PANEL STATIONARY TESTS AGAINST CHANGES IN PERSISTENCE*

Roy Cerqueti
University of Macerata

Mauro Costantini
Brunel University

Luciano Gutierrez
University of Sassari

Joakim Westerlund
Lund University
and
Centre for Financial Econometrics
Deakin University

November 22, 2016

Abstract

In this paper we propose new panel tests to detect changes in persistence. The test statistics are used to test the null hypothesis of stationarity against the alternative of a change in persistence from I(0) to I(1), from I(1) to I(0), and in an unknown direction. The limiting null distributions of the tests are derived and evaluated in small samples by means of Monte Carlo simulations. An empirical illustration is also provided.

JEL Classification: C12; C22.

Keywords: Persistence; stationarity; panel data.

1 Introduction

Over the last two decades, a vast literature has investigated whether economic and financial time series may be characterized by a change in persistence between separate I(1) and I(0)

*Previous versions of the paper were presented at the First International Conference in Memory of Carlo Giannini, "Recent Development in Econometric Methodology", at University of Bergamo, and at seminars at University of Vienna and University of Leicester. The authors would like to thank seminar and conference participants, and in particular David Hendry, Chihwa Kao, Oliver Linton, Robert Kunst, Stephen Pollock, Wojciech Charemza, Panicos Demetriades, Qiang Zhang, Francesco Moscone, and two anonymous referees for many useful comments and suggestions. Westerlund would like to thank the Knut and Alice Wallenberg Foundation for financial support through a Wallenberg Academy Fellowship, and the Jan Wallander and Tom Hedelius Foundation for financial support under research grant number P2014–0112:1.

regimes rather than simply I(1) or I(0) behavior. Changes of this kind in macroeconomic variables are well documented; see the literature reviews in Kim (2000) and Leybourne et al. (2003). A non-exhaustive list of the variables for which such phenomena have been observed includes inflation, real output, budgetary deficits, interest rates and exchange rates. Interestingly, while many data sets are in fact panels of multiple time series, the way that existing tests are constructed requires that the series are tested one at a time. This is wasteful in the sense that each time a test is carried out the information contained in the other series is effectively ignored. The current paper can be seen as a reaction to this. The purpose is to develop tests for changes in persistence that explores the multiplicity of series, and that can be seen as panel extensions of the time series tests of Kim (2000), Kim et al. (2002), and Busetti and Taylor (2004). The tests can be used to flexibly test the null hypothesis of stationarity against the alternative of a change in persistence not only from I(0) to I(1), and from I(1) to I(0), but also when the direction is unknown. The data generating process (DGP) considered is quite general. Some of the allowances are unit-specific constant and trend terms, cross-section heteroskedasticity, error serial correlation and cross-section dependence in the form of common factors. The asymptotic distributions of the tests are derived and evaluated in small samples using Monte Carlo simulation. An empirical illustration is also provided showing how how inflation of 20 developed countries has undergone a shift from I(0) to I(1).

The rest of the paper is organized as follows. Sections 2 and 3 present the model, the test statistics, and their asymptotic distributions, which are evaluated using simulations in Section 4. Section 5 reports the results from the empirical application. Section 6 concludes. Proofs of important results are provided in the Appendix

2 Model and assumptions

Consider the panel data variable $Y_{i,t}$, where $i = 1, ..., N$ and $t = 1, ..., T$ index the time-series and cross-sectional units, respectively. The DGP of this variable is given by

$$Y_{i,t} = \theta_i' D_{t,p} + \lambda_i' F_t + e_{i,t}, \tag{1}$$

$$e_{i,t} = \mu_{i,t} + \varepsilon_{i,t}, \tag{2}$$

where $D_{t,p} = (1, t, ..., t^p)'$ is a p-order trend polynomial such that $D_{t,p} = 0$ is $p = -1$, F_t is an $r \times 1$ vector of common factors with λ_i being the corresponding vector of factor loadings, and $\varepsilon_{i,t}$ is a mean zero and I(0) error term. The following three specifications of $\mu_{i,t}$ are considered, where $1(A)$, $\lfloor x \rfloor$, $\eta_{i,t}$ and $\tau_i^0 \in [0,1]$ denote the indicator function of the event A, the integer part of x, a mean zero I(0) error term, and the break fraction, respectively:

MU1. I(0) $\to$ I(1): $\mu_{i,t} = \mu_{i,t-1} + 1(t > \lfloor T\tau_i^0 \rfloor)\eta_{i,t}$.

MU2. I(1) $\to$ I(0): $\mu_{i,t} = \mu_{i,t-1} + 1(t \leq \lfloor T\tau_i^0 \rfloor)\eta_{i,t}$.

MU3. Unknown direction: I(0) $\to$ I(1) or I(1) $\to$ I(0).

Under MU1 $Y_{i,t}$ is I(0) up to and including time $\lfloor \tau_i^0 T \rfloor$ but is I(1) after the break, provided that $\sigma_{\eta,i}^2 = \text{var}(\eta_{i,t}) > 0$. Under MU2 $Y_{i,t}$ is I(1) up to and including time $\lfloor \tau_i^0 T \rfloor$ but it is I(0) after the break, provided again that $\sigma_{\eta,i}^2 > 0$. Therefore, the hypothesis of stationarity against a shift in persistence from I(0) to I(1) or viceversa can be stated as $H_0 : \sigma_{\eta,1}^2 = ... = \sigma_{\eta,N}^2 = 0$ versus $H_1 : \sigma_{\eta,i}^2 > 0$ for at least some i. Whenever the alternative is I(1) $\to$ I(0) we write "H_1 : I(1) $\to$ I(0)", whereas if the alternative is I(1) $\to$ I(0), we write "H_1 : I(1) $\to$ I(0)".

The conditions placed on the above DGP are given in Assumption 1, where $C < \infty$, $tr(A)$, $||A|| = \sqrt{tr(A'A)}$, $\to_p$ and $\mathcal{F}_{i,t}$ denote a generic positive constant, the trace and Euclidean norm of the (generic) matrix A, convergence in probability, and the sigma-field generated by $\{(\varepsilon_{i,n}, \eta_{i,n})\}_{n=1}^t$, respectively.

Assumption 1.

(i) $\varepsilon_{i,t} = \gamma_i(L)v_{i,t}$, where $v_{i,t}$ is independent and identically distribution (iid) with $E(v_{i,t}) = 0$, $E(v_{i,t}^2) = 1$, $E(v_{i,t}^8) \leq C$, $\gamma_i(L) = \sum_{j=0}^\infty \gamma_{ji} L^j$, $\sum_{j=0}^\infty j||\gamma_{ji}|| \leq C$ and $\gamma_i(1)^2 > 0$;

(ii) $\eta_{i,t} = \phi_i(L)w_{i,t}$, where $w_{i,t}$ is iid with $E(w_{i,t}) = 0$, $E(w_{i,t}^2) = 1$, $E(w_{i,t}^8) \leq C$, $\phi_i(L) = \sum_{j=0}^\infty \phi_{ji} L^j$, $\sum_{j=0}^\infty j||\phi_{ji}|| \leq C$ and $\phi_i(1)^2 > 0$;

(iii) F_t is I(0) such that $E(||F_t||^4) \leq C$ and $T^{-1}\sum_{t=1}^T F_t F_t' \to_p \Sigma_F > 0$;

(iv) $\varepsilon_{i,t}$, $\eta_{i,t}$ and F_t are mutually independent;

(v) $\mu_{1,0} = ... = \mu_{N,0} = 0$;

(vi) λ_i is deterministic such that $||\lambda_i||^4 \leq C$, $N^{-1}\sum_{i=1}^N \lambda_i \lambda_i' \to \Sigma_\lambda > 0$ as $N \to \infty$.

Remark 1. Assumption 1 puts restrictions on the time series and cross-sectional properties of $\varepsilon_{i,t}$ and $\eta_{i,t}$. The restrictions are very similar to the ones of Bai and Ng (2004), and we therefore refer to this other paper for a detailed discussion. The main difference when compared to Bai and Ng (2004) is that here F_t cannot be I(1). Thus, while $Y_{i,t}$ may be cross-correlated, it cannot be affected by common stochastic trends. However, we would like to point out that this assumption is mainly for ease of interpretation of the test outcome, for if F_t is allowed to be I(1) the persistence of $Y_{i,t}$ cannot be inferred from $e_{i,t}$ alone, and in the present paper we focus on the testing of $e_{i,t}$. Hence, analogous to the PANIC approach of Bai and Ng (2004), if F_t is permitted to be I(1), then we also need to test this variable.

3 The test statistics

The general testing idea is to first purge the effect of F_t, and then to submit the resulting residuals to a test for a change in persistence. The implementation of the first step depends on whether F_t is known or not.

3.1 F_t known

Consider the generic variable $X_{i,t}$. The detrended version of this variables is henceforth denoted $X_{i,t}^p = X_{i,t} - \sum_{n=1}^T X_{i,n} a_{n,t,p}$, where $a_{n,k,p} = D_{n,p}'(\sum_{t=1}^T D_{t,p} D_{t,p}')^{-1} D_{k,p}$ and $p \geq 0$. If $p = -1$, then we define $X_{i,t}^p = X_{i,t}$. In this notation, the detrended and defactored version of $Y_{i,t}$ is given by $\hat{e}_{i,t} = Y_{i,t}^p - \hat{\lambda}_i' F_t^p$, where $\hat{\lambda}_i$ is the least squares (LS) slope estimator in a regression of $Y_{i,t}^p$ onto F_t^p. Thus, while in this section F_t is assumed to be known, λ_i is still treated as unknown. Consider the following test statistic, which is suitable for testing if cross-section unit i is I(0) versus I(1) $\to$ I(0) (see, for example, Kim, 2000; Kim et al., 2002; Busetti and Taylor, 2004):

$$K_{i,T}(\tau) = \frac{(\lfloor T\tau \rfloor)^2}{(T - \lfloor T\tau \rfloor)^2} \frac{\sum_{t=\lfloor T\tau \rfloor+1}^T S_{i,t}^1(\tau)^2}{\sum_{t=1}^{\lfloor T\tau \rfloor} S_{i,t}^0(\tau)^2},$$

where $\tau \in [0,1]$, $S_{i,t}^0(\tau) = \sum_{n=1}^t \hat{e}_{i,n}$ and $S_{i,t}^1(\tau) = \sum_{n=\lfloor T\tau \rfloor+1}^t \hat{e}_{i,n}$. The error sequences $\{\hat{e}_{i,n}\}_{n=1}^{\lfloor T\tau \rfloor}$ and $\{\hat{e}_{i,n}\}_{n=\lfloor T\tau \rfloor+1}^T$ come from two separate regressions; while the former uses only the first $\lfloor T\tau \rfloor$ observations, the latter uses only the last $\lfloor T(1-\tau) \rfloor$ observations.

Remark 2. The $K_{i,T}(\tau)$ test considered here is in the spirit of Kwiatkowski et al. (1992) in which the constant I(0) null is tested versus the constant I(1) alternative. An alternative approach is to follow Banerjee et al. (1992) and Leybourne et al. (2003) who use the Dickey–Fuller statistic, in which the null and the alternative hypotheses are reversed. Panel variants of these can be constructed in the same way as the one suggested below for $K_{i,T}(\tau)$ (see Demetrescu and Hanck, 2013, for such a proposal).

Let $\mathcal{C} = [\tau_{min}, \tau_{max}] \subseteq (0,1)$. In this paper, we consider three transformations to eliminate the dependence on τ in $K_{i,T}(\tau)$ (see, for example, Kim, 2000);

T1. The maximum-Chow transformation:

$$K_{i,T}^{1} = \max_{s=\lfloor T\tau_{min}\rfloor,\ldots,\lfloor T\tau_{max}\rfloor} K_i(s/T).$$

T2. The mean-exponential transformation:

$$K_{i,T}^{2} = \ln\left((\lfloor T(\tau_{max} - \tau_{min})\rfloor + 1)^{-1} \sum_{s=\lfloor T\tau_{min}\rfloor}^{\lfloor T\tau_{max}\rfloor} \exp[K_i(s/T)] \right).$$

T3. The mean score transformation:

$$K_{i,T}^{3} = (\lfloor T(\tau_{max} - \tau_{min})\rfloor + 1)^{-1} \sum_{s=\lfloor T\tau_{min}\rfloor}^{\lfloor T\tau_{max}\rfloor} K_i(s/T).$$

In Appendix (Proof of Theorem 1), we show that $K_{i,T}(\tau) \to_w K_i(\tau)$ as $T \to \infty$, where $\to_w$ signifies weak convergence and $K_i(\tau)$ is a certain ratio of stochastic integrals. Since $K_1(\tau), \ldots, K_N(\tau)$ are iid, we may define $\mu_{K,j} = E(K_i^j)$ and $\sigma_{K,j}^2 = \text{var}(K_i^j)$ for $j \in \{1,2,3\}$. Numerical values of $\mu_{K,j}$ and $\sigma_{K,j}$ are reported in Table 1. The proposed panel test statistic for testing H_0 versus H_1 : I(0) $\to$ I(1) is given by

$$K_{NT}^{j} = \frac{1}{\sigma_{K,j}\sqrt{N}} \sum_{i=1}^{N} (K_{i,T}^{j} - \mu_{K,j}).$$

For testing if cross-section unit i is I(0) versus I(1) $\to$ I(0), the following "reverse" test statistic can be used (see Kim, 2000; Kim et al., 2002; Busetti and Taylor, 2004):

$$R_i(\tau) = (K_i(\tau))^{-1},$$

which can be transformed using T1–T3 to eliminate the dependence on τ. The resulting transformed statistic is written in an obvious notation as R_i^j. Based on this test statistic, we may define $R_{NT}^{j} = \sigma_{R,j}^{-1} N^{-1/2} \sum_{i=1}^{N} (R_{i,T}^{j} - \mu_{R,j})$ with obvious definitions of $\sigma_{R,j}^2$ and $\mu_{R,j}$. When the direction of the persistency is unknown, the following maximum statistic may be used:

$$M_{i,T}^{j} = \max\{K_{i,T}^{j}, R_{i,T}^{j}\},$$

which can again be normalized to obtain $M^j_{NT} = \sigma^{-1}_{M,j} N^{-1/2} \sum_{i=1}^N (M^j_{i,T} - \mu_{M,j})$.

Theorem 1. *Under H_0 and Assumption 1, as N, $T \to \infty$ with $N/T \to 0$,*

$$K^j_{NT},\ R^j_{NT},\ M^j_{NT} \to_d N(0,1),$$

where $\to_d$ signifies convergence in distribution.

Remark 3. While the test statistics considered here are independent of $\tau^0_1, ..., \tau^0_N$, in applications it is sometimes useful to be able to estimate these parameters. This can be accomplished using the proposal of Kim (2000, Section 3.2), which basically amounts to setting $\hat{\tau}^0_i$ equal to the suitably maximizing or minimizing value of $K_{i,T}(\tau)$, depending on whether it is I(0) $\to$ I(1) or I(1) $\to$ I(0) that is being tested. Alternatively, we may follow Busetti and Taylor (2004, Section 6.2), who suggest setting $\hat{\tau}^0_i$ equal to the value of τ^0_i that minimizes the sum of squares of $\hat{e}_{i,t}$.

Remark 4. The requirement that $N/T \to 0$ is sufficient but not necessary and is needed to make sure that certain remainder terms are negligible. However, the order of these terms is not the sharpest possible. A more elaborate asymptotic analysis would be required to obtain the exact order. In Section 4, we use Monte Carlo simulation to evaluate the effect of N/T in small samples.

3.2 F_t unknown

The estimation of F_t can be performed in two ways; (i) unrestrictedly, or (ii) restricted under H_0. In both cases, we follow the bulk of the previous literature and use the principal components method (see, for example, Bai and Ng, 2004). The restricted estimator of $F = (F_1, ..., F_T)'$, denoted $\hat{F}^0 = (\hat{F}^0_1, ..., \hat{F}^0_T)'$, is $\sqrt{T}$ times the eigenvectors corresponding to the first r largest eigenvalues of the $T \times T$ matrix $Y^p(Y^p)'$, where $Y^p = (Y^p_1, ..., Y^p_N)$ and $Y^p_i = (Y^p_{i,1}, ..., Y^p_{i,T})'$ are $T \times N$ and $T \times 1$, respectively. Under the normalization $T^{-1}\hat{F}^0(\hat{F}^0)' = I_r$, the estimated loading matrix is $(\hat{\lambda}^0)' = (\hat{\lambda}^0_1, ..., \hat{\lambda}^0_N) = T^{-1}(\hat{F}^0)'Y^p$. The restricted estimator of $e_{i,t}$ that we will be considering can now be constructed as

$$\hat{e}^0_{i,t} = Y^p_{i,t} - (\hat{\lambda}^0_i)'\hat{F}^0_t. \tag{3}$$

Let $X^{p-1}_{i,t}$ be $X_{i,t}$ when detrended using a trend polynomial of order $p-1$. Hence, $X^{p-1}_{i,t} = X_{i,t}$ if $p = 0$. Let $f_t = \Delta F_t$ and $y_{i,t} = \Delta Y_{i,t}$ (for $t = 2, ..., T$). The unrestricted estimators $\hat{f}^1_t$ and $\hat{\lambda}^1_i$

of (the space spanned by) f_t^{p-1} and λ_i are $\hat{F}_t^0$ and $\hat{\lambda}_i^0$, respectively, but with $Y_{i,t}^p$ replaced by $y_{i,t}^{p-1}$. Let

$$\tilde{e}_{i,t}^1 = \sum_{n=2}^{t} [y_{i,n}^{p-1} - (\hat{\lambda}_i^1)' \hat{f}_n^1], \tag{4}$$

where $\tilde{e}_{i,1}^1 = 0$. The unrestricted estimator $\hat{e}_{i,t}^1$ of $e_{i,t}$ is given by $\hat{e}_{i,t}^1 = (\tilde{e}_{i,t}^1)^p$. The appropriate test statistics to consider when F_t is unknown, henceforth denoted K_{hNT}^j, R_{hNT}^j and M_{hNT}^j for $h \in \{0, 1\}$, are given by K_{NT}^j, R_{NT}^j and M_{NT}^j, respectively, with $\hat{e}_{i,t}$ replaced by $\hat{e}_{i,t}^h$.

Theorem 2. *Under H_0 and Assumptions 1, as N, $T \to \infty$ with $N/T \to 0$,*

$$K_{hNT}^j,\ R_{hNT}^j,\ M_{hNT}^j \to_d N(0,1).$$

As Theorem 2 makes clear, the factors can be unknown and still the asymptotic distributions of the test statistics are $N(0, 1)$. This is in agreement with the results reported by Bai and Ng (2004) for their pooled panel unit root tests.

4 Monte Carlo simulations

A small-scale Monte Carlo study was conducted to investigate the properties of the new tests in small samples. The DGP is given by a restricted version of (1)–(2) that sets $\varepsilon_{i,t} \sim N(0, \sigma_{\varepsilon,i}^2)$, $\eta_{i,t} \sim N(0, \sigma_\eta^2)$, $\sigma_\eta \in \{0, 0.25, 0.5\}$, $\tau_i^0 \sim U(0.3, 0.7)$, $r = 1$, and $F_t = \rho F_{t-1} + v_t$, where $v_t \sim N(0, 1)$ and $\rho \in \{0.3, 0.6\}$ (see, for example, Gengenbach et al., 2010, for a similar parametrization). For $\sigma_{\varepsilon,i}$, we consider two cases. In the first, $\sigma_{\varepsilon,i} = 1$ for all i, while in the second, $\sigma_{\varepsilon,i} \sim U(1, 2)$. Since a more volatile idiosyncratic error will make F_t more difficult to discern, we expect that the results for the second case will deteriorate when compared to the first. All results are based on 1,000 replications of samples of size $N \in \{5, 10, 20\}$ and $T \in \{50, 100\}$. Also, following Kim (2000), $\mathcal{C} = [0.20, 0.80]$. Results were obtained for $p \in \{0, 1\}$, although in this paper we focus on the results for the empirically most common specification with $p = 0$ (a constant but no trend). The results for $p = 1$ (constant and trend) can be obtained upon request. Both the restricted and unrestricted factor estimation methods were simulated. Interestingly, the restricted method led to better results in terms of both size accuracy and power. In this paper, we therefore only report the results for the restricted method, where the number of common factors is determined using the IC_2 criterion of Bai and Ng (2002) with a maximum of three factors.[1]

The 5% size and power results are reported in Tables 2–5. While Tables 2 ($\rho = 0.3$) and 3 ($\rho = 0.6$) contain the results for the tests of I(0) $\rightarrow$ I(1), Tables 4 ($\rho = 0.3$) and 5 ($\rho = 0.6$) contain the corresponding results for I(1) $\rightarrow$ I(0). The information content of these tables may be summarized as follows.

- All tests have good size accuracy when $\sigma_{\varepsilon,i} = 1$ and $\rho = 0.3$. This is true for all constellations of T and N considered, although the distortions do have a tendency to increase slightly in N, which is consistent with the previous panel unit root literature (see Westerlund and Breitung, 2013, for a discussion). While there are no big differences, the best size accuracy is generally obtained by using K^2_{NT}, R^2_{NT} and M^2_{NT}, whereas K^3_{NT}, R^1_{NT} and R^3_{NT} generally leads to the worst accuracy.

- As expected, increases in ρ and/or $\sigma_{\varepsilon,i}$ generally lead to reduced size accuracy, although the distortions are never very large. This is true regardless of the direction of the change in persistence. In fact, the results are remarkably stable, given that the test statistics do not require any corrections to account for nuisance parameters.

- All tests perform quite well in terms of power, and there are clear improvements as N and/or T increases. The fact that power is not only increasing in T, but also in N illustrates the advantage of accounting for the cross-sectional variation of the data. Power is also increasing in the distance to the null, as measured by σ_η, which is again just as expected.

5 Empirical illustration

The question of whether inflation should be considered as I(0) or I(1) has been subject to a long debate. According to recent studies (see, for example, Kim, 2000; Busetti and Taylor, 2004), however, inflation may be better characterized by a change in persistence between separate I(1) and I(0) regimes rather than simply I(1) or I(0) behavior. The purpose of this illustration is to test this hypothesis using a large panel of quarterly CPI inflation data covering 20 countries (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany,

[1] See Westerlund and Mishra (2016) for a more elaborate selection approach that uses a data-driven penalty.

Greece, Italy, Japan, Korea, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, the UK and the US) between 1970:1 and 2013:4. All data are taken from OECD Main Economic Indicators.

The number of common factors is determined in the same way as in the simulations. As is customary when dealing with inflation (see, for example, Leybourne et al., 2003), the tests are fitted with a constant but no trend. The results are reported in Table 6. The first thing to note is that while in case of K_{NT}^1, K_{NT}^2 and K_{NT}^3 there is no evidence against the I(0) null, R_{NT}^1, R_{NT}^2 and R_{NT}^3 all lead to a clear rejection. This is true even at the most conservative 1% level. We therefore conclude that inflation has been subject to a change in persistence from I(1) to I(0), which is in agreement with the recent empirical literature based on US data (see, for example, Busetti and Taylor, 2004; Harvey et al., 2006). A common explanation for the observed change in persistence of inflation in the US is that it is due to the stock market collapse of the late 1980's and the recession that followed it. One interpretation of the results reported in the current paper is therefore that they reflect the worldwide recession of the early 1990's, which was to a large extent triggered by the recession in the US. Another possibility is that the results reflect in part monetary policy shifts (see, for example, Davig and Doh, 2014, and the references provided therein).

6 Conclusion

This paper develops panel tests that are suitable for testing the null hypothesis of stationarity against the alternative of a change in persistence from I(0) to I(1), from I(1) to I(0), or when the direction is unknown. The DGP used for this purpose is quite general and allows unit-specific constant and trend terms, cross-section heteroskedasticity, error serial correlation and cross-section dependence in the form of common factors.

References

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

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

Banerjee, A., R. Lumsdaine and J. Stock (1992). Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. *Journal of Business and Economics Statistics* **10**, 271–288.

Busetti, F., and R. A. M. Taylor (2004). Tests of stationarity against a change in persistence. *Journal of Econometrics* **123**, 33–66.

Davig, T., and T. Doh (2014). Monetary policy regime shifts and inflation persistence. *Review of Economics and Statistics* **96**, 862–875.

Demetrescu, M., and C. Hanck (2013). Nonlinear IV panel unit root testing under structural breaks in the error variance. *Statistical Papers* **54**, 1043–1066.

Harvey, D. I., S. J. Leybourne and R. A. M. Taylor (2006). Modified tests for a change in persistence. *Journal of Econometrics* **134**, 441–469.

Gengenbach, C., F. Palm and J.-P. Urbain (2010). Panel unit root tests in the presence of cross-sectional dependencies: Comparison and implications for modelling. *Econometric Reviews* **29**, 111–145.

Kim, J. Y. 2000. Detection of change in persistence of a linear time series. *Journal of Econometrics* **95**, 97–116.

Kim, J. Y., J. B. Franch and R. B. Amador (2002). Corrigendum to "Detection of change in persistence of a linear time series". *Journal of Econometrics* **109**, 389–392.

Leybourne, S. J., T. Kim, P. Newbold and V. Smith (2003). Tests for a change in persistence against the null of difference-stationarity. *Econometrics Journal* **6**, 291–311.

Westerlund, J., and R. Larsson (2009). A note on the pooling of individual PANIC unit root tests. *Econometric Theory* **25**, 1851–1868.

Westerlund, J., and J. Breitung (2013). Lessons from a decade of IPS and LLC. *Econometric Reviews* **32**, 547–591.

Westerlund, J., and S. Mishra (2016). On the determination of the number of factors using information criteria with data-driven penalty. Forthcoming in *Statistical Papers*.

Appendix: Proofs

The proofs of Theorems 1 and 2 are established for K^j_{NT}; the proofs for R^j_{NT} and M^j_{NT} are entirely analogous.

Proof of Theorem 1.

Under MU1, $\mu_{i,t} = \sum_{k=1}^{t} 1(k > \lfloor T\tau \rfloor)\eta_{i,k}$, and by further invoking H_0, $\mu_{i,t} = 0$, giving

$$Y_{i,t} = \theta_i' D_t + \lambda_i' F_t + \mu_{i,t} + \varepsilon_{i,t} = \theta_i' D_t + \lambda_i' F_t + \varepsilon_{i,t}, \tag{A1}$$

It follows that

$$Y^p_{i,t} = \lambda_i' F^p_t + \varepsilon^p_{i,t}, \tag{A2}$$

with obvious definitions of F^p_t and $\varepsilon^p_{i,t}$, which in turn implies

$$\hat{e}_{i,t} = Y^p_{i,t} - \hat{\lambda}_i' F^p_t = \varepsilon^p_{i,t} - (\hat{\lambda}_i - \lambda_i)' F^p_t, \tag{A3}$$

Therefore,

$$T^{-1/2} \sum_{n=1}^{t} \hat{e}_{i,n} = T^{-1/2} \sum_{n=1}^{t} \varepsilon^p_{i,n} - (\hat{\lambda}_i - \lambda_i)' T^{-1/2} \sum_{n=1}^{t} F^p_n. \tag{A4}$$

Under H_0 and with F_t known $Y_{i,t} = \theta_i' D_t + \lambda_i' F_t + \varepsilon_{i,t}$ is just an ordinary time series regression in I(0) variables with exogenous regressors. It follows that $\sqrt{T}(\hat{\lambda}_i - \lambda_i) = O_p(1)$, and therefore, since $T^{-1/2} \sum_{n=1}^{t} F^p_n = O_p(1)$,

$$T^{-1/2} \sum_{n=1}^{t} \hat{e}_{i,n} = T^{-1/2} \sum_{n=1}^{t} \varepsilon^p_{i,n} + O_p(T^{-1/2}). \tag{A5}$$

Hence, using $\overline{K}_{i,T}(\tau)$ to denote $K_{i,T}(\tau)$ with $\hat{e}_{i,n}$ replaces by $e_{i,n}$, we have

$$K_{i,T}(\tau) = \overline{K}_{i,T}(\tau) + O_p(T^{-1/2}), \tag{A6}$$

where the first term on the right is the same as in Harvey et al. (2006). It follows from their results that

$$K_{i,T}(\tau) \to_w \overline{K}_i(\tau) = \frac{\overline{A}_i(\tau)}{\overline{B}_i(\tau)}, \tag{A7}$$

as $T \to \infty$, where $\to_w$ signifies weak convergence, and

$$\begin{aligned}
\overline{A}_i(\tau) &= (1-\tau)^{-2} \int_\tau^1 \overline{a}_i(r)^2 dr, \\
\overline{B}_i(\tau) &= \tau^{-2} \int_0^\tau \overline{b}_i(r)^2 dr, \\
\overline{a}_i(\tau) &= W_{\varepsilon,i}(\tau) - W_{\varepsilon,i}(r) - \int_\tau^1 dW_{\varepsilon,i}(r) D_p(r)' \left(\int_\tau^1 D_p(r) D_p(r)' dr \right)^{-1} \int_\tau^r D_p(s) ds, \\
\overline{b}_i(\tau) &= W_{\varepsilon,i}(r) - \int_0^\tau dW_{\varepsilon,i}(r) D_p(r)' \left(\int_0^\tau D_p(r) D_p(r)' dr \right)^{-1} \int_0^r D_p(s) ds,
\end{aligned}$$

with $W_{\varepsilon,i}(r)$ being a standard Brownian motion, and $D_p(r)$ is such that $Q_T^{-1} D_{\lfloor Tr \rfloor, p} \to D_p(r)$, where $Q_T = \text{diag}(1, T, ..., T^p)$. Note in particular how $D_0(r) = 1$ and $D_1(r) = (1, r)'$. Therefore, by the continuous mapping theorem, and writing $K_{i,T}^j = H_j(K_{i,T}(\tau))$ and $\overline{K}_{i,T}^j = H_j(\overline{K}_{i,T}(\tau))$ as in Busetti and Taylor (2004),

$$K_{i,T}^j = \overline{K}_{i,T}^j(\tau) + O_p(T^{-1/2}) \to_w H_j(\overline{K}_i(\tau)) = \overline{K}_i^j. \tag{A8}$$

Let us now consider K_{NT}^j. By using the previous result

$$K_{NT}^j = \frac{1}{\sigma_{K,j}\sqrt{N}} \sum_{i=1}^N (K_{i,T}^j - \mu_{K,j}) = \frac{1}{\sigma_{K,j}\sqrt{N}} \sum_{i=1}^N (\overline{K}_{i,T}^j - \mu_{K,j}) + O_p(\sqrt{N} T^{-1/2}) \tag{A9}$$

where $O_p(\sqrt{N} T^{-1/2}) = o_p(1)$ under our assumption that $N/T = o(1)$. We now use the same steps as in Moon and Phillips (2000, page 994) to verify that $(\overline{K}_{i,T}^j - \mu_{K,j})$ satisfies conditions (i)–(iv) of the central limit theorem of Phillips and Moon (1999, Theorem 2). In so doing we follow their notation and write $Q_{i,T} = (\overline{K}_{i,T}^j - \mu_{K,j})$, which is iid with mean zero and variance $\sigma_{K,j}^2 \leq C$. We have already shown that $\overline{K}_{i,T}^j \to_w \overline{K}_i^j$ as $T \to \infty$, which implies $Q_{i,T} \to_d Q_i = (\overline{K}_i^j - \mu_{K,j})$, and it is also not difficult to verify that $E(Q_{i,T}^2) \to E(Q_i^2) = \sigma_{K,j}^2$. This verifies conditions (i), (ii) and (iv). Condition (iv) follows from noting that, by the continuous mapping theorem, $Q_{i,T}^2 \to_w Q_i^2$. It follows that

$$\begin{aligned}
K_{NT}^j &= \frac{1}{\sigma_{K,j}\sqrt{N}} \sum_{i=1}^N (K_{i,T}^j - \mu_{K,j}) \\
&= \frac{1}{\sigma_{K,j}\sqrt{N}} \sum_{i=1}^N (\overline{K}_{i,T}^j - \mu_{K,j}) + O_p(\sqrt{N} T^{-1/2}) \to_d N(0,1) \qquad \text{(A10)}
\end{aligned}$$

as N, $T \to \infty$ with $N/T \to 0$. ■

Proof of Theorem 2.

We begin by considering the case when the estimator of $e_{i,t}$ is based on the restricted estimators of λ_i and F_t under H_0. As in Proof of Theorem 1, under MU1 and H_0, $Y_{i,t} = \theta_i' D_{i,t} + \lambda_i' F_t + \varepsilon_{i,t}$. In order to capture the fact that λ_i and F_t are not separately identifiable we introduce the $r \times r$ rotation matrix H such that

$$\hat{e}^0_{i,t} = Y^p_{i,t} - (\hat{\lambda}^0_i)'\hat{F}^0_t = \varepsilon^p_{i,t} - \lambda_i' H^{-1}(\hat{F}^0_t - HF^p_t) - (\hat{\lambda}_i - (H^{-1})'\lambda_i)'\hat{F}^0_t. \tag{A11}$$

Hence,

$$\begin{aligned} T^{-1/2}\sum_{n=1}^{t}\hat{e}^0_{i,n} &= T^{-1/2}\sum_{n=1}^{t}\varepsilon^p_{i,n} - \lambda_i' H^{-1} T^{-1/2}\sum_{n=1}^{t}(\hat{F}^0_n - HF^p_n) \\ &- (\hat{\lambda}_i - (H^{-1})'\lambda_i)' T^{-1/2}\sum_{n=1}^{t}\hat{F}^0_n. \end{aligned} \tag{A12}$$

By Lemmas 1 (c) and 2 of Bai and Ng (2004), $||\hat{\lambda}_i - (H^{-1})'\lambda_i|| = O_p(N^{-1}) + O_p(T^{-1/2})$ and $||T^{-1/2}\sum_{n=1}^{t}(\hat{F}^0_n - HF^p_n)|| = O_p(N^{-1/2}) + O_p(T^{-3/4})$, where the latter result holds uniformly in t. Hence, since

$$T^{-1/2}\sum_{n=1}^{t}\hat{F}^0_n = HT^{-1/2}\sum_{n=1}^{t}F^p_n + T^{-1/2}\sum_{n=1}^{t}(\hat{F}^0_n - HF^p_n) = O_p(1), \tag{A13}$$

we can show that

$$T^{-1/2}\sum_{n=1}^{t}\hat{e}^0_{i,n} = T^{-1/2}\sum_{n=1}^{t}\varepsilon^p_{i,n} + O_p(N^{-1/2}) + O_p(T^{-1/2}). \tag{A14}$$

Hence, as in the case when F_t is known (see Proof of Theorem 1), the estimation and removal of the common component do not affect the asymptotic distribution of the test statistic. Specifically, using $K^j_{0i,T}$ to denote $K^j_{i,T}$ with $\hat{e}^0_{i,n}$ in place of $\hat{e}_{i,n}$, we get

$$|K^j_{0i,T} - K^j_{i,T}| = O_p(N^{-1/2}) + O_p(T^{-1/2}), \tag{A15}$$

which holds uniformly in (j, i). In order to show that the resulting panel statistic, K^j_{0NT} say, converges to $N(0, 1)$, we may use the same argument as in Westerlund and Larsson (2009).

Consider the unrestricted estimator of $e_{i,t}$. We have $\tilde{e}^1_{i,t} = \sum_{n=2}^{t}[y^{p-1}_{i,n} - (\hat{\lambda}^1_i)'\hat{f}^1_n]$, where, under H_0, $y_{i,t} = \Delta Y_{i,t} = \theta_i'\Delta D_t + \lambda_i' f_t + \Delta\varepsilon_{i,t}$ with $f_t = \Delta F_t$. It follows that $y^{p-1}_{i,t} = \lambda_i' f^{p-1}_t + (\Delta\varepsilon_{i,t})^{p-1}$, and therefore

$$\begin{aligned} \tilde{e}^1_{i,t} &= \sum_{n=2}^{t}[y^{p-1}_{i,n} - (\hat{\lambda}^1_i)'\hat{f}^1_n] \\ &= \sum_{n=2}^{t}[(\Delta\varepsilon_{i,t})^{p-1} - \lambda_i' H^{-1}(\hat{f}^1_t - Hf^{p-1}_t) - (\hat{\lambda}_i - (H^{-1})'\lambda_i)'\hat{f}^1_t]. \end{aligned} \tag{A16}$$

Consider $\sum_{n=2}^{t}(\hat{f}_t^1 - Hf_t^{p-1})$. From Proof of Theorem 3 in Bai (2003), using V to denote a diagonal matrix consisting of the first r eigenvalues of $(NT)^{-1}y^{p-1}(y^{p-1})'$ in decreasing order,

$$\begin{aligned}
&\sum_{n=2}^{t}(\hat{f}_t^1 - Hf_t^{p-1}) \\
&= N^{-1/2}V^{-1}T^{-1}\sum_{n=2}^{T}\hat{f}_t^1(f_t^{p-1})'N^{-1/2}\sum_{i=1}^{N}\lambda_i\sum_{n=2}^{t}(\Delta\varepsilon_{i,t})^{p-1} + O_p(N^{-1}) + O_p(T^{-1}) \\
&= N^{-1/2}V^{-1}HT^{-1}\sum_{n=2}^{T}f_t^{p-1}(f_t^{p-1})'N^{-1/2}\sum_{i=1}^{N}\lambda_i(\varepsilon_{i,t}^{p-1} - \varepsilon_{i,1}^{p-1}) \\
&+ N^{-1/2}V^{-1}T^{-1}\sum_{n=2}^{T}(\hat{f}_t^1 - Hf_t^{p-1})(f_t^{p-1})'N^{-1/2}\sum_{i=1}^{N}\lambda_i(\varepsilon_{i,t}^{p-1} - \varepsilon_{i,1}^{p-1}) \\
&+ O_p(N^{-1}) + O_p(T^{-1}). \qquad \text{(A17)}
\end{aligned}$$

where we have made use of the fact that $\sum_{n=2}^{t}(\Delta\varepsilon_{i,n})^{p-1} = \varepsilon_{i,t}^{p-1} - \varepsilon_{i,1}^{p-1}$. Now, $||V||$ and $||N^{-1/2}\sum_{i=1}^{N}\lambda_i(\varepsilon_{i,t}^{p-1} - \varepsilon_{i,1}^{p-1})||$ are both $O_p(1)$. Moreover, by Lemma A.1 of Bai (2003),

$$\begin{aligned}
\left|\left|T^{-1}\sum_{n=2}^{T}(\hat{f}_t^1 - Hf_t^{p-1})(f_t^{p-1})'\right|\right| &\leq \left(T^{-1}\sum_{n=2}^{T}||\hat{f}_t^1 - Hf_t^{p-1}||^2\right)^{1/2}\left(T^{-1}\sum_{n=2}^{T}||f_t^{p-1}||^2\right)^{1/2} \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}),
\end{aligned}$$

from which it follows that

$$\left|\left|\sum_{n=2}^{t}(\hat{f}_t^1 - Hf_t^{p-1})\right|\right| = O_p(N^{-1/2}) + O_p(T^{-1}). \qquad \text{(A18)}$$

By using this and $\hat{F}_t^1 = \sum_{n=2}^{t}\hat{f}_t^1 = H(F_t^{p-1} - F_1^{p-1}) + \sum_{n=2}^{t}(\hat{f}_t^1 - Hf_t^{p-1})$, we obtain

$$\begin{aligned}
\tilde{e}_{i,t}^1 &= \sum_{n=2}^{t}(\Delta\varepsilon_{i,t})^{p-1} - \lambda_i'H^{-1}\sum_{n=2}^{t}(\hat{f}_t^1 - Hf_t^{p-1}) - (\hat{\lambda}_i - (H^{-1})'\lambda_i)'\sum_{n=2}^{t}\hat{f}_t^1 \\
&= \varepsilon_{i,t}^{p-1} - \varepsilon_{i,1}^{p-1} - \lambda_i'H^{-1}\sum_{n=2}^{t}(\hat{f}_t^1 - Hf_t^{p-1}) \\
&- (\hat{\lambda}_i - (H^{-1})'\lambda_i)'H(F_t^{p-1} - F_1^{p-1}) - (\hat{\lambda}_i - (H^{-1})'\lambda_i)'\sum_{n=2}^{t}(\hat{f}_t^1 - Hf_t^{p-1}) \\
&= \varepsilon_{i,t}^{p-1} - \varepsilon_{i,1}^{p-1} + O_p(N^{-1/2}) + O_p(T^{-1/2}). \qquad \text{(A19)}
\end{aligned}$$

suggesting that for $p \geq 0$,

$$\hat{e}_{i,t}^1 = (\tilde{e}_{i,t}^1)^p = \varepsilon_{i,t}^p + O_p(N^{-1/2}) + O_p(T^{-1/2}). \qquad \text{(A20)}$$

When appropriately normalized by $T^{-1/2}$, taking partial sums do not affect the order of the remainder terms. Hence, again, the estimation and removal of the common component do not affect the asymptotic distribution of the test statistic. ■

Table 1: Simulated mean and standard deviation normalization factors.

T	K_{NT}^1	K_{NT}^2	K_{NT}^3	R_{NT}^1	R_{NT}^2	R_{NT}^3	M_{NT}^1	M_{NT}^2	M_{NT}^3
Mean, $p = 0$ (constant)									
50	1.839	1.626	6.218	1.825	1.612	6.190	2.792	2.633	9.218
100	1.795	1.563	6.387	1.811	1.566	6.401	2.742	2.536	9.419
150	1.801	1.560	6.525	1.793	1.543	6.487	2.735	2.516	9.568
500	1.795	1.546	6.801	1.802	1.560	6.856	2.738	2.521	9.996
Standard deviation, $p = 0$ (constant)									
50	1.607	2.355	5.960	1.575	2.262	5.799	1.757	2.883	6.821
100	1.528	2.135	5.755	1.528	2.082	5.661	1.663	2.594	6.478
150	1.530	2.129	5.842	1.521	2.088	5.750	1.664	2.599	6.585
500	1.541	2.098	5.966	1.540	2.121	6.027	1.683	2.599	6.797
Mean, $p = 1$ (constant and trend)									
50	2.498	2.586	9.317	1.058	0.757	3.618	2.719	2.816	10.066
100	2.448	2.574	9.897	1.081	0.786	3.935	2.699	2.841	10.791
150	2.475	2.684	10.569	1.062	0.764	3.970	2.711	2.928	11.386
500	2.367	2.527	9.9038	1.069	0.828	3.903	2.916	1.196	10.339
Standard deviation, $p = 1$ (constant and trend)									
50	1.447	2.528	6.288	0.707	0.879	2.973	1.327	2.477	6.022
100	1.416	2.578	6.622	0.761	0.967	3.332	1.294	2.531	6.322
150	1.421	2.625	6.853	0.723	0.906	3.291	1.291	2.563	6.519
500	1.480	2.645	6.447	0.799	0.910	3.247	1.442	2.724	6.571

Notes: Let $Q_{NT}^j = \sigma_{Q,j}^{-1} N^{-1/2} \sum_{i=1}^{N} (Q_{i,T}^j - \mu_{Q,j})$ be one of the nine test statistics considered, where $j \in \{1, 2, 3\}$ and $Q \in \{K, R, M\}$. The values reported in the table refer to the appropriate mean and standard deviation correction factors, $\mu_{Q,j}$ and $\sigma_{Q,j}$, respectively, needed to construct Q_{NT}^j.

Table 2: 5% size and power when testing I(0) → I(1) and $\rho = 0.3$.

T	N	σ_η	K_{NT}^1	K_{NT}^2	K_{NT}^3	R_{NT}^1	R_{NT}^2	R_{NT}^3	M_{NT}^1	M_{NT}^2	M_{NT}^3
						$\sigma_{\varepsilon,i} = 1$					
50	5	0.000	0.055	0.040	0.061	0.073	0.058	0.077	0.046	0.042	0.040
50	5	0.250	0.103	0.099	0.132	0.090	0.046	0.085	0.082	0.073	0.094
50	5	0.500	0.210	0.205	0.270	0.162	0.027	0.148	0.133	0.132	0.165
50	10	0.000	0.070	0.050	0.073	0.075	0.050	0.064	0.057	0.047	0.036
50	10	0.250	0.141	0.126	0.171	0.142	0.035	0.092	0.112	0.099	0.122
50	10	0.500	0.338	0.343	0.467	0.306	0.112	0.285	0.225	0.227	0.284
50	20	0.000	0.055	0.038	0.067	0.103	0.053	0.075	0.044	0.041	0.059
50	20	0.250	0.204	0.177	0.252	0.295	0.098	0.146	0.124	0.117	0.149
50	20	0.500	0.539	0.553	0.724	0.596	0.439	0.645	0.295	0.319	0.419
100	5	0.000	0.077	0.057	0.079	0.088	0.070	0.074	0.070	0.055	0.057
100	5	0.250	0.278	0.288	0.374	0.186	0.066	0.146	0.225	0.237	0.283
100	5	0.500	0.432	0.450	0.540	0.326	0.059	0.292	0.342	0.363	0.431
100	10	0.000	0.099	0.071	0.080	0.112	0.065	0.087	0.086	0.067	0.051
100	10	0.250	0.359	0.361	0.462	0.342	0.162	0.241	0.313	0.329	0.385
100	10	0.500	0.633	0.680	0.815	0.547	0.348	0.583	0.482	0.532	0.646
100	20	0.000	0.100	0.067	0.077	0.115	0.078	0.087	0.090	0.067	0.049
100	20	0.250	0.506	0.551	0.634	0.510	0.297	0.380	0.387	0.430	0.484
100	20	0.500	0.884	0.924	0.968	0.748	0.639	0.854	0.706	0.762	0.847
						$\sigma_{\varepsilon,i} \sim U(1,2)$					
50	5	0.000	0.071	0.059	0.097	0.090	0.060	0.100	0.062	0.058	0.070
50	5	0.250	0.073	0.069	0.100	0.080	0.049	0.089	0.071	0.057	0.071
50	5	0.500	0.169	0.154	0.206	0.110	0.040	0.119	0.094	0.085	0.132
50	10	0.000	0.074	0.051	0.065	0.088	0.054	0.088	0.059	0.046	0.044
50	10	0.250	0.093	0.078	0.107	0.095	0.051	0.071	0.085	0.064	0.083
50	10	0.500	0.208	0.204	0.319	0.233	0.093	0.202	0.146	0.150	0.199
50	20	0.000	0.047	0.030	0.068	0.109	0.062	0.084	0.044	0.035	0.046
50	20	0.250	0.118	0.085	0.142	0.169	0.069	0.102	0.093	0.071	0.088
50	20	0.500	0.328	0.317	0.477	0.405	0.251	0.405	0.171	0.176	0.246
100	5	0.000	0.093	0.069	0.092	0.099	0.060	0.089	0.068	0.064	0.056
100	5	0.250	0.176	0.162	0.218	0.142	0.048	0.104	0.146	0.139	0.166
100	5	0.500	0.371	0.388	0.485	0.298	0.056	0.247	0.275	0.293	0.353
100	10	0.000	0.084	0.069	0.069	0.104	0.055	0.070	0.070	0.056	0.043
100	10	0.250	0.232	0.245	0.307	0.223	0.083	0.118	0.199	0.204	0.230
100	10	0.500	0.543	0.577	0.727	0.479	0.274	0.497	0.421	0.462	0.558
100	20	0.000	0.076	0.059	0.075	0.105	0.062	0.073	0.089	0.073	0.053
100	20	0.250	0.329	0.354	0.427	0.384	0.181	0.245	0.246	0.273	0.315
100	20	0.500	0.754	0.805	0.916	0.669	0.552	0.757	0.560	0.614	0.706

Notes: σ_η and $\sigma_{\varepsilon,i}$ refer to the standard deviation of $\eta_{i,t}$ and $\varepsilon_{i,t}$, respectively, while ρ refers to the autoregressive coefficient of F_t. The results are based on setting $p = 0$ (constant) and using the restricted factor estimation method, which assumes that the null hypothesis is true.

Table 3: 5% size and power when testing I(0) → I(1) and $\rho = 0.6$.

T	N	σ_η	K^1_{NT}	K^2_{NT}	K^3_{NT}	R^1_{NT}	R^2_{NT}	R^3_{NT}	M^1_{NT}	M^2_{NT}	M^3_{NT}
					$\sigma_{\varepsilon,i} = 1$						
50	5	0.000	0.049	0.042	0.062	0.080	0.062	0.071	0.058	0.047	0.041
50	5	0.250	0.106	0.098	0.134	0.084	0.045	0.082	0.084	0.077	0.097
50	5	0.500	0.221	0.203	0.256	0.151	0.031	0.144	0.146	0.146	0.164
50	10	0.000	0.069	0.045	0.073	0.079	0.046	0.065	0.053	0.038	0.039
50	10	0.250	0.153	0.135	0.181	0.135	0.039	0.096	0.106	0.082	0.121
50	10	0.500	0.342	0.348	0.452	0.313	0.138	0.304	0.246	0.247	0.293
50	20	0.000	0.064	0.040	0.071	0.111	0.053	0.090	0.044	0.036	0.054
50	20	0.250	0.203	0.179	0.255	0.314	0.121	0.178	0.132	0.110	0.154
50	20	0.500	0.543	0.543	0.706	0.591	0.446	0.641	0.307	0.314	0.408
100	5	0.000	0.080	0.060	0.077	0.088	0.068	0.068	0.064	0.057	0.057
100	5	0.250	0.265	0.285	0.353	0.202	0.067	0.151	0.231	0.236	0.264
100	5	0.500	0.405	0.432	0.531	0.338	0.061	0.281	0.333	0.350	0.403
100	10	0.000	0.114	0.079	0.084	0.110	0.062	0.082	0.092	0.068	0.056
100	10	0.250	0.343	0.347	0.439	0.339	0.153	0.229	0.285	0.292	0.348
100	10	0.500	0.582	0.640	0.785	0.545	0.354	0.559	0.481	0.522	0.626
100	20	0.000	0.094	0.062	0.066	0.113	0.073	0.083	0.088	0.072	0.055
100	20	0.250	0.502	0.532	0.618	0.535	0.332	0.415	0.374	0.416	0.454
100	20	0.500	0.875	0.909	0.960	0.765	0.638	0.854	0.708	0.752	0.827
					$\sigma_{\varepsilon,i} \sim U(1,2)$						
50	5	0.000	0.075	0.058	0.092	0.092	0.058	0.099	0.063	0.058	0.068
50	5	0.250	0.087	0.073	0.107	0.084	0.042	0.089	0.073	0.059	0.073
50	5	0.500	0.168	0.157	0.200	0.118	0.042	0.120	0.112	0.105	0.136
50	10	0.000	0.074	0.047	0.071	0.097	0.058	0.092	0.049	0.047	0.046
50	10	0.250	0.109	0.084	0.115	0.103	0.053	0.086	0.070	0.063	0.091
50	10	0.500	0.209	0.200	0.315	0.216	0.098	0.194	0.145	0.149	0.205
50	20	0.000	0.054	0.029	0.074	0.111	0.064	0.094	0.049	0.036	0.041
50	20	0.250	0.133	0.105	0.171	0.182	0.069	0.117	0.104	0.072	0.097
50	20	0.500	0.320	0.309	0.453	0.382	0.247	0.392	0.205	0.197	0.279
100	5	0.000	0.093	0.076	0.096	0.106	0.061	0.095	0.083	0.069	0.067
100	5	0.250	0.186	0.173	0.214	0.142	0.056	0.114	0.160	0.146	0.153
100	5	0.500	0.322	0.352	0.448	0.283	0.065	0.238	0.272	0.280	0.331
100	10	0.000	0.088	0.064	0.069	0.098	0.050	0.070	0.070	0.059	0.043
100	10	0.250	0.226	0.230	0.275	0.225	0.085	0.124	0.203	0.199	0.218
100	10	0.500	0.504	0.542	0.677	0.458	0.261	0.449	0.398	0.422	0.519
100	20	0.000	0.089	0.069	0.069	0.106	0.065	0.075	0.080	0.063	0.051
100	20	0.250	0.328	0.347	0.401	0.388	0.181	0.233	0.249	0.260	0.294
100	20	0.500	0.678	0.734	0.872	0.623	0.498	0.701	0.529	0.578	0.668

Notes: See Table 2 for an explanation.

Table 4: 5% size and power when testing I(1) → I(0) and $\rho = 0.3$.

T	N	σ_η	K^1_{NT}	K^2_{NT}	K^3_{NT}	R^1_{NT}	R^2_{NT}	R^3_{NT}	M^1_{NT}	M^2_{NT}	M^3_{NT}
						$\sigma_{\varepsilon,i} = 1$					
50	5	0.000	0.055	0.040	0.061	0.073	0.058	0.077	0.046	0.042	0.040
50	5	0.250	0.068	0.063	0.094	0.150	0.118	0.105	0.111	0.107	0.076
50	5	0.500	0.079	0.042	0.111	0.209	0.182	0.184	0.142	0.135	0.109
50	10	0.000	0.070	0.050	0.073	0.075	0.050	0.064	0.057	0.047	0.036
50	10	0.250	0.109	0.077	0.110	0.149	0.117	0.105	0.124	0.118	0.084
50	10	0.500	0.176	0.046	0.158	0.426	0.413	0.313	0.303	0.302	0.171
50	20	0.000	0.055	0.038	0.067	0.103	0.053	0.075	0.044	0.041	0.059
50	20	0.250	0.182	0.105	0.129	0.137	0.127	0.097	0.139	0.126	0.081
50	20	0.500	0.514	0.290	0.463	0.670	0.674	0.630	0.487	0.478	0.360
100	5	0.000	0.077	0.057	0.079	0.088	0.070	0.074	0.070	0.055	0.057
100	5	0.250	0.126	0.080	0.103	0.291	0.283	0.224	0.254	0.249	0.160
100	5	0.500	0.227	0.059	0.220	0.492	0.474	0.454	0.403	0.400	0.351
100	10	0.000	0.099	0.071	0.080	0.112	0.065	0.087	0.086	0.067	0.051
100	10	0.250	0.258	0.121	0.188	0.403	0.418	0.316	0.360	0.370	0.221
100	10	0.500	0.452	0.164	0.423	0.798	0.797	0.752	0.694	0.699	0.592
100	20	0.000	0.100	0.067	0.077	0.115	0.078	0.087	0.090	0.067	0.049
100	20	0.250	0.371	0.258	0.228	0.547	0.601	0.388	0.541	0.588	0.346
100	20	0.500	0.646	0.418	0.589	0.922	0.919	0.913	0.843	0.834	0.762
						$\sigma_{\varepsilon,i} \sim U(1,2)$					
50	5	0.000	0.071	0.059	0.097	0.090	0.060	0.100	0.062	0.058	0.070
50	5	0.250	0.071	0.061	0.099	0.122	0.081	0.105	0.076	0.070	0.066
50	5	0.500	0.077	0.050	0.092	0.167	0.136	0.157	0.122	0.111	0.086
50	10	0.000	0.074	0.051	0.065	0.088	0.054	0.088	0.059	0.046	0.044
50	10	0.250	0.079	0.052	0.087	0.103	0.071	0.072	0.086	0.066	0.064
50	10	0.500	0.148	0.059	0.125	0.221	0.202	0.163	0.147	0.142	0.093
50	20	0.000	0.047	0.030	0.068	0.109	0.062	0.084	0.044	0.035	0.046
50	20	0.250	0.113	0.073	0.100	0.135	0.107	0.099	0.116	0.090	0.084
50	20	0.500	0.297	0.141	0.259	0.397	0.395	0.325	0.266	0.266	0.144
100	5	0.000	0.093	0.069	0.092	0.099	0.060	0.089	0.068	0.064	0.056
100	5	0.250	0.117	0.090	0.094	0.213	0.201	0.139	0.201	0.183	0.106
100	5	0.500	0.162	0.051	0.149	0.420	0.415	0.359	0.323	0.319	0.252
100	10	0.000	0.084	0.069	0.069	0.104	0.055	0.070	0.070	0.056	0.043
100	10	0.250	0.180	0.092	0.136	0.283	0.262	0.183	0.245	0.258	0.160
100	10	0.500	0.352	0.107	0.284	0.724	0.734	0.659	0.643	0.649	0.506
100	20	0.000	0.076	0.059	0.075	0.105	0.062	0.073	0.089	0.073	0.053
100	20	0.250	0.275	0.154	0.158	0.319	0.324	0.200	0.291	0.300	0.158
100	20	0.500	0.655	0.376	0.556	0.928	0.932	0.903	0.858	0.860	0.721

Notes: See Table 2 for an explanation.

Table 5: 5% size and power when testing I(1) → I(0) and $\rho = 0.6$.

T	N	σ_η	K^1_{NT}	K^2_{NT}	K^3_{NT}	R^1_{NT}	R^2_{NT}	R^3_{NT}	M^1_{NT}	M^2_{NT}	M^3_{NT}
						$\sigma_{\varepsilon,i} = 1$					
50	5	0.000	0.049	0.042	0.062	0.080	0.062	0.071	0.058	0.047	0.041
50	5	0.250	0.070	0.059	0.089	0.145	0.119	0.106	0.098	0.102	0.078
50	5	0.500	0.099	0.047	0.115	0.190	0.171	0.177	0.141	0.134	0.109
50	10	0.000	0.069	0.045	0.073	0.079	0.046	0.065	0.053	0.038	0.039
50	10	0.250	0.098	0.074	0.106	0.167	0.115	0.114	0.129	0.099	0.077
50	10	0.500	0.177	0.045	0.172	0.411	0.393	0.315	0.304	0.299	0.173
50	20	0.000	0.064	0.040	0.071	0.111	0.053	0.090	0.044	0.036	0.054
50	20	0.250	0.182	0.097	0.130	0.157	0.124	0.097	0.141	0.129	0.086
50	20	0.500	0.483	0.278	0.450	0.613	0.611	0.584	0.434	0.428	0.336
100	5	0.000	0.080	0.060	0.077	0.088	0.068	0.068	0.064	0.057	0.057
100	5	0.250	0.133	0.084	0.108	0.269	0.262	0.217	0.271	0.261	0.160
100	5	0.500	0.209	0.060	0.212	0.483	0.467	0.447	0.392	0.386	0.325
100	10	0.000	0.114	0.079	0.084	0.110	0.062	0.082	0.092	0.068	0.056
100	10	0.250	0.264	0.138	0.182	0.377	0.383	0.285	0.355	0.370	0.212
100	10	0.500	0.437	0.150	0.387	0.750	0.743	0.710	0.628	0.643	0.539
100	20	0.000	0.094	0.062	0.066	0.113	0.073	0.083	0.088	0.072	0.055
100	20	0.250	0.364	0.283	0.252	0.484	0.528	0.350	0.518	0.575	0.320
100	20	0.500	0.639	0.410	0.590	0.909	0.901	0.887	0.800	0.797	0.719
						$\sigma_{\varepsilon,i} \sim U(1,2)$					
50	5	0.000	0.075	0.058	0.092	0.092	0.058	0.099	0.063	0.058	0.068
50	5	0.250	0.084	0.067	0.100	0.118	0.080	0.106	0.078	0.070	0.076
50	5	0.500	0.082	0.046	0.107	0.155	0.132	0.153	0.120	0.109	0.101
50	10	0.000	0.074	0.047	0.071	0.097	0.058	0.092	0.049	0.047	0.046
50	10	0.250	0.074	0.045	0.086	0.116	0.080	0.084	0.083	0.066	0.072
50	10	0.500	0.149	0.063	0.125	0.236	0.213	0.167	0.165	0.154	0.106
50	20	0.000	0.054	0.029	0.074	0.111	0.064	0.094	0.049	0.036	0.041
50	20	0.250	0.124	0.076	0.111	0.162	0.102	0.104	0.109	0.093	0.090
50	20	0.500	0.297	0.152	0.246	0.371	0.356	0.294	0.257	0.254	0.161
100	5	0.000	0.093	0.076	0.096	0.106	0.061	0.095	0.083	0.069	0.067
100	5	0.250	0.110	0.080	0.081	0.204	0.179	0.138	0.182	0.179	0.113
100	5	0.500	0.160	0.056	0.154	0.402	0.390	0.341	0.325	0.330	0.237
100	10	0.000	0.088	0.064	0.069	0.098	0.050	0.070	0.070	0.059	0.043
100	10	0.250	0.177	0.079	0.122	0.287	0.268	0.192	0.247	0.231	0.127
100	10	0.500	0.343	0.108	0.261	0.659	0.664	0.598	0.576	0.581	0.458
100	20	0.000	0.089	0.069	0.069	0.106	0.065	0.075	0.080	0.063	0.051
100	20	0.250	0.279	0.177	0.175	0.299	0.303	0.167	0.279	0.297	0.154
100	20	0.500	0.588	0.339	0.494	0.875	0.896	0.831	0.781	0.793	0.654

Notes: See Table 2 for an explanation.

Table 6: Empirical test results.

Statistic	Unrestricted	Restricted
K_{NT}^1	-0.921	-0.881
K_{NT}^2	-0.658	-0.645
K_{NT}^3	-0.985	-0.947
R_{NT}^1	4.002***	2.439**
R_{NT}^2	4.565***	2.636***
R_{NT}^3	4.844***	3.156***
M_{NT}^1	3.027***	1.662*
M_{NT}^2	3.293***	1.744*
M_{NT}^3	3.862***	2.293**

Notes: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. While the restricted factor estimation method assumes that the null hypothesis is true, the unrestricted method does not.