Purchasing Power Parity: Three Stakes Through the Heart of the Unit Root Null†

Matthew Higgins
Research Department
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045
e-mail: Matthew.Higgins@ny.frb.org

Egon Zakrajšek
Division of Monetary Affairs, MS84
Board of Governors of the Federal Reserve System
20th Street & Constitution Avenue, NW
Washington D.C., 20551
e-mail: Egon.Zakrajsek@frb.gov

April 2000

†We thank Linda Goldberg, Jim Harrigan, Ken Kuttner, Jonathan McCarthy, Gabriel Perez Quiros, Mihael Perman, Simon Potter, Rob Rich, Eric van Wincoop, Kei-Mu Yi, and seminar participants at the University of Ljubljana for their valuable comments and suggestions. All remaining errors and omissions are our own responsibility. The opinions expressed in this paper do not necessarily reflect views of the Federal Reserve Bank of New York, Board of Governors, nor the Federal Reserve System.

Abstract

We provide a comprehensive analysis of the purchasing power parity hypothesis, relying on a linear panel data framework. First, we consider two panel unit root tests, based on transformations of country-specific statistics, which allow for parameter heterogeneity across countries. Using GLS techniques, we modify the two tests to eliminate the upward size distortion induced by cross-sectional dependence among contemporaneous real exchange rate innovations. Second, we consider two tests based on a fixed-effects specification: these tests allow for cross-sectional dependence but impose parameter homogeneity. Three of the four tests provide emphatic support for real exchange rate stationary during the post-Bretton Woods era among relatively open economies. Monte Carlo experiments indicate that the three tests have considerable power against the unit root null. One test allowing parameter heterogeneity provides mixed support for stationarity, but has only limited power against the null.

JEL Classification: C15, C33, F15, F31
Keywords: purchasing power parity, panel data unit root tests

1 Introduction

Purchasing power parity (PPP) is a key building block of many models in international macroeconomics. Yet it is by now well-known that the null hypothesis that real exchange rates contain a unit root cannot generally be rejected using univariate tests for data from the post-Bretton Woods era.[1] Frankel (1986, 1990) and Froot and Rogoff (1995) show that failure to reject the unit root null may be driven by the low power of univariate tests against persistent alternatives. Recently, several authors have turned to panel data unit root tests in an attempt to gain statistical power. This work uniformly finds evidence of real exchange rate stationarity among developed countries during the recent float.[2]

Recent work by O'Connell (1998) challenges the emerging consensus of real exchange rate stationarity. O'Connell shows that the most widely used panel unit root test, introduced by Levin and Lin (1992), suffers from substantial upward size distortion in the presence of cross-sectional dependence among contemporaneous real exchange rate innovations. Using GLS techniques, he develops a new panel unit root test that eliminates the upward size distortion. Using the new test (OC–GLS hereafter), O'Connell fails to reject the unit root null for several country panels, thus undermining earlier panel data evidence in favor of reversion to purchasing power parity.

We provide a comprehensive analysis of the PPP hypothesis, relying on a linear panel data framework. First, to eliminate the upward size distortion induced by contemporaneous cross-sectional dependence, we use GLS transformations to modify the panel unit root tests introduced by Im, Pesaran, and Shin (1995), as well as a second test introduced by Levin and Lin (1993). (We call the modified tests IPS–GLS and LL2–GLS, respectively.) These two tests are based on statistics generated from country-specific augmented Dickey–Fuller (ADF) regressions. Importantly, the tests allow the rate of convergence to PPP and the pattern of serial correlation among real exchange rate innovations to vary across countries.

Second, we exploit a restricted version of the SUR-based panel unit root test recently introduced by Sarno and Taylor (1998a) (SUR–GLS hereafter) and re-examine O'Connell's (1998) benchmark GLS test. These two tests are based on a fixed-effects ADF specification, which assumes a common rate of convergence to long-run equilibrium and a common pattern of temporal dependence in the real exchange rate innovation process.

Using the IPS–GLS and SUR–GLS tests, we find overwhelming evidence of real exchange

[1]See, for example, Meese and Rogoff (1988), Mark (1990), and Pappel and Theodoridis (1998).

[2]Examples include MacDonald (1996), Frankel and Rose (1996), Oh (1996), Wu (1996), and Pappel and Theodoridis (1998). The rapidly growing body of work which applies panel data techniques to real exchange rate behavior also includes Abuaf and Jorion (1990), Wei and Parsley (1996), Flood and Taylor (1996), Frankel and Rose (1996), Jorion and Sweeney (1996), and Koedijk, Schotman, and Dijk (1998). Of these papers, only Koedijk, Schotman, and Dijk (1998) address the issue of cross-sectional dependence of interest here. The authors' approach, however, differs from our own in treating PPP as the null rather than the alternative hypothesis.

rate stationarity during the post-Bretton Woods era for three panels of relatively open economies: *i*) 32 countries classified as economically open by Sachs and Warner (1995); *ii*) the 25 countries belonging to the OECD as of 1995; and *iii*) 19 European countries. For the European and OECD samples, we also find strong support for real exchange rate stationarity using O'Connell's original GLS test. For all three samples, we find mixed support for stationarity using the LL2–GLS test.

Monte Carlo experiments indicate that the IPS–GLS, SUR–GLS, and OC–GLS tests have considerable power against the unit root null. The exact power ranking depends on the cross-sectional and time-series dimensions of the panel, as well as the degree of temporal persistence among real exchange rate innovations. The high power attained by the IPS–GLS test is notable, given that it allows for complete parameter heterogeneity across countries. The LL2–GLS test, in contrast, has limited power against the unit root null, especially in the presence of serially correlated innovation process.

Bias-corrected estimates of the rate of convergence to long-run equilibrium indicate that deviations from PPP erode more quickly for real exchange rates defined using wholesale rather than consumer price indices. This result is consistent with the higher share of tradable goods in the wholesale category.

2 PPP and the Unit Root Hypothesis

In its simplest and strongest form, purchasing power parity states that the price of similar goods sold in two countries should be equal when expressed in a common currency. This hypothesis, known as absolute purchasing power parity, implies the equality of national price levels, provided that national price indices assign a common set of weights to all goods.

Empirical work, on the other hand, has focused on the weaker concept of relative purchasing power parity: a higher rate of domestic inflation—relative to a numeraire country—should be systematically offset by depreciation against the numeraire currency.[3] Relative purchasing power parity implies that real exchange rates—the ratio of a country's nominal exchange rate to its relative price index—should be stationary.

Define the (natural) log of the real exchange rate, q, as:

$$q \equiv (e - e^*) - (p - p^*),$$

where e denotes the log of the nominal ($US) exchange rate of the domestic country, e^* is the log of the nominal exchange rate of the country used as a numeraire, p is the log of the

[3]Relative PPP allows for a constant, unobservable differential between different countries' consumption (or production) baskets. Consequently, tests of relative PPP can rely on data concerning country price indices, rather than (generally unavailable) data concerning absolute country price levels.

domestic price level, and p^* is the log of the foreign price level. Note that if p and p^* are measured using consumer price indices (CPIs), the real exchange rate expresses the price of the consumption basket of the numeraire country in terms of the domestic consumption basket. In our analysis, we always treat the U.S. as the numeraire country, so that e^* is equal to 0.

Under the null hypothesis, it is assumed that *each* country's real exchange rate, q, contains a unit root, and that the first difference of the real exchange rate, Δq, is stationary. Under the alternative hypothesis, *all* real exchange rates are assumed to be stationary. Formally, the evolution of the real exchange rate for country i in period t, $q_{i,t}$ is described by the following data generating process (DGP):

$$\Delta q_{i,t} = \alpha_i + \rho_i q_{i,t-1} + \eta_{i,t}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{1}$$

where α_i denotes a fixed country-specific effect, ρ_i is the parameter measuring the rate of reversion to the relative PPP for country i, and $\eta_{i,t}$ is the real exchange rate innovation for country i in period t..

Under the null hypothesis of a unit root, $\alpha_i = 0$ and $\rho_i = 0$, for all i.[4] Therefore, innovations to the real exchange rate, $\eta_{i,t}$, have a permanent effect on the level of the real exchange rate $q_{i,t}$. Under the alternative hypothesis of stationarity $\rho_i < 0$, for all i, so that innovations to the real exchange rate decay at the rates ρ_i. Moreover, under the alternative, the inclusion of fixed individual effects allows the unconditional mean of $q_{i,t}$ to differ across countries. The long-run equilibrium value of country i's real exchange rate, therefore, is given by $q_i^{LR} = -\alpha_i/\rho_i$. [5]

The simple data generating process described by equation (1) can easily be extended to allow for general ARMA representation of the innovation process under the null hypothesis. For example, one or more lagged real exchange rate changes $(\Delta q_{i,t-1}, \Delta q_{i,t-2}, \ldots)$ can be added as regressors to equation (1) to control for serial correlation in the innovation process, yielding a panel data extension of the standard ADF framework. In addition, disturbances

[4]Banerjee et al. (1996) describe how common factor restrictions can be used to write the null hypothesis in this composite form. Under the composite null, the country-specific fixed effect α_i is absorbed into the country-specific initial value $q_{i,0}$, so that $\alpha_i = 0$, for all i.

[5]Unlike some authors, we do not address the possibility that the DGP for real exchange rate changes might contain a linear time trend. That is,

$$\Delta q_{i,t} = \alpha_{1,i} + \alpha_{2,i} t + \rho_i q_{i,t-1} + \eta_{i,t}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T.$$

The null of a unit root in the above equation implies that $\alpha_{1,i} = \alpha_{2,i} = 0$ and $\rho_i = 0$, for all i. Under the stationary alternative, $\rho_i < 0$ and $\alpha_{1,i}$ and $\alpha_{2,i}$ are unrestricted. That is, under the alternative hypothesis, real exchange rates are stationary around a country-specific deterministic trend. Consequently, the real exchange rates for a given country pair i and j can drift apart indefinitely at the rate $\alpha_{2,i} - \alpha_{2,j}$. Our view is that this implication is at odds with the economic content of the relative purchasing power parity hypothesis—that common-currency price levels should not drift apart indefinitely—but do not consider the issue further here.

that affect all countries equally in a given period t—a pattern of *homogeneous* cross-sectional dependence—can easily be accommodated by adding fixed time effects to equation (1), or equivalently, by expressing all variables as deviations from their time-specific means.

3 Panel Data Unit Root Tests

In this section, we review the most commonly used linear panel unit root tests. A common assumption behind all the tests considered in this paper is that the real exchange rate DGP can be described by a panel data extension of the univariate ADF framework. The tests can then be classified according to the restrictions imposed on the assumed DGP. For instance, some tests require the rate of convergence to long-run equilibrium under the alternative to be the same across all countries—that is $\rho_i = \rho$, for all i—while others allow the rate of convergence to vary across countries.

Similarly, some tests impose a common autoregressive error structure on the real exchange rate innovations, while others allow it to differ across countries. Some of the tests are designed to accommodate only a homogeneous pattern of cross-sectional correlation among real exchange rate innovations, while others are designed to accommodate an arbitrary pattern of contemporaneous cross-sectional dependence. We show below that simple GLS transformation techniques can be used to render the first class of tests suitable for panels characterized by an arbitrary pattern of contemporaneous cross-sectional dependence.

3.1 Levin and Lin (1992)

The first widely used panel unit root test, developed by Levin and Lin (1992), is a direct extension of a univariate ADF test to the panel data setting. The test (LL1 hereafter) restricts the speed of convergence to long-run equilibrium under the alternative of stationarity to be the same for all countries. As in the standard ADF framework, serial correlation among real exchange rate innovations is controlled for by adding one or more lags of the change in the real exchange rate as explanatory variables to the regression; autoregressive parameters at each lag and the lag length itself are restricted to be the same for all countries. Formally, the LL1 specification is given by the standard fixed effects (country-dummy) representation:

$$\Delta q_{i,t} = \alpha_i + \rho q_{i,t-1} + \sum_{k=1}^{m} \lambda_k \Delta q_{i,t-k} + \eta_{i,t}; \quad i = 1, \ldots, N; \quad t = 1, \ldots, T. \qquad (2)$$

Levin and Lin (1992) show that under the null hypothesis of a unit root the t-statistic for $\hat{\rho}$, $t_{\hat{\rho}}$, diverges to $-\infty$ at the rate $\sqrt{N}$.[6] A simple transformation of $t_{\hat{\rho}}$, however, converges

[6]Intuitively, the divergence occurs because the presence of country-specific fixed effects induces a downward bias, for finite T, on the least squares estimator of ρ; see, for example, Nickell (1981). Thus, the

to a standard normal variate as $\frac{\sqrt{N}}{T} \to 0$. For given N and T, appropriate critical values can be derived using Monte Carlo simulations, and critical values for various sample sizes are reported by the authors.

The LL1 test is implemented by estimating equation (2) jointly for all countries. The test can be extended to accommodate homogeneous cross-sectional dependence by adding fixed time effects to equation (2), or equivalently, by expressing all variables as deviations from their time-specific means. Such a transformation has no effect on the limiting distribution of $t_{\hat{\rho}}$, although in finite samples, there is some loss in statistical power for given N and T.

3.2 Heterogeneous cross-sectional dependence: A critique

In a recent paper, O'Connell (1998) shows that the LL1 test suffers from significant upward size distortion in the presence of heterogeneous correlation among contemporaneous real exchange rate innovations. The inclusion of fixed time effects is, at best, only a partial solution to the problem of heterogeneous cross-sectional dependence. Essentially, if the true cross-sectional covariance matrix of the real exchange rate innovations exhibits substantial heterogeneity in its off-diagonal elements, the removal of the time-mean from each individual series will do little to reduce the amount of cross-sectional dependence present in the data.[7] The failure to control for this feature of the data renders problematic any inference based on the LL1 test.

To address this problem, O'Connell (1998) proposes a GLS modification of the LL1 test (OC–GLS hereafter). The use of GLS transformation produces an estimator with critical values invariant to the actual pattern of contemporaneous cross-sectional correlation among real exchange rate innovations. Like the LL1 test, the OC–GLS test imposes the rate of convergence to long run equilibrium under the alternative to be the same for all countries—that is, $\rho_i = \rho$, for all i—and imposes a common serial correlation pattern on the real exchange rate innovations, determined by the autoregressive parameters λ_k, $k = 1, \ldots, m$.

In practice, the OC–GLS test relies on the first differences of raw data to estimate the contemporaneous cross-sectional covariance matrix of real exchange rate innovations.

estimation of fixed individual effects shifts the asymptotic mean and variance of the regression estimator of ρ. In the case of panel data, averaging across N preserves the shift in the mean, so that $\hat{\rho}$ converges to a non-central normal distribution. Because an increase in the cross-sectional dimension N reduces the sample variance—but for fixed T has no effect on the downward bias of $\hat{\rho}$—the t-statistic, $t_{\hat{\rho}}$, approaches $-\infty$ at the rate $\sqrt{N}$.

[7] As noted by Hakkio (1984) in the context of purchasing power parity, the real exchange rate data strongly favor the presence of heterogeneous cross-sectional correlation among contemporaneous real exchange rate innovations. Note that by construction, real exchange rates for countries i and j, $q_{i,t}$ and $q_{j,t}$, contain two common components: independent variation in the value of the dollar and independent variation in the U.S. price index. Other likely sources of correlation between $q_{i,t}$ and $q_{j,t}$ are regional shocks, if countries i and j are in the same region (e.g., Benelux, EMS, etc.). For example, a shock which originates in country i may propagate to country j but not to country k.

In particular, the covariance of exchange rate innovations between countries i and j is estimated as:

$$\hat{\omega}_{ij} = \left(\frac{1}{T-1}\right) \sum_{t=2}^{T} \left[(\Delta q_{i,t} - \overline{\Delta q_i})(\Delta q_{j,t} - \overline{\Delta q_j}) \right], \tag{3}$$

where $\overline{\Delta q_i} = \frac{1}{T-1} \sum_{t=2}^{T} \Delta q_{i,t}$. Under the null hypothesis of a unit root, the first difference of the real exchange rate and its innovation are equivalent (i.e., $\Delta q_{i,t} = \eta_{i,t}$), and the procedure yields consistent estimates of the innovation covariances. The estimate of the $N \times N$ contemporaneous cross-sectional covariance matrix of real exchange rate innovations, $\widehat{\Omega}$, is then defined as $\widehat{\Omega} \equiv [\hat{\omega}_{ij}]$, $i, j = 1, \ldots, N$.

To implement the OC–GLS test, the data matrices corresponding to equation (1) are transformed, using the estimated covariance matrix $\widehat{\Omega}$, rendering the error term $\eta_{i,t}$ cross-sectionally homoscedastic. Formally, equation (1) is pre-multiplied by the $NT \times NT$ GLS-transformation matrix $\Gamma \equiv P \otimes I_T$, where P is the lower triangular matrix from the Cholesky decomposition of $\widehat{\Omega}^{-1}$, I_T is an identity matrix of dimension T, and $\otimes$ denotes the Kronecker product. The transformed specification is then estimated by least squares. Critical values for the OC–GLS estimator are derived via Monte Carlo simulations: artificial data consistent with the null hypothesis are generated, and the lower 1-, 5- and 10-percent tails of the t-statistic for the hypothesis that $\rho = 0$, denoted by $t_{\hat{\rho}}$, are recorded.

To accommodate serial correlation among real exchange rate innovations, the OC–GLS test relies on a four-step procedure. In the first "pre-whitening" stage, differences of the real exchange rate are regressed on one or more lags of itself:

$$\Delta q_{i,t} = \sum_{k=1}^{m} \lambda_k \Delta q_{i,t-k} + u_{i,t}. \tag{4}$$

To ensure consistency with the assumptions of the LL1 test and achieve a parsimonious specification, equation (4) is estimated jointly for all countries, imposing a common maximum lag length m and a common serial correlation pattern, determined by the autoregressive parameters λ_k, $k = 1, \ldots, m$.

Under the null hypothesis of a unit root, $\Delta q_{i,t} - \sum_{k=1}^{m} \lambda_k \Delta q_{i,t-k} = \eta_{i,t}$, and the autoregressive parameters λ_k, $k = 1, \ldots, m$, are estimated consistently. In the second stage, the estimated residuals from the first-stage regression, $\hat{u}_{i,t}$ and $\hat{u}_{j,t}$, are used in place of $\Delta q_{i,t}$ and $\Delta q_{j,t}$ in equation (3) to obtain a consistent estimate of the contemporaneous cross-sectional covariance matrix Ω. In the third stage, the estimated autoregressive parameters $\hat{\lambda}_k$, $k = 1, \ldots, m$, are used to quasi-difference equation (1) to eliminate serial correlation.[8]

[8] For example, if first-order serial correlation is assumed, the quasi-differenced model is given by $(\Delta q_{i,t} - \hat{\lambda}_1 \Delta q_{i,t-1}) = \alpha_i(1 - \hat{\lambda}_1) + \rho(q_{i,t-1} - \hat{\lambda}_1 q_{i,t-2}) + (\eta_{i,t} - \hat{\lambda}_1 \eta_{i,t-1})$. Under the null, the quasi-differenced error term, $(\eta_{i,t} - \hat{\lambda}_1 \eta_{i,t-1})$, is serially uncorrelated.

Finally, a GLS transformation is performed on the quasi-differenced data, ρ is estimated by least squares, and $t_{\hat{\rho}}$ is compared with the appropriate critical values.

Although consistent with classical hypothesis testing, a potential problem with the OC–GLS test is that it yields inconsistent estimates of the autoregressive parameters λ_k, $k = 1, \ldots, m$, under the alternative.[9] The inconsistency of the estimated autoregressive parameters under the alternative implies that the contemporaneous cross-sectional covariance matrix of innovations Ω is also inconsistent. However, it can be shown through direct calculation or Monte Carlo simulations that $E(\widehat{\Omega} - \Omega)$ remains negligible for plausible values of ρ. More important, however, Monte Carlo experiments, reported below, show that the inconsistency of the estimated autoregressive parameters can lead to a sizeable loss in statistical power.

3.3 The SUR approach to cross-sectional dependence

When testing for the presence of unit roots in a panel data setting, the time-series dimension of the panel T typically exceeds the cross-sectional dimension N. This feature of the data can be exploited in a SUR framework to accommodate an arbitrary pattern of contemporaneous cross-sectional correlation. In this section, we describe a SUR-based feasible GLS unit root test (SUR–GLS hereafter) as an alternative to the OC–GLS test, which controls for heterogeneous cross-sectional dependence, is simple to compute, and avoids the aforementioned problems with the OC–GLS test.

In spirit, our SUR–GLS test is similar to the OC–GLS test. The fact that $T > N$ allows us to replace the fixed-effects specification of the OC–GLS test given in equation (2) with a SUR-system of N ADF regressions with country-specific intercepts.[10] The key difference between the OC–GLS test and the SUR–GLS test is that all parameters of the N equation system are estimated jointly, including, of course, the parameters of the contemporaneous

[9] This point is easiest to illustrate in the case of first-order serial correlation ($m = 1$). Note that, under the alternative, a regression of $\Delta q_{i,t}$ on its own lag involves the following term in the numerator of the least squares estimator of λ_1:

$$\left(\frac{1}{T-1}\right)\sum_{t=2}^{T}[(\alpha_i + \rho q_{i,t-1} + \eta_{i,t})(\alpha_i + \rho q_{i,t-2} + \eta_{i,t-1})]$$

Note that $q_{i,t-1}$ directly contains $\eta_{i,t-1}$, while $q_{i,t-1}$ and $q_{i,t-2}$ share the innovation terms $\eta_{i,t-2}, \eta_{i,t-3}, \ldots$. In a case when there is no serial correlation ($\lambda_1 = 0$), it can be shown that the estimated autoregressive parameter $\hat{\lambda}_1$ converges to $\rho/2$ (instead of zero) as $T \to \infty$.

This inconsistency also applies to all the parameters estimated for higher-order processes. Suppose, for example, that an AR(2) model is estimated when in fact there is no serial correlation among real exchange rate innovations. It can be shown that the AR(1) parameter $\hat{\lambda}_1$ converges in probability to $\frac{1-\rho}{2-\rho}\rho$, while the AR(2) parameter $\hat{\lambda}_2$ converges to $\frac{\rho}{2-\rho}$. Note that, with $\rho < 0$, the asymptotic bias of $\hat{\lambda}_1$ increases as additional lags of $\Delta q_{i,t}$ are added to the ADF regression.

[10] Like the OC–GLS test, our test imposes a common decay parameter ρ on all countries, and restricts the temporal process of real exchange rate innovations to be the same across countries.

cross-sectional covariance matrix Ω and the autoregressive parameters λ_k, $k = 1, \ldots, m$.

In particular, we rely on an iterative GLS procedure, using the residuals at each iteration to generate $\widehat{\Omega}$ for the subsequent round of estimation.[11] In contrast to the OC–GLS test, estimates of the autoregressive parameters λ_k, $k = 1, \ldots, m$, are consistent under both the null and alternative hypotheses. Under our approach, the estimates are in fact biased in finite samples under both the null and alternative hypotheses. Monte Carlo experiments, however, indicate that the degree of finite-sample bias is small. More importantly, Monte Carlo experiments indicate that extending the test to accommodate serial correlation results in a relatively small loss in statistical power, relative to the OC–GLS test.

The SUR test described above differs from the test recently introduced by Sarno and Taylor (1998a) only in minor respects. Because all the slope and intercept parameters—as well as the $N(N+1)/2$ elements of the covariance matrix Ω—for the system of N equations are estimated simultaneously, the computational cost of the iterative GLS procedure is high. Sarno and Taylor (1998a) achieve computational feasibility by limiting their sample to four countries (i.e., $N = 4$). In contrast, we achieve computational feasibility by imposing parameter homogeneity in the rate of convergence to long-run equilibrium and in the parameters determining the pattern of serial correlation among real exchange rate innovations. Consequently, we are able to consider panels with considerably larger cross-sectional dimensions.

Second, as noted earlier, we rely on an iterative GLS procedure, whereas Sarno and Taylor (1998a) rely on a two-step GLS procedure. The choice of an iterative rather than two-step procedure affects the finite-sample but not the asymptotic properties of the test.

3.4 Levin and Lin (1993)

The strong parameter homogeneity assumptions of the LL1 test have led those authors to develop a panel unit root test that places fewer restrictions on the DGP than their original test. The second test developed by Levin and Lin (1993) (LL2 hereafter), allows the rate of convergence to long-run equilibrium under the alternative to vary across countries. In addition, the autoregressive parameters at all lags—as well as the lag-length m itself—are allowed to vary across individual cross-sectional units. The empirical model, then, corresponds to an unrestricted ADF specification:

$$\Delta q_{i,t} = \alpha_i + \rho_i q_{i,t-1} + \sum_{k=1}^{m_i} \lambda_{i,k} \Delta q_{i,t-k} + \eta_{i,t}; \quad i = 1, \ldots, N; \quad t = 1, \ldots, T \tag{5}$$

[11] The iteration terminates when the covariance matrix of equation errors changes less than the specified convergence criterion (.001 in our case); see Davidson and MacKinnon (1993) for discussion.

As with the LL1 test, a homogeneous pattern of contemporaneous cross-sectional dependence can be accommodated by including fixed time effects to equation (5), or by expressing all variables as deviations from their time-specific means.

The computational procedure of the LL2 test is elaborate, requiring the calculation of various statistics for each country. The ultimate test statistic has a limiting $N(0,1)$ distribution as $N \to \infty$, $T \to \infty$, and $\frac{N}{T} \to 0$. In finite samples, however, Monte Carlo simulations are required to estimate mean and variance adjustment factors, which preserve a standard normal distribution under the unit root null. The test procedure is described in detail in the Appendix A.1.

3.5 Im, Pesaran, and Shin (1995)

Im, Pesaran and Shin (1995) propose a unit root test for heterogeneous dynamic panels based on the mean-group approach recently advanced in Pesaran and Smith (1995) and Pesaran, Smith, and Im (1996). The test (IPS hereafter) is equivalent to the LL2 test, in the sense that it is valid in the presence of heterogeneity across cross-sectional units, as well as of residual serial correlation across time periods. As with the LL2 test, homogeneous cross-sectional dependence can be accommodated by including fixed time effects in the ADF specification.

The IPS test statistic is a simple function of the average t-statistic for $\hat{\rho}_i$ from the N individual ADF regressions. The authors show that this simple function converges to a standard normal variate as $N \to \infty$, $T \to \infty$, and $\frac{N}{T} \to 0$. As with the LL2 test, Monte Carlo simulations are required to estimate mean and variance adjustment factors, which preserve a $N(0,1)$ distribution in finite samples. The test procedure is described in detail in the Appendix A.2.

3.6 Cross-sectional dependence and the LL2 and IPS tests

As noted above, the LL2 and IPS tests assume that contemporaneous innovations across different cross-sectional units are uncorrelated—that is, $E(\eta_{i,t}\eta_{j,t}) = 0$, for all $i \neq j$. Homogeneous cross-sectional dependence can be accommodated by expressing all variables as deviations from their time-specific means. Both tests, however, remain invalid in the presence of heterogeneous cross-sectional correlation. Monte Carlo experiments (not reported here) indicate that, like the LL1 test, the LL2 and IPS tests suffer from substantial upward size distortion in the presence of heterogeneous cross-sectional dependence.

This section describes a simple GLS procedure which renders LL2 and IPS tests valid when contemporaneous innovations display heterogeneous cross-sectional correlation. Following O'Connell (1998), we first "pre-whiten" raw real exchange rate differences to eliminate serial correlation, as in equation (4). To maintain consistency with the LL2 and IPS tests, the autoregressive parameters and the lag length itself are allowed to vary across countries. The residuals from the individual-specific pre-whitening regressions are used to compute an estimate of the cross-sectional covariance matrix Ω. We then pre-multiply the data matrices corresponding to N equations by the $NT \times NT$ GLS-transformation matrix $\Gamma \equiv P \otimes I_T$, where P is the lower triangular matrix from the Cholesky decomposition of $\widehat{\Omega}^{-1}$.

The result of this transformation is that contemporaneous real exchange rate innovations are now cross-sectionally uncorrelated. The assumptions underlying the LL2 and IPS tests are now satisfied, and the tests can be applied to the transformed data in the usual manner; we refer to the modified tests as LL2–GLS and IPS–GLS, respectively.

Note that the pre-whitening regression is used only to derive serially uncorrelated residuals, which are used to estimate the contemporaneous cross-sectional covariance matrix Ω. The estimated country-specific autoregressive parameters $\lambda_{i,k}$, $i = 1, \dots, N$, $k = 1, \dots, m_i$, are discarded, and the possibility of serial correlation among real exchange rate innovations is accommodated by adding m_i lags of $\Delta q_{i,t}$ to the country-specific regression equation.

4 Monte Carlo Design

Implementing the OC–GLS and SUR–GLS tests requires calculation of appropriate critical values using Monte Carlo simulations. The process is straightforward. First, we generate 10,000 artificial real exchange rate panels that match the time-series and cross-sectional dimensions of the relevant samples. The artificial panels are consistent with the null hypothesis of a unit root, so that $q_{i,t} = \sum_{s=1}^{t} \epsilon_{i,s}$, where $\epsilon_{i,t}$ is a Gaussian random number with unit variance.[12] For each of the two tests, we calculate the t-statistic for the hypothesis that $\rho = 0$, denoted as $t_{\hat{\rho}}$ and record the 1-, 5- and 10-percent critical values.

For the OC–GLS test, we tailor critical values to the number of autoregressive terms, m, in the first-stage pre-whitening regression. Similarly for the SUR–GLS test, we calculate critical values corresponding to the number of lags of $\Delta q_{i,t}$ in the SUR system of N ADF regressions. For both tests, however, empirical critical values are essentially invariant to the number of lags of $\Delta q_{i,t}$ included in the specification (see tables A.5–A.8 in the Appendix Tables).

Implementing the LL2–GLS and IPS–GLS tests requires calculation of appropriate mean and variance adjustment factors. As above, we generate 10,000 artificial real exchange rate panels, consistent with the unit root null, which match the time-series and cross-sectional dimensions of the relevant samples. We then calculate the mean and variance of the relevant

[12] In all Monte Carlo simulations, we drop the first 25 observations for each "country" to minimize initial-value bias. After these initial observations are discarded, the time series dimension of the artificial real exchange rate panel matches that of the panel of interest.

test statistics absent adjustment (see Appendices A.1 and A.2) and derive the adjustment factors needed to transform the test statistics to a standard normal distribution in finite samples. For both tests, we allow the adjustment factors to vary with the number of autoregressive terms, m_i, included in the country-specific ADF regression.[13]

In generating artificial data, we assume that there is in fact no cross-sectional dependence among contemporaneous real exchange rate innovations, and that the innovations are serially uncorrelated. These assumptions are innocuous. Monte Carlo experiments (not reported) show that, within the range of experimental error, estimated critical values and adjustment factors do not depend on the actual contemporaneous cross-sectional covariance matrix Ω or on the actual serial correlation parameters $\lambda_{i,k}$, $i = 1, \ldots, N$, $k = 1, \ldots, m_i$.[14]

5 Empirical Results

Our focus is on the behavior of real exchange rates during the post-Bretton Woods era. We apply the four panel unit root tests described above to three different samples. The three samples consist of three, partly overlapping groups of countries: *i*) 32 countries classified as economically "open" by Sachs and Warner (1995); *ii*) the 25 countries belonging to the OECD as of the end of 1995; and *iii*) 19 European countries. We use quarterly, seasonally unadjusted data and define real exchange rates using consumer price indices (CPIs). As a check on the robustness of our results, we also study the behavior of real exchange rates defined using wholesale price indices (WPIs), with minor changes in the sample composition and sample range. Details concerning country membership and data availability are reported in the tables A.1 and A.3 in the Appendix Tables.

5.1 Results using consumer price indices

As a benchmark, we first estimate univariate ADF tests for the 36 countries included in our various samples. In line with most previous research, the U.S. is treated as the numeraire country. We rely on the Akaike information criterion (AIC) to choose the lag length m_i for the country-specific ADF regressions—in every case, we assume a minimum lag length of

[13]In their original test, Im, Pesaran, and Shin (1995) tailor mean and variance adjustment factors to the number of lags of $\Delta q_{i,t}$ in the ADF regressions. Levin and Lin (1993), on the other hand, do not allow adjustment factors to vary with the number of lags. Our Monte Carlo results, however, indicate that doing so is necessary in order to leave the LL2 or LL2–GLS test statistics with the desired $N(0,1)$ distribution in finite samples. The adjustment factors for the LL2–GLS and IPS–GLS tests for all samples and various ADF specifications are reported in tables A.9–A.12 in the Appendix Tables.

[14]This result should not be surprising. The covariance matrix Ω is estimated in order to transform the data, rendering the error term cross-sectionally homoscedastic. The estimation error separating Ω and $\widehat{\Omega}$ does not depend on the actual degree or pattern of cross-sectional dependence. Similarly, the autoregressive parameters are estimated in order to render the equation errors serially uncorrelated. The estimation error separating the $\lambda_{i,k}$ and $\hat{\lambda}_{i,k}$ does not depend on the actual parameter values.

four. According to the AIC, a four-lag ADF specification appears adequate to capture any serial dependence among real exchange rate innovations for 34 out of 36 countries; for only two countries (Barbados and Mauritius) the AIC is minimized by a five-lag specification.[15]

The univariate ADF tests provide essentially no support for real exchange rate stationarity (Appendix Table A.2). We are able to reject the unit root null at the 10-percent significance level for only five of the 36 countries—little better than would be expected to happen by chance. We are unable to reject the unit root null at the five-percent level for even a single country.

These results come as no surprise. Numerous studies fail to reject the unit root null for the recent float period using univariate ADF tests. Does this lack of support for (relative) purchasing power parity stem from the well-known low statistical power of univariate unit root tests, or does it reveal that real exchange rates during the post-Bretton Woods era are, in fact, non-stationary?

To answer this question, we apply the panel unit root tests described above to the Open, OECD, and Europe samples. Our baseline results are based on a four-lag ADF specification (i.e., $m_i = m = 4$, for all i), corresponding to the minimum AIC indicated by the data. The results indicate that low statistical power lies behind the lack of support from univariate ADF tests for real exchange rate stationarity (Table 1).

The evidence is strongest using the IPS–GLS and SUR–GLS tests. Both tests reject the unit root null at the one-percent significance level for *all* three samples. Support for real exchange rate stationarity also comes from the OC–GLS test: the unit root null is rejected at the one-percent level for the Europe and OECD samples, although it is not rejected at conventional significance levels for the larger Open sample.[16] The evidence from the LL2–GLS test is also consistent with the relative PPP hypothesis: the unit-root null is rejected at the five-percent level for the Europe and OECD samples and at the 10-percent level for the Open sample.

Taken together, the four tests provide overwhelming support of real exchange rate stationarity for the Europe and OECD samples and strong support for the Open sample. We think of the results using the IPS–GLS and SUR–GLS tests as two stakes through the heart of the unit root null. The somewhat less emphatic results from the OC–GLS and LL2–GLS tests together comprise a third stake.

[15] For each sample of interest, we also estimate the fixed effects (country-dummy) specification of the ADF regression, so that $\rho_i = \rho$ and $\lambda_{i,k} = \lambda_k$, for all i. We allow the covariance matrix of the regression errors to exhibit unrestricted contemporaneous cross-sectional dependence and estimate the resulting specification with restricted maximum likelihood; see, for instance, Diggle, Liang, and Zeger (1995). Using the AIC, the results indicate that for all three samples four lags are sufficient to capture the serial correlation pattern of real exchange rate innovations.

[16] O'Connell (1998) is unable to reject the unit root null for a similar Europe panel. The divergence in results may reflect differences in panel membership; the availability of two additional years of data; or the fact that O'Connell relies on a nine-lag pre-whitening regression (see also Section 5.3 below).

To ensure that the results are not driven by failure to accommodate higher-order serial correlation, we derive a second set of results with $m_i = m = 6$, for all i (bottom of Table 1). The results are largely unchanged, except that the LL2–GLS test now fails to reject the null for every sample.

5.2 Results using wholesale price indices

As a check on the robustness of our results, we perform similar analysis on real exchange rates defined with wholesale rather than consumer prices indices. The new tests also allow us to ask whether the higher weight of tradable goods in wholesale price indices brings faster reversion to purchasing power parity (Section 5.4 below).

As before, we begin by estimating univariate ADF tests. The country-specific lag lengths accommodating serial correlation are again chosen using the AIC; as before, we assume a minimum lag length of four (Appendix Table A.4). For 27 of the 28 countries with complete wholesale price data, including four lags of $\Delta q_{i,t}$ in the ADF regression appears sufficient to capture all serial correlation among real exchange rate innovations; for one country (Taiwan), a six-lag ADF specification appears necessary.

Again, the univariate ADF tests provide essentially no support for the relative PPP hypothesis. We are able to reject the unit root null at the 10-percent level for only two of the 28 countries (Ireland and Mexico), less often than would be expected to happen by chance.

The picture changes dramatically when we turn to panel unit root tests. According to the top of Table 2, *all* four tests reject the unit root null at the one-percent level, for *every* sample in the case of a four-lag ADF specification. As before, to ensure that the results are not driven by failure to accommodate higher-order serial correlation, we derive a second set of results with $m_i = m = 6$, for all i (bottom of Table 2). As in the case of CPI-defined real exchange rates, the main difference is that the LL2–GLS test does not reject the unit root null at conventional significance levels for any of the three samples. The three remaining tests, however, reject the null at five-percent (or better) significance levels for all three samples.

5.3 Power analysis

The empirical results presented above build a strong case for real exchange rate stationarity during the recent floating-rate period. The results, however, differ significantly across the four tests. For instance, using CPI-defined real exchange rates, the baseline IPS–GLS test rejects the unit root null at the one-percent level for all three panels; the LL2–GLS test, on the other hand, rejects the null only at the five- or ten-percent level. In addition, using a six-lag ADF specification, the LL2–GLS test fails to reject the unit root null in every case.

In this section, we examine whether these differences reflect differences in the statistical power of the four tests.

To address this question, we conduct additional Monte Carlo experiments. The design of the experiments is the same as above, except that the artificial data are generated under the alternative hypothesis of stationarity. In particular, we generate 5,000 artificial real exchange rate panels that match the time-series and cross-sectional dimensions of our three actual samples.

We assume that real exchange rate innovations decay at the rate $\rho = -.04$—that is four percent per quarter—so that the level of the real exchange rate is given by $q_{i,t} = \sum_{s=0}^{t-1}(1+\rho)^s \epsilon_{i,t-s}$, where $\epsilon_{i,t}$ is a Gaussian random number with unit variance.[17] We then apply the OC–GLS, SUR–GLS, LL2–GLS, and IPS–GLS tests in the usual manner, relying on the critical values and adjustment factors calculated under the null hypothesis in Section 4.

It is important to examine the effect of the number of lags m in the ADF specification on test power, or in the case of the OC–GLS estimator, the number of lags included in the pre-whitening regression. Hence, we conduct Monte Carlo experiments for zero-, four-, and eight-lag ADF specifications.

The effect of lag length m on test power is of interest for two reasons. First, as noted earlier, the pre-whitening regression in the OC–GLS test yields inconsistent estimates of the autoregressive parameters under the alternative. The test, therefore, may suffer from a relatively large loss in power as the number of lags included in the pre-whitening regression increases.

Second, as noted earlier, two of the tests (IPS–GLS and LL2–GLS) allow the autoregressive parameters to vary across panel members, while the OC–GLS and SUR–GLS tests impose parameter homogeneity. The cost of this greater generality may be a greater loss in test power as the number of autoregressive terms in the ADF regression increases. In particular, adding an additional lag of $\Delta q_{i,t}$ to the ADF specification involves estimating only a single additional parameter for the SUR–GLS and OC–GLS tests, but N additional parameters for the LL2–GLS and IPS–GLS tests.

Our Monte Carlo experiments indicate that the OC–GLS, SUR–GLS, and IPS–GLS tests have very high power against the unit root null. Consider the Europe panel (Table 3). In the zero-lag case, for example, the OC–GLS test has power of 59 percent for a one-percent test size and 94 percent for a ten-percent test size. The power of the SUR–GLS is slightly lower, with the IPS–GLS test close behind. The power of the LL2–GLS test, however, is considerably lower, at only six percent for a one-percent test size and 45 percent for a 10 percent test.[18]

[17] As is customary in power calculations, we set the nuisance parameters, $\alpha_i = 0$, for all i.

[18] For reasons of computational feasibility, we have not attempted to assess the power performance of

The power of the OC–GLS test, on the other hand, falls rapidly as the lag-length m increases. Consider moving from a zero-lag to an eight-lag ADF specification while remaining within the Europe panel. At the five-percent significance level, the power of the OC–GLS test falls by nearly a half, from 86 to 45 percent. The power of the SUR–GLS test falls much less, from 80 to 62 percent. The same pattern holds across all panels. Evidently, the power of the SUR–GLS test suffers only from the number of additional parameters estimated. The power of the OC–GLS test, on the other hand, suffers from the number of additional parameters estimated, and also from the fact that the additional parameters are estimated inconsistently.

The IPS–GLS test suffers a relatively modest reduction in test power—from 77 to 53 percent at the five-percent test size—when moving from a zero- to an eight-lag ADF specification. The power of the LL2–GLS test, on the other hand, essentially disappears, falling from 28 to zero percent, when moving from a zero- to an eight-lag ADF specification. The high power attained by the IPS–GLS test—under both the zero- and eight-lag specifications—is notable, given that it allows for heterogeneity across countries in the rate of convergence to long-run equilibrium and in the parameters describing serial correlation of the innovation process.

The power loss arising from the estimation of additional autoregressive parameters becomes more pressing if the method for choosing the lag length m tends to select a high-order ADF specification. For example, O'Connell's (1998) procedure is to regress raw real exchange rate changes on lagged changes, beginning with a maximum lag length of 12. If the deepest lag is significant, according to a likelihood ratio test, that lag length is selected. If not, the number of lags is reduced by one, and the likelihood ratio test is re-applied.

Because auto-regressions involving raw real exchange rate changes yield inconsistent parameter estimates under the alternative of stationarity, this procedure may point to "significant" temporal dependence where none exists. For example, with $T = 95$ and $N = 19$, as in the Europe sample, and $\rho = -.04$, a specification containing eight or more lags will be chosen 68 percent of the time, in the absence of any true temporal dependence and using a ten-percent test to assess statistical significance.[19] The OC–GLS procedure not only loses power quickly as the number of estimated autoregressive parameters increases, but furthermore, tends to choose a high-order ADF specification.

The power of the various tests also depends on N, the cross-sectional dimension of the panel. Note that the power of the OC–GLS test, relative to SUR–GLS, improves as

an unrestricted SUR estimator for panels of the relatively large cross-sectional dimensions considered here. Initial experiments with $N = 4$, however, suggest that relaxing our parameter restrictions results in only a moderate loss in test power.

[19]This conclusion is based on a Monte Carlo experiment involving 5,000 replications. If statistical significance is assessed using a five-percent test, a specification containing eight or more lags will be chosen 39 percent of the time.

the cross-sectional dimension of the panel increases. Indeed, for a panel corresponding to our Open sample (with CPI-defined real exchange rates), the power of the OC–GLS test exceeds the power of SUR–GLS, even under an eight-lag ADF specification. This result is not surprising. As the cross-sectional dimension of the panel rises, the number of elements in the covariance matrix Ω to be estimated rises with it. This places a premium on the small-sample efficiency of the OC–GLS estimator, outweighing the large-sample inconsistency of the estimated autoregressive parameters.

The lesson is that the OC–GLS grows increasingly attractive relative to SUR–GLS as $\frac{T}{N}$ declines. However, we are unable to provide a firm metric for choosing between the two tests. The power of the IPS–GLS test also rises relative to that of SUR–GLS as the cross-sectional dimension of the panel rises, but the increase is more modest.

One should also note that we report lower test power using the SUR–GLS test than do Sarno and Taylor (1998a) using a similar test.[20] The lower test power reported here reflects an important difference in the design of Monte Carlo experiments. Sarno and Taylor estimate the country-specific drift terms—α_i, $i = 1, \ldots, N$, in equation (2)—using univariate regressions. The estimated drift parameters are then included in the DGP used to generate artificial data under the null hypothesis.[21]

The estimated critical values are valid only if the estimates of the drift parameters α_i, $i = 1, \ldots, N$, match those present in the true real exchange rate DGP. Monte Carlo experiments (results available on request) show that that the initial estimates of the country-specific drift terms are biased both under the null hypothesis and local alternatives—in particular, the α_i's are typically overestimated in absolute value. Moreover, the estimates are subject to very wide dispersion under both the null and local alternatives.

The experiments also show that the distribution of test statistics under the null hypothesis is very sensitive to the values of the drift terms included in the DGP used to generate artificial data. In particular, small increases (in absolute value) in the α_i's bring sizeable reductions in empirical critical values. As a result, including estimated drift terms in the process used to generate artificial data will generally leave nominal test size well above true test size, boosting apparent test power.[22] Because the true α_i's are unknown and cannot be estimated accurately, we follow the more conservative procedure of setting such nuisance parameters to zero when generating the artificial data.[23]

[20] All else equal, we should find higher test power than Sarno and Taylor, because our SUR-based test imposes two (assumed) true restrictions: *i*) the rate of convergence is the same for all countries; and *ii*) the autoregressive parameters are the same for all countries.

[21] As noted earlier, some interpretations of the null hypothesis require that $\alpha_i = 0$, for all i; see Banerjee et al. (1996).

[22] It is worth emphasizing that the authors' empirical results imply rejection of the unit root null for the G-5 countries even using the critical values that would be calculated by setting $\alpha_i = 0$, for all i.

[23] The points made here are not specific to SUR-based panel unit root tests. For example, adding a drift term of -.01 to a univariate unit root process with an innovation standard deviation of .05 changes the

5.4 The rate of convergence to long-run equilibrium

Although the results presented above provide strong evidence of real exchange rate stationarity during the post-Bretton Woods era, they do not provide direct evidence as to the rate of convergence to long-run equilibrium. The actual point estimates of ρ from the various tests, in fact, imply fairly rapid mean reversion. For example, applying the SUR–GLS test to CPI-defined real exchange rates yields the estimated rate of convergence of more than six percent per quarter, or 25 percent per annum, for the Europe and OECD samples and almost five percent per quarter for the Open sample (Table 5).

These point estimates, however, are biased downward under both the null and alternative hypotheses, because of the presence of fixed country effects.[24] Indeed, under the null hypothesis that $\rho = 0$, the expected value of $\hat{\rho}$ comes to about -.04 for each of the three samples. The degree of bias is somewhat smaller under the alternative, because shocks to the real exchange rate erode at the rate ρ, reducing the covariance between $\overline{\epsilon}_i$ and $\overline{q}_{i,-1}$.

We rely on Monte Carlo simulations to calculate the bias-corrected rate of convergence to long-run equilibrium implied by our empirical results. In particular, we generate 5,000 artificial real exchange rate panels, matching cross-sectional and time-series dimensions of our three samples. The artificial real exchange rate series are generated under the alternative of stationarity, with the decay parameter ρ taking on the following values: -.050, -.045, -.040, -.035, -.030, -.025, -.020, -.015, -.010, -.005, and -.001. For each Monte Carlo experiment—corresponding to the assumed values of ρ in the DGP used to generate the artificial data—we average the *estimated* rate of convergence over 5,000 replications. If the estimate of ρ from the actual data lies between two assumed values, we use a linear interpolation to derive a bias-corrected estimate of ρ.

Our bias-corrected estimates imply that real-exchange rates return to their long-run equilibrium values rather slowly. Consider the results for the SUR–GLS test using CPI-defined real exchange rates. For the Europe sample, the bias-corrected estimated rate of convergence comes to 3.2 percent quarter, or 12.2 percent per annum, implying a half-life for real exchange rate equilibrium deviations of just over five years. The estimated rate of convergence is even slightly lower for the OECD sample. For the larger Open sample, convergence apparently occurs at a snail's pace, only 0.2 percent per quarter, implying a half-life for equilibrium deviations of more than 85 years! One could well interpret this result as supporting the absence of mean reversion in any economically relevant sense.

Our estimates point to substantially faster convergence to long-run equilibrium for WPI-defined real exchange rates, which assign a greater weight to tradable goods. For the Europe

estimated five-percent critical value for the ADF test from -2.89 to -2.52. With $\rho = -.04$ and $T = 100$, the power of the univariate test would appear to rise from 9.8 to 20.8 percent.

[24]The transformed right-hand-side variable $(q_{i,t} - \overline{q}_{i,-1})$, and the transformed error term, $(\epsilon_{i,t} - \overline{\epsilon}_i)$, are correlated through their mean components. The bias, of course, goes to zero at the rate $1/T$ as $T \to \infty$.

sample, the bias-corrected speed of convergence is 5.7 percent for quarter, or 21 percent per year, implying a half-life for equilibrium deviations of just under three years. The estimated rate of convergence for the OECD sample is about the same. The sharpest contrast is found for the larger Open sample: the estimated speed of convergence is 7.0 percent per quarter—against only 0.2 percent in the case of CPI-defined real exchange rates—implying a half life of less than 2.5 years.

The much faster rate of convergence for the WPI-defined real exchange rates in the Open sample is not due to the fact that fewer countries report wholesale price data (26 vs. 32 for consumer prices). If we remove countries that do not report wholesale price data from the Open sample, the bias-corrected rate of convergence for CPI-defined real exchange rates remains very low, at about 0.4 percent per quarter.

We also generate biased-corrected estimates of the rate of convergence to PPP using the IPS–GLS test (Table 6). This test implies somewhat faster mean reversion for CPI-defined real exchange rates. The contrast is sharpest for the Open sample: the estimated rate of convergence is now a slow but economically meaningful 1.5 percent per quarter, implying a half life for equilibrium deviations of about 11.5 years.

The faster mean reversion for WPI-defined real exchange rate is also evident according to the IPS–GLS test, although the contrast is less stark than with the SUR–GLS test. The fact that the bias-corrected estimates of the decay parameter ρ differ significantly across the two tests and the three samples is, in our view, *prima facie* evidence against the assumption of a common rate of reversion across countries.

6 Conclusion

We have provided a systematic analysis of the PPP hypothesis, using four linear panel unit root tests. Two of the tests, IPS–GLS and LL2–GLS, are based on statistics derived from individual-specific ADF regressions and allow for complete parameter heterogeneity across countries. One of our contributions has been to modify the two tests to eliminate the upward size distortion induced by contemporaneous cross-sectional dependence among real exchange rate innovations. The remaining two tests, SUR–GLS and OC–GLS, are explicitly designed to accommodate cross-sectional dependence in the innovation process, but allow for only limited parameter heterogeneity across countries.

Our empirical results provide strong support for real exchange rate stationarity during the recent floating-rate period. Three of the four tests (SUR–GLS, OC–GLS, and IPS–GLS) emphatically reject the unit root null for our three panels of relatively open economies. Because our results using the LL2–GLS test are less than emphatic, we think of the paper as placing three stakes through the heart of the unit root null.

Our results are consistent with other recent work. Cheung and Lai (1998) apply two efficient univariate tests to post-Bretton Woods real exchange rates for the G-5 countries (the U.S., Japan, Germany, France, and the UK). The authors are able to reject the unit root null for most country pairs. Using an unrestricted version of our SUR–GLS test, Sarno and Taylor (1998b) find strong evidence of real exchange rate stationarity for the same sample. The Sarno-Taylor bottom line is the same as our own: controlling for cross-sectional dependence does not undermine panel data evidence in favor of real exchange rate stationarity among relatively open economies.

References

ABUAF, N., AND P. JORION (1990): "Purchasing Power Parity in the Long Run," *Journal of Finance*, 45, 157–174.

ANDREWS, D. W. K. (1991): "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

BANERJEE, A., J. DOLADO, J. GALBRAITH, AND D. HENDRY (1996): *Cointegration, Error Correction and the Econometric Analysis of Non-Stationary Data.* Oxford University Press, Oxford, UK.

CHEUNG, Y.-W., AND K. S. LAI (1998): "Parity Reversion in Real Exchange Rates During the Post-Bretton Woods Era," *Journal of International Money and Finance*, 17, 597–614.

DAVIDSON, R., AND J. G. MACKINNON (1993): *Estimation and Inference in Econometrics.* Oxford University Press, Oxford, UK.

DIGGLE, P. J., K.-Y. LIANG, AND S. L. ZEGER (1995): *Analysis of Longitudinal Data.* Clarendon Press, Oxford, UK.

FLOOD, R. P., AND M. P. TAYLOR (1996): "Exchange Rate Economics: What's Wrong with the Conventional Macro Approach?," in *The Microstructure of Foreign Exchange Markets*, ed. by J. A. Frankel, G. Galli, and A. Giovannini, pp. 261–294. The University of Chicago Press, Chicago.

FRANKEL, J. A. (1986): "International Capital Mobility and Crowding Out in the U.S. Economy: Imperfect Integration of Financial Markets or of Goods Markets," in *How Open is the U.S. Economy?*, ed. by R. Hafer, pp. 33–67. Lexington Books, Lexington, MA.

——— (1990): "Zen and the Art of Modern Macroeconomics: The Search for Perfect Nothingness," in *Monetary Policy for a Volatile Global Economy*, ed. by W. Haraf, and T. Willett, pp. 117–123. American Enterprise Institute, Washington D.C.

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

FROOT, K. A., AND K. ROGOFF (1995): "Perspectives on PPP and Long-Run Real Exchange Rates," in *Handbook of International Economics, Vol. III*, ed. by G. Grossman, and K. Rogoff, pp. 1647–1688. Elsevier Science Publishers, Amsterdam, The Netherlands.

HAKKIO, C. S. (1984): "A Re-examination of Purchasing Power Parity," *Journal of International Economics*, 17, 265–277.

IM, K. S., M. H. PESARAN, AND Y. SHIN (1995): "Testing for Unit Roots in Heterogeneous Panels," Unpublished paper, Dept. of Applied Economics, University of Cambridge.

JORION, P., AND R. J. SWEENEY (1996): "Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting," *Journal of International Money and Finance*, 15, 535–550.

KOEDIJK, K., P. SCHOTMAN, AND M. DIJK (1998): "The Re-Emergence of PPP in the 1990s," *Journal of International Money and Finance*, 17, 51–61.

LEVIN, A., AND C.-F. LIN (1992): "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties," Unpublished paper, Dept. of Economics, University of California San Diego.

——— (1993): "Unit Root Tests in Panel Data: New Results," Unpublished paper, Dept. of Economics, University of California San Diego.

MACDONALD, R. (1996): "Panel Unit Root Tests and Real Exchange Rates," *Economic Letters*, 50, 7–11.

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

MEESE, R., AND K. ROGOFF (1988): "Was It Real? The Exchange Rate Interest Differential Relation over the Modern Floating Exchange Rate Period," *Journal of Finance*, 43, 933–948.

NEWEY, W. K., AND K. D. WEST (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

NICKELL, S. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 1417–1426.

O'CONNELL, P. G. J. (1998): "The Overvaluation of Purchasing Power Parity," *Journal of International Economics*, 44, 1–20.

OH, K.-Y. (1996): "Purchasing Power Parity and Unit Root Tests Using Panel Data," *Journal of International Money and Finance*, 15, 405–418.

PAPPEL, D. H., AND H. THEODORIDIS (1998): "Increasing Evidence of Purchasing Power Parity Over the Current Float," *Journal of International Money and Finance*, 17, 41–50.

PARSLEY, D., AND S.-J. WEI (1996): "Convergence to the Law of One Price Without Trade Barriers or Currency Fluctuations," *Quarterly Journal of Economics*, 111, 1211–1236.

PESARAN, M. H., AND R. SMITH (1995): "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68, 79–113.

PESARAN, M. H., R. SMITH, AND K. S. IM (1996): "Dynamic Linear Models for Heterogeneous Panels," in *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, ed. by L. Matyas, and P. Sevestre, pp. 145–195. Kluwer Academic Publishers, Norwell, MA.

SACHS, J., AND A. WARNER (1995): "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity*, 1, 1–95.

SARNO, L., AND M. P. TAYLOR (1998a): "The Behavior of the Real Exchange Rates During the Post-Bretton Woods Period," *Journal of International Economics*, 46, 281–312.

——— (1998b): "Real Exchange Rates Under the Recent Float: Unequivocal Evidence of Mean Reversion," *Economic Letters*, 60, 131–137.

WU, Y. (1996): "Are Real Exchange Rates Nonstationary? Evidence from a Panel Data Test," *Journal of Money, Credit, and Banking*, 28, 54–63.

TABLE 1

Panel Data Unit Root Tests of PPP
CPI Based Real Exchange Rates
Empirical Size Properties

Four-Lag ADF Specification ($m = 4$)

Unit Root Test	SAMPLE EUROPE	OECD	OPEN
OC–GLS	0.01	0.01	nr
SUR–GLS	0.01	0.01	0.01
LL2–GLS	0.05	0.05	0.10
IPS–GLS	0.01	0.01	0.01
N	19	25	32
$\min T_i$	-	-	71
$\max T_i$	-	-	95
$\overline{T}$	95	93	93
Obs.	1,805	2,375	2,975

Six-Lag ADF Specification ($m = 6$)

Unit Root Test	SAMPLE EUROPE	OECD	OPEN
OC–GLS	0.01	0.01	nr
SUR–GLS	0.05	0.05	0.05
LL2–GLS	nr	nr	nr
IPS–GLS	0.01	0.01	0.01
N	19	25	32
$\min T_i$	-	-	69
$\max T_i$	-	-	93
$\overline{T}$	93	93	91
Obs.	1,767	2,325	2,911

Notes: The entries in the table correspond to the significance level at which the null hypothesis of a unit root can be rejected; nr indicates that the null hypothesis cannot be rejected at the 10 percent (or better) significance level. All empirical significance levels are based on 10,000 Monte Carlo replications of a panel with the cross-sectional dimension N and time series dimension $\overline{T}$.

TABLE 2

Panel Data Unit Root Tests of PPP
WPI Based Real Exchange Rates
Empirical Size Properties

Four-Lag ADF Specification ($m = 4$)

	SAMPLE		
Unit Root Test	EUROPE	OECD	OPEN
OC–GLS	0.01	0.01	0.01
SUR–GLS	0.01	0.01	0.01
LL2–GLS	0.01	0.01	0.01
IPS–GLS	0.01	0.01	0.01
N	16	22	26
$\min T_i$	67	67	49
$\max T_i$	95	95	95
$\overline{T}$	90.8	92	90
Obs.	1,453	2,023	2,340

Six-Lag ADF Specification ($m = 6$)

	SAMPLE		
Unit Root Test	EUROPE	OECD	OPEN
OC–GLS	0.01	0.01	0.01
SUR–GLS	0.01	0.01	0.01
LL2–GLS	nr	nr	nr
IPS–GLS	0.05	0.01	0.01
N	16	22	26
$\min T_i$	65	65	47
$\max T_i$	93	93	93
$\overline{T}$	88.8	90	88
Obs.	1,421	1,979	2,288

Notes: The entries in the table correspond to the significance level at which the null hypothesis of a unit root can be rejected; nr indicates that the null hypothesis cannot be rejected at the 10 percent (or better) significance level. All empirical significance levels are based on 10,000 Monte Carlo replications of a panel with the cross-sectional dimension N and time series dimension $\overline{T}$.

TABLE 3

Panel Data Unit Root Tests of PPP
CPI Based Real Exchange Rates
Empirical Power Properties

	Zero Lag ADF Specification ($m = 0$)								
	EUROPE Test Size			OECD Test Size			OPEN Test Size		
Unit Root Test	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
OC–GLS	0.59	0.86	0.94	0.74	0.93	0.97	0.82	0.95	0.98
SUR–GLS	0.51	0.80	0.91	0.55	0.85	0.94	0.48	0.84	0.94
LL2–GLS	0.06	0.28	0.45	0.07	0.31	0.51	0.10	0.31	0.51
IPS–GLS	0.46	0.77	0.88	0.59	0.85	0.93	0.67	0.87	0.94

	Four Lag ADF Specification ($m = 4$)								
	EUROPE Test Size			OECD Test Size			OPEN Test Size		
Unit Root Test	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
OC–GLS	0.34	0.70	0.83	0.50	0.79	0.89	0.64	0.86	0.94
SUR–GLS	0.37	0.72	0.86	0.46	0.79	0.90	0.43	0.76	0.89
LL2–GLS	0.00	0.03	0.08	0.00	0.03	0.09	0.00	0.03	0.07
IPS–GLS	0.35	0.65	0.80	0.48	0.76	0.86	0.52	0.79	0.89

	Eight Lag ADF Specification ($m = 8$)								
	EUROPE Test Size			OECD Test Size			OPEN Test Size		
Unit Root Test	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
OC–GLS	0.16	0.45	0.63	0.22	0.58	0.74	0.36	0.68	0.82
SUR–GLS	0.27	0.62	0.77	0.33	0.67	0.83	0.25	0.63	0.81
LL2–GLS	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.01	0.02
IPS–GLS	0.22	0.53	0.68	0.31	0.61	0.76	0.37	0.67	0.80

Notes: All power calculations are based on 5,000 Monte Carlo replications. The data for each panel dimension are generated under the alternative hypothesis of stationarity: $q_{i,t} = 0.96q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$.

TABLE 4

Panel Data Unit Root Tests of PPP
WPI Based Real Exchange Rates
Empirical Power Properties

Unit Root Test	EUROPE Test Size 1.0%	EUROPE 5.0%	EUROPE 10.0%	OECD Test Size 1.0%	OECD 5.0%	OECD 10.0%	OPEN Test Size 1.0%	OPEN 5.0%	OPEN 10.0%
Zero Lag ADF Specification ($m = 0$)									
OC–GLS	0.47	0.75	0.87	0.63	0.87	0.95	0.69	0.90	0.95
SUR–GLS	0.40	0.74	0.86	0.47	0.82	0.92	0.49	0.81	0.92
LL2–GLS	0.04	0.21	0.37	0.06	0.23	0.41	0.07	0.28	0.46
IPS–GLS	0.37	0.67	0.80	0.48	0.78	0.89	0.51	0.81	0.95
Four Lag ADF Specification ($m = 4$)									
OC–GLS	0.25	0.55	0.71	0.37	0.70	0.84	0.44	0.75	0.87
SUR–GLS	0.27	0.63	0.78	0.41	0.73	0.86	0.41	0.75	0.87
LL2–GLS	0.00	0.02	0.06	0.00	0.03	0.08	0.00	0.02	0.07
IPS–GLS	0.27	0.58	0.73	0.37	0.67	0.80	0.39	0.68	0.82
Eight Lag ADF Specification ($m = 8$)									
OC–GLS	0.10	0.35	0.54	0.19	0.48	0.66	0.22	0.53	0.70
SUR–GLS	0.18	0.53	0.70	0.27	0.61	0.78	0.26	0.60	0.76
LL2–GLS	0.00	0.00	0.02	0.00	0.01	0.02	0.00	0.01	0.01
IPS–GLS	0.17	0.46	0.62	0.25	0.55	0.70	0.28	0.59	0.74

Notes: All power calculations are based on 5,000 Monte Carlo replications. The data for each panel dimension are generated under the alternative hypothesis of stationarity: $q_{i,t} = 0.96q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$.

TABLE 5

Rate of Convergence to PPP
Actual and Bias-Corrected SUR–GLS Estimates of ρ
Four-Lag ADF Specification

CPI Based Real Exchange Rates					
EUROPE		OECD		OPEN	
Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}$	B-C $\hat{\rho}$
-0.064	-0.032	-0.061	-0.020	-0.047	-0.002

WPI Based Real Exchange Rates					
EUROPE		OECD		OPEN	
Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}$	B-C $\hat{\rho}$
-0.087	-0.057	-0.086	-0.056	-0.095	-0.070

Notes: Bias-corrected (B-C) estimates of the decay parameter ρ are computed by matching the actual empirical estimates of ρ with Monte Carlo estimates generated using various assumed values for ρ. For each replication, N stationary processes of length $T + 30$ are generated as $q_{i,t} = (1+\rho)\, q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ and $\rho = -.050, -.045, \ldots, -.001$. The first 25 observations are discarded to minimize the effect of the initial value bias. The following N-equation system of ADF regressions is estimated in a SUR framework with iterative GLS: $\Delta q_{i,t} = \alpha_i + \rho q_{i,t-1} + \sum_{k=1}^{4} \lambda_k \Delta q_{i,t-k} + \eta_{i,t}$, $t = 1, \ldots, T$. There were 5,000 Monte Carlo replications for each assumed value of ρ.

TABLE 6

Rate of Convergence to PPP
Actual and Bias-Corrected IPS–GLS Estimates of ρ
Four-Lag ADF Specification

CPI Based Real Exchange Rates

EUROPE		OECD		OPEN	
Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}_{MG}$	B-C $\hat{\rho}$
-0.071	-0.035	-0.070	-0.035	-0.059	-0.015

WPI Based Real Exchange Rates

EUROPE		OECD		OPEN	
Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}$	B-C $\hat{\rho}$	Actual $\hat{\rho}$	B-C $\hat{\rho}$
-0.077	-0.035	-0.082	-0.045	-0.083	-0.045

Notes: Bias-corrected (B-C) estimates of the decay parameter ρ are computed by matching the actual empirical estimates of ρ with Monte Carlo estimates generated using various assumed values for ρ. For each replication, N stationary processes of length $T+30$ are generated as $q_{i,t} = (1+\rho)\,q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ and $\rho = -.050, -.045, \ldots, -.001$. The first 25 observations are discarded to minimize the effect of the initial value bias. A GLS transformation, rendering the error term cross-sectionally homoscedastic, is performed on the data, and for each of the N countries, the following ADF regressions is estimated by OLS: $\Delta q_{i,t} = \alpha_i + \rho_i q_{i,t-1} + \sum_{k=1}^{4} \lambda_{i,k} \Delta q_{i,t-k} + \eta_{i,t}$, $t = 1, \ldots, T$. The mean-group (MG) estimate of ρ is computed as $\hat{\rho} = \frac{1}{N}\sum_{i=1}^{N} \hat{\rho}_i$. There were 5,000 Monte Carlo replications for each assumed value of ρ.

A LL2 and IPS Tests

A.1 LL2: Levin and Lin (1993)

The LL2 test can be summarized in four steps. First, two country-specific auxiliary regressions are estimated using OLS to obtain orthogonalized residuals $\hat{e}_{i,t}$ and $\hat{\nu}_{i,t-1}$, defined as

$$\hat{e}_{i,t} = \Delta q_{i,t} - \hat{\alpha}_i - \sum_{k=1}^{m_i} \hat{\gamma}_{i,k} \Delta q_{i,t-k}; \tag{1}$$

$$\hat{\nu}_{i,t-1} = q_{i,t-1} - \hat{\alpha}_i - \sum_{k=1}^{m_i} \hat{\theta}_{i,k} \Delta q_{i,t-k}. \tag{2}$$

To control for heterogeneity across countries, the residuals $\hat{e}_{i,t}$ and $\hat{\nu}_{i,t-1}$ are normalized, where the normalization factor is equal to the standard error of the country-specific regression of $\hat{e}_{i,t}$ on $\hat{v}_{i,t-1}$. That is, the heterogeneity-corrected orthogonalized residuals $\tilde{e}_{i,t}$ and $\tilde{v}_{i,t-1}$ are defined as

$$\tilde{e}_{i,t} = \frac{\hat{e}_{i,t}}{\hat{\sigma}_{i,T}} \quad \text{and} \quad \tilde{v}_{i,t-1} = \frac{\hat{v}_{i,t-1}}{\hat{\sigma}_{i,T}}, \tag{3}$$

where

$$\hat{\sigma}^2_{i,T} = \left(\frac{1}{T - m_i - 1}\right) \sum_{t=m_i+2}^{T} (\hat{e}_{i,t} - \hat{\delta}_i \hat{v}_{i,t-1})^2 \tag{4}$$

and $\hat{\delta}_i$ is the OLS estimator of δ_i in the country-specific regression of $\hat{e}_{i,t}$ on $\hat{v}_{i,t-1}$.

Second, Levin and Lin estimate the ratio of long-run to short-run (innovation) standard deviation for each country-specific ADF regression and calculate the average of this ratio for the entire panel as:

$$\widehat{S}_{NT} = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{\sigma}_{\eta_i,T}}{\hat{\sigma}_{i,T}}, \tag{5}$$

where $\hat{\sigma}^2_{\eta_i,T}$ is the estimator of the long-run variance $\sigma^2_{\eta_i,T}$, obtained as

$$\hat{\sigma}^2_{\eta_i,T} = \frac{1}{T-1} \sum_{t=2}^{T} (\Delta q_{i,t})^2 + 2 \sum_{j=1}^{L} w(j,L) \left(\frac{1}{T-1} \sum_{t=2+j}^{T} \Delta q_{i,t} \Delta q_{i,t-j} \right), \tag{6}$$

where $w(j, L)$ are the sample covariance weights to ensure a non-negative value of $\hat{\sigma}^2_{\eta_i,T}$.[25]

Third, the authors consider the following pooled regression:

$$\tilde{e}_{i,t} = \beta \tilde{v}_{i,t-1} + error, \quad i = 1, \ldots, N; \quad t = m_i + 2, \ldots, T, \tag{7}$$

where $\tilde{e}_{i,t}$ and $\tilde{v}_{i,t-1}$ denote the heterogeneity-corrected normalized residuals from step one. The null hypothesis of a unit root is tested against the alternative of stationarity by considering the t-statistic for testing the hypothesis that $\beta = 0$:

$$t_{\hat{\beta}_{NT}} = \frac{\hat{\beta}_{NT}}{RSE(\hat{\beta}_{NT})}, \tag{8}$$

[25] Levin and Lin (1993) advocate the use of Bartlett weights, $w(j, L) = j/(L+1)$, proposed by Newey and West (1987); in this case, the estimator $\hat{\sigma}^2_{\eta_i,T}$ is consistent if the lag truncation parameter L grows exponentially at a rate less than T.

where $\hat{\beta}_{NT}$ is the pooled OLS estimator of β,

$$RSE(\hat{\beta}_{NT}) = \hat{\sigma}_{NT}\left[\sum_{i=1}^{N}\sum_{t=2+m_i}^{T}\tilde{v}_{i,t-1}^2\right]^{-\frac{1}{2}} \tag{9}$$

is the standard error of $\hat{\beta}_{NT}$, and $\hat{\sigma}_{NT}$ is the standard error of the regression, computed as the square root of

$$\hat{\sigma}_{NT}^2 = \frac{1}{N\widetilde{T}}\sum_{i=1}^{N}\sum_{t=2+m_i}^{T}(\tilde{e}_{i,t} - \hat{\beta}_{NT}\tilde{v}_{i,t-1})^2, \tag{10}$$

where $\widetilde{T} = T - \overline{m} - 1$, and $\overline{m} = \frac{1}{N}\sum_{i=1}^{N} m_i$.

Finally, the authors show that as $N \to \infty$, $T \to \infty$, and $\frac{N}{T} \to 0$, a function of the t-statistic $t_{\hat{\beta}_{NT}}$ converges to a standard normal variate. In particular, the specific functional form for the Levin and Lin (1993) test statistic is

$$LL2 = \frac{t_{\hat{\beta}_{NT}} - N\widetilde{T}\widehat{S}_{NT}\hat{\sigma}_{NT}^{-2}RSE(\hat{\beta}_{NT})\mu_{\widetilde{T}}^*}{\sigma_{\widetilde{T}}^*} \overset{a}{\sim} N(0,1), \tag{11}$$

where $\mu_{\widetilde{T}}^*$ and $\sigma_{\widetilde{T}}^*$ are the mean and standard deviation adjustment factors obtained by the authors via stochastic simulations, so that the LL2 test statistic retains a $N(0,1)$ distribution.

The adjustment factors are tabulated for $N = 250$ and different values of $\widetilde{T}$. In particular, at each Monte Carlo replication, Gaussian random numbers with unit variance are used to generate 250 independent unit root processes of time series dimension $\widetilde{T}+1$. The panel is then used to compute the sample statistics $\hat{\beta}_{NT}$, $t_{\hat{\beta}_{NT}}$, $\widehat{S}_{NT}$, $\hat{\sigma}_{NT}$, and $RSE(\hat{\beta}_{NT})$, using the lag truncation parameter $L = 3.21T^{\frac{1}{3}}$ (see Andrews (1991)) and $m_i = 0$, for all i. Finally, based on 25,000 replications, $\mu_{\widetilde{T}}^*$ is computed as the mean of $t_{\hat{\beta}_{NT}}/N\widetilde{T}\widehat{S}_{NT}\hat{\sigma}_{NT}^{-2}RSE(\hat{\beta}_{NT})$, and $\sigma_{\widetilde{T}}^*$ is computed as the standard deviation of $t_{\hat{\beta}_{NT}} - N\widetilde{T}\widehat{S}_{NT}\hat{\sigma}_{NT}^{-2}RSE(\hat{\beta}_{NT})\mu_{\widetilde{T}}^*$.

A.2 IPS: Im, Pesaran, and Shin (1995)

The IPS t-bar statistic, $\bar{t}_{NT}$, is based on the average of the t-statistics for $\hat{\rho}_i$, obtained from country-specific estimation of the standard ADF regression. The t-bar statistic is defined as:

$$\bar{t}_{NT} = \frac{1}{N}\sum_{i=1}^{N} t_i(T_i, m_i, \hat{\rho}_i), \tag{1}$$

where $t_i(T_i, m_i, \hat{\rho}_i)$ denotes the t-statistic of the OLS estimator of ρ_i, based on the sample of length T_i and with m_i lags in the ADF regression. Im, Pesaran, and Shin (1995) show that as $N \to \infty$, $T \to \infty$, and $\frac{N}{T} \to 0$, the standardized $\bar{t}_{NT}$ statistic converges to the $N(0,1)$ distribution. That is, the IPS test statistic is given by:

$$IPS = \frac{\sqrt{N}(\bar{t}_{NT} - a_{NT})}{\sqrt{b_{NT}}} \overset{a}{\sim} N(0,1), \tag{2}$$

where a_{NT} and b_{NT} are the mean and variance adjustment factors for the t-bar statistic $\bar{t}_{NT}$ obtained via stochastic simulations.

Im, Pesaran, and Shin tabulate the adjustment factors a_{NT} and b_{NT} for $N = 1$ and different values of T and m. In particular, at each Monte Carlo replication, Gaussian random numbers with unit variance are used to generate a random walk of dimension $T + m + 1$. Next, the authors estimate the standard ADF regression to obtain the t-statistic for testing $\rho = 0$, denoted by $t^j(T, m, 0)$, where j stands for the j-th Monte Carlo replication. This procedure is repeated R times, and the mean

$$a_{NT} = E_R[t(T, m, 0)] = \frac{1}{R} \sum_{j=1}^{R} t^j(T, m, 0), \tag{3}$$

and the variance:

$$b_{NT} = V_R[t(T, m, 0)] = \frac{1}{R} \sum_{j=1}^{R} \left[t^j(T, m, 0) - E_R[t(T, m, 0)] \right]^2 \tag{4}$$

are computed.[26]

[26] The adjustment factors reported by Im, Pesaran, and Shin (1995) are based on 50,000 replications.

A.3 Appendix Tables

TABLE A.1

Sample Membership
CPI Based Real Exchange Rates

Country	Sample: EUROPE	Sample: OECD	Sample: OPEN	Start	End	T
Australia		√	√	73Q1	97Q4	100
Austria	√	√	√	73Q1	97Q4	100
Barbados			√	73Q1	97Q4	100
Belgium	√	√	√	73Q1	97Q4	100
Botswana			√	79Q1	97Q4	76
Canada		√	√	73Q1	97Q4	100
Chile			√	76Q1	97Q4	88
Cyprus	√		√	73Q1	97Q4	100
Denmark	√	√	√	73Q1	97Q4	100
Finland	√	√	√	73Q1	97Q4	100
France	√	√	√	73Q1	97Q4	100
Germany	√	√	√	73Q1	97Q4	100
Greece	√	√	√	73Q1	97Q4	100
Hong Kong			√	76Q1	97Q4	88
Iceland	√	√		73Q1	97Q4	100
Indonesia			√	73Q1	97Q4	100
Ireland	√	√	√	73Q1	97Q4	100
Italy	√	√	√	73Q1	97Q4	100

TABLE A.1 (Continued)

Country	Sample: EUROPE	OECD	OPEN	Start	End	T
Japan		√	√	73Q1	97Q4	100
Jordan			√	76Q1	97Q3	87
Korea		√	√	73Q1	97Q4	100
Luxembourg	√	√	√	73Q1	97Q4	100
Malaysia			√	73Q1	97Q4	100
Mauritius			√	73Q1	97Q4	100
Mexico		√	√	73Q1	97Q4	100
Netherlands	√	√	√	73Q1	97Q4	100
New Zealand		√		73Q1	97Q4	100
Norway	√	√	√	73Q1	97Q4	100
Portugal	√	√	√	73Q1	97Q4	100
Singapore			√	73Q1	97Q4	100
Spain	√	√	√	73Q1	97Q4	100
Sweden	√	√	√	73Q1	97Q4	100
Switzerland	√	√	√	73Q1	97Q4	100
Thailand			√	73Q1	97Q3	99
Turkey		√		73Q1	97Q4	100
United Kingdom	√	√	√	73Q1	97Q4	100

TABLE A.2

Univariate ADF Unit Root Test of PPP
CPI Based Real Exchange Rates

Country	AIC Lag Selection $m=4$	$m=5$	$m=6$	$m=7$	$m=8$	p-value	T
United Kingdom	$\surd$					0.09	95
Austria	$\surd$					0.17	95
Belgium	$\surd$					0.18	95
Denmark	$\surd$					0.13	95
France	$\surd$					0.16	95
Germany	$\surd$					0.16	95
Italy	$\surd$					0.16	95
Luxembourg	$\surd$					0.22	95
Netherlands	$\surd$					0.13	95
Norway	$\surd$					0.23	95
Sweden	$\surd$					0.28	95
Switzerland	$\surd$					0.12	95
Canada	$\surd$					0.57	95
Japan	$\surd$					0.28	95
Finland	$\surd$					0.08	95
Greece	$\surd$					0.20	95
Iceland	$\surd$					0.16	95
Ireland	$\surd$					0.09	95

TABLE A.2 (Continued)

Country	AIC Lag Selection[a] $m=4$	$m=5$	$m=6$	$m=7$	$m=8$	p-value[b]	T
Portugal	√					0.40	95
Spain	√					0.21	95
Turkey	√					0.61	95
Australia	√					0.36	95
New Zealand	√					0.18	95
Chile	√					0.39	83
Mexico	√					0.09	95
Barbados		√				0.09	94
Cyprus	√					0.16	95
Jordan	√					0.65	82
Hong Kong	√					0.81	80
Indonesia	√					0.95	95
South Korea	√					0.42	95
Malaysia	√					0.86	95
Singapore	√					0.13	95
Thailand	√					0.64	94
Botswana	√					0.16	71
Mauritius		√				0.60	94

Notes: Each country-specific ADF test includes m lags of the change in the real exchange rate, Δq_t, where m is determined by the minimum AIC.

[a]√ indicates the ADF specification with the minimum AIC.

[b]Probability value for the univariate ADF(m) test of the null hypothesis that the real exchange rate contains a unit root.

TABLE A.3

Sample Membership
WPI Based Real Exchange Rates

Country	Sample			Start	End	T
	EUROPE	OECD	OPEN			
Australia		√	√	73Q1	97Q4	100
Austria	√	√	√	73Q1	97Q4	100
Belgium	√	√	√	80Q2	97Q4	72
Canada		√	√	73Q1	97Q4	100
Chile			√	76Q1	97Q4	88
Denmark	√	√	√	73Q1	97Q4	100
Finland	√	√	√	73Q1	97Q4	100
France	√	√	√	73Q1	97Q4	100
Germany	√	√	√	73Q1	97Q4	100
Greece	√	√	√	73Q1	97Q4	100
Indonesia			√	73Q1	97Q4	100
Ireland	√	√	√	73Q1	97Q2	98
Italy	√	√	√	73Q1	97Q4	100
Japan		√	√	73Q1	97Q4	100
Korea		√	√	73Q1	97Q4	100
Luxembourg	√	√	√	80Q1	97Q4	72
Malaysia			√	84Q1	97Q2	54
Mexico		√	√	73Q1	97Q4	100
Netherlands	√	√	√	73Q1	97Q4	100
New Zealand		√		73Q1	97Q4	100
Norway	√	√	√	73Q1	97Q3	99
Singapore			√	74Q1	97Q4	96
Spain	√	√	√	75Q1	97Q4	92
Sweden	√	√	√	73Q1	97Q4	100
Switzerland	√	√	√	73Q1	97Q4	100
Taiwan			√	73Q1	97Q4	100
Thailand			√	73Q1	97Q2	99
United Kingdom	√	√	√	73Q1	97Q4	100

TABLE A.4

Univariate ADF Unit Root Test of PPP
WPI Based Real Exchange Rates

Country	AIC Lag Selection[a] $m = 4$	$m = 5$	$m = 6$	$m = 7$	$m = 8$	p-value[b]	T
United Kingdom	√					0.25	95
Austria	√					0.11	95
Belgium	√					0.23	67
Denmark	√					0.20	95
France	√					0.11	95
Germany	√					0.15	95
Italy	√					0.18	95
Luxembourg	√					0.28	67
Netherlands	√					0.22	95
Norway	√					0.23	94
Sweden	√					0.18	95
Switzerland	√					0.11	95
Canada	√					0.17	95
Japan	√					0.25	95
Finland	√					0.14	95
Greece	√					0.13	95
Ireland	√					0.10	93
Spain	√					0.21	87
Australia	√					0.17	95
New Zealand	√					0.27	95
Chile	√					0.27	83
Mexico	√					0.06	95
Indonesia	√					0.95	95
South Korea	√					0.90	95
Malaysia	√					0.47	49
Singapore	√					0.54	91
Thailand	√					0.30	94
Taiwan			√			0.13	93

Notes: Each country-specific ADF test includes m lags of the change in the real exchange rate, Δq_t, where m is determined by the minimum AIC.
[a]√ indicates the ADF specification with the minimum AIC.
[b]Probability value for the univariate ADF(m) test of the null hypothesis that the real exchange rate contains a unit root.

TABLE A.5

Monte Carlo Critical Values
Panel Data Unit Root Test: OC–GLS
CPI Based Real Exchange Rates

Four-Lag Pre-whitening Regression

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-7.21	-6.60	-6.27	-7.84	-7.22	-6.90	-8.39	-7.79	-7.43
$N = 19$, $T = 95$			$N = 25$, $T = 95$			$N = 32$, $T = 93$		

Six-Lag Pre-whitening Regression

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-7.22	-6.63	-6.30	-7.84	-7.23	-6.87	-8.36	-7.78	-7.44
$N = 19$, $T = 93$			$N = 25$, $T = 93$			$N = 32$, $T = 91$		

Eight-Lag Pre-whitening Regression

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-7.20	-6.60	-6.27	-7.91	-7.22	-6.89	-8.40	-7.74	-7.39
$N = 19$, $T = 91$			$N = 25$, $T = 91$			$N = 32$, $T = 89$		

Notes: All critical values are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T + 26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. The following specification is estimated using O'Connell's (1998) GLS procedure: $\Delta q_{i,t} = \alpha_i + \rho q_{i,t-1} + \eta_{i,t}$, $t = 1, \ldots, T$, and the 1st, 5th, and the 10th percentile of the distribution of the t-statistic $t_{\hat{\rho}}$ for the hypothesis $\rho = 0$ are reported.

TABLE A.6

Monte Carlo Critical Values
Panel Data Unit Root Test: OC–GLS
WPI Based Real Exchange Rates

Four-Lag Pre-whitening Regression

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-6.79	-6.25	-5.93	-7.54	-6.94	-6.60	-7.92	-7.32	-6.96
$N = 16$, $T = 91$			$N = 22$, $T = 92$			$N = 26$, $T = 90$		

Six-Lag Pre-whitening Regression

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-6.79	-6.22	-5.89	-7.52	-6.91	-6.57	-7.89	-7.27	-6.93
$N = 16$, $T = 89$			$N = 23$, $T = 90$			$N = 26$, $T = 88$		

Eight-Lag Pre-whitening Regression

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-6.80	-6.20	-5.87	-7.51	-6.92	-6.58	-7.93	-7.31	-6.95
$N = 16$, $T = 87$			$N = 22$, $T = 88$			$N = 26$, $T = 86$		

Notes: All critical values are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T + 26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. The following specification is estimated using O'Connell's (1998) GLS procedure: $\Delta q_{i,t} = \alpha_i + \rho q_{i,t-1} + \eta_{i,t}$, $t = 1, \ldots, T$, and the 1st, 5th, and the 10th percentile of the distribution of the t-statistic $t_{\hat{\rho}}$ for the hypothesis $\rho = 0$ are reported.

TABLE A.7

Monte Carlo Critical Values
Panel Data Unit Root Test: SUR–GLS
CPI Based Real Exchange Rates

Four-Lag ADF Specification ($m = 4$)

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-8.32	-7.65	-7.29	-9.60	-8.9	-8.52	-11.27	-10.54	-10.14
$N = 19$, $T = 95$			$N = 25$, $T = 95$			$N = 32$, $T = 93$		

Six-Lag ADF Specification ($m = 6$)

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-8.31	-7.67	-7.29	-9.64	-8.92	-8.55	-11.46	-10.63	-10.19
$N = 19$, $T = 93$			$N = 25$, $T = 93$			$N = 32$, $T = 91$		

Eight-Lag ADF Specification ($m = 8$)

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-8.39	-7.69	-7.32	-9.70	-8.96	-8.57	-11.57	-10.67	-10.22
$N = 19$, $T = 91$			$N = 25$, $T = 91$			$N = 32$, $T = 89$		

Notes: All critical values are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T + m + 26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. The following N-equation system of ADF regressions is estimated in a SUR framework with iterative GLS: $\Delta q_{i,t} = \alpha_i + \rho q_{i,t-1} + \sum_{k=1}^{m} \lambda_k \Delta q_{i,t-k} + \eta_{i,t}$, $t = 1, \ldots, T$, and the 1st, 5th, and the 10th percentile of the distribution of the t-statistic $t_{\hat{\rho}}$ for the hypothesis $\rho = 0$ are reported.

TABLE A.8

Monte Carlo Critical Values
Panel Data Unit Root Test: SUR–GLS
WPI Based Real Exchange Rates

Four-Lag ADF Specification ($m = 4$)

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-7.74	-7.04	-6.71	-8.96	-8.30	-7.94	-9.86	-9.18	-8.84
$N = 16$, $T = 91$			$N = 22$, $T = 92$			$N = 26$, $T = 90$		

Six-Lag ADF Specification ($m = 6$)

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-7.73	-7.06	-6.49	-8.99	-8.36	-7.95	-10.03	-9.25	-8.86
$N = 16$, $T = 89$			$N = 22$, $T = 90$			$N = 26$, $T = 88$		

Eight-Lag ADF Specification ($m = 8$)

EUROPE			OECD			OPEN		
Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)			Critical Value ($t_{\hat{\rho}}$)		
1.0%	5.0%	10.0%	1.0%	5.0%	10.0%	1.0%	5.0%	10.0%
-7.82	-7.07	-6.69	-9.13	-8.37	-7.97	-10.08	-9.31	-8.92
$N = 16$, $T = 87$			$N = 22$, $T = 88$			$N = 26$, $T = 86$		

Notes: All critical values are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T + m + 26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. The following N-equation system of ADF regressions is estimated in a SUR framework with iterative GLS: $\Delta q_{i,t} = \alpha_i + \rho q_{i,t-1} + \sum_{k=1}^{m} \lambda_k \Delta q_{i,t-k} + \eta_{i,t}$, $t = 1, \ldots, T$, and the 1st, 5th, and the 10th percentile of the distribution of the t-statistic $t_{\hat{\rho}}$ for the hypothesis $\rho = 0$ are reported.

TABLE A.9

Mean and Standard Deviation Adjustment Factors
Panel Data Unit Root Test: LL2–GLS
CPI Based Real Exchange Rates

Four-Lag ADF Specification ($m = 4$)

EUROPE		OECD		OPEN	
$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$
-0.485	0.892	-0.484	0.889	-0.480	0.947
$N = 19, T = 95$		$N = 25, T = 95$		$N = 32, T = 93$	

Six-Lag ADF Specification ($m = 6$)

EUROPE		OECD		OPEN	
$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$
-0.471	0.903	-0.470	0.931	-0.465	0.990
$N = 19, T = 93$		$N = 25, T = 93$		$N = 32, T = 91$	

Eight-Lag ADF Specification ($m = 8$)

EUROPE		OECD		OPEN	
$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$
-0.457	0.951	-0.454	0.988	-0.450	1.030
$N = 19, T = 91$		$N = 25, T = 91$		$N = 32, T = 89$	

Notes: The mean and the variance adjustment factors used to standardize the LL2–GLS statistic are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T + m + 26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. As described in the text, first a GLS transformation is performed on the data, rendering the error term cross-sectionally homoscedastic, then the computational test procedure follows Levin and Lin (1993).

TABLE A.10

Mean and Standard Deviation Adjustment Factors
Panel Data Unit Root Test: LL2–GLS
WPI Based Real Exchange Rates

Four-Lag ADF Specification $(m = 4)$

EUROPE		OECD		OPEN	
$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$
-0.486	0.869	-0.482	0.902	-0.480	0.921
$N = 16$, $T = 91$		$N = 22$, $T = 92$		$N = 26$, $T = 90$	

Six-Lag ADF Specification $(m = 6)$

EUROPE		OECD		OPEN	
$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$
-0.467	0.926	-0.469	0.934	-0.466	0.958
$N = 16$, $T = 89$		$N = 22$, $T = 90$		$N = 26$, $T = 88$	

Eight-Lag ADF Specification $(m = 8)$

EUROPE		OECD		OPEN	
$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$	$\mu^*_{\tilde{T}}$	$\sigma^*_{\tilde{T}}$
-0.453	0.972	-0.452	0.973	-0.451	0.993
$N = 16$, $T = 87$		$N = 22$, $T = 88$		$N = 26$, $T = 86$	

Notes: The mean and the standard deviation adjustment factors used to standardize the LL2–GLS statistic are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T+m+26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. As described in the text, first a GLS transformation is performed on the data, rendering the error term cross-sectionally homoscedastic, then the computational test procedure follows Levin and Lin (1993).

TABLE A.11

Mean and Variance Adjustment Factors
Panel Data Unit Root Test: IPS–GLS
CPI Based Real Exchange Rates

Four-Lag ADF Specification ($m = 4$)					
EUROPE		OECD		OPEN	
a_{NT}	b_{NT}	a_{NT}	b_{NT}	a_{NT}	b_{NT}
-1.433	0.825	-1.414	0.836	-1.394	0.857
$N = 19, T = 95$		$N = 25, T = 95$		$N = 32, T = 93$	
Six-Lag ADF Specification ($m = 6$)					
EUROPE		OECD		OPEN	
a_{NT}	b_{NT}	a_{NT}	b_{NT}	a_{NT}	b_{NT}
-1.412	0.847	-1.391	0.866	-1.362	0.887
$N = 19, T = 93$		$N = 25, T = 93$		$N = 32, T = 91$	
Eight-Lag ADF Specification ($m = 8$)					
EUROPE		OECD		OPEN	
a_{NT}	b_{NT}	a_{NT}	b_{NT}	a_{NT}	b_{NT}
-1.382	0.882	-1.359	0.901	-1.330	0.921
$N = 19, T = 91$		$N = 25, T = 91$		$N = 32, T = 89$	

Notes: The mean and the variance adjustment factors used to standardize the IPS–GLS statistic are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T + m + 26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. As described in the text, first a GLS transformation is performed on the data, rendering the error term cross-sectionally homoscedastic, then the computational test procedure follows Im, Pesaran, and Shin (1995).

TABLE A.12

Mean and Variance Adjustment Factors
Panel Data Unit Root Test: IPS–GLS
WPI Based Real Exchange Rates

Four-Lag ADF Specification ($m = 4$)					
EUROPE		OECD		OPEN	
a_{NT}	b_{NT}	a_{NT}	b_{NT}	a_{NT}	b_{NT}
-1.437	0.823	-1.424	0.835	-1.406	0.846
$N = 16, T = 91$		$N = 22, T = 92$		$N = 26, T = 90$	
Six-Lag ADF Specification ($m = 6$)					
EUROPE		OECD		OPEN	
a_{NT}	b_{NT}	a_{NT}	b_{NT}	a_{NT}	b_{NT}
-1.412	0.851	-1.392	0.869	-1.377	0.880
$N = 16, T = 89$		$N = 22, T = 90$		$N = 26, T = 88$	
Eight-Lag ADF Specification ($m = 8$)					
EUROPE		OECD		OPEN	
a_{NT}	b_{NT}	a_{NT}	b_{NT}	a_{NT}	b_{NT}
-1.384	0.879	-1.363	0.899	-1.346	0.916
$N = 16, T = 87$		$N = 22, T = 88$		$N = 26, T = 86$	

Notes: The mean and the variance adjustment factors used to standardize the IPS–GLS statistic are based on 10,000 Monte Carlo replications. For each replication, N random walks of length $T + m + 26$ are generated as $q_{i,t} = q_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t} \sim N(0,1)$ The first 25 observations are discarded to minimize the effect of the initial value bias. As described in the text, first a GLS transformation is performed on the data, rendering the error term cross-sectionally homoscedastic, then the computational test procedure follows Im, Pesaran, and Shin (1995).