A New Look at Panel Testing of Stationarity and the PPP Hypothesis

Jushan Bai[*]   Serena Ng[†]

October 2001

Abstract

This paper uses a decomposition of the data into common and idiosyncratic components to develop procedures that test if these components satisfy the null hypothesis of stationarity. The decomposition also allows us to construct pooled tests that satisfy the cross-section independence assumption. In simulations, tests on the components separately generally have better properties than testing the observed series. However, the results are less than satisfactory, especially in comparison with similar procedures developed for unit root tests. The problem can be traced to the properties of the stationarity test, and is not due to the weakness of the common-idiosyncratic decomposition. We apply both panel stationarity and unit root tests to real exchange rates. We found evidence in support of a large stationary common factor. Rejections of PPP are likely due to non-stationarity of country-specific variations.

Keywords: Common factors, common trends, principal components, unit root, cointegration

[*]Dept. of Economics, Boston College, Chestnut Hill, MA 02467 Email Jushan.Bai@bc.edu
[†]Dept. of Economics, Johns Hopkins University, Baltimore, MD 21218 Email: Serena.Ng@jhu.edu

1   Introduction

Statistical hypotheses are typically set up in a way to favor the null hypothesis unless there is substantial evidence against it. For many economic problems, the null hypothesis of stationarity is more natural than the null hypothesis of a unit root. For instance, purchasing power parity (PPP) implies that real exchange rates are stationary. Testing PPP has been a motivating example for many papers in the literature on testing stationarity.[1] This is because if PPP is rejected for the stationarity null, then the evidence against PPP would be stronger than merely not rejecting the unit root null hypothesis. Testing PPP has also been the primary motivation for developing panel testing procedures, the idea being that time series of cross sections should contain more information about the dynamic properties of the data than a single series, and panel tests should be expected to have more power than testing the series individually. Testing stationarity in a panel of data will also the pursuit of the present paper. Our work distinguishes itself from the large literature on this topic by making a distinction between common and idiosyncratic variations. Tests are applied to these components rather than the observed data themselves.

Testing whether a panel of data is I(1) or I(0) has been considered by many researchers. Quah (1994), Levin and Lin (1993), Im, Pesaran and Shin (1997), Hardi (2000), and Pedroni (1995) among others derived asymptotic distributions for panel unit root and cointegration tests under various assumptions about fixed effects and heterogeneous time trends. The statistics are usually based on averaging of tests or estimated parameters obtained for the individual units. But the average of statistics could be misleading especially when the distribution of the test is skewed. Maddala and Wu (1999), and Choi (2001) considered averaging the p-values associated with the individual tests. Pooling p values has the added advantage that panel tests can be constructed whether non-stationarity or stationarity is taken as the null hypothesis. However, regardless of the maintained hypothesis, most (if not all) of these pooled tests assume cross-sectional independence. This assumption, while convenient, is unlikely to be true for macroeconomic time series. For example, real exchange rates are often defined using the same base country; cross-sectional correlation arises almost by construction. In cross-country and sectoral analysis, the independence assumption rules out shocks that affect more than one country or one sector. More generally, shocks that are common to a large number of units in a panel imply the presence of strong cross-section correlation that cannot be aggregated away.

The assumption of cross-section independence can potentially create two problems. First, a non-rejection of the unit root null would imply the existence of N independent unit roots when in

---

[1]See e.g. Papell (1997), Culver and Papell (1999), Kuo and Mikkola (1999), and the references therein.

fact fewer than N exist because of common stochastic trends. Second, it is under the independence assumption that pooled tests have been shown to be more powerful than testing the series one by one. O'Connell (1998) showed that the pooled tests will over-reject the null hypothesis when the independence assumption is violated, whether the null hypothesis is a unit root or stationarity. Size distortions could be misread as higher power. Banerjee, Marcellino and Osbat (2001) argued against use of panel unit root test because of this potential problem.

A useful framework, and one that will be used in the present study, is to model strong cross-section correlation via a factor model:

$$X_{it} = D_{it} + \lambda_i' F_t + e_{it} \quad (i = 1, 2, ..., N, \ t = 1, 2, ..., T). \tag{1}$$

In a factor model, the data can be decomposed into three components:- D_{it} is the deterministic component, F_t $(r \times 1)$ is a vector of unobservable common factors, and e_{it} is a unit-specific stochastic term. The presence of F_t implies non-trivial cross-sectional correlation. Associated with these common factors are loadings denoted λ_i. These represent the exposure of cross-section i to the common factors. Some λ_i's are allowed to be zero, so that some cross-sections may not be influenced by the common factors. As is common assumed in factor analysis, the specific component e_{it} is cross-sectionally uncorrelated and is independent of the common component $\lambda_i' F_t$.

The factor model makes the revealing point that stationarity of an observed series X_{it} requires stationarity of F_t (or $\lambda_i = 0$) and of e_{it}. Non-stationarity, on the other hand, could arise because of a unit root in any one of the r factors, or in e_{it}. Existing testing procedures could have difficulty in establishing whether X_{it} is I(1) or I(0) if the data admit a factor structure. This is because when F_t is $I(0)$ and e_{it} is $I(1)$, X_{it} becomes the sum of two series with different orders of integration. Thus, even though X_{it} is fundamentally I(1), univariate stationarity tests will have low power while most unit root tests will have distorted sizes when the I(0) component F_t is much larger than the I(1) idiosyncratic error. In the same manner, when one or more of the r factors F_t are I(1) and e_{it} is I(0), X_{it} is I(1) for every i. Stationarity tests will lack power when e_{it} is large, while unit root tests will reject too often.

In Bai and Ng (2001), we proposed a set of unit root tests for use in panel analysis. The present paper follows the same theme and develops panel tests for stationarity. Whereas existing tests in the literature are applied to the observed data, we test the common and idiosyncratic components separately. This can be expected to be more precise than testing the sum of two components, especially when the two pieces are integrated of different order. Removing the common components from the observed data then allows us to construct valid pooled test statistics that satisfy the cross-section independence assumption.

Since we only observe X_{it} but not F_t or e_{it}, the underlying components cannot be identified from a single time series alone. The key to our analysis lies in consistent estimation of the common and the idiosyncratic components from large dimensional panels (i.e. when N and T are both large). Loosely speaking, the large N is necessary to identify variations that are common in the cross-section, while a large T is necessary to consistently estimate terms that are idiosyncratic. As in Stock and Watson (1998), Bai and Ng (2001, 2002), we estimate λ_i and F_t by the method of principal components. Consistency of the estimator when N and T go to infinity follows from the results of Bai (2001a) assuming F_t are all I(0), and Bai (2001b) when F_t are I(1). Section 2 proposes a suite of tests for stationarity. As will become clear, the stationarity test being considered bears relation to a specific unit root test. Accordingly, Section 3 offers results for the particular panel unit root. Simulations are presented in Section 4, and tests are applied to real exchange rates in Section 5.

In the analysis to follow, we assume D_{it} is a polynomial in time of order p and present results for $p = 0$ (in which case $D_{it} = c_i$) and $p = 1$ (in which case $D_{it} = c_i + \beta_i t$). We assume the invariance principle holds so that for a series x_t $(t = 1, \ldots T)$ satisfying mixing conditions,

$$\frac{1}{\sqrt{T}\sigma_x} \sum_{s=1}^{[Tr]} x_s \Rightarrow B(r),$$

where $B(r)$ is a standard Brownian motion and σ_x^2 is the spectral density of x_t at frequency zero. If $\widetilde{x}_t = x_t - \bar{x}$, where $\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t$, then

$$\frac{1}{\sqrt{T}\sigma_x} \sum_{s=1}^{[Tr]} \widetilde{x}_t \Rightarrow B(r) - rB(1) \equiv V(r)$$

is a Brownian bridge. Furthermore, if $\breve{x}_t$ is the residual from a regression of x_t on a constant and a time trend,

$$\frac{1}{\sqrt{T}\sigma_x} \sum_{s=1}^{[Tr]} \breve{x}_t \Rightarrow B(r) - rB(1) - 6(r^2 - r)\int_0^1 (s - \frac{1}{2})dB(s) \equiv U(r)$$

is a second level Brownian bridge.

Our analysis permits some, none, or all of the factors to be non-stationary. We assume

$$F_{mt} = \alpha_m F_{mt-1} + u_{mt} \quad m = 1, \ldots k \tag{2}$$

$$e_{it} = \rho_i e_{it-1} + \epsilon_{it} \quad i = 1, \ldots N, \tag{3}$$

where ϵ_{it} and u_{mt} are iid and mutually independent. The results hold even when these errors are weakly dependent. Factor m is non-stationary if $\alpha_m = 1$. The idiosyncratic component is stationary

if $\rho_i < 1$ and has a unit root if $\rho_i = 1$. We consider the KPSS test developed in Kwiatkowski, Phillips, Schmidt and Shin (1992), the most commonly used test for stationarity. If x is the series to be tested, the KPSS test is

$$KPSS_x = \frac{\frac{1}{T}\sum_{j=1}^{T}(\frac{1}{\sqrt{T}}\sum_{j=1}^{t}x_j)^2}{\omega_x^2}$$

where ω_x^2 is a consistent kernel estimate of σ_x^2. In what follows, we propose a framework that allows for improved inference taking the properties of the univariate KPSS test as given. The proofs in the Appendix can be amended to accommodate other consistent stationarity tests of choice.

2 Panel Stationarity Tests

In this section, procedures are proposed to test

$$H_0 \quad : \quad \rho_i < 0, \quad \text{or } e_{it} \text{ is } I(0) \quad \text{for all } i.$$
$$H_1 \quad : \quad \rho_i = 1, \quad \text{or } e_{it} \text{ is } I(1) \quad \text{for some } i.$$

Since the objective is to test if the level of F_t and e_{it} are stationary, it would seem natural to obtain principal component estimates of F_t and e_{it} from (1). These estimates would, however, be consistent only under the null hypothesis that $\rho_i < 1$. Under the alternative hypothesis when the idiosyncratic errors are non-stationary, the principal components estimator applied to the non-differenced data cannot guarantee a consistent estimation of F_t. In consequence, the estimated common factors will be non-stationary even though the true factors are stationary. We therefore consider applying the principal components method to the data in first differenced form. This guarantees consistent estimation of the common factors up to a location shift and an invertible transformation. When the e_{it} are I(0), Δe_{it} are over-differenced. But over-differencing is permitted in the theory developed by Bai and Ng (2002) and Bai and Ng (2001). Estimation of the differenced model thus provides consistent estimates e_{it} and F_t under both the null and the alternative hypothesis.

Estimation of the differenced model yields estimates of Δe_{it} and ΔF_t. Our interest is in testing stationarity of e_{it} and F_t in level form. The construction of the test depends on whether or not there is a linear time trend.

2.1 The Intercept Only Case: $p = 0$

When $p = 0$, $X_{it} = c_i + \lambda_i' F_t + e_{it}$. The model in differenced form is:

$$\Delta X_{it} = \lambda_i' \Delta F_t + \Delta e_{it}. \tag{4}$$

Let ΔX be the $(T-1) \times N$ data matrix in differences such that the ith column is $(\Delta X_{i2}, \Delta X_{i3}, ..., \Delta X_{iT})'$ $(i = 1, 2, ..., N)$. Let $\Delta F = (\Delta F_2, \Delta F_3, ..., \Delta F_T)'$ and $\Lambda = (\lambda_1, ..., \lambda_N)'$. The estimated factors (in differences), $\widehat{\Delta F_{1t}}, ... \widehat{\Delta F_{kt}}$, are the k eigenvectors corresponding to the first k largest eigenvalues of the $(T-1) \times (T-1)$ matrix $\Delta X \cdot \Delta X'$. The estimated loading matrix, $\widehat{\Lambda}$, is equal to $\widehat{\Lambda} = \Delta X' \cdot \widehat{\Delta F}$. Finally, let $\widehat{\Delta e_{it}} = \Delta X_{it} - \widehat{\lambda_i'} \widehat{\Delta F_t}$ $(t = 2, ..., T, \ i = 1, 2, ..., N)$.

The steps to test stationarity of the common factors and the idiosyncratic components can be summarized as follows:

1: Estimate ΔF_t and λ_i by the method of principal components, as described above.

2: Given $\widehat{\Delta F_t}$, construct the following partial sum process for each $m = 1, ... k$,

$$\widehat{F}_{mt} = \sum_{s=2}^{t} \widehat{\Delta F_{ms}}.$$

Test the null hypothesis that $\widehat{F}_{mt}$ is stationary for each $m = 1, ... k$ using the KPSS test with demeaning. Denote this test by $S_F^c(m)$.

3: For each i, construct the partial sum $\widetilde{e}_{it} = \sum_{s=2}^{t} \widehat{\Delta e_{is}}$, $t = 2, ... T$.

(a) If F_{mt} is I(0) for every $m = 1, ... k$, for each $i = 1, ... N$, apply the KPSS test to $\{\widehat{e}_{it}^0\}_{t=1}^T$, where $\widehat{e}_{it}^0$ is $\widetilde{e}_{it}$ after demeaning.[2] Denote the test statistic by $S_{e0}^c(i)$.

(b) If $\bar{k}$ of the $F_t s$ are I(1), let $\widehat{e}_{it}^1$ be the residuals from a projection of of $\widetilde{e}_{it}$ on 1 and $\widehat{F}_{1t}, ... \widehat{F}_{kt}$. For each i, apply the test to $\{\widehat{e}_{it}^1\}_{t=1}^T$ to give $S_{e1}^c(i)$.

Theorem 1 *(p = 0) Suppose the KPSS statistic developed in Kwiatkowski et al. (1992) is used to test stationarity and assume that $N, T \to \infty$. Let V_{um} and V_{ei} $(i = 1, ... N)$, which are $N + k$ mutually independent Brownian bridges.*

1. *If F_{mt} is stationary, then*
$$S_F^c(m) \Rightarrow \int_0^1 V_{um}(r)^2 dr.$$

2. *If F_{mt} is I(0) for every m, then for each $i = 1, ... N$,*
$$S_{e0}^c(i) \Rightarrow \int_0^1 V_{ei}(r)^2 dr.$$

3. *If $\bar{k}$ of the factors are I(1), then $S_{e1}^c(i)$ has the same limiting distribution as the statistic developed in Shin (1994) for testing the null hypothesis of cointegration with $\bar{k}$ integrated regressors and a constant.*

---

[2] That is, $\widetilde{e}_{it} - \overline{\widetilde{e}}_i$ with $\overline{\widetilde{e}}_i$ being the sample mean of $\widetilde{e}_{it}$.

Bai (2001a) showed that the condition $\sqrt{T}/N \to 0$ is sufficient for consistent estimation of the factors, which will clearly be true if N is large. Since $\widehat{F}_t$ can be consistently estimated under the assumptions of the analysis, stationarity tests can treat the estimated factors as though they were known. The S_F^c test has the same distributions as derived in Kwiatkowski et al. (1992) for the constant only case. At the 5% level, the critical value is 0.463.

The limiting distributions for the test on $\widehat{e}_{it}$ depend on whether F_t is I(1) or I(0). If every F_{mt} is I(0), the test on $\widehat{e}_{it}$ has the same limit as the KPSS stationary test. At the 5% level, the critical value is also .463. If $\bar{k}$ factors are I(1), stationarity of e_{it} implies cointegration between X_i and a subset of F of dimension $\bar{k}$. Then test of the estimated idiosyncratic components has the same limiting distribution as resported in Shin (1994) developed for testing the null hypothesis of cointegration. At the 5% level, the critical values are 0.324 and .225 for $\bar{k} = 1$ and 2, respectively. In each case, the null hypothesis is rejected when the test statistic exceeds the critical value.

2.2 The Case with a Linear Trend: $p = 1$

When $p = 1$, $X_{it} = c_i + \beta_i t + \lambda_i' F_t + e_{it}$. The model in differenced form is:

$$\Delta X_{it} = \beta_i + \lambda_i' \Delta F_t + \Delta e_{it}. \tag{5}$$

Let $\widetilde{\Delta X}$ be the $(T-1) \times N$ matrix such that the ith column is the ith cross-section series (in differences) with demeaning. That is, the ith column of $\widetilde{\Delta X}$ is $(\Delta X_{i2} - \overline{\Delta X_i}, ..., \Delta X_{iT} - \overline{\Delta X_i})'$, where $\overline{\Delta X_i} = \frac{1}{T-1}\sum_{t=2}^{T} \Delta X_{it}$ $(i = 1, 2, ..., N)$. Let $\widehat{\Delta F}$ be the k eigenvectors corresponding to the k largest eigenvalues of the $(T-1) \times (T-1)$ matrix $\widetilde{\Delta X} \cdot \widetilde{\Delta X}'$ and $\widehat{\Lambda} = \widetilde{\Delta X} \cdot \widehat{\Delta F}$. Finally, define $\widehat{\Delta e_{it}} = \Delta X_{it} - \overline{\Delta X_i} - \widehat{\lambda}_i' \widehat{\Delta F_t}$.

The steps to test stationarity of the common factors and the idiosyncratic components are as follows:

1: Estimate ΔF_t and λ_i by the method of principal components, as described above.

2: Given $\widehat{\Delta F_t}$, construct the following partial sum process for each $m = 1, \ldots k$,

$$\widehat{F}_{mt} = \sum_{s=2}^{t} \widehat{\Delta F_{ms}}.$$

Test the null hypothesis that $\widehat{F}_{mt}$ is stationary for each $m = 1, \ldots k$ using the KPSS test with demeaning and detrending. Denote this test by $S_F^\tau(m)$.

3: For each i, construct the partial sum $\widetilde{e}_{it} = \sum_{s=2}^{t} \widehat{\Delta e_{is}}$, $t = 2, \ldots T$.

(a) If F_{jt} is I(0) for every $j = 1, \ldots k$, then for each $i = 1, \ldots N$, apply the KPSS test to $\{\widehat{e}_{it}^0\}_{t=1}^T$, where $\widehat{e}_{it}^0$ are the residuals from a projection of $\widetilde{e}_{it}$ on a constant and a time trend. Denote the test by $S_{e0}^\tau(i)$.

(b) If $\bar{k}$ of the $F_t s$ are I(1), let $\widehat{e}_{it}^1$ be the residuals from a projection of $\widetilde{e}_{it}$ on a constant, a time trend, and $\widehat{F}_{1t}, \ldots \widehat{F}_{\bar{k}t}$. The test statistic for the series $\{\widehat{e}_{it}^1\}_{t=1}^T$ is denoted by $S_{e1}^\tau(i)$.

Theorem 2 *(p = 1) Suppose the KPSS statistic developed in Kwiatkowski et al. (1992) is used to test stationarity and assume that N and T both approach infinity. Let U_{um} (m = 1, \ldots k) and U_{ei} be $N + k$ mutually independent second level Brownian bridges.*

1. If F_{mt} is stationary, then

$$S_F^\tau(m) \Rightarrow \int_0^1 U_{um}(r)^2 dr.$$

2. If F_{mt} is I(0) for every m, then for each $i = 1, \ldots N$,

$$S_{e0}^\tau(i) \Rightarrow \int_0^1 U_{ei}(r)^2 dr.$$

3. If $\bar{k}$ of the factors are I(1), then $S_{e1}^\tau(i)$ has the same limiting distribution as the statistic developed in Shin (1994) for testing the null hypothesis of cointegration in an equation with $\bar{k}$ integrated regressors and a time trend.

Because of the explicit estimation of intercepts and linear trends, the test statistics are invariant to coefficients on the intercepts and the linear trends. In particular, the test statistics will be the same whether or not the common factors F_t have a non-zero mean and a linear trend. That is, if $F_t = \mu + \pi t + \xi_t$, where μ and π are $k \times 1$ vector of coefficients; and ξ_t is a vector $(k \times 1)$ of zero-mean stationary processes (under the null) or a vector of non-drifting I(1) processes (under the alternative), we can simply treat ξ_t as our F_t. The limiting distribution of $S_F^\tau(m)$ coincides with that of the KPSS test derived for the linear trend case. At the 5% level, the critical value is .149. As in the case when $p = 0$, the properties of $S_e^\tau(i)$ depends on whether F_t is I(1) or I(0). Under stationarity, the limiting distribution is identical to that of $S_F^\tau(m)$ and thus also has a 5% critical value of .149. When $\bar{k}$ of the common factors are I(1), testing stationarity of e_{it} is the same as testing the null hypothesis of cointegration. As shown in Shin (1994), the limiting distribution depends on functionals of the I(1) regressors. The critical values thus depend on the rank of these regressors. For $\bar{k} = 1$ and 2, these are .122 and .100, respectively.

Pooling is valid when the limiting distribution of the test on unit i does not contain terms that are common across i. Theorems 1 and 2 show that tests on X_{it} do not have this property.

However, for testing stationarity of the idiosyncratic components, the limiting distributions are cross-sectionally independent when the F_t's are stationary or trend stationary. More precisely, test statistics based on $\widehat{e}^0$ are asymptotically independent over i. Using the same argument as in Maddala and Wu (1999) and Choi (2001), we have the following result:

Corollary 1 *Let $q(i)$ be the p-value associated with the $S_{e0}^c(i)$ test (or $S_{e0}^\tau(i)$ test). Consider pooled tests defined by $Q = -2\sum_{i=1}^N \log q(i)$. If F_{mt} is I(0) for every $m = 1, \ldots k$, then*

$$\frac{Q - 2N}{\sqrt{4N}} \Rightarrow N(0,1).$$

In contrast, I(1) factors have non-vanishing effect on the projection residuals $\widehat{e}_{it}$. Test statistics based on $\widehat{e}_{it}^1$ have limiting distributions that depend on the I(1) common factors and thus not independent over i, making pooling invalid.

3 A Panel Unit Root test

In Bai and Ng (2001), we proposed a suite of test procedures which we referred to as PANIC:-panel analysis of nonstationarity of the idioysncratic and common components. Results were derived assuming the Dickey-Fuller test was used to test the null hypothesis of a unit root. But the key to PANIC is consistent estimation of F_t and λ_i, and applicability of the results is not limited to the Dickey-Fuller test. In this section, we present results for another unit root test. Specifically, consider testing for a unit root in the series $\{x_t\}$ using the statistic:

$$MSB_x = \frac{T^{-2}\sum_{t=1}^T x_{t-1}^2}{s_w^2}, \tag{6}$$

where s_w^2 is an autoregressive estimate of σ_w^2, the spectrum at frequency zero of $\{\Delta x_t\}$.[3] The MSB test is the square of the SB statistic developed in Sargan and Bhargava (1983) for iid errors. It is extended to the case of weakly dependent errors by Stock (1990), leading to the MSB (modified Sargan-Bhargava) test as defined above.

Under the null hypothesis that $\rho_i = 1$ for every i, we estimate the factor model using the first differenced data when $p = 0$, and the demeaned first-differenced data when $p = 1$. This yields $\widehat{\Delta e_{it}}$ and $\widehat{\Delta F_t}$. Cumulating these series leads to $\widetilde{e}_{it}$ and $\widehat{F}_{1t} \ldots \widehat{F}_{mt}$ as defined in Section 2. Now for each

---

[3]The autoregressive estimate of the spectrum is $s_w^2 = \widehat{\sigma}_{wk}^2/(1 - \sum_{j=1}^k \widehat{b}_j)^2$, where $\widehat{b}_j$ and $\widehat{\sigma}_{wk}^2$ are obtained from the regression

$$\Delta x_t = b_0 x_{t-1} + b_1 \Delta x_{t-1} + \ldots + b_k \Delta x_{t-k} + w_{tk}$$

with $\widehat{\sigma}_{wk}^2 = \frac{1}{T}\sum_{t=k+1}^T \widehat{w}_t^2$. The estimator s_ω^2 is consistent for σ_ω^2 under the null hypothesis of a unit root and bounded under the alternative. As discussed in Perron and Ng (1998), this is required for the class of MSB to be consistent. The test is a member of a class of tests analyzed in Perron and Ng (1996).

$i = 1, \ldots N$ and $m = 1, \ldots k$, apply the MSB test to $\tilde{e}_{it}$ and $\hat{F}_{mt}$. Denote these tests by $M_e^c(i)$ and $M_F^c(m)$ when $p = 0$, and by $M_e^\tau(i)$ and $M_F^\tau(m)$ when $p = 1$.

Theorem 3 *Suppose the MSB statistic is used to test the unit root null hypothesis. Suppose also that $N, T \to \infty$.*

1. *When $p = 0$,*

$$M_F^c(m) \;\Rightarrow\; \int_0^1 B_{um}(r)^2 dr \tag{7}$$

$$M_e^c(i) \;\Rightarrow\; \int_0^1 B_{\epsilon i}^2(r) dr. \tag{8}$$

2. *When $p = 1$,*

$$M_F^\tau(m) \;\Rightarrow\; \int_0^1 V_{um}(r)^2 dr \tag{9}$$

$$M_e^\tau(i) \;\Rightarrow\; \int_0^1 V_{\epsilon i}^2(r) dr. \tag{10}$$

The results stated in (8) and (10) hold whether F_t is I(0) or I(1). These evidently do not depend on F_t, and thus the test statistics are poolable across i.

Corollary 2 *For $(i = 1, 2, ..., N)$, let $q(i)$ be the p-value associated with the $M_e^{c,\tau}(i)$ test using $\{\tilde{e}_{it}\}_{t=1}^T$. Define $Q = -2\sum_{i=1}^N \log q(i)$. Under the null hypothesis that $\rho_i = 1 \ \forall i$*

$$\frac{Q - 2N}{\sqrt{4N}} \Rightarrow N(0, 1).$$

Examination of the results reveals that the limiting distribution of $M_e^\tau(i)$ (corresponding to $p = 1$) is the same as $S_{e0}^c(i)$ (corresponding to $p = 0$). All the distributions derived in the preceding theorems belong to the family of generalized Cramér-von Mises distributions. As Harvey (2001) pointed out, unit root and stationarity tests with such limiting distributions can be studied in a unified framework. Whereas inference about a unit root is based on the upper tail of a Cramér-von Mises distribution, stationarity tests are based on the lower tail.

Nonetheless, the $KPSS$ and MSB tests in the present context are fundamentally different in three ways. First, stationarity tests for e_{it} depend on whether the common factors are I(1) or I(0), and in practice, pretesting of F_t will be necessary. However, a unit root test on the idiosyncratic errors is invariant to the properties of F_t and is thus immuned to inference problems that might arise in pretests. Second, the stationarity test is based on explicit detrending of $\tilde{e}_{it}$ according to whether p is 0 or 1. In contrast, the unit root test is based $\tilde{e}_{it}$ detrended according to the first-differenced model. The deterministic terms will likely have a larger effect on the stationarity than

9

the unit root test. Third, the stationarity test is based on the partial sum of the series, while the unit root test is based on the level of the series itself. Errors from estimation of the factors can be expected to have a larger impact on the stationarity test.

3.1 Monte Carlo Simulations

Data are generated according to (1)-(3) with a single common factor ($k = 1$). In addition, λ_i are i.i.d. $N(1, 1)$, ϵ_{it} are i.i.d. $N(0, 1)$, and $u_t \sim N(0, \sigma_F^2)$. Let α be the autoregressive parameter in the common factor process F_t and let ρ be the (common) autoregressive parameter in the idiosyncratic error processes e_{it}. The following parameters are considered:

- σ_F^2=10, 1, and .5.
- (ρ, α)={(.5,.8),(.8,.5),(0,.9),(.9,0),(1,0),(1,.5),(1,.8),
  (1,.9),(1,.95),(0,1), (.5,1),(.8,1),(.9,1),(.95,1)}.

Because the factor model is estimated in differenced form, the tests are invariant to the value of c_i in (1) and thus is set to zero. Similarly, when the differenced data are demeaned when $p = 1$, the tests are also invariant to β_i in (1) and thus also set to zero. We report results for $T = 200$ and $N = 20$.[4] Asymptotic critical values at the 5% significance level are used. These are obtained by first approximating the standard Brownian motion as the partial sum of 500 N(0,1) errors. The critical values for the individual tests are the percentiles from 10,000 simulations of the limiting distributions reported in Theorems 1 and 2. The pooled tests depend on the p values associated with either the stationarity or the unit root test. Approximate p-values are obtained by creating a look-up table that contains 300 percentage points of the asymptotic distributions simulated earlier. In particular, 100 points is used to approximate the upper tail, 100 to approximate the lower tail, and 100 points for the middle part of the asymptotic distributions. The p values match up very well with Table 3 of MacKinnon (1994), whenever they are available. These look-up tables are available from the authors.

Tables 1 and 2 report the rejection rates over 1000 replications. The column labeled $\widehat{F}$ is the rejection rate of the tests applied to the estimated common factor. The columns labeled X and e are the average rejection rates, where the average is taken across N units over 1000 trials. Results for a particular i are similar.

We first report in Table 1a results for the modified Sargan-Bhargava (MSB) unit root test. These rejection rates represent size in one of three cases: (i) when $\widehat{F}_t$ is tested and $\alpha = 1$, or (ii) when $\widehat{e}_t$ is tested and $\rho = 1$, or (iii) when X is tested and either $\alpha = 1$ or $\rho = 1$. Other entries

---

[4]The results change little for larger N. As expected, power is higher when T is large.

represent power.[5] The first thing to note is that the results for $p = 0$ are similar to those for $p = 1$. When both F and e are stationary, the MSB test has more power when applied to the data X directly, as indicated by the first five rows of Table 1a. But when F is nonstationary (implying X is nonstationary but its first difference has a negative moving average component), the MSB test on X is oversized. However, separate tests on F and e are much more accurate. As shown in rows with $\alpha = 1$, the rejection rates on F are close to the nominal size of 5%, while the test also has power in rejecting a unit root in F. Similarly, when only e is I(1), the test also has good size and power. The results thus show that testing the components separately is more precise than testing the sum of two series, even when the components have to be estimated from cross sections with only 20 units.

Table 1b reports the rejection rates for the pooled unit root test. The entries are given size and power interpretation as described in the previous paragraph. Pooled tests based on X are invalid because of cross-sectional dependence. Only pooling on $\widehat{e}$ is permitted by our theory. This is confirmed by the simulations. Consistent with the findings of O'Connell (1998), the pooled test applied to X rejects the unit root hypothesis too often. Size distortions are significantly smaller when tests based on $\widehat{e}$ are pooled. A motivation for considering pooled tests is higher power. Indeed, the power of the pooled tests on $\widehat{e}$ is remarkably higher than the univariate tests reported in Table 1a. When a linear trend is in the model, the size of the pooled tests is inflated somewhat, but have good properties overall. In all, Tables 1a and 1b show that the idiosyncratic-common decomposition is effective. More accurate univariate and powerful pooled tests can be obtained.

We now turn to the stationarity tests. As discussed in Caner and Kilian (2002) and many others, the KPSS test is extremely sensitive to the choice of the bandwidth used in kernel estimation. We experimented with a large number of bandwidths and report results for $\text{int}[12(T/100)^{1/4}]$ lags, found in many studies to yield the best results. Table 2a reports results for testing $\{X_{it}\}, \{\widehat{F}_t\}$, and $\{\widehat{e}_{it}\}$ unit by unit. These rejection rates represent power in one of three cases:- (i) when $\widehat{F}_t$ is tested and $\alpha = 1$, or (ii) when $\widehat{e}_t$ is tested and $\rho = 1$, or (iii) when X is tested and either $\alpha = 1$ or $\rho = 1$. All other entries represent size. Notably, the results are less satisfactory than the unit root tests. The statistic for testing if X is stationary should have rejection rates close to the nominal size of .05 when $\alpha < 1$ and $\rho < 1$. In simulations, this is true only when both F_t and e_{it} are iid, and the test is imprecise when the common factor or the idiosyncratic error is persistent. It rejects with probability .22 when $(\rho, \alpha) = (0, .9)$, and has a rejection rate of .18 when $(\rho, \alpha) = (.9, 0)$. The test also has size problems when applied to the components separately, though they are more precise

---

[5]The MSB test necessitates the choice of the lag length for the autoregressive spectrum. This is set to 6 for X and 2 for $\widehat{e}$. This is based on analysis in Ng and Perron (2001a) that a longer lag is necessary when there is a negative moving component, and a shorter lag should be used to preserve power otherwise.

than testing X directly when either e or F is iid. For example, when $(\rho, \alpha) = (0, .9)$, the test on $\hat{e}^0$ has a rejection rate of .08. When $(\rho, \alpha) = (.9, 0)$, the rejection rate on F is .05. In general, the KPSS test rejects stationarity too often. Thus, even though power appears high when one of the components indeed has a unit root, they are likely inflated by the size problem. As with many others, we experimented with different bandwidths and kernels but failed to find one that is robust over a large range of parameter values.

The simulation results show that the KPSS test is substantially oversized except in the uninteresting case when the common factors or idiosyncratic errors are very weakly serially correlated. Because of size distortion in the individual tests, the pooled tests become difficult to interpret. Nonetheless, Table 2 reveals some useful results that support the theoretical analysis. Our theory predicts that when $\alpha = 1$, a test on the stationarity of the idiosyncratic errors e_{it} should be based on $\hat{e}^1$, while the test based on $\hat{e}^0$ is invalid. Indeed, by examining the rows of Table 2a with $\alpha = 1$, tests based on $\hat{e}^1$ have less size distortion than those based on $\hat{e}^0$. Similarly, when $\alpha < 1$, theory suggests that $\hat{e}^0$ should be used. The first five rows in Table 2a show less size distortion when using $\hat{e}^0$ than using $\hat{e}^1$. However, when $\rho = 1$ and $\alpha < 1$, using $\hat{e}^0$ has less power than using $\hat{e}^1$. These results suggest that it would be useful in practice to pretest F, and then decide whether to use $\hat{e}^0$ or $\hat{e}^1$. It is conceivable that better size and power can be achieved.

To the extent that the panel unit root test works very well, and both stationarity and unit root tests are based on estimation of the same factor model in first differenced form, the problem is not likely due to the decomposition of the data into common and idiosyncratic components. It is nonetheless possible that the errors from estimation of the factors are magnified by the stationarity test considered, which is based on partial sums. The present analysis evidently provides no solution to the size problem with the KPSS, which has previously been documented by Caner and Kilian (2002), Hobijn, Franses and Ooms (1998), among others.[6] However, our analysis is useful in understanding size distortion arising from pooling, vis-á-vis size distortion due to the univariate test itself. As we have shown, even when the problem of cross-section dependence is solved, the panel stationarity test will still tend to over-reject stationarity.

4 Application to PPP

Under PPP, real exchange rates should be mean reverting and thus stationary. Because real exchange rates are often defined using the same base country, cross-section correlation arises almost by construction, even in the absence of global shocks. Strong cross-section correlation amounts to

---

[6]The Leybourne and McCabe (1994) statistic is also used to test stationarity but also has problems similar to the KPSS test, as pointed out by Caner and Kilian (2002).

a common factor that cannot be aggregated away. As O'Connell (1998) found, standard panel unit root tests are biased towards the alternative hypothesis and thus also suffer from size distortions. O'Connell suggests removing the cross-section correlation by a GLS transformation of the data. This requires that the common component be stationary, which need not be the case. Also, constructing a consistent $N \times N$ covariance matrix estimator is not easy when N is allowed to go to infinity. The decomposition approach of this paper offers a useful alternative. It also allows us to discern the source of nonstationarity.

Quarterly data for nominal exchange rates and the consumer price indices are obtained from the International Finance Statistics. We use data from 1974:1-1997:4 for 21 countries:- Canada, Australia, New Zealand, Austrai, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, UK, Japan, Korea, Singapore, and Thailand. The U.S. is used as the numeraire country. Since the nominal exchange rates are expressed as the national currency per US dollar, an increase in the real currency means a real depreciation for the home country vis-a-́vis the US dollar. To proceed with statistical analysis, we take logarithms of the data, which are then demeaned and standarized to have unit variance.

We begin with testing for stationarity and unit roots for the observed data X_{it} (real exchange rates). The results are reported in Table 3, the column under X. We tag a series with a '-' if the KPSS test rejects stationarity. A '+' is used for series that cannot reject a unit root. According to the column labeled X, the KPSS statistic rejects the null hypothesis of stationarity in 5 of the 21 observed series:- Canada, Australia, Ireland, Japan, and Thailand. The MSB rejects the unit root null hypothesis for all but Ireland and Japan.

Our earlier simulations show that testing the common and idiosyncratic components separately is more accurate than testing on the observed data. We then proceed to estimate the factors and the loadings using the method of principal components. The number of factors k is unknown. Bai and Ng (2002) proposed an information based procedure that can consistently estimate k. Using the penalty $(N + T) \log(N + T)/NT$, the criterion selects one factor. The factor associated with largest eigenvalue explains 58% of the variations in the data, while the second factor explains only 14% of the variations. We proceed with estimation assuming there is one common factor. The MSB test on the common factor is .053, very close to the critical value of .057, and is weak evidence for a unit root. The KPSS test for the common factor process is 0.172 with a p value of .783. We cannot reject the null hypothesis that the common component is stationary. In light of fact that the KPSS tends to over-reject stationarity, this non-rejection is rather strong evidence for stationarity.

As a first diagnostic, we first get some idea of whether the real exchange rates appear consistent with a factor structure. Columns 1 and 2 of Table 3 report the ratio of the standard deviation of the

idiosyncratic component (based on one factor) to the standard deviation of the differenced data, as well as the standard deviation of the common to the idiosyncratic component. If all variations are idiosyncratic, the first statistic should be close to one and the second should be small. The Asian countries and Canada have real exchange rate variations dominated by the idiosyncratic components, as $\frac{\text{var}(\widehat{\Delta e})}{\text{var}\Delta X}$ exceeds .9 for all these countries. But real exchange rate variations of the 14 European countries are apparently dominated by the common components. Overall, the evidence suggest that there are indeed non-trivial common variations in the data.

We then test the idiosyncratic errors assuming F_t is stationary. For the KPSS test, the relevant column is thus $\widehat{e}^0$. It suggests that 13 of the 21 series are non-stationary. The MSB test also finds 13 non-stationary series. Interestingly, unit root tests on X cannot reject the null hypothesis in all eight cases that the idiosyncratic errors cannot reject a unit root. Since a series with no tagged symbol is judged stationary by both tests, while a series with a '-' and corresponding '+' are judged non-stationary by both tests, of the 21 idiosyncratic series, the KPSS and MSB are in agreement over 13 series. Six series (New Zealand, Denmark, Norway, the U.K., Korea, and Singapore) are stationary, while 7 series (Australia, Austria, Germany, Ireland, Sweden, Switzerland, Japan) are non-stationary. At the 10% level, both tests also suggest Italy is non-stationary.

The pooled stationarity test strongly rejects the hypothesis that the observed series, or their idiosyncratic components, are stationary. In fact, the pooled statistic is the same for both X and $\widehat{e}^0$. The pooled unit root statistic on X rejects the hypothesis that all the observed series are I(1), but does not reject the hypothesis that all idiosyncratic series are I(1). This last finding is consistent with the simulation results that cross-section correlation will bias the pooled test towards rejecting the unit root null. Both pooled tests suggest that the underlying source of non-stationarity in the observed data is likely the idiosyncratic components.

5   Conclusion

When a series is the sum of two components with possibly different dynamic properties, testing whether the components are I(1) or I(0) should be more accurate than testing the series itself. The motivation of this paper is to exploit the fact that common and idiosyncratic components can be consistently estimated from a factor model. We develop procedures to test if these components satisfy the null hypothesis of stationarity. The decomposition into common and idiosyncratic components also allows us to develop pooled tests that satisfy the cross-section independence assumption. In simulations, tests on the components are indeed more accurate than testing the summed series. However, the results are less than satisfactory, especially in comparison with similar procedures developed for unit root tests. The problem can be traced to the properties of the univariate sta-

tionarity test, and is not due to the weakness of the common-idiosyncratic decomposition. We look forward to the development of new stationary tests with more robust properties.[7]

A primary interest in stationarity tests is the PPP hypothesis. We take our procedures to the data. Evidence from both panel unit root and stationarity tests suggest the presence of one common, stationarity factor. In view of the tendency of the KPSS test to over-reject the null hypothesis, this non-rejection can be seen as strong evidence for stationarity. However, the results also find that a large number of real exchange rates have non-stationary idiosyncratic components. Understanding the structural source of this non-stationarity is perhaps a promising way to understand why the evidence tends to pile up against PPP.

---

[7]Jansson (2001) suggests using covariates to improve the power of stationarity tests. This will not resolve the size problem.

Appendix

We first explain why the estimated $\widehat{F}_t$ can be treated as the true F_t process. In the literature on large dimensional factor analysis, as in Stock and Watson (1998), Bai and Ng (2002), and Bai (2001ab), it is shown that $\widehat{F}_t$ is consistent for HF_t, where H is a $k \times k$ matrix of full rank. It is clear that F_t is stationary if and only if HF_t is stationary. That is, an invertible matrix transformation does not alter its stationarity property. Furthermore, a transformation of the regressors will not alter the regression residuals. Thus, whether one uses F_t or HF_t ($t = 1, 2, ..., T$) as regressors, the same residuals will be obtained. Of course, $\widehat{F}_t$ is not exactly equal to HF_t because of estimation errors. But the estimation errors are negligible if N is large. This is due to the following lemma:

Lemma 1 *Consistency of $\widehat{F}_t$.*

- *Suppose F_t is I(0) and Assumptions A to G of Bai (2001a) hold. Then $\widehat{F}_t$ is $\sqrt{N}$ consistent if $\sqrt{N}/T \to 0$. If $\sqrt{N}/T \to \tau > 0$, $\widehat{F}_t$ is T consistent.*

- *Suppose F_t is I(1) and Assumptions A-F of Bai (2001b) hold. Then $\widehat{F}_t$ is $\sqrt{N}$ consistent if $N/T^3 \to 0$. If $N/T^3 \to \tau > 0$, $\widehat{F}_t$ is consistent at rate $T^{3/2}$.*

One actually does not need a large N relative to T. A further result, given in Bai (2001a), shows that if $\sqrt{N}/T \to 0$, then $\widehat{F}_t$ can be treated as though it were F_t in any time series regression. It is possible to give a rigorous proof for Theorems 1 and 2 that explicitly allows for estimation errors in F_t. But in view of Lemma 1, we can assume F_t is known with little loss of generality.

Proof of the Theorem 1

When $p = 0$, the model in level and first differenced forms are

$$
\begin{aligned}
X_{it} &= c_i + \lambda_i' F_t + e_{it} \\
\Delta X_{it} &= \lambda_i' \Delta F_t + \Delta e_{it}.
\end{aligned}
$$

By Lemma 1, the method of principal components applied to ΔX_{it} will give consistent estimates of ΔF_t. For large N, ΔF_t can be treated as known. Furthermore, $\widehat{\lambda}_i$ will be $\sqrt{T}$ consistent. Thus, we have

$$
\Delta X_{it} = \widehat{\lambda}_i \Delta F_t + \widehat{\Delta e_{it}}.
$$

From $\widehat{\Delta e_{it}} = \Delta e_{it} - (\widehat{\lambda}_i - \lambda_i)\Delta F_t$, the partial sum of this series is

$$
\begin{aligned}
\widetilde{e}_{it} &= \sum_{j=2}^{t} \widehat{\Delta e_{it}} \\
&= e_{it} - e_{i1} - (\widehat{\lambda}_i - \lambda_i)'(F_t - F_1).
\end{aligned}
$$

16

which depends on e_{i1} and $(\widehat{\lambda}_i - \lambda_i)'(F_t - F_1)$. Removal of these effects depend on whether F_t is I(0) or I(1). If F_t is I(0), we demean $\widetilde{e}_{it}$. Let $\bar{e}_i = \frac{1}{T}\sum_{t=2}^{T} e_{it}$, $\bar{F} = \frac{1}{T}\sum_{t=2}^{T} F_t$. Then

$$\overline{\widetilde{e}_i} = \frac{1}{T}\sum_{t=2}^{T}\widetilde{e}_{it} = \bar{e}_i - e_{i1} - (\widehat{\lambda}_i - \lambda_i)'(\bar{F} - F_1),$$

from which it follows that

$$\widehat{e}_{it}^0 = \widetilde{e}_{it} - \overline{\widetilde{e}_i} = e_{it} - \bar{e}_i - (\widehat{\lambda}_i - \lambda_i)'(F_t - \bar{F}).$$

Consider now the scaled partial sum of $\widehat{e}_{it}^0$. We have

$$\frac{1}{\sqrt{T}}\sum_{s=1}^{t}\widehat{e}_{is}^0 = \frac{1}{\sqrt{T}}\sum_{s=1}^{t}(e_{is} - \bar{e}_i) - \sqrt{T}(\widehat{\lambda}_i - \lambda_i)'\frac{1}{T}\sum_{k=1}^{t}(F_s - \bar{F})$$

$$= \frac{1}{\sqrt{T}}\sum_{s=1}^{t}(e_{is} - \bar{e}_i) + o_p(\frac{1}{\sqrt{T}}).$$

Thus if F_t is I(0),

$$\frac{1}{\sqrt{T}}\sum_{k=1}^{[Tr]}\widehat{e}_{ik}^0 \Rightarrow \sigma_i[B_{\epsilon,i}(r) - rB_{\epsilon,i}(1)] \equiv V_{\epsilon i}(r)$$

$$\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{k=1}^{t}\widehat{e}_{it}^0\right)^2 \Rightarrow \sigma_i^2\int_0^1 V_{\epsilon i}(r)^2 dr.$$

where σ_i^2 is the long-run variance of e_{it}. The limiting distribution is independent across i and can thus be pooled.

If F_t is I(1), demeaning alone is not sufficient to purge the effect of $e_{i1} + (\widehat{\lambda}_i - \lambda_i)'(F_t - F_1)$. We must project $\widetilde{e}_{it}$ on $[1\ F_t]$ to obtain new residuals $\widehat{e}_{it}^1$. Because $\widetilde{e}_{it} = e_{it} - e_{i1} - (\widehat{\lambda}_i - \lambda_i)'(F_t - F_1)$, $\widehat{e}_{it}$ are equivalent to those obtained by projecting e_{it} on the regressors. The KPSS test on such a residual process is studied in Shin (1994), where the limiting distributions are also derived. Thus the details are omitted. Finally, because the limiting distributions across i depend on the common stochastic trends F_t, they are not independent across i. This implies that these statistics cannot be pooled.

Proof of the Theorems 2

For $p = 1$, $\Delta X_{it} = \beta_i + \lambda_i'\Delta F_t + \Delta e_{it}$, and $\overline{\Delta X_i} = \frac{1}{T-1}\sum_{t=2}^{T}\Delta X_{it} = \beta_i + \lambda_i'\overline{\Delta F} + \overline{\Delta e_i}$, where $\overline{\Delta F} = \frac{1}{T-1}\sum_{t=1}^{T}\Delta F_t = \frac{F_T - F_1}{T-1}$ and $\overline{\Delta e_i} = \frac{e_{iT} - e_{i1}}{T-1}$. Thus

$$\Delta X_{it} - \overline{\Delta X_i} = \lambda_i'(\Delta F_t - \overline{\Delta F}) + \Delta e_{it} - \overline{\Delta e_i} \tag{11}$$

17

The principal components estimator based on the data $\Delta X_{it} - \overline{\Delta X_i}$ $(i = 1, 2, ..., N, t = 2, ..., T)$ will provide estimates of λ_i and $\Delta F_t - \overline{\Delta F}$, respectively. Because $\widehat{\Delta F_t}$ is root-N consistent for $\Delta F_t - \overline{\Delta F}$, when N is large relative to T, the estimation error is negligible and we can simply assume $\widehat{\Delta F_t} = \Delta F_t - \overline{\Delta F}$. This implies that

$$\widehat{F_t} = \sum_{s=2}^{t} \widehat{\Delta F_s} = \sum_{s=2}^{t} (\Delta F_s - \overline{\Delta F}) = F_t - F_1 - \frac{F_T - F_1}{T - 1}(t - 1)$$

The residual from projecting $\widehat{F_t}$ on $[1, t]$ will remove $F_1 + \frac{F_T - F_1}{T-1}(t - 1)$. This projection residual is asymptotically equivalent to the residual by projecting the true process F_t on $[1, t]$. Thus the KPSS test based on such residuals has a second level Brownian Bridge as its limiting distribution, as shown in Kwiatkowski et al. (1992). This proves part 1 of Theorem 2.

By the definition of $\widehat{\Delta e_{it}}$,

$$\Delta X_{it} - \overline{\Delta X_i} = \widehat{\lambda_i}' \widehat{\Delta F_t} + \widehat{\Delta e_{it}} \tag{12}$$

Subtracting (12) from (11) and noting $\widehat{\Delta F_t} = \Delta F_t - \overline{\Delta F}$, we have

$$\widehat{\Delta e_{it}} = \Delta e_{it} - \overline{\Delta e_i} - (\widehat{\lambda_i} - \lambda_i)(\Delta F_t - \overline{\Delta F})$$

Then $\widetilde{e}_{it} = \sum_{s=2}^{t} \widehat{\Delta e_{is}}$ is given by

$$\widetilde{e}_{it} = e_{it} - e_{i1} - \frac{(e_{iT} - e_{i1})}{T}(t - 1) - (\widehat{\lambda_i} - \lambda_i) \left[F_t - F_1 - \frac{(F_T - F_1)}{T}(t - 1) \right]. \tag{13}$$

Because $\widehat{\lambda_i} - \lambda_i = O_p(T^{-1/2})$, the last term of (13) is negligible if F_t is I(0). By projecting $\widetilde{e}_{it}$ on $[1, t]$, the projection residual will further remove the effects due to $e_{i1} + \frac{(e_{iT} - e_{i1})}{T}(t - 1)$. Thus the KPSS test based on the demeaned and detrended $\widehat{e}_{it}^0$ is asymptotically equivalent to the one based on the residual from a projection of e_{it} on $[1, t]$. Thus the limiting distribution is a second level Brownian bridge. This proves part 2 of Theorem 2.

If F_t is I(1), the last term of (13) is no longer negligible. We need to project $\widetilde{e}_{it}$ on $[1, \ t; F_{1t}, \ldots F_{\bar{k}t}]$. The projection will purge the effect of F_t, the linear trends, the term e_{i1}, as well as $(\widehat{\lambda_i} - \lambda_i)' F_1$ in (13). The resulting residual is asymptotically equal to the residual by projecting the true process e_{it} on $[1, \ t; F_{1t}, \ldots F_{\bar{k}t}]$. The limiting distribution of the KPSS test on such residuals is derived in Shin (1994).

Proof of Theorem 3. The MSB test is based on $\widetilde{e}_{it}$ in equation (13) (no further demeaning and detrending). Because $(\widehat{\lambda_i} - \lambda_i) = O_p(T^{-1/2})$, the last of term of (13) is $O_p(1)$ whether F_t is I(1) or I(0). Thus

$$\widetilde{e}_{it} = e_{it} - e_{iT} \frac{t - 1}{T} + O_p(1)$$

Under the null hypothesis that e_{it} is I(1), for $t = [Tr]$,

$$\frac{\tilde{e}_{it}}{\sqrt{T}} = \frac{e_{it}}{\sqrt{T}} - \frac{e_{iT}}{\sqrt{T}}\left(\frac{t-1}{T}\right) + \frac{1}{\sqrt{T}}O_p(1) \Rightarrow \sigma_{i,\epsilon}[B_i(r) - rB_i(1)]$$

where $\sigma_{i,\epsilon}^2$ is the long-run variance of $\Delta e_{it} = \epsilon_{it}$ and $B_i(r)$ is a Brownian motion process. It follows that

$$\frac{1}{T^2}\sum_{t=1}^{T}\tilde{e}_{it}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\tilde{e}_{it}}{\sqrt{T}}\right)^2 \Rightarrow \sigma_{i,\epsilon}^2\int_0^1 V_i(r)^2 dr$$

where $V_i(r) = B_i(r) - rB_i(1)$. Dividing the above by a consistent estimator of $\sigma_{i,\epsilon}^2$ leads to the desired result.

Table 1a: Rejection rates for the null hypothesis of a unit Root

T	N	ρ	α	$\sigma^F = \sqrt{10}$			$\sigma^F = 1$			$\sigma^F = \sqrt{.5}$		
				X	$\widehat{F}$	$\widehat{e}^0$	X	$\widehat{F}$	$\widehat{e}^0$	X	$\widehat{F}$	$\widehat{e}^0$
							$p = 0$					
200	20	0.00	0.00	0.99	0.80	0.79	0.99	0.81	0.80	0.99	0.78	0.79
200	20	0.50	0.80	0.99	0.92	0.90	1.00	0.91	0.90	1.00	0.92	0.90
200	20	0.80	0.50	1.00	0.90	0.92	0.99	0.89	0.92	1.00	0.91	0.92
200	20	0.00	0.90	0.94	0.85	0.79	0.92	0.85	0.79	0.92	0.87	0.79
200	20	0.90	0.00	0.90	0.78	0.85	0.92	0.78	0.85	0.92	0.78	0.85
200	20	1.00	0.00	0.07	0.75	0.06	0.06	0.47	0.06	0.06	0.34	0.06
200	20	1.00	0.50	0.21	0.88	0.06	0.08	0.67	0.06	0.07	0.55	0.06
200	20	1.00	0.80	0.38	0.90	0.05	0.13	0.75	0.06	0.10	0.61	0.06
200	20	1.00	0.90	0.43	0.86	0.06	0.16	0.72	0.06	0.12	0.61	0.06
200	20	1.00	0.95	0.34	0.55	0.06	0.15	0.46	0.06	0.12	0.42	0.06
200	20	0.00	1.00	0.09	0.06	0.66	0.08	0.04	0.65	0.10	0.05	0.64
200	20	0.50	1.00	0.08	0.05	0.84	0.13	0.05	0.84	0.16	0.07	0.83
200	20	0.80	1.00	0.12	0.06	0.88	0.20	0.06	0.88	0.24	0.06	0.87
200	20	0.90	1.00	0.11	0.06	0.82	0.22	0.06	0.82	0.27	0.06	0.82
200	20	0.95	1.00	0.12	0.08	0.57	0.19	0.06	0.57	0.24	0.06	0.57
200	20	1.00	1.00	0.07	0.06	0.06	0.07	0.06	0.06	0.07	0.05	0.06
							$p = 1$					
200	20	0.00	0.00	0.96	0.74	0.74	0.95	0.76	0.74	0.96	0.73	0.75
200	20	0.50	0.80	0.95	0.85	0.90	0.96	0.85	0.90	0.97	0.86	0.89
200	20	0.80	0.50	0.98	0.90	0.85	0.96	0.89	0.85	0.96	0.90	0.85
200	20	0.00	0.90	0.77	0.61	0.73	0.73	0.60	0.73	0.73	0.62	0.73
200	20	0.90	0.00	0.70	0.77	0.61	0.72	0.71	0.61	0.74	0.69	0.62
200	20	1.00	0.00	0.03	0.69	0.06	0.02	0.49	0.06	0.03	0.34	0.06
200	20	1.00	0.50	0.12	0.87	0.06	0.03	0.74	0.06	0.03	0.62	0.06
200	20	1.00	0.80	0.26	0.84	0.06	0.06	0.75	0.06	0.05	0.65	0.06
200	20	1.00	0.90	0.28	0.63	0.06	0.08	0.55	0.06	0.05	0.51	0.06
200	20	1.00	0.95	0.18	0.27	0.06	0.07	0.26	0.06	0.06	0.26	0.06
200	20	0.00	1.00	0.03	0.06	0.64	0.04	0.06	0.63	0.06	0.06	0.63
200	20	0.50	1.00	0.04	0.06	0.85	0.08	0.06	0.86	0.10	0.06	0.85
200	20	0.80	1.00	0.05	0.05	0.83	0.13	0.05	0.83	0.16	0.06	0.82
200	20	0.90	1.00	0.05	0.05	0.60	0.14	0.06	0.60	0.17	0.06	0.60
200	20	0.95	1.00	0.06	0.06	0.29	0.10	0.07	0.29	0.13	0.05	0.28
200	20	1.00	1.00	0.03	0.07	0.06	0.03	0.05	0.06	0.03	0.06	0.06

Table 1b: Pooled Tests: Rejection rates for the null hypothesis of a unit root

T	N	ρ	α	$\sigma^F = \sqrt{10}$		$\sigma^F = 1$		$\sigma^F = \sqrt{.5}$	
				X	$\widehat{e}^0$	X	$\widehat{e}^0$	X	$\widehat{e}^0$
				$p = 0$					
200	20	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.50	0.80	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.80	0.50	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.00	0.90	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.90	0.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.00	0.28	0.08	0.13	0.09	0.13	0.10
200	20	1.00	0.50	0.86	0.09	0.23	0.09	0.18	0.09
200	20	1.00	0.80	0.99	0.08	0.56	0.09	0.37	0.10
200	20	1.00	0.90	0.99	0.10	0.70	0.08	0.49	0.08
200	20	1.00	0.95	0.95	0.07	0.68	0.09	0.54	0.08
200	20	0.00	1.00	0.29	1.00	0.33	1.00	0.37	1.00
200	20	0.50	1.00	0.33	1.00	0.44	1.00	0.51	1.00
200	20	0.80	1.00	0.36	1.00	0.58	1.00	0.66	1.00
200	20	0.90	1.00	0.39	1.00	0.62	1.00	0.72	1.00
200	20	0.95	1.00	0.39	1.00	0.58	1.00	0.72	1.00
200	20	1.00	1.00	0.30	0.10	0.24	0.09	0.22	0.09
				$p = 1$					
200	20	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.50	0.80	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.80	0.50	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.00	0.90	0.98	1.00	1.00	1.00	0.99	1.00
200	20	0.90	0.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.00	0.02	0.11	0.00	0.13	0.00	0.13
200	20	1.00	0.50	0.38	0.11	0.01	0.13	0.00	0.11
200	20	1.00	0.80	0.81	0.12	0.08	0.13	0.02	0.12
200	20	1.00	0.90	0.80	0.12	0.17	0.11	0.05	0.13
200	20	1.00	0.95	0.56	0.13	0.15	0.11	0.06	0.13
200	20	0.00	1.00	0.13	1.00	0.14	1.00	0.16	1.00
200	20	0.50	1.00	0.13	1.00	0.22	1.00	0.26	1.00
200	20	0.80	1.00	0.17	1.00	0.33	1.00	0.44	1.00
200	20	0.90	1.00	0.15	1.00	0.36	1.00	0.46	1.00
200	20	0.95	1.00	0.17	1.00	0.30	1.00	0.39	1.00
200	20	1.00	1.00	0.09	0.12	0.03	0.13	0.01	0.12

Table 2a: Rejection rates for the null hypothesis of stationarity

				$\sigma^F = \sqrt{10}$				$\sigma^F = 1$				$\sigma^F = \sqrt{.5}$			
T	N	ρ	α	X	$\widehat{F}$	$\widehat{e}^0$	$\widehat{e}^1$	X	$\widehat{F}$	$\widehat{e}^0$	e^1	X	$\widehat{F}$	$\widehat{e}^0$	$\widehat{e}^1$
							$p = 0$								
200	20	0.00	0.00	0.05	0.06	0.05	0.12	0.05	0.05	0.04	0.12	0.04	0.03	0.05	0.12
200	20	0.50	0.80	0.12	0.13	0.06	0.14	0.10	0.11	0.06	0.13	0.09	0.12	0.05	0.13
200	20	0.80	0.50	0.08	0.06	0.11	0.23	0.10	0.06	0.11	0.23	0.10	0.06	0.11	0.23
200	20	0.00	0.90	0.22	0.23	0.08	0.11	0.20	0.22	0.08	0.11	0.20	0.23	0.09	0.11
200	20	0.90	0.00	0.18	0.05	0.23	0.38	0.22	0.14	0.24	0.39	0.23	0.15	0.23	0.38
200	20	1.00	0.00	0.76	0.51	0.79	0.89	0.79	0.75	0.79	0.84	0.79	0.77	0.79	0.81
200	20	1.00	0.50	0.71	0.26	0.79	0.89	0.78	0.65	0.79	0.85	0.79	0.71	0.79	0.83
200	20	1.00	0.80	0.61	0.18	0.80	0.89	0.75	0.47	0.79	0.86	0.77	0.57	0.79	0.84
200	20	1.00	0.90	0.57	0.28	0.79	0.88	0.74	0.41	0.79	0.87	0.76	0.48	0.79	0.85
200	20	1.00	0.95	0.58	0.42	0.80	0.86	0.72	0.47	0.79	0.85	0.75	0.52	0.79	0.84
200	20	0.00	1.00	0.79	0.80	0.51	0.05	0.77	0.80	0.52	0.05	0.76	0.80	0.52	0.05
200	20	0.50	1.00	0.78	0.79	0.33	0.06	0.76	0.81	0.33	0.06	0.72	0.78	0.32	0.06
200	20	0.80	1.00	0.79	0.81	0.22	0.10	0.72	0.81	0.22	0.11	0.69	0.80	0.23	0.11
200	20	0.90	1.00	0.77	0.80	0.28	0.20	0.68	0.78	0.27	0.21	0.66	0.79	0.28	0.21
200	20	0.95	1.00	0.74	0.77	0.41	0.35	0.69	0.78	0.41	0.35	0.66	0.78	0.41	0.36
200	20	1.00	1.00	0.78	0.79	0.79	0.72	0.80	0.80	0.79	0.70	0.80	0.79	0.80	0.70
							$p = 1$								
200	20	0.00	0.00	0.04	0.04	0.04	0.10	0.05	0.06	0.04	0.10	0.04	0.04	0.04	0.10
200	20	0.50	0.80	0.13	0.14	0.05	0.11	0.10	0.11	0.05	0.10	0.08	0.11	0.05	0.10
200	20	0.80	0.50	0.08	0.05	0.13	0.23	0.11	0.06	0.13	0.22	0.11	0.06	0.13	0.22
200	20	0.00	0.90	0.27	0.28	0.08	0.09	0.26	0.29	0.09	0.09	0.25	0.29	0.09	0.09
200	20	0.90	0.00	0.21	0.05	0.28	0.43	0.27	0.13	0.29	0.43	0.28	0.17	0.29	0.43
200	20	1.00	0.00	0.70	0.25	0.77	0.85	0.76	0.61	0.77	0.83	0.76	0.69	0.77	0.82
200	20	1.00	0.50	0.61	0.14	0.77	0.85	0.74	0.46	0.77	0.84	0.75	0.55	0.76	0.82
200	20	1.00	0.80	0.48	0.16	0.77	0.85	0.68	0.30	0.76	0.83	0.72	0.39	0.77	0.83
200	20	1.00	0.90	0.49	0.28	0.76	0.83	0.66	0.35	0.76	0.82	0.70	0.41	0.77	0.82
200	20	1.00	0.95	0.60	0.53	0.77	0.81	0.69	0.53	0.76	0.79	0.72	0.55	0.77	0.80
200	20	0.00	1.00	0.73	0.75	0.35	0.06	0.72	0.77	0.36	0.05	0.69	0.75	0.36	0.06
200	20	0.50	1.00	0.75	0.77	0.18	0.06	0.68	0.76	0.18	0.06	0.69	0.79	0.19	0.06
200	20	0.80	1.00	0.72	0.76	0.16	0.14	0.66	0.78	0.17	0.15	0.61	0.78	0.17	0.14
200	20	0.90	1.00	0.71	0.76	0.29	0.29	0.63	0.77	0.29	0.29	0.58	0.76	0.30	0.30
200	20	0.95	1.00	0.72	0.76	0.50	0.49	0.66	0.75	0.50	0.48	0.64	0.77	0.50	0.49
200	20	1.00	1.00	0.76	0.75	0.77	0.73	0.77	0.78	0.77	0.73	0.77	0.76	0.76	0.74

Table 2b: Pooled Tests: Rejection rates for the null hypothesis of stationarity

T	N	ρ	α	$\sigma^F = \sqrt{10}$			$\sigma^F = 1$			$\sigma^F = \sqrt{.5}$		
				X	$\widehat{e}^0$	$\widehat{e}^1$	X	$\widehat{e}^0$	$\widehat{e}^1$	X	$\widehat{e}^0$	$\widehat{e}^1$
							$p = 0$					
200	20	0.00	0.00	0.26	0.07	0.75	0.19	0.05	0.75	0.16	0.07	0.77
200	20	0.50	0.80	0.43	0.16	0.80	0.37	0.15	0.78	0.37	0.13	0.77
200	20	0.80	0.50	0.30	0.61	0.99	0.48	0.65	0.99	0.55	0.65	0.99
200	20	0.00	0.90	0.60	0.32	0.65	0.59	0.33	0.69	0.60	0.33	0.67
200	20	0.90	0.00	0.90	0.98	1.00	0.98	0.99	1.00	0.99	0.99	1.00
200	20	1.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.00	1.00	0.96	0.91	0.24	0.96	0.91	0.23	0.95	0.92	0.24
200	20	0.50	1.00	0.96	0.81	0.29	0.96	0.82	0.28	0.95	0.82	0.29
200	20	0.80	1.00	0.96	0.88	0.57	0.96	0.87	0.58	0.97	0.89	0.57
200	20	0.90	1.00	0.96	1.00	0.87	0.98	0.99	0.87	0.99	1.00	0.88
200	20	0.95	1.00	0.97	1.00	0.99	1.00	1.00	0.99	1.00	1.00	0.99
200	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
							$p = 1$					
200	20	0.00	0.00	0.29	0.12	0.71	0.24	0.12	0.71	0.20	0.13	0.73
200	20	0.50	0.80	0.48	0.21	0.75	0.47	0.20	0.72	0.44	0.19	0.71
200	20	0.80	0.50	0.40	0.87	0.99	0.67	0.85	1.00	0.75	0.85	1.00
200	20	0.00	0.90	0.70	0.44	0.61	0.69	0.46	0.61	0.70	0.47	0.61
200	20	0.90	0.00	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.80	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	0.00	1.00	0.95	0.88	0.33	0.95	0.88	0.31	0.95	0.88	0.35
200	20	0.50	1.00	0.96	0.75	0.38	0.95	0.72	0.36	0.96	0.77	0.38
200	20	0.80	1.00	0.96	0.93	0.83	0.97	0.94	0.84	0.97	0.93	0.83
200	20	0.90	1.00	0.98	1.00	1.00	0.99	1.00	0.99	1.00	1.00	1.00
200	20	0.95	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3: Application to Real Exchange Rates

Country	$\frac{\mathrm{var}(\widehat{\Delta e})}{\mathrm{var}(\Delta X)}$	$\frac{\sigma(\widehat{\lambda_i'}\widehat{F_t})}{\sigma(\widehat{e^0})}$	KPSS		SHIN	MSB	
			X	$\widehat{e}^0$	$\widehat{e}^1$	X	$\widehat{e}$
CANADA	0.995	0.044	0.476^-	0.475^-	0.522^-	0.030	0.141^+
AUSTRALIA	0.891	0.357	0.605^-	0.794^-	0.837^-	0.041	0.181^+
NEW ZEALAND	0.724	0.844	0.233	0.224	0.161	0.007	0.037
AUSTRIA	0.081	3.378	0.308	0.691^-	0.677^-	0.038	0.213^+
BELGIUM	0.093	2.330	0.155	0.639^-	0.715^-	0.033	0.131^+
DENMARK	0.092	5.001	0.190	0.177	0.110	0.030	0.057
FINLAND	0.316	1.364	0.098	0.149	0.156	0.018	0.107^+
FRANCE	0.113	3.920	0.140	0.633^-	0.651^-	0.037	0.566^+
GERMANY	0.086	2.972	0.167	0.580^-	0.565^-	0.035	0.275^+
IRELAND	0.158	1.761	0.515^-	0.531^-	0.768^-	0.061^+	0.119^+
ITALY	0.337	1.614	0.281	0.374	0.367^-	0.039	0.173^+
NETHERLANDS	0.076	2.871	0.142	0.723^-	0.696^-	0.029	0.054
NORWAY	0.173	3.438	0.121	0.451	0.402^-	0.047	0.015
SPAIN	0.385	1.264	0.296	0.496^-	0.622^-	0.031	0.165^+
SWEDEN	0.347	1.387	0.201	0.724^-	0.845^-	0.036	0.267^+
SWITZERLAND	0.219	2.142	0.407	0.706^-	0.751^-	0.031	0.511^+
UK	0.415	1.312	0.246	0.262	0.270	0.024	0.029
JAPAN	0.560	0.594	0.837^-	0.992^-	0.968^-	0.089^+	0.567^+
KOREA	0.988	0.187	0.117	0.101	0.072	0.001	0.000
SINGAPORE	0.548	0.618	0.276	0.258	0.272	0.035	0.038
THAILAND	0.938	0.278	0.797^-	0.897^-	0.881^-	0.021	0.037
5% CV			.463	.463	.320	.057	.057
10% CV			.343	.343	.235	.076	.076
Pooled			3.129	3.129	10.123	4.259	.341

The stationarity tests are based on 12 lags of the Parzen kernel. The unit root test is based on 4 lags in estimation of the autoregressive spectral density.

A '-' denotes rejection of stationarity, and a '+' indicates non-rejection of the unit root hypothesis.

References

Bai, J. and Ng, S. (2001), A PANIC Attack on Unit Roots and Cointegration, mimeo, Boston College.

Bai, J. and Ng, S. (2002), Determining the Number of Factors in Approximate Factor Models, *forthcoming in Econometrica*.

Bai, J. S. (2001a), Inference on Factor Models of Large Dimensions, mimeo, Boston College.

Bai, J. S. (2001b), Estimating Cross-Section Common Stochastic Trends in Non-Stationary Panel Data, mimeo, Boston College.

Banerjee, A., Marcellino, M. and Osbat, C. (2001), Testing for PPP: Should we use Panel Methods?, mimeo, European University Institute.

Caner, M. and Kilian, L. (2002), Size Distortions of Tests of Null Hypothesis of Stationarity: Evidence and Implications for the PPP Debate, *Journal of International Money and Finance* p. forthcoming.

Choi, I. (2001), Unit Root Tests for Panel Data, *Journal of International Money and Finance* **20**, 249–272.

Culver, S. and Papell, D. (1999), Long Run Purchasing Power Parity with Short Run Data: Evidence with a Null Hypothesis of Stationarity, *Journal of International Money and Finance* **18:5**, 751–768.

Hardi, K. (2000), Testing for Stationarity in Heterogeneous Panel Data, *Econometrics Journal* **3**, 148–161.

Harvey, A. C. (2001), A Unified Approach to Testing for Stationarity and Unit Roots, mimeo, University of Cambridge.

Hobijn, B., Franses, P. and Ooms, M. (1998), Generaltions of the KPSS Test for Stationarity, Econometric Institute Report 9802, Erasmus Unversity.

Im, K., Pesaran, M. and Shin, Y. (1997), Testing for Unit Roots in Heterogeneous Panels, mimeo, Cambridge University.

Jansson, M. (2001), Stationarity Testing with Covariates, mimeo, U.C. Berkeley.

Kuo, B. and Mikkola, A. (1999), Re-examining Long Run Purchasing Power Parity, *Journal of International Money and Finance* **18:2**, 251–266.

Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, Y. (1992), (Testing the Null Hypothesis of Stationarity Against the Alternative Of a Unit Root, *Journal of Econometrics* **54**, 159–178.

Levin, A. and Lin, C. F. (1993), Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. UCSD Discussion Paper 93-56.

Leybourne, S. and McCabe, P. (1994), A Consistent Test for a Unit Root, *Journal of Business and Economic Statistics* **12:2**, 157–166.

MacKinnon, J. (1994), Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests, *Journal of Business and Economic Statistics* **12**, 167–177.

Maddala, G. S. and Wu, S. (1999), A Comparative Study of Unit Root Tests with Panel data and a New Simple Test, *Oxford Bulletin of Economics and Statistics* **Special Issue**, 631–652.

Ng, S. and Perron, P. (2001a), Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power, *Econometrica, forthcoming.*

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

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

Pedroni, P. (1995), Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis: New Results, Indiana University Working Paper 95-013.

Perron, P. and Ng, S. (1996), Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties, *Review of Economic Studies* **63**, 435–465.

Perron, P. and Ng, S. (1998), An Autoregressive Spectral Density Estimator at Frequency Zero for Nonstationarity Tests, *Econometric Theory* **14**, 560–603.

Quah, D. (1994), Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Panels, *Economics Letters* **44**, 1–9.

Sargan, J. D. and Bhargava, A. (1983), Testing for Residuals from Least Squares Regression Being Generated by Gaussian Random Walk, *Econometrica* **51**, 153–174.

Shin, Y. (1994), A Residuals Based Test for the Null of Cointegration Against the Alternative of No Cointegration, *Econometric Theory* **10:1**, 91–115.

Stock, J. H. (1990), A Class of Tests for Integration and Cointegration, manuscript, Harvard University.

Stock, J. H. and Watson, M. W. (1998), Diffusion Indexes, NBER Working Paper 6702.