	−2.052	1	−0.083	−2.063	2
Germany	−0.072	−2.360	3	–	–	–	−0.044	−1.935	2	−0.077	−2.375	2
Italy	−0.072	−2.252	1	−0.052	−2.043	3	−0.067	−2.382	1	−0.092	−2.127	2
Japan	−0.037	−1.689	1	−0.051	−2.294	2	–	–	–	−0.060	−2.234	1
Netherlands	−0.066	−2.092	1	−0.087	−2.292	8	−0.037	−1.684	1	−0.070	−2.059	2
Sweden	−0.054	−1.880	1	−0.102	−2.554	1	−0.039	−1.823	1	−0.048	−1.485	1
Switzerland	−0.072	−2.283	1	−0.078	−3.289**	2	−0.092	−2.652*	1	−0.112	−2.945**	1
U.K.	−0.087	−2.232	1	−0.077	−2.375	2	−0.060	−2.234	1	–	–	–
U.S.	–	–	–	−0.072	−2.360	3	−0.037	−1.689	1	−0.087	−2.232	1

[a] The ADF test regression is $\Delta q_t = \delta + \alpha q_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta q_{t-j} + \epsilon_t$, with α is the rate of mean reversion, t_α is the t-statistic for the hypothesis $\alpha = 0$ and p is the lag order. An *(**)[***] indicates rejection of the null of non-stationarity at the 10% (5%) [1%] significance level based on the appropriate critical values from MacKinnon (1991).

Table 4: Multivariate unit root test results for G10 real exchange rates, 1973.1-1997.4[a]

Relative to:	U.S.	Germany	Japan	U.K.
	$\hat{\alpha}_i$	$\hat{\alpha}_i$	$\hat{\alpha}_i$	$\hat{\alpha}_i$
Canada	−0.025	−0.036	−0.033	−0.035
France	−0.065	−0.197	−0.090	−0.057
Germany	−0.072	—-	−0.106	−0.073
Italy	−0.075	−0.050	−0.100	−0.086
Japan	−0.065	−0.058	—-	−0.067
Netherlands	−0.072	−0.093	−0.089	−0.060
Sweden	−0.055	−0.065	−0.075	−0.046
Switzerland	−0.079	−0.056	−0.145	−0.096
U.K.	−0.056	−0.089	−0.083	—-
U.S.	—-	−0.041	−0.040	−0.038
	Likelihood Ratio Unit Root Tests			
$LR_{\Phi=0}$	42.196	45.431	69.884	45.723
	(0.052)	(0.028)	(0.000)	(0.026)
$CLR_{\Phi=0}$	40.682	42.769	67.642	44.099
	(0.069)	(0.047)	(0.000)	(0.036)

[a] ISURE estimates of α_i in (22) equal $\hat{\alpha}_i$. "$LR_{\Phi=0}$" and "$CLR_{\Phi=0}$" are likelihood ratio statistics for the null of N unit roots, with the corresponding p-values within parentheses (see Appendix B).

Table 5: Mean reversion speeds across G10 real exchange rates, 1973.1-1997.4[a]

Relative to:	U.S.		Germany		Japan		U.K.	
	50%	90%	50%	90%	50%	90%	50%	90%
Canada	27 (12-59)	33 (15-112)	21 (9-35)	33 (13-72)	23 (8-35)	56 (18-91)	22 (8-33)	51 (16-80)
France	12 (6-15)	29 (13-38)	6 (4-8)	8 (6-16)	9 (5-10)	20 (12-25)	13 (6-17)	38 (15-54)
Germany	11 (6-13)	27 (13-34)	—- —-	—- —-	7 (5-8)	19 (12-22)	10 (6-12)	28 (15-35)
Italy	10 (5-13)	25 (11-35)	15 (3-36)	42 (10-118)	8 (5-9)	19 (10-23)	9 (4-11)	23 (9-33)
Japan	12 (5-19)	26 (9-45)	13 (5-21)	31 (10-55)	—- —-	—- —-	11 (5-18)	26 (9-45)
Netherlands	11 (6-12)	27 (13-34)	13 (3-21)	21 (11-33)	8 (5-10)	23 (13-28)	12 (6-16)	34 (15-46)
Sweden	14 (5-23)	35 (12-60)	11 (4-22)	31 (9-63)	10 (5-13)	26 (13-34)	16 (6-28)	46 (15-84)
Switzerland	10 (5-12)	25 (12-33)	13 (3-27)	38 (11-86)	5 (4-7)	13 (7-17)	8 (4-10)	21 (10-29)
U.K.	13 (5-23)	36 (11-66)	8 (4-14)	22 (7-40)	9 (5-13)	21 (8-32)	—- —-	—- —-
U.S.	—- —-	—- —-	18 (10-26)	26 (13-50)	19 (7-27)	47 (15-72)	20 (7-28)	47 (15-71)

[a] The columns labeled with "50%" ("90%") report the number of quarters after which 50% (90%) of a shock in the real exchange rate has been reversed, calculated through (25). The corresponding 95% confidence intervals, based on 10,000 parametric bootstrap simulations, are reported in parentheses.