Panel Stationarity Tests for Purchasing Power Parity with Cross-sectional Dependence

David Harris
Department of Economics
University of Melbourne

Stephen Leybourne[1]
School of Economics
University of Nottingham

Brendan McCabe
School of Management
University of Liverpool

30 August 2004

We investigate the purchasing power parity hypothesis for a group of 17 countries using a new panel based test of stationarity that allows for arbitrary cross-sectional dependence. We treat the short run time series dynamics non-parametrically and thus avoid the need to fit separate, and potentially misspecified, models for the individual series. The statistic is simple to compute and uses standard Normal critical values, even in the presence of a wide range of deterministic components. We also show how the test can be applied using an approximate factor model for cross sectional dependence. Taken together, these features provide a generally applicable solution to the problem of testing for stationarity versus unit roots in macro-panel based data. The tests find significant evidence against the purchasing power parity hypothesis being true.

[1] Corresponding author, School of Economics, University of Nottingham, Nottingham NG7 2RD, UK. Email: steve.leybourne@nottingham.ac.uk

1 Introduction

Relatively long time series of many core macroeconomic variables are now available for the majority of developed economies. The use of panel data, and in particular unit root or stationarity tests, to empirically validate various important macroeconomic theories has become a rapid growth area of applied econometric research. For example, panel tests have been used to assess the evidence for the hypotheses of purchasing power parity (PPP), for convergence of growth rates, for mean reversion of inflation rates and for the real interest rate parity hypothesis. These tests attempt to exploit the potential power gains that are offered by analyzing a time series panel as opposed to individual series and, as such, they have the potential to provide more compelling evidence for, or against, certain models of economic behaviour. Recent tests have been proposed by, *inter alia*, O'Connell (1998), Maddala and Wu (1999), Hadri (2000), Choi (2001, 2002), Chang and Song (2002), Levin, Lin and Chu (2002), Shin and Snell (2002), Chang (2003) and Im, Pesaran and Shin (2003).

One of the major factors that any panel test needs to be able to address, if reliable inference is to be made in practical situations, is cross-sectional dependence. Cross-sectional dependencies are likely to be the rule rather than the exception in studying cross-country data due to the existence of strong inter-economy linkages. The tests of Hadri (2000), Choi (2001), Levin, Lin and Chu (2002), Shin and Snell (2002) and Im, Pesaran and Shin (2003) all assume independence across the panel and their size properties are uncertain when this rather unrealistic assumption does not hold.[2] The test of O'Connell (1998) allows for cross-sectional dependence, but this is restricted to the innovation term driving an assumed finite order AR process in the model. Choi (2002) permits cross-sectional dependence but only after imposing a common additive error component across the panel. The testing approach adopted by Chang and Song (2002) provides, at least in theory, a general treatment of the problem of cross-sectional dependence but their procedure relies on user-supplied parameters, whose values are a function of the dependence structure itself, which rather limits its practical appeal. Maddala and Wu (1999) and Chang (2003) approach the problem indirectly, relying on bootstrap procedures but the underlying tests are not pivotal. Regarding time series dynamics, with the exception of the test of Hadri (2000), all of these tests rely on fitting an appropriately specified time series regression model to each individual series in the panel (a tedious and error prone undertaking unless the cross-sectional dimension is relatively small). For tests that allow cross-sectional dependence, this is a doubly vital requirement, as any notion of these tests' robustness to cross-sectional dependence is intimately reliant on the correct modelling of the time series dynamics.

Recently, Bai and Ng (2004a,b) have suggested the use of an approximate factor model to account for cross sectional dependence in panel data. Their idea is to orthogonally decompose a panel of time series into a fixed number of independent common factors and remaining idiosyncratic components which are independent or weakly dependent. Bai and Ng (2004a,b) respectively construct Dickey-Fuller and KPSS tests for estimates of the unobserved components although they do not explicitly provide tests for the observed series. While the performance of the Dickey-Fuller approach may be deemed satisfactory, that of the KPSS procedure is much less so due to significant problems with the size of the tests.

It would seem, then, that none of the extant approaches offers a totally satisfactory solution to the problem of testing for stationarity, when the cross-sectional dependence structure and time series dynamics are both unknown. In contrast, the new stationarity test statistic we suggest in this paper is constructed so as to overcome both these problems. We allow for arbitrary unknown cross-sectional dependence between the series in the panel; the series may be contemporaneously

[2] O'Connell (1998) shows that the test of Levin, Lin and Chu (2002) can suffer severe size distortions if applied to panels where independence does not hold.

or cross-serially dependent. We also permit a wide range of heterogeneous stationary time series dynamics, which includes the conventional $ARMA$ class.

Our statistic is based on a vector version of the stationarity test of Harris, McCabe and Leybourne (2003) (HML) (rather than a KPSS-type stationarity test as in Hadri (2000)). The statistic is, in essence, the sum of the lag-k sample autocovariances across the panel, suitably studentized, where we allow k to be a simple increasing function of the time dimension. By controlling k in such a way, we remove any need to explicitly model the time series dynamics of each series in the panel, even though their time series dynamics may be quite heterogeneous. At the same time, the studentization automatically robustifies the statistic to the presence of any form of cross-sectional dependence. Our statistic is simple to construct and, conveniently, possesses a limiting null distribution which is standard normal under quite general linear process assumptions.[3] Asymptotic normality also holds when the statistic is calculated using residuals from deterministic regression models fitted to each series. These may include polynomial trends or even structural breaks and there is no requirement that the same deterministic model be fitted to each series. As such, the test can be applied across a range of empirically relevant modelling situations without reference to model-dependent null critical values, or the need to compute bootstrap critical values. We also show how our new statistic can be applied to the factor model should such a model be deemed appropriate. The test, when applied to the factor model, is for stationarity of the observed series and has significantly more power than the original when the factor model is correct. We also show how to construct a KPSS test for the observed series (as opposed to the components) in the factor model and compare its performance with the new test by means of some Monte Carlo experiments.

The plan of the paper is as follows. In the next section we introduce our statistic by explaining how it can be used to distinguish between stationarity and unit roots in the panel context. We also derive its asymptotic properties and show how to incorporate deterministic regression effects. In Section 3 we show how our test may be applied to the observed series in the context of Bai and Ng's (2004a) factor model. We also demonstrate how a KPSS test may be constructed for the observed data. Section 4 reports the results of a number of Monte Carlo experiments to gauge the empirical size and power of the tests. The results are very encouraging. In particular, the robustness of the new test's size to different patterns of cross-sectional dependence and time series dynamics stands out as a prominent characteristic. In addition, there are useful power gains available by using our new statistic in combination with Bai and Ng's (2004a) factor model when appropriate. Finally, Section 5 assesses the evidence for PPP in a panel of US Dollar real exchange rates using the new test. We include the structural breaks version of PPP, as expounded by Papell (2002), in the analysis and conclude by finding no evidence in favour of the hypothesis.

2 A Panel Test of Stationarity with Cross Sectional Correlation

Consider a panel of N time series $\{z_{i,t}, t = 1, ..., T\}$ generated by the processes

$$\begin{aligned} z_{i,t} &= \phi_i z_{i,t-1} + \varepsilon_{i,t} \\ i &= 1, 2, ..., N \ , \quad t = 1, 2, ..., T \end{aligned} \tag{1}$$

where each zero mean disturbance term $\{\varepsilon_{i,t}, t = 1, ..., T\}$ is $I(0)$ and $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$ may be correlated for any i and j. Throughout, we consider N to be fixed and we shall let T grow in our limit theory.[4]

[3] Our asymptotics are based on a fixed cross-section dimension, and passing the time series dimension to infinity. For many macroeconomic applications, the assumption of a fixed cross-section dimension would appear reasonable.

[4] It is possible to allows the number of observations to vary with the individual time series involved but we use a single T for notational convenience.

We wish to test the null hypothesis of joint stationarity

$$H_0 : |\phi_i| < 1 \text{ for all } i$$

against the unit root alternative

$$H_1 : \phi_i = 1 \text{ for at least one } i.$$

2.1 Motivation

To motivate our statistic, fix i and suppose for simplicity that $\{\varepsilon_{i,t}, t = 1, ..., T\}$ in (1) is i.i.d. with variance σ_i^2 and, following Section 2.1 of HML, consider a test statistic for the variable $z_{i,t}$ based on the scaled first order sample autocovariance $C_{i,1} = T^{-1/2} \sum_{t=2}^{T} z_{i,t} z_{i,t-1}$. Suitably centered and studentized, this is the Dickey-Fuller statistic for testing the null hypothesis $\phi_i = 1$. However, for testing $H_0 : |\phi_i| < 1$, the difficulty is that $E(C_{i,1}) \simeq T^{1/2}\sigma_i^2 \phi_i / \left(1 - \phi_i^2\right)$, so that the null distribution depends on ϕ_i. We consider instead the corresponding k^{th} order autocovariance

$$C_{i,k} = (T-k)^{-1/2} \sum_{t=k+1}^{T} z_{i,t} z_{i,t-k}.$$

If k is chosen so that $k \to \infty$ and $k/T \to 0$ as $T \to \infty$ then $E(C_{i,k}) \simeq T^{1/2}\sigma_i^2 \phi_i^k / \left(1 - \phi_i^2\right) \to 0$ as $T \to \infty$, so that the dependence of $E(C_{i,k})$ on ϕ_i disappears asymptotically. It is shown in Section 5 of HML that $C_{i,k}$ is asymptotically standard normal when suitably studentized and hence any dependence of the null distribution on ϕ_i (and σ_i^2) is removed. When $\phi_i = 1$, $E(C_{i,k}) \simeq T^{3/2}\sigma_i^2$ and it can be shown the studentized statistic is divergent. Hence, for a single $z_{i,t}$, a consistent test with standard normal asymptotic null distribution can be based on $C_{i,k}$ with $k \to \infty$.

To construct a joint test of $H_0 : |\phi_i| < 1$ *for all* i, we follow the literature on panel Dickey-Fuller tests and consider the sum of the individual test statistics

$$C_k = \sum_{i=1}^{N} C_{i,k}.$$

Under H_0, it is easily seen that

$$E[C_k] \simeq T^{1/2} \sum_{i=1}^{N} \sigma_i^2 \phi_i^k / (1 - \phi_i^2)$$

and, setting k as before, $E[C_k] \to 0$ as $T \to \infty$ since N is fixed. This eliminates the dependence of $E(C_k)$ on *all* of the ϕ_i simultaneously. In fact C_k, when suitably studentized, has an asymptotic standard normal distribution under H_0. Under H_1, suppose without loss of generality that $\phi_i = 1$ for $i = 1, \dots, sN$ for some $s \le 1$ and $|\phi_i| < 1$ for any i such that $sN < i \le N$. Then

$$E[C_k] \simeq T^{3/2} \sum_{i=1}^{sN} \sigma_i^2 + T^{1/2} \sum_{i=sN+1}^{N} \sigma_i^2 \phi_i^k / (1 - \phi_i^2),$$

so that the leading right hand side term once more suggests that the test should be consistent.

2.2 The Test Statistic and its Distribution Theory

We assume that $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, ..., \varepsilon_{N,t})'$ follows the stationary linear process assumption (Assumption LP) of HML. This assumption permits cross-sectional correlation of any form between the series in the panel; this correlation may be contemporaneous or cross-serial. In addition, it allows for heterogeneity in the dynamics and variation across the panel. The series may exhibit a range of individual temporal dependence structures, including those of stationary $ARMA$ processes.

Defining $a_{k,t} = \sum_{i=1}^{N} z_{i,t} z_{i,t-k}$, the statistic C_k can be written $C_k = T^{-1/2} \sum_{t=k+1}^{T} a_{k,t}$. The studentized version of C_k is then given by

$$S_k = \frac{C_k}{\hat{\omega}\left\{a_{k,t}\right\}},$$

where $\hat{\omega}^2\left\{a_t\right\}$ is the generic long-run variance estimator (LRV) of a sequence of variables $a_1, ..., a_T$:

$$\hat{\omega}^2\{a_t\} = \hat{\gamma}_0\{a_t\} + 2\sum_{j=1}^{l}\left(1 - \frac{j}{l+1}\right)\hat{\gamma}_j\{a_t\}, \quad \hat{\gamma}_j\{a_t\} = T^{-1}\sum_{t=j+1}^{T} a_t a_{t-j}. \tag{2}$$

Written this way, the statistic S_k can be seen to be the standardized mean of the constructed series $a_{k,t}$ divided by its long run standard deviation, which provides some intuition for the following result.

Lemma 1 *If the conditions of Theorems FCLT and LRV of HML hold and $k = O\left(T^{1/2}\right)$ then, as $T \to \infty$,*
(i) $S_k \Rightarrow N[0,1]$ under H_0,
(ii) S_k diverges to $+\infty$ under H_1.

The Lemma is a special case of Theorem 1 below and so its proof is omitted. Theorems FCLT and LRV of HML impose conditions on the choice of k and l, the truncation parameter of the LRV. In this paper, we require $k = \left\lceil (\delta T)^{1/2} \right\rceil$ for some constant $\delta > 0$ and $l/k \to 0$ as $T \to \infty$.

The role of the appropriate specification of k and $\hat{\omega}^2\left\{a_{k,t}\right\}$ is to remove the effects of the temporal dependence in individual series from the asymptotic null distribution of S_k. Since S_k depends on $z_{i,t}$ and $z_{i,t-k}$ only through the cross section sum $a_{k,t} = \sum_{i=1}^{N} z_{i,t} z_{i,t-k}$, any valid estimate of the long run variance of $\{a_{k,t}\}$ will automatically correct for any pattern of cross sectional dependence in $z_{i,t}$. Hence, there is no need to parametrically model the dynamic structure of each series or their cross-sectional dependencies. A similar LRV approach is adopted by Driscoll and Kraay (1998) to GMM inference for panel data with cross sectional dependence but without the added complication of $k \to \infty$.

2.3 Deterministic Regression Effects

The statistic S_k is generally not feasible because in practice each $z_{i,t}$ will be estimated from a regression on some deterministic terms. Such vectors of deterministics are denoted $\mathbf{x}_{i,t}$ and may be different for each i if required. In place of (1), we consider the model given by

$$\begin{aligned} y_{i,t} &= \beta_i' \mathbf{x}_{i,t} + z_{i,t}. \\ z_{i,t} &= \phi_i z_{i,t-1} + \varepsilon_{i,t} \\ i &= 1, 2, ..., N, \quad t = 1, 2, ..., T. \end{aligned} \tag{3}$$

Let $\hat{z}_{i,t}$ denote an OLS residual from the regression (3) i.e. $\hat{z}_{i,t} = y_{i,t} - \hat{\beta}_i' \mathbf{x}_{i,t}$ where $\hat{\beta}_i$ is the usual OLS estimator. It is desirable that the statistic be invariant to relative rescaling of the series, so in place of $z_{i,t}$ in the definition of S_k we will use instead the standardized residuals

$$\tilde{z}_{i,t} = \hat{z}_{i,t}/s_i, \tag{4}$$

where s_i is the sample standard deviation of $\hat{z}_{i,t}$. The resulting statistic is

$$\hat{S}_k = \frac{\tilde{C}_k}{\hat{\omega}\left\{\tilde{a}_{k,t}\right\}}$$

where $\tilde{C}_k = T^{-1/2}\sum_{t=k+1}^{T}\tilde{a}_{k,t}$ and $\tilde{a}_{k,t} = \sum_{i=1}^{N}\tilde{z}_{i,t}\tilde{z}_{i,t-k}$.

It can be shown that $\hat{S}_k$ has an asymptotic standard normal null distribution and, when dealing with a small number of series, this often proves to be an adequate approximation for the finite sample distribution. However, if the panel dimension is not relatively small, individual finite sample errors that arise from the estimation of the regression models combine in the construction of the aggregate numerator $\tilde{C}_k = \sum_{i=1}^{N}\tilde{C}_{i,k}$ and can significantly affect the finite sample null distribution of $\hat{S}_k$. To see how this arises, write the numerator of $\hat{S}_k$ as $\tilde{C}_k = \sum_{i=1}^{N}\tilde{C}_{i,k}$ where $\tilde{C}_{i,k} = T^{-1/2}\sum_{t=k+1}^{T}\tilde{z}_{i,t}\tilde{z}_{i,t-k}$. After some manipulation, $\tilde{C}_{i,k}$ may be expressed as

$$\begin{aligned}\tilde{C}_{i,k} &= \frac{1}{s_i^2}T^{-1/2}\sum_{t=k+1}^{T} z_{i,t}z_{i,t-k} \\ &\quad -T^{-1/2}\left[\frac{1}{s_i^2}T^{-1/2}\sum_{t=1}^{T} z_{i,t}\mathbf{x}_{i,t}'\left(T^{-1}\sum_{t=1}^{T}\mathbf{x}_{i,t}\mathbf{x}_{i,t}'\right)^{-1}T^{-1/2}\sum_{t=1}^{T}\mathbf{x}_{i,t}z_{i,t}\right] + o_p(T^{-1/2})\end{aligned} \tag{5}$$

The first term in this expression is asymptotically normal under the null and the effect of the regression estimation error is captured by (5). Under H_0, the term in square brackets in (5) is $O_p(1)$ and so the whole term is $O_p(T^{-1/2})$. Because the term (5) is clearly negative it induces a negative finite sample error into each individual statistic and the amplification of this problem is obvious when we subsequently consider $\tilde{C}_{i,k}$ summed over N. Since we are conducting upper tail tests, *ceteris paribus*, the effect of this is to *reduce* the finite sample size of the test[5]. It is possible to produce a finite sample correction for this regression estimation error by subtracting an estimate of the term (5) from $\tilde{C}_{i,k}$. It is shown in the proof of Theorem 1 below that the expectation of the term in square brackets in (5) can be consistently estimated by

$$\tilde{c}_i = \operatorname{tr}\left[\left(T^{-1}\sum_{t=1}^{T}\mathbf{x}_{i,t}\mathbf{x}_{i,t}'\right)^{-1}\hat{\mathbf{\Omega}}\{\mathbf{x}_{i,t}\tilde{z}_{i,t}\}\right] \tag{6}$$

where for any vector sequence $\mathbf{a}_1, ..., \mathbf{a}_T$,

$$\hat{\mathbf{\Omega}}\{\mathbf{a}_t\} = \hat{\mathbf{\Gamma}}_0\{\mathbf{a}_t\} + \sum_{j=1}^{l}\left(1-\frac{j}{l+1}\right)\left(\hat{\mathbf{\Gamma}}_j\{\mathbf{a}_t\} + \hat{\mathbf{\Gamma}}_j\{\mathbf{a}_t\}'\right), \quad \hat{\mathbf{\Gamma}}_j\{\mathbf{a}_t\} = T^{-1}\sum_{t=j+1}^{T}\mathbf{a}_t\mathbf{a}_{t-j}'. \tag{7}$$

This is the usual matrix version of the scalar long run variance $\hat{\omega}\{.\}^2$. Specializing to the case of a constant, $\mathbf{x}_{i,t} = 1$, we can compute $\tilde{c}_i = \hat{\omega}\{\tilde{z}_{it}\}^2$ and in the case of $\mathbf{x}_{i,t} = [1, w_{i,t}]'$, $\tilde{c}_i = \hat{\omega}\{\tilde{z}_{it}\}^2 + \hat{\omega}\{\tilde{z}_{it}\tilde{w}_{i,t}\}^2$ where $\tilde{w}_{i,t} = (w_{i,t} - \bar{w}_i)/s_{w,i}$ and $s_{w,i}$ is the standard deviation of the variable $w_{i,t}$.

[5]Simulation results, available from the authors, emphatically confirm this for moderate values of N and T.

To summarize, our recommended statistic for application is

$$\tilde{S}_k = \frac{\tilde{C}_k + \tilde{c}}{\hat{\omega}\left\{\tilde{a}_{k,t}\right\}} \tag{8}$$

where $\tilde{c} = (T-k)^{-1/2}\sum_{i=1}^{N}\tilde{c}_i$, $\tilde{c}_i$ is given in (6), $\tilde{C}_k = (T-k)^{-1/2}\sum_{t=k+1}^{T}\tilde{a}_{k,t}$ where $\tilde{a}_{k,t} = \sum_{i=1}^{N}\tilde{z}_{i,t}\tilde{z}_{i,t-k}$ and $\tilde{z}_{i,t}$ is obtained from (4) and $\hat{\omega}\left\{\tilde{a}_{k,t}\right\}$ is obtained from (2). The proof of the following theorem is given in the Appendix.

Theorem 1 *If the conditions of Theorem RES of HML hold and $k = (\delta T)^{1/2}$ (for some constant $\delta > 0$) then as $T \to \infty$*
(i) $\tilde{S}_k \Rightarrow N(0,1)$ under H_0,
(ii) $\tilde{S}_k$ diverges to $+\infty$ under H_1.

Theorem RES of HML specifies quite general conditions on $\mathbf{x}_{i,t}$, allowing for a wide range of deterministic regression functions including constants, linear and polynomial trends, dummy variables, structural breaks and various other models. The theorem shows that a consistent test is obtained by rejecting H_0 for values of $\tilde{S}_k$ greater than the appropriate upper tail critical value from the standard normal distribution. We implement the test in Sections 4 and 5 using $l = \left\lceil 12\,(T/100)^{1/4}\right\rceil$ in $\hat{\omega}\left\{\tilde{a}_{k,t}\right\}$ and $k = \left\lceil (3T)^{1/2}\right\rceil$.

3 Panel Tests for Stationarity in Factor Models

An alternative approach to panel stationarity testing in the presence of cross sectional correlation is based on the factor model of Bai and Ng (2004a,b). Instead of the nonparametric model (1), consider the factor model

$$y_{i,t} = \mu_i + z_{i,t}, \tag{9}$$

$$z_{i,t} = \boldsymbol{\lambda}_i'\mathbf{f}_t + e_{i,t}, \tag{10}$$

where μ_i is a constant term[6], $\mathbf{f}_t$ is an $r \times 1$ vector of latent factors, $\boldsymbol{\lambda}_i$ is an $r \times 1$ vector of loading parameters and $e_{i,t}$ is an idiosyncratic component for each i. It is assumed that $\mathbf{f}_t = (f_{1,t}, \ldots, f_{r,t})'$ and $e_{i,t}$ satisfy

$$f_{j,t} = \alpha_j f_{j,t-1} + u_{j,t},\ j = 1, \ldots, r, \tag{11}$$

$$e_{i,t} = \rho_i e_{i,t-1} + v_{i,t},\ i = 1, \ldots, N, \tag{12}$$

for $t = 1, \ldots, T$, where $\{u_{j,t}, t = 1, ..., T\}$ and $\{v_{i,t}, t = 1, ..., T\}$ are mutually independent $I(0)$ disturbances[7] for all i and j.

Bai and Ng (2004a) give a detailed analysis of the estimation of the unobserved components $f_{j,t}$ and $e_{i,t}$ by principal components, showing that it is necessary to carry out the principal components analysis on $\Delta z_{i,t}$ rather than $z_{i,t}$ for consistent estimation under both null and alternative hypotheses[8]. To briefly summarize their estimation method, let ΔY be the $(T-1) \times N$ matrix of observations on $\Delta y_{i,t}$ and let $\widehat{\Delta F}$ be the $(T-1) \times r$ matrix of eigenvectors corresponding to the

[6] Other deterministic components may be included, with some complications noted below.

[7] The asymptotic theory of the tests based on this model allows weak correlation among $u_{j,t}$ and $\varepsilon_{i,t}$, see Assumption C of Bai and Ng (2004a). Independence is maintained here for ease of interpretation and is imposed in the simulations below.

[8] In contrast to our asymptotic theory for the $\tilde{S}_k$ test, this consistency follows as both N and T go to infinity. Although this consistency is not required for our results, we adopt their estimation strategy as it provides a test with good finite sample properties.

largest r eigenvalues of $\Delta Y \Delta Y'$ (i.e. the r largest principal components of ΔY). The estimated idiosyncratic components $\widehat{\Delta E}$ in first difference form are the residuals from an OLS regression of ΔY on $\widehat{\Delta F}$. Letting $\widehat{\Delta f_{j,t}}$ and $\widehat{\Delta e_{i,t}}$ denote the individual elements of $\widehat{\Delta F}$ and $\widehat{\Delta E}$ respectively, the estimated components are found by partial summation, that is $\hat{f}_{j,t} = \sum_{s=2}^{t} \widehat{\Delta f_{j,s}}$ and $\hat{e}_{i,t} = \sum_{s=2}^{t} \widehat{\Delta e_{i,s}}$ for $t = 2, \ldots, T$ ($j = 1, \ldots, r$ and $i = 1, \ldots, N$). Since r is unknown in practice, it is estimated using the Bai and Ng (2002) BIC type criterion

$$\hat{r} = \arg \min_{0 \leq r \leq r_{\max}} \left\{ \log \hat{\sigma}^2(r) + r \log \left(\frac{NT}{N+T} \right) \left(\frac{N+T}{NT} \right) \right\} \tag{13}$$

where $\hat{\sigma}^2(r) = (NT)^{-1} \sum_{t=2}^{T} \sum_{i=1}^{N} \hat{e}_{i,t}^2$ to give an estimate of r. Bai and Ng (2002) prove consistency of $\hat{r}$ when $r \leq r_{\max}$ and $N, T \to \infty$.

Bai and Ng (2004a) construct unit root tests for the estimated components and Bai and Ng (2004b) provide stationarity tests based on the well known univariate stationarity test of Kwiatkowski, Phillips, Schmidt and Shin (1992, KPSS). By testing the unit root or stationarity properties of the components rather than the observable series, Bai and Ng are able to avoid the problems arising from cross sectional correlations among the series and to deduce much detailed information about the time series properties of the panel. However, in this paper we aim to find a single test for the null hypothesis that the series $\{z_{i,t}, t = 1, ..., T\}$ are stationary for all i.

The null hypothesis for the panel stationarity test based on (9)–(12) is

$$H_0 : |\alpha_j| < 1, \ |\rho_i| < 1 \text{ for all } i, j,$$

against the alternative hypothesis

$$H_1 : \alpha_j = 1 \text{ for at least one } j \text{ and/or } \rho_i = 1 \text{ for at least one } i.$$

The approach is apply a stationarity test to all of the estimated components jointly.

3.1 A Stationarity Test for the Estimated Components

The $\tilde{S}_k$ test can be calculated for the estimated components $\hat{f}_{j,t}$ and $\hat{e}_{i,t}$. Let $\tilde{f}_{j,t}$ and $\tilde{e}_{i,t}$ denote $\hat{f}_{j,t}$ and $\hat{e}_{i,t}$ each individually demeaned and standardized to have unit standard deviation. The statistic (8) can be calculated for these standardized estimated components by redefining $\tilde{z}_{i,t}$ to be the i^{th} element of the $(N + \hat{r}) \times 1$ vector $\left(\tilde{f}_{1,t}, \ldots, \tilde{f}_{\hat{r},t}, \tilde{e}_{1,t}, \ldots, \tilde{e}_{N,t} \right)'$ and the resulting statistic is denoted $\tilde{S}_k^F$. The asymptotic properties of $\tilde{S}_k^F$ follow from Theorem 2 below, which shows that a consistent test is obtained by rejecting H_0 for values of $\tilde{S}_k^F$ greater than the appropriate upper tail critical value from the standard normal distribution.

Equation (9) can be extended to a general deterministic regression of the form

$$y_{i,t} = \beta_i' \mathbf{x}_{i,t} + z_{i,t}. \tag{14}$$

This deterministic formulation extends that of Bai and Ng (2004a) who allowed only for a constant and trend. For example, $\mathbf{x}_{i,t}$ in Section 5 contains trend functions with structural breaks whose dates differ for each i. The computation of $\tilde{S}_k^F$ is outlined where $\mathbf{x}_{i,t} = \mathbf{x}_t$ for all i, while details of how to handle fully general deterministic terms are available in Appendix 6.2. The estimation of the factor model (14), (10)–(12) in first differences begins with the estimation of the OLS regressions of $\Delta y_{i,t}$ on $\Delta \mathbf{x}_t$[9] for each $i = 1, \ldots, N$ to obtain residuals $\Delta \hat{z}_{i,t}$ which are arranged in the $(T-1) \times N$ matrix $\widehat{\Delta Z}$. Equation (10) is estimated by principal components as discussed in

[9]If $\mathbf{x}_t$ contains a constant then the corresponding element of $\Delta \mathbf{x}_t$ is deleted.

the previous section, with ΔY replaced by $\widehat{\Delta Z}$. The resulting estimated components, $\hat{f}_{j,t}$ and $\hat{e}_{i,t}$ for $j = 1, \ldots, \hat{r}$ and $i = 1, \ldots, N$ are individually regressed on $\mathbf{x}_t$ to give residuals which are each standardized to have unit variance. Again $\tilde{z}_{i,t}$ is redefined to be the i^{th} element of the $(N + \hat{r}) \times 1$ vector $\left(\tilde{f}_{1,t}, \ldots, \tilde{f}_{\hat{r},t}, \tilde{e}_{1,t}, \ldots, \tilde{e}_{N,t}\right)'$ of standardized residuals, and the statistic (8) is calculated from this vector to give $\tilde{S}_k^F$.

Theorem 2 *Under the conditions of Theorem 1,*
(i) $\tilde{S}_k^F \Rightarrow N(0,1)$ *under* H_0*,*
(ii) $\tilde{S}_k^F$ *diverges to* $+\infty$ *under* H_1*.*

Note that the theorem is proved for general deterministic regressions as outlined in Appendix 6.2. It follows essentially because $\tilde{S}_k^F$ is computed on a linear transformation (determined by the principal components analysis) of the observed data, provided the linear transformation is appropriately de-trended. The asymptotic theory is not informative about the relative merits of $\tilde{S}_k^F$ and $\tilde{S}_k$, but the simulations in Section 4 reveal that $\tilde{S}_k^F$ has superior finite sample power in the factor model.

3.2 A KPSS Test for the Estimated Components

A point of comparison for the $\tilde{S}_k^F$ and $\tilde{S}_k$ tests is provided by a simple adaptation of the pooled KPSS test of Bai and Ng (2004b)[10]. It is also a variation of Hadri (2000) where the pooled KPSS test was applied directly to $y_{i,t}$ assuming no cross sectional correlation. For the factor model (9)–(12), let $\tilde{f}_{j,t}$ and $\tilde{e}_{i,t}$ be standardized versions of $\hat{f}_{j,t}$ and $\hat{e}_{i,t}$ obtained from the principal components estimation. The individual KPSS statistics for the estimated components are

$$\eta_{f,j} = \frac{T^{-2} \sum_{t=2}^{T} \left(\sum_{s=2}^{t} \tilde{f}_{j,s}\right)^2}{\hat{\omega}^2 \left\{\tilde{f}_{j,t}\right\}}, \quad \eta_{e,i} = \frac{T^{-2} \sum_{t=2}^{T} \left(\sum_{s=2}^{t} \tilde{e}_{i,s}\right)^2}{\hat{\omega}^2 \left\{\tilde{e}_{i,t}\right\}}$$

for $j = 1, \ldots, \hat{r}$ and $i = 1, \ldots, N$. The pooled test of $H_0 : |\alpha_j| < 1,\ |\rho_i| < 1$ for all i, j is

$$\bar{\eta}_\mu = \frac{1}{c_2 \sqrt{N + \hat{r}}} \left(\sum_{j=1}^{\hat{r}} (\eta_{f,j} - c_1) + \sum_{i=1}^{N} (\eta_{e,i} - c_1) \right) \tag{15}$$

where $c_1 = 0.167$ and $c_2 = 0.149$ are asymptotic constants whose values are taken from Hadri (2000). The μ subscript denotes the inclusion of a constant in (9). It follows from Theorem 1 of Bai and Ng (2004b) and Theorem 1 of Hadri (2000) that $\bar{\eta}_\mu$ has an asymptotic standard normal null distribution as N and T approach infinity. Under the alternative, those individual statistics corresponding to nonstationary components diverge to $+\infty$, so the pooled test is defined to reject the null when $\bar{\eta}_\mu$ is larger than the appropriate upper tail critical value from the standard normal distribution.

The consistency of the test when both N and T approach infinity requires consideration. If an individual component ($f_{j,t}$ or $e_{i,t}$) is $I(1)$ then its corresponding statistic diverges to $+\infty$ at rate[11] $O_p(T/l)$ where l is the bandwidth parameter in the long run variance $\hat{\omega}^2\{.\}$. If any of

[10] Bai and Ng (2004b) provide a test of the stationarity of the individual factors and a separate pooled test for the idiosyncratic components, but testing the components separately obviously introduces a multiple testing problem for our null.

[11] This follows from the consistency of the component estimates under the alternative shown by Bai and Ng (2004a) and then from the well known divergence rate of the KPSS statistic.

Table 1: Finite sample constants for $\bar{\eta}_\mu$ statistic

			$\hat{c}_{1,T}$					$\hat{c}_{2,T}$		
T	30	50	75	150	300	30	50	75	150	300
QS	0.312	0.235	0.207	0.184	0.175	0.182	0.105	0.108	0.121	0.130
Parzen	0.211	0.190	0.182	0.174	0.170	0.098	0.109	0.116	0.127	0.137

$f_{j,t}$ are $I(1)$ or if a fixed number of $e_{i,t}$ are $I(1)$ then $\bar{\eta}_\mu$ is of order $O_p\left(T/\left(l\sqrt{N}\right)\right)$, which thus provides a condition on T, N and l for the consistency of the test against a fixed number of $I(1)$ components. We impose $l = O\left(T^{1/4}\right)$ in this paper, so consistency against a fixed number of $I(1)$ components requires T and N to satisfy $T^{3/2}/N \to \infty$. In practical terms, a sufficiently long time series relative to the cross section dimension is required to consistently detect individual $I(1)$ components. An important motivating economic problem for the panel unit root and stationarity testing literature is the PPP hypothesis, and panel data sets for this problem typically have T considerably larger than N; for example, the data set analyzed in Section 5 below has $T = 312$ and $N = 17$. We conjecture that $\bar{\eta}_\mu$ should have non-negligible power against fixed numbers of $I(1)$ components in such panels, and we evaluate this conjecture in the simulations of Section 4. If a fixed *fraction* of the $e_{i,t}$ are $I(1)$ then $\bar{\eta}_\mu$ is of order $O_p\left(T\sqrt{N}/l\right)$, so it is clearly consistent against such an alternative and should exhibit some power advantage over univariate tests.

In common with the experience of Yin and Wu (2000), we find it necessary to replace the asymptotic constants c_1 and c_2 in $\bar{\eta}_\mu$ with finite sample estimates to obtain reasonable finite sample properties. We obtain these constants by generating 2000 samples from (10) – (12) with $r = 0$ and $\rho_i = 0$ for all i, so that $z_{i,t} \sim$ i.i.d.$N(0,1)$ for all i and t. The Parzen and Quadratic Spectral (QS) lag windows were used (following Yin and Wu (2000) and Bai and Ng (2004b) respectively). For each combination of sample sizes N and T, constants were computed from the means and standard deviations of the sampling distributions of $\bar{\eta}_\mu$. The constants have important variation with T (and also with the choice of lag window and l), but little variation over the values of N considered here. The constants, denoted $\hat{c}_{1,T}$ and $\hat{c}_{2,T}$, are reported in Table 1 because the test cannot reasonably be applied with the asymptotic constants.

In principle, the $\bar{\eta}_\mu$ test could be extended to allow general deterministic regression of the form (14). For example, if $\mathbf{x}_{i,t} = (1,t)'$ for each i then the required principal components estimation is summarized in Section 2.2 of Bai and Ng (2004b) and the statistics $\eta_{f,j}$ and $\eta_{e,i}$ use detrended versions of $\hat{f}_{j,t}$ and $\hat{e}_{i,t}$. The asymptotic constants in the resulting pooled statistic $\bar{\eta}_\tau$ are $c_1 = 0.067$ and $c_2 = 0.042$ (Hadri, 2000, equation 25), although in practice these constants would need to be replaced by finite sample ones analogous to those in Table 1. The inclusion of more general deterministics would require the computation of further constants and could become unwieldy in practical situations. For example, if structural breaks were to be included then the constants would depend on the number of breaks and the break dates. Allowing different deterministic specifications for each i (e.g. different break dates for each i as in Section 5 below) would be possible by computing constants $c_{1,T,i}$ and $c_{2,T,i}$ for standardization of individual $\eta_{e,i}$ statistics, although it is not clear how best to handle the $\eta_{f,j}$ statistics[12]. As a result of this difficulty, and also because of the inferior finite sample performance of the $\bar{\eta}_\mu$ test reported in Section 4, we do not pursue the KPSS approach in our application in Section 5.

[12] Asymptotically, the standardisation of the $\eta_{f,j}$ does not matter since $\hat{r}$ is bounded as $N \to \infty$. Obviously, there would be some finite sample effect.

3.3 Comparison of the two approaches

A theoretical difference between testing via $\tilde{S}_k$ (or $\tilde{S}_k^F$) and $\bar{\eta}_\mu$ is that $\tilde{S}_k$ and $\tilde{S}_k^F$ use an asymptotic approximation with $T \to \infty$ and N fixed, while $\bar{\eta}_\mu$ uses both T and N approaching infinity. We consider the resulting difference in interpretation of how each approach handles the cross sectional correlation.

Since $\tilde{S}_k$ holds N fixed, it may be regarded as being parametric with respect to cross sectional correlation since there are a fixed number $(N(N-1)/2)$ of correlations[13]. However, these correlations have no structure imposed upon them so the approach is otherwise completely flexible. Any lagged cross correlations between individual series are modelled nonparametrically within the multivariate linear process assumption (HML, Assumption LP) and their effects on the test statistic are implicitly catered for by the long run variance in the denominator.

Since $\bar{\eta}_\mu$ is based upon a factor model with a fixed number of factors, it may also be regarded as being parametric with respect to the cross sectional correlation, although there is also a nonparametric aspect to the approach since the factor loadings are unrestricted for each $i = 1, \ldots, N$. The requirement that $N \to \infty$ is important for the consistent estimation of the factors and idiosyncratic components in model (9). Compared to $\tilde{S}_k$ and $\tilde{S}_k^F$, we believe the parametric aspect of the $\bar{\eta}_\mu$ test imposes the more restrictive assumption because model (9) can potentially be misspecified if insufficient factors are included or if a factor model is inappropriate. Note that $\tilde{S}_k^F$ remains valid in the presence of misspecification of the factor model. We include some simulations in Section 4 that address the possible practical consequences of such a misspecification.

4 Finite Sample Properties

In this section we report the results of a simulation experiment to evaluate the finite sample properties of the three tests considered in this paper. To our knowledge there is no other existing panel test for stationarity that is valid in the presence of cross sectional correlation that can be included in the experiment. All experiments include a constant term in the deterministic specification.

For all experiments the data generating process is the factor model (9) – (12) with disturbances $(u_{1,t}, \ldots, u_{r,t}, \varepsilon_{1,t}, \ldots, \varepsilon_{N,t})'$ drawn from the $N(0, I_{N+r})$ distribution. The sample sizes are $N = 10, 20, 30, 40$ and $T = 30, 50, 75, 150, 300$, which are chosen to include realistic sample sizes for macroeconometric applications. The number of factors included are $r = 0, 2$ and 7. The estimation of the factor model is implemented using $r_{\max} = 6$ when choosing the number of factors, so the factor model can be correctly specified when $r = 0$ or 2, but is misspecified when $r = 7$. We include the $r = 7$ case to give some idea of the consequences of misspecifying the factor model, which in practice could be due to underspecification of the number of factors or to a model with a fixed number of factors being inappropriate. There is no cross sectional correlation when $r = 0$. When $r = 2, 7$ we consider $\lambda_i \sim \text{i.i.d.}N(\kappa, \kappa^2)$ (fixed across replications) with $\kappa = 3$, which is the same as used by Bai and Ng (2004b) except for the multiplicative constant κ. Since $u_{j,t}$ and $\varepsilon_{i,t}$ are standard normal variates, κ is used to control the relative standard deviations of the factors compared to the idiosyncratic components. The value of κ turns out to be unimportant for $r = 2$, but is important for $r = 7$ when the factor model is misspecified because it determines the relative magnitude of the factor[14] that is omitted. The choice of $\kappa = 3$ is within the range of empirical standard deviation ratios reported in Table 4 of Bai and Ng (2004b) for a panel of quarterly real

[13] We conjecture that the approach remains valid if $N \to \infty$ sufficiently slowly as $T \to \infty$, and hence may be classed fully nonparametric, but we defer this technical issue for other research.

[14] We find that $\hat{r} = 6$ in every replication in this case, so one factor can be considered to be omitted. Obviously finding $\hat{r} = r_{\max}$ in practice would lead to some model respecification, with at least an increase in $r_{\max}$. The point here is to evaluate the effect of a model misspecification.

exchange rates. Predictably, unreported simulations reveal that the consequences of the model specification are worse when κ is increased from 3, and better when κ is decreased.

After preliminary experimentation, we implement the $\tilde{S}_k$ and $\tilde{S}_k^F$ statistics using $k = \left\lceil (3T)^{1/2} \right\rceil$ and the long run variances are estimated using a Bartlett lag window with $l = \left\lceil 12\,(T/100)^{1/4} \right\rceil$. Following Bai and Ng (2004b), the long run variances are computed using the Quadratic Spectral[15] lag window with $l = \left\lceil 12\,(T/100)^{1/4} \right\rceil$ giving $\bar{\eta}_\mu^{QS}$.

All tests are taken to reject the null hypothesis at the 5% significance level when the test statistic is larger than 1.65. All experiments use 5000 replications.

4.1 Size

Under the null hypothesis, the values of α_i and ρ_j in (11) and (12) considered are $(\alpha_j, \rho_i) = (0, 0)$, $(0.4, 0)$, $(0.8, 0)$, $(0, 0.4)$, $(0, 0.8)$ for all i and j. Any difference in the effect of autocorrelation in the factors and idiosyncratic components can be evaluated. We also consider α_j and ρ_i drawn from independently from a $U\,[0, 0.8]$ distribution for all i and j (fixed across replications) to include some heterogeneity across the components.

Table 2(a) contains estimated finite sample sizes for $r = 0$; the case of no cross-correlation. The results for $\rho_i = 0$ reveal actual sizes near to the asymptotic 0.05 level. The only exception is some mild oversizing for the $\tilde{S}_k$ and $\tilde{S}_k^F$ tests for $T = 30$, although it is not surprising that an asymptotic approximation based on $T \to \infty$ with N fixed is not very accurate when T is small and N is about the same magnitude. Increasing ρ_i to introduce idiosyncratic autocorrelation reveals that $\tilde{S}_k$ and $\tilde{S}_k^F$ are well behaved for $\rho_i = 0.4$ and somewhat undersized for small T and $\rho_i = 0.8$. The $\bar{\eta}_\mu^{QS}$ test is undersized for $\rho_i = 0.4$ and small T and displays a worrying mixture of undersizing and oversizing when $\rho_i = 0.8$.

Table 2(b) gives the results for the cross-correlated case where $r = 2$. The $\tilde{S}_k$ and $\tilde{S}_k^F$ tests have good size properties in all cases and quickly approach correct levels for realistic macroeconomic sample sizes. The $\bar{\eta}_\mu^{QS}$ is somewhat undersized for small T in the presence of strong autocorrelation, then becomes oversized as T increases, even with $T = 300$.

Table 2(c) gives the results for $r = 7$. The properties of $\tilde{S}_k$ and $\tilde{S}_k^F$ are essentially the same as for $r = 2$, illustrating the robustness to cross correlation and factor model misspecification of these tests. The interest is in $\bar{\eta}_\mu^{QS}$ since it relies on the correct specification of the model — across all values of the autoregressive parameters, it is clear that $\bar{\eta}_\mu^{QS}$ becomes progressively very oversized as T increases. This gives some quantitative idea of the possible price of relying on a misspecified model for this testing problem.

In summary, it is clear that $\tilde{S}_k$ and $\tilde{S}_k^F$ provide generally much superior finite sample size properties across a range of data generating processes.

4.2 Power

Under the alternative hypothesis, we set $\rho_i = 1$ for $i = 1, \ldots, sN$ and $\rho_i = 0$ for $i = sN+1, \ldots, N$ where $s = 0.1$, 0.2, 0.4 and 0.6. When $r = 2$, we also consider an alternative of the form $\alpha_1 = \alpha_2 = 1$ with $\rho_i = 0$ for all i. Results for $r = 7$ are not reported for brevity and since the model for computing $\bar{\eta}_\mu^{QS}$ is misspecified and size control for that test is poor. Powers are reported in Table 3. Table 4 reports size adjusted powers (using the finite sample critical values calculated with $s = 0$ and $\alpha_1 = \alpha_2 = 0$ in the case of $r = 2$), which may have some theoretical interest but very limited practical relevance since they are based on infeasible tests.

[15] Unreported simulations show the Parzen lag window suggested by Yin and Wu (2000) for the test of Hadri (2000) to be inferior to the QS lag window in this case.

Table 3(a) reports powers for the non cross-correlated case $r = 0$. The results for $\tilde{S}_k$ and $\tilde{S}_k^F$ reveal power increasing with T as expected, and interestingly power also increasing with N. The pattern is similar for $\bar{\eta}_\mu^{QS}$, although the effect of increasing N is not as uniform. Table 2(a) shows the sizes of all three tests for $\rho_i = 0$ are very similar for $T = 150$ and 300, so comparing powers for these sample sizes reveals a considerable power advantage for the $\tilde{S}_k$ and $\tilde{S}_k^F$ tests. There is little difference between these latter two tests.

Table 3(b) reports powers for $r = 2$, and shows a significant change from the results in Table 3(a). The $\tilde{S}_k$ test is now least powerful by some margin against idiosyncratic unit roots while $\tilde{S}_k^F$ is occasionally most powerful and $\bar{\eta}_\mu^{QS}$ is frequently most powerful. The $\tilde{S}_k$ test does, however, have highly competitive against common factors with unit roots (the $s = 0$, $\alpha_j = 1$ alternative). Overall though, these results show a clear power advantage for $\tilde{S}_k^F$ over $\tilde{S}_k$, so it is well worth the preliminary estimation of the factor model to calculate $\tilde{S}_k^F$. The size adjusted powers in Table 4 again confirm the superiority of $\tilde{S}_k^F$ over $\tilde{S}_k$ when $r = 2$.

Overall, the size and power results show $\tilde{S}_k^F$ to be the preferred panel stationarity test for factor models. The $\tilde{S}_k$ has good size properties but inferior power against idiosyncratic unit roots in the presence of stationary common factors. The $\bar{\eta}_\mu^{QS}$ test has poor size control in the presence of strongly autocorrelated idiosyncratic components in any model and is susceptible to misspecification of the factor model, so it cannot be recommended for use from this comparison.

5 Testing the Purchasing Power Parity Hypothesis

In this section, we empirically test the purchasing power parity hypothesis, which is a fundamental ingredient of macroeconomic models of bilateral exchange rate behaviour. The validity of the PPP hypothesis has been an issue that has attracted a vast amount of attention in recent times and has been tested extensively using different panel unit root tests. In general, little evidence in support of PPP has been uncovered. For example, Papell (1997), Cheung and Lai (2000), Wu and Wu (2001) and Chang and Song (2002) are unable to provide strong evidence against the unit root null.[16] A failure to reject this null does not, however, provide compelling evidence against the PPP hypothesis, not least because low test power may be an issue here since real exchange rates tend to be highly correlated as they are typically constructed using a common numeraire currency and price index. Conversely, even if a rejection of the unit root null were to be obtained, this could not be interpreted as evidence for PPP holding in the entire panel because it may be that only a subset of the real exchange rates are stationary. In view of this, it makes some sense to apply our panel stationarity tests, $\tilde{S}_k$ and $\tilde{S}_k^F$, in this context. Here, the PPP hypothesis is represented by the stationary null and a rejection can, *ceteris paribus*, fairly unambiguously be interpreted as evidence against the PPP hypothesis being true.

We consider monthly real exchange rates against the US Dollar for the following countries: Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. The real exchange rate data was constructed from raw nominal exchange rate and consumer price index data taken from the IMF International Financial Statistics database. It covers the period of the recent float, 1973.01 to 1998.12. We have $N = 17$ and $T = 312$. In our notation we take $y_{i,t}$ to be the natural log of the real exchange rate, each standardized to have unit variance[17].

[16] In fact, what little empirical evidence there is in support of PPP has mainly arisen from application of tests that do not account for cross-sectional dependence at all; see Oh (1996) and Wu (1996).

[17] The data files and Gauss program for this application may be downloaded from www.economics.unimelb.edu.au/dharris

5.1 PPP with a Mean

As usual in PPP analysis, we first hypothesize that $y_{i,t}$ has a constant mean represented by

$$y_{i,t} = \mu_i + z_{i,t},$$

where the null hypothesis of PPP requires that $z_{i,t}$ is $I(0)$ and the alternative hypothesis is that $z_{i,t}$ is $I(1)$. The $\tilde{S}_k$ test statistics for each of the 17 series individually (i.e. each test has $N = 1$) are given in Table 5 under the "Constant" heading, together with the p-value of each test. At the 0.05 and 0.10 levels, the null hypothesis is rejected for 4 and 10 countries respectively, so the evidence against PPP is mixed.

The panel $\tilde{S}_k$ statistic pooled across the 17 countries gives $\tilde{S}_k = 1.93$ $(p = 0.027)$. The estimation of the factor model (9)–(12), for this data, gives $\hat{r} = 2$ when $r_{\max} = 5$, and the resulting test statistic is $\tilde{S}_k^F = 3.75$ $(p = 0.000)$. Thus the panel tests clearly reject the null hypothesis of PPP.

5.2 PPP with Structural Breaks

Papell (2002) suggests that rejections of PPP, such as that we have just found, are due to an unusual one-off episode in the 1980's when there was a large unexplained appreciation of the US dollar followed by an equally large offsetting depreciation. In econometric terms, this translates to a generalization of the deterministic specification so that

$$y_{i,t} = \mu_{i,t} + z_{i,t}, \tag{16}$$

where

$$\mu_{i,t} = \begin{cases} \beta_{1,i}, & t \leq \tau_{1,i}, \\ \beta_{2,i} + \beta_{3,i}t, & \tau_{1,i} \leq t \leq \tau_{2,i}, \\ \beta_{4,i} + \beta_{5,i}t, & \tau_{2,i} \leq t \leq \tau_{3,i}, \\ \beta_{1,i}, & \tau_{3,i} \leq t, \end{cases} \tag{17}$$

and all the $\beta_{j,i}$ and $\tau_{j,i}$ have to be estimated. Papell (2002) suggests that PPP holds around a constant long run real exchange rate before break point $\tau_{1,i}$ and after break point $\tau_{3,i}$; note that the same mean, $\beta_{1,i}$, applies in these two time periods and that this restriction is imposed on our test. The middle two time periods correspond to the great appreciation $(\tau_{1,i} \leq t \leq \tau_{2,i})$ and the great depreciation $(\tau_{2,i} \leq t \leq \tau_{3,i})$. The null hypothesis of PPP in this setting is that $y_{i,t}$ has representation (16) where $\mu_{i,t}$ satisfies (17) and $z_{i,t}$ is $I(0)$.

As in Papell (2002), note that the representation of $\mu_{i,t}$ in (17) is constrained to be continuous in t and it is therefore convenient to reparameterize $\mu_{i,t}$ as

$$\mu_{i,t} = \alpha_{1,i} + \alpha_{2,i}x_{1,i,t} + \alpha_{3,i}x_{2,i,t} + \alpha_{4,i}x_{3,i,t},$$

where for $h = 1, 2, 3$

$$x_{h,i,t} = (t - \tau_{h,i}) \cdot 1(t > \tau_{h,i}),$$

subject to the additional restrictions that $\alpha_{2,i} + \alpha_{3,i} + \alpha_{4,i} = 0$ (so that there is no trend for $t > \tau_{3,i}$) and $\alpha_{2,i}(\tau_{3,i} - \tau_{1,i}) + \alpha_{3,i}(\tau_{3,i} - \tau_{2,i})$ (so that the constant means for $t \leq \tau_{1,i}$ and $t \geq \tau_{3,i}$ are equal). The substitution of these restrictions gives

$$\mu_{i,t} = \alpha_{1,i} + \alpha_{2,i}x_{i,t}. \tag{18}$$

where

$$x_{i,t} = \left(x_{1,i,t} - \frac{\tau_{3,i} - \tau_{1,i}}{\tau_{3,i} - \tau_{2,i}} x_{2,i,t} + \frac{\tau_{2,i} - \tau_{1,i}}{\tau_{3,i} - \tau_{2,i}} x_{3,i,t} \right). \tag{19}$$

To implement the test it is first necessary to estimate the break dates $\tau_{1,i}, \tau_{2,i}$ and $\tau_{3,i}$ for each i. The number of breaks is set to three by the null hypothesis and does not need to be estimated, so it is computationally possible to use a three dimensional grid search to consistently estimate the break dates by least squares on the first difference of (16) where $\mu_{i,t}$ is given by (18). The estimated break dates for each country are shown in Table 5[18] and the estimated trend functions are shown graphically in Figures 1–3. Using the estimated break dates for each country i, denoted $\hat{\tau}_{1,i}$, $\hat{\tau}_{2,i}$ and $\hat{\tau}_{3,i}$, regressors $\hat{x}_{i,t}$ as in (19) can be constructed. This gives a regression[19] of the form $y_{i,t} = \alpha_{1,i} + \alpha_{2,i}\hat{x}_{i,t} + z_{i,t}$ for each i, which can be used to calculate $\tilde{S}_k$ and $\tilde{S}_k^F$.

The results from this analysis are given in Table 5 under the heading "Structural breaks". The individual tests do not reject the null hypothesis and the $\tilde{S}_k$ panel test gives $\tilde{S}_k = 1.12$ ($p = 0.132$). However, estimation of the factor model (with $\hat{r} = 2$ again found) gives $\tilde{S}_k^F = 3.20$ ($p = 0.001$), so that the null hypothesis is rejected. A quite plausible interpretation of this result, arising from our simulation evidence above, is that the $\tilde{S}_k$ test may have relatively low power in the presence of cross correlation (recall $\hat{r} = 2$) while the $\tilde{S}_k^F$ test retains rather better power in this case.

Finally, we note that a deterministic specification such as (17) nicely illustrates the versatility of the $\tilde{S}_k$ and $\tilde{S}_k^F$ tests. The approach used by Papell (2002) requires that the break dates be the same for all countries and even then requires bootstrap critical values for the panel unit root test. In contrast, the $\tilde{S}_k$ and $\tilde{S}_k^F$ tests permit arbitrary regression functions for each i with asymptotic critical values in all cases taken from the standard normal distribution.

[18] The break date estimates $\hat{\tau}_{1,i}$ for Canada and $\hat{\tau}_{3,i}$ for Japan lie outside the observed sample. Such estimates are possible because the means for $t \leq \tau_{1,i}$ and $\tau_{3,i} \leq t$ are constrained to be equal.

[19] Strictly speaking, this is not a deterministic regression because the break points are estimated. However, under the null hypothesis the estimated break points can be reexpressed as estimated break fractions in the usual way, and these are consistent. Theorem RES of HML can then be adapted in a straightforward manner to show that $\hat{x}_{i,t}$ may be replaced by $x_{i,t}$ which is defined in terms of the true breaks $\tau_{1,i}$,$\tau_{2,i}$,$\tau_{3.i}$ without change to the asymptotics of Theorems 1 and 2.

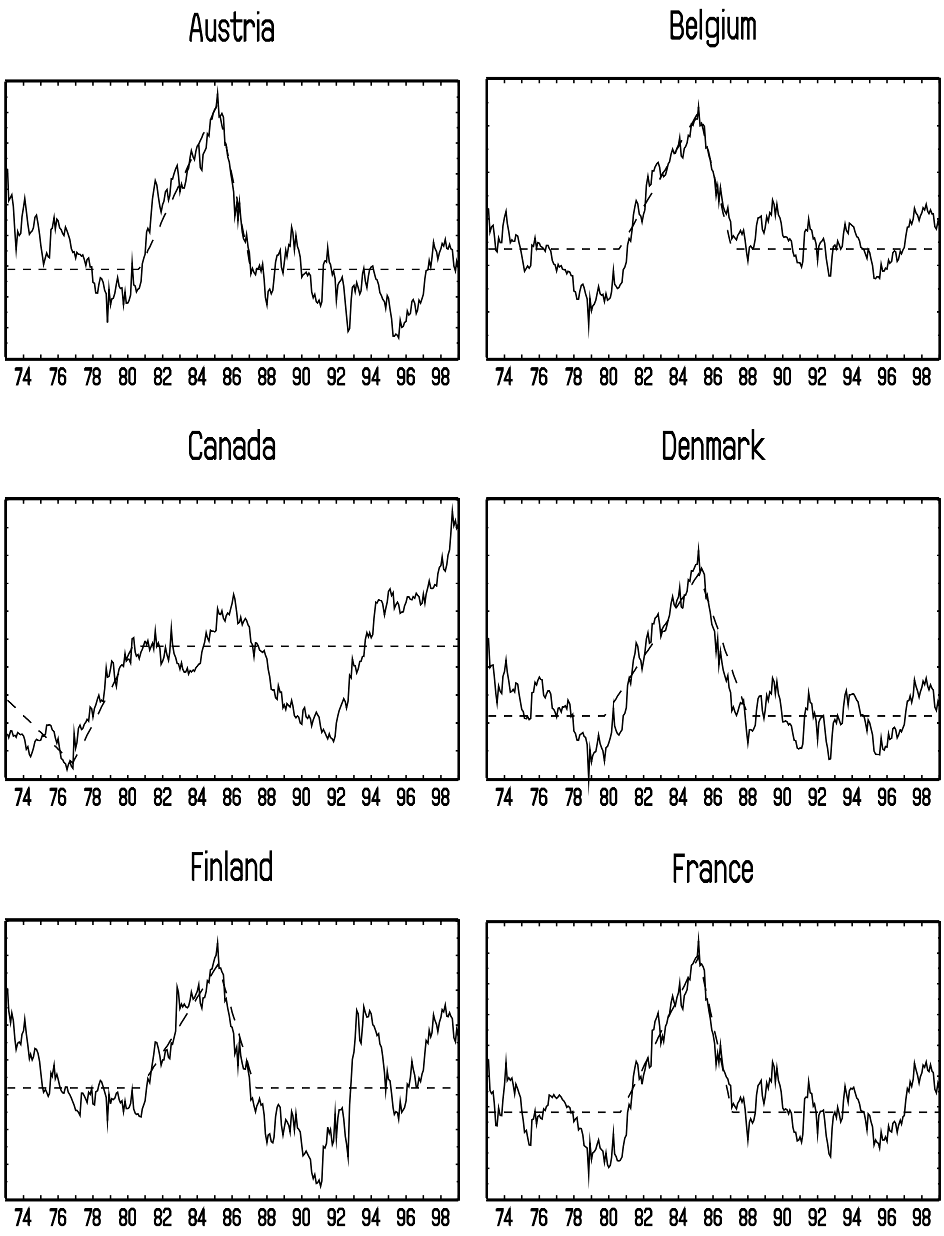

Figure 1. Real exchange rates with fitted structural breaks.

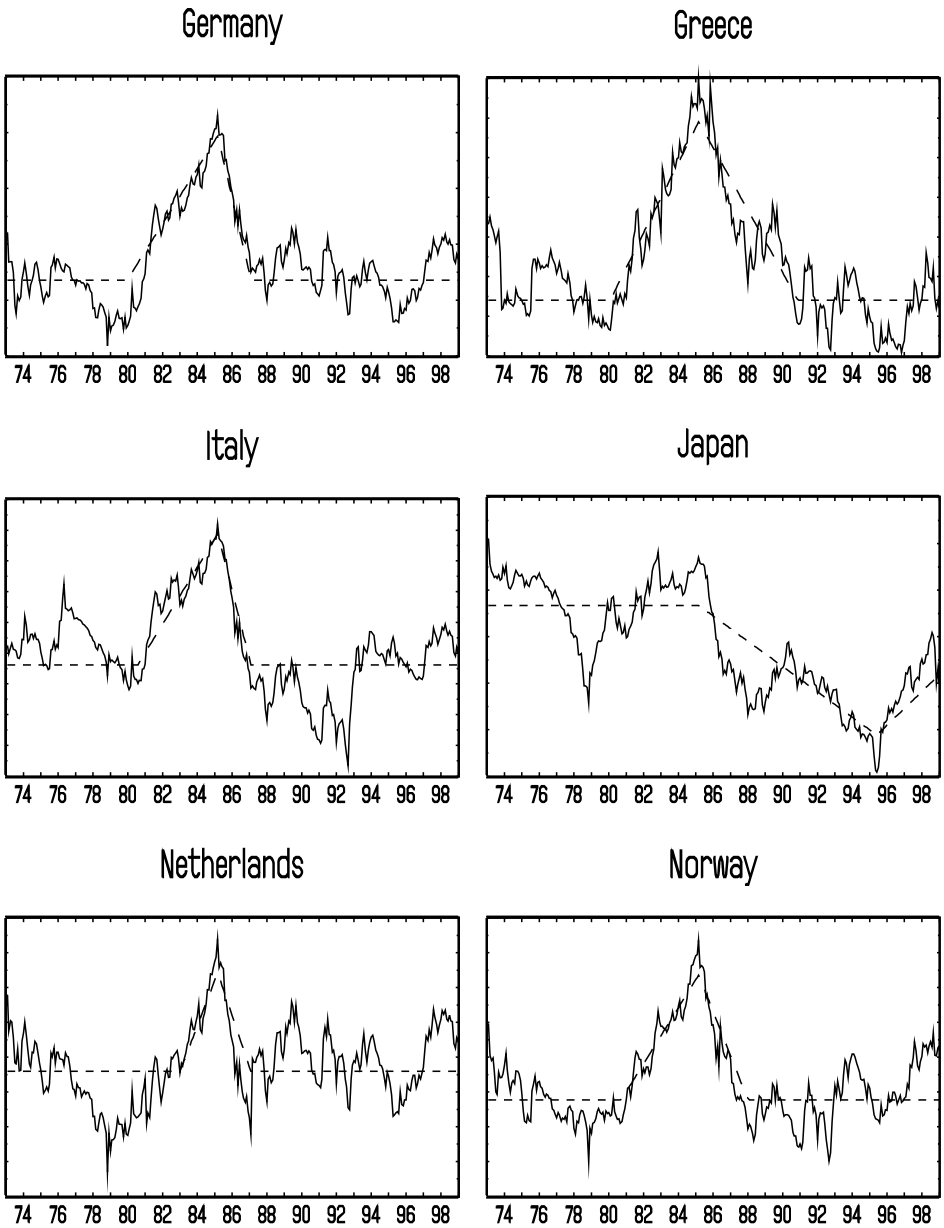

Figure 2. Real exchange rates with fitted structural breaks.

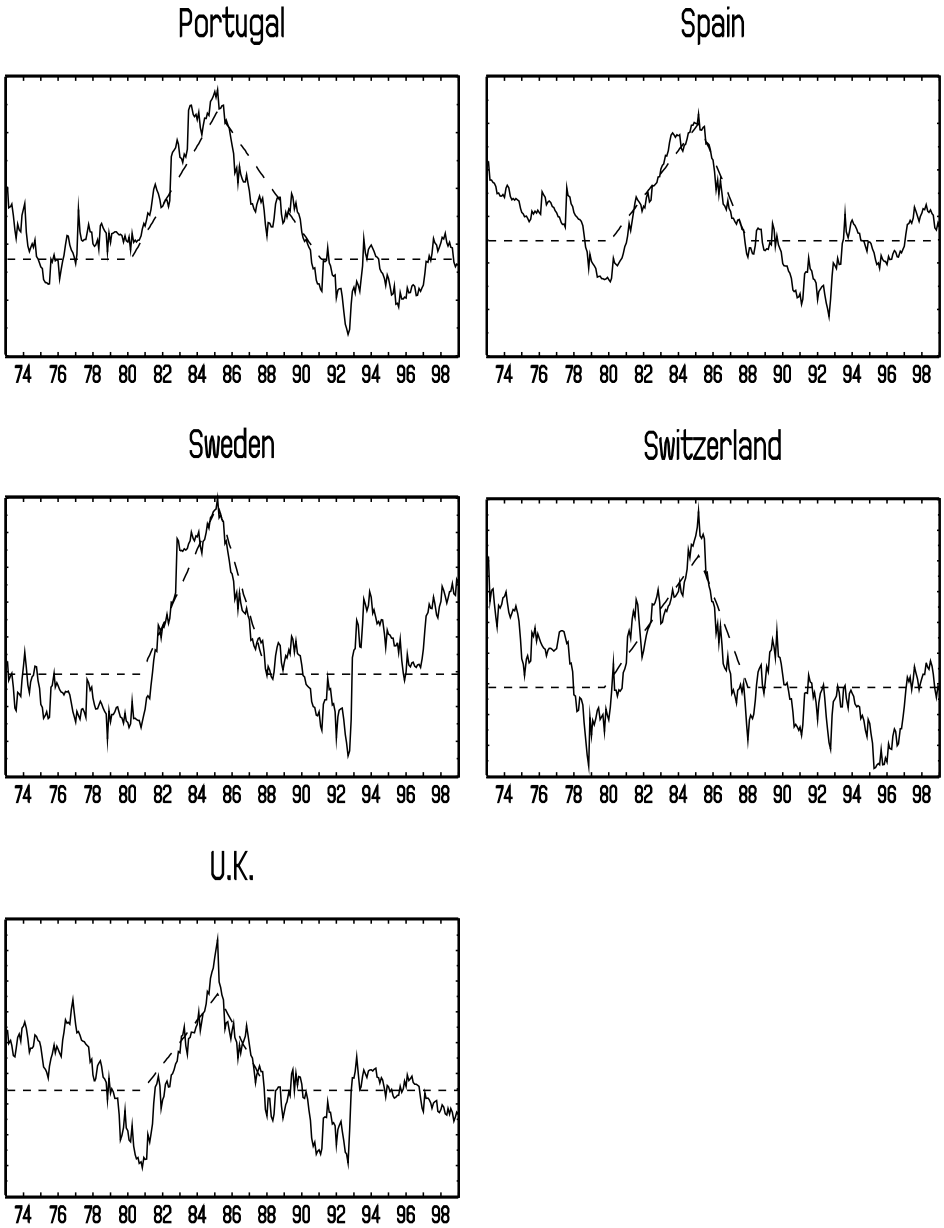

Figure 3. Real exchange rates with fitted structural breaks.

REFERENCES

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, 59, 817–858.

Bai, J. and S. Ng, 2002. Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.

Bai, J. and S. Ng, 2004a. A PANIC attach on unit roots and cointegration. *Econometrica* 72, 1127–1177.

Bai, J. and S. Ng, 2004b. A new look at panel testing of stationarity and the PPP hypothesis. In *Identification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Andrews, D. and J.H. Stock (eds) Cambridge University Press.

Bai, J. and P. Perron, 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.

Chang, Y., 2003. Bootstrap unit root tests in panels with cross-sectional dependency. Forthcoming, *Journal of Econometrics.*

Chang, Y. and Song, W., 2002. Panel unit root tests in the presence of cross-sectional dependency and heterogeneity. Mimeo., Rice University.

Cheung, Y.W. and Lai, K., 2000. On cross-country differences in the persistence of real exchange rates. *Journal of International Economics,* 50, 375-397.

Choi, I., 2001. Unit root tests for panel data. *Journal of International, Money and Finance,* 20, 219-247.

Choi, I., 2002. Combination unit root tests for cross-sectionally correlated panels. Mimeo., Hong Kong University.

Driscoll, J.C. and Kraay, A.C., 1998, Consistent Covariance Estimation with Spatially Dependent Panel Data, *Review of Economics and Statistics*, 80, 549-560.

Hadri, K., 2000. Testing for stationarity in heterogeneous panel data. *Econometrics Journal,* 3, 148-161.

Harris, D., McCabe, B.P.M. and Leybourne, S.J., 2003. Some limit theory for infinite order autocovariances. *Econometric Theory,* 19, 829-864.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159–178.

Im, K.S., Pesaran, M.H. and Shin Y., 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics,* 115, 53-74.

Levin A., Lin C.F. and Chu C.S., 2002. Unit root tests in panel data: Asymptotic and finite sample properties. *Journal of Econometrics*, 108, 1-24.

Maddala, G.S. and Wu, S., 1999. A comparative study of unit root tests with panel data and a new simple approach. *Oxford Bulletin of Economics and Statistics,* 61, 631-652.

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

Oh., K.Y., 1996. Purchasing power parity and unit root tests using panel data. *Journal of International, Money and Finance,* 15, 405-418.

Papell, D.H., 1997. Searching for stationarity: Purchasing power parity under the current float. *Journal of International Economics,* 43, 313-332.

Papell, D.H., 2002. The Great Appreciation, the Great Depreciation and the Purchasing Power Parity Hypothesis. *Journal of International Economics* 57, 51–82.

Shin, Y. and A. Snell (2002) Mean Group Tests for Stationarity in Heterogeneous Panels. Manuscript, University of Edinburgh.

Wu, Y., 1996. Are real exchange rates nonstationary? Evidence from a panel-data test. *Journal of Money, Credit and Banking,* 28, 54-63.

Wu, S. and Wu, J.L., 2001. Is purchasing power parity overvalued ? *Journal of Money, Credit and Banking,* 33, 804-812.
Yin, Y. and S. Wu, 2000. Stationarity tests in heterogeneous panels. In *Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Baltagi, B. (ed.) Elsevier Science.

6 Appendix

6.1 Proof of Theorem 1.

(i) Let $\hat{\mathbf{z}}_t = (\hat{z}_{1,t}, ..., \hat{z}_{N,t})'$ and $\tilde{\mathbf{z}}_t = (\tilde{z}_{1,t}, ..., \tilde{z}_{N,t})'$. Then $\tilde{C}_k$ can be written

$$\tilde{C}_k = \mathbf{d}' \operatorname{vec}\left[T^{-1/2}\sum_{t=k+1}^{T}\tilde{\mathbf{z}}_t\tilde{\mathbf{z}}'_{t-k}\right]$$

for a selector vector $\mathbf{d}$ defined as $\mathbf{d} = \operatorname{vec}[I_{N^2}]$. Now,

$$\begin{aligned}\operatorname{vec}\left[T^{-1/2}\sum_{t=k+1}^{T}\tilde{\mathbf{z}}_t\tilde{\mathbf{z}}'_{t-k}\right] &= \operatorname{vec}\left[T^{-1/2}\hat{\mathbf{G}}_0^{-1}\sum_{t=k+1}^{T}\hat{\mathbf{z}}_t\hat{\mathbf{z}}'_{t-k}\hat{\mathbf{G}}_0^{-1}\right] \\ &= (\hat{\mathbf{G}}_0^{-1}\otimes\hat{\mathbf{G}}_0^{-1})\operatorname{vec}\left[T^{-1/2}\sum_{t=k+1}^{T}\hat{\mathbf{z}}_t\hat{\mathbf{z}}'_{t-k}\right]\end{aligned}$$

where $\hat{\mathbf{G}}_0 = \operatorname{diag}[\hat{\gamma}_0\{\hat{z}_{1,t}\}^{1/2}, ..., \hat{\gamma}_0\{\hat{z}_{N,t}\}^{1/2}]$. It follows from HML, Theorem 8, that

$$\operatorname{vec}\left[T^{-1/2}\sum_{t=k+1}^{T}\hat{\mathbf{z}}_t\hat{\mathbf{z}}'_{t-k}\right] \Rightarrow N[0, \mathbf{\Omega}]$$

on noting that $\boldsymbol{\eta}_t = \operatorname{diag}[(1-\phi_1 L)^{-1}, .., (1-\phi_N L)^{-1}]\boldsymbol{\varepsilon}_t$ also satisfies the conditions of Assumption LP of HML. Moreover, since $\hat{\mathbf{G}}_0 \xrightarrow{p} \mathbf{G}_0 = \operatorname{diag}[E(\boldsymbol{\eta}_t\boldsymbol{\eta}'_t)^{1/2}]$,

$$\tilde{C}_k \Rightarrow N[0, \mathbf{d}'(\mathbf{G}_0^{-1}\otimes\mathbf{G}_0^{-1})\mathbf{\Omega}(\mathbf{G}_0^{-1}\otimes\mathbf{G}_0^{-1})\mathbf{d}]$$

by the continuous mapping theorem (CMT). Next, with $a_t = \sum_{i=1}^{N}\tilde{z}_{i,t}\tilde{z}_{i,t-k}$ and $\mathbf{b}_t = \operatorname{vec}[\tilde{\mathbf{z}}_t\tilde{\mathbf{z}}'_{t-k}]$ and $\mathbf{c}_t = \operatorname{vec}[\hat{\mathbf{z}}_t\hat{\mathbf{z}}'_{t-k}]$, we may write

$$\hat{\omega}\{a_t\}^2 = \mathbf{d}'\hat{\mathbf{\Omega}}\{\mathbf{b}_t\}\mathbf{d} = \mathbf{d}'(\hat{\mathbf{G}}_0^{-1}\otimes\hat{\mathbf{G}}_0^{-1})\hat{\mathbf{\Omega}}\{\mathbf{c}_t\}(\hat{\mathbf{G}}_0^{-1}\otimes\hat{\mathbf{G}}_0^{-1})\mathbf{d}$$

where $\hat{\mathbf{\Omega}}\{.\}$ is given in (7). From HML, Theorem 8, for a specified matrix $\mathbf{\Omega}$, $\hat{\mathbf{\Omega}}\{\mathbf{c}_t\} \xrightarrow{p} \mathbf{\Omega}$ and hence, by the CMT,

$$\hat{\omega}\{a_t\}^2 \xrightarrow{p} \mathbf{d}'(\mathbf{G}_0^{-1}\otimes\mathbf{G}_0^{-1})\mathbf{\Omega}(\mathbf{G}_0^{-1}\otimes\mathbf{G}_0^{-1})\mathbf{d}$$

so that

$$\frac{\tilde{C}_k}{\hat{\omega}\{a_t\}} \Rightarrow N[0, 1]. \tag{20}$$

To deal with the bias correction $\tilde{c}$, the expectation of the estimation error under H_0 using the standardized residuals $\tilde{z}_{i,t}$ can be written

$$c_i = \operatorname{tr} E\left[\frac{1}{s_i^2}\left(T^{-1/2}\sum_{t=1}^{T}\tilde{\mathbf{x}}_{i,t}z_{i,t}\right)\left(T^{-1/2}\sum_{t=1}^{T}\tilde{\mathbf{x}}_{i,t}z_{i,t}\right)'\right]$$

where $\tilde{\mathbf{x}}_{i,t} = \left(T^{-1}\sum_{t=1}^{T}\mathbf{x}_{i,t}\mathbf{x}_{i,t}'\right)^{-1/2}\mathbf{x}_{i,t}$. In this form c_i is clearly $O(1)$ and moreover by Theorem 1 of Andrews (1991) c_i is consistently estimated by $\tilde{c}_i$. Thus $\tilde{S}_k = \hat{\omega}\{a_t\}^{-1}\left(\tilde{C}_k + \sum_{i=1}^{N}T^{-1/2}\tilde{c}_i\right) = \hat{\omega}\{a_t\}^{-1}\tilde{C}_k + O_p\left(T^{-1/2}\right) \Rightarrow N[0,1]$ from (20).

(ii) Suppose, without loss of generality, that $\phi_i = 1$ for $i = 1, ..., sN$, $0 < s \leq 1$ and $\phi_i < 1$ for $i = sN+1, ..., N$ (with the obvious modification for $s = 1$). Now

$$T^{-1/2}\tilde{C}_k = \sum_{i=1}^{sN}T^{-1}\sum_{t=k+1}^{T}\tilde{z}_{i,t}\tilde{z}_{i,t-k} + \sum_{i=sN+1}^{N}T^{-1}\sum_{t=k+1}^{T}\tilde{z}_{i,t}\tilde{z}_{i,t-k},$$

and the second term is $O_p\left(T^{-1/2}\right)$ from the proof of Theorem 1. Noting that $T^{-1}\hat{\gamma}_0\{\hat{z}_{i,t}\} = O_p(1)$ for $i = 1, \ldots, M$, the first term satisfies

$$\begin{aligned}
\sum_{i=1}^{sN}T^{-1}\sum_{t=k+1}^{T}\tilde{z}_{i,t}\tilde{z}_{i,t-k} &= \sum_{i=1}^{sN}\frac{1}{T^{-1}\hat{\gamma}_0\{\hat{z}_{it}\}}T^{-2}\sum_{t=k+1}^{T}\hat{z}_{i,t}\hat{z}_{i,t-k} \\
&= \sum_{i=1}^{sN}\frac{1}{T^{-1}\hat{\gamma}_0\{\hat{z}_{i,t}\}}T^{-2}\sum_{t=1}^{T}\hat{z}_{i,t}^2 + o_p(1) \\
&= sN + o_p(1).
\end{aligned}$$

Thus,

$$\tilde{C}_k = T^{1/2}sN + o_p\left(T^{1/2}\right). \tag{21}$$

Next, with $a_t = \sum_{i=1}^{N}\tilde{z}_{i,t}\tilde{z}_{i,t-k}$,

$$l^{-1}\hat{\omega}\{a_t\}^2 = l^{-1}\left(\hat{\gamma}_0\{a_t\} + 2\sum_{j=1}^{l}\left(1 - \frac{j}{l}\right)\hat{\gamma}_j\{a_t\}\right) \leq 3\hat{\gamma}_0\{a_t\} \tag{22}$$

where $\hat{\gamma}_0\{a_t\} = T^{-1}\sum_{t=k+1}^{T}\left(\sum_{i=1}^{N}\tilde{z}_{i,t}\tilde{z}_{i,t-k}\right)^2$.Thus

$$\hat{\gamma}_0\{a_t\} \leq \sum_{i=1}^{N}\sum_{j=1}^{N}T^{-1}\sum_{t=1}^{T}\tilde{z}_{i,t}^2\tilde{z}_{j,t}^2 = O_p(1) \tag{23}$$

since for $i, j = sN+1, \ldots, N$,

$$T^{-1}\sum_{t=k+1}^{T}\tilde{z}_{i,t}^2\tilde{z}_{j,t}^2 = \frac{1}{\hat{\gamma}_0\{\hat{z}_{i,t}\}.\hat{\gamma}_0\{\hat{z}_{j,t}\}}T^{-1}\sum_{t=k+1}^{T}\hat{z}_{i,t}^2\hat{z}_{j,t}^2 = O_p(1)$$

for $i, j = 1, \ldots, sN$,

$$T^{-1}\sum_{t=k+1}^{T}\tilde{z}_{i,t}^2\tilde{z}_{j,t}^2 = \frac{1}{T^{-1}\hat{\gamma}_0\{\hat{z}_{i,t}\}.T^{-1}\hat{\gamma}_0\{\hat{z}_{j,t}\}}T^{-3}\sum_{t=k+1}^{T}\hat{z}_{i,t}^2\hat{z}_{j,t}^2 = O_p(1)$$

and for $i = sN+1, \ldots, N$ and $j = 1, \ldots, sN$,

$$T^{-1}\sum_{t=k+1}^{T}\tilde{z}_{i,t}^2\tilde{z}_{j,t}^2 = \frac{1}{\hat{\gamma}_0\{\hat{z}_{i,t}\}.T^{-1}\hat{\gamma}_0\{\hat{z}_{j,t}\}}T^{-2}\sum_{t=k+1}^{T}\hat{z}_{i,t}^2\hat{z}_{j,t}^2 = O_p(1).$$

(and similarly for $j = sN+1,\ldots,N$ and $i = 1,\ldots,sN$). Combining (21), (22) and (23) we find for any critical value c, as $T \to \infty$,

$$P\left[\frac{\tilde{C}_k}{\hat{\omega}\{a_t\}} > c\right] = P\left[\frac{sN + o_p(1)}{\frac{l}{T^{1/2}} O_p(1)} > c\right] \to 1$$

since $l = o(T^{1/2})$. The bias correction $\tilde{c}$ obviously does not affect test consistency since each $T^{-1/2}\tilde{c}_i$ is non-negative.

6.2 Deterministic Regressions in the Factor Model

In standard notation, the model (14) (where $\mathbf{x}_{i,t}$ is an $m_i \times 1$ vector of regressors) can be written

$$\underset{(NT\times1)}{\mathbf{y}} = \underset{(NT\times m)}{\mathbf{X}}\underset{(m\times1)}{\boldsymbol{\beta}} + \underset{(NT\times1)}{\mathbf{z}}, \tag{24}$$

where $\mathbf{y} = (\mathbf{y}_1',\ldots,\mathbf{y}_N')'$, $\mathbf{X} = \operatorname{diag}(\mathbf{X}_1,\ldots,\mathbf{X}_N)$, $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1',\ldots,\boldsymbol{\beta}_N'\right)'$, $\mathbf{z} = (\mathbf{z}_1',\ldots,\mathbf{z}_N')'$, $m = \sum_{i=1}^N m_i$ and $\mathbf{y}_i = (y_{i,1},\ldots,y_{i,T})'$, $\mathbf{X}_i = (\mathbf{x}_{i,1},\ldots,\mathbf{x}_{i,T})'$, $\mathbf{z}_i = (z_{i,1},\ldots,z_{i,T})'$. The factor model (10) can be written

$$\underset{(T\times N)}{\mathbf{Z}} = \underset{(T\times r)}{\mathbf{F}}\underset{(r\times N)}{\boldsymbol{\Lambda}} + \underset{(T\times N)}{\mathbf{E}} \tag{25}$$

where $\mathbf{Z} = (\mathbf{z}_1,\ldots,\mathbf{z}_N)$, $\mathbf{F} = (\mathbf{f}_1,\ldots,\mathbf{f}_r)'$, $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_N)$ and $\mathbf{E} = (\mathbf{e}_1,\ldots,\mathbf{e}_N)$ where $\mathbf{e}_i = (e_{i,1},\ldots,e_{i,T})'$ for each $i = 1,\ldots,N$.

The estimation of the factor model in first differences proceeds as follows. The first difference of equation (24) can be written

$$\underset{(N(T-1)\times1)}{\Delta\mathbf{y}} = \underset{(N(T-1)\times m_\Delta)}{(\Delta\mathbf{X}\mathbf{C})}\underset{(m_\Delta\times1)}{\boldsymbol{\beta}_C} + \underset{(N(T-1)\times1)}{\Delta\mathbf{z}},$$

where $\boldsymbol{\beta}_C = \left(\mathbf{C}'\mathbf{C}\right)^{-1}\mathbf{C}'\boldsymbol{\beta}$ and $\mathbf{C}$ is an $m \times m_\Delta$ matrix chosen to exclude columns of $\Delta\mathbf{X}$ corresponding to constant terms in $\mathbf{X}$ so that $\Delta\mathbf{X}\mathbf{C}$ has full column rank $m_\Delta \le m$. The residuals from this regression can be written $\widehat{\Delta\mathbf{z}} = \Delta\mathbf{y} - (\Delta\mathbf{X}\mathbf{C})\,\hat{\boldsymbol{\beta}}_C$, where $\hat{\boldsymbol{\beta}}_C = \left(\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{X}\mathbf{C}\right)^{-1}\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{y}$ is the usual OLS estimator, and the $(T-1)\times N$ matrix $\widehat{\Delta\mathbf{Z}}$ is defined to satisfy $\widehat{\Delta\mathbf{z}} = \operatorname{vec}\left(\widehat{\Delta\mathbf{Z}}\right)$. The estimated factors can then be written $\widehat{\Delta\mathbf{F}} = \widehat{\Delta\mathbf{Z}}\hat{\boldsymbol{\Gamma}}$ where $\hat{\boldsymbol{\Gamma}}$ is the $N \times r$ matrix of eigenvectors corresponding to the largest r eigenvalues of $\widehat{\Delta\mathbf{Z}}'\widehat{\Delta\mathbf{Z}}$. The estimated idiosyncratic components are then $\widehat{\Delta\mathbf{E}} = \widehat{\Delta\mathbf{Z}} - \widehat{\Delta\mathbf{F}}\hat{\boldsymbol{\Lambda}}$ where $\hat{\boldsymbol{\Lambda}} = \left(\widehat{\Delta\mathbf{F}}'\widehat{\Delta\mathbf{F}}\right)^{-1}\widehat{\Delta\mathbf{F}}'\widehat{\Delta\mathbf{Z}}$. Taking the partial sums of $\widehat{\Delta\mathbf{F}}$ and $\widehat{\Delta\mathbf{E}}$ gives the component estimates $\hat{\mathbf{F}}$ and $\hat{\mathbf{E}}$, with corresponding $r(T-1)\times 1$ and $N(T-1)\times 1$ vectors $\hat{\mathbf{f}} = \operatorname{vec}\left(\hat{\mathbf{F}}\right)$ and $\hat{\mathbf{e}} = \operatorname{vec}\left(\hat{\mathbf{E}}\right)$.

The deterministic regressions for the estimated factors $\hat{\mathbf{f}}$ proceed as follows. Since $\hat{\mathbf{f}}$ can be written $\hat{\mathbf{f}} = \left(\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\widehat{\Delta\mathbf{z}}$, it is necessary to regress $\hat{\mathbf{f}}$ on $\left(\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{X}$. However this regressor matrix may not have full column rank, in which case it is sufficient to regress $\hat{\mathbf{f}}$ on $\mathbf{X}_f = \left(\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{X}\mathbf{C}_f$ where $\mathbf{C}_f$ is a matrix chosen such that $\mathbf{X}_f$ has full column rank and its columns form a basis for the vector space containing the columns of $\left(\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{X}$. In practice, if $\left(\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{X}$ has less than full column rank, a simple choice for $\mathbf{C}_f$ is the matrix of eigenvectors corresponding to the non-zero eigenvalues of $\mathbf{X}'\left(\hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{X}$. The residuals from the regression of $\hat{\mathbf{f}}$ on $\mathbf{X}_f$ are denoted $\hat{\hat{\mathbf{f}}}$. The corresponding $(T-1)\times r$ matrix $\hat{\hat{\mathbf{F}}}$ is defined to satisfy

$\hat{\hat{\mathbf{f}}} = \text{vec}\left(\hat{\hat{\mathbf{F}}}\right)$. Each of the r columns of $\hat{\hat{\mathbf{F}}}$ is standardized by its sample standard deviation to give the $(T-1)\times r$ matrix $\tilde{\mathbf{F}}$, whose individual elements are denoted $\tilde{f}_{j,t}$, $j = 1,\ldots,r$ and $t = 2,\ldots,T$.

The deterministic regressions for the estimated idiosyncratic components $\hat{\mathbf{e}}$ proceed similarly. We can write $\widehat{\Delta\mathbf{E}} = \widehat{\Delta\mathbf{Z}} - \widehat{\Delta\mathbf{Z}}\left(\hat{\mathbf{\Gamma}}\hat{\mathbf{\Lambda}}\right)$ since $\widehat{\Delta\mathbf{F}} = \widehat{\Delta\mathbf{Z}}\hat{\mathbf{\Gamma}}$, so $\hat{\mathbf{e}} = \left(\mathbf{I}_{N(T-1)} - \left(\hat{\mathbf{\Lambda}}'\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\right)\hat{\mathbf{z}}$. Therefore it is necessary to regress $\hat{\mathbf{e}}$ on $\left(\mathbf{I}_{N(T-1)} - \left(\hat{\mathbf{\Lambda}}'\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\right)\mathbf{X}$. If this regressor matrix does not have full rank then it is replaced by $\mathbf{X}_e = \left(\mathbf{I}_{N(T-1)} - \left(\hat{\mathbf{\Lambda}}'\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\right)\mathbf{X}\mathbf{C}_e$ where, like $\mathbf{C}_f$, $\mathbf{C}_e$ is chosen (by principal components or other means) so that the columns of $\mathbf{X}_e$ provide a basis for those of $\left(\mathbf{I}_{N(T-1)} - \left(\hat{\mathbf{\Lambda}}'\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\right)\mathbf{X}$. The residuals from the regression of $\hat{\mathbf{e}}$ on $\mathbf{X}_e$ are denoted $\hat{\hat{\mathbf{e}}}$ and the columns of the corresponding matrix $\hat{\hat{\mathbf{E}}}$ (i.e. $\hat{\hat{\mathbf{e}}} = \text{vec}\left(\hat{\hat{\mathbf{E}}}\right)$) and standardized by their respective standard deviations to give $\tilde{\mathbf{E}}$ with individual elements $\tilde{e}_{i,t}$.

To calculate the statistic (8), define $\tilde{a}_{k,t} = \sum_{j=1}^{r}\tilde{f}_{j,t}\tilde{f}_{j,t-k} + \sum_{i=1}^{N}\tilde{e}_{i,t}\tilde{e}_{i,t-k}$. Then $\tilde{C}_k$ and $\hat{\omega}\left\{\tilde{a}_{k,t}\right\}$ are calculated as described following equation (8). The bias correction term $\tilde{c}$ in (8) is given by

$$\tilde{c} = (T-k)^{-1/2}\,\text{tr}\left[\left(\mathbf{X}_f'\mathbf{X}_f/T\right)^{-1}\hat{\mathbf{\Omega}}\left\{w_{f,s}\right\} + \left(\mathbf{X}_e'\mathbf{X}_e/T\right)^{-1}\hat{\mathbf{\Omega}}\left\{w_{e,s}\right\}\right],$$

where $\mathbf{w}_f = \mathbf{X}_f \odot \hat{\mathbf{f}}\boldsymbol{\iota}_{m_f}' = \left\{w_{f,s}'\right\}_{s=1}^{r(T-1)}$ and $\mathbf{w}_e = \mathbf{X}_e \odot \hat{\mathbf{e}}\boldsymbol{\iota}_{m_e}' = \left\{w_{e,s}'\right\}_{s=1}^{N(T-1)}$, where $\odot$ is the Hadamard product, $\boldsymbol{\iota}_m$ is an $m \times 1$ vector of ones, m_f and m_e are the column dimensions of $\mathbf{X}_f$ and $\mathbf{X}_e$ respectively.

6.3 Proof of Theorem 2

In the notation defined in Section 6.2, the residuals $\widehat{\Delta\mathbf{z}}$ can be written

$$\widehat{\Delta\mathbf{z}} = \Delta\mathbf{z} - \Delta\mathbf{X}\mathbf{C}\left(\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{X}\mathbf{C}\right)^{-1}\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{z}.$$

Taking partial sums of $\widehat{\Delta\mathbf{z}}$ gives

$$\hat{\mathbf{z}} = \mathbf{z} - \mathbf{X}\mathbf{C}\left(\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{X}\mathbf{C}\right)^{-1}\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{z}$$

apart from some asymptotically negligible initial value effects, so that $\mathbf{z}$ and $\mathbf{X}$ now have $N(T-1)$ rows.

The partial sum of $\widehat{\Delta\mathbf{f}} = \text{vec}\left(\widehat{\Delta\mathbf{F}}\right) = \left(\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\widehat{\Delta\mathbf{z}}$ gives

$$\hat{\mathbf{f}} = \left(\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{z} - \mathbf{X}_f\mathbf{B}_f$$

where $\mathbf{X}_f = \left(\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{X}\mathbf{C}_f$ and $\mathbf{B}_f = \left(\mathbf{C}_f'\mathbf{C}_f\right)^{-1}\mathbf{C}_f'\mathbf{C}\left(\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{X}\mathbf{C}\right)^{-1}\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{z}$. Thus regressing $\hat{\mathbf{f}}$ on $\mathbf{X}_f$ will remove the $\mathbf{X}_f\mathbf{B}_f$ term from the residuals, giving

$$\hat{\hat{\mathbf{f}}} = \bar{\mathbf{P}}_f\hat{\mathbf{f}} = \bar{\mathbf{P}}_f\left(\hat{\mathbf{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{z}, \tag{26}$$

where $\bar{\mathbf{P}}_f = \mathbf{I}_{r(T-1)} - \mathbf{X}_f\left(\mathbf{X}_f'\mathbf{X}_f\right)^{-1}\mathbf{X}_f' = \mathbf{I}_{r(T-1)} - \mathbf{P}_f$. The corresponding matrix $\hat{\hat{\mathbf{F}}}$ satisfies $\hat{\hat{\mathbf{f}}} = \text{vec}\left(\hat{\hat{\mathbf{F}}}\right)$ and the standardized matrix $\tilde{\mathbf{F}}$ is found by $\tilde{\mathbf{F}} = \hat{\hat{\mathbf{F}}}\hat{\mathbf{G}}_f^{-1}$, where $\hat{\mathbf{G}}_f$ is an $r \times r$ diagonal matrix containing the sample standard deviations of the columns of $\hat{\hat{\mathbf{F}}}$ on the diagonal. Thus $\tilde{\mathbf{f}} = \left(\hat{\mathbf{G}}_f^{-1} \otimes \mathbf{I}_{T-1}\right)\hat{\hat{\mathbf{f}}}$.

The partial sum of $\widehat{\Delta \mathbf{e}} = \hat{\mathbf{A}}\widehat{\Delta \mathbf{z}}$ where $\hat{\mathbf{A}} = \mathbf{I}_{N(T-1)} - \left(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)$ gives

$$\hat{\mathbf{e}} = \hat{\mathbf{A}}\mathbf{z} - \mathbf{X}_e\mathbf{B}_e$$

where $\mathbf{X}_e = \left(\mathbf{I}_{N(T-1)} - \left(\hat{\boldsymbol{\Lambda}}'\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\right)\mathbf{X}\mathbf{C}_e$ and $\mathbf{B}_e = \left(\mathbf{C}_e'\mathbf{C}_e\right)^{-1}\mathbf{C}_e'\mathbf{C}\left(\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{X}\mathbf{C}\right)^{-1}\mathbf{C}'\Delta\mathbf{X}'\Delta\mathbf{z}$ so regressing $\hat{\mathbf{e}}$ on $\mathbf{X}_e$ will remove $\mathbf{X}_e$. This leaves

$$\hat{\hat{\mathbf{e}}} = \bar{\mathbf{P}}_e\hat{\mathbf{e}} = \bar{\mathbf{P}}_e\hat{\mathbf{A}}\mathbf{z} \tag{27}$$

where $\bar{\mathbf{P}}_e$ is the orthogonal projection on $\mathbf{X}_e$. The corresponding matrix $\hat{\hat{\mathbf{E}}}$ satisfies $\hat{\hat{\mathbf{e}}} = \operatorname{vec}\left(\hat{\hat{\mathbf{E}}}\right)$ and the standardized matrix $\tilde{\mathbf{E}}$ is found by $\tilde{\mathbf{E}} = \hat{\hat{\mathbf{E}}}\hat{\mathbf{G}}_e^{-1}$, where $\hat{\mathbf{G}}_e$ is an $r \times r$ diagonal matrix containing the sample standard deviations of the columns of $\hat{\hat{\mathbf{E}}}$ on the diagonal. These steps show that the appropriate regressions of $\hat{\mathbf{f}}$ on $\mathbf{X}_f$ and $\hat{\mathbf{e}}$ on $\mathbf{X}_e$ remove the effects of the initial deterministic regression in first differences.

Since the model is estimated in differences, it follows that under both null and alternative $\widehat{\Delta \mathbf{Z}}'\widehat{\Delta \mathbf{Z}}/T = \boldsymbol{\Sigma}_{\Delta\Delta} + O_p\left(T^{-1/2}\right)$ and hence $\hat{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma} + O_p\left(T^{-1/2}\right)$ where $\boldsymbol{\Gamma}$ is the matrix of eigenvectors corresponding to the largest r eigenvalues of $\boldsymbol{\Sigma}_{\Delta\Delta}$. Thus $\hat{\boldsymbol{\Lambda}} = \left(\hat{\boldsymbol{\Gamma}}'\widehat{\Delta \mathbf{Z}}'\widehat{\Delta \mathbf{Z}}\hat{\boldsymbol{\Gamma}}\right)^{-1}\hat{\boldsymbol{\Gamma}}'\widehat{\Delta \mathbf{Z}}'\widehat{\Delta \mathbf{Z}} = \left(\boldsymbol{\Gamma}'\boldsymbol{\Sigma}_{\Delta\Delta}\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\Delta\Delta} + O_p\left(T^{-1/2}\right)$. Recalling the definitions of $\hat{\hat{\mathbf{f}}}$ and $\hat{\hat{\mathbf{e}}}$ in (26) and (27) respectively, consider $\bar{\mathbf{f}} = \left(\hat{\boldsymbol{\Gamma}}' \otimes \mathbf{I}_{T-1}\right)\mathbf{z}$ and $\bar{\mathbf{e}} = \hat{\mathbf{A}}\mathbf{z}$. We can write $\left(\bar{\mathbf{f}}', \bar{\mathbf{e}}'\right)' = \operatorname{vec}\left(\mathbf{W}\hat{\mathbf{C}}\right)$ where $\mathbf{W} = \mathbf{Z}\boldsymbol{\Sigma}_{\Delta\Delta}^{-1/2}$ and $\hat{\mathbf{C}} = \boldsymbol{\Sigma}_{\Delta\Delta}^{1/2}\left(\hat{\boldsymbol{\Gamma}}, \hat{\mathbf{P}}\right)$ where $\hat{\mathbf{P}} = \mathbf{I}_N - \hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{\Lambda}}$ and $\boldsymbol{\Sigma}_{\Delta\Delta} = \mathbf{E}\left(\Delta\mathbf{z}_t\Delta\mathbf{z}_t'\right)$. Now $\hat{\mathbf{C}} = \mathbf{C} + O_p\left(T^{-1/2}\right)$ where $\mathbf{C} = \boldsymbol{\Sigma}_{\Delta\Delta}^{1/2}\left(\boldsymbol{\Gamma}, \mathbf{P}\right)$ and $\mathbf{P} = \mathbf{I}_N - \boldsymbol{\Gamma}\left(\boldsymbol{\Gamma}'\boldsymbol{\Sigma}_{\Delta\Delta}\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\Delta\Delta}$. Note that $\mathbf{C}$ has rank N since (i) $\boldsymbol{\Gamma}$ has full column rank r, (ii) $\mathbf{P}$ has rank[20] $N-r$ and (iii) $\boldsymbol{\Gamma}'\boldsymbol{\Sigma}_{\Delta\Delta}\mathbf{P} = \mathbf{0}$. This shows that $\left(\bar{\mathbf{f}}', \bar{\mathbf{e}}'\right)' = \operatorname{vec}\left(\mathbf{W}\hat{\mathbf{C}}\right)$ is, even asymptotically, a rank N linear transformation of $\mathbf{Z}$ and hence that all series of $\left(\bar{\mathbf{f}}', \bar{\mathbf{e}}'\right)'$ are $I(0)$ if all series of $\mathbf{Z}$ are $I(0)$ and also that $\left(\bar{\mathbf{f}}', \bar{\mathbf{e}}'\right)'$ must contain $I(1)$ elements if $\mathbf{Z}$ does. Therefore Theorem 1 can be applied to $\left(\bar{\mathbf{f}}', \bar{\mathbf{e}}'\right)'$. The effect of the detrending regressions represented by the projections $\bar{\mathbf{P}}_f$ and $\bar{\mathbf{P}}_e$ can be handled as in Theorem RES of HML. The preceding arguments hold for every $r = 0, 1, \ldots, r_{\max}$ and hence hold when r is replaced by $\hat{r} \in \{0, 1, \ldots, r_{\max}\}$.

[20] $P' = \Sigma_{\Delta\Delta}^{1/2}P^*\Sigma_{\Delta\Delta}^{-1/2}$ where $P^* = I_N - \Sigma_{\Delta\Delta}^{1/2}\Gamma\left(\Gamma'\Sigma_{\Delta\Delta}\Gamma\right)^{-1}\Gamma'\Sigma_{\Delta\Delta}^{1/2}$ is idempotent with $\operatorname{tr} P^* = N - r$ and hence has rank $N-r$.

Table 2: Simulated Size
(a) $r = 0$.

	$\rho_i = 0.0$			$\rho_i = 0.4$			$\rho_i = 0.8$			$\rho_i = U[0, 0.8]$		
N	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$
							$T = 30$					
10	0.07	0.09	0.03	0.08	0.07	0.00	0.02	0.02	0.00	0.07	0.06	0.00
20	0.11	0.11	0.04	0.08	0.08	0.01	0.01	0.01	0.00	0.04	0.04	0.01
30	0.11	0.11	0.05	0.05	0.05	0.01	0.00	0.00	0.00	0.04	0.04	0.01
40	0.12	0.11	0.06	0.05	0.05	0.01	0.00	0.00	0.00	0.03	0.03	0.01
							$T = 50$					
10	0.06	0.06	0.03	0.07	0.07	0.01	0.06	0.07	0.00	0.07	0.08	0.01
20	0.07	0.07	0.05	0.07	0.07	0.01	0.03	0.03	0.00	0.07	0.07	0.01
30	0.07	0.07	0.03	0.06	0.06	0.00	0.01	0.01	0.00	0.06	0.06	0.00
40	0.10	0.10	0.05	0.08	0.08	0.01	0.01	0.01	0.00	0.06	0.06	0.01
							$T = 75$					
10	0.05	0.06	0.05	0.07	0.07	0.03	0.06	0.06	0.04	0.07	0.07	0.03
20	0.07	0.07	0.03	0.08	0.08	0.02	0.05	0.05	0.04	0.07	0.07	0.02
30	0.07	0.07	0.04	0.06	0.06	0.01	0.03	0.03	0.03	0.06	0.06	0.02
40	0.08	0.08	0.05	0.07	0.07	0.02	0.02	0.02	0.04	0.06	0.06	0.03
							$T = 150$					
10	0.05	0.05	0.05	0.06	0.06	0.05	0.06	0.06	0.10	0.06	0.06	0.05
20	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.05	0.13	0.06	0.06	0.06
30	0.05	0.05	0.06	0.07	0.07	0.04	0.04	0.04	0.17	0.06	0.06	0.06
40	0.06	0.06	0.06	0.07	0.07	0.04	0.04	0.04	0.17	0.05	0.05	0.05
							$T = 300$					
10	0.05	0.05	0.06	0.05	0.05	0.06	0.05	0.05	0.13	0.05	0.05	0.08
20	0.05	0.05	0.06	0.05	0.05	0.06	0.05	0.05	0.15	0.05	0.05	0.07
30	0.05	0.05	0.06	0.06	0.06	0.05	0.06	0.06	0.18	0.05	0.05	0.06
40	0.05	0.05	0.04	0.05	0.05	0.04	0.05	0.05	0.17	0.06	0.06	0.05

(b) $r = 2$.

	$\alpha_j = 0.0$, $\rho_i = 0.0$			$\alpha_j = 0.4$, $\rho_i = 0.0$			$\alpha_j = 0.8$, $\rho_i = 0.0$			$\alpha_j = 0.0$, $\rho_i = 0.4$			$\alpha_j = 0.0$, $\rho_i = 0.8$			$\alpha_j = U[0, 0.8]$, $\rho_i = U[0, 0.8]$		
N	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$
									$T = 30$									
10	0.03	0.08	0.03	0.05	0.10	0.02	0.07	0.07	0.01	0.04	0.08	0.01	0.06	0.03	0.00	0.06	0.06	0.00
20	0.03	0.10	0.02	0.05	0.10	0.02	0.08	0.07	0.01	0.03	0.07	0.01	0.06	0.01	0.00	0.06	0.05	0.00
30	0.03	0.11	0.03	0.05	0.09	0.02	0.07	0.07	0.01	0.04	0.05	0.01	0.04	0.00	0.00	0.06	0.02	0.00
40	0.02	0.11	0.03	0.05	0.10	0.02	0.06	0.07	0.01	0.04	0.05	0.00	0.06	0.00	0.00	0.07	0.03	0.00
									$T = 50$									
10	0.03	0.07	0.04	0.05	0.07	0.03	0.08	0.07	0.05	0.04	0.07	0.01	0.05	0.06	0.01	0.05	0.06	0.00
20	0.04	0.09	0.02	0.05	0.09	0.02	0.07	0.09	0.06	0.04	0.08	0.01	0.06	0.02	0.00	0.06	0.07	0.00
30	0.04	0.09	0.01	0.06	0.09	0.02	0.09	0.09	0.05	0.05	0.07	0.00	0.05	0.02	0.00	0.07	0.06	0.00
40	0.03	0.09	0.02	0.05	0.09	0.02	0.08	0.09	0.07	0.03	0.07	0.00	0.06	0.01	0.00	0.05	0.06	0.00
									$T = 75$									
10	0.05	0.06	0.05	0.05	0.06	0.05	0.07	0.06	0.08	0.04	0.06	0.03	0.05	0.07	0.06	0.05	0.06	0.04
20	0.05	0.07	0.03	0.06	0.07	0.04	0.07	0.08	0.11	0.05	0.08	0.01	0.06	0.05	0.03	0.06	0.07	0.01
30	0.05	0.07	0.03	0.05	0.06	0.04	0.07	0.08	0.12	0.04	0.08	0.01	0.05	0.04	0.03	0.05	0.07	0.02
40	0.04	0.09	0.03	0.06	0.08	0.05	0.07	0.08	0.16	0.05	0.09	0.01	0.05	0.02	0.02	0.06	0.06	0.03
									$T = 150$									
10	0.05	0.05	0.06	0.06	0.04	0.08	0.06	0.05	0.11	0.04	0.06	0.06	0.04	0.07	0.09	0.05	0.06	0.08
20	0.05	0.06	0.05	0.06	0.06	0.05	0.06	0.05	0.14	0.05	0.06	0.04	0.05	0.04	0.11	0.05	0.07	0.05
30	0.04	0.06	0.05	0.05	0.05	0.06	0.06	0.06	0.17	0.04	0.05	0.04	0.05	0.04	0.13	0.06	0.06	0.07
40	0.04	0.05	0.05	0.06	0.07	0.05	0.06	0.07	0.18	0.04	0.06	0.04	0.04	0.04	0.14	0.05	0.05	0.05
									$T = 300$									
10	0.04	0.05	0.07	0.05	0.04	0.06	0.06	0.05	0.12	0.05	0.05	0.06	0.04	0.06	0.11	0.06	0.05	0.08
20	0.04	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.13	0.05	0.05	0.05	0.05	0.06	0.15	0.05	0.05	0.09
30	0.05	0.05	0.05	0.05	0.05	0.06	0.06	0.05	0.12	0.05	0.05	0.04	0.05	0.05	0.16	0.06	0.05	0.07
40	0.05	0.05	0.04	0.05	0.06	0.04	0.06	0.06	0.12	0.04	0.05	0.04	0.05	0.06	0.16	0.06	0.06	0.06

(c) $r = 7$.

	$\alpha_j = 0.0$, $\rho_i = 0.0$			$\alpha_j = 0.4$, $\rho_i = 0.0$			$\alpha_j = 0.8$, $\rho_i = 0.0$			$\alpha_j = 0.0$, $\rho_i = 0.4$			$\alpha_j = 0.0$, $\rho_i = 0.8$			$\alpha_j = U[0, 0.8]$, $\rho_i = U[0, 0.8]$		
N	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$
									$T = 30$									
10	0.03	0.08	0.03	0.06	0.08	0.01	0.07	0.04	0.00	0.03	0.09	0.01	0.04	0.07	0.01	0.05	0.07	0.01
20	0.04	0.09	0.03	0.07	0.08	0.01	0.07	0.06	0.00	0.04	0.09	0.01	0.04	0.06	0.00	0.07	0.06	0.00
30	0.04	0.07	0.06	0.08	0.08	0.02	0.07	0.09	0.00	0.05	0.08	0.03	0.04	0.08	0.00	0.07	0.08	0.01
40	0.04	0.06	0.07	0.08	0.08	0.02	0.05	0.07	0.00	0.04	0.09	0.02	0.04	0.08	0.00	0.07	0.07	0.00
									$T = 50$									
10	0.03	0.06	0.06	0.06	0.07	0.04	0.09	0.08	0.04	0.04	0.07	0.03	0.04	0.08	0.02	0.06	0.07	0.03
20	0.04	0.05	0.12	0.04	0.07	0.10	0.07	0.08	0.12	0.04	0.07	0.05	0.05	0.08	0.00	0.06	0.07	0.05
30	0.04	0.05	0.15	0.06	0.06	0.15	0.07	0.10	0.17	0.04	0.07	0.09	0.04	0.07	0.01	0.07	0.07	0.12
40	0.04	0.06	0.20	0.06	0.06	0.16	0.08	0.09	0.19	0.04	0.06	0.11	0.04	0.08	0.01	0.06	0.08	0.15
									$T = 75$									
10	0.05	0.05	0.07	0.06	0.06	0.08	0.07	0.07	0.12	0.04	0.07	0.05	0.05	0.08	0.05	0.06	0.06	0.09
20	0.05	0.05	0.15	0.06	0.06	0.17	0.06	0.08	0.25	0.05	0.06	0.12	0.05	0.07	0.05	0.06	0.07	0.15
30	0.05	0.04	0.17	0.06	0.07	0.19	0.08	0.07	0.26	0.05	0.05	0.10	0.05	0.07	0.05	0.05	0.06	0.19
40	0.05	0.05	0.24	0.05	0.06	0.22	0.07	0.09	0.32	0.05	0.05	0.18	0.05	0.07	0.09	0.06	0.06	0.20
									$T = 150$									
10	0.04	0.06	0.07	0.04	0.06	0.09	0.06	0.06	0.19	0.05	0.06	0.06	0.05	0.06	0.09	0.06	0.06	0.09
20	0.05	0.05	0.17	0.05	0.05	0.19	0.06	0.07	0.27	0.05	0.05	0.15	0.05	0.04	0.11	0.06	0.05	0.20
30	0.05	0.06	0.19	0.04	0.06	0.24	0.06	0.06	0.29	0.05	0.05	0.16	0.06	0.05	0.14	0.06	0.06	0.24
40	0.05	0.04	0.22	0.06	0.05	0.23	0.07	0.07	0.31	0.04	0.05	0.18	0.06	0.05	0.17	0.06	0.05	0.23
									$T = 300$									
10	0.05	0.05	0.10	0.05	0.05	0.11	0.06	0.06	0.20	0.05	0.05	0.10	0.05	0.05	0.13	0.05	0.05	0.13
20	0.05	0.05	0.17	0.05	0.05	0.21	0.06	0.06	0.26	0.05	0.05	0.14	0.05	0.05	0.13	0.05	0.05	0.23
30	0.05	0.04	0.18	0.05	0.05	0.20	0.06	0.06	0.29	0.04	0.05	0.14	0.05	0.05	0.14	0.05	0.05	0.24
40	0.05	0.05	0.18	0.05	0.06	0.22	0.06	0.06	0.28	0.05	0.05	0.17	0.05	0.04	0.15	0.05	0.05	0.25

Table 3: Simulated Power

(a) $r = 0$.

	$s=0.1$			$s=0.2$			$s=0.4$			$s=0.6$		
N	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$
							$T=30$					
10	0.09	0.09	0.03	0.15	0.15	0.01	0.17	0.13	0.00	0.16	0.11	0.00
20	0.15	0.14	0.03	0.15	0.14	0.03	0.15	0.14	0.02	0.12	0.10	0.01
30	0.14	0.14	0.05	0.15	0.14	0.04	0.12	0.11	0.02	0.10	0.08	0.01
40	0.15	0.14	0.04	0.14	0.13	0.04	0.10	0.10	0.02	0.07	0.06	0.01
							$T=50$					
10	0.06	0.06	0.04	0.15	0.25	0.34	0.30	0.40	0.46	0.41	0.47	0.48
20	0.14	0.14	0.06	0.25	0.25	0.11	0.41	0.41	0.24	0.55	0.55	0.46
30	0.18	0.18	0.05	0.29	0.29	0.09	0.48	0.48	0.25	0.64	0.64	0.52
40	0.20	0.20	0.09	0.34	0.34	0.16	0.55	0.55	0.42	0.70	0.70	0.73
							$T=75$					
10	0.06	0.06	0.05	0.22	0.36	0.46	0.55	0.67	0.77	0.74	0.81	0.87
20	0.17	0.17	0.08	0.43	0.43	0.21	0.79	0.79	0.63	0.93	0.93	0.91
30	0.27	0.27	0.11	0.58	0.58	0.32	0.90	0.90	0.83	0.98	0.98	0.98
40	0.32	0.32	0.17	0.66	0.66	0.44	0.95	0.95	0.92	1.00	1.00	1.00
							$T=150$					
10	0.05	0.05	0.05	0.37	0.38	0.26	0.83	0.86	0.88	0.96	0.96	0.98
20	0.27	0.27	0.17	0.73	0.73	0.56	0.98	0.98	0.97	1.00	1.00	1.00
30	0.46	0.46	0.28	0.89	0.89	0.76	1.00	1.00	1.00	1.00	1.00	1.00
40	0.59	0.59	0.38	0.95	0.95	0.87	1.00	1.00	1.00	1.00	1.00	1.00
							$T=300$					
10	0.05	0.05	0.06	0.67	0.67	0.48	0.98	0.98	0.95	1.00	1.00	1.00
20	0.52	0.52	0.30	0.96	0.96	0.85	1.00	1.00	1.00	1.00	1.00	1.00
30	0.79	0.79	0.52	1.00	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00
40	0.92	0.92	0.65	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00

(b) $r = 2$.

	$s=0.1$, $\alpha_j=0$			$s=0.2$, $\alpha_j=0$			$s=0.4$, $\alpha_j=0$			$s=0.6$, $\alpha_j=0$			$s=0$, $\alpha_j=1$		
N	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$
							$T=30$								
10	0.02	0.10	0.03	0.02	0.14	0.01	0.05	0.14	0.00	0.08	0.11	0.00	0.19	0.14	0.00
20	0.03	0.14	0.01	0.04	0.13	0.01	0.08	0.10	0.00	0.09	0.07	0.00	0.18	0.13	0.00
30	0.03	0.14	0.02	0.04	0.11	0.01	0.05	0.07	0.01	0.07	0.05	0.00	0.19	0.14	0.00
40	0.03	0.13	0.02	0.04	0.11	0.01	0.04	0.06	0.00	0.08	0.03	0.00	0.18	0.14	0.01
							$T=50$								
10	0.04	0.07	0.03	0.05	0.27	0.32	0.08	0.41	0.42	0.11	0.48	0.43	0.38	0.35	0.43
20	0.03	0.21	0.23	0.05	0.34	0.36	0.05	0.49	0.52	0.07	0.57	0.57	0.35	0.38	0.55
30	0.05	0.24	0.23	0.07	0.36	0.32	0.11	0.52	0.48	0.15	0.64	0.60	0.39	0.39	0.57
40	0.05	0.23	0.15	0.05	0.38	0.40	0.09	0.57	0.57	0.12	0.68	0.73	0.36	0.40	0.62
							$T=75$								
10	0.04	0.06	0.05	0.06	0.43	0.59	0.10	0.70	0.79	0.19	0.81	0.88	0.55	0.54	0.71
20	0.06	0.29	0.43	0.10	0.60	0.69	0.15	0.87	0.86	0.34	0.95	0.95	0.60	0.60	0.77
30	0.06	0.39	0.49	0.09	0.69	0.72	0.21	0.93	0.92	0.32	0.99	0.99	0.59	0.61	0.81
40	0.06	0.45	0.59	0.08	0.78	0.81	0.17	0.97	0.96	0.25	1.00	1.00	0.59	0.63	0.82
							$T=150$								
10	0.05	0.05	0.07	0.08	0.61	0.74	0.25	0.89	0.96	0.59	0.96	0.99	0.77	0.76	0.89
20	0.08	0.54	0.69	0.17	0.84	0.92	0.42	0.99	1.00	0.67	1.00	1.00	0.79	0.81	0.93
30	0.08	0.63	0.81	0.18	0.94	0.97	0.47	1.00	1.00	0.76	1.00	1.00	0.78	0.81	0.93
40	0.07	0.83	0.92	0.17	0.98	0.99	0.53	1.00	1.00	0.87	1.00	1.00	0.78	0.81	0.93
							$T=300$								
10	0.05	0.06	0.06	0.11	0.83	0.88	0.39	0.99	0.99	0.83	1.00	1.00	0.92	0.95	0.97
20	0.11	0.79	0.87	0.32	0.99	0.99	0.88	1.00	1.00	1.00	1.00	1.00	0.96	0.95	0.98
30	0.16	0.89	0.94	0.46	1.00	1.00	0.93	1.00	1.00	1.00	1.00	1.00	0.95	0.96	0.98
40	0.16	0.98	0.98	0.49	1.00	1.00	0.95	1.00	1.00	1.00	1.00	1.00	0.96	0.97	0.98

Table 4: Simulated Size Adjusted Power

(a) $r = 0$.

N	$s = 0.1$			$s = 0.2$			$s = 0.4$			$s = 0.6$		
	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$
	$T = 30$											
10	0.05	0.06	0.03	0.11	0.11	0.01	0.15	0.10	0.00	0.14	0.08	0.00
20	0.09	0.08	0.01	0.11	0.09	0.02	0.11	0.08	0.01	0.08	0.06	0.01
30	0.07	0.07	0.03	0.10	0.09	0.02	0.07	0.06	0.01	0.04	0.04	0.01
40	0.08	0.07	0.02	0.08	0.07	0.02	0.05	0.04	0.01	0.02	0.02	0.01
	$T = 50$											
10	0.04	0.04	0.03	0.12	0.21	0.32	0.26	0.36	0.43	0.38	0.43	0.45
20	0.10	0.10	0.03	0.19	0.19	0.07	0.34	0.34	0.18	0.46	0.46	0.36
30	0.13	0.13	0.03	0.22	0.22	0.05	0.39	0.39	0.16	0.54	0.54	0.41
40	0.13	0.13	0.05	0.24	0.24	0.08	0.41	0.41	0.29	0.55	0.55	0.60
	$T = 75$											
10	0.05	0.05	0.04	0.19	0.33	0.45	0.51	0.65	0.77	0.72	0.79	0.86
20	0.14	0.14	0.06	0.37	0.37	0.17	0.75	0.75	0.58	0.91	0.91	0.89
30	0.21	0.21	0.08	0.51	0.51	0.26	0.87	0.87	0.78	0.97	0.97	0.98
40	0.26	0.26	0.12	0.59	0.59	0.37	0.93	0.93	0.89	0.99	0.99	1.00
	$T = 150$											
10	0.05	0.05	0.05	0.37	0.38	0.26	0.83	0.86	0.88	0.96	0.96	0.98
20	0.26	0.26	0.16	0.72	0.72	0.55	0.98	0.98	0.97	1.00	1.00	1.00
30	0.42	0.42	0.25	0.88	0.88	0.74	1.00	1.00	1.00	1.00	1.00	1.00
40	0.53	0.53	0.33	0.94	0.94	0.85	1.00	1.00	1.00	1.00	1.00	1.00
	$T = 300$											
10	0.05	0.05	0.06	0.66	0.66	0.48	0.98	0.98	0.95	1.00	1.00	1.00
20	0.50	0.50	0.29	0.96	0.96	0.84	1.00	1.00	1.00	1.00	1.00	1.00
30	0.77	0.77	0.49	1.00	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00
40	0.91	0.91	0.63	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00

(b) $r = 2$.

	$s=0.1$, $\alpha_j=0$			$s=0.2$, $\alpha_j=0$			$s=0.4$, $\alpha_j=0$			$s=0.6$, $\alpha_j=0$			$s=0$, $\alpha_j=1$		
N	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$	$\tilde{S}_k$	$\tilde{S}_k^F$	$\bar{\eta}_\mu^{QS}$
							$T=30$								
10	0.06	0.05	0.02	0.05	0.11	0.01	0.09	0.11	0.00	0.13	0.08	0.00	0.21	0.11	0.00
20	0.05	0.08	0.01	0.06	0.07	0.01	0.10	0.05	0.00	0.11	0.03	0.00	0.19	0.07	0.00
30	0.05	0.07	0.01	0.05	0.06	0.01	0.07	0.03	0.00	0.09	0.02	0.00	0.20	0.07	0.00
40	0.07	0.07	0.01	0.07	0.06	0.01	0.08	0.02	0.00	0.11	0.01	0.00	0.20	0.06	0.00
							$T=50$								
10	0.05	0.05	0.03	0.07	0.24	0.30	0.10	0.37	0.39	0.13	0.44	0.39	0.41	0.31	0.40
20	0.05	0.16	0.20	0.07	0.28	0.32	0.07	0.41	0.45	0.09	0.49	0.51	0.37	0.31	0.51
30	0.06	0.17	0.18	0.08	0.27	0.26	0.12	0.41	0.39	0.17	0.51	0.50	0.41	0.28	0.52
40	0.06	0.15	0.09	0.06	0.26	0.30	0.09	0.41	0.47	0.13	0.50	0.60	0.38	0.26	0.57
							$T=75$								
10	0.05	0.05	0.05	0.07	0.41	0.58	0.11	0.67	0.78	0.21	0.80	0.87	0.57	0.52	0.70
20	0.06	0.26	0.41	0.10	0.57	0.67	0.16	0.84	0.85	0.36	0.94	0.94	0.61	0.57	0.76
30	0.07	0.31	0.46	0.10	0.61	0.68	0.22	0.89	0.90	0.34	0.98	0.98	0.60	0.54	0.80
40	0.07	0.37	0.54	0.08	0.71	0.78	0.18	0.94	0.95	0.26	0.99	1.00	0.60	0.56	0.81
							$T=150$								
10	0.06	0.05	0.07	0.09	0.61	0.74	0.27	0.89	0.96	0.63	0.96	0.99	0.79	0.76	0.89
20	0.08	0.49	0.68	0.17	0.81	0.92	0.43	0.99	1.00	0.68	1.00	1.00	0.79	0.78	0.92
30	0.09	0.61	0.81	0.19	0.94	0.97	0.48	1.00	1.00	0.76	1.00	1.00	0.79	0.79	0.93
40	0.07	0.82	0.92	0.18	0.97	0.99	0.55	1.00	1.00	0.88	1.00	1.00	0.79	0.80	0.93
							$T=300$								
10	0.05	0.06	0.06	0.11	0.83	0.88	0.39	0.99	0.99	0.82	1.00	1.00	0.91	0.95	0.97
20	0.12	0.78	0.87	0.33	0.99	0.99	0.89	1.00	1.00	1.00	1.00	1.00	0.96	0.95	0.98
30	0.15	0.88	0.94	0.45	1.00	1.00	0.92	1.00	1.00	1.00	1.00	1.00	0.95	0.96	0.98
40	0.16	0.97	0.98	0.49	1.00	1.00	0.95	1.00	1.00	1.00	1.00	1.00	0.96	0.96	0.98

Table 5: PPP tests

	Constant		Structural breaks		Break dates		
	$\tilde{S}_k$	p-value	$\tilde{S}_k$	p-value	$\hat{\tau}_{1,i}$	$\hat{\tau}_{2,i}$	$\hat{\tau}_{3,i}$
Austria	1.47	0.071	0.67	0.251	1980:09	1985:03	1987:02
Belgium	1.36	0.087	0.88	0.190	1980:08	1985:03	1987:02
Canada	1.99	0.023	1.28	0.100	1968:12	1976:11	1980:04
Denmark	1.26	0.104	−0.47	0.680	1979:10	1985:03	1988:01
Finland	0.34	0.368	0.16	0.437	1980:10	1985:03	1987:05
France	1.00	0.158	−0.40	0.657	1980:09	1985:03	1987:02
Germany	1.31	0.095	0.25	0.400	1979:12	1985:03	1987:02
Greece	2.36	0.009	−0.54	0.706	1980:02	1985:03	1990:11
Italy	1.07	0.142	1.50	0.067	1980:08	1985:03	1987:02
Japan	2.72	0.003	−0.41	0.660	1985:03	1995:06	2003:03
Netherlands	0.11	0.458	0.13	0.448	1983:01	1985:03	1987:02
Norway	1.09	0.138	0.62	0.266	1980:09	1985:03	1988:01
Portugal	2.57	0.005	1.31	0.095	1980:03	1985:03	1991:03
Spain	1.59	0.056	0.93	0.175	1980:02	1985:03	1987:01
Sweden	1.35	0.089	0.49	0.312	1980:10	1985:03	1988:01
Switzerland	1.46	0.072	0.88	0.188	1979:10	1985:03	1988:01
UK	0.57	0.284	0.91	0.182	1980:11	1985:03	1988:01
Panel, $\tilde{S}_k$	1.93	0.027	1.12	0.132			
Panel, $\tilde{S}_k^F$ $(\hat{r} = 2)$	3.75	0.000	3.20	0.001			