Testing for Stationarity in Large Panels with Cross-Dependence, and US Evidence on Unit Labor Cost*

Matei Demetrescu, Uwe Hassler† and Adina I. Tarcolea
Goethe University Frankfurt‡

This version: May 7, 2009

Abstract

A new stationarity test for heterogeneous panel data with large cross-sectional dimension is developed and used to examine a panel with growth rates of unit labor cost in the US. The test allows for strong cross-unit dependence in the form of unbounded long-run correlation matrices, for which a simple parameterization is proposed. A KPSS-type distribution results asymptotically if letting $T \to \infty$ be followed by $N \to \infty$. Some evidence against stationarity (short memory) is found for the examined series.

Key words: Panel KPSS-type test, cross-correlation, inflation dynamics
JEL Classification: C23, E31

---

*Earlier versions of this paper were presented at the 13th International Conference on Computing in Economics and Finance (June 2007, HEC Montreal), the workshop "Persistence in Economic and Financial Time Series" (June 2007, Frankfurt), at the DAGStat Meeting (March 2007, Bielefeld), and at the Annual Meeting of the German Economic Association (October 2007, Munich). We thank participants of those meetings and in particular Anindya Banerjee and Helmut Lütkepohl for many comments. Further, we are grateful to an anonymous referee for many constructive suggestions and very helpful comments. Financial support by the Deutsche Forschungsgemeinschaft (DFG) through HA-3306/2-1 is gratefully acknowledged.

†Corresponding author: hassler@wiwi.uni-frankfurt.de

‡Statistics and Econometric Methods, Goethe University Frankfurt, Grüneburgplatz 1, D-60323 Frankfurt, Germany, Tel: +49.69.798.34762, Fax: +49.69.798.35014.

1 Introduction

The question of whether inflation rates should be treated as stationary or not is of continued interest, and the international evidence is rather mixed. Some studies apply panel unit root tests to collect support against the null hypothesis of nonstationarity, see e.g. Culver and Papell (1997) and Lee and Wu (2001) using data for 13 OECD countries. We add two aspects to this evidence. First, we examine the null hypothesis of stationarity (or more precisely of integration of order zero) instead of a unit root. Second, we do not examine inflation data directly, but analyze nominal unit labor cost [ULC] series, one of the driving forces of inflation, since it would be economically implausible that inflation and unit labor cost, or their respective growth rates, have different orders of integration. Furthermore, using detailed data on ULC for 50 US states and the District of Columbia [D.C.] enables us to apply more powerful panel methods for inference. This data set, however, has two particular characteristics: on the one hand, the number of cross-sections, N, is large compared to the number of time observations, T; on the other hand, US states are strongly dependent contemporaneously.

The literature on panel stationarity testing is scarce, unlike the literature on unit root testing in panels, see Breitung and Pesaran (2008) for a recent overview of panel unit root tests. They distinguish between first generation tests and second generation tests: the first generation assumes independent units, whereas the second generation allows for different degrees of cross-sectional dependence.

In what concerns stationarity testing, the first generation of panel tests is represented by Hadri (2000), who proposed a Lagrange multiplier [LM] panel test for the null hypothesis that individual time series are stationary around a deterministic level or trend (in fact a panel extension of the KPSS test, thus named after Kwiatkowski *et al.*, 1992). Using sequential asymptotics, $T \to \infty$ followed by $N \to \infty$, asymptotic standard normality of the test statistic is established. Hadri and Larsson (2005) extend this result for any finite T, while Shin and Snell (2006) allow for joint asymptotics, $N, T \to \infty$, with $N/T \to 0$. Hlouskova and Wagner (2006) provide a detailed simulation study regarding the properties of first-generation panel tests (unit root as well as stationarity tests.)

Contributing to the second generation as well, Shin and Snell (2006) also suggest a way to deal with particular types of cross-sectional dependence, namely to subtract cross-sectional means before applying Hadri's test. Nyblom and Harvey (2000) on the other hand do allow for any type of cross-unit dependence in a time series context: they derive a test for the number of common stochastic trends underlying a multivariate random walk, delivering a generalization of the KPSS test. The Nyblom-Harvey test, however, relies on estimation of long-run covariance matrices, which cannot be accomplished reliably without parameter restrictions unless N is much smaller than T. Harris, Leybourne and McCabe (2005) discuss a cross-sectional sum of suitably standardized autocovariances for finite N and obtain limiting normality as $T \to \infty$. Bai and Ng (2004a) suggest for stationarity testing an approach similar to PANIC (Bai and Ng, 2004b), but find its small-sample properties to be less satisfactory. A further possibility would be to combine p values of dependent test statistics from individual units. This is found to work well for moderate N in the case of panel unit root tests (see Demetrescu, Hassler and Tarcolea, 2006, as well as Hartung, 1999).

We address the gap in the existing literature on panel stationarity and tackle the problem of stationarity tests in panels with large N, small to moderate T, and persistent cross-dependence. By "persistent" we understand the norm of the long-run correlation matrix to be unbounded as $N \to \infty$. To this end we propose a simple parameterization of the long-run correlation matrix. This parameterization motivates panel test statistics based on individual-unit KPSS statistics. Using sequential asymptotics, we establish the suggested test statistics to follow Cramér-von Mises type distributions (of which KPSS distributions are a particular case.) These test statistics can also be used to test for panel (co)integration.

This paper is structured as follows. We first discuss the assumed model. Section 3 contains the asymptotic analysis of the suggested stationarity test. Then, the finite sample properties are studied by means of Monte Carlo simulations, while Section 5 is dedicated to the analysis of the growth rates of US unit labor cost. The final section concludes.

2 Model and Assumptions

Let $\{\mathbf{y}_t\}_{t=1,2,\ldots,T}$ denote the observed panel, $\mathbf{y}_t = (y_{1,t}, \ldots, y_{N,t})'$. In the simplest case, the KPSS framework assumes for each unit a component representation,

$$y_{i,t} = r_{i,t} + u_{i,t}, \quad t = 1, 2, \ldots, T, \ i = 1, 2, \ldots, N,$$

where $r_{i,t}$ is a random walk with starting value $r_{i,0}$ and increments having variance σ^2_{ri}, and $u_{i,t}$ is a zero-mean process integrated of order 0, $I(0)$. For simplicity, we let for now $r_{i,0} = 0$. The KPSS test is actually derived as an LM test for the null hypothesis $\sigma^2_{ri} = 0$. Since under this restriction one has $r_{i,t} = 0$ almost surely, $t = 1, 2, \ldots, T$, the null hypothesis is often re-written as

$$y_{i,t} = u_{i,t}, \quad t = 1, 2, \ldots, T, \ i = 1, 2, \ldots, N, \tag{1}$$

where $\mathbf{y}_t = \mathbf{u}_t = (u_{1,t}, \ldots, u_{N,t})'$ is integrated of order 0, see the following assumption. Technical conditions ensuring Assumption 1 are given for instance in Phillips and Durlauf (1986).

Assumption 1 *Let $\Omega = (\omega_{ij})_{1 \le i,j \le N}$ denote the long-run covariance matrix of $\mathbf{u}_t$,*

$$\Omega = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{h=-T+1}^{T-1} E\left(\mathbf{u}_t \mathbf{u}'_{t+h}\right)$$

and assume Ω is finite and positive definite. Further, let weak convergence to multivariate Brownian motion hold as $T \to \infty$ for the partial sums of the process $\mathbf{u}_t$:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} \mathbf{u}_t \Rightarrow \Omega^{0.5} \mathbf{W}(s), \quad s \in [0,1],$$

where "$\Rightarrow$" stands for weak convergence in a suitable metric space of random functions defined on $[0,1]^N$ and $\mathbf{W}(s) = (W_1(s), \ldots, W_N(s))'$ is an N-dimensional vector of independent standard Wiener processes.

Let Ξ denote the long-run correlation matrix of $\mathbf{u}_t$,

$$\Xi = \left(\frac{\omega_{ij}}{\sqrt{\omega_{ii}\omega_{jj}}} \right)_{1 \le i,j \le N} = (\rho_{ij})_{1 \le i,j \le N} .$$

Should the number of units N be of the same magnitude as the number of time observations T, the estimation of Ω without (zero) restrictions is unreliable. Hadri (2000) assumed independent units implying the zero restrictions $\omega_{ij} = 0$ for $i \neq j$. We relax this assumption by allowing for non-zero off-diagonal elements in order to capture cross-dependence. In fact, we impose restrictions on Ξ rather than on Ω, see our next assumption[1].

Assumption 2 *For $i \neq j$, let $\rho_{ij} = \rho$, and $\rho \in (N^{-1}, 1)$.*

This parameterization has been used before, see O'Connell (1998) or the method of combining p values by Hartung (1999); in our context, it allows for a considerable degree of heterogeneity in the panel, single units being allowed to have different long-run variances. Dynamics are also quite general, autocovariance matrices not being directly restricted – only Ξ is.

[1]The constant long-run correlation ρ must belong to the interval $\left(-(N-1)^{-1}, 1\right)$ to ensure positive definiteness of Ξ. Our assumption is slightly stronger and ensures that ρ is nonzero, so that we can divide by ρ in eq. (5).

At the same time we model persistent cross-correlation in the sense that the spectral norm of Ξ is unbounded and growing with N, the largest eigenvalue of Ξ under Assumption 2 being $O(N)$. We understand that Assumption 2 is overly restrictive to be met exactly in practice. As N becomes large, however, it becomes increasingly important to account for persistent cross-correlation, this being the dominating feature of the data. It can in fact be shown that Hadri's (2000) test falsely rejects with probability one in such cases. Thus, even if the long-run correlation matrix does not obey the constant-correlation specification, we expect a considerable degree of the cross-dependence to be captured by allowing for homogeneous non-zero long-run correlation across units, and a test allowing for constant $\rho \neq 0$ will outperform a test assuming constant $\rho = 0$ recovering the case studied by Hadri (2000).

Moreover, we can relax Assumption 2 allowing for a "local random effect model," i.e. we allow for heterogeneous ρ_{ij}, although this heterogeneity may not be permanent:

Assumption 3 *Let* $\rho_{ij} = \rho + N^{-\alpha}\gamma_{ij}$ *for all pairs* $\{i, j\}$ *with* $i \neq j$*, where* $\rho \in (N^{-1}, 1)$, $\alpha > 0$ *and* $\gamma_{ij} = O_p(1)$.

3 Panel stationarity test

3.1 Previous results for independence or finite N

The first approach due to Hadri (2000) and improved by Hadri and Larsson (2005) assumes N independent units and $N \to \infty$. Hadri (2000) basically considers single KPSS test statistics,

$$\kappa_i = \frac{1}{T^2}\sum_{t=1}^{T} S_{i,t}^2, \text{ where } S_{i,t} = \widehat{\omega}_{ii}^{-0.5}\sum_{j=1}^{t} y_{i,j} = \sum_{j=1}^{t} \widetilde{y}_{i,j},$$

with $\widehat{\omega}_{ii}$ a consistent estimator[2] of ω_{ii} (as $T \to \infty$). Under the null hypothesis (1), Assumption 1 and the Continuous Mapping Theorem imply as $T \to \infty$

$$\kappa_i \xrightarrow{d} \int_0^1 W_i^2 (s) \, ds \sim \mathcal{C}v\mathcal{M}_0 \, , \tag{2}$$

where "$\xrightarrow{d}$" stands for convergence in distribution. This limiting distribution is a Cramér-von Mises type functional[3], $\mathcal{C}v\mathcal{M}_0$, see Harvey (2001) for a unifying discussion. Assuming independent units, Hadri (2000) obtains

$$\kappa = \frac{1}{\sqrt{N\sigma^2}} \sum_{i=1}^{N} (\kappa_i - \mu) \xrightarrow{d} \mathcal{N}(0,1)$$

as $N \to \infty$, where $\mu = E\left(\int_0^1 W^2(s)\, ds\right)$, $\sigma^2 = Var\left(\int_0^1 W^2(s)\, ds\right)$. Shin and Snell (2006) argue that, by subtracting cross-sectional means before computing κ, invariance with respect to some forms of nuisance cross-correlation is attained asymptotically.

The second approach allows for dependent panels with arbitrary Ω or Ξ, but it works under finite N. It relies on the vector $\mathbf{S}_t = (S_{1,t}, S_{2,t}, \ldots, S_{N,t})'$, where Assumption 1 implies

$$\frac{1}{\sqrt{T}} \mathbf{S}_{[sT]} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} \widetilde{\mathbf{y}}_t \Rightarrow \Xi^{0.5} \mathbf{W}(s) \, . \tag{3}$$

Consequently, it holds for the trace statistic (which can be expressed as $\kappa_\perp$)

[2]Hadri (2000) works with a non-parametric estimator, see Newey and West (1987) and Andrews (1991), while Shin and Snell (2006) use a semi-parametric estimation method. In fact, Hadri (2000) assumes and estimates a *common* (long-run) variance, which is overly restrictive.

[3]The index 0 reminds us of the fact that no deterministic terms have been fitted i.e. we have a functional over a standard Wiener process.

by Nyblom and Harvey (2000) as $T \to \infty$:

$$\kappa_{\perp} = \frac{1}{T^2} \sum_{t=1}^{T} \mathbf{S}_t' \widehat{\Xi}^{-1} \mathbf{S}_t \xrightarrow{d} \int_0^1 \mathbf{W}(s)' \mathbf{W}(s)\, ds. \tag{4}$$

For finite (or small) N, consistent estimation of Ξ is a routine problem, yielding a limiting distribution free of nuisance parameters in (4).

3.2 Results for large N and constant ρ

In this subsection we propose a sequential asymptotic theory where $T \to \infty$ is followed by $N \to \infty$. Motivated by $\kappa_{\perp}$ from (4) we consider the product $\mathbf{S}_t'\mathbf{S}_t$ without normalization. Consequently the matrix Ξ shows up in the limiting distribution for finite N. More precisely, it holds for any positive integer N as $T \to \infty$ that

$$\begin{aligned} \frac{1}{T^2} \sum_{t=1}^{T} \mathbf{S}_t'\mathbf{S}_t &\xrightarrow{d} \int_0^1 \mathbf{W}(s)' \Xi \mathbf{W}(s)\, ds \\ &= (1-\rho) \int_0^1 \mathbf{W}(s)' \mathbf{W}(s)\, ds + \rho \int_0^1 \mathbf{W}(s)' \iota\iota' \mathbf{W}(s)\, ds\,, \end{aligned}$$

where the equality with $\iota = (1, \ldots, 1)' \in \mathbb{R}^N$ is a consequence of Assumption 2. This motivates the modified test statistic[4]

$$\widetilde{\kappa} = \frac{1}{NT^2\widehat{\rho}} \sum_{t=1}^{T} \mathbf{S}_t'\mathbf{S}_t - \frac{1}{2}\, \frac{1-\widehat{\rho}}{\widehat{\rho}} \tag{5}$$

[4]Note that the test statistic can be written as sum of single test statistics. Hence, it is straightforward to modify the test for unbalanced panels with $T_i \neq T_j$:

$$\frac{1}{N\widehat{\rho}} \sum_{i=1}^{N} \sum_{t=1}^{T_i} \frac{S_{i,t}^2}{T_i^2} - \frac{1}{2}\, \frac{1-\widehat{\rho}}{\widehat{\rho}}\,.$$

with some consistent estimator of ρ from Assumption 2. The limiting distribution of $\widetilde{\kappa}$ coincides with the well-known KPSS-type distribution $\mathcal{C}v\mathcal{M}_0$ given in (2).

Proposition 1 *Given model (1), it holds for $\widetilde{\kappa}$ from (5) under Assumptions 1 and 2*

$$\widetilde{\kappa} \xrightarrow{d} \mathcal{C}v\mathcal{M}_0,$$

as $T \to \infty$ followed by $N \to \infty$, if a consistent (as $N \to \infty$) estimator $\widehat{\rho}$ is used.

Proof: See Appendix.

Consistency of the panel test based on $\widetilde{\kappa}$ is guaranteed if there is at least one stochastic trend in the panel, be it unit-specific or common to several units. Indeed, it follows from Nyblom and Harvey (2000) for finite N that $\widetilde{\kappa} \to \infty$ in probability in such cases. And due to the use of sequential asymptotics, $T \to \infty$ followed by $N \to \infty$, this holds in the limit as well.

Remark 1 *The suggested test has power against fractional alternatives, as it inherits the properties of the univariate KPSS test, and it also has power against structural breaks, see Lee and Schmidt (1996), and Lee, Huang and Shin (1997), respectively, for the corresponding analyses of the KPSS test.*

3.3 Consistent estimation of ρ

At this stage, an algorithm for the consistent estimation of ρ is required. We adopt a proposal by Hartung (1999), which is based on the limiting normality of the partial sum process $\mathbf{S}_T$ as $T \to \infty$ implied by Assumption 1, see (3):

$$\frac{1}{\sqrt{T}}\mathbf{S}_T \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \Xi\right). \tag{6}$$

Based on properties of quadratic forms of a multivariate normal random vector with covariance matrix Ξ, Hartung (1999) puts forward the following steps:

1. Compute the sample variance q of the N-dimensional vector $T^{-0.5}\mathbf{S}_T$,[5]

$$q = \frac{1}{(N-1)\,T} \sum_{i=1}^{N} \left(S_{i,T} - \overline{S}_T\right)^2 , \tag{7}$$

   where this quadratic form is related to a χ^2 distribution with $N-1$ degrees of freedom:

$$\frac{N-1}{1-\rho}\, q \sim \chi^2(N-1).$$

2. Let[6]

$$\widehat{\rho}^{*} = \max\left(N^{-0.5}, 1-q\right).$$

It then holds $\widehat{\rho}^{*} = \rho + O_p\left(N^{-0.5}\right)$, since

$$E(q) = 1-\rho \text{ and } Var(q) = \frac{2\,(1-\rho)^2}{N-1}.$$

The estimator $\widehat{\rho}^{*}$ is consistent under Assumption 2. For estimates $\widehat{\rho}^{*} \approx 0$ one may work with Hadri's (2000) test in practice. Further, we adopt Hartung's (1999) correction that improves the performance in small samples:

$$\widehat{\rho} = \max\left(N^{-0.5}, 1-q\right) + 0.2\sqrt{\frac{2}{N-1}}\left(1 - \max\left(N^{-0.5}, 1-q\right)\right). \tag{8}$$

[5] In heterogeneous panels with differing time dimensions T_i, one may modify q as follows:

$$q = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{S_{i,T_i}}{\sqrt{T_i}} - \frac{1}{N} \sum_{i=1}^{N} \frac{S_{i,T_i}}{\sqrt{T_i}} \right)^2 .$$

[6] Hartung (1999) originally proposed $\widehat{\rho}^{*} = \max\left(-(N-1)^{-1}, 1-q\right)$. Our modification avoids $\widehat{\rho}^{*} = 0$, see footnote 1.

Although this procedure was developed for Ξ as given in Assumption 2, the local effects allowed by Assumption 3 influence neither the consistency of $\widehat{\rho}$ nor the asymptotic distribution of the suggested test statistic. The following proposition formalizes this result.

Proposition 2 *Under the assumptions of Proposition 1 but with Assumption 2 replaced by Assumption 3 and with $\widehat{\rho}$ from (8), the limiting theory from Proposition 1 continues to hold.*

Proof: See Appendix.

3.4 Deterministic components

We now generalize (1) with respect to deterministic components under the null hypothesis of integration of order zero. In particular, a constant or a linear time trend are considered ($t = 1, 2, \ldots, T$):

$$y_{i,t} = c_i + u_{i,t} \, , \tag{9}$$

$$y_{i,t} = c_i + \gamma_i \, t + u_{i,t} \, . \tag{10}$$

In the presence of deterministic components, estimating ω_{ii}, as well as computing the cumulated sums $S_{i,t}$, is simply based on demeaned or detrended observations, respectively. The estimation of ρ, however, requires additional care. Classical demeaning or detrending of the data by means of ordinary least squares [OLS] does not yield valid estimators of ρ in (8), because these rely on normality of the partial sums, see (6). To circumvent this problem and retain normality, we suggest to demean or detrend the data recursively for estimation of ρ, where recursive demeaning and detrending means[7]

$$\begin{aligned} y_{i,t}^{\mu} &= y_{i,t} - \frac{1}{t}\sum_{j=1}^{t} y_{i,j} \\ y_{i,t}^{\tau} &= y_{i,t} + \frac{2}{t}\sum_{j=1}^{t} y_{i,j} - \frac{6}{t\,(t+1)}\sum_{j=1}^{t} j y_{i,j}\,. \end{aligned}$$

Now, we standardize with the respective long-run variance estimates $\widehat{\omega}_{ii}$ to obtain the vectors $\widetilde{\mathbf{y}}_t^{\mu}$ and $\widetilde{\mathbf{y}}_t^{\tau}$ (note that $S_{i,t}$ and $\widehat{\omega}_{ii}$ are still computed using usual demeaning or detrending). Then, we have the following lemma.

Lemma 3 *Denote $\widetilde{y}_{i,t}^{\mu} = \widehat{\omega}_{ii}^{-0.5} y_{i,t}^{\mu}$ and $\widetilde{y}_{i,t}^{\tau} = \widehat{\omega}_{ii}^{-0.5} y_{i,t}^{\tau}$. Then, given the models (9) and (10) and Assumption 1, it holds as $T \to \infty$ that*

$$\frac{1}{\sqrt{T}}\mathbf{S}_T^{\mu} := \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \widetilde{\mathbf{y}}_t^{\mu} \xrightarrow{d} \mathcal{N}(0,\Xi)\,,$$

$$\frac{1}{\sqrt{T}}\mathbf{S}_T^{\tau} := \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \widetilde{\mathbf{y}}_t^{\tau} \xrightarrow{d} \mathcal{N}(0,\Xi)\,.$$

Proof: See Appendix.

Now, we can use $\mathbf{S}_T^{\mu}$ or $\mathbf{S}_T^{\tau}$ defined implicitly in Lemma 3 to replace $\mathbf{S}_T$ in (7). This defines q^{μ} or q^{τ}, and replacing q in (8), we obtain $\widehat{\rho}^{\mu}$ or $\widehat{\rho}^{\tau}$, respectively. Relying on the limiting normality with covariance matrix Ξ, those estimators of ρ are consistent. The resulting test statistics are

$$\widetilde{\kappa}^{\mu} = \frac{1}{NT^2\widehat{\rho}^{\mu}}\sum_{t=1}^{T} \mathbf{S}_t'\mathbf{S}_t - \frac{1}{6}\,\frac{1-\widehat{\rho}^{\mu}}{\widehat{\rho}^{\mu}} \qquad (11)$$

[7] One could allow for other deterministic components than "just" a mean or a linear trend, e.g. for a shift in the mean. Along these lines, Demetrescu (2009) analyzes the properties of recursive adjustment for general deterministic components.

and

$$\widetilde{\kappa}^{\tau} = \frac{1}{NT^2\widehat{\rho}^{\tau}} \sum_{t=1}^{T} \mathbf{S}_t'\mathbf{S}_t - \frac{1}{15}\,\frac{1-\widehat{\rho}^{\tau}}{\widehat{\rho}^{\tau}}, \tag{12}$$

respectively, where $\mathbf{S}_t$ must not be confused with $\mathbf{S}_T^{\mu}$ or $\mathbf{S}_T^{\tau}$. $\mathbf{S}_t$ constructed from data demeaned or detrended the usual way leads to LBI tests, while the use of $\mathbf{S}_T^{\mu}$ or $\mathbf{S}_T^{\tau}$ would not. The limiting distributions involve Cramér-von Mises type functionals with standard (first level) Brownian bridges and so-called second level Brownian bridges, respectively, in short $\mathcal{C}v\mathcal{M}_1$ and $\mathcal{C}v\mathcal{M}_2$. Detailed critical values are given in the first two columns of Table 2 in MacNeill (1978) (see also Anderson and Darling, 1952, or Kwiatkowski *et al.*, 1992). Further, the expected values of $\mathcal{C}v\mathcal{M}_1$ and $\mathcal{C}v\mathcal{M}_2$ are given as $\frac{1}{6}$ and $\frac{1}{15}$, respectively, in MacNeill (1978), which explains the corresponding factors in (11) and (12).

Proposition 4 *Under the assumptions of Lemma 3 together with Assumption 3 it holds for* $\widetilde{\kappa}^{\mu}$ *and* $\widetilde{\kappa}^{\tau}$ *that*

$$\widetilde{\kappa}^{\mu} \xrightarrow{d} \mathcal{C}v\mathcal{M}_1,$$

$$\widetilde{\kappa}^{\tau} \xrightarrow{d} \mathcal{C}v\mathcal{M}_2,$$

provided that $T \to \infty$ *followed by* $N \to \infty$.

Proof: Lemma 3 ensures consistent estimators of ρ. The rest of the proof follows the arguments establishing Propositions 1 and 2.

3.5 Testing for unit roots

Our method can also be applied to test for a unit root in cross-correlated panels, since Lee and Schmidt (1996) established consistency of the KPSS test against integration of order -1. This suggests to build differences of order one of the data,[8] to build the test statistic $\widetilde{\kappa}$ defined in (5) and reject for *too*

[8]A similar idea was put forward in a different context by Phillips and Ouliaris (1990).

small values of the test statistic. Naturally, the estimator $\widehat{\rho}$ from (8) can be used, if computed with differenced data. The result similar to Propositions 1 and 2 is obvious.

Corollary 5 *If $\mathbf{y}_t \sim I(1)$, i.e. $\Delta \mathbf{y}_t = \mathbf{u}_t$ from Assumptions 1 and 2 (or 3, respectively), it holds for $\widetilde{\kappa}$ based on $\Delta \mathbf{y}_t$ as $T \to \infty$ followed by $N \to \infty$ that*

$$\widetilde{\kappa} \xrightarrow{d} \mathcal{C}v\mathcal{M}_0,$$

if a consistent (as $N \to \infty$) estimator $\widehat{\rho}$ is used.

Proof: obvious and omitted.

Remark 2 *Deterministic components are easier to deal with than in the case of stationarity testing, since, due to differencing, a non-zero mean cancels out, and a linear trend turns into a non-zero mean.*

Remark 3 *Panel cointegration testing under cross-unit dependence is still a thorny issue. Our panel stationarity test could be applied to residual series from static regressions in single units. If using* fully modified OLS *or* dynamic OLS *when estimating the cointegration residuals, this should lead to test statistics for the null of cointegration having the same asymptotic properties as $\widetilde{\kappa}$ defined in (5). However, a rigorous treatment of this topic is beyond the scope of this paper.*

4 Small-sample behaviour

For studying the small-sample properties of our test, we simulate panels with $N \in \{10, 20, 50, 100, 500\}$ units and $T \in \{20, 50, 100, 250\}$ time observations. All simulations are carried out in R with 1000 replications for each case. The data is generated as follows. Each unit follows an MA process with the same parameter $\theta = 0.5$, and the respective innovations $\epsilon_t = (\epsilon_{1,t}, \ldots, \epsilon_{N,t})'$ are standard normal pseudo-random numbers exhibiting cross-correlation, $\Xi = (\rho_{ij})$ with $\rho_{ii} = 1$ and $\rho_{ij} = \rho \in \{0.2, 0.5, 0.8\}$ for $i \neq j$:

$$\mathbf{u}_t = \epsilon_t + \theta \epsilon_{t-1}. \tag{13}$$

For the case where a deterministic mean is to be accounted for, we examine our test ($\widetilde{\kappa}^\mu$) together with the test due to Hadri (2000) with cross-sectional demeaning following Shin and Snell (2006, κ^μ) and the test obtained by combining significance of single KPSS statistics from each unit (κ^p), see Demetrescu, Hassler and Tarcolea (2006). The p values are obtained from the empirical cumulative distribution function of a KPSS statistic with demeaning simulated for $T = 100$ with 100000 replications, no short-run correlation and the usual variance estimator.

Estimation of single-unit long-run variances ω_{ii} is done nonparametrically. We use the quadratic spectral kernel shown by Andrews (1991) to have certain optimality properties. Following Kwiatkowski *et al.* (1992), who use a bandwidth $b = \left[4\,(T/100)^{0.25}\right]$, we choose the bandwidth as a (slightly modified) deterministic function of T. More precisely, we set $b = \left[4\,(T/100)^{0.2}\right]$, accounting for the optimal rate $b = O\left(T^{0.2}\right)$ established by Andrews (1991).[9] The results for the 5% significance level are as follows.

Examining Tables 1 through 3, one observes the modification of Hadri's test (2000) due to Shin and Snell (2006) to be oversized in all cases.[10] While the distortions do not depend on the magnitude of the cross-correlation ρ, they diminish with increasing T and increase with growing N. This suggests that independence of ρ is attained, but the $\mathcal{C}v\mathcal{M}_1$ distribution is not a very good approximation of the small-sample distribution of single tests based on cross-sectionally demeaned data. Given the form of κ^μ, this most likely

[9] We avoid data-driven methods of bandwidth choice, since these are not as reliable as deterministic rules with KPSS-type tests, see Hlouskova and Wagner (2006).

[10] The test not accounting for cross-correlation suffers from tremendous size distortions towards the alternative and we do not report its behavior.

Table 1. Empirical size of $\widetilde{\kappa}^{\mu}$, κ^{μ} and κ^{p} for $\rho = 0.2$

	$T = 20$			$T = 50$			$T = 100$			$T = 250$		
N	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}
10	5.8	15.5	1.1	9.1	12.7	2.0	9.9	10.0	1.5	16.7	9.1	2.8
20	3.6	23.2	1.0	5.8	13.2	2.3	9.6	11.8	2.5	13.4	10.9	3.4
50	2.2	41.2	0.3	5.7	19.3	3.3	5.4	12.6	3.1	10.0	9.8	4.1
100	1.7	66.9	0.2	4.1	32.0	2.7	5.8	18.5	5.0	7.7	12.5	5.0
500	0.7	100.	0.2	3.3	84.1	2.8	4.3	55.1	6.7	5.8	38.9	10.1

Note: $\widetilde{\kappa}^{\mu}$ denotes the test statistic from (11), κ^{μ} Hadri's (2000) test with Shin and Snell's (2006) cross-sectional demeaning and κ^{p} the test obtained by combining p values of KPSS statistics from each unit; nominal size is 5%.

Table 2. Empirical size of $\widetilde{\kappa}^{\mu}$, κ^{μ} and κ^{p} for $\rho = 0.5$

	$T = 20$			$T = 50$			$T = 100$			$T = 250$		
N	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}
10	5.1	16.3	2.6	6.0	11.2	3.7	7.0	9.8	5.5	8.0	9.1	5.6
20	3.0	23.8	1.4	4.5	15.9	5.1	5.3	9.1	5.2	6.1	9.8	6.5
50	2.3	43.5	2.1	3.0	20.8	5.6	4.2	14.3	6.4	5.3	12.6	9.4
100	1.2	64.5	1.9	3.5	31.9	6.1	3.2	15.6	6.8	5.7	14.7	11.2
500	1.6	99.8	1.9	3.6	85.7	5.0	3.3	53.0	7.7	6.3	35.8	14.5

Note: see Table 1 for details

arises from the fact that the moments (and in particular the mean) are not approximated precisely enough; the effects of this imprecision amplify with growing N.

Regarding the approach of combining p values, the respective test is undersized for very small T. For larger sample sizes, the size is better held as ρ grows, with the exception of the case $N = 500$, where κ^{p} is oversized for some of the studied cases. As a rule of thumb, the larger ρ, the better the behavior of the test, see Table 3.

The same holds for our test, $\widetilde{\kappa}^{\mu}$, which also works best in the presence of large cross-correlation ($\rho = 0.8$), although it is somewhat undersized for

Table 3. Empirical size of $\widetilde{\kappa}^\mu$, κ^μ and κ^p for $\rho = 0.8$

	$T=20$			$T=50$			$T=100$			$T=250$		
N	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p
10	2.8	17.3	2.4	5.2	10.7	5.7	3.8	9.5	4.0	4.7	10.4	5.1
20	3.4	24.9	3.4	3.2	13.8	2.9	4.7	12.1	4.9	4.5	9.7	4.6
50	2.9	43.9	2.7	4.2	20.5	4.5	4.4	14.3	4.5	5.2	9.8	5.5
100	3.5	67.3	3.6	4.8	32.9	5.1	5.6	19.6	5.7	4.2	15.8	4.6
500	2.5	100.	2.6	4.1	83.3	4.5	4.5	52.5	4.8	5.7	33.4	5.9

Note: see Table 1 for details

$T = 20$. This can be explained by the fact that the variance of $\widehat{\rho}$ is smaller when ρ lies closer to 1, see Section 3.3. For small ρ, the test's behavior is less reliable: $\widetilde{\kappa}^\mu$ is undersized for $N \gg T$ and oversized for $T \gg N$, see Table 1. In the case of medium cross-correlation, the properties of the test are good, except for the case with $T = 20$ and large N, see Table 2.

In case of negative MA roots, e.g. $\theta = -0.5$ in (13), all tests become very conservative in that the experimental significance level is considerably below the nominal one. Further results are not reported here but are available upon request.

To gauge Proposition 3 under Assumption 3, we next allow for varying correlation depending on unit i and j, specifically

$$\rho_{ij} = 0.4 + 0.6(1 - |i - j|/N)$$

with considerable variation between $0.4 + 0.6\,\frac{1}{N} \le \rho_{ij} \le 1$. In Table 4 one finds simulation findings comparable to Tables 1 through 3. The resulting size properties are as good as (or even better than) in case of constant ρ.

In what concerns the alternative, we add to each unit a random walk component $r_{i,t}$,

$$\mathbf{y}_t = \mathbf{r}_t + \mathbf{u}_t, \quad t = 1, 2, \ldots, T, \tag{14}$$

where the components $r_{i,t}$ of $\mathbf{r}_t$ are independent and their increments have

Table 4. Empirical size of $\widetilde{\kappa}^{\mu}$, κ^{μ} and κ^{p} for varying ρ_{ij}

	$T=20$			$T=50$			$T=100$			$T=250$		
N	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}
10	4.6	21.9	4.1	5.8	18.1	4.9	5.9	17.4	5.2	6.6	17.2	6.1
20	4.2	28.5	3.9	5.6	24.2	5.5	5.8	22.6	5.7	6.5	21.8	6.3
50	4.3	35.7	4.2	5.2	31.0	5.5	5.7	27.4	5.9	6.4	28.2	7.0
100	4.4	40.7	4.5	5.6	34.5	6.1	5.9	31.0	6.4	6.5	30.8	7.3
500	4.1	48.2	4.5	5.7	41.1	6.4	6.3	37.9	6.8	5.9	35.6	7.1

Note: see Table 1 for details

very small variance relative to the $I(0)$ components $u_{i,t}$, $\sigma_{ri}^2 = 0.01$. The simulation results are provided in Tables 5 to 7.

Table 5. Empirical power of $\widetilde{\kappa}^{\mu}$, κ^{μ} and κ^{p} for $\rho = 0.2$

	$T=20$			$T=50$			$T=100$			$T=250$		
N	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}	$\widetilde{\kappa}^{\mu}$	κ^{μ}	κ^{p}
10	11.8	29.4	1.8	65.2	73.3	24.6	98.5	98.7	86.9	100	100	100
20	8.5	46.8	1.7	74.0	92.2	33.0	99.9	100	97.7	100	100	100
50	5.6	75.5	0.9	89.0	100	49.9	100	100	100	100	100	100
100	5.6	96.2	0.8	96.2	100	63.2	100	100	100	100	100	100
500	4.6	100	0.0	99.9	100	80.2	100	100	100	100	100	100

Note: Data exhibits weak stochastic trends, see (14) and Table 1 for further details

Due to size distortions of κ^{μ}, the results regarding its power are not meaningful. Our test, $\widetilde{\kappa}^{\mu}$, exhibits altogether good power properties, with the exception of the case $T = 20$, where the rejection frequencies are only marginally higher than the empirical size (which should not come as a surprise, considering that the studied alternative lies very close to the null). Otherwise, $\widetilde{\kappa}^{\mu}$ is significantly more powerful than κ^{p}, which is best seen for medium sample sizes; however, the power advantage of $\widetilde{\kappa}^{\mu}$ over κ^{p} somewhat decreases with increasing ρ.

To sum up, our test works increasingly well with growing N and ρ. For

Table 6. Empirical power of $\widetilde{\kappa}^\mu$, κ^μ and κ^p for $\rho = 0.5$

	$T = 20$			$T = 50$			$T = 100$			$T = 250$		
N	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p
10	7.7	36.5	3.8	53.2	85.7	22.1	97.3	99.9	84.0	100	100	100
20	7.2	55.3	3.8	63.3	98.8	31.0	99.9	100	96.9	100	100	100
50	6.9	89.3	4.4	76.1	100	43.4	100	100	100	100	100	100
100	3.6	99.3	2.2	79.0	100	48.2	100	100	100	100	100	100
500	4.3	100	3.3	87.0	100	53.4	100	100	100	100	100	100

Note: See Table 5 for details

Table 7. Empirical power of $\widetilde{\kappa}^\mu$, κ^μ and κ^p for $\rho = 0.8$

	$T = 20$			$T = 50$			$T = 100$			$T = 250$		
N	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p	$\widetilde{\kappa}^\mu$	κ^μ	κ^p
10	4.6	58.3	3.6	46.0	98.6	22.3	97.7	100	85.3	100	100	100
20	3.8	83.4	3.8	48.5	100	25.3	99.8	100	94.0	100	100	100
50	4.8	99.3	4.7	55.0	100	34.4	100	100	99.5	100	100	100
100	4.3	100	4.7	59.0	100	40.4	100	100	99.8	100	100	100
500	4.0	100	4.6	65.2	100	41.4	100	100	99.9	100	100	100

Note: See Table 5 for details

some combinations of N and T, the approach of combining p values also has merits, while Shin and Snell's (2006) robustification to cross-correlation of Hadri's (2000) test leads to spurious rejection of the null for panels of usual sizes.

5 Growth rates of US unit labor cost

Let us turn our attention to the question of whether nominal US unit labor cost should be treated as $I(0)$ or not. We test the null hypothesis in a panel of 50 US states and District of Columbia with annual observations from 1977 to 1997. The data was obtained from Fritsche and Kuzin (2007), and is based on data sets of the Bureau of Economic Analysis. Observations from

1998 to 2000 are unfortunately not available due to a change of the industrial classification system in 1997.

The nominal unit labor cost is computed as

$$Y = \frac{\text{nominal compensation of employees}}{\text{real gross state product}}.$$

The growth rates examined are log-differences,

$$y_{i,t} = \Delta \log \left(Y_{i,t} \right).$$

Figure 1 plots the growth rates, 1978 to 1997, and shows a considerable degree of co-movement. The obvious "outlier" is Alaska, having a significantly larger variance than other US states.

The test statistics are computed from demeaned growth rates with the quadratic spectral window and bandwidth $b = \left[4\left(T/100\right)^{0.2}\right]$, or $b = 2$ given $T = 20$. The empirical finding of our new test is

$$\widetilde{\kappa}^{\mu} = 0.690,$$

where an estimated $\widehat{\rho}^{\mu} = 0.753$ emerges. The corresponding 95% and 99% percentiles from the Cramér-von Mises distribution are $P(\mathcal{C}v\mathcal{M}_1 \leq 0.461361) = 0.95$ and $P(\mathcal{C}v\mathcal{M}_1 \leq 0.743458) = 0.99$, indicating significant evidence against integration of order zero at the 5%, but not at the 1% level.[11] Note, however, that we can not tell how pervasive the deviation from the null hypothesis is. To get evidence which or how many series violate the null, one might try to adopt a multiple testing strategy advised by Hanck (2009) for panel unit root tests.

In order to make sure our results are robust to choice of bandwidth b, we repeated the analysis with $b = 1$ and $b = 3$. The test based on $\widetilde{\kappa}^{\mu}$ is

[11] We do not report the other studied tests since these are not sufficiently reliable in terms of size, see Section 4.

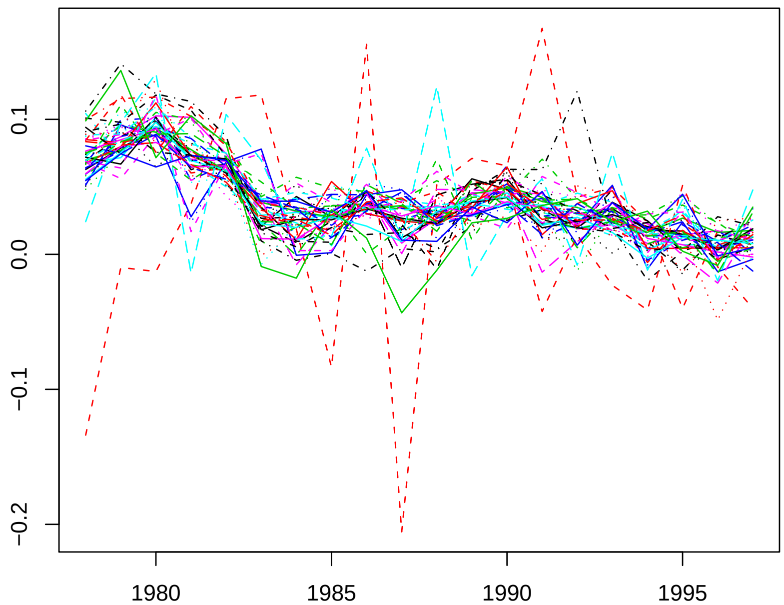

Figure 1. ULC annual growth rates in 50 US states and D.C., 1978-1997

significant at the 1% level for $b = 1$, $\widetilde{\kappa}^{\mu} = 1.40$, but only at the 5% level for $b = 3$, $\widetilde{\kappa}^{\mu} = 0.502$. Repeating the analysis without Alaska, we did not find any qualitative differences. Note, however, that in the empirical example we have $N = 51$ and $T = 20$, where we know from the previous tables that the test is slightly undersized under the null hypothesis and not very powerful under the alternative. This may be the reason why rejection at the 1% level is not robust with respect to the choice of the bandwidth.

Our findings do not necessarily mean the growth rates are integrated of order one. This would be inconsistent with the results of Fritsche and Kuzin (2007), who, by means of PANIC analysis of the unit labor cost data, find no significant evidence against integration of order one. Rather, the alternative

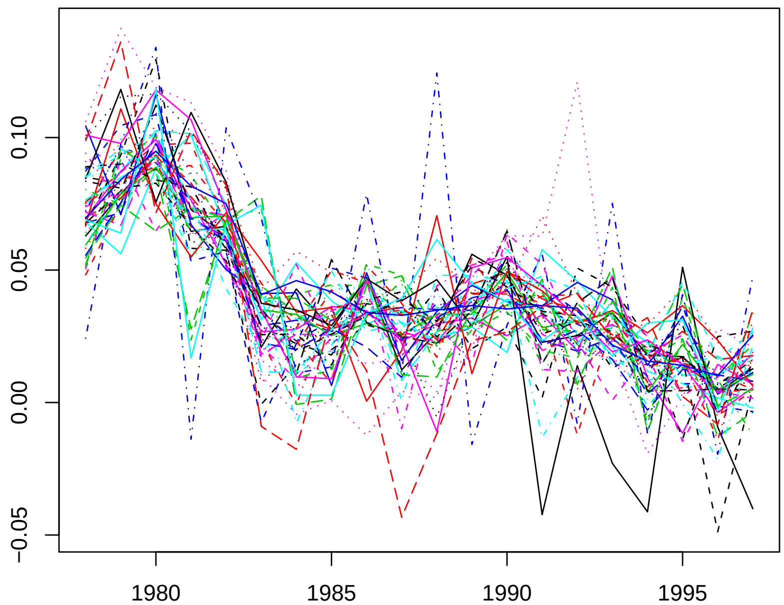

Figure 2. ULC annual growth rates in 49 US states (without Alaska) and D.C., 1978-1997

might be fractional integration, against which our test has power.

Another perhaps more realistic explanation may be a change of the overall economic environment; recall that the late 70's and early 80's were marked by stagflation with high inflation rates, which decreased in the late 80's and even more in the 90's. To capture this aspect, we tested the detrended growth rates. Indeed, a trend is more visible in Figure 2 without Alaska than in Figure 1. In this case, the evidence against (trend-) stationarity is much weaker: $\widetilde{\kappa}^{\tau}$ yields a value of 0.111, which isn't significant, even at the 10% level ($P(\mathcal{C}v\mathcal{M}_2 \leq 0.119220) = 0.90$, $P(\mathcal{C}v\mathcal{M}_2 \leq 0.147891) = 0.95$). Varying the bandwidth, some evidence against stationarity appears only for $b = 1$

($\widetilde{\kappa}^{\tau} = 0.168$, significant at 5%). The same picture arises when testing the panel without Alaska.

6 Summary

In order to examine unit labor cost data for the US states, a panel stationarity test for panels with large cross-sectional dimension in presence of cross-unit correlation was proposed and analyzed. More precisely, the test has the null hypothesis of integration of order zero.

By using sequential asymptotics, $T \to \infty$ followed by $N \to \infty$, a distribution belonging to the Cramér-von Mises family of distributions (of which the KPSS distribution is a member, too) was established for our test. The test allows for persistent cross-correlation in form of unbounded norm of the long-run correlation matrix of the panel and performs well in small samples compared to other panel stationarity tests allowing for cross-sectional dependence.

Applying this test to annual unit labor cost data for 50 US states and D.C., evidence against stationarity of their growth rates was found for the time period 1977-1997. The strength of the evidence depends on the deterministic component we modeled; allowing for a time trend in growth rates to capture changes in overall economic environment weakens the case against (trend-) stationarity of unit labor cost (and thus inflation) growth rates.

Appendix

Proof of Proposition 1

It holds

$$\frac{1}{T^2}\sum_{t=1}^{T}\mathbf{S}_t'\mathbf{S}_t \xrightarrow{d} (1-\rho)\int_0^1 \mathbf{W}(s)'\,\mathbf{W}(s)\,ds + \rho\int_0^1 \mathbf{W}(s)'\,\iota\iota'\mathbf{W}(s)\,ds. \quad (15)$$

The first term on the right-hand side amounts to a sum of squares. Consequently, a weak Law of Large Numbers applies,

$$\frac{1}{N}\int_0^1 \mathbf{W}(s)'\,\mathbf{W}(s)\,ds = \frac{1}{N}\sum_{i=1}^{N}\int_0^1 W_i^2(s)\,ds \xrightarrow{p} \mu := E\left(\int_0^1 W_i^2(s)\,ds\right),$$

where the expectation μ equals $\frac{1}{2}$. To study the second term in (15), we observe that

$$\frac{\mathbf{W}(s)'\,\iota\iota'\mathbf{W}(s)}{N} = \frac{\left(\sum_{i=1}^{N} W_i(s)\right)^2}{N} = \widetilde{W}^2(s)$$

where

$$\widetilde{W}(s) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} W_i(s).$$

Note that $\widetilde{W}(s)$ is a standard Wiener process for any $N \in \mathbb{N}$, when W_i are independent, since $\widetilde{W}(0) = 0$ with probability one, the increments $\widetilde{W}(s_2) - \widetilde{W}(s_1) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(W_i(s_2) - W_i(s_1))$ follow a normal distribution with mean 0 and variance $s_2 - s_1$, and the increments $\widetilde{W}(s_2) - \widetilde{W}(s_1)$ are independent for non-overlapping time intervals $[s_1, s_2]$ and $[s_1^*, s_2^*]$. Hence,

$$\frac{1}{N}\int_0^1 \mathbf{W}(s)'\,\iota\iota'\mathbf{W}(s)\,ds = \int_0^1 \widetilde{W}^2(s)\,ds = \mathcal{C}v\mathcal{M}_0.$$

For $T \to \infty$ we thus obtain

$$\widetilde{\kappa} \xrightarrow{d} \frac{1-\rho}{\widehat{\rho}} \frac{1}{N} \sum_{i=1}^{N} \int_0^1 W_i^2(s) ds + \frac{\rho}{\widehat{\rho}} \mathcal{C} v \mathcal{M}_0 - \frac{1}{2} \frac{1-\widehat{\rho}}{\widehat{\rho}}.$$

At this stage, plugging in a consistent (as $N \to \infty$) estimator of ρ poses no problems, and the proof is completed by sequential asymptotics.

Proof of Proposition 2

The consistency of the estimator $\widehat{\rho}$ follows directly from Demetrescu, Hassler and Tarcolea (2006, Proposition 1).

For the result to hold, we only need to show that the relation

$$\frac{1}{T^2} \sum_{t=1}^{T} \mathbf{S}_t' \mathbf{S}_t \xrightarrow{d} (1-\rho) \int_0^1 \mathbf{W}(s)' \mathbf{W}(s)\, ds + \rho \int_0^1 \mathbf{W}(s)' \iota\iota' \mathbf{W}(s)\, ds + o_p(N) \tag{16}$$

holds under Assumption 3.

Denote $\Gamma = (\gamma_{ij})_{i,j}$ and $\Xi_\rho = \{\rho\}_{i \neq j}$ the correlation matrix from Assumption 2. Then, $\Xi = \Xi_\rho + N^{-\alpha}\Gamma$ and

$$\frac{1}{T^2} \sum_{t=1}^{T} \mathbf{S}_t' \mathbf{S}_t \xrightarrow{d} \int_0^1 \mathbf{W}(s)' \Xi_\rho \mathbf{W}(s)\, ds + N^{-\alpha} \int_0^1 \mathbf{W}(s)' \Gamma \mathbf{W}(s)\, ds.$$

Let $\|\cdot\|$ denote the spectral matrix norm. Note that $\|\Gamma\| = O(N)$, since the used spectral norm is bounded by the square root of the product of the column-sum norm and the row-sum norm, and these are obviously $O(N)$. Further, $\|\Gamma^2\| < \|\Gamma\| \|\Gamma\| = O(N^2)$. Since $\mathbf{W}(s)$ follows a multivariate normal distribution with mean zero and sI_N covariance matrix, where I_N denotes the $N \times N$ unity matrix, the quadratic form $\mathbf{W}(s)' \Gamma \mathbf{W}(s)$ will have variance

$2tr\left(s^2\Gamma I_N \Gamma I_N\right)$, which translates for finite s in

$$\mathbf{W}(s)' \Gamma \mathbf{W}(s) = O_p\left(\sqrt{\|\Gamma^2\|}\right) = O_p(N).$$

It follows

$$\int_0^1 \mathbf{W}(s)' \Gamma \mathbf{W}(s)\, ds = O_p(N),$$

which establishes (16), α being positive and the behaviour of $\int_0^1 \mathbf{W}(s)' \Xi_\rho \mathbf{W}(s)\, ds$ being established in Proposition 1.

Proof of Lemma 3

It holds in the univariate case

$$T^{-0.5} \sum_{t=1}^{T} \widetilde{y}_t^{\mu} = \frac{1}{T^{0.5}} \sum_{t=1}^{T} \left(\widetilde{y}_t - \frac{1}{t} \sum_{j=1}^{t} \widetilde{y}_j \right) = \frac{1}{T^{0.5}} \sum_{t=1}^{T} \widetilde{y}_t \left(1 - \sum_{j=t}^{T} \frac{1}{j} \right).$$

Since $\sum_{j=1}^{p} \frac{1}{j} = C + \ln p + O\left(\frac{1}{p}\right)$, with C Euler's constant, we can write for $t \geq 2$

$$\begin{aligned} \sum_{j=t}^{T} \frac{1}{j} &= \sum_{j=1}^{T} \frac{1}{j} - \sum_{j=1}^{t-1} \frac{1}{j} = \ln T + O\left(\frac{1}{T}\right) - \ln(t-1) - O\left(\frac{1}{t-1}\right) \\ &= \ln\left(\frac{T}{t-1}\right) + O\left(\frac{1}{T} - \frac{1}{t-1}\right). \end{aligned}$$

It follows

$$\frac{1}{T^{0.5}} \sum_{t=1}^{T} \widetilde{y}_t^{\mu} = \frac{1}{T^{0.5}} \sum_{t=2}^{T} \widetilde{y}_t \left(1 + \ln \frac{t-1}{T}\right) + O_p\left(\frac{1}{T^{0.5}} \sum_{t=2}^{T} \frac{\widetilde{y}_t}{t-1}\right).$$

For the second term, we obviously have that $\widetilde{y}_t = O_p(1)$. Then,

$$\begin{aligned}\frac{1}{T^{0.5}}\sum_{t=2}^{T}\frac{\widetilde{y}_t}{t-1} &= O_p\left(\frac{1}{T^{0.5}}\sum_{t=2}^{T}\frac{1}{t-1}\right) = O_p\left(\frac{\ln T}{T^{0.5}}\right), \\ &= o_p(1).\end{aligned}$$

For the first term, convergence to a Riemann-Stieltjes integral holds due to weak convergence to Wiener process and consistency of $\widehat{\omega}_{ii}$,

$$\frac{1}{T^{0.5}}\sum_{t=2}^{T}\widetilde{y}_t\left(1+\ln\frac{t-1}{T}\right) \Rightarrow \int_0^1 (1+\ln s)\,dW(s).$$

In the multivariate case, this becomes

$$\frac{1}{T^{0.5}}\sum_{t=2}^{T}\widetilde{\mathbf{y}}_t\left(1+\ln\frac{t-1}{T}\right) \Rightarrow \Xi^{0.5}\int_0^1 (1+\ln s)\,d\mathbf{W}(s).$$

Then, the covariance matrix of the stochastic integral is given by[12]

$$\Xi\int_0^1 (1+\ln s)^2\,ds = \Xi,$$

as needed to obtain the result for recursive demeaning.

For recursive detrending, we have

$$\frac{1}{T^{0.5}}\sum_{t=2}^{T}\widetilde{y}_t^{\tau} = \frac{1}{T^{0.5}}\sum_{t=2}^{T}\left(\widetilde{y}_{t-1} + \frac{2}{t}\sum_{i=1}^{t}\widetilde{y}_i - \frac{6}{t(t+1)}\sum_{i=1}^{t} i\,\widetilde{y}_i\right).$$

[12]Integration by parts yields

$$\int \ln^2(s)ds = s\ln^2(s) - \int 2\ln(s)ds \quad\text{or}\quad \int (1+\ln(s))^2 ds = s(1+\ln^2(s)).$$

Rearranging terms leads to

$$T^{0.5}\sum_{t=2}^{T}\widetilde{y}_t^{\tau}=\frac{1}{T^{0.5}}\sum_{t=2}^{T}\widetilde{y}_{t-1}\left(1+2\sum_{i=t}^{T}\frac{1}{i}-6t\sum_{i=t}^{T}\frac{1}{i(i+1)}\right).$$

Since

$$\sum_{i=t}^{T}\frac{1}{i(i+1)}=\sum_{i=t}^{T}\left(\frac{1}{i}-\frac{1}{i+1}\right)=\frac{1}{t}-\frac{1}{T+1},$$

it follows similarly to the demeaning case

$$\frac{1}{T^{0.5}}\sum_{t=1}^{T}\widetilde{y}_{t-1}^{\tau}=\frac{1}{T^{0.5}}\sum_{t=1}^{T}y_{t-1}\left(1-2\ln\frac{t}{T}-6+6\frac{t}{T+1}\right)+o_p\left(1\right),$$

from which the desired result follows with (see footnote 9)

$$\int_0^1\left(1-2\ln s-6+6s\right)^2 ds=1.$$

References

Anderson, T.W. and D.A. Darling (1952), Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes, *Annals of Mathematical Statistics* 23, 193-212.

Andrews, D.W.K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* 59, 817-858.

Bai, J. and S. Ng (2004a), A New Look at Panel Testing of Stationarity and the PPP Hypothesis, in: Andrews, D.W.K. and J. Stock (eds.), *Indentification and Inference in Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Cambridge University Press.

Bai, J. and S. Ng (2004b), A PANIC Attack on Unit Roots and Cointegration, *Econometrica* 72, 1127-1177.

Breitung, J. and M.H. Pesaran (2008), Unit Roots and Cointegration in Panels, in: Mátyás, L. and P. Sevestre (eds.), *The Econometrics of Panel Data*, 3rd edition, Ch. 9, pp. 279-322.

Culver, S.E. and D.H. Papell (1997), Is There a Unit Root in the Inflation Rate? Evidence from Sequential Break and Panel Data Models, *Journal of Applied Econometrics* 12, 435-444.

Demetrescu, M. (2009), Recursive Adjustment for General Deterministic Components and Improved Tests for the Cointegration Rank, Working Paper, Goethe University Frankfurt.

Demetrescu, M., U. Hassler and A.I. Tarcolea (2006), Combining Significance of Correlated Statistics with Application to Panel Data, *Oxford Bulletin of Economics and Statistics* 68, 647-633.

Fritsche, U. and V. Kuzin (2007), Unit Labor Cost Growth Differentials in the Euro Area, Germany, and the US: Lessons from PANIC and Cluster Analysis, *DEP Discussion Papers, Macroeconomics and Finance Series, 3/2007.*

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

Hadri, K. and R. Larsson (2005), Testing for Stationarity in Heterogeneous Panel Data where the Time Dimension is Finite, *Econometrics Journal* 8, 55-69.

Hanck, Ch. (2009), For Which Countries did PPP hold? A Multiple Testing Approach, *Empirical Economics* forthcoming.

Harris, D., S. Leybourne and B. McCabe (2005), Panel Stationarity Tests for Purchasing Power Parity with Cross-sectional Dependence, *Journal of Business and Economic Statistics* 23, 395-409.

Hartung, J. (1999), A Note on Combining Dependent Tests of Significance, *Biometrical Journal* 41, 849-855.

Harvey, A.C. (2001), Testing in Unobserved Component Models, *Journal of Forecasting* 20, 1-19.

Hlouskova, J. and M. Wagner (2006), The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study, *Econometric Reviews* 25, 85-116.

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.

Lee, D. and P. Schmidt (1996), On the Power of the KPSS Test of Stationarity against Fractionally-integrated Alternatives, *Journal of Econometrics* 73, 285-302.

Lee, H.Y. and J.L. Wu (2001) Mean Reversion of Inflation Rates: Evidence from 13 OECD Countries, *Journal of Macroeconomics* 23, 477-487.

Lee, J., C.J. Huang and Y. Shin (1997), On Stationary Tests in the Presence of Structural Breaks, *Economics Letters* 55, 165-172.

MacNeill, I.B. (1978), Properties of Sequences of Partial Sums of Polynomial Regression Residuals with Applications to Tests for Change of Regression at Unknown Times, *The Annals of Statistics* 6, 422-433.

Newey, W.K. and K.D. West (1987), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.

Nyblom, J. and A.C. Harvey (2000), Tests of Common Stochastic Trends, *Econometric Theory* 16, 176-199.

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

Phillips, P.C.B. and S.N. Durlauf (1986), Multiple Time Series Regressions with Integrated Processes, *Review of Economic Studies* 53, 473-495.

Phillips, P.C.B. and S. Ouliaris (1990), Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica* 58, 165-193.

Shin, Y. and A. Snell (2006), Mean Group Tests for Stationarity in Heterogeneous Panels, *Econometrics Journal* 9, 123-158.