OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 68, SUPPLEMENT (2006) 0305-9049

Cointegration Testing in Panels with Common Factors*

CHRISTIAN GENGENBACH, FRANZ C. PALM and JEAN-PIERRE URBAIN

Department of Quantitative Economics, Universiteit Maastricht, Maastricht, The Netherlands (e-mail: c.gengenbach@ke.unimaas.nl; f.palm@ke.unimaas.nl; j.urbain@ke.unimaas.nl)

Abstract

Panel unit-root and no-cointegration tests that rely on cross-sectional independence of the panel unit experience severe size distortions when this assumption is violated, as has, for example, been shown by Banerjee, Marcellino and Osbat [*Econometrics Journal* (2004), Vol. 7, pp. 322–340; *Empirical Economics* (2005), Vol. 30, pp. 77–91] via Monte Carlo simulations. Several studies have recently addressed this issue for panel unit-root tests using a common factor structure to model the cross-sectional dependence, but not much work has been done yet for panel no-cointegration tests. This paper proposes a model for panel no-cointegration using an unobserved common factor structure, following the study by Bai and Ng [*Econometrica* (2004), Vol. 72, pp. 1127–1177] for panel unit roots. We distinguish two important cases: (i) the case when the non-stationarity in the data is driven by a reduced number of common stochastic trends, and (ii) the case where we have common and idiosyncratic stochastic trends present in the data. We discuss the homogeneity restrictions on the cointegrating vectors resulting from the presence of common factor cointegration. Furthermore, we study the asymptotic behaviour of some existing

*Previous versions of this paper were presented at the 12th International Panel Conference in Copenhagen, at the Conference on Frontiers in Times Series Analysis in Olbia, at a seminar at the Econometric Institute in Rotterdam and at the NSF/NBER Time Series Conference in Heidelberg. The authors would like to thank conference and seminar participants and in particular Jörg Breitung, Josep Carrion-i-Silvestre, Peter Pedroni, Joakim Westerlund and three unknown referees for helpful comments and suggestions on a previous draft. The usual disclaimer applies.

JEL Classification numbers: C32, C33.

residual-based panel no-cointegration tests, as suggested by Kao [*Journal of Econometrics* (1999), Vol. 90, pp. 1–44] and Pedroni [*Econometric Theory* (2004a), Vol. 20, pp. 597–625]. Under the data-generating processes (DGP) used, the test statistics are no longer asymptotically normal, and convergence occurs at rate T rather than $\sqrt{N}T$ as for independent panels. We then examine the possibilities of testing for various forms of no-cointegration by extracting the common factors and individual components from the observed data directly and then testing for no-cointegration using residual-based panel tests applied to the defactored data.

I. Introduction

The effect on panel unit-root tests of persistent cross-sectional dependence has been analysed in some detail in Monte Carlo simulations (Banerjee, Marcellino and Osbat, 2005) or by asymptotic analysis (Lyhagen, 2000; Pedroni and Urbain, 2001). First-generation panel unit-root tests are found to display dramatic size distortions or even worse to diverge with the cross-sectional dimension of the panel. To overcome these problems, new panel unit-root tests have been proposed that model the possibly persistent cross-sectional dependence using common factor models (see Breitung and Pesaran, 2005, for a recent overview).

For tests for the null of no-cointegration, few studies have so far been carried out. Banerjee, Marcellino and Osbat (2004) conduct an extensive Monte Carlo study where they conclude that while all statistics investigated (residual-based tests or likelihood-based trace-type test) are affected, the presence of cross-sectional cointegration appears much less harmful for single-equation tests than for the panel version of the Johansen test. In many cases, in the presence of cointegration between units of the panel, these tests cannot discriminate between cointegration among the units and cointegration for a single unit of the panel. Bai and Kao (2004) and Banerjee and Carrion-i-Silvestre (2006) study tests for panel no-cointegration with cross-sectional dependence using residual-based tests for a single cointegration relationship. The error term of the cointegrating equation follows a common factor structure as in Bai and Ng (2004). Urbain (2004), on the other hand, studies analytically the issue of spurious regression in panels when the units are cointegrated along the cross-sectional dimension, i.e. when there is cross-member cointegration. In contrast with the spurious regression results for independent panels studied by Phillips and Moon (1999), Pedroni (1995) or Kao (1999), these estimators are often not consistent and in fact converge to non-degenerate limiting distributions once the observed non-stationarity is generated by a reduced number of common stochastic trends.

This paper builds on these results to study panel tests for no-cointegration when the cross-sectional dependence in the panel is modelled by a common factor structure as in Bai and Ng (2004). Two different cases are considered that we believe are of theoretical and empirical relevance: (i) the case where the observed non-stationarity in the variables originates from cross-sectional common trends only; (ii) the case where we have both cross-sectional common and idiosyncratic stochastic trends. The spurious regression analysis for the first case reported in Urbain (2004) corresponds to the cross-member cointegration case. The second case is considered by Moon and Perron (2004) and Pesaran (2006) in the context of panel unit-root analysis and excludes the existence of cross-unit cointegration in the panel because both components are $I(1)$.

For both classes of data-generating processes (DGPs), we discuss the homogeneity restrictions on the cointegrating vectors resulting from the presence of common factor cointegration. These implications of the common factor cointegration are important reasons for proposing a sequential approach whereby the data are decomposed into common and idiosyncratic components and (no-)cointegration is tested for these components separately. Then, we study analytically the behaviour of several tests for panel cointegration including Kao's (1999) and Pedroni's (1999, 2004a) residual-based panel no-cointegration tests that have been widely used in empirical studies. For example, when the number of common factors generating the non-stationarity in the panel is kept fixed while the cross-sectional dimension of the panel increases, the Gaussian limiting results derived for the independent case are no longer valid. Tests that are based on pooled or least-squares dummy variable (LSDV) estimation of the underlying panel cointegration static regression in some cases diverge with $\sqrt{N}$ and hence important size distortions can occur for moderate values of N. Group-mean statistics are also affected and not asymptotically Gaussian. These results complement and help to have a better understanding of some of the Monte Carlo results reported by Banerjee *et al.* (2004). We then examine the possibilities of testing for no-cointegration, using residual-based panel tests applied to the defactored data.

The paper is organized as follows: in section II we present our model for panel no-cointegration with a common factor structure. In section III we examine the asymptotic behaviour of some residual-based panel no-cointegration tests when the data are generated by our DGP. Section IV discusses defactoring the data prior to testing for various forms of no-cointegration when the data contain unobserved common factors. The finite sample behaviour of the proposed approach is analysed in section V. Conclusions are drawn in section VI.

A note on notation: throughout the text, M is used to denote a generic positive number, not depending on T or N. For a matrix $\mathbf{A}$, $\mathbf{A} > 0$ denotes that $\mathbf{A}$

is positive-definite. Furthermore, $\|\mathbf{A}\| = \mathrm{trace}(\mathbf{A}'\mathbf{A})^{\frac{1}{2}}$. We write the integral $\int_0^1 B(r)\,\mathrm{d}r$ as $\int B$, and $\int_0^1 B(r)B(r)'\mathrm{d}r$ as $\int BB'$. Furthermore, '$\Rightarrow$' denotes weak convergence, and '$\xrightarrow{p}$' denotes convergence in probability. For any number x, $\lfloor x \rfloor$ denotes the largest integer smaller than x. For any variable $X_{i,t}$,

$$\tilde{X}_{i,t} = X_{i,t} - \frac{1}{T}\sum_{s=1}^{T} X_{i,s}.$$

Similarly, for any Brownian motion B, $\tilde{B} = B - \int B$. Throughout the paper, we employ sequential limit theory,[1] where we consider $T \to \infty$ followed by $N \to \infty$.

II. The model

We consider balanced panels with N cross-sectional units and T time-series observations, indexed by $i = 1, \ldots, N$ and $t = 1, \ldots, T$, respectively. For each unit in the panel, we observe a $(1 + m)$-dimensional vector of variables $Z_{i,t} = (Y_{i,t}, X'_{i,t})'$, where $Y_{i,t}$ is a scalar time series and $X_{i,t}$ an m-vector time series.[2] We assume that the DGP for $Z_{i,t}$ has a common factor structure as, e.g. in Bai and Ng (2004), and we assume the presence of k common factors in the data. Furthermore, we assume the number of common factors to be fixed as T, $N \to \infty$ throughout the paper. Our model is given by

$$Z_{i,t} = D_{i,t} + \Lambda_i F_t + E_{i,t}, \qquad (1)$$

$t = 1, \ldots, T$, $i = 1, \ldots, N$. $D_{i,t}$ is an unobserved deterministic component such that either $D_{i,t} = 0$ for all i and t if there are no deterministic components present, $D_{i,t} = d_{0i}$ for all t if the data contain individual-specific fixed effects, or $D_i = d_{0i} + d_{1i}t$ if the data contain individual-specific deterministic linear time trends, where the coefficients d_{0i} and d_{1i} depend on i only. For the remainder of the paper we assume $D_{i,t} = 0$, unless mentioned otherwise. The common component in $Z_{i,t}$ is given by F_t in equation (1). F_t is a k-vector of common $I(1)$ factors given by

$$F_t = F_{t-1} + f_t, \qquad (2)$$

where $f_t = \Phi(L)\eta_t$, η_t is a sequence of $(k \times 1)$ identically and independently distributed (i.i.d.) $(0, I_k)$ random vectors, and

[1]Although sequential limits are sometimes restrictive, they correspond to joint limits under certain restrictions (see, e.g. Phillips and Moon, 1999). Furthermore, sequential asymptotic theory is well established in the literature.

[2]We assume that, e.g. economic theory leads to a natural choice of Y in such a way that swapping some X for Y would not make sense. Nevertheless, the choice of Y is an interesting topic in cointegration analysis, but beyond the scope of this work.

$$\Phi(L) = \sum_{j=0}^{\infty} \Phi_j L^j.$$

The $(1 + m) \times k$ matrix of factor loadings Λ_i is assumed to be of full rank and block-diagonal, with block diagonality corresponding to the partition of $Z_{i,t}$, and diagonal blocks denote a λ_{1i}' and λ_{2i}' for the upper left and lower right block, respectively.

As for the vectors of observations $Z_{i,t}$, we have partitions for the unobserved vector of common factors $F_t = (F_t^{Y'}, F_t^{X'})'$, where F_t^Y and F_t^X have k_Y and k_X elements, respectively, and the partition of F_t corresponds to the structure of Λ_i, such that λ_{1i} is $k_Y \times 1$ and λ_{2i} is $k_X \times m$. The block-diagonal structure for the factor loadings is necessary to ensure that $Y_{i,t}$ and $X_{i,t}$ are not cointegrated when the non-stationarity in the data is driven by the common factors alone. When the idiosyncratic components are non-stationary as well, this assumption on Λ_i might be relaxed and a more general structure can be considered.

For the idiosyncratic component in equation (1), $E_{i,t}$, we distinguish two cases, namely stationary and non-stationary idiosyncratic components. For the former case we have

$$E_{i,t} = e_{i,t}, \tag{3}$$

while in the latter case we assume

$$E_{i,t} = E_{i,t-1} + e_{i,t}, \tag{4}$$

where the stationary vector $e_{i,t} = \Gamma_i(L)\varepsilon_{i,t}$ with $\varepsilon_{i,t}$ being a sequence of i.i.d.$(0, \Sigma_i)$ random vectors,

$$\Gamma_i(L) = \sum_{j=0}^{\infty} \Gamma_{ij} L^j.$$

Again, we partition $E_{i,t}$ conformable with the data $Z_{i,t}$, such that $E_{i,t} = (E_{i,t}^Y, E_{i,t}^{X'})'$, where $E_{i,t}^Y$ is a scalar time series and $E_{i,t}^X$ has m elements.

For the above-given model we specify the following assumptions, where M denotes a generic positive real number:

Assumption 1. Common factors: (i) $\eta_t \sim$ i.i.d.$(0, I_k)$ with finite fourth moments, (ii) there is an M such that

$$\sum_{j=0}^{\infty} j \cdot \|\Phi_j\| < M,$$

(iii) $\text{rank}(\Phi(1)) = k$, (iv) $E\|F_0\| \leq M$.

Assumption 2. Factor loadings: (i) for non-random λ_{1i} and λ_{2i}, $\|\lambda_{1i}\| \le M$ and $\|\lambda_{2i}\| \le M$; for random λ_{1i} and λ_{2i}, $E\|\lambda_{1i}\|^4 \le M$ and $E\|\lambda_{2i}\|^4 \le M$,

(ii) $$N^{-1}\sum_{i=1}^{N}\Lambda_i'\Lambda_i \xrightarrow{p} \Sigma_\Lambda > 0,$$

(iii) for non-random λ_{1i} and λ_{2i},

$$N^{-1}\sum_{i=1}^{N}\lambda_{1i} \neq 0 \quad \text{and} \quad N^{-1}\sum_{i=1}^{N}\lambda_{2i} \neq 0;$$

for random λ_{1i} and λ_{2i}, $E(\lambda_{1i}) \neq 0$ and $E(\lambda_{2i}) \neq 0$.

Assumption 3. Idiosyncratic components: for each $i = 1, \ldots, N$, (i) $\varepsilon_{i,t} \sim$ i.i.d.$(0, \Sigma_i)$ with finite eighth moments, and $\varepsilon_{i,t}$ and $\varepsilon_{j,s}$ are independent for any t, s and $i \neq j$, (ii) $E\|\varepsilon_{i,0}\| < M$, (iii) $\Gamma_i(L)$ fulfils the random coefficients and summability conditions from Phillips and Moon (1999, Assumptions 1 and 2 on p. 1060 and p. 1061 respectively), (iv) $\text{rank}(\Gamma_i(1)) = m + 1$, $\forall i$.

Assumption 4. The errors, η_t, $\varepsilon_{i,t}$, and the factor loadings Λ_i form mutually independent groups.

Under the conditions of Assumption 1, the common factors F_t form a k-dimensional $I(1)$ process and the possibility of cointegration between the common factors is excluded. The full-rank assumption on the long-run covariance matrix of F_t could in fact be relaxed, as long as the diagonal blocks corresponding to the long-run covariances of F_t^Y and F_t^X each have at least rank 1. The long-run covariance matrix of the common factors is given by (see e.g. Phillips and Durlauf, 1986)

$$\Omega = \Phi(1)\Phi(1)' = \Xi + \Theta + \Theta',$$

where

$$\Xi = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E(f_t f_t') \quad \text{and} \quad \Theta = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E(f_t F_{t-1}').$$

Furthermore, an invariance principle holds such that

$$T^{-1/2}F_{\lfloor rT\rfloor} \Rightarrow B_F(r) \quad \text{as} \quad T \to \infty, \tag{5}$$

where B_F is a k-vector Brownian motion with covariance matrix Ω. Assumptions 2(i) and 2(ii) are standard assumptions for factor models and ensure that the factor loadings are identifiable. Assumption 2(iii) is needed for the spurious regression results when the non-stationarity in the data is only driven by the common factors. Assumption 3(iii) specifies that a panel functional central limit theorem holds for $S_{i,t} = \sum_{s=1}^{t} e_{i,t}$, which corresponds to $E_{i,t}$ in case the idiosyncratic components are non-stationary as in equation (4), or to its cumulative sum if equation (3) is true. The long-run covariance matrix of $S_{i,t}$ is given by

$$\Psi_i = \Gamma_i(1)\Sigma_i\Gamma_i(1)' = \Upsilon_i + \Delta_i + \Delta_i',$$

where

$$\Upsilon_i = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E(e_{i,t}e_{i,t}') \quad \text{and} \quad \Delta_i = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E(e_{i,t}S_{i,t-1}'),$$

and an invariance principle ensures that

$$T^{-1/2}S_{i,\lfloor rT\rfloor} \Rightarrow B_i(r) \quad \text{as} \quad T\to\infty, \tag{6}$$

where B_i is a randomly scaled $(1 + m)$-vector Brownian motion with covariance matrix Ψ_i. Assumption 3(iv) ensures that the idiosyncratic terms do not cointegrate in case these are $I(1)$ vectors.

The implications of these assumptions are best understood by considering the Beveridge–Nelson (BN) decomposition for F_t and for $E_{i,t} = \sum_{s=1}^{t} e_{i,s}$:

$$F_t = \Phi(1)\sum_{s=1}^{t}\eta_s + \Phi^*(L)(\eta_t - \eta_0) + F_0, \tag{7}$$

$$E_{i,t} = \Gamma_i(1)\sum_{s=1}^{t}\varepsilon_{i,s} + \Gamma_i^*(L)(\varepsilon_{i,t} - \varepsilon_{i,0}) + E_{i,0}, \tag{8}$$

where

$$\Phi^*(L) = \sum_{j=0}^{\infty}\Phi_j^* L^j \quad \text{with} \quad \Phi_j^* = -\sum_{l=j+1}^{\infty}\Phi_l,$$

$$\Gamma_i^*(L) = \sum_{j=0}^{\infty}\Gamma_{i,j}^* L^j \quad \text{with} \quad \Gamma_{i,j}^* = -\sum_{l=j+1}^{\infty}\Gamma_{i,l},$$

$\Phi^*(L)(\eta_t - \eta_0)$ and $\Gamma_i^*(L)(\varepsilon_{i,t} - \varepsilon_{i,0})$ are stationary with finite fourth-order moments and F_0 and $E_{i,0}$ are $O_p(1)$ by assumption.

If equation (3) is true, the idiosyncratic data components are $I(0)$, and the $I(1)$ trends of the common factors contained in $\Lambda_i\Phi(1)\sum_{s=1}^{t}\eta_s$ drive the non-stationarity in the data. Then, we might observe *cross-member cointegration* between some $Y_{i,t}$ and $Y_{j,t}$, and between some $X_{i,t}$ and $X_{j,t}$ for some i, j, $i \neq j$, the exact cointegration structure depending on the individual loadings. The assumption on the block-diagonal structure of the factor loadings Λ_i in turn implies that we have separation in a cointegrating system (see Hecq, Palm and

Urbain, 2002). Note that the assumption of cointegration between $Y_{i,t}$ and $X_{i,t}$ would only be possible if the common factors F_t^Y and F_t^X would cointegrate, which is ruled out by Assumption 1 from which the full rank of the long-run covariance matrix of F_t follows.

When $E_{i,t}$ is given by equation (4), both common and idiosyncratic data components are non-stationary. Furthermore, the idiosyncratic components do not cointegrate along the cross-section. Hence, we do not have cointegration 'within' units, e.g. between $Y_{i,t}$ or $X_{i,t}$. The BN decomposition of $Z_{i,t}$ is easily obtained from equations (1) and (7–8) and shows that the non-stationarity of $Z_{i,t}$ stems from the term $\Lambda_i\Phi(1)\sum_{s=1}^{t}\eta_s + \Gamma_i(1)\sum_{s=1}^{t}\varepsilon_{i,s}$.

Remark 1. To investigate tests for no-cointegration we need to maintain the assumption that there does not exist a full-column rank matrix β_i' such that $\beta_i'Z_{it} \sim I(0)$. Different cases can be considered. Two cases are important, namely one with cross-member cointegration where we have $I(1)$ common factors and $I(0)$ idiosyncratic terms and one where the panel units contain common stochastic trends, but do not cointegrate even along the cross-sectional dimension so that both the common and the idiosyncratic components are $I(1)$.

Remark 2: Heterogeneity and cross-sectional dependence. With $I(1)$ common factors as well as $I(1)$ idiosyncratic components, we actually have two different sets of possible cointegrating vectors that would annihilate the idiosyncratic and the common $I(1)$ stochastic trends, respectively (see also the discussion in Gregoir, 2005; Breitung and Pesaran, 2005). Combining equations (1) and (7)–(8), the resulting BN representation of $Z_{i,t}$ shows that it will not be easy to annihilate both. In particular, cointegrating vector(s), say δ, that annihilate the common $I(1)$ components should lie in the left null space of Λ_i, that is $\delta\Lambda_i\Phi(1) = 0$ as $\Phi(1)$ is of full rank by Assumption 1, while those for the idiosyncratic components, say γ_i' would have to lie in the left null space of $\Gamma_i(1)$, i.e. $\gamma_i'\Gamma_i(1) = 0$. If the intersection of these left null spaces is empty, there does not exist a cointegrating relationship that annihilates both the unit roots from the common stochastic trends and those of the idiosyncratic terms. In this case, none of the $Z_{i,t}$ vectors is cointegrated. The components taken in isolation could be cointegrated though.

In fact, there is an important trade-off between the degree of heterogeneity that can be allowed for and the existence of cross-sectional dependence modelled by common factors. Without loss of generality, consider the following simple bivariate DGP where we have a single $I(1)$ common factor in Y and a single $I(1)$ common factor in X:

$$Y_{i,t} = \lambda_{1,i}F_t^Y + E_{i,t}^Y, \tag{9}$$

$$X_{i,t} = \lambda_{2,i} F_t^X + E_{i,t}^X, \tag{10}$$

from which we see that any linear combination can be written as

$$Y_{i,t} - \beta_i X_{i,t} = \lambda_{1,i} \left(F_t^Y - \frac{\beta_i \lambda_{2,i}}{\lambda_{1,i}} F_t^X \right) + E_{i,t}^Y - \beta_i E_{i,t}^X. \tag{11}$$

For the linear combination $(1, -\beta_i)'$ to be a cointegrating vector such that $Y_{i,t} - \beta_i X_{i,t} \sim I(0)$, two conditions need to hold, namely

(i) $\left(F_t^Y - \frac{\beta_i \lambda_{2,i}}{\lambda_{1,i}} F_t^X \right) \sim I(0)$
(ii) $(E_{i,t}^Y - \beta_i E_{i,t}^X) \sim I(0)$.

Given that here we have only two $I(1)$ common factors, there can be at most a single linear cointegrating combination between these factors and hence $\beta_i \lambda_{2,i} / \lambda_{1,i}$ should be the same $\forall i$. Three different cases are compatible with a constant (over i) ratio:

1 With *homogeneity* of the factor loadings and of β_i the ratio $\beta_i \lambda_{2,i} / \lambda_{1,i} = \beta \lambda_2 / \lambda_1$ does not depend on i. A similar restriction is considered by Gregoir (2005). Another possibility is homogeneity of β_i and constancy of the ratios of the factor loadings $\lambda_{2,i} / \lambda_{1,i}$ for all i which is also excluded by Assumptions 1–4.
2 The second case allows for some degree of heterogeneity: the factor loadings vary with β_i such that the ratio $\beta_i \lambda_{2,i} / \lambda_{1,i}$ is constant across i. This is excluded by Assumptions 1–4 where the loadings and Ψ_i are assumed to vary independently of each other.
3 A third case arises when for all i the variables $Y_{i,t}$ and $X_{i,t}$ have a single common source of nonstationarity F_t only. The idiosyncratic component is assumed to be stationary (or could be cointegrated with cointegrating vector β_i). In this case, $Y_{i,t}$ and $X_{i,t}$ are cointegrated with $\beta_i = \lambda_{1,i} / \lambda_{2,i}$. This is ruled out by the assumption of block diagonality of Λ_i, but it would be a natural alternative hypothesis to the null of no-cointegration. Homogeneity of the cointegrating vector then arises if $\lambda_{1,i} / \lambda_{2,i}$ is constant across entities i.

To conclude, if we allow for almost unrestricted (under Assumptions 1–4) heterogeneity, the existence of cointegrating relations that annihilate *both* the common and idiosyncratic $I(1)$ stochastic trends is very unlikely. The consequences of this for testing of the null of no-cointegration in this factor set-up will be mentioned in section IV.

Remark 3. A similar framework is also, independently of the present study, proposed by Dees *et al.* (2005) for the study of macroeconomic linkages within the Euro area. The purpose of their study was however different, as no attempt to discuss tests for cointegration is made. This study is thus complementary to theirs.

III. The behaviour of panel residual-based tests

The purpose of this section is to study, given the set-up introduced in the preceding section, the asymptotic behaviour of some standard and popular panel tests for no-cointegration. The statistics we consider are designed to test for the presence of a single cointegration relationship between $Y_{i,t}$ and $X_{i,t}$.[3] Kao (1999) considers a homogenous cointegrating vector, whereas Pedroni (1999) allows for heterogeneity. However, both rely on the cross-sectional independence of the panel unit to derive asymptotic normality for their test statistics.

3.1. Kao (1999)

Kao (1999) proposes estimating the homogeneous cointegrating relationship by pooled regression allowing for individual fixed effects. The regression equation is given by

$$Y_{i,t} = \alpha_i + \beta X_{i,t} + u_{i,t}, \tag{12}$$

where β and $X_{i,t}$ are row and column vectors, respectively, and $u_{i,t}$ is a regression error. The LSDV estimator for β is

$$\tilde{\beta} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T} \tilde{Y}_{i,t}\tilde{X}'_{i,t}\right)\left(\sum_{i=1}^{N}\sum_{t=1}^{T} \tilde{X}_{i,t}\tilde{X}'_{i,t}\right)^{-1},$$

where

$$\tilde{Y}_{i,t} = Y_{i,t} - \frac{1}{T}\sum_{s=1}^{T} Y_{i,s} \quad \text{and} \quad \tilde{X}_{i,t} = X_{i,t} - \frac{1}{T}\sum_{s=1}^{T} X_{i,s}.$$

The residuals from this first-stage regression $\tilde{u}_{i,t} = \tilde{Y}_{i,t} - \tilde{\beta}\tilde{X}_{i,t}$ will still contain a unit root under the null hypothesis of no cointegration. We now estimate a pooled Dickey–Fuller (DF) regression

$$\Delta\tilde{u}_{i,t} = (\rho - 1)\tilde{u}_{i,t-1} + v_{i,t}, \tag{13}$$

where the pooled ordinary least squares (POLS) estimator of $(\rho - 1)$ is given by

[3]This is a restrictive assumption that we, however, will make in what follows. Approaches that allow for more than one cointegrating vector are reviewed in Breitung and Pesaran (2005).

$$(\tilde{\rho}-1)=\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\Delta\tilde{u}_{i,t}\tilde{u}_{i,t-1}\right)\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\tilde{u}_{i,t-1}^{2}\right)^{-1}.$$

Kao's (1999) tests are based on $\tilde{\rho}$ and the corresponding t-statistic

$$t_{\tilde{\rho}}=(\tilde{\rho}-1)\left(\hat{s}_{\tilde{u}}^{2}\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\tilde{u}_{i,t-1}^{2}\right)^{-1}\right)^{-\frac{1}{2}},$$

where

$$\hat{s}_{\tilde{u}}^{2}=N^{-1}T^{-1}\sum_{i=1}^{N}\sum_{t=2}^{T}(\Delta\tilde{u}_{i,t-1}-(\tilde{\rho}-1)\tilde{u}_{i,t-1})^{2},$$

corrected for endogeneity and serial correlation. When the panel units are cross-sectionally independent, the test statistics are asymptotically normally distributed as $T\to\infty$ followed by $N\to\infty$. However, for the model given by equations (1), (2) and (3) or (4), this assumption is clearly violated. Using the results reported in Lemmas 1–3 in Appendix A, we obtain the following limit results:

$$\begin{aligned}
\mathrm{vec}\left(\int \mathrm{d}B_{F\Lambda}B_{F\Lambda}'\right)&=\breve{\mathbf{\Lambda}}\,\mathrm{vec}\left(\int \mathrm{d}B_{F}B_{F}'\right),\quad \mathrm{vec}(\Theta_{F\Lambda})=\breve{\mathbf{\Lambda}}\,\mathrm{vec}(\Theta),\\
\mathrm{vec}\left(\int B_{F\Lambda}B_{F\Lambda}'\right)&=\breve{\mathbf{\Lambda}}\,\mathrm{vec}\left(\int B_{F}B_{F}'\right),\\
\mathrm{vec}\left(\int \mathrm{d}B_{F\Lambda}\tilde{B}_{F\Lambda}'\right)&=\breve{\mathbf{\Lambda}}\,\mathrm{vec}\left(\int \mathrm{d}B_{F}\tilde{B}_{F}'\right),\\
\mathrm{vec}\left(\int \tilde{B}_{F\Lambda}\tilde{B}_{F\Lambda}'\right)&=\breve{\mathbf{\Lambda}}\,\mathrm{vec}\left(\int \tilde{B}_{F}\tilde{B}_{F}'\right),\quad\text{and}\quad \breve{\mathbf{\Lambda}}=\operatorname*{p\,lim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}(\Lambda_{i}\otimes\Lambda_{i}).
\end{aligned}$$

Ψ^{YX} is the long-run average covariance between the idiosyncratic errors in $Y_{i,t}$ and $X_{i,t}$, Ψ^{XX} is the long-run average covariance matrix of the idiosyncratic errors in $X_{i,t}$, and B_F and B_i are given in equations (5) and (6), respectively.

Proposition 1. Given Assumptions 1, 2, 3 and 4:

(A) Consider the model given by equations (1), (2) and (3),

(a) $\tilde{\beta}\Rightarrow(\int\tilde{B}_{F\Lambda}^{Y}\tilde{B}_{F\Lambda}^{X'})(\int\tilde{B}_{F\Lambda}^{X}\tilde{B}_{F\Lambda}^{X'})^{-1}=\tilde{\mathbf{b}}_{A}$ as T, $N\to\infty$ sequentially,

(b) $T(\tilde{\rho}-1)\Rightarrow\dfrac{(1,-\tilde{\mathbf{b}}_{A})\left(\int \mathrm{d}B_{F\Lambda}\tilde{B}_{F\Lambda}'+\Theta_{F\Lambda}+\gamma_{1}-\Upsilon\right)(1,-\tilde{\mathbf{b}}_{A})'}{(1,-\tilde{\mathbf{b}}_{A})\left(\int\tilde{B}_{F\Lambda}\tilde{B}_{F\Lambda}'\right)(1,-\tilde{\mathbf{b}}_{A})'}$

as $T,N\to\infty$ sequentially, where $\gamma_1=E(\gamma_{i1})$ and $\gamma_{i1}=E(\tilde{e}_{i,t-1}\tilde{e}_{i,t}')$,

(c) $t_{\tilde{\rho}}$ diverges at rate $\sqrt{N}$ as T, $N\to\infty$ sequentially.

(B) Consider the model given by equations (1), (2) and (4),

(a) $\tilde{\beta} \Rightarrow (\int \tilde{B}^{Y}_{F\Lambda}\tilde{B}^{X'}_{F\Lambda} + \frac{1}{6}\Psi^{YX})(\int \tilde{B}^{X}_{F\Lambda}\tilde{B}^{X'}_{F\Lambda} + \frac{1}{6}\Psi^{XX})^{-1} = \tilde{\mathbf{b}}_B$ as T, $N \to \infty$ sequentially,

(b) $T(\tilde{\rho} - 1) \Rightarrow \dfrac{(1, -\tilde{\mathbf{b}}_B)\left(\int \mathrm{d}B_{F\Lambda}\tilde{B}'_{F\Lambda} + \Theta_{F\Lambda} - \frac{1}{2}\Psi + \Delta\right)(1, -\tilde{\mathbf{b}}_B)'}{(1, -\tilde{\mathbf{b}}_B)\left(\int \tilde{B}_{F\Lambda}\tilde{B}'_{F\Lambda} + \frac{1}{6}\Psi\right)(1, -\tilde{\mathbf{b}}_B)'}$ as $T, N \to \infty$ sequentially,

(c) $t_{\tilde{\rho}}$ diverges at rate $\sqrt{N}$ as T, $N \to \infty$ sequentially.

Proof. See Appendix. ■

The results summarized in Proposition 1 are clearly in contrast to the asymptotic normality Kao (1999) derives for the test statistics for independent panels, although we have not yet considered corrections for serial correlation and endogenous regressors. Results A(a) and B(a) are similar to those derived by Urbain (2004) for the pooled least squares estimator (PLS). This is in sharp contrast with the $\sqrt{N}$ consistency of the LSDV estimator in the case of a spurious regression estimated from independent panel data (see Phillips and Moon, 1999). The statistics proposed by Kao (1999) rely on this consistency, namely on the fact that $\tilde{\beta} \xrightarrow{p} \Psi^{YX}\Psi^{XX-1}$ where Ψ^{YX} is the long-run average covariance between the errors driving $X_{i,t}$ and those driving $Y_{i,t}$ and Ψ^{XX} is the long-run average covariance matrix of the $X_{i,t}$ values. The presence of common factors destroys this property and consequently the asymptotic normality of these estimators and of the statistics relying on this result. For the case of stationary idiosyncratic components, our findings are similar to the spurious regression results from time-series analysis. With non-stationary idiosyncratic components we obtain some mixture of time-series and panel spurious regression results in the limiting distributions. The tests are inconsistent when the data have a common factor structure, and size distortions have to be expected which will increase with N. The nuisance parameters in the limiting distributions given in Proposition 1 introduced by the serial correlation in the common factors and idiosyncratic components can be corrected for non-parametrically, i.e. the composite effect of $\Theta_{F\Lambda} + \gamma_1 - \Upsilon$ or $\Theta_{F\Lambda} + \Delta$ can be accounted for. However, it is not possible to identify nuisance parameters associated with the common factors or the idiosyncratic components individually. So, the covariance of $\tilde{B}_{F\Lambda}$ as well as the long-run average covariance matrix of idiosyncratic stochastic trends, Ψ, will in general remain in the limits. The limit of $t_{\tilde{\rho}}$ will be the product of $\sqrt{N}$, the limit of $(\tilde{\rho} - 1)$ and the limit of the standard deviation of $(\tilde{\rho} - 1)$. Whereas the standard deviation is positive, the driving factor of the limiting distribution of $(\tilde{\rho} - 1)$ is $\int \mathrm{d}B_{F\Lambda}\tilde{B}'_{F\Lambda}/\int \tilde{B}_{F\Lambda}\tilde{B}'_{F\Lambda}$ which has a negative expected value. Thus, $t_{\tilde{\rho}}$ can be expected to diverge to $-\infty$.

3.2. Pedroni (1999)

Pedroni (1999) allows for heterogeneity of the slope coefficient β in the cointegration relationship (12), which thus becomes β_i. He proposes estimating a first-stage regression individually for each panel member to obtain an estimate of β_i,

$$\tilde{\beta}_i = \left(\sum_{t=1}^{T} \tilde{Y}_{i,t}\tilde{X}'_{i,t}\right)\left(\sum_{t=1}^{T} \tilde{X}_{i,t}\tilde{X}'_{i,t}\right)^{-1}. \tag{14}$$

Pedroni (1999) proposes two classes of statistics, namely those based on the within-dimension denoted as 'panel' statistics, and those based on the between-dimension denoted as 'group mean' statistics. For the former group, the residuals from the first-stage regression, $\tilde{u}_{i,t} = \tilde{Y}_{i,t} - \tilde{\beta}_i\tilde{X}_{i,t}$, are stacked and a pooled DF regression is estimated as in equation (13).[4] The group mean statistics are based on averages of individual unit-root statistics, derived from

$$\Delta\tilde{u}_{i,t} = (\rho_i - 1)\tilde{u}_{i,t-1} + v_{i,t}, \tag{15}$$

to obtain

$$(\tilde{\rho}_i - 1) = \left(\sum_{t=2}^{T} \Delta\tilde{u}_{i,t}\tilde{u}_{i,t-1}\right)\left(\sum_{t=2}^{T} \tilde{u}^2_{i,t-1}\right)^{-1}.$$

Consider now the panel-rho statistic denoted by $Z_{\tilde{\rho}_{NT}-1}$ and the group-mean rho statistic $\tilde{Z}_{\tilde{\rho}_{NT}-1}$ given by

$$Z_{\tilde{\rho}_{NT}-1} = \left(\sum_{i=1}^{N}\sum_{t=2}^{T}(\Delta\tilde{u}_{i,t}\tilde{u}_{i,t-1} - \hat{\boldsymbol{\lambda}}_i)\right)\left(\sum_{i=1}^{N}\sum_{t=2}^{T} \tilde{u}^2_{i,t-1}\right)^{-1}, \tag{16}$$

and

$$\tilde{Z}_{\tilde{\rho}_{NT}-1} = \sum_{i=1}^{N}\left(\left(\sum_{t=2}^{T}(\Delta\tilde{u}_{i,t}\tilde{u}_{i,t-1} - \hat{\boldsymbol{\lambda}}_i)\right)\left(\sum_{t=2}^{T} \tilde{u}^2_{i,t-1}\right)^{-1}\right), \tag{17}$$

with

$$\hat{\boldsymbol{\lambda}}_i = T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}\tilde{v}_{i,t}\tilde{v}_{i,t-s}$$

where $\tilde{v}_{i,t}$ are the residuals of the second-stage regression, and J and ω_{sJ} are suitable bandwidth and kernel functions, respectively. For these two statistics, we obtain the following limiting results:

[4]Note that although the estimated DF equation is the same for Kao (1999) and Pedroni (1999), the residuals used in the estimation are obtained from individual regressions instead of a pooled one.

Proposition 2. Given Assumptions 1, 2, 3 and 4:

(A) Consider the model given by equations (1), (2) and (3),

$$\text{(a)}\ \ \tilde{\beta}_i \Rightarrow \left(\lambda'_{1i}\left(\int \tilde{B}_F^Y \tilde{B}_F^{X'}\ \lambda_{2i}\right)\left(\lambda'_{2i}\left(\int \tilde{B}_F^X \tilde{B}_F^{X'}\right)\ \lambda_{2i}\right)^{-1} = \tilde{\mathbf{b}}_{iA} \quad \text{as } T \to \infty,$$

$$\text{(b)}\ \ TZ_{\tilde{\rho}_{NT}-1} \Rightarrow \frac{\sum_{i=1}^{N} \lambda'_{1i} L'_{11} \int \mathrm{d}Q_F \tilde{Q}'_F L_{11} \lambda_{1i}}{\sum_{i=1}^{N} \lambda'_{1i} L'_{11} \int \tilde{Q}_F \tilde{Q}'_F L_{11} \lambda_{i1}} \quad \text{as } T \to \infty,$$

$$\text{(c)}\ \ T\tilde{Z}_{\tilde{\rho}_{NT}-1} \Rightarrow \sum_{i=1}^{N} \frac{\lambda'_{1i} L'_{11} \int \mathrm{d}Q_F \tilde{Q}'_F L_{11} \lambda_{1i}}{\lambda'_{1i} L'_{11} \int \tilde{Q}_F \tilde{Q}'_F L_{11} \lambda_{i1}} \quad \text{as } T \to \infty,$$

where

$$\tilde{Q} = \tilde{W}_F^Y - \left(\int \tilde{W}_F^Y \tilde{W}_F^{X\prime}\right)\left(\int \tilde{W}_F^X \tilde{W}_F^{X\prime}\right)^{-1} \tilde{W}_F^X,$$

$\tilde{W}_F$ is a demeaned *k*-vector standard Brownian motion, and L_{11} is the upper left element of L, the block triangular decomposition of $\Omega = L'L$.

(B) Consider the model given by (1), (2) and (4),

$$\text{(a)}\ \ \tilde{\beta}_i \Rightarrow \left(\lambda'_{1i} \int \tilde{B}_F^Y \tilde{B}_F^{X'}\ \lambda_{2i} + \int \tilde{B}_i^Y \tilde{B}_i^{X'} + \lambda'_{1i} \int \tilde{B}_F^Y \tilde{B}_i^{X'} + \int \tilde{B}_i^Y \tilde{B}_F^{X'}\ \lambda_{2i}\right)$$
$$\times\left(\lambda'_{2i} \int \tilde{B}_F^X \tilde{B}_F^{X'}\ \lambda_{2i} + \int \tilde{B}_i^X \tilde{B}_i^{X'} + \lambda'_{2i} \int \tilde{B}_F^X \tilde{B}_i^{X'} + \int \tilde{B}_i^X \tilde{B}_F^{X'}\ \lambda_{2i}\right)^{-1} = \tilde{\mathbf{b}}_{iB}$$
as $T \to \infty$,

$$\text{(b)}\ \ TZ_{\tilde{\rho}_{NT}-1} \Rightarrow \frac{\sum_{i=1}^{N}(1-\tilde{\mathbf{b}}_{iB})\left(\Lambda'_i\left(\int \mathrm{d}B_F \tilde{B}'_F\right)\Lambda'_i + \int \mathrm{d}B_i \tilde{B}'_i + \Lambda_i \int \mathrm{d}B_F \tilde{B}'_i + \int \mathrm{d}B_i \tilde{B}'_F \Lambda'_i\right)(1-\tilde{\mathbf{b}}_{iB})'}{\sum_{i=1}^{N}(1-\tilde{\mathbf{b}}_{iB})\left(\Lambda_i \int \tilde{B}_F \tilde{B}'_F \Lambda'_i + \int \tilde{B}_i \tilde{B}'_i + \Lambda_i \int \tilde{B}_F \tilde{B}'_i + \int \tilde{B}_i \tilde{B}'_F \Lambda'_i\right)(1-\tilde{\mathbf{b}}_{iB})'}$$

as $T \to \infty$,

$$\text{(c)}\ \ T\tilde{Z}_{\tilde{\rho}_{NT}-1} \Rightarrow \sum_{i=1}^{N} \frac{(1-\tilde{\mathbf{b}}_{iB})\left(\Lambda'_i\left(\int \mathrm{d}B_F \tilde{B}'_F\right)\Lambda'_i + \int \mathrm{d}B_i \tilde{B}'_i + \Lambda_i \int \mathrm{d}B_F \tilde{B}'_i + \int \mathrm{d}B_i \tilde{B}'_F \Lambda'_i\right)(1-\tilde{\mathbf{b}}_{iB})'}{(1-\tilde{\mathbf{b}}_{iB})\left(\Lambda_i \int \tilde{B}_F \tilde{B}'_F \Lambda'_i + \int \tilde{B}_i \tilde{B}'_i + \Lambda_i \int \tilde{B}_F \tilde{B}'_i + \int \tilde{B}_i \tilde{B}'_F \Lambda'_i\right)(1-\tilde{\mathbf{b}}_{iB})'}$$

as $T \to \infty$,

Proof. See Appendix. ■

For the panel-rho and group-mean-rho statistics, Pedroni (1999, 2004a) derives asymptotic normality when they are properly standardized. In particular, $\sqrt{N}TZ_{\tilde{\rho}_{NT}-1} - \sqrt{N}\theta_2\theta_1^{-1}$ and $N^{-\frac{1}{2}}T\tilde{Z}_{\tilde{\rho}_{NT}-1} - \sqrt{N}\tilde{\theta}_1$ are asymptotically normally distributed for independent panels, where θ_1, θ_2 and $\tilde{\theta}_1$ are means of functionals of Brownian motions (for details, see Pedroni, 2004a). The results from Proposition 2 indicate that under the DGP we consider, $TZ_{\tilde{\rho}_{NT}-1}$ and $T\tilde{Z}_{\tilde{\rho}_{NT}-1}$ converge, so that the two test statistics diverge at rate $\sqrt{N}$ when standardized as above. Furthermore, because of the presence of the common factors, the individual statistics will not be independent along the cross-section, so that the use of a Central Limit Theorem (CLT) to derive asymptotic normality of the average statistic will be invalid. The result is similar to that derived by Lyhagen (2000) for the Im–Pesaran–Shin (IPS) statistics. Moreover, for independent panels the distributions of $Z_{\tilde{\rho}_{NT}-1}$ and $\tilde{Z}_{\tilde{\rho}_{NT}-1}$ will be nuisance parameter-free. For the DGP we consider, this is not true in general. Although the composite effect of serial correlation in the common factors and idiosyncratic components can be corrected for non-parametrically, nuisance parameters coming only from the common factors or from the idiosyncratic components cannot be identified. So, the limiting distributions will in general depend on the long-run covariances of the common and/or idiosyncratic stochastic trends. A special case arises when there is a single common factor in $Y_{i,t}$ and the idiosyncratic components are stationary. Then, $\lambda_{1i}L_{11}$ will cancel from the limits given in Proposition 2A(b) and (c).

IV. A two-step procedure to test for (no-)cointegration in the presence of common factors

As shown in section III, standard panel tests for the null of no-cointegration suffer from serious problems when applied to data with a common factor structure. To tackle the problem we propose a simple approach based on the Bai and Ng (2004) PANIC methodology.[5]

A related, albeit different, idea is exploited in the work of Banerjee and Carrion-i-Silvestre (2006), who assume a factor structure for the *disturbance* of a panel static regression model:

$$Y_{i,t} = \alpha_i + \beta_i X_{i,t} + u_{i,t}$$
$$u_{i,t} = \gamma_i' F_t + E_{i,t},$$

where F_t and $E_{i,t}$ are the common factors and the idiosyncratic components, respectively, that can be either $I(1)$ or $I(0)$. A similar framework is used by Bai and Kao (2004) for the estimation of a cointegrating relationship in the presence of common factors. Under some conditions that bound the possible heterogeneity, this framework leads to panel statistics for the null of no-cointegration that have the same distribution as panel unit-root tests and hence are not affected by the number of regressors.[6]

Consider the simple bivariate DGP (9)–(10)[7] and address the issue of no-cointegration at three different levels.

[5]Wagner and Müller-Fürstenberger (2004) use similar ideas in an empirical study of the Kuznets curve.

[6]A similar set-up is retained by Westerlund (2005) who proposes Durbin–Hausman tests for cointegration in panels.

[7]The discussion extends to a more general set-up.

(i) Testing for idiosyncratic component no-cointegration: This would mean testing the null hypothesis that $(E^Y_{i,t} - \beta_i E^X_{i,t}) \sim I(1)$ against $(E^Y_{i,t} - \beta_i E^X_{i,t}) \sim I(0)$.

(ii) Testing for common factor no-cointegration: This would boil down to testing the null hypothesis that $(F^Y_t - \delta F^X_t) \sim I(1)$ against $(F^Y_t - \delta F^X_t) \sim I(0)$.

(iii) Testing for panel no-cointegration: This is testing the null hypothesis that $Y_{i,t} - \beta_i X_{i,t} \sim I(1)$ against $Y_{i,t} - \beta_i X_{i,t} \sim I(0)$. Rejecting the null of no-cointegration requires evidence of idiosyncratic component cointegration with cointegrating vector $(1, -\beta_i)'$ as well as of common factor cointegration with cointegrating vector $(1, -\beta_i \lambda_{2,i}/\lambda_{1,i})'$ which should be constant across the individuals i.

Provided the components have been extracted from the data, case (i) is tested using standard panel tests for no-cointegration given in equations (16) and (17). Case (ii) can be investigated using standard time-series no-cointegration tests such as the Johansen rank test. Case (iii) is slightly more problematic since rejecting the null of panel no-cointegration requires not only factor and idiosyncratic cointegration, but also cointegrating vector(s) for the factors of a very specific form. The restrictions between the cointegrating coefficients result from the common factor structure and from the condition that the left null spaces of the common factor and idiosyncratic component cointegration must have a non-empty intersection.

There is however a useful indirect way of addressing this question. Consider equation (11) and write $(F^Y_t - (\beta_i \lambda_{2,i}/\lambda_{1,i})F^X_t) \equiv G_t$ and $(E^Y_{i,t} - \beta_i E^X_{i,t}) \equiv E^*_{i,t}$ such that equation (11) becomes:

$$Y_{i,t} - \beta_i X_{i,t} = \lambda_{1,i} G_t + E^*_{i,t} \tag{18}$$

which is nothing but the parametrization considered in Banerjee and Carrion-i-Silvestre (2006). Under this parametrization, $(1, -\beta_i \lambda_{2,i}/\lambda_{1,i})$ will be a cointegrating vector for the common factors if and only if $G_t \sim I(0)$. One may consequently investigate the hypothesis of panel cointegration using the approach proposed by these authors.

Now we shall outline a sequential testing procedure based on the factor structure under equations (1), (2) and (3) or (4) that does not restrict the heterogeneity. The approach starts with a decomposition of the data into common factors and idiosyncratic components as in Bai and Ng (2004). It investigates the cointegration properties of the extracted factors and components.

Step 1

Conduct a PANIC analysis of each variable $X_{i,t}$ and $Y_{i,t}$ individually to extract the common factors, e.g. using the principal components approach advocated by Bai and Ng (2004). Test for unit-roots in both the factors and the idiosyncratic components using the Bai and Ng (2004) or the Breitung and Das (2005) approach.

Step 2

a. If $I(1)$ common factors and $I(0)$ idiosyncratic components are detected, we face the situation of *cross-member cointegration* and consequently the nonstationarity in the panel is entirely due to a reduced number of common stochastic trends. Cointegration between $Y_{i,t}$ and $X_{i,t}$ can occur only if the common factors for $Y_{i,t}$ cointegrate with those of $X_{i,t}$. The null of no-cointegration between these estimated factors can be tested using a Johansen type of likelihood ratio test for example.

b. If $I(1)$ common factors and $I(1)$ idiosyncratic components are detected, we carry out step (2a) on the estimated common factors and we will work with *defactored* series. In contrast to Banerjee and Carrion-i-Silvestre (2006), however, who defactor the residuals from a static regression (11) we defactor separately $Y_{i,t}$ and $X_{i,t}$. The defactored $Y_{i,t}$ (e.g. the estimated idiosyncratic components) is simply obtained as

$$\hat{E}^Y_{i,t} = \sum_{s=1}^{t} \hat{e}^Y_{i,s} = \sum_{s=1}^{t} (\Delta Y_{i,s} - \hat{\lambda}'_{1,i} \hat{f}_s)$$

where $\hat{f}_s$ is a consistent factor estimate of f_t in equation (2) and $\hat{\lambda}'_{1,i}$ a consistent estimate of the loading. Testing for no-cointegration between the defactored data can be conducted using standard panel tests for no-cointegration such as those of Pedroni (1999, 2004a) given in equations (16) and (17).

The rejection of no-cointegration between $Y_{i,t}$ and $X_{i,t}$ occurs only if the tests for both common factor and idiosyncratic no-cointegration reject. However, this is a necessary condition. If the three restrictions mentioned under (iii) hold as well for the cointegrating vectors, panel cointegration will hold. If the outcome of step (2b) is that both the common factors and the idiosyncratic components cointegrate one might want to jointly or sequentially test the restrictions on the cointegrating vectors. The required tests are not available with the exception of a homogeneity test on the idiosyncratic component cointegrating vectors proposed by Pedroni (2004b). Comparing point estimates of the parameters involved could yield further insight into the structure of the model. Formal testing of panel no-cointegration could be done using the Banerjee and Carrion-i-Silvestre (2006) test.

Remark 4. The sequential panel no-cointegration test outlined in steps 1 and 2 is a multiple comparison procedure. Panel no-cointegration is rejected if both the hypotheses of common factor no-cointegration and idiosyncratic component no-cointegration are rejected and the restrictions between the cointegrating vector parameters are not rejected. An approximate test of the joint hypothesis could use the Bonferroni procedure (see, e.g. Savin, 1980). In a Monte Carlo simulation, the joint hypothesis test of factor and idiosyncratic component (no-)cointegration is found to be undersized because of the idiosyncratic component (no-)cointegration. Its power properties are shown to be fine. The results are available upon request.

Remark 5. The theoretical justification for this sequential procedure is analogous to that of the PANIC panel unit-root analysis. As the DGP implies that all series have a Bai and Ng (2004) representation, we proceed by analogy with the results derived in Bai and Ng (2004). Provided the number of common factors is known or consistently selected using one of the consistent selection procedures discussed in Bai and Ng (2004), then it holds that

$$T^{-1/2}\sum_{s=2}^{t}\hat{e}_{i,s}^{Y} = T^{-1/2}\sum_{s=2}^{t}e_{i,s}^{Y} + O_p(C_{NT}^{-1})$$

where $\hat{e}_{i,t}^{Y}$ is the estimated idiosyncratic component, $\hat{e}_{i,t}^{Y} = \Delta Y_{i,s} - \hat{\lambda}_{1,i}'\hat{f}_s$, $\hat{f}_s$ a consistent factor estimate of f_t, $\hat{\lambda}_{1,i}'$ a consistent estimate of the loading and $C_{NT}^{-1} = \min(N^{1/2}, T^{1/2})$. Consequently,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor \mathrm{Tr} \rfloor}\hat{e}_{i,t}^{Y} \Rightarrow B_i^Y(r), \forall i$$

where $B_i^Y(r)$ is the first element of the $(1 + m)$-vector Brownian motion $B_i(r)$. $B_i^Y(r)$ and $B_j^Y(r)$ are uncorrelated Brownian motions for $i \neq j$. The same holds for $X_{i,t}$. Consequently, standard panel no-cointegration tests derived under the maintained assumption of independent panel units, such as those proposed by Pedroni (2004a), can be used on the defactored observations.

Remark 6. This approach requires both large N and T which is one of the important limitations. Moreover, this approach will have finite sample properties that can, *at best*, be close to those observed for the tests when applied to a panel data set with independent units.

Remark 7. If the rank of the long-run covariance matrix of the factors turns out to be smaller than k, that is if the factors cointegrate, then a further step is needed to assess overall lack of cointegration between $Y_{i,t}$ and $X_{i,t}$. No cointegration then requires separability in cointegration as discussed and analysed in detail in Hecq *et al.* (2002).

V. Some Monte Carlo evidence

The theoretical foundation of the approach proposed in the preceding section requires both large N and T which is not always met in typical applications of panel cointegration techniques. A Monte Carlo analysis of some of its finite sample properties is called for. We focus on the empirical size properties of the proposed approach, namely testing for no-cointegration using defactored data, as it was shown that tests designed for cross-sectionally independent data may suffer from dramatic size distortions when applied to panels with cross-member cointegration for example as pointed out by Banerjee *et al.* (2004). The DGP is a simple bivariate process (i.e. $m = 1$) with $k = 2$ common factors that obeys the representation (1)–(4).

$$
\begin{aligned}
Z_{i,t} &= \Lambda_i F_t + E_{i,t}, \quad E_{i,t} = e_{i,t} \;\text{ or }\; E_{i,t} = E_{i,t-1} + e_{i,t},\\
e_{i,t} &= \varepsilon_{i,t} + \Gamma_i \varepsilon_{i,t-1},\\
F_t &= F_{t-1} + f_t, \quad f_t = \eta_t + \Phi_1 \eta_{t-1},
\end{aligned}
$$

where $\varepsilon_{i,t} \sim$ i.i.d. $N(0, \Sigma_i)$, $\eta_t \sim$ i.i.d. $N(0, I_2)$. The loading matrix has a diagonal structure with diagonal elements $\lambda_{1i}, \lambda_{2i} \sim U[-1, 3]$, where U denotes uniform distributions. The remaining parameters are also drawn from independent uniform distributions to allow for some degree of heterogeneity: $\Phi_{11,22} \sim U[0.5, 0.7]$, $\Phi_{12,21} \sim U[0, 0.5]$, $\Sigma_{i,11,22} \sim U[1, 1.4]$, $\Gamma_{11,22} \sim U[0.5, 0.7]$, $\Phi_{12,21} \sim U[0, 0.5]$ and $\Sigma_{i,12,21} \sim U[0, 0.2]$. The sample size has been set to $T \in \{50, 100, 250\}$ and the number of units in the panel is set to $N \in \{25, 50, 100\}$. We consider the rejection frequencies based on 1,000 replications[8] for Kao's pooled normalized coefficient (the ρ test) and pooled ADF test, and Pedroni's panel-t, panel-ρ, group-mean t and group-mean ρ statistics based on raw data. Furthermore, we consider Pedroni's panel ρ and Pedroni's group-mean ρ statistics applied to the defactored data, and Johansen trace test for the estimated common factors, using the information criterion of Aznar and Salvador (2002) to select the lag length of the Vector Error Correction Model (VECM).

For the last two statistics based on the defactored data, we estimate the number of common factors k using the IC_1 criterion of Bai and Ng (2002) with $k_{\max} = 4$. For the Augmented Dickey–Fuller (ADF)-type tests the lag length is selected using the Akaike's Information Criterion (AIC). For the non-parametric correction for serial correlation, we use a quadratic spectral kernel with a bandwidth of $3.21T^{\frac{1}{3}}$ (see Andrews, 1991).

The two polar cases that we consider in the simulations are the cases discussed earlier, namely the case of cross-member cointegration in which the common factors are $I(1)$ and the idiosyncratic components are $I(0)$, and the case where both common factors and idiosyncratic components are $I(1)$. In addition,

[8] All experiments are carried out using GAUSS 6.0.

TABLE 1

$k = 2$ common factors; non-stationary common factors F_t with I(0) or I(1) idiosyncratic component $E_{i,t}$

$E_{i,t}$	N:	25	50	100	25	50	100	25	50	100
Raw data										
	T	*Kao – ρ**			*Kao – ADF*			*Pedroni – Panel-ρ*		
I(0)	50	0.27	0.32	0.35	0.48	0.50	0.54	0.68	0.90	0.88
	100	0.39	0.47	0.49	0.54	0.62	0.62	0.84	0.96	0.95
	250	0.52	0.54	0.55	0.64	0.67	0.69	0.93	1.00	0.96
I(1)	50	0.17	0.17	0.23	0.59	0.65	0.69	0.00	0.00	0.00
	100	0.23	0.28	0.36	0.63	0.74	0.75	0.02	0.02	0.03
	250	0.34	0.39	0.45	0.74	0.81	0.80	0.10	0.08	0.14
	T	*Pedroni – Panel-t*			*Pedroni – Group-ρ*			*Pedroni – Group-t*		
I(0)	50	0.76	0.92	0.91	0.33	0.67	0.62	0.52	0.79	0.77
	100	0.83	0.96	0.94	0.67	0.94	0.89	0.67	0.89	0.85
	250	0.92	0.99	0.95	0.88	1.00	0.94	0.78	0.95	0.88
I(1)	50	0.03	0.02	0.04	0.00	0.00	0.00	0.03	0.02	0.04
	100	0.06	0.04	0.08	0.00	0.00	0.01	0.04	0.03	0.06
	250	0.13	0.10	0.18	0.02	0.01	0.04	0.07	0.05	0.10
Estimated components										
	T	*Idiosyncratic–Panel-ρ*			*Idiosyncratic – Group-ρ*			*Aznar/Johansen*		
I(0)	50	1.00	1.00	1.00	1.00	1.00	1.00	0.12	0.12	0.11
	100	1.00	1.00	1.00	1.00	1.00	1.00	0.09	0.11	0.09
	250	1.00	1.00	1.00	1.00	1.00	1.00	0.08	0.10	0.08
I(1)	50	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.14	0.12
	100	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.12	0.09
	250	0.02	0.01	0.00	0.00	0.00	0.00	0.11	0.10	0.09

Note: Rejection frequencies are based on 5% asymptotic critical values.

we consider cases where only the common factors are cointegrated but the idiosyncratic components are not cointegrated, non-cointegrated common factors are combined with cointegrated idiosyncratic components, and cointegration in both the common factors and the idiosyncratic components.

Tables 1 and 2 present simulation results for the five cases with MA(1) dynamics in the error terms and $k = 2$ common factors, one common factor in $Y_{i,t}$ and one in $X_{i,t}$. Furthermore, the number of common factors is estimated using the IC_1 criterion of Bai and Ng (2002) with $k_{max} = 4$. Note that the criterion always picks the correct number of common factors. Both Kao test statistics show strong size distortions when either the common factors or the idiosyncratic components (or both) cointegrate. The Pedroni tests exhibit very strong size distortions in the cross-member cointegration case (Table 1). When non-stationary idiosyncratic components are combined with non-cointegrated or cointegrated common factors (Tables 1 and 2) size distortions are reduced.

TABLE 2

$k = 2$ common factors; cointegration in either F_t or $E_{i,t}$ or both. NC denotes no cointegration, C cointegration

F_t	$E_{i,t}$	*N:*	*25*	*50*	*100*	*25*	*50*	*100*	*25*	*50*	*100*
Raw data											
		T	*Kao-ρ**			*Kao-ADF*			*Pedroni Panel-ρ*		
C	NC	50	0.18	0.17	0.24	0.55	0.60	0.65	0.01	0.00	0.01
		100	0.25	0.29	0.39	0.59	0.69	0.73	0.07	0.03	0.11
		250	0.37	0.40	0.50	0.70	0.77	0.78	0.18	0.10	0.32
NC	C	50	0.26	0.28	0.33	0.57	0.60	0.64	0.05	0.13	0.10
		100	0.36	0.43	0.46	0.62	0.72	0.72	0.16	0.39	0.32
		250	0.46	0.52	0.55	0.71	0.77	0.78	0.31	0.53	0.48
C	C	50	0.41	0.43	0.47	0.57	0.58	0.61	0.58	0.63	0.69
		100	0.50	0.56	0.59	0.60	0.67	0.69	0.84	0.91	0.97
		250	0.60	0.63	0.63	0.69	0.72	0.74	0.96	0.99	1.00
		T	*Pedroni Panel-t*			*Pedroni Group-ρ*			*Pedroni Group-t*		
C	NC	50	0.05	0.03	0.06	0.00	0.00	0.00	0.03	0.02	0.05
		100	0.09	0.06	0.19	0.05	0.01	0.10	0.07	0.03	0.15
		250	0.17	0.11	0.33	0.25	0.13	0.57	0.19	0.10	0.42
NC	C	50	0.14	0.23	0.23	0.01	0.03	0.03	0.10	0.17	0.18
		100	0.20	0.42	0.38	0.06	0.32	0.18	0.14	0.35	0.29
		250	0.29	0.52	0.50	0.21	0.68	0.40	0.24	0.55	0.44
C	C	50	0.69	0.77	0.84	0.74	0.81	0.91	0.86	0.94	0.98
		100	0.87	0.94	0.98	1.00	1.00	1.00	1.00	1.00	1.00
		250	0.95	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Estimated components											
		T	*Idiosyncratic Panel-ρ*			*Idiosyncratic Group-ρ*			*Aznar/Johansen*		
C	NC	50	0.00	0.00	0.00	0.00	0.00	0.00	0.61	0.61	0.74
		100	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.78	0.85
		250	0.02	0.00	0.00	0.00	0.00	0.00	0.68	0.85	0.87
NC	C	50	0.92	0.99	1.00	0.94	1.00	1.00	0.12	0.12	0.11
		100	1.00	1.00	1.00	1.00	1.00	1.00	0.10	0.11	0.10
		250	1.00	1.00	1.00	1.00	1.00	1.00	0.09	0.10	0.08
C	C	50	0.98	1.00	1.00	0.99	1.00	1.00	0.67	0.65	0.75
		100	1.00	1.00	1.00	1.00	1.00	1.00	0.79	0.83	0.89
		250	1.00	1.00	1.00	1.00	1.00	1.00	0.78	0.87	0.92

Notes: If F_t is cointegrated, $F_t^Y = \sum_{s=1}^{t} f_s^Y$, $F_t^X = F_t^Y + f_t^X$. If $E_{i,t}$ is cointegrated, $E_{i,t}^Y = \sum_{s=1}^{T} e_{i,s}^Y$, $E_{i,t}^X = E_{i,t}^Y + e_{i,s}^X$, where f_t and $e_{i,t}$ are MA processes generated as described in section V.

Rejection frequencies are based on 5% asymptotic critical values.

The tests are even undersized for some combinations of N and T. When both the common factors and the idiosyncratic components cointegrate (Table 2), the Pedroni tests have rejection frequencies of up to 1. However, as the factor loadings are heterogeneous, panel cointegration is not present (see the discussion in section IV).

The tests applied to the estimated idiosyncratic components show rejection frequencies of (close to) 1 when they are stationary or cointegrated. When the idiosyncratic components are not cointegrated, the idiosyncratic panel-ρ and idiosyncratic group-ρ tests are undersized. The Aznar/Johansen test applied to the estimated common factors is slightly oversized when the common factors do not cointegrate with rejection frequencies between 8% and 15%. When there is cointegration among the common factors, the test has a power between 61% and 92%.

We also perform simulations where we have introduced a second factor in $X_{i,t}$, such that $k = 3$ now.[9] Again estimating the number of common factors using the IC_1 criterion of Bai and Ng (2002), we note that the second common factor of $X_{i,t}$ is not picked up.[10] Nevertheless, simulation results for the Kao and Pedroni tests applied to the raw data and the Aznar/Johansen test applied to the extracted common factors do not change qualitatively compared with the results obtained for $k = 2$. However, the idiosyncratic panel-ρ and idiosyncratic group-ρ applied to the estimated common components exhibit reduced power when the common components are cointegrated, in particular when $T = 50$.

VI. Conclusions

We have considered the problem of testing for (no-)cointegration in panel data characterized by strong cross-sectional dependencies resulting from common factors as in the study of Bai and Ng (2004). We focus on two polar cases that we believe are of empirical relevance.

For both classes of DGPs, we discuss the homogeneity restrictions for the cointegrating vectors resulting from the presence of common factor cointegration. We study analytically the behaviour of several tests for panel cointegration including Kao (1999) and Pedroni's (1999, 2004a) residual-based panel no-cointegration tests that have been widely used in empirical work in the recent years. The results complement and help to understand some of the Monte Carlo results reported by Banerjee *et al.* (2004), such as the loss of Gaussian limiting results and occurrence of size distortions resulting from the presence of cross-sectional dependence.

These observations provide sufficient reason to propose a two-step procedure for testing for no-cointegration in panels with common factors. Our procedure is similar in spirit and complementary to the work of Banerjee and Carrion-i-Silvestre (2006). It has the advantages of covering many sub-cases of interest and allowing us to get a clear picture of the common and

[9]Tables with the results for these simulations are included in the working paper version of this paper (Gengenbach *et al.*, 2005).

[10]Similarly, the PC_1 or BIC_3 criteria from Bai and Ng (2002) only select a single common factor for $X_{i,t}$.

idiosyncratic components in the panel and about the homogeneity requirements for common factor cointegration. The procedure is simple to apply and makes use of existing tools. Simulation results show the procedure to have reasonable size properties.

While being attractive due, among other things, to its ease of application and nice properties, some limitations are inherent in this approach. The theoretical validity of the proposed procedure, and that of Banerjee and Carrion-i-Silvestre (2006), relies on both large N and large T which may be unrealistic for applications with 'moderate' N and large T. The performance of the proposed procedures, in particular the power properties, in such situations needs to be further studied even if the size properties reported in the Monte Carlo section are promising. If a large N analysis is inappropriate for the problem under study, an alternative could be to adopt the nonlinear Instrumental Variables (IV) testing approach of Demetrescu and Tarcolea (2005) or use bootstrapping techniques that seem to work well from an empirical point of view (see Fachin, 2005). Future work should study the merits of these alternative approaches both theoretically and empirically.

A second limitation lies in the fact that the approach is residual-based and hence suffers from the usual critiques against residual-based tests such as the maintained assumptions of a single cointegrating relationship (if it exists) as well as the imposition of the common factor restriction. Nothing however precludes extension of the ideas developed in this paper to other cointegration techniques that would not suffer from these drawbacks.

References

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

Aznar, A. and Salvador, M. (2002). 'Selecting the rank of the cointegration space and the form of the intercept using an information criterion', *Econometric Theory*, Vol. 18, pp. 926–947.

Bai, J. and Kao, C. (2004). 'On the estimation and inference of a panel cointegration model with cross-sectional dependence', mimeo, Presented at ESEM2004.

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

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

Banerjee, A. and Carrion-i-Silvestre, J. L. (2006). 'Cointegration in panel data with breaks and cross-section dependence', ECB Working Paper No. 591, Frankfurt.

Banerjee, A., Marcellino, M. and Osbat, C. (2004). 'Some cautions on the use of panel methods for integrated series of macro-economic data', *Econometrics Journal*, Vol. 7, pp. 322–340.

Banerjee, A., Marcellino, M. and Osbat, C. (2005). 'Testing for PPP: should we use panel methods?', *Empirical Economics*, Vol. 30, pp. 77–91.

Breitung, J. and Das, S. (2005). 'Testing for unit roots in panels with a factor structure', University of Bonn.

Breitung, J. and Pesaran, H. M. (2005). 'Unit roots and cointegration in panels', in Matyas L. and Sevestre P. (eds), *The Econometrics of Panel Data*, Kluwer Academic Publishers, Dordrecht (forthcoming).

Dees, S., di Mauro, F., Pesaran, M. H. and Smith, L. V. (2005). *Exploring the International Linkages of the Euro Area: A Global VAR Analysis*, CESifo Working Paper No. 1425.

Demetrescu, M. and Tarcolea, A.-I. (2005). *Panel Cointegration Testing using Nonlinear Instruments*, Working Paper, Goethe-University Frankfurt.

Fachin, S. (2005). 'Long-run trends in internal migrations in Italy: a study in panel cointegration with dependent units', mimeo, University of Rome 'La Sapienza'.

Gengenbach, C., Palm, F. C. and Urbain, J.-P. (2005). 'Cointegration testing in panels with common factors', METEOR Research Memorandum, Universiteit Maastricht.

Gregoir, S. (2005). *Representation and Statistical Analysis of Weakly Linearly Exchangeable Dynamic Panels*, Working Paper, CREST.

Hecq, A., Palm, F. C. and Urbain, J.-P. (2002). 'Separation, weak exogeneity and P–T decomposition in cointegrated VAR systems with common features', *Econometric Reviews*, Vol. 21, pp. 273–307.

Kao, C. (1999). 'Spurious regression and residual-based tests for cointegration in panel data', *Journal of Econometrics*, Vol. 90, pp. 1–44.

Lyhagen, J. (2000). 'Why not use standard panel unit root test for testing PPP', mimeo, Stockholm School of Economics.

Moon, H. R. and Perron, B. (2004). 'Testing for a unit root in panels with dynamic factors', *Journal of Econometrics*, Vol. 122, pp. 81–126.

Pedroni, P. (1995). *Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests, with an Application to the PPP Hypothesis*, Indiana University Working Papers in Economics, No. 95-013.

Pedroni, P. (1999). 'Critical values for cointegration tests in heterogeneous panels with multiple regressors', *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 653–670.

Pedroni, P. (2004a). 'Panel cointegration, asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis', *Econometric Theory*, Vol. 20, pp. 597–625.

Pedroni, P. (2004b). 'Social capital, barriers to production and capital shares; implications for the importance of parameter heterogeneity from a nonstationary panel approach', mimeo, Williams College.

Pedroni, P. and Urbain, J.-P. (2001). 'Cross member cointegration in non-stationary panels', mimeo, Universtiteit Maastricht.

Pesaran, M. H. (2006). 'A simple panel unit root test in the presence of cross section dependence', Cambridge University DAE Working Paper No. 0346.

Phillips, P. C. B. and Durlauf, S. (1986). 'Multiple time series regression with integrated processes', *Review of Economic Studies*, Vol. 53, pp. 473–495.

Phillips, P. C. B. and Moon, H. R. (1999). 'Linear regression limit theory for nonstationary panel data', *Econometrica*, Vol. 67, pp. 1057–1111.

Savin, N. E. (1980). 'The Bonferroni and Scheffe multiple comparison procedures', *Review of Economic Studies*, Vol. 47, pp. 255–273.

Urbain, J.-P. (2004). 'Spurious regression in nonstationary panels with cross-member cointegration', mimeo, Presented at ESEM2004.

Wagner, M. and Müller-Fürstenberger, G. (2004). *The Carbon Kuznets Curve: A Cloudy Picture Emitted by Bad Econometrics?*, Discussion Paper 4.18, University of Bern.

Westerlund, J. (2005). *New Simple Tests for Panel Cointegration*, Working Paper 2005:8, Lund University, Department of Economics.

Appendix

Given Assumptions 1–4, we can summarize some convergence results. In the following lemmas, M is used to denote a generic positive number, not depending on T or N. For a matrix $\mathbf{A}$, $\mathbf{A} > 0$ denotes that $\mathbf{A}$ is positive-definite. Furthermore, $\|\mathbf{A}\| = \mathrm{trace}(\mathbf{A}'\mathbf{A})^{\frac{1}{2}}$. We write the integral $\int_0^1 B(r)\,\mathrm{d}r$ as $\int B$, and $\int_0^1 B(r)B(r)'\mathrm{d}r$ as $\int BB'$. Furthermore, '$\Rightarrow$' denotes weak convergence, and '$\xrightarrow{p}$' denotes convergence in probability. For any number x, $\lfloor x \rfloor$ denotes the largest integer smaller than x. For any variable $X_{i,t}$,

$$\tilde{X}_{i,t} = X_{i,t} - \frac{1}{T}\sum_{s=1}^{T} X_{i,s}.$$

Similarly, for any Brownian motion B, $\tilde{B} = B - \int B$. Throughout the paper, we employ sequential limit theory, where we consider $T \to \infty$ followed by $N \to \infty$. Furthermore, for non-random factor loadings,

$$\bar{\Lambda} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \Lambda_i,$$

while for random factor loadings $\bar{\Lambda} = E(\Lambda_i)$, $\Psi = E(\Psi_i)$ and $\Delta = E(\Delta_i)$.

Lemma 1 presents convergence results for the common data component $\Lambda_i F_t$. The limiting distributions are functionals of Brownian motions weighted by the factor loadings, even as $N \to \infty$. These results are intuitive, as we assume a fixed number of common factors. Lemma 2 summarizes the convergence for the idiosyncratic components, where we recover the panel spurious regression results for Phillips and Moon (1999). In Lemma 3, the limits for the cross-products of the common and individual-specific components are given. It is evident that these cross-products will only affect limiting distributions for finite N, but as $N \to \infty$ these effects will vanish because of the independence of the shock driving F_t and $E_{i,t}$.

Lemma 1: Common component. Given Assumptions 1, 2 and 4,

(a) $\frac{1}{T}\sum_{t=1}^{T} \Lambda_i f_t F'_{t-1}\Lambda'_i \Rightarrow \Lambda_i(\int \mathrm{d}B_F B'_F + \Theta)\Lambda'_i \quad \text{as } T \to \infty,$

and $\frac{1}{N}\sum_{i=1}^{N} \Lambda_i(\int \mathrm{d}B_F B'_F + \Theta)\Lambda'_i \xrightarrow{p} \int \mathrm{d}B_{F\Lambda} B'_{F\Lambda} + \Theta_{F\Lambda} \quad \text{as } N \to \infty,$

(b) $\frac{1}{T^2}\sum_{t=1}^{T} \Lambda_i F_t F'_t \Lambda'_i \Rightarrow \Lambda_i(\int B_F B'_F)\Lambda'_i \quad \text{as } T \to \infty,$

and $\frac{1}{N}\sum_{i=1}^{N} \Lambda_i(\int B_F B'_F)\Lambda'_i \xrightarrow{p} \int B_{F\Lambda} B'_{F\Lambda} \quad \text{as } N \to \infty,$

(c) $\frac{1}{T}\sum_{t=1}^{T}\Lambda_i f_t \tilde{F}'_{t-1}\Lambda'_i \Rightarrow \Lambda_i(\int \mathrm{d}B_F\tilde{B}'_F+\Theta)\Lambda'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i(\int \mathrm{d}B_F\tilde{B}'_F+\Theta)\Lambda'_i \xrightarrow{p} \int \mathrm{d}B_{F\Lambda}\tilde{B}'_{F\Lambda}+\Theta_{F\Lambda} \quad \text{as } N\to\infty,$

(d) $\frac{1}{T^2}\sum_{t=1}^{T}\Lambda_i\tilde{F}_t\tilde{F}'_t\Lambda'_i \Rightarrow \Lambda_i(\int \tilde{B}_F\tilde{B}'_F)\Lambda'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i(\int \tilde{B}_F\tilde{B}'_F)\Lambda'_i \xrightarrow{p} \int \tilde{B}_{F\Lambda}\tilde{B}'_{F\Lambda} \quad \text{as } N\to\infty,$

where

$$\begin{aligned}
\mathrm{vec}\left(\int \mathrm{d}B_{F\Lambda}B'_{F\Lambda}\right) &= \check{\mathbf{\Lambda}}\,\mathrm{vec}\left(\int \mathrm{d}B_F B'_F\right), \quad \mathrm{vec}(\Theta_{F\Lambda}) = \check{\mathbf{\Lambda}}\,\mathrm{vec}(\Theta),\\
\mathrm{vec}\left(\int B_{F\Lambda}B'_{F\Lambda}\right) &= \check{\mathbf{\Lambda}}\,\mathrm{vec}\left(\int B_F B'_F\right),\\
\mathrm{vec}\left(\int \mathrm{d}B_{F\Lambda}\tilde{B}'_{F\Lambda}\right) &= \check{\mathbf{\Lambda}}\mathrm{vec}\left(\int \mathrm{d}B_F\tilde{B}'_F\right) \quad \text{and} \quad \mathrm{vec}\left(\int \tilde{B}_{F\Lambda}\tilde{B}'_{F\Lambda}\right)\\
&= \check{\mathbf{\Lambda}}\,\mathrm{vec}\left(\int \tilde{B}_F\tilde{B}'_F\right), \text{ and} \quad \check{\mathbf{\Lambda}} = \mathrm{p}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}(\Lambda_i\otimes\Lambda_i).
\end{aligned}$$

Lemma 2: Idiosyncratic components. Given Assumption 3,

(a) $\frac{1}{T}\sum_{t=1}^{T}e_{i,t}S'_{i,t-1} \Rightarrow \left(\int \mathrm{d}B_iB'_i+\Delta_i\right) \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\left(\int \mathrm{d}B_iB'_i+\Delta_i\right) \xrightarrow{p} \Delta \quad \text{as } N\to\infty,$

(b) $\frac{1}{T^2}\sum_{t=1}^{T}S_{i,t}S'_{i,t} \Rightarrow \int B_iB'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\int B_iB'_i \xrightarrow{p} \frac{1}{2}\Psi \quad \text{as } N\to\infty,$

(c) $\frac{1}{T}\sum_{t=1}^{T}e_{i,t}\tilde{S}'_{i,t-1} \Rightarrow \left(\int \mathrm{d}B_i\tilde{B}'_i+\Delta_i\right) \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\left(\int \mathrm{d}B_i\tilde{B}'_i+\Delta_i\right) \xrightarrow{p} -\frac{1}{2}\Psi+\Delta \quad \text{as } N\to\infty,$

(d) $\frac{1}{T^2}\sum_{t=1}^{T}\tilde{S}_{i,t}\tilde{S}'_{i,t} \Rightarrow \int \tilde{B}_i\tilde{B}'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\int \tilde{B}_i\tilde{B}'_i \xrightarrow{p} \frac{1}{6}\Psi \quad \text{as } N\to\infty.$

Lemma 3. Given Assumptions 1, 2, 3 and 4

(a) $\frac{1}{T}\sum_{t=1}^{T}\Lambda_i F_{t-1}e'_{i,t} \Rightarrow \Lambda_i \int B_F \; \mathrm{d}B'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i \int B_F \; \mathrm{d}B'_i \xrightarrow{p} 0 \quad \text{as } N\to\infty,$

(b) $\frac{1}{T}\sum_{t=1}^{T}\Lambda_i f_t S'_{i,t-1} \Rightarrow \Lambda_i \int \mathrm{d}B_F B'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i \int \mathrm{d}B_F B'_i \xrightarrow{p} 0 \quad \text{as } N\to\infty,$

(c) $\frac{1}{T^2}\sum_{t=1}^{T}\Lambda_i F_t S'_{i,t} \Rightarrow \Lambda_i \int B_F B'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i \int B_F B'_i \xrightarrow{p} 0 \quad \text{as } N\to\infty,$

(d) $\frac{1}{T}\sum_{t=1}^{T}\Lambda_i \tilde{F}_{t-1}e'_{i,t} \Rightarrow \Lambda_i \int \tilde{B}_F \; \mathrm{d}B'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i \int \tilde{B}_F \; \mathrm{d}B'_i \xrightarrow{p} 0 \quad \text{as } N\to\infty,$

(e) $\frac{1}{T}\sum_{t=1}^{T}\Lambda_i \tilde{F}_{t-1}\tilde{e}'_{i,t} \Rightarrow \Lambda_i \int \tilde{B}_F \; \mathrm{d}B'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i \int \tilde{B}_F \; \mathrm{d}B'_i \xrightarrow{p} 0 \quad \text{as } N\to\infty,$

(f) $\frac{1}{T}\sum_{t=1}^{T}\Lambda_i f_t \tilde{S}'_{i,t-1} \Rightarrow \Lambda_i \int \mathrm{d}B_F \tilde{B}'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i \int \mathrm{d}B_F \tilde{B}'_i \xrightarrow{p} 0 \quad \text{as } N\to\infty,$

(g) $\frac{1}{T^2}\sum_{t=1}^{T}\Lambda_i \tilde{F}_t \tilde{S}'_{i,t} \Rightarrow \Lambda_i \int \tilde{B}_F \tilde{B}'_i \quad \text{as } T\to\infty,$

and $\frac{1}{N}\sum_{i=1}^{N}\Lambda_i \int \tilde{B}_F \tilde{B}'_i \xrightarrow{p} 0 \quad \text{as } N\to\infty.$

The proofs of Lemma 1–3 are omitted here. They are included in the working paper version of this article (Gengenbach *et al.*, 2005).

Proof of Proposition 1(a): convergence of $\tilde{\beta}$

The LSDV estimator of β is given by

$$\tilde{\beta} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}'_{i,t}\right)\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}'_{i,t}\right)^{-1}.$$

Consider the numerator

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}'_{i,t}=\sum_{i=1}^{N}\sum_{t=1}^{T}(\lambda'_{1i}\tilde{F}^{Y}_{t}\tilde{F}^{X\prime}_{t}\lambda_{21}+\tilde{E}^{Y}_{i,t}\tilde{E}^{X\prime}_{i,t}+\lambda'_{1i}\tilde{F}^{Y}_{t}\tilde{E}^{X\prime}_{i,t}+\tilde{E}^{Y}_{i,t}\tilde{F}^{X\prime}_{t}\lambda_{21}). \tag{19}$$

If the idiosyncratic term is given by equation (3), we have

$$\sum_{i=1}^{N}(O_p(T^2)+O_p(T)+O_p(T)+O_p(T))$$

in equation (19). So, as $T\rightarrow\infty$,

$$\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}'_{i,t}\Rightarrow\sum_{i=1}^{N}\lambda'_{1i}\int\tilde{B}^{Y}_{F}\tilde{B}^{X\prime}_{F}\lambda_{2i}$$

from the first result of Lemma 1(d). Now, using the second result we obtain

$$\frac{1}{N}\sum_{i=1}^{N}\lambda'_{1i}\int\tilde{B}^{Y}_{F}\tilde{B}^{X\prime}_{F}\lambda_{2i}\xrightarrow{p}\int\tilde{B}^{Y}_{F\Lambda}\tilde{B}^{X}_{F\Lambda}{}'\quad\text{as } N\rightarrow\infty,$$

where $\int\tilde{B}^{Y}_{F\Lambda}\tilde{B}^{X}_{F\Lambda}{}'$ is the $1\times m$ upper right block of $\int\tilde{B}_{F\Lambda}\tilde{B}'_{F\Lambda}$ defined in Lemma 1.

If the idiosyncratic terms are also $I(1)$, such that the DGP includes equation (4), all terms in equation (19) are $O_p(T^2)$ when summed over T. However, the cross-products of the common factors and idiosyncratic components will vanish in the limit as $N\rightarrow\infty$. Using Lemmas 1(d), 2(d) and 3(g) we find as $T\rightarrow\infty$ followed by $N\rightarrow\infty$,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}'_{i,t}\Rightarrow\int\tilde{B}^{Y}_{F\Lambda}\tilde{B}^{X}_{F\Lambda}{}'+\frac{1}{6}\Psi^{YX},$$

where Ψ^{YX} is the upper right $1\times m$ block of Ψ.

Now the denominator of $\tilde{\beta}$ is given by

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}'_{i,t}=\sum_{i=1}^{N}\sum_{t=1}^{T}(\lambda'_{2i}\tilde{F}^{X}_{t}\tilde{F}^{X\prime}_{t}\lambda_{21}+\tilde{E}^{X}_{i,t}\tilde{E}^{X\prime}_{i,t}+\lambda'_{2i}\tilde{F}^{X}_{t}\tilde{E}^{X\prime}_{i,t}+\tilde{E}^{X}_{i,t}\tilde{F}^{X\prime}_{t}\lambda_{21}). \tag{20}$$

Similar to the results for the numerator, the terms in equation (20) are

$$\sum_{i=1}^{N}\left(O_p(T^2)+O_p(T)+O_p(T)+O_p(T)\right),$$

if the DGP contains equation (3). Hence,

$$\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}_{i,t}' \Rightarrow \sum_{i=1}^{N}\lambda_{2i}'\int \tilde{B}_F^X\tilde{B}_F^{X\prime}\lambda_{2i} \quad \text{as } T\to\infty.$$

Furthermore, the remaining term is $O_p(N)$, and we obtain

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}_{i,t}' \Rightarrow \int \tilde{B}_{F\Lambda}^X\tilde{B}_{F\Lambda}^{X}{}' \quad \text{as } T\to\infty$$

followed by $N\to\infty$, where $\int \tilde{B}_{F\Lambda}^X\tilde{B}_{F\Lambda}^{X}{}'$ is the lower right $m\times m$ block of $\int \tilde{B}_{F\Lambda}\tilde{B}_{F\Lambda}'$.

If the true DGP contains equation (4), all terms in the summation over T in equation (20) are $O_p(T^2)$ and as above, the cross-products between common and idiosyncratic components will vanish in the cross-sectional average as $N\to\infty$. We find as $T\to\infty$ followed by $N\to\infty$,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}_{i,t}' \Rightarrow \int \tilde{B}_{F\Lambda}^X\tilde{B}_{F\Lambda}^{X}{}' + \frac{1}{6}\Psi^{XX},$$

where Ψ^{XX} is the lower right $m\times m$ block of Ψ.

Combining the results given above yields Proposition 1A(a) and B(a). ■

Proof of Proposition 1(b): convergence of $\tilde{\rho}$

The residuals from the first-stage PLS regression are given by $\tilde{u}_{i,t} = (1,-\tilde{\beta})Z_{i,t} = Y_{i,t} - \tilde{\beta}X_{i,t}$. For the pooled regression given in equation (13) we have

$$(\tilde{\rho}-1) = \left(\sum_{i=1}^{N}\sum_{t=2}^{T}(1\ -\tilde{\beta})\Delta Z_{i,t}\tilde{Z}_{i,t-1}'(1\ -\tilde{\beta})'\right) \times \left(\sum_{i=1}^{N}\sum_{t=2}^{T}(1\ -\tilde{\beta})\tilde{Z}_{i,t-1}\tilde{Z}_{i,t-1}'(1\ -\tilde{\beta})'\right)^{-1}. \tag{21}$$

For the numerator consider

$$\sum_{i=1}^{N}\sum_{t=2}^{T}\Delta Z_{i,t}\tilde{Z}_{i,t-1}' = \sum_{i=1}^{N}\sum_{t=2}^{T}(\Lambda_i f_t\tilde{F}_{t-1}'\Lambda_i' + \Delta E_{i,t}\tilde{E}_{i,t-1}' + \Lambda_i f_t\tilde{E}_{i,t-1}' + \Delta E_{i,t}\tilde{F}_{t-1}'\Lambda_i'). \tag{22}$$

From Lemma 1(c),

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\Lambda_i f_t\tilde{F}_{t-1}\Lambda_i' \Rightarrow \int \mathrm{d}B_{F\Lambda}\tilde{B}_{F\Lambda}' + \Theta_{F\Lambda} \quad \text{as } T\to\infty$$

followed by $N \to \infty$. If the idiosyncratic terms are $I(0)$, i.e. the true DGP is given by equation (3),

$$\sum_{i=1}^{N}\sum_{t=2}^{T}\Delta E_{i,t}\tilde{E}'_{i,t-1} = \sum_{i=1}^{N}\sum_{t=2}^{T}\left((e_{i,t}-e_{i,t-1})e'_{i,t-1}-(e_{i,t}-e_{i,t-1})\bar{e}'_i\right),$$

where

$$\bar{e}_i = \frac{1}{T}\sum_{t=1}^{T}e_{i,t}.$$

Now,

$$\frac{1}{T}\sum_{t=2}^{T}e_{i,t}e'_{i,t-1}\xrightarrow{p}\gamma_{i1}\quad \text{as } T\to\infty,$$

with

$$\gamma_{i1} = \lim_{T\to\infty}\frac{1}{T}\sum_{i=1}^{T}E(e_{i,t}e'_{i,t-1}),\quad \text{and}\quad \frac{1}{N}\sum_{i=1}^{N}\gamma_{i1}\xrightarrow{p}\gamma_1 \text{ as } N\to\infty,$$

with $\gamma_1 \equiv E(\gamma_{i1})$. Moreover,

$$\frac{1}{T}\sum_{t=2}^{T}e_{i,t-1}e'_{i,t-1}\xrightarrow{p}\Upsilon_i\quad \text{as } T\to\infty$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\Upsilon_i\xrightarrow{p}\Upsilon\quad \text{as } N\to\infty.$$

Furthermore,

$$\frac{1}{T}\sum_{t=2}^{T}e_{i,t}\bar{e}'_i\xrightarrow{p}0\quad \text{and}\quad \frac{1}{T}\sum_{t=2}^{T}e_{i,t-1}\bar{e}'_i\xrightarrow{p}0\quad \text{as } T\to\infty.$$

Hence,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\Delta E_{i,t}\tilde{E}'_{i,t-1}\xrightarrow{p}\gamma_1-\Upsilon\quad \text{as } T\to\infty$$

followed by $N \to \infty$.

For the third term in equation (22) we have as $T \to \infty$,

$$\frac{1}{T}\sum_{t=2}^{T}\Lambda_i f_t\tilde{E}'_{i,t-1} = \frac{1}{T}\sum_{t=2}^{T}\Lambda_i f_t e'_{i,t-1} - \frac{1}{T}\sum_{t=2}^{T}\Lambda_i f_t\bar{e}'_i\xrightarrow{p}0.$$

Finally, as $T \to \infty$,

$$\frac{1}{T}\sum_{t=2}^{T} \Delta E_{i,t}\tilde{F}'_{t-1}\Lambda'_i = \frac{1}{T}e_{i,T}\tilde{F}'_{T-1}\Lambda' - \frac{1}{T}e_{i,1}\tilde{F}'_1\Lambda' - \frac{1}{T}\sum_{t=2}^{T} e_{i,t-1}f'_{t-1}\Lambda'_i \xrightarrow{p} 0.$$

Hence, as $T \to \infty$ followed by $N \to \infty$,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T} \Delta Z_{i,t}\tilde{Z}'_{i,t-1} \Rightarrow \int \mathrm{d}B_{F\Lambda}\tilde{B}'_{F\Lambda} + \Theta_{F\Lambda} + \gamma_1 - \Upsilon.$$

If the idiosyncratic components are $I(1)$ and their true DGP includes equation (4), such that $\Delta E_{i,t} = e_{i,t}$ and $\tilde{E}_{i,t-1} = \tilde{S}_{i,t-1}$, using Lemmas 1(c), 2(c) and 3(d) and 3(f), we obtain as $T \to \infty$ followed by $N \to \infty$,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T} \Delta Z_{i,t}\tilde{Z}'_{i,t-1} \Rightarrow \int \mathrm{d}B_{F\Lambda}\tilde{B}_{F\Lambda} + \Theta_{F\Lambda} - \frac{1}{2}\Psi + \Delta.$$

For the denominator in equation (21) consider

$$\sum_{i=1}^{N}\sum_{t=2}^{T} \tilde{Z}_{i,t-1}\tilde{Z}'_{i,t-1} = \sum_{i=1}^{N}\sum_{t=2}^{T} \big(\Lambda_i\tilde{F}_{t-1}\tilde{F}'_{t-1}\Lambda'_i + \tilde{E}_{i,t-1}\tilde{E}'_{i,t-1} + \Lambda_i\tilde{F}_{t-1}\tilde{E}'_{i,t-1} + \tilde{E}_{i,t-1}\tilde{F}'_{t-1}\Lambda'_i\big). \tag{23}$$

If the idiosyncratic components are given by equation (3), we find,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=2}^{T} \tilde{Z}_{i,t-1}\tilde{Z}'_{i,t-1} \Rightarrow \int \tilde{B}_{F\Lambda}\tilde{B}'_{F\Lambda} \quad \text{as } T \to \infty$$

followed by $N \to \infty$.

For $I(1)$ idiosyncratic components given by equation (4), we find using Lemmas 1(d), 2(d) and 3(g)

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=2}^{T} \tilde{Z}_{i,t-1}\tilde{Z}'_{i,t-1} \Rightarrow \int \tilde{B}_{F\Lambda}\tilde{B}'_{F\Lambda} + \frac{1}{6}\Psi$$

as $T \to \infty$ followed by $N \to \infty$. Combining the above given results with those of A(a) or B(a) yields Proposition 1A(b) and B(b). ■

Proof of Proposition 1(c): divergence of $t_{\tilde{\rho}}$

The t-statistic for $\tilde{\rho} = 1$ is given by

$$t_{\tilde{\rho}} = (\tilde{\rho} - 1)s^{-1}\left(\sum_{i=1}^{N}\sum_{t=2}^{T} \tilde{u}^2_{i,t-1}\right)^{\frac{1}{2}},$$

where

$$s^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\left(\Delta\tilde{u}_{i,t}^2 + O_p(1)\right).$$

As

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}\Delta\tilde{u}_{i,t}^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=2}^{T}(1 \ -\tilde{\beta})\Delta Z_{i,t}\Delta\tilde{Z}_{i,t}'(1 \ -\tilde{\beta})',$$

which is $O_p(1)$ whether the idiosyncratic components are $I(0)$ or $I(1)$, s^2 is $O_p(1)$. Furthermore, $T(\tilde{\rho}-1)$ and $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=2}^{T}\tilde{u}_{i,t-1}\tilde{u}_{i,t-1}'$ are $O_p(1)$ as well whether $E_{i,t}$ is given by equations (3) or (4), as shown above. Hence,

$$t_{\tilde{\rho}} = \sqrt{N}T(\tilde{\rho}-1)s^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=2}^{T}\tilde{u}_{i,t-1}\tilde{u}_{i,t-1}'\right)^{\frac{1}{2}} = \sqrt{N}O_p(1),$$

which diverges at rate $\sqrt{N}$ as $T\to\infty$ followed by $N\to\infty$. ■

Proof of Proposition 2(a): convergence of $\tilde{\boldsymbol{\beta}}_i$

For each panel unit i, the estimator of β_i is given by

$$\tilde{\beta}_i = \left(\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}_{i,t}'\right)\left(\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}_{i,t}'\right)^{-1}.$$

Consider the numerator

$$\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}_{i,t}' = \sum_{t=1}^{T}(\lambda_{1i}'\tilde{F}_t^Y\tilde{F}_t^{X\prime}\lambda_{21} + \tilde{E}_{i,t}^Y\tilde{E}_{i,t}^{X\prime} + \lambda_{1i}'\tilde{F}_t^Y\tilde{E}_{i,t}^{X\prime} + \tilde{E}_{i,t}^Y\tilde{F}_t^{X\prime}\lambda_{21}). \quad (24)$$

If the idiosyncratic term is given by equation (3), we have $O_p(T^2)$ + $O_p(T)$ + $O_p(T)$ + $O_p(T)$ in equation (24). So, as $T\to\infty$,

$$\frac{1}{T^2}\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}_{i,t}' \Rightarrow \lambda_{1i}'\int\tilde{B}_F^Y\tilde{B}_F^{X\prime}\lambda_{2i}$$

from the first result of Lemma 1(d).

If the idiosyncratic terms are also $I(1)$, such that the DGP includes equation (4), all terms in equation (24) are $O_p(T^2)$ when summed over T. Using Lemmas 1(d), 2(d) and 3(g), we find as $T\to\infty$,

$$\frac{1}{T^2}\sum_{t=1}^{T}\tilde{Y}_{i,t}\tilde{X}_{i,t}' \Rightarrow \left(\lambda_{1i}'\int\tilde{B}_F^Y\tilde{B}_F^{X\prime}\lambda_{2i} + \int\tilde{B}_i^Y\tilde{B}_i^{X\prime} + \lambda_{1i}'\int\tilde{B}_F^Y\tilde{B}_i^{X\prime} + \int\tilde{B}_i^Y\tilde{B}_F^{X\prime}\lambda_{21}\right).$$

Now the denominator of $\tilde{\beta}_i$ is given by

$$\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}'_{i,t}=\sum_{t=1}^{T}(\lambda'_{2i}\tilde{F}_t^X\tilde{F}_t^{X\prime}\lambda_{21}+\tilde{E}_{i,t}^X\tilde{E}_{i,t}^{X\prime}+\lambda'_{2i}\tilde{F}_t^X\tilde{E}_{i,t}^{X\prime}+\tilde{E}_{i,t}^X\tilde{F}_t^{X\prime}\lambda_{21}). \quad (25)$$

Similar to the results for the numerator, the terms in equation (25) are $O_p(T^2)$ + $O_p(T)$ + $O_p(T)$ + $O_p(T)$, if the DGP contains equation (3). Hence,

$$\frac{1}{T^2}\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}'_{i,t}\Rightarrow\lambda'_{2i}\int\tilde{B}_F^X\tilde{B}_F^{X\prime}\lambda_{2i}\quad\text{as } T\to\infty.$$

If the true DGP contains equation (4), all terms in equation (25) are $O_p(T^2)$ and we have, as $T\to\infty$,

$$\frac{1}{T^2}\sum_{t=1}^{T}\tilde{X}_{i,t}\tilde{X}'_{i,t}\Rightarrow\left(\lambda'_{2i}\int\tilde{B}_F^X\tilde{B}_F^{X\prime}\lambda_{2i}+\int\tilde{B}_i^X\tilde{B}_i^{X\prime}+\lambda'_{2i}\int\tilde{B}_F^X\tilde{B}_i^{X\prime}+\int\tilde{B}_i^X\tilde{B}_F^{X\prime}\lambda_{21}\right).$$

Combining the results given above yields Proposition 2A(a) and B(a). ■

Proposition 2(b): convergence of $Z_{\tilde{\rho}_{NT}-1}$ and $\tilde{Z}_{\tilde{\rho}_{NT}-1}$

The residuals from the individual first-stage regression are given by $\tilde{u}_{i,t}=(1,-\tilde{\beta}_i)Z_{i,t}=Y_{i,t}-\tilde{\beta}_iX_{i,t}$. Consider first

$$\sum_{t=2}^{T}\Delta\tilde{u}_{i,t}\tilde{u}_{i,t-1}=\sum_{t=2}^{T}(1,\ -\tilde{\beta}_i)\Delta Z_{i,t}\tilde{Z}'_{i,t-1}(1,\ -\tilde{\beta}_i)'. \quad (26)$$

Now,

$$\begin{aligned}\sum_{t=2}^{T}\Delta Z_{i,t}\tilde{Z}'_{i,t-1}&=\sum_{t=2}^{T}(\Lambda_i f_t+\Delta E_{i,t})(\Lambda_i\tilde{F}_{t-1}+\tilde{E}_{i,t-1})'\\&=\sum_{t=2}^{T}(\Lambda_i f_t\tilde{F}'_{t-1}\Lambda'_i+\Delta E_{i,t}\tilde{E}'_{i,t-1}+\Lambda_i f_t\tilde{E}'_{i,t-1}+\Delta E_{i,t}\tilde{F}'_{t-1}\Lambda'_i).\end{aligned} \quad (27)$$

From Lemma 1(c),

$$\frac{1}{T}\sum_{t=2}^{T}\Lambda_i f_t\tilde{F}_{t-1}\Lambda'_i\Rightarrow\int\Lambda_i(\mathrm{d}B_F\tilde{B}_F+\Theta)\Lambda'_i\quad\text{as } T\to\infty.$$

If the idiosyncratic terms are $I(0)$, i.e. the true DGP is given by equation (3),

$$\sum_{t=2}^{T}\Delta E_{i,t}\tilde{E}'_{i,t-1}=\sum_{t=2}^{T}\left((e_{i,t}-e_{i,t-1})e'_{i,t-1}-(e_{i,t}-e_{i,t-1})\bar{e}_i\right),$$

where

$$\bar{e}_i = \frac{1}{T}\sum_{t=1}^{T} e_{i,t}.$$

Now,

$$\frac{1}{T}\sum_{t=2}^{T} e_{i,t}e'_{i,t-1} \xrightarrow{p} \gamma_{i1} \quad \text{as } T \to \infty,$$

with

$$\gamma_{i1} = \lim_{T\to\infty}\frac{1}{T}\sum_{i=1}^{T} E(e_{i,t}e_{i,t-1}).$$

Moreover,

$$\frac{1}{T}\sum_{t=2}^{T} e_{i,t-1}e'_{i,t-1} \xrightarrow{p} \Upsilon_i \quad \text{as } T \to \infty.$$

Furthermore,

$$\frac{1}{T}\sum_{t=2}^{T} e_{i,t}\bar{e}'_i \xrightarrow{p} 0 \quad \text{and} \quad \frac{1}{T}\sum_{t=2}^{T} e_{i,t-1}\bar{e}'_i \xrightarrow{p} 0 \quad \text{as } T \to \infty.$$

Hence,

$$\frac{1}{T}\sum_{t=2}^{T} \Delta E_{i,t}\tilde{E}'_{i,t-1} \xrightarrow{p} \gamma_{i1} - \Upsilon_i \quad \text{as } T \to \infty.$$

For the third term in equation (27) we have, as $T \to \infty$,

$$\frac{1}{T}\sum_{t=2}^{T} \Lambda_i f_t \tilde{E}'_{i,t-1} = \frac{1}{T}\sum_{t=2}^{T} \Lambda_i f_t e'_{i,t-1} - \frac{1}{T}\sum_{t=2}^{T} \Lambda_i f_t \bar{e}'_i \xrightarrow{p} 0.$$

Finally, as $T \to \infty$,

$$\frac{1}{T}\sum_{t=2}^{T} \Delta E_{i,t}\tilde{F}'_{t-1}\Lambda'_i = \frac{1}{T}e_{i,T}\tilde{F}'_{T-1}\Lambda' - \frac{1}{T}e_{i,1}\tilde{F}'_1\Lambda' - \frac{1}{T}\sum_{t=2}^{T} e_{i,t-1}f'_{t-1}\Lambda'_i \xrightarrow{p} 0.$$

Hence, as $T \to \infty$,

$$\frac{1}{T}\sum_{t=2}^{T} \Delta Z_{i,t}\tilde{Z}'_{i,t-1} \Rightarrow \Lambda_i\left(\int \mathrm{d}B_F\tilde{B}_F + \Theta\right)\Lambda'_i + \gamma_{i1} - \Upsilon_i.$$

If the idiosyncratic components are $I(1)$ and their true DGP includes equation (4), such that $\Delta E_{i,t} = e_{i,t}$ and $\tilde{E}_{i,t-1} = \tilde{S}_{i,t-1}$, using Lemmas 1(c), 2(c) and 3(d) and 3(f), we obtain, as $T \to \infty$,

$$\frac{1}{T}\sum_{t=2}^{T} \Delta Z_{i,t}\tilde{Z}'_{i,t-1} \Rightarrow \left(\Lambda'_i\left(\int \mathrm{d}B_F\tilde{B}'_F + \Theta\right)\Lambda'_i + \int \mathrm{d}B_i\tilde{B}'_i + \Delta_i + \Lambda_i \int \mathrm{d}B_F\tilde{B}'_i + \int \mathrm{d}B_i\tilde{B}'_F\Lambda'_i\right).$$

Furthermore, note that the residuals $\tilde{v}_{i,t} = \Delta\tilde{u}_{i,t} + o_p(1)$ regardless of whether they were obtained from the poooled regression equation (13) or the individual regression equation (15). Now,

$$\begin{aligned}\hat{\boldsymbol{\lambda}}_i &= T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}\tilde{v}_{i,t}\tilde{v}_{i,t-s}\\ &= T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}\Delta\tilde{u}_{i,t}\Delta\tilde{u}_{i,t-s} + o_p(1)\\ &= T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}(1,\ -\tilde{\beta}_i)\Delta\tilde{Z}_{i,t}\Delta\tilde{Z}'_{i,t-s}(1,\ -\tilde{\beta}_i)' + o_p(1).\end{aligned}$$

Expanding $\Delta\tilde{Z}_{i,t}\Delta\tilde{Z}'_{i,t-s}$ in terms of the common factors and unobserved components we obtain the following four terms and convergence results for suitable choices of bandwidth J and kernel function ω_{sJ}. First,

$$T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}\Lambda_i\tilde{f}_{i,t}\tilde{f}'_{i,t-s}\Lambda'_i \xrightarrow{p} \Lambda_i\Omega\Lambda'_i. \tag{28}$$

Next,

$$T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}\Lambda_i\tilde{f}_{i,t}\Delta\tilde{E}'_{i,t-s} \xrightarrow{p} 0, \tag{29}$$

and

$$T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}\Delta\tilde{E}_{i,t}\tilde{f}'_{i,t-s}\Lambda_i \xrightarrow{p} 0, \tag{30}$$

because of the independence of common factors and idiosyncratic components. Finally,

$$T^{-1}\sum_{s=1}^{J}\omega_{sJ}\sum_{t=s+1}^{T}\Delta\tilde{E}_{i,t}\Delta\tilde{E}'_{i,t-s} \xrightarrow{p} \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E(e_{i,t}\tilde{E}_{i,t}), \tag{31}$$

which is $\gamma_{1i} - \Upsilon_i$ if the idiosyncratic components are stationary, and Δ_i if they are $I(1)$.

Now consider

$$\sum_{t=2}^{T} \Delta\tilde{u}_{i,t}\tilde{u}_{i,t-1} = \sum_{t=2}^{T}(1,\ -\tilde{\beta}_i)\Delta Z_{i,t}\tilde{Z}'_{i,t-1}(1,\ -\tilde{\beta}_i)'. \tag{32}$$

We have

$$\sum_{t=2}^{T} \tilde{Z}_{i,t-1}\tilde{Z}'_{i,t-1} = \sum_{t=2}^{T} \big(\Lambda_i\tilde{F}_{t-1}\tilde{F}'_{t-1}\Lambda'_i + \tilde{E}_{i,t-1}\tilde{E}'_{i,t-1} + \Lambda_i\tilde{F}_{t-1}\tilde{E}'_{i,t-1} + \tilde{E}_{i,t-1}\tilde{F}'_{t-1}\Lambda'_i\big). \tag{33}$$

If the idiosyncratic components are given by equation (3), when summed over T the first term in equation (33) is $O_p(T^2)$, while the remaining three are $O_p(T)$. So,

$$\frac{1}{T^2}\sum_{t=2}^{T} \tilde{Z}_{i,t-1}\tilde{Z}'_{i,t-1} \Rightarrow \Lambda_i \int \tilde{B}_F\tilde{B}'_F\Lambda'_i \quad \text{as } T \to \infty.$$

For $I(1)$ idiosyncratic components given by equation (4), we find using Lemmas 1(d), 2(d) and 3(g), as $T \to \infty$,

$$\frac{1}{T^2}\sum_{t=2}^{T} \tilde{Z}_{i,t-1}\tilde{Z}'_{i,t-1} \Rightarrow \left(\Lambda_i \int \tilde{B}_F\tilde{B}'_F\Lambda'_i + \int \tilde{B}_i\tilde{B}'_i + \Lambda_i\int \tilde{B}_F\tilde{B}'_i + \int \tilde{B}_i\tilde{B}'_F\Lambda'_i\right).$$

We use the block-triangular decomposition of the long-run covariance matrix of the common non-stationary factors Ω, such that $\Omega = L'L$ with $L_{11} = \Omega_{11} - \Omega'_{21}\Omega_{22}^{-1}\Omega_{21})^{\frac{1}{2}}$, $L_{21} = \Omega_{22}^{-\frac{1}{2}}\Omega_{21}$, and $L_{22} = \Omega_{22}^{\frac{1}{2}}$, where blocks are conformable to the partition of $B_F = (B_F^{Y\prime}, B_F^{X\prime})'$. Note that $\Omega_{22} > 0$ by Assumption 1.

Now, $\tilde{B}_F = L'\tilde{W}_F$, where $\tilde{W}_F$ is a demeaned k-vector standard Brownian motion. Furthermore, denote

$$\boldsymbol{\eta}'_i = (1, -\tilde{\mathbf{b}}_{iA}) \quad \text{and} \quad \boldsymbol{\kappa}' = \left(\mathbf{I}_{k_Y}, -\left(\int \tilde{W}_F^Y\tilde{W}_F^{X\prime}\right)\left(\int \tilde{W}_F^X\tilde{W}_F^{X\prime}\right)^{-1}\right).$$

Then,

$$L\Lambda'_i\boldsymbol{\eta}_i = \boldsymbol{\kappa}L_{11}\lambda_{1i} \quad \text{and} \quad \boldsymbol{\eta}'_i\tilde{B}_F = \lambda'_{1i}L'_{11}\tilde{Q}_F,$$

with

$$\tilde{Q}_F = \tilde{W}_F^Y - \left(\int \tilde{W}_F^Y\tilde{W}_F^{X\prime}\right)\left(\int \tilde{W}_F^X\tilde{W}_F^{X\prime}\right)^{-1}\tilde{W}_F^X.$$

Finally,

$$\boldsymbol{\eta}_i' \int \mathrm{d}B_F \tilde{B}_F' \boldsymbol{\eta}_i = \lambda_{1i}' L_{11}' \int \mathrm{d}Q_F \tilde{Q}_F' L_{11} \lambda_{1i},$$

and

$$\boldsymbol{\eta}_i' \int \tilde{B}_F \tilde{B}_F' \boldsymbol{\eta}_i = \lambda_{1i}' L_{11}' \int \tilde{Q}_F \tilde{Q}_F' L_{11} \lambda_{1i}.$$

Combining the above given results with those of A(a) or B(a) yields the convergence results for $Z_{\tilde{\rho}_{NT}-1}$ and $\tilde{Z}_{\tilde{\rho}_{NT}-1}$. ■