Testing for Panel Cointegration using Common Correlated Effects Estimators*

Anindya Banerjee
Department of Economics
University of Birmingham
Edghaston, Birmingham, B15 2TT, UK
E-mail: a.banerjee@bham.ac.uk

Josep Lluís Carrion-i-Silvestre
AQR-IREA Research group
Department of Econometrics, Statistics and Spanish Economy
University of Barcelona
Av. Diagona, 690
08034, Barcelona
E-mail: carrion@ub.edu

January 17, 2017

*Carrion-i-Silvestre acknowledges the financial support from Spanish Ministerio de Ciencia y Tecnología under grant ECO2014-58991-C3-1-R. We thank Rob Taylor and two referees, George Kapetanios, Hashem Pesaran and Jean-Pierre Urbain, and the participants at the 17th Panel Data Conference in Montreal, the Conference in Honor of H. M. Pesaran in Cambridge and the XIV Encuentro de Economía Aplicada in Huelva for their helpful comments. Responsibility for any remaining errors lies with us.

Abstract

Spurious regression analysis in panel data when the time series are cross-section dependent is analyzed in the paper. We show that consistent estimation of the long-run average parameter is possible once we control for cross-section dependence using cross-section averages in the spirit of the common correlated effects approach in Pesaran (2006). This result is used to design a panel cointegration test statistic accounting for cross-section dependence. The performance of the proposal is investigated in comparison with factor-based methods to control for cross-section dependence when strong, semi-weak and weak cross-section dependence may be present.

Keywords: panel cointegration, cross-section dependence, common factors, spatial econometrics

JEL codes: C12, C22

1 Introduction

During the last twenty years the analysis of macroeconomic panels has experienced a vast and rapid development. This has been primarily due to two reasons: first, the easy availability of statistical information concerning panels of data where the time dimension is augmented by the use of cross-section variation (for example, across countries, regions or industrial sectors), and, second, the belief that combining these two sources of information would lead to better statistical inference.

The recent literature has seen many efforts, in particular to design procedures aimed at estimating long-run relationships among economic variables using macro-panel data techniques. Testing for cointegration in panel data has been a particular area of focus, since it constitutes the analysis that needs to be conducted prior to estimating long-run relationships. The early papers in this area such as the ones in Kao (1999) and Pedroni (2000) assumed cross-section independence among the units of the panel data, a situation that is rarely found in empirical economic analyses. Cross-section dependence appears naturally when studying economic data due to, for instance, market integration processes, globalization of economic activity, offshoring processes or because of the presence of common shocks like oil price shocks. More recent papers have therefore devoted considerable attention to devising procedures relaxing the assumption of cross-section independence – see, for example, Breitung and Pesaran (2008) for an overview.

There may be different sources of cross-section dependence, exerting different degrees of dependence intensity. On the one hand, we may have pervasive cross-section dependence due to the presence of a dominant unit in the panel data setup, a situation that can be interpreted as if there were common factors affecting all the time series in the panel. On the other hand, cross-section dependence may be important only among some neighbours. The notion of 'neighbour' does not of course necessarily need to be defined in terms of physical contiguity, such as neighbouring regions or cities, but may also be defined *inter alia* in terms of economic distance, usually, trade partnerships. This characterization of cross-section dependence has given rise to the notions of weak and strong dependence as discussed prominently by Chudik, Pesaran and Tosetti (2011).

In this paper we investigate the performance of panel cointegration tests in the presence of weak and/or strong cross-section dependence in the sense to be defined below. In particular, our contribution is to develop the theory of a panel cointegration test based on the common correlated effects (CCE) estimation procedure proposed by Pesaran (2006). Holly, Pesaran and Yamagata (2010) use a CCE-based procedure in their investigation without developing a formal test for cointegration. Our proposal fills this gap. We show, drawing upon arguments developed by Phillips and Moon (1999) and Kao, Trapani and Urga (2011) that in a panel spurious regression, the pooled CCE estimator provides a consistent estimate of the long-run average coefficient. The Kao et al. (2011) paper is particularly useful for the theoretical derivations since it allows us to work with the invariant σ-field generated by the factors which drive the cross-section dependence. Once consistency is proved, this result can then be used as the basis for a panel cointegration test.

Two other papers are worth mentioning to put our work in context. First, Urbain and Westerlund (2011) look at the asymptotics of least squares regressions in spurious and cointegrated

panels with cross-section dependence driven by common and idiosyncratic stochastic trends or factors which may be integrated or stationary. Second, Gengenbach, Urbain and Westerlund (2016) study a panel cointegration test based on the error correction model approach where the cross-section dependence is modelled via the use of cross-section averages. While the techniques used are similar, neither paper deals with the estimation of a long-run average coefficient à la Phillips and Moon (1999) in the construction of a cointegration test based on this estimator.

An important feature of our test is that it can be linked directly to the Pesaran (2007) and Pesaran, Smith and Yamagata (2013) panel data unit root tests with which it is asymptotically equivalent depending upon the number of integrated common factors.

The procedures that are proposed in the paper are investigated through Monte Carlo simulations to evaluate the potential benefits of using the new proposal compared to alternative approaches existing in the literature. In this respect we compare the size and power properties of the pooled CCE-based test with the factor-based cointegration testing procedure developed in Banerjee and Carrion-i-Silvestre (2015). Factor-based approaches to model dependence are typically thought of as being an alternative to the use of CCE estimators. The advantages of the latter method include the ability to allow for a more detailed look at the nature of the dependence, notably their decomposition into I(1) and I(0) components. This allows for cointegration not only among the variables themselves but also between the variables and the non-stationary factors. The disadvantages include the potential inconsistency of the factor estimates and their loadings when the dependence is weak, and of the difficulty of estimating both the number of factors and the factors themselves especially when the cross-section dimension is relatively small. The convenience of using CCE-based tests is also an advantage that needs to be taken into account.

Having developed the theory, we explore these trade-offs through Monte Carlo simulations in order to evaluate the potential benefits of using the newly proposed test is one of the main purposes of our paper. The simulation experiments cover a wide spectrum of cross-section dependence patterns. First, we consider strong, semi-strong and weak cross-section dependence through a factor model specification. Second, weak dependence is also analyzed using the typical spatial econometrics configurations to control for the presence of cross-section dependence**.** It is not our purpose here to offer a broad evaluation of the size and power properties of a range of panel tests for cointegration. Rather this paper is intended to propose and evaluate our new test and to compare its performance with a factor-based competitor.[1]

The application of the procedures proposed in this paper is illustrated using two different empirical examples, emphasizing different forms of dependence. In both cases, we find that cross-section dependence is present among the units of the panels. First, we focus on the estimation of a model to explain the behaviour of house prices in 48 US states and the District of Columbia, complementing the analysis carried out in Holly et al. (2010). Second, we proceed to estimate a production function for a set of OECD developed countries which represents a typical macro panel. Here the likely presence of an I(1) productivity trend puts it within the framework of models with strong dependence among the units of the panel. Both these examples

[1] The only other paper we are aware of that makes this direct comparison is by Urbain and Westerlund (2015). However, since their framework is entirely stationary, they do not deal with issues relating to cointegration, focusing instead only on cross-section dependence.

serve to illustrate and interpret further the results of the simulation analysis.

The paper is organized as follows. Section 2 describes the model upon which the panel cointegration test statistic proposed in the paper is based. We derive consistency results for the pooled CCE estimator under different specifications of the deterministic terms. Next, Section 3 defines the panel cointegration test statistic using the CCE estimator. The finite sample performance under different sources of cross-section dependence is investigated in Section 4. Section 5 presents the results of two empirical applications. Finally, Section 6 concludes. All the proofs are contained in the appendix at the end of the paper.

2 The model

Let $Y_{i,t} = (y_{i,t}, x'_{i,t})'$ a $(1+k)$-vector of $I(1)$ stochastic processes with the following data generating process (DGP):

$$\begin{aligned} Y_{i,t} &= D_{i,t} + \pi_i F_t + U_{i,t} && (1) \\ (I-L) F_t &= v_t = C(L) w_t && (2) \\ (I-L) U_{i,t} &= e_{i,t} = H_i(L) \varepsilon_{i,t}, && (3) \end{aligned}$$

where $D_{i,t}$ denotes the deterministic part of the model that is given by either the absence of deterministic elements $D_{i,t} = 0 \ \forall i$ (Model 0), a vector of constant terms, $D_{i,t} = \mu_i = \left(\mu_{i,0}, \mu_{i,1}, \ldots, \mu_{i,k}\right)'$ (Model 1), or a vector linear time trends, $D_{i,t} = \delta_i (1,t)'$, with $\delta_i = \left(\delta'_{i,0}, \delta'_{i,1}, \ldots, \delta'_{i,k}\right)'$, $\delta_{i,j} = \left(\mu_{i,j}, \eta_{i,j}\right)'$, $j = 0, 1, \ldots, k$, (Model 2). The F_t component denotes a $(r \times 1)$-vector of common factors and π_i the $((k+1) \times r)$ matrix of factor loadings, and denote by K an invariant σ-field generated by F_t so that, conditionally on K, $U_{i,t} = \left(U_{y_i,t}, U'_{x_i,t}\right)'$ are independent across i. The disturbance terms v_t and $e_{i,t}$ are assumed to be I(0) stationary processes, $i = 1, \ldots, N$, $t = 1, \ldots, T$, $j = 1, \ldots, r$.

Our analysis is based on the same set of assumptions as in Bai and Ng (2004) and Banerjee and Carrion-i-Silvestre (2015). Let $M < \infty$ be a generic positive number, not depending on T and N. Further, the Euclidean norm of a generic matrix A is defined as $\|A\| = trace\,(A'A)^{1/2}$. Then:

Assumption A: (i) for non-random π_i, $\|\pi_i\| \leq M$; for random π_i, $E\|\pi_i\|^4 \leq M$, (ii) $\frac{1}{N}\sum_{i=1}^{N} \pi'_i \pi_i \xrightarrow{p} \Sigma_\Pi$, a $(r \times r)$ positive definite matrix.

Assumption B: (i) $v_t = C(L) w_t$, $w_t \sim iid(0, \Sigma_w)$, $E\|w_t\|^4 \leq M$, and (ii) $Var(\Delta F_t) = \sum_{j=0}^{\infty} C_j \Sigma_w C'_j > 0$, (iii) $\sum_{j=0}^{\infty} j \|C_j\| < M$; and (iv) $C(1)$ has rank r_1, $0 \leq r_1 \leq r$.

Assumption C: (i) for each i, $e_{i,t} = H_i(L)\varepsilon_{i,t}$, $\varepsilon_{i,t} \sim iid\left(0, \sigma^2_{\varepsilon,i}\right)$, $E|\varepsilon_{i,t}|^8 \leq M$, $\sum_{j=0}^{\infty} j |H_{i,j}| < M$, $\omega_i^2 = H_i(1)^2 \sigma^2_{\varepsilon,i} > 0$; (ii) $E(\varepsilon_{i,t}\varepsilon_{j,t}) = \tau_{i,j}$ with $\sum_{i=1}^{N} |\tau_{i,j}| \leq M$ for all j;
(iii) $E\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[\varepsilon_{i,s}\varepsilon_{i,t} - E(\varepsilon_{i,s}\varepsilon_{i,t})\right]\right|^4 \leq M$, for every (t,s).

Assumption D: The errors $\varepsilon_{i,t}$, w_t and the loadings π_i are three mutually independent groups across i, t and $(1+k)$ dimensions.

Assumption E: $E\|F_0\| \leq M$, and for every $i = 1, \ldots, N$, $E\|U_{i,0}\| \leq M$.

The model specification considers the case where the stochastic regressors $x_{i,t}$ are assumed

to be either cross-section independent – imposing all, but the first, rows of π_i to be equal to zero – or cross-section dependent with dependence driven by a set of observable common factors F_t. Furthermore, it is possible to assume that the set of common factors affecting the endogenous variable $y_{i,t}$ is different from those affecting the stochastic regressors $x_{i,t}$, a situation that is covered if we define π_i to be a block-diagonal matrix.

Despite the presence of the operator $(I-L)$ in equation (2), F_t does not have to be I(1). In fact, F_t can be I(0), I(1), or a combination of both, depending on the rank of $C(1)$. If $C(1)=0$, then F_t is I(0). If $C(1)$ is of full rank, then each component of F_t is I(1). If $C(1)\neq 0$, but not full rank, then some components of F_t are I(1) and some are I(0).

Note that although this framework is very flexible, it implies a change in the standard definition of cointegration. The usual definition of cointegration among $Y_{i,t}=(y_{i,t},x'_{i,t})'$ requires F_t to be I(0), so that the observable variables capture all the common stochastic trends. However, allowing F_t to be I(1) is also relevant from an empirical point of view since F_t might be accounting for effects that are not captured by $Y_{i,t}$ alone. In such a situation, cointegration among the elements in $Y_{i,t}$ up to the inclusion of I(1) factors is both possible and desirable economically, which will imply that $H_i(1)\neq 0$ but is not full rank.[2] This leads to a broader concept of cointegration, where the observable variables in $Y_{i,t}$ alone do not generate a cointegrating relationship. Instead, common factors are required to enter in the model in order to define a long-run relationship. In this regard, the framework is close to Bai, Kao and Ng (2009) where cointegration is assumed allowing for the possibility of I(1) factors.

Panel spurious regression has been tackled in Phillips and Moon (1999). Contrary to what is found at the unit level analysis – see Granger and Newbold (1974) and Phillips (1986) – pooled estimation of the parameters affecting the stochastic regressors leads to consistent estimates of the so-called long-run average coefficient β. For ease of exposition, let us assume that there is no deterministic component (Model 0) and specify the model that relates the dependent variable $y_{i,t}$ and the explanatory variables $x_{i,t}$ in matrix notation as:

$$y_i = x_i\beta_i + F\lambda_i + \xi_i, \tag{4}$$

where y_i denotes the $(T\times 1)$ vector of the dependent variable, x_i is a $(T\times k)$ matrix of explanatory variables, β_i is a $(k\times 1)$ vector of parameters, F is a $(T\times r)$ matrix of common factors, λ_i is a $(r\times 1)$ vector of factor loadings, and ξ_i is a $(T\times 1)$ vector collecting the idiosyncratic errors. Note that we can write (4) in terms of the elements in (1) and obtain:

$$\begin{aligned}
y_i &= x_i\beta_i + F\lambda_i + \xi_i \\
(F\pi_i^y + U_{y_i}) &= (F\pi_i^x + U_{x_i})\,\beta_i + F\lambda_i + \xi_i \\
U_{y_i} &= U_{x_i}\beta_i + F\left(\lambda_i - \pi_i^y + \pi_i^x\beta_i\right) + \xi_i \\
U_{y_i} &= U_{x_i}\beta_i + F\lambda_i^* + \xi_i,
\end{aligned}$$

[2] Note that in this case, the common factors will be accounting for misspecification errors in the model, due, for instance, to the omission of relevant non-stationary stochastic regressors. Further, the presence of I(1) common factors will capture global stochastic trends that are difficult to proxy using observable data – see the discussion on the production function empirical example below. Finally, I(1) common factors will also help to capture the existence of cross-cointegration relationship among the same variable for the different units of the panel data – see Banerjee, Marcellino and Osbat (2005).

so that β_i can be estimated as the vector of parameters that capture the relationship among the idiosyncratic component of the variables.

Let us define the projection matrix $M_D = I_T - D\left(D'D\right)^{-1}D'$ that removes the effect of the deterministic component on the variables of the model, where $D = \iota$ for Model 1 and $D = [\iota\ \tau]$ for Model 2 – it should be understood that $M_D = I_T$ for Model 0. Then we have:

$$\begin{aligned} y_i &= D_i + x_i\beta_i + F\lambda_i + \xi_i \\ M_D y_i &= M_D x_i\beta_i + M_D F\lambda_i + M_D\xi_i \\ \tilde{y}_i &= \tilde{x}_i\beta_i + \tilde{F}\lambda_i + \tilde{\xi}_i, \end{aligned}$$

and

$$\begin{aligned} M_{\tilde{F}}\tilde{y}_i &= M_{\tilde{F}}\tilde{x}_i\beta_i + M_{\tilde{F}}\tilde{\xi}_i \\ y_i^* &= x_i^*\beta_i + \xi_i^*, \end{aligned} \tag{5}$$

with $M_{\tilde{F}} = I_T - \tilde{F}\left(\tilde{F}'\tilde{F}\right)^{-1}\tilde{F}'$, where the superscript "$*$" indicates that the corresponding variable has been detrended and defactored. Note that at this stage, we have assumed that the common factors are observable.[3] The pooled estimator is defined as:

$$\begin{aligned} \hat{\beta} &= \left[\sum_{i=1}^{N}\left(x_i^{*\prime}x_i^*\right)\right]^{-1}\sum_{i=1}^{N}\left(x_i^{*\prime}y_i^*\right) \\ &= \left[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}x_{i,t}^*x_{i,t}^{*\prime}\right]^{-1}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}x_{i,t}^*y_i^*. \end{aligned} \tag{6}$$

Theorem 1 *Let $Y_{i,t}$ be a vector of $(1+k)$ stochastic processes with DGP given by (1)-(3). Under the assumption that $H_i(1)$ is positive definite almost surely for all i (spurious regression), the pooled estimator given in (6) converges as $(T,N)\to\infty$ jointly to*

$$\hat{\beta} \xrightarrow{p} \beta = \Omega_{U_xU_x}^{-1}\Omega_{U_xU_y},$$

where β denotes the long-run average regression coefficient, and $\Omega_{U_xU_x}$ and $\Omega_{U_xU_y}$ are long-run average covariance matrices of the respective idiosyncratic components defined in the companion appendix.

The proof is provided in the companion appendix. It is also possible to derive the limiting distribution of the estimated long-run average parameter as in Phillips and Moon (1999). However, for the purposes of testing for cointegration pursued in this paper, we are only required to show consistency of the estimator as stated in Theorem 1.

[3] Although the projection against the deterministic component and the factors can be done in one step, proceeding in two stages facilitates the derivations below.

3 Panel CCE cointegration test with cross-section dependence

So far, the cross-section dependence has been assumed to be driven by a set of observable common factors. However, in most cases this situation is infeasible from an empirical point of view, and we need to devise procedures to estimate (or proxy for) the unobserved common factors.

There are two popular approaches in the literature that address this issue. First, the Bai and Ng (2002, 2004) proposal, which uses principal components to estimate the common factors and panel information criteria to chose the number of common factors. Second, we can use the cross-section average method suggested in Pesaran (2006, 2007) and Pesaran et al. (2013), which employs cross-section averages as convenient proxies to capture the common factors without requiring the estimation of their number. This paper looks in detail at this second approach and also establishes a comparison with testing procedures based on Bai and Ng (2002, 2004). Note, however, that the derivations obtained in the previous section follows if we also estimate the common factors and loadings using the approach described in Bai and Ng (2004).

To see how the CCE procedure works, we define the average of (1) as:

$$\bar{Y}_t = \bar{D}_t + \bar{\pi} F_t + \bar{U}_t,$$

where

$$\begin{aligned} \bar{Y}_t &= \frac{1}{N}\sum_{i=1}^{N} Y_{i,t}; \quad \bar{D}_t = \bar{\delta} d_t; \quad \bar{\delta} = \frac{1}{N}\sum_{i=1}^{N} \delta_i \\ \bar{\pi} &= \frac{1}{N}\sum_{i=1}^{N} \pi_i; \quad \bar{U}_t = \frac{1}{N}\sum_{i=1}^{N} U_{i,t}, \end{aligned}$$

with $d_t = 0$ for Model 0, $d_t = 1$ and $\delta_i = \left(\mu_{i,0}, \mu_{i,1}, \ldots, \mu_{i,k}\right)'$ for Model 1, and $d_t = (1, t)'$ and $\delta_i = \left(\delta'_{i,0}, \delta'_{i,1}, \ldots, \delta'_{i,k}\right)'$, $\delta_{i,j} = \left(\mu_{i,j}, \eta_{i,j}\right)'$, $j = 0, 1, \ldots, k$, for Model 2.

Assumption F: Let us assume that $rank\left(\bar{\pi}\right) = r \leq (1 + k)$ for all N as $N \to \infty$.

If the rank condition established in Assumption F is met, we have

$$F_t = \left(\bar{\pi}'\bar{\pi}\right)^{-1} \bar{\pi}' \left(\bar{Y}_t - \bar{D}_t - \bar{U}_t\right).$$

Provided that $\bar{U}_t = N^{-1}\sum_{i=1}^{N} U_{i,t} \xrightarrow{p} 0$ as $N \to \infty$ for all t, and , $\bar{\pi} \xrightarrow{p} E\left(\pi_i\right) = \pi$ as $N \to \infty$, we have that

$$F_t - \left(\bar{\pi}'\bar{\pi}\right)^{-1} \bar{\pi}' \left(\bar{Y}_t - \bar{D}_t\right) \xrightarrow{p} 0 \text{ as } N \to \infty,$$

which indicates that, for sufficiently large N, the observable averages $\bar{h}_t = \left(\bar{D}_t, \bar{Y}'_t\right)'$ can be used to proxy the unobserved factors.

Following Holly et al. (2010), let us specify the cross-section augmented regression:

$$y_{i,t} = D_{i,t} + x'_{i,t}\beta + \bar{z}'_t\lambda_i + \upsilon_{i,t}, \tag{7}$$

where $\bar{z}_t = \left(\bar{y}_t, \bar{x}'_t\right)'$ collects the cross-section averages of the dependent and the stochastic

regressors of the model. In order to estimate the β parameters in (7), Holly et al. (2010) use the pooled CCE estimator (PCCE) in Pesaran (2006), which is given by:

$$\hat{\beta}_{PCCE} = \left(\sum_{i=1}^{N} x_i' \bar{M} x_i \right)^{-1} \left(\sum_{i=1}^{N} x_i' \bar{M} y_i \right), \tag{8}$$

where $x_i = [x_{i,1}\ x_{i,2}\ \ldots\ x_{i,k}]$, $x_{i,j} = (x_{i,j,1}, x_{i,j,2}, \ldots, x_{i,j,T})'$, denotes the $(T \times k)$ matrix of regressors of interest, y_i is the $(T \times 1)$ vector of the dependent variable for the i-th unit, and $\bar{M} = I - \bar{H}\left(\bar{H}'\bar{H}\right)^{-1}\bar{H}'$, $\bar{H} = [\bar{z}]$ for Model 0, $\bar{H} = [\iota\ \bar{z}]$ for Model 1 and $\bar{H} = [\iota\ \tau\ \bar{z}]$ for Model 2, with $\iota = (1, 1, \ldots, 1)'$ a vector of ones, $\tau = (1, 2, \ldots, T)'$ a linear time trend and $\bar{z} = [\bar{x}\ \bar{y}]$ the $(T \times (k+1))$ matrix of cross-section averages.

One interesting feature is that the PCCE estimator is easy to compute and does not require the estimation of the factors driving the cross-section dependence. The main drawback is that consistency has been proved by Kapetanios et al. (2011) only under the maintained hypothesis that cointegration exists, an hypothesis that needs to be tested. Therefore, in order to assess the validity of the testing procedure that we apply, we need to show whether the PCCE estimator is consistent under the null hypothesis of no cointegration. This result is provided in the following Theorem.

Theorem 2 *Let $Y_{i,t}$ be a vector of $(1+k)$ stochastic processes with DGP given by (1)-(3). Under the assumption that $H_i(1)$ is positive definite almost surely for all i (spurious regression) and rank $(\bar{\pi}) = r \leq (1+k)$ for all N as $T, N \to \infty$, the pooled estimator given in (8) converges as $(T, N) \to \infty$ jointly to*

$$\hat{\beta}_{PCCE} \xrightarrow{p} \beta = \Omega_{U_x U_x}^{-1} \Omega_{U_x U_y}.$$

The proof is provided in the companion appendix.

3.1 A test for cointegration based on the PCCE estimator

Using Theorem 2 and following on from the contributions of Pesaran (2007), Holly et al. (2010) and Pesaran et al. (2013), in this section we propose a panel cointegration test statistic that is based on the PCCE estimator. It is worth mentioning that the use of the Pesaran approach requires us to constrain the DGP that has been used so far in the sense to be described below. Thus, Pesaran (2007) and Pesaran et al. (2013) specify the following DGP:

$$y_{i,t} = D_{i,t} + x_{i,t}'\beta_i + \zeta_{i,t} \tag{9}$$

$$\zeta_{i,t} = \theta_i \zeta_{i,t-1} + f_t'\lambda_i + \epsilon_{i,t}, \tag{10}$$

with f_t a $(r \times 1)$-vector of I(0) common factors and $\epsilon_{i,t}$ an I(0) idiosyncratic disturbance term.

As can be seen, (10) can be written as $\zeta_{i,t} = 1/(1-\theta_i L)(f_t'\lambda_i + \epsilon_{i,t})$ so that the null hypothesis of spurious regression, $\theta_i = 1\ \forall i$, implies that $\zeta_{i,t} = \sum_{j=1}^{t}\left(f_j'\lambda_i + \epsilon_{i,j}\right) = F_{1,t}'\lambda_i + \xi_{i,t}$, with $F_{1,t} \sim I(1)$ and $\xi_{i,t} \sim I(1)$. Under the alternative hypothesis, we have $|\theta_i| < 1$ for $i = 1, \ldots N_1$; $\theta_i = 1$ for $i = N_1 + 1, \ldots N$, with $N_1/N \to \delta$, $0 < \delta \leq 1$ as $N \to \infty$, so that

cointegration exists for N_1 units. Further, note that under the alternative hypothesis, for the N_1 units for which $\zeta_{i,t} = \sum_{j=0}^{\infty} \theta_i^j \left(f'_{t-j} \lambda_i + \epsilon_{i,t-j} \right) = F'_{0,t} \lambda_i + \xi_{i,t}$, $F_{0,t} \sim I(0)$ and $\xi_{i,t} \sim I(0)$, but for the remaining $N - N_1$ units, $F_{1,t} \sim I(1)$ and $\xi_{i,t} \sim I(1)$.

The Pesaran approach requires us to assume the same order of integration for all common factors and the idiosyncratic component for each specific unit of the panel, namely I(1) under the null hypothesis of spurious regression for all units, and I(0) for N_1 units and I(1) for the remaining $N - N_1$ units under the alternative hypothesis. This is a substantial restriction of the general framework used above for estimation, where no assumption on the order of integration properties of the common factors was needed in relation to that of the idiosyncratic components. It is worth noticing that the same stochastic processes f_t that generates the common factors, can play different roles under the alternative hypothesis for the different units of the panel data – i.e., under the alternative hypothesis f_t can generate either $F_{0,t}$ or $F_{1,t}$, depending on the specific unit of the panel.

The PCCE-based cointegration test begins by estimating the long-run average coefficient using the PCCE method. Given the consistency of the PCCE estimator under the null hypothesis of spurious regression, in the second stage we use the PCCE estimated parameters to define the variable:

$$\tilde{y}_{i,t} = y_{i,t} - x'_{i,t} \hat{\beta}_{PCCE}, \tag{11}$$

for which the following model is estimated using ordinary least squares (OLS) estimation method:

$$\tilde{y}_{i,t} = D_{i,t} + v_{i,t},$$

and the OLS residuals are then computed as $\hat{v}_{i,t} = \tilde{y}_{i,t} - \hat{D}_{i,t}$. The null hypothesis of no cointegration is tested analyzing the order of integration of $\hat{v}_{i,t}$ through the application of the cross-section augmented Dickey-Fuller cointegration (CADF) statistic:

$$CADF_P = N^{-1} \sum_{i=1}^{N} t_{\hat{\alpha}_{i,0}},$$

where $t_{\hat{\alpha}_{i,0}}$ denotes the pseudo t-ratio of the estimated $\alpha_{i,0}$ parameter in the regression:

$$\Delta \hat{v}_{i,t} = \alpha_{i,0} \hat{v}_{i,t-1} + \sum_{j=1}^{p} \alpha_{i,j} \Delta \hat{v}_{i,t-j} + \varphi_i \overline{\hat{v}}_{t-1} + \sum_{j=0}^{p} \kappa_{i,j} \Delta \overline{\hat{v}}_{t-j} + \nu_{i,t}, \tag{12}$$

when there is one common factor and

$$\Delta \hat{v}_{i,t} = \alpha_{i,0} \hat{v}_{i,t-1} + \sum_{j=1}^{p} \alpha_{i,j} \Delta \hat{v}_{i,t-j} + \varphi'_i \overline{\hat{A}}_{t-1} + \sum_{j=0}^{p} \kappa'_{i,j} \Delta \overline{\hat{A}}_{t-j} + \nu_{i,t}, \tag{13}$$

with $\overline{\hat{A}}_t = \left(\overline{\hat{v}}_t, \bar{x}_{1,t}, \ldots, \bar{x}_{k,t} \right)'$ the vector of cross-section averages augmentation terms for the $1 + k$ multiple common factors case.

It is worth noticing that (12) and (13) define the two extreme cases i.e., the case where there is just one common factor and the case where the rank condition established at Assumption F is met with equality. For those intermediate cases where there are fewer common factors than

observables – i.e., $r < 1 + k$ – the vector $\overline{\hat{A}}_t$ in (13) will be defined with $\overline{\hat{v}}_t$ and $r - 1$ elements of the cross-section averages of the stochastic regressors. It is important to emphasize that in empirical applications the number of common factors (r) does not need to equal the total number of observables $(1 + k)$ of the model so that the intermediate cases are relevant from an empirical point of view. If analysts have knowledge, for example based on economic theory, about the number of common factors to include in the model, one can impose the restriction of the number of common factors and use the critical values that involves $(1 + k)$ observable variables but $r < 1 + k$ common factors. In this regard, we could follow the strategy in Pesaran et al. (2013) and compute the test statistic using all possible combinations of r cross-section averages available in the system as a way of obtaining robust conclusions.

When the number of common factors is not known, we can follow a conservative strategy and assume that the rank condition is satisfied with equality and base inference on the estimation of (13). The price that we would pay if the true number of common factors is $r < 1 + k$ but we impose $r = 1 + k$ is to have a test statistic with empirical size smaller than the nominal size accompanied by loss of power. The advantage is to allow us to remain agnostic about the number of integrated stochastic trends driving the data.[4]

In order to derive the critical values appropriate for the PCCE-based cointegration test, note that we can substitute (7) in (11) and obtain

$$\begin{aligned} \tilde{y}_{i,t} &= D_{i,t} + \bar{z}_t'\lambda_i + \xi_{i,t} - x_{i,t}'(\hat{\beta}_{PCCE} - \beta) \\ &= \tilde{y}_{i,t}^P - x_{i,t}'(\hat{\beta}_{PCCE} - \beta), \end{aligned}$$

where $\tilde{y}_{i,t}^P$ is the unit root part of the process analyzed by Pesaran (2007) – when $r = 1$ – and by Pesaran et al. (2013) when $r > 1$. It is worth noticing that testing for panel cointegration is asymptotically equivalent to testing for the panel unit root hypothesis addressed in Pesaran (2007) and Pesaran et al. (2013). Using Theorem 2 it is possible to show that as $T, N \to \infty$, $x_{i,t}'(\hat{\beta}_{PCCE} - \beta)$ has negligible effect on the unit root test of $\tilde{y}_{i,t}$ so that, as $T, N \to \infty$, the cointegration test $t_{\hat{\alpha}_{i,0}}$ is defined by $\tilde{y}_{i,t}^P$ and $\overline{\tilde{y}}_t^P$, which are the same elements that define the limiting distributions in Pesaran (2007) – see his Theorem 3.2 for the $CADF_{if}$ statistic when $r = 1$ – and Pesaran et al. (2013) – see their Theorem 2.1 when $r > 1$.[5]

Although the limiting distributions of the test statistics proposed in this paper and the ones reported in Pesaran (2007) and Pesaran et al. (2013) are equivalent, it is the case that there are slight differences for panel data sets of small T and/or N dimensions. In order to save space, we only report critical values for the pooled test ($CADF_P$), although critical values for the individual $t_{\hat{\alpha}_{i,0}}$ test statistic can be computed using a GAUSS program available upon request. Tables 1 and 2 present the critical values for the $CADF_P$ test statistic for Model 1 and Model

[4] An alternative strategy in the case of unknown number of factors, as followed by Pesaran et al. (2013), is to undertake the testing for all permissible values of r (using all combinations of r cross-section averages for each choice of r). The size properties of such a procedure are not clear nor are the likely conclusions if one accepts the null hypothesis for some values of r and rejects for others. This is a topic for further research by us based on multiple or repeated tests.

[5] The limiting distributions are obtained using sequential and joint limits assuming that $N/T \to k > 0$. Since consistency only requires that N and T tend to infinity jointly, the condition that $N/T \to k > 0$ does not pose any difficulty. The limiting distribution of $CADF_{if}$ can also be derived under sequential limits provided $N \to \infty$ before $T \to \infty$.

2, respectively, when there is one common factor ($r = 1$) – i.e., the rank condition is met with inequality – whereas Tables 3 and 4 collect the critical values for the multiple common factor case ($r > 1$) – in this case, we impose that the rank condition is met with equality.[6]

The computation of the critical values is based on Pesaran et al. (2013), generating the dependent variable as $y_{i,t} = y_{i,t-1} + \varepsilon_{1,i,t}$ and a vector of k explanatory variables $x_{i,t} = x_{i,t-1} + \varepsilon_{2,i,t}$, where $\varepsilon_{i,t} = \left(\varepsilon_{1,i,t}, \varepsilon'_{2,i,t}\right)' \sim iid\ N(0, I_{k+1})$, $i = 1, 2, \ldots, N$, $t = -50, -49, \ldots, T$, and $y_{i,-50} = x_{i,-50} = 0$. Using these independent time series we have computed the PCCE estimator and retrieved the $\hat{e}_{i,t}$ residuals that are used to estimate the regression equation in (12) and obtain the individual and $CADF_P$ statistics. The simulation uses 50,000 replications using different combinations of T and N. As can be seen, the critical values for the one common factor case are close to the ones computed in Pesaran (2007) when T is large, although they differ in finite samples – for example, for Model 2 compare Table 2 of our paper with Table 1b of Pesaran (2007) when $T = 200$. Note also that for large T and N the critical values do not depend on the number of regressors, since the consistency property of the $\hat{\beta}_{PCCE}$ estimator implies that the $CADF_P$ statistic behaves like the Pesaran (2007) panel unit root statistic, making our critical values applicable to cases where there are more than two regressors. A similar feature is found when comparing the critical values in Pesaran et al. (2013) and the ones computed in this paper for the multiple common factor case.

Pesaran (2007) also proposes a truncated version of the $CADF_{if}$ statistic in order to ensure that the statistic has finite moments. In our case, the truncation takes the following form – see Pesaran (2007) pp. 277:

$$\begin{cases} t^*_{\hat{\alpha}_{i,0}} = t_{\hat{\alpha}_{i,0}} & \text{if } -d_1 < t_{\hat{\alpha}_{i,0}} < d_2 \\ t^*_{\hat{\alpha}_{i,0}} = -d_1 & \text{if } t_{\hat{\alpha}_{i,0}} \leq -d_1 \\ t^*_{\hat{\alpha}_{i,0}} = d_2 & \text{if } t_{\hat{\alpha}_{i,0}} \geq d_2 \end{cases},$$

where $(d_1, d_2) = (6.19, 2.61)$ for Model 1 and $(d_1, d_2) = (6.42, 1.70)$ for Model 2. Note that we use the same threshold values as in Pesaran (2007) given that the limiting distributions of our test statistic and that of Pesaran are the same. The unreported computations that we have carried out show that the critical values of the truncated and untruncated versions of the test statistic coincide exactly for all values of $T > 15$ so that, in order to save space, we have not presented these critical values on the paper. Truncation can also be applied to the multiple common factor case, although Pesaran et al. (2013) do not provide the values of the upper and lower limits for the different number of common factors, although they can be easily obtained – a GAUSS program can be used to compute the threshold values for the truncated version of the statistic for any number of common factors.

As discussed briefly in the introduction above, the approach proposed in this paper for testing panel cointegration differs from a common-factor-based approach. An example of the latter is contained in Banerjee and Carrion-i-Silvestre (2015), who deal with the same model specification that is used in this paper but where the common factors and factor loadings are

[6] We do not report the critical values of all possible combinations where the rank condition is satisfied with inequality, although a GAUSS program is available from the authors to compute the critical values for any desired combination.

estimated using principal components. In addition to accounting for cross-section dependence in two different ways, the testing procedures also differ in one other crucial aspect, namely the computation of the estimate of β_i in the individual units of the panel. In order for Theorem 2 to apply, β needs to be a pooled estimator, i.e., the potential heterogeneity of the β_i's across the units of the panel is not taken into account in computing the CCE-based test. This is in contrast to the common-factor-based test which, since it is computed after first differencing the data, allows for heterogeneity in the β_i parameter and the test statistic for the idiosyncratic component (which is most directly comparable to the CCE-based test) is based on a mean-group test constructed by averaging across the unit-specific standardized t-statistics. In principle this therefore adds to the flexibility of the common-factor based approach, although such flexibility is unnecessary if either strict homogeneity holds or the β_i coefficients are generated by means of a random effects-type specification. However to counter this flexibility there is also the disadvantage of the need to estimate more parameters in order to construct the corresponding test statistic.

The Monte Carlo simulations reported below specify homogeneous β_i in order to present the most favorable scenario from the point of view of the use of pooled estimators while disadvantaging the factor-based tests.

4 Finite sample performance

4.1 Common factor model: weak and strong dependence

This section looks at the performance of the CCE-based tests for cointegration in comparison with the factor-based approach under several different specifications of cross-section dependence, both strong and weak. It should be noted that under some specifications of weak or semi-strong dependence to be noted below, the factor approaches are no longer optimal and do not provide consistent estimates of the factors or their loadings as typically Assumption A(ii) is violated in such circumstances. It is nevertheless of importance to compare the results of the two approaches, since at an empirical level it is often not clear what form the cross-section dependence takes in the data. It is therefore interesting and important to note within the context of the simulation exercises the better performance of the factor-based tests despite worries about the consistency of the procedures when dependence is only weak. Many of the features of the DGP used below are influenced by the empirical examples, which help us to interpret better the results arising from the estimation of the models.

4.1.1 Strong dependence

Let us first consider the DGP defined by:

$$y_{i,t} = x_{i,t} + \lambda_i' F_t + \xi_{i,t} \tag{14}$$

$$\Delta x_{i,t} = \upsilon_{i,t} \tag{15}$$

$$F_{j,t} = \rho F_{j,t-1} + w_{j,t} \tag{16}$$

$$\xi_{i,t} = \phi_i \xi_{i,t-1} + \varepsilon_{i,t}, \tag{17}$$

where $r = \{1, 2\}$, $\lambda_{i,j} \sim N(1, 1)$, $v_{i,t} \sim N(0, 1)$, $w_{j,t} \sim N\left(0, \sigma_F^2\right)$, $j = 1, 2$, and $\varepsilon_{i,t} \sim N(0, 1)$ are four mutually independent groups. Under the null hypothesis of no cointegration we specify $\phi_i = 1\ \forall i$, whereas under the alternative hypothesis of cointegration we have $|\phi_i| < 1$ for some i. Note that the definition of cointegration that we are testing for only focuses on the idiosyncratic component, regardless of the order of integration of the common factor. Thus, if $F_t \sim I(1)$ cointegration exists among $(y_{i,t}, x_{i,t}, F_t)$ but not between $(y_{i,t}, x_{i,t})$. It is worth noticing that the definition of the loadings implies that $\sum_{i=1}^{N} |\lambda_i| = O_p(N)$, so that we are facing the case of strong dependence.

The simulations focus on Model 2 using the following setup. The empirical size is analyzed using $\phi_i = 1$, whereas the empirical power is investigated using $\phi_i = \{0.99, 0.95, 0.9\}$.[7] As for the common factor component, we consider one and two common factors with autoregressive parameter given by $\rho = \{1, 0.99, 0.95\}$ with different importance, which is modelled through the following values for the variance $\sigma_F^2 = \{0.5, 1, 10\}$. The time dimension is set at $T = \{50, 100, 250\}$ and the cross-section dimension is $N = \{10, 20, 50\}$. The nominal size is set at 5% and the critical values tabulated in the previous section are used. The simulations are performed using GAUSS with 1,000 replications. In order to save space, we only report the results for $\sigma_F^2 = 1$ where the number of common factors is estimated. The results for the remaining two values of σ_F^2 are qualitatively very similar and the full set of tables can be found in Banerjee and Carrion-i-Silvestre (2011).

The simulations conducted in this subsection distinguish among three different situations. First, we cover the case where there is one common factor, and use the critical values that are computed for the true number of common factors – in this case the rank condition is met with inequality, i.e., $\hat{r} = 1$. Second, we consider the case of two common factors using the critical values that are computed for the true number of common factors – in this case the rank condition is met with equality, i.e., $\hat{r} = 2$. Finally, we focus on the one common factor case but where we assume that there are two common factors – the rank condition is satisfied with inequality but we use the critical values that are appropriate when it is satisfied with equality. This case is discussed in order to mimic the scenario of conservative inference.

One common factor and $\hat{r} = 1$ Before presenting the results for the empirical size and power of the panel cointegration test statistic that is proposed in this paper, we have conducted a small Monte Carlo simulation to show that the consistency property obtained in Theorem 2 gives a proper approximation in finite samples. Table A.1 in the supplementary material reports the results of the mean, median and root mean square error of the $\hat{\beta}_{PCCE}$ estimator with $N = \{10, 20, 50, 100\}$ and $T = \{50, 100, 250\}$ for the one common factor case. As can be seen, the mean and the median are close to the true value of the parameter – i.e., $\beta = 1$ in (14) – regardless of the values of ϕ_i and ρ. This emphasizes the value of the approach since it is possible to obtain consistent estimates of β when there is no cointegration ($\phi_i = 1$) and when there is cointegration ($|\phi_i| < 1$). Moreover, because the factor is being controlled for adequately,

[7] Note that for the empirical power analysis we impose an homogeneity restriction in order to fully control the degree of temporal dependence. Although we do not expect significant changes in the picture that would be obtained, it would be possible to conduct the analysis allowing for heterogeneous parameters, but the degree of temporal dependence will be different for each unit.

whether the factor is integrated ($\rho = 1$) or stationary ($|\rho| < 1$) does not affect the root mean square errors. Under spurious regression, root mean square error decreases with N. This result is supported by the theoretical derivations shown in Theorem 2 for the limiting distribution of $\hat{\beta}_{PCCE}$. When there is cointegration, it decreases with both N and T. Finally, for a given combination of N and T, the root mean square error is larger under the spurious regression case than when there is cointegration.

Table 5 presents the empirical size and power for the $CADF_P$ panel cointegration test statistic for $N = \{10, 20, 50\}$ for the one common factor case. In each table we also report the results for the test statistics in Banerjee and Carrion-i-Silvestre (2015) – hereafter, Z_τ statistic – for which the number of common factors throughout this section is estimated using the panel BIC information criterion in Bai and Ng (2002) with a maximum of six common factors.

As can be seen, the Z_τ test has the correct size, regardless of the value of the autoregressive parameter of the common factor (ρ), except when both N and T are small. The $CADF_P$ statistic has the correct size when $\rho = 1$, although we observe that the test statistic tends to be conservative (underrejects) as ρ moves away from 1 and T gets large. Note that this can be explained by the fact that this setup violates the common factor restriction that is required by Pesaran's (2007) framework, namely that $\phi_i = \rho$ – i.e., the dynamic of the idiosyncratic component should be the same as the one driving the common factor component.

As for empirical power, we observe that the $CADF_P$ statistic does not out-perform the Z_τ statistic for any of the cases shown here. However, the empirical power of the two statistics is almost equivalent for large T which may be taken as good grounds for preferring the use of the CCE-based test when T is reasonably large. It is worth mentioning that even in those cases where the $CADF_P$ statistic becomes conservative due to the violation of the common factor restriction, it still shows good power.

So far, we have compared the panel data test statistics that are computed using the estimated idiosyncratic component. The procedure in Banerjee and Carrion-i-Silvestre (2015) also allows us to analyze the stochastic properties of the estimated common factors. The ADF statistic that is computed using the estimated common factor is reported in the columns labelled as $t_{\tilde{F}}$. As can be seen, the $t_{\tilde{F}}$ has the correct size under the null hypothesis that $\rho = 1$, with empirical power that increases, as expected, as ρ moves away from 1 and T gets large.

To sum up, for this simple scenario, the principal components-based panel cointegration test in Banerjee and Carrion-i-Silvestre (2015) shows better overall performance, with empirical size close to the nominal size and empirical power higher than those demonstrated by the CCE-based statistics. However, both approaches tend to provide the same empirical power when the time dimension is large, and the convenience of the CCE-based approach needs also to be taken into account when assessing the relative merits of these alternative testing procedures.

Finally, it could be stressed that the procedure in Banerjee and Carrion-i-Silvestre (2015) is more informative, as it allows to obtain a fuller picture of the stochastic properties of all the specified components affecting the model. As noted earlier, from an empirical point of view, assessing the stochastic properties of the common factors is particularly important since this allows us to interpret whether $(y_{i,t}, x_{i,t})$ cointegrate alone or whether we need to consider $(y_{i,t}, x_{i,t}, F_t)$ to get a cointegrating relationship.

Two common factors and $\hat{r} = 2$ The simulations conducted in this section are based upon the DGP given by (14) to (17) using two common factors, but where instead of using (15) for the generation of the stochastic regressors $x_{i,t}$, they are defined according to

$$x_{i,t} = \lambda_i^{x'} F_t + \sum_{j=1}^{t} \upsilon_{i,j}, \tag{18}$$

where $\lambda_{i,j}^{x} \sim N(1,1)$, $j = 1, 2$. Note that now we are considering that both the dependent variable and the stochastic exogenous regressors are affected by the common factors. In this case, the rank condition is satisfied with equality provided that the number of observables $(1 + k = 2)$ equals the number of common factors $(r = 2)$, which also equals the number of cross-section averages that are used in the computation of the statistics – i.e., we assume that $\hat{r} = 2$.

Table 6 reports in the columns labelled as Equality the empirical size and power for the $CADF_P$ test statistic, when $N = \{20, 50\}$. As can be seen, the test statistic has a liberal empirical size for $N = 20$, regardless of the order of integration of the common factors. However, the empirical size equals the nominal size when the number of units of the panel increases up to $N = 50$. It is worth noticing that in this case the test statistic features under-rejection problems when the common factors are I(0), a situation that violates the assumptions made in our framework. As for the empirical power, the $CADF_P$ test statistic shows decent power figures, which tends to one as T gets large, regardless of the value of N.

One common factor and $\hat{r} = 2$ From an empirical point of view, it is more interesting to analyze the effects that might have on the empirical size and power of the $CADF_P$ test statistic when practitioners use more cross-section averages than common factors present in the model. In this case, we have specified the DGP given by (14), (16), (17) and (18), but considering just one common factor $(r = 1)$. Thus, by using all cross-section averages available in the system we are covering the situation where the assumed number of common factors $(\hat{r} = 2)$ is larger than the true number of common factors $(r = 1)$. Note that now the rank condition is met with inequality $(r < 1 + k)$.

Table 6 reports in the columns labelled as Inequality the empirical size and power for the $CADF_P$ test statistic, when $N = \{20, 50\}$. As can be seen, the empirical size is close to the nominal empirical size when the common factor is I(1), regardless of the value of N. As expected, the test statistic becomes conservative as the common factor becomes I(0) – see the comments above. As for the empirical power, the $CADF_P$ test statistic has good power, which tends to one as T gets large. An interesting feature is that for a given combination of (ϕ_i, ρ), the figures for the empirical power are smaller than the ones obtained when the correct number of cross-section averages are used to capture the effects of the common factors – see the values for the empirical power offered in Table 5 compared with the ones in the columns labelled as Inequality of Table 6. This result is something to be expected, since we are including more regressors than needed in the regression equation in which the test statistic bases on, so that a fall in the power will be produced.

4.1.2 Semi-strong dependence and weak dependence

Semi-strong dependence In the previous simulation experiment we defined a common factor model where the sum of the loadings $\sum_{i=1}^{N}|\lambda_i| = O_p(N)$, a condition that is required in order to get a consistent estimate of the space generated by the common factors. However, it is interesting to analyze the behaviour of the test statistics when we consider departures from this specification, leading to so called semi-strong (or semi-weak) and weak cross-section dependence.

For example, following Chudik et al. (2011), we may specify the loadings as

$$\lambda_i = \frac{\varrho_i}{\sqrt{3\sum_{i=1}^{N}\varrho_i^2}}; \quad \varrho_i \sim N(1,1),$$

so that in this case $\sum_{i=1}^{N}|\lambda_i| = O_p\left(N^{1/2}\right)$.

Weak dependence Alternatively we may consider the case where the loadings of the common factors are such that $\sum_{i=1}^{N}|\lambda_i| = O_p(1)$, so that we face the case of weak dependence through the loadings. In this regard, we may also follow Chudik et al. (2011) and specify the loadings as

$$\lambda_i = \frac{\varrho_i}{2\sum_{i=1}^{N}\varrho_i}; \quad \varrho_i \sim N(1,1),$$

with the rest of the parameters of the DGP as defined in the previous section. It should be noted that in this case the factor structure is not identified, so that the application of principal components would not lead to consistent estimates of either the common factors or the factor loadings. The use of the test statistic in Banerjee and Carrion-i-Silvestre (2015) is thus strictly speaking not justified.

Table 7 reports the results for $N = \{10, 20, 50\}$ when the number of common factors is estimated. As can be seen, most of the features that were outlined in the previous section are still valid. However, there are some important differences that could be noted. First, the empirical size of the CCE-based statistic is close to the nominal one even for the case where the common factor is I(0), so that we do not see any under-size distortions in this case. This may be a reflection of the fact that the data generation processes here are better suited to the CCE approach. Second, except where we have semi-strong dependence with $\sigma_F^2 = 10$, the panel BIC information criterion does not detect any common factor since the conditions for consistent estimation are not satisfied. We therefore report the results only for the tests on the idiosyncratic component and show that the difficulties of applying the factor approach here notwithstanding, the Z_τ statistic remains more powerful than the CCE-based statistic even when $T = 50$, although they again perform equally well in terms of power as T gets large.

4.2 Spatial autocorrelation

Our final specification of the DGP follows Baltagi, Bresson and Pirotte (2007) and introduces weak cross-section dependence in the panel data setup using a spatial error model. The DGP

is given by

$$\begin{aligned} y_{i,t} &= x_{i,t} + \xi_{i,t} \\ \Delta x_{i,t} &= v_{i,t} \\ \xi_{i,t} &= \phi_i \xi_{i,t-1} + \varepsilon_{i,t}, \end{aligned}$$

where the error component can follow one of these three different spatial models. First, we consider the spatial autoregressive (SAR) specification:

$$\varepsilon_t = \vartheta W_N \varepsilon_t + \epsilon_t = (I_N - \vartheta W_N)^{-1} \epsilon_t,$$

with $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \ldots, \varepsilon_{N,t})'$, W_N is an $(N \times N)$ known spatial weights matrix, ϑ is the spatial autoregressive parameter and ϵ_t is an $(N \times 1)$ error vector assumed to be distributed independently across cross-section dimension with constant variance σ_ϵ^2. Second, it is possible to define a spatial moving average (SMA) specification:

$$\varepsilon_t = \epsilon_t + \vartheta W_N \epsilon_t = (I_N + \vartheta W_N) \epsilon_t,$$

where now ϑ is the spatial moving average parameter. Finally, we also use the spatial error component (SEC) specification:

$$\varepsilon_t = \epsilon_t + \vartheta W_N \psi_t,$$

where ϵ_t is an $(N \times 1)$ vector of local error components and ψ_t is an $(N \times 1)$ vector of spillover error components. The two component vectors are assumed to consist of *iid* terms with respective variances σ_ϵ^2 and σ_ψ^2, and are uncorrelated.

Of special interest is the SEC specification since we can relate the spatial model with the common factor model that has been investigated in the previous section. We can specify:

$$\varepsilon_t = \epsilon_t + \vartheta W_N \Gamma F_t,$$

where now $\psi_t = \Gamma F_t$, with $\Gamma = (\gamma_1', \gamma_2', \ldots, \gamma_N')'$ the $(N \times r)$ matrix of loadings. Further, if we set $\vartheta = 1$ and $W_N = I_N$ we get the common factor representation used above. This allows us to specify different models depending on the degree of weak correlation that we want to allow. For instance, if the spatial weight matrix is now $V_N = I_N + W_N$ with $\psi_t = \Gamma F_t$ and $\vartheta \neq 0$, the common factors will not only affect each unit, but also their neighbours.

The simulations that are reported in this section follow the setup in Baltagi et al. (2007), who use two different values for $\vartheta = \{0.4, 0.8\}$ and the spatial weight matrix W_N given by the sparse weight matrix $W(1,1)$ that defines the '1 ahead and 1 behind' matrix with the i-th row $(1 < i < N)$ of this $N \times N$ matrix having non-zero elements in positions $i+1$ and $i-1$. The W_N matrix has been normalized so that the sum of the elements of each row equals one. Other sparse weight matrices $W(j,j)$, $j = 2, 3, \ldots, 10$, were used in Baltagi et al. (2007), although they claimed that qualitatively similar results were obtained. Therefore and in order to save space, we only use the $W(1,1)$ matrix as a way to illustrate the effect of spatial dependence on the panel data cointegration tests that we consider in the paper.

The simulation experiment has been conducted for $N = \{10, 20, 50\}$ and, in general, qualitatively similar results are obtained regardless of the number of cross-section units. Consequently, in what follows our discussion focuses on the results reported in Table 8 for $N = 20$, given that this panel dimension is closer to the ones used in the empirical practice – Tables A.2 and A.3 in the supplementary material present the results of the empirical size and power of the panel data cointegration test statistics for $N = \{10, 50\}$, respectively.

When the SAR specification is used, both $CADF_P$ and $Z_\tau$statistics show the correct size when $\vartheta = 0.4$. However, size distortions (over-rejection problems) are observed when $\vartheta = 0.8$, being the size distortions comparable for both test statistics – in some cases, size distortions are larger for the $Z_\tau$statistic ($N = 10$ and $N = 20$). In general, the Z_τ statistic is more powerful, although in some cases this might be due to the effects of the size distortions. As expected, the empirical power of both test statistics tends to one as N and/or T increase.

When the spatial dependence is driven by a SMA specification, both test statistics have the correct empirical size for $\vartheta = 0.4$, but show over-rejection problems when $\vartheta = 0.8$. In this regard, the size distortions are less important for the $CADF_P$ statistic, although the distortions almost disappear for both statistics when $N = 50$. As for the empirical power, we observe that the Z_τ statistic is more powerful than the $CADF_P$ statistic in all cases.

The three SEC specifications that we have considered lead to similar qualitative results. For $N = 20$ and $N = 50$, the empirical size of the two statistics is close to the nominal one regardless of the value of T and ϑ. Only mild overrejection problems are found for the Z_τ statistic when $N = 10$, while the $CADF_P$ statistic shows good performance. The Z_τ statistic is more powerful than the $CADF_P$ statistic when $N = 20$ and $N = 50$, but the performance of the $CADF_P$ statistic for $N = 10$ is very good if one bear in mind that the empirical size is controlled. Finally, the empirical power of both statistics tends to one as T gets large.

In summary we may conclude from the results of the simulation experiments that there is some evidence in favour of the dominance of factor-based procedures over the CCE approach. However, there may be circumstances where the factor approach is not strictly applicable (such as in the semi-strong or weak specifications, and when N is really small). Allied to the convenience of the CCE approach and equivalent performances for large T, these are good reasons to propose the use of our new test for cointegration in panels.

5 Empirical illustrations

5.1 House prices in the US

Holly et al. (2010) analyze the long-run relationship between the logarithm of the real house price index ($p_{i,t}$) and the logarithm of the real per capita disposable income ($y_{i,t}$) for 48 US States and the District of Columbia ($N = 49$) using annual data between 1975 and 2003 ($T = 29$) – see Figures 1 and 2. The model under investigation is given by

$$p_{i,t} = \alpha_i + \beta y_{i,t} + u_{i,t},$$

where note that slope homogeneity is imposed. The computation of the CD test statistic in Pesaran (2004) leads us to reject the null hypothesis of no cross-section correlation for the panels of the variables involved in the model, which indicates that panel cointegration test statistics that account for the presence of cross-section dependence have to be used – see Table 10.

For this example we do not undertake a comparison with the test statistic in Banerjee and Carrion-i-Silvestre (2015) since the T dimension is too small relative to N for our needs (in order to enable consistent computation of the factors). However it may be seen as an advantage of the CCE-based approach that a feasible test for cointegration can be constructed in the presence of cross-section dependence for reasonably small N and T – see tables for size and power properties.

We have computed the individual CCE test statistics proposed in this paper using up to four lags for the autoregressive correction in (12) and, as in Holly et al. (2010), considering the presence of one common factor. Table 9 shows that the null hypothesis of no cointegration is rejected at the 5% level of significance in 3 ($p = 0$), 8 ($p = 1$), 13 ($p = 2$), 18 ($p = 3$) and 18 ($p = 4$) cases out of 49 – if the level of significance is set at the 10%, rejection happens in 6 ($p = 0$), 13 ($p = 1$), 19 ($p = 2$), 25 ($p = 3$) and 20 ($p = 4$) cases out of 49. The same results are obtained regardless of whether the truncated or untruncated version of the statistic is used. Therefore, even in the most favorable situation, evidence in favor of cointegration is found for only half of the units. It would be the case that pooling the individual information will lead to better statistical inference, provided that the assumption of cross-section independence of $\hat{e}_{i,t}$, $i = 1, 2, \ldots, N$, in (12) is met. The computation of the $CADF_P$ statistic gives $CADF_P = -1.85$ ($p = 0$), $CADF_P = -2.56$ ($p = 1$) and $CADF_P = -2.78$ ($p = 2$), depending on the order of the autoregressive correction that is used. As can be seen, when we compare the values of the $CADF_P$ statistic with the critical values in Table 1 we conclude that, except for $p = 0$, the null hypothesis of no cointegration is rejected at the 5% level of significance. However, it should borne in mind that rejection of the null hypothesis does not necessarily imply that cointegration holds for all units.

5.2 Production function

The second empirical application focuses on the estimation of a production function using the data in Banerjee, Eberhardt and Reade (2010) taken from the Penn World Table database (version 6.3). We define a panel data set of developed countries that includes Australia, Austria, Belgium, Canada, Denmark, Finland, France, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom and the United States. The selection of these countries allows us to have a balanced panel data set covering the period between 1951 and 2007. Notice also that our data set includes almost all EU-15 countries – we have not been able to include Germany because of lack of information between 1951 and 1969 – and almost all G7 countries – the exception is Japan, for which we do not have information for the whole period. Therefore, we deal with a panel data set of dimension $T = 57$ and $N = 19$, which fits the requirement of having a panel with T larger than N. The model that is estimated is given by:

$$y_{i,t} = \alpha_i + \beta_1 l_{i,t} + \beta_2 k_{i,t} + u_{i,t},$$

where $y_{i,t}$ denotes the logarithm of the real GDP per capita, $l_{i,t}$ is the logarithm of the population and $k_{i,t}$ is the logarithm of the real capital stock per capita. As before, the CD test statistic in Pesaran (2004) rejects the null hypothesis of no cross-section correlation for the panels of the variables involved in the model, which indicates that panel cointegration test statistics that account for the presence of cross-section dependence have to be used – see Table 10. The CCE estimation of the slope parameters equals $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)' = (0.8, 0.78)'$.

Table 11 presents the individual CCE $t_{\hat{\alpha}_{i,0}}$ statistics, $i = 1, 2, \ldots, 19$, considering that there is one common factor. As can be seen, using the untruncated version of the statistic the null hypothesis of no cointegration can be rejected at the 5% level of significance in 2 ($p = 0$), 1 ($p = 1$), 2 ($p = 2$), 1 ($p = 3$) and 1 ($p = 4$) cases out of 19 – we use the critical values for $N = 20$ and $T = 50$. If the level of significance is set at the 10% level, the rejection of the null hypothesis of no cointegration happens in 3 ($p = 0$), 3 ($p = 1$), 2 ($p = 2$), 3 ($p = 3$) and 3 ($p = 4$) cases out of 19. If we use the truncated version of the statistic, the results that are obtained are almost identical, with the marginal exception for Spain with $p = 3$, where now the null hypothesis of no cointegration cannot be rejected at the 10% level of significance. Thus, using the individual based statistics we find little evidence against the null hypothesis of no cointegration. The individual information can be combined computing the $CADF_P$ statistic, which produces $CADF_P = -1.68$ ($p = 0$), $CADF_P = -1.71$ ($p = 1$), and $CADF_P = -1.67$ ($p = 2$), depending on the order of the autoregressive correction that is used. As can be seen, when we compare the values of the $CADF_P$ statistic with the critical values in Table 1 for $N = 20$ and $T = 50$ we conclude that the null hypothesis of no cointegration cannot be rejected at the 5% level of significance, regardless of the order of autocorrelation that is considered.

The results of the test statistic in Banerjee and Carrion-i-Silvestre (2015) with up to six common factors are reported in Table 12. We present two different sets of results depending on whether or not the variables are divided by their standard deviations when using principal components – see Banerjee and Carrion-i-Silvestre (2015) for further details. Without this transformation, the panel BIC information criterion in Bai and Ng (2002) leads to selection of the maximum number of factors that is allowed. In this case, all the estimated common factors are non-stationary. Once transformed, the panel BIC indicates that there is only one integrated common factor. However, regardless of the number of common factors or the transformation, the statistics in Banerjee and Carrion-i-Silvestre (2015) indicate that the idiosyncratic disturbance terms are stationary. It is worth mentioning that rejection of the null hypothesis of no cointegration does not necessarily mean that all cross-section units are cointegrated. Therefore we cannot conclude that the variables in the vector $Y_{i,t} = (y_{i,t}, l_{i,t}, k_{i,t})'$ are cointegrated, since at least one non-stationary common factor is detected. Cointegration is possible only by the inclusion of common factors in the model.

6 Conclusions

The paper has shown that consistent estimate of the long-run average coefficient is obtained when time series in the panel data are cross-section dependence, which is accounted for using a common factor model approach. The estimation procedure that is applied is based on the CCE

approach in Pesaran (2006). Our result contributes to the literature of non-stationary panel data analysis, where consistent estimation of the parameters of the model is feasible in a spurious regression framework. The paper conducts an extensive simulation exercise to study the finite sample performance of the statistic that has been proposed, allowing for weak and strong cross-section dependence. The two empirical applications illustrate the effectiveness of the respective approaches. Where a weak dependence structure is plausible such as in the house prices example, the use of CCE-based tests provides satisfactory and confirmatory results. Where however an integrated trend may be relevant, the restriction of being unable to decompose between common and idiosyncratic components (especially to have different degrees of persistence) handicaps somewhat the CCE-based tests in relation to common factor approaches. This is especially seen in the empirical example where the cointegration possibility is found to be not among the original variables (between output, labour and income) but between the original variables and an integrated stochastic common trend.

References

[1] Bai, J. and Ng, S. (2004): A PANIC attack on unit roots and cointegration. *Econometrica* 72, 1127-1177.

[2] Bai, J., Kao, C. and Ng, S. (2009): Panel cointegration with global stochastic trends. *Journal of Econometrics* 149, 82-99.

[3] Baltagi, B. H., Bresson, G. and Pirotte, A. (2007): Panel unit root tests and spatial dependence. *Journal of Applied Econometrics* 22, 339-360.

[4] Banerjee, A. and Carrion-i-Silvestre, J. L. (2015): Cointegration in panel data with breaks and cross-section dependence. *Journal of Applied Econometrics* 30, 1-23.

[5] Banerjee, A. and Carrion-i-Silvestre, J. L. (2011): Testing for Panel Cointegration using Common Correlated Effects Estimators. *Working Paper*, University of Barcelona.

[6] Banerjee, A., Eberhardt, M. and Reade, J. J. (2010): Panel estimation for worriers. *Mimeo*, Department of Economics, University of Birmingham.

[7] Banerjee, A., Marcellino, M. and Osbat, C. (2005): Testing for PPP: Should we use panel methods? *Empirical Economics* 30, 77–91.

[8] Breitung, J. and Pesaran, M. H. (2008): Unit Roots and Cointegration in Panels. In Matyas, L., and P. Sevestre (Eds.) The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice, 279–322. Kluwer Academic Publishers, Boston.

[9] Chudik, A. and Pesaran, M. H. and Tosetti, E. (2011): Weak and strong cross section dependence and estimation of large panels. *Econometrics Journal* 14, C45-C90.

[10] Dickey, D. A. and Fuller, W. A. (1979): Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427-431.

[11] Gengenbach, C., Urbain, J. P. and Westerlund, J. (2016): Error correction testing in panels with common stochastic trends. *Journal of Applied Econometrics* 31, 982-1004.

[12] Granger, C. and Newbold, P. (1974): Spurious regressions in econometrics. *Journal of Econometrics* 2, 111-20.

[13] Holly, S., Pesaran, M. H. and Yamagata, T. (2010): A spatio-temporal model of house prices in the USA. *Journal of Econometrics* 158, 160-173.

[14] Kapetanios, G., Pesaran, M. H. and Yamagata, T. (2011): Panels with nonstationary multifactor error structures. *Journal of Econometrics* 160, 326-348.

[15] Kao, C. (1999): Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics* 90, 1-44.

[16] Kao, C., Trapani, L. and Urga, G. (2011): Asymptotics for panel models with common shocks. *Econometric Reviews* forthcoming.

[17] Pedroni, P. (2000): Fully modified OLS for heterogeneous cointegrated panels. *Advances in Econometrics* 15, 93-130.

[18] Pesaran, M. H. (2004): General diagnostic tests for cross section dependence in panels. *CESifo Working Papers* No. 1233.

[19] Pesaran, M. H. (2006): Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74, 967-1012.

[20] Pesaran, M. H. (2007): A simple panel unit root test in the presence of cross section dependence. *Journal of Applied Econometrics* 22, 265-312.

[21] Pesaran, M. H., Smith, L. V. and Yamagata, T. (2013): Panel unit root tests in the presence of a multifactor error structure. *Journal of Econometrics* 175, 94-115.

[22] Phillips, P. C. B. (1986): Understanding spurious regressions in econometrics. *Journal of Econometrics* 31, 311-340.

[23] Phillips, P. C. B. and Moon, H. R. (1999): Linear regression limit theory for nonstationary panel data. *Econometrica* 67, 1057-1111.

[24] Phillips, P. C. B. and Moon, H. R. (2000): Nonstationary panel data analysis: An overview of some recent developments. *Econometric Reviews* 19, 263-286.

[25] Urbain, J. P. and Westerlund, J. (2011): Least Squares Asymptotics in Spurious and Cointegrated Panel Regressions with Common and Idiosyncratic Stochastic Trends. *Oxford Bulletin of Economics and Statistics* 73, 119-139.

[26] Urbain, J. P. and Westerlund, J. (2015): Cross sectional averages or principal components? *Journal of econometrics* 185, 372-377.

Table 1: Critical values for the $CADF_P$ test statistic with one common factor for Model 1. The rank condition is met with inequality

				5% level of significance						10% level of significance					
$k+1$	r	p	$T\backslash N$	20	30	50	70	100	200	20	30	50	70	100	200
2	1	0	30	-2.32	-2.27	-2.22	-2.20	-2.18	-2.17	-2.22	-2.18	-2.14	-2.13	-2.12	-2.11
2	1	0	50	-2.27	-2.22	-2.18	-2.16	-2.14	-2.12	-2.18	-2.14	-2.11	-2.09	-2.08	-2.07
2	1	0	70	-2.26	-2.21	-2.16	-2.14	-2.13	-2.11	-2.17	-2.13	-2.09	-2.08	-2.07	-2.05
2	1	0	100	-2.25	-2.20	-2.15	-2.13	-2.12	-2.10	-2.16	-2.12	-2.08	-2.07	-2.06	-2.05
2	1	0	200	-2.23	-2.18	-2.14	-2.12	-2.11	-2.09	-2.15	-2.10	-2.07	-2.06	-2.05	-2.04
2	1	1	30	-2.35	-2.30	-2.25	-2.24	-2.22	-2.20	-2.24	-2.20	-2.17	-2.16	-2.15	-2.14
2	1	1	50	-2.28	-2.24	-2.19	-2.17	-2.16	-2.14	-2.19	-2.15	-2.12	-2.11	-2.09	-2.08
2	1	1	70	-2.26	-2.21	-2.17	-2.15	-2.14	-2.12	-2.17	-2.14	-2.10	-2.08	-2.08	-2.06
2	1	1	100	-2.25	-2.20	-2.15	-2.14	-2.12	-2.10	-2.16	-2.12	-2.09	-2.07	-2.06	-2.05
2	1	1	200	-2.24	-2.18	-2.14	-2.12	-2.11	-2.09	-2.15	-2.11	-2.07	-2.06	-2.05	-2.04
2	1	2	30	-2.31	-2.25	-2.21	-2.20	-2.18	-2.16	-2.20	-2.16	-2.12	-2.12	-2.10	-2.09
2	1	2	50	-2.25	-2.21	-2.17	-2.14	-2.13	-2.11	-2.16	-2.12	-2.09	-2.08	-2.06	-2.05
2	1	2	70	-2.24	-2.19	-2.15	-2.13	-2.12	-2.10	-2.15	-2.11	-2.08	-2.06	-2.06	-2.04
2	1	2	100	-2.24	-2.19	-2.14	-2.12	-2.11	-2.09	-2.15	-2.11	-2.07	-2.06	-2.05	-2.04
2	1	2	200	-2.23	-2.17	-2.13	-2.11	-2.10	-2.08	-2.14	-2.10	-2.06	-2.05	-2.04	-2.03
3	1	0	30	-2.34	-2.28	-2.22	-2.20	-2.18	-2.17	-2.24	-2.19	-2.15	-2.13	-2.12	-2.11
3	1	0	50	-2.29	-2.23	-2.18	-2.16	-2.15	-2.12	-2.20	-2.15	-2.11	-2.09	-2.09	-2.07
3	1	0	70	-2.27	-2.22	-2.16	-2.14	-2.13	-2.11	-2.18	-2.14	-2.10	-2.08	-2.07	-2.06
3	1	0	100	-2.26	-2.21	-2.16	-2.14	-2.12	-2.10	-2.17	-2.13	-2.09	-2.07	-2.06	-2.05
3	1	0	200	-2.25	-2.19	-2.14	-2.12	-2.11	-2.09	-2.16	-2.11	-2.08	-2.06	-2.05	-2.04
3	1	1	30	-2.36	-2.31	-2.26	-2.23	-2.22	-2.20	-2.26	-2.21	-2.18	-2.16	-2.15	-2.14
3	1	1	50	-2.30	-2.24	-2.20	-2.17	-2.16	-2.14	-2.21	-2.16	-2.12	-2.11	-2.10	-2.08
3	1	1	70	-2.28	-2.22	-2.17	-2.15	-2.14	-2.12	-2.19	-2.14	-2.10	-2.09	-2.08	-2.07
3	1	1	100	-2.26	-2.21	-2.16	-2.14	-2.12	-2.10	-2.18	-2.13	-2.09	-2.08	-2.07	-2.05
3	1	1	200	-2.25	-2.19	-2.15	-2.13	-2.11	-2.09	-2.16	-2.12	-2.08	-2.06	-2.05	-2.04
3	1	2	30	-2.31	-2.26	-2.21	-2.19	-2.18	-2.16	-2.20	-2.16	-2.13	-2.11	-2.10	-2.09
3	1	2	50	-2.27	-2.21	-2.17	-2.14	-2.13	-2.11	-2.17	-2.13	-2.09	-2.08	-2.07	-2.05
3	1	2	70	-2.25	-2.20	-2.15	-2.13	-2.12	-2.10	-2.16	-2.12	-2.08	-2.07	-2.06	-2.04
3	1	2	100	-2.24	-2.19	-2.15	-2.13	-2.11	-2.09	-2.16	-2.11	-2.08	-2.06	-2.05	-2.04
3	1	2	200	-2.24	-2.18	-2.14	-2.12	-2.10	-2.08	-2.15	-2.11	-2.07	-2.06	-2.05	-2.03
4	1	0	30	-2.34	-2.28	-2.23	-2.20	-2.18	-2.17	-2.24	-2.20	-2.15	-2.14	-2.12	-2.11
4	1	0	50	-2.30	-2.24	-2.18	-2.16	-2.15	-2.13	-2.21	-2.16	-2.12	-2.10	-2.09	-2.07
4	1	0	70	-2.28	-2.22	-2.17	-2.15	-2.13	-2.11	-2.19	-2.14	-2.10	-2.09	-2.07	-2.06
4	1	0	100	-2.27	-2.21	-2.16	-2.14	-2.12	-2.10	-2.18	-2.13	-2.09	-2.08	-2.06	-2.05
4	1	0	200	-2.26	-2.20	-2.15	-2.13	-2.11	-2.09	-2.17	-2.12	-2.08	-2.07	-2.05	-2.04
4	1	1	30	-2.37	-2.31	-2.26	-2.23	-2.22	-2.20	-2.26	-2.22	-2.18	-2.16	-2.15	-2.14
4	1	1	50	-2.31	-2.25	-2.20	-2.17	-2.16	-2.14	-2.21	-2.16	-2.13	-2.11	-2.10	-2.08
4	1	1	70	-2.29	-2.23	-2.18	-2.16	-2.14	-2.12	-2.19	-2.15	-2.11	-2.09	-2.08	-2.07
4	1	1	100	-2.27	-2.21	-2.16	-2.14	-2.13	-2.11	-2.18	-2.13	-2.09	-2.08	-2.07	-2.05
4	1	1	200	-2.25	-2.20	-2.15	-2.13	-2.11	-2.09	-2.17	-2.12	-2.09	-2.07	-2.05	-2.04
4	1	2	30	-2.31	-2.26	-2.22	-2.19	-2.17	-2.16	-2.21	-2.16	-2.13	-2.11	-2.10	-2.09
4	1	2	50	-2.27	-2.21	-2.17	-2.15	-2.13	-2.11	-2.17	-2.13	-2.09	-2.08	-2.07	-2.05
4	1	2	70	-2.26	-2.21	-2.16	-2.14	-2.12	-2.10	-2.17	-2.12	-2.09	-2.07	-2.05	-2.04
4	1	2	100	-2.25	-2.20	-2.14	-2.13	-2.11	-2.09	-2.16	-2.12	-2.08	-2.06	-2.05	-2.04
4	1	2	200	-2.25	-2.19	-2.14	-2.12	-2.10	-2.08	-2.16	-2.11	-2.08	-2.06	-2.04	-2.03

Table 2: Critical values for the $CADF_P$ test statistic with one common factor for Model 2. The rank condition is met with inequality

				5% level of significance						10% level of significance					
$k+1$	r	p	$T\backslash N$	20	30	50	70	100	200	20	30	50	70	100	200
2	1	0	30	-2.92	-2.86	-2.81	-2.78	-2.76	-2.74	-2.82	-2.78	-2.74	-2.72	-2.70	-2.69
2	1	0	50	-2.83	-2.77	-2.72	-2.70	-2.68	-2.65	-2.74	-2.70	-2.66	-2.64	-2.63	-2.61
2	1	0	70	-2.79	-2.74	-2.69	-2.66	-2.65	-2.62	-2.71	-2.67	-2.63	-2.61	-2.59	-2.58
2	1	0	100	-2.77	-2.71	-2.66	-2.64	-2.62	-2.60	-2.69	-2.65	-2.61	-2.59	-2.57	-2.56
2	1	0	200	-2.74	-2.69	-2.64	-2.62	-2.60	-2.57	-2.67	-2.62	-2.58	-2.56	-2.55	-2.53
2	1	1	30	-2.96	-2.91	-2.86	-2.84	-2.83	-2.81	-2.86	-2.82	-2.79	-2.77	-2.76	-2.74
2	1	1	50	-2.85	-2.80	-2.75	-2.72	-2.71	-2.69	-2.76	-2.72	-2.68	-2.66	-2.65	-2.63
2	1	1	70	-2.80	-2.75	-2.70	-2.68	-2.66	-2.64	-2.72	-2.68	-2.64	-2.62	-2.61	-2.60
2	1	1	100	-2.78	-2.72	-2.67	-2.65	-2.63	-2.61	-2.70	-2.65	-2.61	-2.60	-2.58	-2.57
2	1	1	200	-2.75	-2.69	-2.64	-2.62	-2.60	-2.58	-2.67	-2.63	-2.58	-2.57	-2.55	-2.54
2	1	2	30	-2.90	-2.85	-2.81	-2.79	-2.78	-2.76	-2.79	-2.75	-2.72	-2.70	-2.70	-2.69
2	1	2	50	-2.81	-2.76	-2.72	-2.70	-2.68	-2.66	-2.71	-2.68	-2.65	-2.63	-2.62	-2.60
2	1	2	70	-2.78	-2.72	-2.68	-2.66	-2.64	-2.62	-2.69	-2.65	-2.62	-2.60	-2.59	-2.57
2	1	2	100	-2.76	-2.70	-2.66	-2.64	-2.62	-2.60	-2.67	-2.64	-2.60	-2.58	-2.57	-2.55
2	1	2	200	-2.73	-2.68	-2.63	-2.61	-2.59	-2.57	-2.66	-2.62	-2.58	-2.56	-2.55	-2.53
3	1	0	30	-2.93	-2.86	-2.81	-2.78	-2.76	-2.74	-2.84	-2.78	-2.74	-2.72	-2.71	-2.69
3	1	0	50	-2.84	-2.78	-2.72	-2.70	-2.68	-2.66	-2.76	-2.71	-2.66	-2.64	-2.63	-2.61
3	1	0	70	-2.81	-2.75	-2.69	-2.67	-2.64	-2.62	-2.73	-2.68	-2.63	-2.61	-2.60	-2.58
3	1	0	100	-2.78	-2.72	-2.67	-2.64	-2.62	-2.60	-2.71	-2.66	-2.61	-2.59	-2.57	-2.56
3	1	0	200	-2.76	-2.70	-2.64	-2.62	-2.60	-2.57	-2.68	-2.63	-2.59	-2.57	-2.55	-2.53
3	1	1	30	-2.97	-2.91	-2.87	-2.84	-2.83	-2.81	-2.87	-2.82	-2.79	-2.77	-2.76	-2.74
3	1	1	50	-2.86	-2.80	-2.75	-2.72	-2.71	-2.69	-2.77	-2.72	-2.68	-2.66	-2.65	-2.63
3	1	1	70	-2.82	-2.76	-2.71	-2.68	-2.67	-2.64	-2.73	-2.68	-2.65	-2.62	-2.61	-2.60
3	1	1	100	-2.79	-2.73	-2.68	-2.65	-2.63	-2.61	-2.71	-2.66	-2.62	-2.60	-2.58	-2.57
3	1	1	200	-2.76	-2.70	-2.65	-2.62	-2.61	-2.58	-2.68	-2.63	-2.59	-2.57	-2.56	-2.54
3	1	2	30	-2.90	-2.85	-2.81	-2.79	-2.78	-2.76	-2.79	-2.75	-2.72	-2.71	-2.70	-2.69
3	1	2	50	-2.82	-2.76	-2.72	-2.69	-2.68	-2.66	-2.73	-2.68	-2.65	-2.63	-2.62	-2.60
3	1	2	70	-2.79	-2.73	-2.69	-2.66	-2.64	-2.62	-2.70	-2.65	-2.62	-2.60	-2.59	-2.57
3	1	2	100	-2.77	-2.71	-2.66	-2.64	-2.62	-2.60	-2.69	-2.64	-2.60	-2.58	-2.57	-2.55
3	1	2	200	-2.75	-2.69	-2.64	-2.62	-2.60	-2.57	-2.67	-2.62	-2.58	-2.56	-2.55	-2.53
4	1	0	30	-2.94	-2.87	-2.81	-2.78	-2.76	-2.74	-2.85	-2.79	-2.74	-2.72	-2.70	-2.69
4	1	0	50	-2.85	-2.79	-2.73	-2.70	-2.68	-2.66	-2.76	-2.71	-2.67	-2.65	-2.63	-2.61
4	1	0	70	-2.82	-2.75	-2.69	-2.67	-2.65	-2.62	-2.73	-2.68	-2.64	-2.62	-2.60	-2.58
4	1	0	100	-2.79	-2.73	-2.67	-2.65	-2.62	-2.60	-2.71	-2.66	-2.61	-2.59	-2.58	-2.56
4	1	0	200	-2.76	-2.70	-2.65	-2.62	-2.60	-2.58	-2.69	-2.64	-2.59	-2.57	-2.55	-2.53
4	1	1	30	-2.98	-2.92	-2.87	-2.84	-2.82	-2.81	-2.88	-2.83	-2.79	-2.77	-2.75	-2.74
4	1	1	50	-2.86	-2.81	-2.75	-2.73	-2.71	-2.69	-2.77	-2.73	-2.69	-2.67	-2.65	-2.64
4	1	1	70	-2.83	-2.76	-2.71	-2.69	-2.67	-2.64	-2.74	-2.69	-2.65	-2.63	-2.61	-2.60
4	1	1	100	-2.80	-2.74	-2.68	-2.66	-2.63	-2.62	-2.71	-2.66	-2.62	-2.60	-2.58	-2.57
4	1	1	200	-2.76	-2.71	-2.65	-2.63	-2.60	-2.58	-2.69	-2.64	-2.60	-2.58	-2.56	-2.54
4	1	2	30	-2.91	-2.86	-2.81	-2.79	-2.78	-2.76	-2.80	-2.76	-2.72	-2.71	-2.69	-2.68
4	1	2	50	-2.82	-2.77	-2.72	-2.70	-2.68	-2.66	-2.73	-2.69	-2.65	-2.63	-2.62	-2.61
4	1	2	70	-2.79	-2.74	-2.68	-2.66	-2.65	-2.63	-2.71	-2.66	-2.62	-2.60	-2.59	-2.57
4	1	2	100	-2.77	-2.71	-2.66	-2.64	-2.62	-2.60	-2.69	-2.64	-2.60	-2.59	-2.57	-2.56
4	1	2	200	-2.75	-2.70	-2.64	-2.62	-2.60	-2.57	-2.68	-2.63	-2.59	-2.57	-2.55	-2.53

Table 3: Critical values for the $CADF_P$ test statistic with multiple factors for Model 1. The rank condition is met with equality

				5% level of significance						10% level of significance					
$k+1$	r	p	$T\backslash N$	20	30	50	70	100	200	20	30	50	70	100	200
2	2	0	30	-2.51	-2.45	-2.40	-2.38	-2.36	-2.34	-2.41	-2.36	-2.32	-2.30	-2.29	-2.27
2	2	0	50	-2.50	-2.44	-2.40	-2.37	-2.36	-2.33	-2.40	-2.36	-2.32	-2.30	-2.29	-2.27
2	2	0	70	-2.50	-2.44	-2.40	-2.37	-2.35	-2.33	-2.40	-2.36	-2.32	-2.30	-2.29	-2.27
2	2	0	100	-2.49	-2.44	-2.39	-2.38	-2.35	-2.33	-2.40	-2.36	-2.32	-2.31	-2.29	-2.27
2	2	0	200	-2.49	-2.44	-2.39	-2.37	-2.36	-2.34	-2.40	-2.35	-2.32	-2.31	-2.29	-2.28
2	2	1	30	-2.54	-2.47	-2.41	-2.39	-2.37	-2.35	-2.41	-2.36	-2.32	-2.31	-2.29	-2.27
2	2	1	50	-2.50	-2.45	-2.40	-2.37	-2.36	-2.33	-2.40	-2.35	-2.32	-2.29	-2.28	-2.27
2	2	1	70	-2.50	-2.44	-2.40	-2.37	-2.35	-2.33	-2.40	-2.35	-2.32	-2.30	-2.29	-2.27
2	2	1	100	-2.49	-2.44	-2.39	-2.37	-2.35	-2.33	-2.40	-2.35	-2.32	-2.30	-2.29	-2.27
2	2	1	200	-2.49	-2.44	-2.39	-2.37	-2.35	-2.34	-2.40	-2.35	-2.32	-2.31	-2.29	-2.28
2	2	2	30	-2.48	-2.40	-2.35	-2.32	-2.30	-2.27	-2.34	-2.28	-2.24	-2.22	-2.20	-2.18
2	2	2	50	-2.46	-2.40	-2.35	-2.32	-2.31	-2.29	-2.35	-2.30	-2.27	-2.24	-2.23	-2.21
2	2	2	70	-2.46	-2.41	-2.36	-2.34	-2.32	-2.29	-2.36	-2.32	-2.28	-2.26	-2.25	-2.23
2	2	2	100	-2.47	-2.41	-2.37	-2.35	-2.33	-2.31	-2.37	-2.33	-2.29	-2.28	-2.26	-2.24
2	2	2	200	-2.48	-2.42	-2.38	-2.36	-2.34	-2.32	-2.39	-2.34	-2.31	-2.29	-2.28	-2.26
3	3	0	30	-2.73	-2.66	-2.60	-2.58	-2.55	-2.53	-2.62	-2.56	-2.51	-2.49	-2.48	-2.46
3	3	0	50	-2.73	-2.66	-2.61	-2.58	-2.57	-2.54	-2.63	-2.57	-2.53	-2.51	-2.50	-2.48
3	3	0	70	-2.73	-2.67	-2.61	-2.59	-2.57	-2.55	-2.63	-2.58	-2.54	-2.52	-2.50	-2.49
3	3	0	100	-2.73	-2.67	-2.62	-2.59	-2.57	-2.55	-2.64	-2.59	-2.54	-2.52	-2.51	-2.49
3	3	0	200	-2.73	-2.67	-2.62	-2.60	-2.58	-2.56	-2.64	-2.59	-2.55	-2.53	-2.52	-2.50
3	3	1	30	-2.71	-2.64	-2.57	-2.55	-2.52	-2.50	-2.58	-2.52	-2.47	-2.45	-2.43	-2.41
3	3	1	50	-2.71	-2.64	-2.59	-2.55	-2.55	-2.52	-2.60	-2.54	-2.50	-2.47	-2.47	-2.44
3	3	1	70	-2.71	-2.65	-2.59	-2.57	-2.55	-2.53	-2.61	-2.56	-2.51	-2.50	-2.48	-2.47
3	3	1	100	-2.72	-2.65	-2.61	-2.58	-2.56	-2.54	-2.62	-2.57	-2.53	-2.51	-2.49	-2.48
3	3	1	200	-2.73	-2.67	-2.62	-2.59	-2.58	-2.55	-2.63	-2.59	-2.54	-2.52	-2.51	-2.49
3	3	2	30	-2.61	-2.53	-2.46	-2.43	-2.40	-2.37	-2.46	-2.40	-2.34	-2.32	-2.29	-2.27
3	3	2	50	-2.64	-2.56	-2.51	-2.48	-2.46	-2.44	-2.52	-2.45	-2.41	-2.39	-2.38	-2.36
3	3	2	70	-2.66	-2.60	-2.54	-2.52	-2.49	-2.47	-2.55	-2.50	-2.45	-2.44	-2.42	-2.41
3	3	2	100	-2.68	-2.62	-2.57	-2.54	-2.52	-2.50	-2.58	-2.53	-2.49	-2.47	-2.45	-2.44
3	3	2	200	-2.71	-2.65	-2.60	-2.57	-2.56	-2.53	-2.61	-2.57	-2.52	-2.50	-2.49	-2.47
4	4	0	30	-2.91	-2.83	-2.76	-2.74	-2.71	-2.68	-2.79	-2.73	-2.68	-2.65	-2.63	-2.61
4	4	0	50	-2.92	-2.85	-2.79	-2.76	-2.74	-2.72	-2.82	-2.76	-2.71	-2.69	-2.67	-2.65
4	4	0	70	-2.93	-2.86	-2.80	-2.78	-2.76	-2.73	-2.83	-2.77	-2.73	-2.71	-2.69	-2.67
4	4	0	100	-2.95	-2.87	-2.81	-2.79	-2.77	-2.75	-2.85	-2.79	-2.74	-2.72	-2.70	-2.68
4	4	0	200	-2.95	-2.88	-2.83	-2.80	-2.78	-2.76	-2.85	-2.80	-2.76	-2.73	-2.72	-2.70
4	4	1	30	-2.84	-2.75	-2.68	-2.65	-2.63	-2.60	-2.70	-2.63	-2.57	-2.55	-2.53	-2.51
4	4	1	50	-2.87	-2.80	-2.73	-2.71	-2.69	-2.66	-2.75	-2.70	-2.65	-2.62	-2.60	-2.59
4	4	1	70	-2.89	-2.82	-2.77	-2.74	-2.72	-2.69	-2.79	-2.73	-2.69	-2.66	-2.65	-2.62
4	4	1	100	-2.92	-2.84	-2.79	-2.76	-2.74	-2.72	-2.82	-2.75	-2.71	-2.69	-2.67	-2.65
4	4	1	200	-2.93	-2.87	-2.81	-2.78	-2.77	-2.75	-2.84	-2.79	-2.74	-2.72	-2.70	-2.69
4	4	2	30	-2.75	-2.64	-2.54	-2.50	-2.46	-2.43	-2.57	-2.49	-2.40	-2.37	-2.34	-2.32
4	4	2	50	-2.75	-2.68	-2.62	-2.59	-2.56	-2.54	-2.63	-2.57	-2.52	-2.49	-2.47	-2.46
4	4	2	70	-2.81	-2.74	-2.69	-2.66	-2.63	-2.61	-2.70	-2.64	-2.60	-2.57	-2.56	-2.54
4	4	2	100	-2.86	-2.79	-2.73	-2.71	-2.69	-2.66	-2.76	-2.69	-2.65	-2.63	-2.61	-2.59
4	4	2	200	-2.90	-2.84	-2.79	-2.76	-2.74	-2.72	-2.81	-2.76	-2.71	-2.69	-2.67	-2.66

Table 4: Critical values for the $CADF_P$ test statistic with multiple factors for Model 2. The rank condition is met with equality

				5% level of significance						10% level of significance					
$k+1$	r	p	$T\backslash N$	20	30	50	70	100	200	20	30	50	70	100	200
2	2	0	30	-2.97	-2.90	-2.86	-2.82	-2.80	-2.78	-2.87	-2.82	-2.78	-2.75	-2.74	-2.72
2	2	0	50	-2.95	-2.89	-2.84	-2.81	-2.79	-2.77	-2.86	-2.81	-2.77	-2.75	-2.73	-2.72
2	2	0	70	-2.94	-2.88	-2.83	-2.81	-2.79	-2.77	-2.85	-2.81	-2.77	-2.75	-2.73	-2.71
2	2	0	100	-2.94	-2.88	-2.83	-2.81	-2.79	-2.76	-2.85	-2.81	-2.77	-2.75	-2.73	-2.71
2	2	0	200	-2.93	-2.88	-2.83	-2.80	-2.78	-2.76	-2.85	-2.81	-2.76	-2.75	-2.73	-2.71
2	2	1	30	-3.00	-2.94	-2.89	-2.86	-2.84	-2.81	-2.88	-2.83	-2.79	-2.77	-2.75	-2.73
2	2	1	50	-2.96	-2.90	-2.84	-2.82	-2.80	-2.78	-2.86	-2.81	-2.77	-2.75	-2.74	-2.72
2	2	1	70	-2.94	-2.89	-2.84	-2.81	-2.79	-2.77	-2.85	-2.80	-2.76	-2.75	-2.73	-2.72
2	2	1	100	-2.94	-2.88	-2.83	-2.81	-2.79	-2.77	-2.85	-2.81	-2.77	-2.75	-2.73	-2.71
2	2	1	200	-2.93	-2.88	-2.83	-2.80	-2.79	-2.76	-2.85	-2.80	-2.76	-2.75	-2.73	-2.71
2	2	2	30	-2.94	-2.86	-2.81	-2.77	-2.75	-2.72	-2.80	-2.74	-2.70	-2.67	-2.65	-2.63
2	2	2	50	-2.90	-2.85	-2.79	-2.77	-2.75	-2.73	-2.80	-2.75	-2.71	-2.69	-2.68	-2.66
2	2	2	70	-2.91	-2.85	-2.80	-2.77	-2.75	-2.73	-2.81	-2.76	-2.72	-2.71	-2.69	-2.67
2	2	2	100	-2.91	-2.86	-2.81	-2.79	-2.76	-2.74	-2.82	-2.78	-2.73	-2.72	-2.70	-2.69
2	2	2	200	-2.92	-2.86	-2.81	-2.79	-2.77	-2.75	-2.84	-2.79	-2.75	-2.73	-2.72	-2.70
3	3	0	30	-3.14	-3.06	-3.00	-2.98	-2.95	-2.92	-3.03	-2.97	-2.92	-2.90	-2.88	-2.86
3	3	0	50	-3.13	-3.06	-3.01	-2.98	-2.96	-2.93	-3.04	-2.98	-2.93	-2.91	-2.89	-2.88
3	3	0	70	-3.13	-3.07	-3.01	-2.98	-2.96	-2.94	-3.04	-2.99	-2.94	-2.92	-2.90	-2.89
3	3	0	100	-3.13	-3.07	-3.01	-2.99	-2.97	-2.94	-3.04	-2.99	-2.94	-2.93	-2.91	-2.89
3	3	0	200	-3.13	-3.07	-3.02	-2.99	-2.98	-2.95	-3.05	-3.00	-2.95	-2.93	-2.92	-2.90
3	3	1	30	-3.12	-3.05	-2.99	-2.96	-2.94	-2.91	-2.99	-2.93	-2.88	-2.86	-2.84	-2.82
3	3	1	50	-3.11	-3.04	-2.99	-2.96	-2.94	-2.91	-3.01	-2.94	-2.91	-2.88	-2.87	-2.85
3	3	1	70	-3.11	-3.05	-3.00	-2.97	-2.95	-2.93	-3.01	-2.96	-2.92	-2.90	-2.88	-2.86
3	3	1	100	-3.12	-3.05	-3.00	-2.98	-2.96	-2.93	-3.03	-2.98	-2.93	-2.91	-2.89	-2.88
3	3	1	200	-3.12	-3.06	-3.01	-2.99	-2.97	-2.95	-3.04	-2.99	-2.95	-2.93	-2.91	-2.89
3	3	2	30	-3.01	-2.93	-2.85	-2.82	-2.79	-2.76	-2.85	-2.78	-2.72	-2.70	-2.68	-2.65
3	3	2	50	-3.02	-2.95	-2.90	-2.87	-2.85	-2.83	-2.91	-2.85	-2.81	-2.78	-2.77	-2.75
3	3	2	70	-3.05	-2.99	-2.93	-2.90	-2.88	-2.86	-2.94	-2.90	-2.85	-2.83	-2.81	-2.79
3	3	2	100	-3.08	-3.01	-2.95	-2.94	-2.91	-2.89	-2.98	-2.93	-2.88	-2.87	-2.85	-2.83
3	3	2	200	-3.10	-3.04	-2.99	-2.97	-2.95	-2.92	-3.01	-2.97	-2.92	-2.90	-2.89	-2.87
4	4	0	30	-3.28	-3.19	-3.13	-3.10	-3.07	-3.04	-3.16	-3.09	-3.04	-3.02	-2.99	-2.97
4	4	0	50	-3.29	-3.22	-3.16	-3.13	-3.10	-3.08	-3.19	-3.13	-3.08	-3.06	-3.04	-3.02
4	4	0	70	-3.30	-3.23	-3.17	-3.14	-3.12	-3.09	-3.20	-3.15	-3.10	-3.08	-3.05	-3.04
4	4	0	100	-3.31	-3.24	-3.18	-3.15	-3.13	-3.11	-3.22	-3.16	-3.11	-3.09	-3.07	-3.05
4	4	0	200	-3.32	-3.25	-3.19	-3.16	-3.14	-3.12	-3.23	-3.18	-3.13	-3.11	-3.09	-3.07
4	4	1	30	-3.21	-3.12	-3.04	-3.01	-2.98	-2.95	-3.05	-2.99	-2.92	-2.90	-2.88	-2.86
4	4	1	50	-3.23	-3.16	-3.10	-3.07	-3.04	-3.02	-3.11	-3.06	-3.01	-2.99	-2.97	-2.95
4	4	1	70	-3.26	-3.19	-3.13	-3.10	-3.08	-3.05	-3.16	-3.10	-3.05	-3.02	-3.01	-2.99
4	4	1	100	-3.28	-3.21	-3.15	-3.12	-3.10	-3.08	-3.18	-3.13	-3.08	-3.05	-3.04	-3.02
4	4	1	200	-3.30	-3.24	-3.18	-3.15	-3.13	-3.11	-3.21	-3.16	-3.11	-3.09	-3.07	-3.05
4	4	2	30	-3.15	-3.02	-2.92	-2.87	-2.83	-2.78	-2.94	-2.85	-2.77	-2.73	-2.69	-2.66
4	4	2	50	-3.09	-3.03	-2.96	-2.93	-2.90	-2.88	-2.97	-2.91	-2.86	-2.84	-2.82	-2.80
4	4	2	70	-3.16	-3.09	-3.03	-3.00	-2.98	-2.96	-3.05	-3.00	-2.95	-2.93	-2.91	-2.89
4	4	2	100	-3.21	-3.15	-3.09	-3.06	-3.04	-3.02	-3.12	-3.06	-3.01	-2.99	-2.97	-2.95
4	4	2	200	-3.27	-3.21	-3.15	-3.12	-3.10	-3.08	-3.18	-3.13	-3.08	-3.06	-3.04	-3.02

Table 5: Empirical size and power of the panel cointegration test statistics with strong cross-section dependence

			$N = 10$			$N = 20$			$N = 50$		
ϕ_i	ρ	T	Z_τ	$t_{\tilde{F}}$	$CADF_P$	Z_τ	$t_{\tilde{F}}$	$CADF_P$	Z_τ	$t_{\tilde{F}}$	$CADF_P$
1	1	50	0.124	0.010	0.051	0.061	0.061	0.047	0.052	0.066	0.050
1	1	100	0.072	0.046	0.051	0.058	0.053	0.050	0.059	0.056	0.053
1	1	250	0.066	0.051	0.045	0.053	0.055	0.043	0.048	0.049	0.052
1	0.99	50	0.123	0.010	0.053	0.060	0.061	0.047	0.053	0.064	0.050
1	0.99	100	0.072	0.048	0.051	0.059	0.056	0.045	0.059	0.058	0.049
1	0.99	250	0.065	0.064	0.042	0.053	0.061	0.038	0.048	0.062	0.038
1	0.95	50	0.120	0.010	0.047	0.056	0.074	0.040	0.049	0.077	0.038
1	0.95	100	0.070	0.076	0.040	0.058	0.092	0.029	0.056	0.095	0.021
1	0.95	250	0.064	0.235	0.025	0.053	0.257	0.012	0.045	0.276	0.006
1	0.9	50	0.118	0.013	0.040	0.056	0.102	0.027	0.047	0.107	0.020
1	0.9	100	0.069	0.167	0.030	0.058	0.200	0.016	0.056	0.211	0.007
1	0.9	250	0.065	0.642	0.021	0.051	0.788	0.006	0.046	0.809	0.001
0.99	1	50	0.127	0.010	0.055	0.064	0.061	0.048	0.062	0.065	0.052
0.99	1	100	0.096	0.047	0.057	0.078	0.053	0.058	0.102	0.057	0.068
0.99	1	250	0.186	0.053	0.085	0.261	0.056	0.107	0.476	0.050	0.160
0.99	0.99	50	0.125	0.009	0.057	0.064	0.061	0.049	0.061	0.064	0.051
0.99	0.99	100	0.095	0.048	0.054	0.079	0.056	0.055	0.101	0.057	0.061
0.99	0.99	250	0.185	0.065	0.078	0.260	0.060	0.096	0.473	0.064	0.132
0.99	0.95	50	0.121	0.009	0.050	0.061	0.073	0.041	0.058	0.078	0.040
0.99	0.95	100	0.090	0.075	0.044	0.077	0.092	0.034	0.098	0.095	0.028
0.99	0.95	250	0.181	0.244	0.050	0.261	0.258	0.038	0.469	0.278	0.040
0.99	0.9	50	0.121	0.014	0.043	0.057	0.102	0.028	0.056	0.107	0.023
0.99	0.9	100	0.089	0.167	0.032	0.077	0.202	0.019	0.097	0.211	0.010
0.99	0.9	250	0.179	0.668	0.044	0.259	0.798	0.022	0.466	0.818	0.016
0.95	1	50	0.179	0.010	0.089	0.217	0.060	0.103	0.416	0.066	0.146
0.95	1	100	0.560	0.044	0.215	0.830	0.054	0.358	0.996	0.057	0.586
0.95	1	250	1.000	0.056	0.950	1	0.054	1.000	1.000	0.051	1.000
0.95	0.99	50	0.179	0.008	0.088	0.218	0.062	0.101	0.415	0.064	0.141
0.95	0.99	100	0.561	0.047	0.215	0.832	0.058	0.354	0.997	0.058	0.581
0.95	0.99	250	1.000	0.069	0.944	1	0.064	1.000	1.000	0.065	1.000
0.95	0.95	50	0.177	0.010	0.085	0.214	0.073	0.084	0.410	0.077	0.112
0.95	0.95	100	0.556	0.078	0.187	0.830	0.093	0.300	0.996	0.097	0.495
0.95	0.95	250	1.000	0.288	0.931	1	0.271	0.999	1.000	0.285	1.000
0.95	0.9	50	0.172	0.014	0.071	0.209	0.101	0.065	0.401	0.109	0.078
0.95	0.9	100	0.557	0.175	0.160	0.826	0.206	0.233	0.996	0.213	0.396
0.95	0.9	250	1.000	0.783	0.923	1	0.83	0.998	1.000	0.836	1.000
0.9	1	50	0.338	0.010	0.213	0.742	0.061	0.318	0.981	0.067	0.532
0.9	1	100	0.971	0.046	0.765	1	0.057	0.974	1.000	0.058	1.000
0.9	1	250	1.000	0.061	1.000	1	0.056	1.000	1.000	0.051	1.000
0.9	0.99	50	0.336	0.009	0.213	0.739	0.063	0.319	0.981	0.065	0.527
0.9	0.99	100	0.970	0.050	0.764	1	0.059	0.973	1.000	0.060	1.000
0.9	0.99	250	1.000	0.076	1.000	1	0.065	1.000	1.000	0.065	1.000
0.9	0.95	50	0.336	0.010	0.200	0.736	0.074	0.296	0.981	0.079	0.492
0.9	0.95	100	0.971	0.081	0.739	1	0.094	0.966	1.000	0.100	1.000
0.9	0.95	250	1.000	0.311	1.000	1	0.280	1.000	1.000	0.284	1.000
0.9	0.9	50	0.335	0.014	0.183	0.731	0.102	0.258	0.981	0.111	0.438
0.9	0.9	100	0.970	0.181	0.710	1	0.209	0.950	1.000	0.219	1.000
0.9	0.9	250	1.000	0.833	1.000	1	0.842	1.000	1.000	0.842	1.000

Table 6: Empirical size and power of the $CADF_P$ panel cointegration test statistic with strong cross-section dependence. Rank condition is satisfied with equality and with inequality

			Equality		Inequality	
ϕ_i	ρ	T	$N = 20$	$N = 50$	$N = 20$	$N = 50$
1	1	50	0.113	0.066	0.052	0.047
1	1	100	0.110	0.068	0.053	0.051
1	1	250	0.112	0.069	0.048	0.051
1	0.99	50	0.109	0.068	0.049	0.047
1	0.99	100	0.109	0.061	0.052	0.052
1	0.99	250	0.110	0.046	0.044	0.042
1	0.95	50	0.108	0.049	0.045	0.043
1	0.95	100	0.100	0.024	0.039	0.031
1	0.95	250	0.107	0.007	0.025	0.013
1	0.9	50	0.101	0.027	0.036	0.032
1	0.9	100	0.090	0.007	0.024	0.017
1	0.9	250	0.161	0.003	0.016	0.007
0.99	1	50	0.114	0.068	0.053	0.050
0.99	1	100	0.118	0.081	0.061	0.063
0.99	1	250	0.169	0.153	0.107	0.125
0.99	0.99	50	0.113	0.068	0.051	0.050
0.99	0.99	100	0.119	0.074	0.061	0.060
0.99	0.99	250	0.169	0.125	0.101	0.112
0.99	0.95	50	0.110	0.050	0.046	0.044
0.99	0.95	100	0.107	0.032	0.047	0.037
0.99	0.95	250	0.177	0.027	0.060	0.045
0.99	0.9	50	0.104	0.029	0.037	0.032
0.99	0.9	100	0.104	0.009	0.030	0.020
0.99	0.9	250	0.262	0.012	0.041	0.022
0.95	1	50	0.162	0.143	0.097	0.111
0.95	1	100	0.273	0.419	0.278	0.410
0.95	1	250	0.716	0.987	0.993	1.000
0.95	0.99	50	0.160	0.141	0.095	0.109
0.95	0.99	100	0.270	0.409	0.279	0.408
0.95	0.99	250	0.741	0.987	0.993	1.000
0.95	0.95	50	0.153	0.118	0.085	0.090
0.95	0.95	100	0.278	0.299	0.239	0.333
0.95	0.95	250	0.889	0.985	0.988	1.000
0.95	0.9	50	0.146	0.076	0.073	0.071
0.95	0.9	100	0.300	0.197	0.198	0.253
0.95	0.9	250	0.973	0.988	0.983	1.000
0.9	1	50	0.275	0.381	0.269	0.377
0.9	1	100	0.584	0.946	0.903	0.997
0.9	1	250	0.944	1.000	1.000	1.000
0.9	0.99	50	0.274	0.379	0.271	0.377
0.9	0.99	100	0.582	0.949	0.902	0.998
0.9	0.99	250	0.959	1.000	1.000	1.000
0.9	0.95	50	0.271	0.346	0.253	0.341
0.9	0.95	100	0.628	0.933	0.887	0.995
0.9	0.95	250	0.997	1.000	1.000	1.000
0.9	0.9	50	0.266	0.281	0.226	0.291
0.9	0.9	100	0.703	0.910	0.863	0.991
0.9	0.9	250	1.000	1.000	1.000	1.000

Table 7: Empirical size and power of the panel cointegration tests with semi-strong and weak cross-section dependence through the loadings

			Semi-strong cross-section dependence						Weak cross-section dependence					
			$N=10$		$N=20$		$N=50$		$N=10$		$N=20$		$N=50$	
ϕ_i	ρ	T	Z_τ	$CADF$	Z_τ	$CADF$	Z_τ	$CADF$	Z_τ	$CADF$	Z_τ	$CADF$	Z_τ	$CADF$
1	1	50	0.105	0.051	0.051	0.050	0.055	0.049	0.098	0.051	0.051	0.047	0.053	0.052
1	1	100	0.061	0.056	0.052	0.053	0.057	0.043	0.059	0.052	0.052	0.051	0.058	0.046
1	1	250	0.053	0.050	0.056	0.046	0.049	0.048	0.053	0.050	0.054	0.047	0.047	0.048
1	0.99	50	0.105	0.051	0.051	0.050	0.055	0.050	0.098	0.051	0.052	0.047	0.053	0.052
1	0.99	100	0.062	0.056	0.053	0.054	0.057	0.043	0.059	0.053	0.052	0.050	0.057	0.047
1	0.99	250	0.059	0.048	0.055	0.045	0.050	0.049	0.056	0.049	0.055	0.047	0.047	0.049
1	0.95	50	0.105	0.051	0.053	0.047	0.056	0.048	0.100	0.051	0.052	0.047	0.052	0.053
1	0.95	100	0.067	0.056	0.055	0.049	0.057	0.042	0.062	0.053	0.052	0.050	0.057	0.048
1	0.95	250	0.065	0.049	0.061	0.040	0.052	0.048	0.054	0.051	0.054	0.045	0.047	0.049
1	0.9	50	0.106	0.052	0.052	0.043	0.057	0.048	0.098	0.051	0.052	0.046	0.052	0.053
1	0.9	100	0.073	0.054	0.058	0.043	0.059	0.044	0.062	0.054	0.052	0.049	0.057	0.048
1	0.9	250	0.066	0.051	0.064	0.040	0.053	0.049	0.054	0.050	0.055	0.045	0.047	0.049
0.99	1	50	0.109	0.052	0.060	0.051	0.066	0.051	0.104	0.053	0.060	0.046	0.066	0.056
0.99	1	100	0.075	0.061	0.078	0.060	0.102	0.057	0.076	0.059	0.079	0.060	0.103	0.061
0.99	1	250	0.164	0.085	0.246	0.111	0.475	0.135	0.170	0.083	0.252	0.100	0.479	0.136
0.99	0.99	50	0.108	0.053	0.059	0.051	0.067	0.051	0.104	0.052	0.060	0.046	0.066	0.056
0.99	0.99	100	0.076	0.061	0.078	0.060	0.103	0.057	0.078	0.058	0.078	0.060	0.102	0.061
0.99	0.99	250	0.171	0.085	0.253	0.102	0.482	0.133	0.173	0.086	0.253	0.100	0.479	0.135
0.99	0.95	50	0.108	0.053	0.059	0.050	0.067	0.050	0.105	0.052	0.059	0.047	0.066	0.056
0.99	0.95	100	0.084	0.058	0.083	0.055	0.106	0.055	0.079	0.057	0.078	0.059	0.103	0.061
0.99	0.95	250	0.192	0.082	0.272	0.088	0.493	0.131	0.175	0.084	0.255	0.097	0.480	0.135
0.99	0.9	50	0.110	0.052	0.061	0.046	0.068	0.049	0.106	0.052	0.060	0.047	0.066	0.055
0.99	0.9	100	0.090	0.059	0.083	0.052	0.110	0.054	0.080	0.058	0.079	0.059	0.104	0.062
0.99	0.9	250	0.200	0.084	0.279	0.089	0.497	0.130	0.175	0.084	0.256	0.097	0.481	0.135
0.95	1	50	0.171	0.077	0.228	0.096	0.450	0.121	0.168	0.077	0.239	0.090	0.450	0.118
0.95	1	100	0.573	0.201	0.842	0.315	0.997	0.502	0.597	0.198	0.857	0.299	0.998	0.492
0.95	1	250	0.999	0.929	1.000	0.999	1.000	1.000	1.000	0.935	1.000	0.999	1.000	1.000
0.95	0.99	50	0.171	0.078	0.229	0.098	0.448	0.120	0.168	0.076	0.239	0.089	0.449	0.119
0.95	0.99	100	0.573	0.202	0.847	0.314	0.997	0.498	0.595	0.201	0.858	0.298	0.998	0.492
0.95	0.99	250	1.000	0.930	1.000	0.999	1.000	1.000	1.000	0.936	1.000	0.999	1.000	1.000
0.95	0.95	50	0.174	0.078	0.234	0.095	0.452	0.117	0.170	0.076	0.240	0.090	0.451	0.118
0.95	0.95	100	0.600	0.201	0.865	0.301	0.998	0.496	0.598	0.199	0.862	0.298	0.998	0.493
0.95	0.95	250	1.000	0.935	1.000	0.998	1.000	1.000	1.000	0.937	1.000	0.999	1.000	1.000
0.95	0.9	50	0.176	0.076	0.242	0.087	0.456	0.111	0.171	0.076	0.239	0.089	0.451	0.117
0.95	0.9	100	0.617	0.200	0.873	0.287	0.998	0.490	0.599	0.197	0.863	0.297	0.998	0.491
0.95	0.9	250	1.000	0.940	1.000	0.999	1.000	1.000	1.000	0.938	1.000	0.999	1.000	1.000
0.9	1	50	0.396	0.172	0.802	0.281	0.991	0.435	0.416	0.173	0.818	0.268	0.991	0.432
0.9	1	100	0.986	0.732	1.000	0.963	1.000	1.000	0.994	0.738	1.000	0.952	1.000	1.000
0.9	1	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.99	50	0.398	0.172	0.801	0.281	0.991	0.434	0.415	0.173	0.818	0.268	0.992	0.431
0.9	0.99	100	0.988	0.733	1.000	0.964	1.000	1.000	0.993	0.740	1.000	0.953	1.000	1.000
0.9	0.99	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.95	50	0.401	0.170	0.811	0.272	0.991	0.433	0.418	0.171	0.818	0.268	0.992	0.430
0.9	0.95	100	0.991	0.734	1.000	0.958	1.000	1.000	0.994	0.740	1.000	0.951	1.000	1.000
0.9	0.95	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.9	50	0.407	0.172	0.821	0.265	0.992	0.434	0.417	0.172	0.819	0.268	0.992	0.430
0.9	0.9	100	0.993	0.740	1.000	0.955	1.000	1.000	0.994	0.740	1.000	0.951	1.000	1.000
0.9	0.9	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 8: Empirical size and power of the panel cointegration tests with normalized spatial dependence, $N = 20$

		SAR				SMA			
		$\vartheta = 0.4$		$\vartheta = 0.8$		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.074	0.065	0.113	0.137	0.064	0.064	0.117	0.087
1	100	0.084	0.073	0.124	0.147	0.070	0.071	0.121	0.090
1	250	0.066	0.067	0.113	0.146	0.067	0.054	0.131	0.085
0.99	50	0.080	0.066	0.114	0.142	0.070	0.066	0.118	0.091
0.99	100	0.111	0.084	0.150	0.162	0.096	0.077	0.142	0.104
0.99	250	0.240	0.119	0.263	0.211	0.239	0.111	0.273	0.145
0.95	50	0.201	0.112	0.191	0.192	0.204	0.105	0.191	0.135
0.95	100	0.691	0.311	0.545	0.365	0.688	0.309	0.517	0.330
0.95	250	1.000	0.998	0.997	0.967	1.000	0.998	0.998	0.992
0.9	50	0.627	0.287	0.411	0.349	0.662	0.280	0.386	0.311
0.9	100	1.000	0.935	0.962	0.852	0.999	0.942	0.955	0.912
0.9	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

		SEC1				SEC2			
		$\vartheta = 0.4$		$\vartheta = 0.8$		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.048	0.046	0.050	0.043	0.054	0.043	0.058	0.049
1	100	0.062	0.052	0.060	0.052	0.062	0.053	0.066	0.053
1	250	0.053	0.049	0.052	0.046	0.049	0.049	0.047	0.048
0.99	50	0.051	0.048	0.055	0.046	0.056	0.045	0.060	0.051
0.99	100	0.084	0.060	0.083	0.058	0.084	0.061	0.091	0.062
0.99	250	0.222	0.096	0.224	0.098	0.221	0.100	0.219	0.100
0.95	50	0.187	0.089	0.183	0.087	0.186	0.086	0.192	0.086
0.95	100	0.708	0.300	0.706	0.296	0.701	0.296	0.706	0.294
0.95	250	1.000	0.999	1.000	0.999	1.000	0.999	1.000	0.999
0.9	50	0.683	0.266	0.678	0.262	0.673	0.263	0.680	0.266
0.9	100	1.000	0.953	1.000	0.954	1.000	0.953	1.000	0.957
0.9	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

		SEC3			
		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.058	0.052	0.071	0.062
1	100	0.063	0.054	0.075	0.069
1	250	0.050	0.050	0.065	0.064
0.99	50	0.068	0.053	0.078	0.066
0.99	100	0.092	0.062	0.103	0.077
0.99	250	0.222	0.107	0.235	0.121
0.95	50	0.202	0.092	0.212	0.111
0.95	100	0.702	0.293	0.701	0.312
0.95	250	1.000	0.999	1.000	0.999
0.9	50	0.676	0.270	0.658	0.284
0.9	100	1.000	0.956	0.998	0.942
0.9	250	1.000	1.000	1.000	1.000

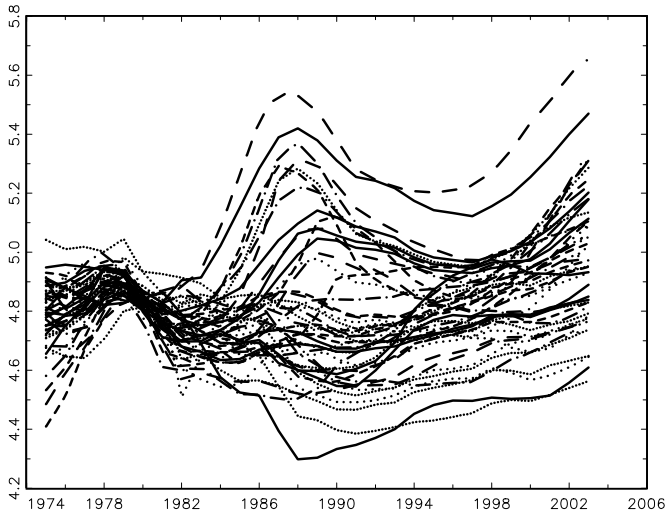

Figure 1: US State real house price

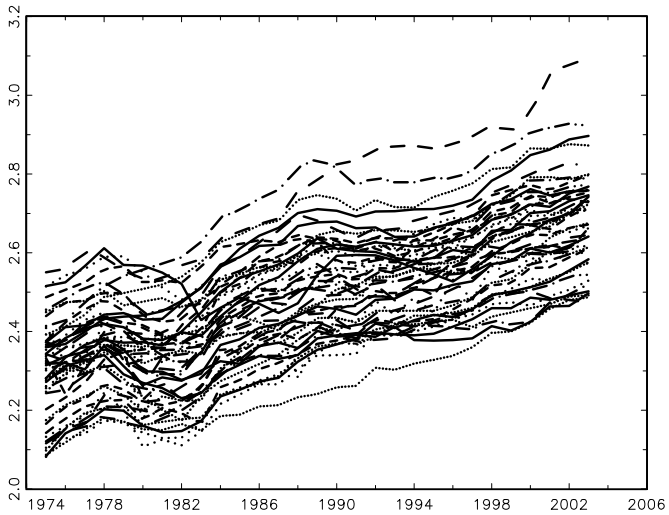

Figure 2: US State real per capita disposable income

Table 9: US Housing price and disposable income relationship. Individual cointegration statistics

	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$
Alabama	-0.827	-1.259	-3.575	-3.585	-1.858
Arkansas	-1.121	-2.557	-3.668	-2.563	-1.772
Arizona	-1.972	-2.157	-2.708	-3.491	-3.633
California	-1.418	-3.526	-3.797	-3.563	-3.555
Colorado	-1.110	-1.963	-1.867	-2.370	-1.633
Connecticut	-0.953	-2.196	-2.905	-5.117	-3.560
District of Columbia	-2.273	-2.606	-3.370*	-1.964	-1.901
Delaware	-2.346	-2.212	-3.025	-3.151*	-2.442
Florida	-2.016	-1.807	-2.664	-1.480	-1.314
Georgia	-3.758**	-3.991	-3.127*	-1.752	-4.663
Iowa	-1.409	-2.559	-2.956	-3.052	-4.393
Idaho	-1.588	-2.302	-3.826	-8.305	-4.972
Illinois	-1.669	-1.775	-1.531	-2.875	-1.736
Indiana	-1.626	-1.980	-1.839	-2.106	-2.472
Kansas	-1.343	-1.662	-1.989	-2.134	-1.758
Kentucky	-3.414*	-1.610	-1.151	-2.955	-1.420
Louisiana	-1.817	-2.932	-4.087**	-6.247**	-3.523**
Massachusetts	-1.210	-2.031	-3.119*	-4.465**	-5.050**
Maryland	-1.348	-3.061	-3.640**	-4.336**	-2.906
Maine	-0.935	-3.192*	-3.108*	-3.334*	-3.553**
Michigan	-1.521	-3.145*	-2.767	-2.450	-2.294
Minnesota	-0.379	-0.416	-2.033	-1.223	-1.121
Missouri	-1.575	-2.957	-3.456*	-1.456	-1.416
Mississippi	-3.266	-2.602	-2.341	-3.788**	-2.325
Montana	-1.578	-1.975	-2.198	-1.949	-1.316
North Carolina	-2.145	-1.814	-1.404	-3.812**	-4.361**
North Dakota	-2.636	-4.622**	-3.935**	-3.433*	-2.408
Nebraska	-1.092	-3.506**	-3.782**	-4.605**	-4.307**
New Hampshire	-0.911	-3.236*	-3.040	-3.716**	-3.744**
New Jersey	-1.798	-2.895	-2.622	-2.651	-2.565
New Mexico	-1.368	-2.165	-3.879**	-3.624**	-3.020
Nevada	-2.136	-1.832	-1.153	-2.097	-1.482
New York	-4.822**	-4.808**	-1.824	-0.230	-1.793
Ohio	-1.625	-1.792	-2.228	-2.707	-3.752**
Oklahoma	-2.114	-4.011**	-4.511**	-2.935	-5.321**
Oregon	-1.243	-1.758	-2.242	-3.802**	-3.739**
Pennsylvania	-1.769	-3.008	-2.425	-3.324*	-2.340
Rhode Island	-1.344	-3.139*	-3.317*	-5.344**	-5.342**
South Carolina	-3.112*	-5.697**	-1.794	-1.547	-1.235
South Dakota	-3.854**	-2.213	-2.229	-2.110	-3.181
Tennessee	-1.230	-1.918	-2.588	-2.714	-1.616
Texas	-2.348	-3.654**	-4.175**	-3.320*	-4.181**
Utah	-1.156	-2.919	-2.828	-3.335*	-2.552
Virginia	-1.730	-1.875	-2.767	-4.731**	-1.550
Vermont	-2.189	-2.460	-3.557	-4.176**	-2.844
Washington	-2.011	-3.409*	-2.503	-3.179*	-3.768**
Wisconsin	-1.939	-0.792	-1.184	-2.969	-2.634
West Virginia	-1.776	-1.442	-1.729	-2.155	-2.913
Wyoming	-1.927	-2.112	-3.677**	-4.694**	-3.326*

Notes: Columns 2 to 6 report the results for different lags. ** and * denote rejection of the null hypothesis of no cointegration at the 5 and 10% levels of significance, respectively.

Table 10: Residual cross correlation of ADF(p) regressions. US Housing price and production function illustrations

US Housing prices				
	ADF(1)	ADF(2)	ADF(3)	ADF(4)
$p_{i,t}$	67.07	63.16	58.99	56.28
$y_{i,t}$	93.41	89.02	88.49	82.62
Production function				
	ADF(1)	ADF(2)	ADF(3)	ADF(4)
$y_{i,t}$	30.28	28.25	28.16	28.32
$l_{i,t}$	12.09	4.63	4.43	5.11
$k_{i,t}$	42.57	24.82	25.13	25.33

Table 11: Production function. Individual CCE cointegration statistics

	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$
AUS	-0.881	1.071	0.149	-0.069	0.483
AUT	-3.189*	-4.439**	-6.153**	-4.667**	-2.858
BEL	1.351	1.942	0.406	0.703	0.111
CAN	-2.005	-2.694	-1.928	-1.136	-1.422
CHE	-1.236	-1.126	-1.220	-1.897	-1.977
DNK	-2.804	-2.430	-2.210	-2.124	-3.386**
ESP	-5.354**	-3.206*	-3.755**^{a}	-3.017*	-3.233*
FIN	-0.808	-2.137	-1.577	-1.482	-1.535
FRA	-2.401	-2.110	-1.451	-0.655	-0.927
GBR	-3.490**	-2.405	-2.674	-2.709	-3.031*
GRC	-2.387	-3.177*	-2.190	-2.226	-2.056
IRL	-0.366	-1.251	-1.202	-1.558	-1.955
ITA	-1.431	-1.139	-0.583	-0.418	0.304
LUX	-1.502	-1.904	-1.819	-1.761	-2.283
NLD	-2.439	-2.521	-2.317	-3.136*	-2.671
NOR	-1.613	-1.824	-1.510	-1.231	-1.087
PRT	-0.747	-0.593	-0.035	0.247	0.390
SWE	-0.225	-1.507	-0.945	-0.668	-1.298
USA	-0.367	-0.950	-0.673	-0.283	-0.451

Notes: Columns 2 to 6 report the results for different lags. ** and * denote rejection of the null hypothesis of no cointegration at the 5 and 10% levels of significance, respectively. a indicates that the null hypothesis is not rejected when using the truncated version of the test statistic.

Table 12: Production function. Banerjee and Carrion-i-Silvestre panel cointegration test

Not transformed							
	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
Z_τ	-3.56	-3.16	-3.09	-2.84	-2.75	-2.46	-2.53
$\hat{r}_1$ (non-parametric MQ test)	-	1	1	3	4	5	6
$\hat{r}_1$ (parametric MQ test)	-	1	2	3	4	5	6
Transformed							
	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
Z_τ	-3.56	-3.79	-3.75	-3.76	-3.88	-4.08	-3.62
$\hat{r}_1$ (non-parametric MQ test)	-	1	2	3	4	5	6
$\hat{r}_1$ (parametric MQ test)	-	1	2	3	4	5	6

Notes: Transformed means that the $y^*_{i,t}$ variables are divided by their standard deviation

Testing for Panel Cointegration using Common Correlated Effects Estimators*
Supporting information

Anindya Banerjee
Department of Economics
University of Birmingham
Edghaston, Birmingham, B15 2TT, UK
E-mail: a.banerjee@bham.ac.uk

Josep Lluís Carrion-i-Silvestre
AQR-IREA Research group
Department of Econometrics, Statistics and Spanish Economy
University of Barcelona
Av. Diagona, 690
08034, Barcelona
E-mail: carrion@ub.edu

February 17, 2017

*Carrion-i-Silvestre acknowledges the financial support from Spanish Ministerio de Ciencia y Tecnología under grant ECO2014-58991-C3-1-R. We thank Rob Taylor and two referees, George Kapetanios, Hashem Pesaran and Jean-Pierre Urbain, and the participants at the 17th Panel Data Conference in Montreal, the Conference in Honor of H. M. Pesaran in Cambridge and the XIV Encuentro de Economía Aplicada in Huelva for their helpful comments. Responsibility for any remaining errors lies with us.

A Supporting information

Lemma 1 *Define the vector of stochastic processes* $V_{i,t} = \left(e'_{i,t}, v'_t\right)'$ *that satisfies the panel functional central limit theorem (CLT):*

$$T^{-1/2}\sum_{t=1}^{[Tr]} V_{i,t} \Rightarrow C_i(1) W_i(r) \qquad \text{as } T \to \infty \text{ for all } i,$$

where $C_i(1)$ *is a* $((1+k+r)\times(1+k+r))$*-matrix of conditional long-run standard deviations.*

Proof: see Lemma 3 in Phillips and Moon (1999).

Lemma 2 *Define the* $(1+k+r)$*-vector* $W_i(r)$ *of standard Brownian motions. The expected value of the cross-product matrix for the demeaned –* $W_i^*(r) = W_i(r) - \int_0^1 W_i(s)\,ds$ *– and detrended –* $W_i^*(r) = W_i(r) - (4-6r)\int_0^1 W_i(s)\,ds - (-6+12r)\int_0^1 sW_i(s)\,ds$ *– vectors of Brownian motions is given by:*
a) Demeaned Brownian motions:

$$E\left(W^*(r)W^*(s)\right) = (r\wedge s) - \frac{2r-r^2}{2} - \frac{2s+s^2}{2} + \frac{1}{3}.$$

b) Detrended Brownian motions:

$$\begin{aligned} E\left(W^*(r)W^*(s)\right) &= (r\wedge s) + 2r^3s - r^3 - 3r^2s + 2r^2 + 2rs^3 - 3rs^2 \\ &\quad + \frac{6}{5}rs - \frac{11}{10}r - s^3 + 2s^2 - \frac{11}{10}s + \frac{2}{15}. \end{aligned}$$

Proof. Constant term. Define a vector of demeaned Brownian motions $W_i^*(r) = W_i(r) - \int_0^1 W_i(s)\,ds$ for which we want to compute $E\left(W_i^*(r)W_i^{*\prime}(s)\right)$. Since the Brownian motions are independent across i, consider the cross-product for one element of the vector, and remove the subscript to simplify notation:

$$\begin{aligned} E\left(W^*(r)W^*(s)\right) &= E\left[\left(W(r) - \int_0^1 W(u)\,du\right)\left(W(s) - \int_0^1 W(u)\,du\right)\right] \\ &= E\left[W(r)W(s) - W(r)\int_0^1 W(u)\,du - W(s)\int_0^1 W(u)\,du + \left(\int_0^1 W(u)\,du\right)^2\right] \\ &= E\left[A1 - A2 - A3 + A4\right]. \end{aligned}$$

The expected value of $A1$ is:

$$E(A1) = E\left(W(r)W(s)\right) = (r\wedge s).$$

For the second element we have:

$$\begin{aligned} E(A2) &= E\left[W(r)\int_0^1 W(u)\,du\right] \\ &= \int_0^r E(W(r)W(u))\,du + \int_r^1 E(W(r)W(u))\,du \\ &= \left(\int_0^r u du + \int_r^1 r du\right) = \left(\frac{r^2}{2} + r(1-r)\right). \end{aligned}$$

The computation of the expected value for the third element is similar:

$$\begin{aligned} E(A3) &= E\left[W(s)\int_0^1 W(u)\,du\right] \\ &= \int_0^s E(W(s)W(u))\,du + \int_s^1 E(W(s)W(u))\,du \\ &= \left(\int_0^s u du + \int_s^1 s du\right) = \left(\frac{s^2}{2} + s(1-s)\right). \end{aligned}$$

Finally, for the fourth element:

$$\begin{aligned} E(A4) &= E\left[\left(\int_0^1 W(u)\,du\right)^2\right] \\ &= E\left[\int_0^1\int_0^1 W(u)W(v)\,dudv\right] = 2\int_0^1\int_0^u E(W(u)W(v))\,dudv \\ &= 2\int_0^1\int_0^u v\,dudv = \frac{1}{3}. \end{aligned}$$

Taken all these elements together, we obtain:

$$E(W^*(r)W^*(s)) = (r\wedge s) - \frac{2r-r^2}{2} - \frac{2s+s^2}{2} + \frac{1}{3}.$$

Time trend case. Define a vector of detrended Brownian motions $W_i^*(r) = W_i(r) - (4-6r)\int_0^1 W_i(s)\,ds - (-6+12r)\int_0^1 sW_i(s)\,ds$ for which we want to compute $E(W_i^*(r)W_i^{*\prime}(s))$. Since the Brownian motions are independent across i, consider the cross-product for one element of the vector, and remove the subscript to simplify notation:

$$\begin{aligned} E(W^*(r)W^*(s)) &= E\Big[(W(r) - (4-6r)\int_0^1 W(u)\,du - (-6+12r)\int_0^1 uW(u)\,du) \\ &\quad (W(s) - (4-6s)\int_0^1 W(u)\,du - (-6+12s)\int_0^1 uW(u)\,du)\Big] \end{aligned}$$

so that

$$\begin{aligned}
E\left(W^{*}\left(r\right)W^{*}\left(s\right)\right) &= E\left[W\left(r\right)W\left(s\right)-\left(4-6s\right)W\left(r\right)\int_{0}^{1}W\left(u\right)du-W\left(r\right)\left(-6+12s\right)\int_{0}^{1}uW\left(u\right)du\right. \\
&\quad -\left(4-6r\right)W\left(s\right)\int_{0}^{1}W\left(u\right)du+\left(4-6r\right)\left(4-6s\right)\left(\int_{0}^{1}W\left(u\right)du\right)^{2} \\
&\quad +\left(4-6r\right)\left(-6+12s\right)\int_{0}^{1}W\left(u\right)du\int_{0}^{1}uW\left(u\right)du \\
&\quad -\left(-6+12r\right)W\left(s\right)\int_{0}^{1}uW\left(u\right)du \\
&\quad +\left(-6+12r\right)\left(4-6s\right)\int_{0}^{1}W\left(u\right)du\int_{0}^{1}uW\left(u\right)du \\
&\quad \left.+\left(-6+12r\right)\left(-6+12s\right)\left(\int_{0}^{1}uW\left(u\right)du\right)^{2}\right] \\
&= E\left[B1-B2-B3-B4+B5+B6-B7+B8+B9\right].
\end{aligned}$$

Let us focus on the expected value of each of these nine elements. For the first element we have:

$$E\left(B1\right)=E\left(W\left(r\right)W\left(s\right)\right)=\left(r\wedge s\right).$$

The second element:

$$\begin{aligned}
E\left(B2\right) &= E\left[\left(4-6s\right)W\left(r\right)\int_{0}^{1}W\left(u\right)du\right] \\
&= \left(4-6s\right)\left[\int_{0}^{r}E\left(W\left(r\right)W\left(u\right)\right)du+\int_{r}^{1}E\left(W\left(r\right)W\left(u\right)\right)du\right] \\
&= \left(4-6s\right)\left(\int_{0}^{r}udu+\int_{r}^{1}rdu\right)=\left(4-6s\right)\left(\frac{r^{2}}{2}+r\left(1-r\right)\right).
\end{aligned}$$

The third element:

$$\begin{aligned}
E\left(B3\right) &= E\left[\left(-6+12s\right)W\left(r\right)\int_{0}^{1}uW\left(u\right)du\right] \\
&= \left(-6+12s\right)\left[\int_{0}^{r}uE\left(W\left(r\right)W\left(u\right)\right)du+\int_{r}^{1}uE\left(W\left(r\right)W\left(u\right)\right)du\right] \\
&= \left(-6+12s\right)\left(\frac{r^{3}}{3}+r\left(\frac{1}{2}-\frac{r^{2}}{2}\right)\right).
\end{aligned}$$

The fourth element – note that is similar to $B2$:

$$\begin{aligned}
E\left(B4\right) &= E\left[\left(4-6r\right)W\left(s\right)\int_{0}^{1}W\left(u\right)du\right] \\
&= \left(4-6r\right)\left[\int_{0}^{s}E\left(W\left(s\right)W\left(u\right)\right)du+\int_{s}^{1}E\left(W\left(s\right)W\left(u\right)\right)du\right] \\
&= \left(4-6r\right)\left(\int_{0}^{s}udu+\int_{s}^{1}sdu\right)=\left(4-6r\right)\left(\frac{s^{2}}{2}+s\left(1-s\right)\right).
\end{aligned}$$

The fifth element:

$$\begin{aligned}
E(B5) &= E\left[(4-6r)(4-6s)\left(\int_0^1 W(u)\,du\right)^2\right] \\
&= (4-6r)(4-6s)\,E\left[\left(\int_0^1 W(u)\,du\right)^2\right] \\
&= (4-6r)(4-6s)\,E\left[\left(\int_0^1\int_0^1 W(u)\,W(v)\,dudv\right)\right] \\
&= (4-6r)(4-6s)\left[2\left(\int_0^1\int_0^u E(W(u)\,W(v))\,dudv\right)\right] \\
&= (4-6r)(4-6s)\left[2\left(\int_0^1\int_0^u v\,dudv\right)\right] = \frac{1}{3}(4-6r)(4-6s).
\end{aligned}$$

The sixth element:

$$\begin{aligned}
E(B6) &= E\left[(4-6r)(-6+12s)\int_0^1 W(u)\,du\int_0^1 uW(u)\,du\right] \\
&= (4-6r)(-6+12s)\,E\left[\int_0^1 W(u)\,du\int_0^1 uW(u)\,du\right] \\
&= (4-6r)(-6+12s)\,E\left[\int_0^1\int_0^1 vW(u)\,W(v)\,dudv\right] \\
&= (4-6r)(-6+12s)\left[\int_0^1\left(\int_0^u vE(W(u)\,W(v))+\int_u^1 vE(W(u)\,W(v))\right)dvdu\right] \\
&= (4-6r)(-6+12s)\left[\int_0^1\left(\int_0^u v^2+\int_u^1 uv\right)dvdu\right] = \frac{5}{24}(4-6r)(-6+12s).
\end{aligned}$$

The seventh element – similar to $B3$:

$$\begin{aligned}
E(B7) &= E\left[(-6+12r)\,W(s)\int_0^1 uW(u)\,du\right] \\
&= (-6+12r)\left[\int_0^s uE(W(s)\,W(u))\,du+\int_s^1 uE(W(s)\,W(u))\,du\right] \\
&= (-6+12r)\left(\frac{s^3}{3}+s\left(\frac{1}{2}-\frac{s^2}{2}\right)\right).
\end{aligned}$$

The eight element – similar to $B6$:

$$\begin{aligned}
E(B8) &= E\left[(4-6s)(-6+12r)\int_0^1 W(u)\,du\int_0^1 uW(u)\,du\right] \\
&= \frac{5}{24}(4-6s)(-6+12r).
\end{aligned}$$

Finally, the ninth element:

$$\begin{aligned}
E(B9) &= E\left[(-6+12r)(-6+12s)\left(\int_0^1 uW(u)\,du\right)^2\right] \\
&= (-6+12r)(-6+12s)\,E\left[\left(\int_0^1 uW(u)\,du\right)^2\right] \\
&= (-6+12r)(-6+12s)\,E\left[\int_0^1\int_0^1 uvW(u)\,W(v)\,dudv\right] \\
&= (-6+12r)(-6+12s)\left[2\int_0^1\int_0^u uvE(W(u)\,W(v))\,dudv\right] \\
&= (-6+12r)(-6+12s)\left[2\int_0^1\int_0^u uv^2 dudv\right] = \frac{2}{15}(-6+12r)(-6+12s).
\end{aligned}$$

Taking all these elements together, we obtain:

$$\begin{aligned}
E(W^*(r)\,W^*(s)) &= (r\wedge s) - (4-6s)\left(\frac{r^2}{2}+r(1-r)\right) - (-6+12s)\left(\frac{r^3}{3}+r\left(\frac{1}{2}-\frac{r^2}{2}\right)\right) \\
&\quad -(4-6r)\left(\frac{s^2}{2}+s(1-s)\right)+\frac{1}{3}(4-6r)(4-6s) \\
&\quad +\frac{5}{24}(4-6r)(-6+12s) \\
&\quad -(-6+12r)\left(\frac{s^3}{3}+s\left(\frac{1}{2}-\frac{s^2}{2}\right)\right) \\
&\quad +\frac{5}{24}(4-6s)(-6+12r) \\
&\quad +\frac{2}{15}(-6+12r)(-6+12s) \\
&= (r\wedge s)+2r^3s-r^3-3r^2s+2r^2+2rs^3-3rs^2 \\
&\quad +\frac{6}{5}rs-\frac{11}{10}r-s^3+2s^2-\frac{11}{10}s+\frac{2}{15}
\end{aligned}$$

■

A.1 Proof of Theorem 1

A.1.1 No deterministic component

In this section we analyze the model specification that does not include any deterministic component, i.e., $D_{i,t}=0\ \forall i$ in (1). For ease of exposition, we start considering that all common factors in the model are I(1), but the derivations also apply if there is a mixture of I(0) and I(1) common factors, or all common factors are I(0) – see below.

Let $M_i(r) = (M_{y_i}(r)', M_{x_i}(r)', M_F(r)')' = (M_{U_i}(r)', M_F(r)')' = C_i(1)\,W_i(r)$, where $M_i(r)$ is a randomly scaled Brownian motion with a conditional covariance matrix $C_i(1)\,C_i(1)'$ that has a well defined expectation provided that $\left\|EC_i(1)\,C_i(1)'\right\| < \infty$ as shown in Lemma 1(d) in Phillips and Moon (1999). Let us define $Z_{i,t} = \left(U_{i,t}', F_t'\right)'$, by the continuous mapping

theorem we have that as $T \to \infty$ for a fixed N

$$T^{-2}\sum_{t=1}^{T} Z_{i,t}Z_{i,t}' \Rightarrow C_i(1)\int_0^1 W_i(r)\,W_i'(r)\,dr\;C_i(1)' = \int_0^1 M_i(r)\,M_i'(r)\,dr.$$

Further, we define the long-run conditional covariance matrix of $Z_{i,t} = \left(U_{i,t}', F_t'\right)' = (U_{y_i,t}, U_{x_{i,1},t}, \ldots, U_{x_{i,k},t}, F_t')'$ as

$$\begin{aligned}\Omega_i &= \begin{bmatrix} \Omega_{U_{y_i}U_{y_i}} & \Omega_{U_{y_i}U_{x_i}} & \Omega_{U_{y_i}F} \\ \Omega_{U_{x_i}U_{y_i}} & \Omega_{U_{x_i}U_{x_i}} & \Omega_{U_{x_i}F} \\ \Omega_{FU_{y_i}} & \Omega_{FU_{x_i}} & \Omega_{FF} \end{bmatrix} = \begin{bmatrix} \Omega_{U_i} & \Omega_{U_iF} \\ \Omega_{FU_i} & \Omega_{FF} \end{bmatrix} \\ &= C_i(1)\,C_i(1)' = \begin{bmatrix} C_{U_{y_i}}(1)\,C_{U_{y_i}}(1)' & C_{U_{y_i}}(1)\,C_{x_i}(1)' & C_{U_{y_i}}(1)\,C_F(1)' \\ C_{U_{x_i}}(1)\,C_{U_{y_i}}(1)' & C_{U_{x_i}}(1)\,C_{U_{x_i}}(1)' & C_{U_{x_i}}(1)\,C_F(1)' \\ C_F(1)\,C_{U_{y_i}}(1)' & C_F(1)\,C_{U_{x_i}}(1)' & C_F(1)\,C_F(1)' \end{bmatrix},\end{aligned}$$

with $C_i(1) = (C_{y_i}(1)', C_{x_i}(1)', C_F(1)')' = (C_{U_i}(1)', C_F(1)')'$ and the long-run average covariance matrix of $Z_{i,t}$ as:

$$\Omega = \begin{bmatrix} \Omega_{U_yU_y} & \Omega_{U_yU_x} & \Omega_{U_yF} \\ \Omega_{U_xU_y} & \Omega_{U_xU_x} & \Omega_{U_xF} \\ \Omega_{FU_y} & \Omega_{FU_x} & \Omega_{FF} \end{bmatrix} = \begin{bmatrix} \Omega_U & \Omega_{UF} \\ \Omega_{FU} & \Omega_{FF} \end{bmatrix} = E\left(C_i(1)\,C_i(1)'\right).$$

Let K be the invariant σ-field generated by F_t, so that $U_{i,t}$ are independent across i conditional on K. Then, we define the expected value of the cross product matrix as

$$\begin{aligned} E\left(\int_0^1 M_i(r)\,M_i'(r)\,dr\right) &= E\left(C_i(1)\,E\left(\int_0^1 W_i(r)\,W_i'(r)\,dr\right)C_i(1)'\right) \\ &= E\left(C_i(1)\,\frac{1}{2}I_{(1+k+r)}C_i(1)'\right) \\ &= \frac{1}{2}\Omega. \end{aligned}$$

Note that averaging across i the cross-products involving $U_{i,t}$ we have, conditional on K,

$$N^{-1}\sum_{i=1}^{N} T^{-2}\sum_{t=1}^{T} U_{i,t}U_{i,t}' \Rightarrow N^{-1}\sum_{i=1}^{N}\left(\int_0^1 M_{U_i}(r)\,M_{U_i}'(r)\,dr\,\middle|\,K\right).$$

Using Lemma 4 in Phillips and Moon (1999) and Theorem 9 in Kao, Trapani and Urga (2011), we have that $E\left\|\int_0^1 M_{U_i}(r)\,M_{U_i}'(r)\,dr\,\middle|\,K\right\|^2 < \infty$, so that as $N \to \infty$ the law of strong numbers gives

$$N^{-1}\sum_{i=1}^{N}\left(\int_0^1 M_{U_i}(r)\,M_{U_i}'(r)\,dr\,\middle|\,K\right) \overset{a.s.}{\to} E\left(\int_0^1 M_{U_i}(r)\,M_{U_i}'(r)\,dr\,\middle|\,K\right) = \frac{1}{2}\Omega_U.$$

We have defined the pooled estimator as

$$\hat{\beta} = \left[\frac{1}{N}\sum_{i=1}^{N} T^{-2}\left(x_i^{*\prime}x_i^{*}\right)\right]^{-1} \frac{1}{N}\sum_{i=1}^{N} T^{-2}\left(x_i^{*\prime}y_i^{*}\right),$$

where, in this case, $x_i^* = M_F x_i$. Note that

$$\begin{aligned} T^{-2}x_i^{*\prime}x_i^{*} &= T^{-2}x_i^{\prime}M_F x_i = T^{-2}U_{x_i}^{\prime}M_F U_{x_i} \\ &= T^{-2}U_{x_i}^{\prime}U_{x_i} - T^{-2}U_{x_i}^{\prime}F\left(T^{-2}F^{\prime}F\right)^{-1}T^{-2}F^{\prime}U_{x_i}, \end{aligned}$$

so that, in the limit,

$$\begin{aligned} T^{-2}x_i^{*\prime}x_i^{*} \Rightarrow\ & C_{U_{x_i}}(1)\left(\int_0^1 W_{U_{x_i}}(r)\,W_{U_{x_i}}^{\prime}(r)\,dr\middle| K\right)C_{U_{x_i}}(1)^{\prime} \\ & -\left[\left(C_{U_{x_i}}(1)\left(\int_0^1 W_{U_{x_i}}(r)\,W_F^{\prime}(r)\,dr\middle| K\right)C_F(1)^{\prime}\right)\right. \\ & \left(C_F(1)\left(\int_0^1 W_F(r)\,W_F^{\prime}(r)\,dr\middle| K\right)C_F(1)^{\prime}\right)^{-1} \\ & \left.\left(C_F(1)\left(\int_0^1 W_F(r)\,W_{U_{x_i}}^{\prime}(r)\,dr\middle| K\right)C_{U_{x_i}}(1)^{\prime}\right)\right]. \end{aligned}$$

Using the fact that $E\left(W_{U_i}(r)\,W_{U_i}^{\prime}(s)\right) = (r\wedge s)\,I_{1+k}$, with $(r\wedge s) = \min\{r,s\}$, we have, conditional on K,

$$\begin{aligned} E\left(T^{-2}x_i^{*\prime}x_i^{*}\right) \quad & \overset{a.s.}{\to} E\left[C_{U_{x_i}}(1)\left(\int_0^1 W_{U_{x_i}}(r)\,W_{U_{x_i}}^{\prime}(r)\,dr\middle| K\right)C_{U_{x_i}}(1)^{\prime}\right] \\ & -E\left[\int_0^1\int_0^1 C_{U_{x_i}}(1)\,W_{U_{x_i}}(r)\,h(r,s)\,W_{U_{x_i}}^{\prime}(s)\,C_{U_{x_i}}(1)^{\prime}\,drds\middle| K\right], \end{aligned}$$

with $h(r,s) = W_F^{\prime}(r)\,C_F(1)^{\prime}\left(C_F(1)\left(\int_0^1 W_F(r)\,W_F^{\prime}(r)\,dr\right)C_F(1)^{\prime}\right)^{-1}C_F(1)\,W_F(s)$. Note that

$$E\left[C_{U_{x_i}}(1)\left(\int_0^1 W_{U_{x_i}}(r)\,W_{U_{x_i}}^{\prime}(r)\,dr\middle| K\right)C_{U_{x_i}}(1)^{\prime}\right] = \left(\int_0^1 r dr\right)\Omega_{U_xU_x} = \frac{1}{2}\Omega_{U_xU_x},$$

and

$$E\left[\int_0^1\int_0^1 C_{U_{x_i}}(1)\,W_{U_{x_i}}(r)\,h(r,s)\,W_{U_{x_i}}^{\prime}(s)\,C_{U_{x_i}}(1)^{\prime}\,drds\middle| K\right] = \left(\int_0^1\int_0^1 (r\wedge s)\,h(r,s)\,dsdr\right)\Omega_{U_xU_x},$$

so that, conditional on K, we have

$$N^{-1}\sum_{i=1}^{N} T^{-2}x_i^{*\prime}x_i^{*} \overset{a.s.}{\to} \left(\frac{1}{2} - \int_0^1\int_0^1 (r\wedge s)\,h(r,s)\,dsdr\right)\Omega_{U_xU_x}.$$

Similarly, for the numerator of the pooled estimator

$$\begin{aligned} T^{-2}x_i^{*\prime}y_i^* &= T^{-2}x_i'M_Fy_i = T^{-2}U_{x_i}'M_FU_{y_i} \\ &= T^{-2}U_{x_i}'U_{y_i} - T^{-2}U_{x_i}'F\left(T^{-2}F'F\right)^{-1}T^{-2}F'U_{y_i}, \end{aligned}$$

so that, conditional on K,

$$N^{-1}\sum_{i=1}^{N}T^{-2}x_i^{*\prime}y_i^* \overset{a.s.}{\to} \left(\frac{1}{2} - \int_0^1\int_0^1 (r\wedge s)\,h(r,s)\,dsdr\right)\Omega_{U_xU_y}.$$

Finally,

$$\begin{aligned} \hat{\beta} \overset{p}{\to} & \left[\left(\frac{1}{2} - \int_0^1\int_0^1 (r\wedge s)\,h(r,s)\,dsdr\right)\Omega_{U_xU_x}\right]^{-1} \\ & \left[\left(\frac{1}{2} - \int_0^1\int_0^1 (r\wedge s)\,h(r,s)\,dsdr\right)\Omega_{U_xU_y}\right] \\ = & \ \Omega_{U_xU_x}^{-1}\Omega_{U_xU_y} = \beta. \end{aligned}$$

Note that the conditioning variables that appear in the numerator and denominator of the estimator cancel out so that the conditional limit of the estimator is also the unconditional limit.

So far, the proof has used sequential limits to show the consistency of the pooled estimator. However and following Phillips and Moon (1999), the same result is achieved if we base our analysis on joint limit theory. By the Beveridge-Nelson (BN) decomposition:

$$Z_{i,t} \overset{a.s.}{=} C_i(1)\,P_{i,t} + \tilde{V}_{i,0} - \tilde{V}_{i,t} + Z_{i,0},$$

with $P_{i,t} = \sum_{t=1}^{[Tr]} S_{i,t}$, $S_{i,t} = \left(w_t', \varepsilon_{i,t}'\right)'$. Then, define

$$N^{-1}\sum_{i=1}^{N}T^{-2}\sum_{t=1}^{T}Z_{i,t}Z_{i,t}' \overset{a.s.}{=} N^{-1}\sum_{i=1}^{N}\left(Q_{i,t} + R_{i,t}\right),$$

where

$$\begin{aligned} Q_{i,t} &= T^{-2}\sum_{t=1}^{T}C_i(1)\,P_{i,t}P_{i,t}'C_i(1) \\ R_{i,t} &= R_{1,i,t} + R_{1,i,t}' + R_{2,i,t} \\ R_{1,i,t} &= T^{-2}\sum_{t=1}^{T}C_i(1)\,P_{i,t}\left(\tilde{V}_{i,0} - \tilde{V}_{i,t} + Z_{i,0}\right)' \\ R_{2,i,t} &= T^{-2}\sum_{t=1}^{T}\left(\tilde{V}_{i,0} - \tilde{V}_{i,t} + Z_{i,0}\right)\left(\tilde{V}_{i,0} - \tilde{V}_{i,t} + Z_{i,0}\right)'. \end{aligned}$$

We need to show that $\|Q_{i,t}\|$ is uniformly integrable in T, provided that then

$$N^{-1}\sum_{i=1}^{N} Q_{i,t} \xrightarrow{p} \frac{1}{2}\Omega,$$

as $(T,N) \to \infty$ jointly. By $\|AB\| \leq \|A\| \|B\|$ and the triangle inequality

$$\|Q_{i,t}\| \leq \|C_i(1)\|^2 T^{-2} \sum_{t=1}^{T} \|P_{i,t}\|^2.$$

Note that as $T \to \infty$

$$T^{-2}\sum_{t=1}^{T} \|P_{i,t}\|^2 \Rightarrow \int_0^1 \|W_{i,t}\|^2 dr,$$

and that

$$E\left(T^{-2}\sum_{t=1}^{T} \|P_{i,t}\|^2\right) = tr\left(T^{-2}\sum_{t=1}^{T} E\left(P_{i,t}, P_{i,t}'\right)\right) \to E\left(\int_0^1 \|W_{i,t}\|^2 dr\right) = \frac{1}{2} tr\left(I_{1+k+r}\right).$$

Then, it follows from Billingsley (1968) that $T^{-2}\sum_{t=1}^{T} \|P_{i,t}\|^2$ is uniformly integrable in T. Since $E\|C_i(1)\|^2 < \infty$, we can conclude that $\|C_i(1)\|^2 T^{-2}\sum_{t=1}^{T} \|P_{i,t}\|^2$ is uniformly integrable in T and, hence, $\|Q_{i,t}\|$ is uniformly integrable in T. Consequently, $N^{-1}\sum_{i=1}^{N} Q_{i,t} \xrightarrow{p} \frac{1}{2}\Omega$ as stated above.

So far, we have assumed that all r common factors are I(1), but it would be the case that there is a subset of r_0 I(0) common factors and a subset of r_1 I(1) common factors, $r = r_0 + r_1$. Let us define $F_t = \left(F_{1,t}', F_{0,t}'\right)'$, with $F_{1,t}$ the $(r_1 \times 1)$-vector of I(1) common factors and $F_{0,t}$ the $(r_0 \times 1)$-vector of I(0) common factors. In this case,

$$\begin{aligned} T^{-2} x_i^{*\prime} x_i^* &= T^{-2} U_{x_i}' M_F U_{x_i} \\ &= T^{-2} U_{x_i}' U_{x_i} - T^{-2} U_{x_i}' F\Psi \left(\Psi F' F \Psi\right)^{-1} \Psi F' U_{x_i}, \end{aligned}$$

with $\Psi = diag\left(\Psi_1, \Psi_0\right)$ a rescaling diagonal matrix defined by the $(r_1 \times 1)$-vector $\Psi_1 = (T^{-1}, \ldots, T^{-1})$ and by the $(r_0 \times 1)$-vector $\Psi_0 = \left(T^{-1/2}, \ldots, T^{-1/2}\right)$, so that $\Psi F' F \Psi = O_p(1)$. Using these elements, we have

$$N^{-1}\sum_{i=1}^{N} T^{-2} x_i^{*\prime} x_i^* \xrightarrow{a.s.} \left(\frac{1}{2} - \int_0^1 \int_0^1 (r \wedge s) h(r,s) dsdr\right) \Omega_{U_x U_x}.$$

The same applies for $N^{-1}\sum_{i=1}^{N} T^{-2} x_i^{*\prime} y_i^* \xrightarrow{a.s.} \left(\frac{1}{2} - \int_0^1 \int_0^1 (r \wedge s) h(r,s) dsdr\right) \Omega_{U_x U_y}$, so that $\hat{\beta} \xrightarrow{p} \Omega_{U_x U_x}^{-1} \Omega_{U_x U_y} = \beta$ as above. Consequently, having a combination of I(0) and I(1) common factors does not alter the result about the consistency of the pooled estimator.

A.1.2 Constant term

In this section we consider the deterministic specification given by Model 1 through the definition of $D_{i,t} = \mu_i = \left(\mu_{i,0}, \mu_{i,1}, \ldots, \mu_{i,k}\right)'$. Using the projection matrix $M_D = I - D\left(D'D\right)^{-1} D'$, where

$D = \iota$ denotes a vector of ones, we define $\tilde{Z}_{i,t} = \left(\tilde{U}'_{i,t}, \tilde{F}'_t\right)'$, where $\tilde{U}_i = M_D U_i$ and $\tilde{F} = M_D F$ are the OLS detrended variables. By the continuous mapping theorem we have that as $T \to \infty$ for a fixed N

$$T^{-2}\sum_{t=1}^{T}\tilde{Z}_{i,t}\tilde{Z}'_{i,t} \Rightarrow C_i(1)\int_0^1 W_i^*(r)\,W_i^{*\prime}(r)\,dr\;C_i(1)' = \int_0^1 M_i^*(r)\,M_i^{*\prime}(r)\,dr,$$

where $W_i^*(r) = W_i(r) - \int_0^1 W_i(s)\,ds$ and $M_i^*(r) = M_i(r) - \int_0^1 M_i(s)\,ds$ are demeaned Brownian motion vectors. As above,

$$\begin{aligned} E\left(\int_0^1 M_i^*(r)\,M_i^{*\prime}(r)\,dr\right) &= E\left(C_i(1)\,E\left(\int_0^1 W_i^*(r)\,W_i^{*\prime}(r)\,dr\right)C_i(1)'\right) \\ &= E\left(C_i(1)\,\frac{1}{6}I_{(1+k+r)}C_i(1)'\right) \\ &= \frac{1}{6}\Omega. \end{aligned}$$

The developments carried out in the previous section follow here replacing $W_i(r)$ by $W_i^*(r)$. Note that now

$$\begin{aligned} T^{-2}x_i^{*\prime}x_i^* &= T^{-2}\tilde{x}_i'M_{\tilde{F}}\tilde{x}_i = T^{-2}\tilde{U}'_{x_i}M_{\tilde{F}}\tilde{U}_{x_i} \\ &= T^{-2}\tilde{U}'_{x_i}\tilde{U}_{x_i} - T^{-2}\tilde{U}'_{x_i}\tilde{F}\left(T^{-2}\tilde{F}'\tilde{F}\right)^{-1}T^{-2}\tilde{F}'\tilde{U}_{x_i}, \end{aligned}$$

so that, in the limit,

$$\begin{aligned} T^{-2}x_i^{*\prime}x_i^* \;\Rightarrow\; & C_{U_{x_i}}(1)\left(\int_0^1 W_{U_{x_i}}^*(r)\,W_{U_{x_i}}^{*\prime}(r)\,dr\middle|K\right)C_{U_{x_i}}(1)' \\ & -\left[\left(C_{U_{x_i}}(1)\left(\int_0^1 W_{U_{x_i}}^*(r)\,W_F^{*\prime}(r)\,dr\middle|K\right)C_F(1)'\right)\right. \\ & \left(C_F(1)\left(\int_0^1 W_F^*(r)\,W_F^{*\prime}(r)\,dr\middle|K\right)C_F(1)'\right)^{-1} \\ & \left.\left(C_F(1)\left(\int_0^1 W_F^*(r)\,W_{U_{x_i}}^{*\prime}(r)\,dr\middle|K\right)C_{U_{x_i}}(1)'\right)\right]. \end{aligned}$$

From Lemma 2, $E\left(W_{U_i}^*(r)\,W_{U_i}^{*\prime}(s)\right) = ((r\wedge s) - \left(2r - r^2\right)/2 - \left(2s - s^2\right)/2 + 1/3)I_{1+k}$, so that, conditional on K, we obtain $N^{-1}\sum_{i=1}^N T^{-2}x_i^{*\prime}x_i^* \xrightarrow{a.s.} (1/6 - \int_0^1\int_0^1((r\wedge s) - \left(2r - r^2\right)/2 - \left(2s - s^2\right)/2 + 1/3)h(r,s)\,dsdr)\Omega_{U_xU_x}$ and $N^{-1}\sum_{i=1}^N T^{-2}x_i^{*\prime}y_i^* \xrightarrow{a.s.} (1/6 - \int_0^1\int_0^1((r\wedge s) - \left(2r - r^2\right)/2 - \left(2s - s^2\right)/2 + 1/3)h(r,s)\,dsdr)\Omega_{U_xU_y}$, where now $h(r,s) = W_F^{*\prime}(r)\,C_F(1)'(C_F(1)(\int_0^1 W_F^*(r)\,W_F^{*\prime}(r)\,dr)C_F(1)')^{-1}C_F(1)\,W_F^*(s)$ with $W_F^*(r) = W_F(r) - \int_0^1 W_F(s)\,ds$. Therefore, $\hat{\beta} \xrightarrow{p} \Omega_{U_xU_x}^{-1}\Omega_{U_xU_y} = \beta$, as above. Following the steps given in the previous subsection, it can be shown that the same result is obtained if we use joint limits, where the only difference is that we use demeaned Brownian motions instead of standard Brownian motions – to be specific, we need to consider that in this case $N^{-1}\sum_{i=1}^N Q_{i,t} \xrightarrow{p} \frac{1}{6}\Omega$ and the rest of the proof applies. As above, note that the conditioning variables that appear in the numerator and denominator of the estimator cancel out so that the conditional limit of the estimator is also the unconditional

limit.

A.1.3 Time trend

In this section we consider the deterministic specification given by Model 2, i.e., $D_{i,t} = (1, t)\,[\delta_{i,0}$, $\delta_{i,1}$, ..., $\delta_{i,k}]$, with $\delta_{i,j} = \left(\mu_{i,j}, \eta_{i,j}\right)'$, $j = 0, 1, \ldots, k$. Using the projection matrix $M_D = I - D\left(D'D\right)^{-1} D'$, where $D = [\iota\ \tau]$ with ι a vector of ones and $\tau = (1, 2, \ldots, T)'$. We define $\tilde{Z}_{i,t} = \left(\tilde{U}'_{i,t}, \tilde{F}'_t\right)'$, where $\tilde{U}_i = M_D U_i$ and $\tilde{F} = M_D F$ are the OLS detrended variables. By the continuous mapping theorem we have that as $T \to \infty$ for a fixed N

$$T^{-2}\sum_{t=1}^{T}\tilde{Z}_{i,t}\tilde{Z}'_{i,t} \Rightarrow C_i(1)\int_0^1 W_i^*(r)\,W_i^{*\prime}(r)\,dr\ C_i(1)' = \int_0^1 M_i^*(r)\,M_i^{*\prime}(r)\,dr,$$

where $W_i^*(r) = W_i(r) - (4-6r)\int_0^1 W_i(s)\,ds - (-6+12r)\int_0^1 sW_i(s)\,ds$ and $M_i^*(r) = M_i(r) - (4-6r)\int_0^1 M_i(s)\,ds - (-6+12r)\int_0^1 sM_i(s)\,ds$ are detrended Brownian motion vectors. In this case,

$$\begin{aligned} E\left(\int_0^1 M_i^*(r)\,M_i^{*\prime}(r)\,dr\right) &= E\left(C_i(1)\,E\left(\int_0^1 W_i^*(r)\,W_i^{*\prime}(r)\,dr\right)C_i(1)'\right) \\ &= E\left(C_i(1)\,\frac{1}{15}I_{(1+k+r)}C_i(1)'\right) \\ &= \frac{1}{15}\Omega. \end{aligned}$$

From Lemma 2, $E\left(W_{U_i}^*(r)\,W_{U_i}^{*\prime}(s)\right) = ((r\wedge s) + 2r^3s - r^3 - 3r^2s + 2r^2 + 2rs^3 - 3rs^2 + \frac{6}{5}rs - \frac{11}{10}r - s^3 + 2s^2 - \frac{11}{10}s + \frac{2}{15})I_{1+k}$, so that, conditional on K, $N^{-1}\sum_{i=1}^N T^{-2}x_i^{*\prime}x_i^* \xrightarrow{a.s.} (\frac{1}{15} - \int_0^1\int_0^1((r\wedge s)+2r^3s-r^3-3r^2s+2r^2+2rs^3-3rs^2+\frac{6}{5}rs-\frac{11}{10}r-s^3+2s^2-\frac{11}{10}s+\frac{2}{15})h(r,s)\,dsdr)\Omega_{U_xU_x}$ and $N^{-1}\sum_{i=1}^N T^{-2}x_i^{*\prime}y_i^* \xrightarrow{a.s.} (\frac{1}{15} - \int_0^1\int_0^1((r\wedge s) + 2r^3s - r^3 - 3r^2s + 2r^2 + 2rs^3 - 3rs^2 + \frac{6}{5}rs - \frac{11}{10}r - s^3 + 2s^2 - \frac{11}{10}s + \frac{2}{15})h(r,s)\,dsdr)\Omega_{U_xU_y}$, with $h(r,s) = W_F^{*\prime}(r)\,C_F(1)'\,(C_F(1)\left(\int_0^1 W_F^*(r)\,W_F^{*\prime}(r)\,dr\right)C_F(1)')^{-1}C_F(1)\,W_F^*(s)$ with $W_F^*(r) = W_F(r) - (4-6r)\int_0^1 W_F(s)\,ds - (-6+12r)\int_0^1 sW_F(s)\,ds$. Consequently, $\hat{\beta} \xrightarrow{p} \Omega_{U_xU_x}^{-1}\Omega_{U_xU_y} = \beta$. Following the steps given above, it can be shown that the same result is obtained if we use joint limits, where the only difference is that we use detrended Brownian motions instead of standard Brownian motions – to be specific, in this case $N^{-1}\sum_{i=1}^N Q_{i,t} \xrightarrow{p} \frac{1}{15}\Omega$ and the rest of the proof applies. As above, note that the conditioning variables that appear in the numerator and denominator of the estimator cancel out so that the conditional limit of the estimator is also the unconditional limit.

A.2 Proof of Theorem 2

In order to prove Theorem 2, we begin by defining the projection matrix $\bar{M} = I - \bar{H}\left(\bar{H}'\bar{H}\right)^{-1}\bar{H}'$, where $\bar{H} = \bar{z}$ for Model 0, $\bar{H} = [\iota\ \bar{z}]$ for Model 1 and $\bar{H} = [\iota\ \tau\ \bar{z}]$ for Model 2, with $\bar{z} = [\bar{x}\ \bar{y}]$ being the $(T \times (k+1))$ matrix of cross-section averages. Further, let us define $M_g = I - G\left(G'G\right)^{-1}G'$ and $M_q = I - Q\left(Q'Q\right)^{+}Q'$.

In the case of Model 0 $G = F$ denotes the $(T \times r)$ matrix of unobserved factors, $Q = G\bar{P}$

with $\bar{P} = \bar{\pi}$, and A^+ is the Moore-Penrose inverse of matrix A. For Model 1, $G = [\iota \; F]$, and

$$\bar{P} = \begin{bmatrix} 1 & \bar{\delta} \\ 0 & \bar{\pi} \end{bmatrix}.$$

Finally, for Model 2 we have $G = [\iota \; \tau \; F]$, and

$$\bar{P} = \begin{bmatrix} I_2 & \bar{\delta} \\ 0 & \bar{\pi} \end{bmatrix}.$$

The pooled estimator is computed as:

$$\begin{aligned} \hat{\beta} &= \left[\frac{1}{N}\sum_{i=1}^{N} T^{-2}\left(x_i^{*\prime} x_i^*\right)\right]^{-1} \frac{1}{N}\sum_{i=1}^{N} T^{-2}\left(x_i^{*\prime} y_i^*\right) \\ &= \left[\frac{1}{N}\sum_{i=1}^{N} T^{-2}\left(x_i^{\prime} \bar{M} x_i\right)\right]^{-1} \frac{1}{N}\sum_{i=1}^{N} T^{-2}\left(x_i^{\prime} \bar{M} y_i\right). \end{aligned}$$

In order to show consistency of the pooled estimator, we need to establish that the quadratic form involving the projection matrix using the cross-section averages is asymptotically equivalent to that defined by using the true factors. That is, we need to analyze

$$T^{-2}\left\|Y_i^{\prime}\bar{M}Y_i - Y_i^{\prime}M_q Y_i\right\|,$$

where

$$Y_i = F\pi_i^{\prime} + U_i$$

denotes the model defined in (1) in matrix notation, assuming no deterministic terms to simplify the notation.

Note that

$$\begin{aligned} T^{-2}\left\|Y_i^{\prime}\bar{M}Y_i - Y_i^{\prime}M_q Y_i\right\| &\leq T^{-2}\left\|\left(Y_i^{\prime}\bar{H} - Y_i^{\prime}Q\right)\left(\bar{H}^{\prime}\bar{H}\right)^{-1}\bar{H}^{\prime}Y_i\right\| \\ &\quad + T^{-2}\left\|Y_i^{\prime}Q\left(\left(\bar{H}^{\prime}\bar{H}\right)^{-1} - \left(Q^{\prime}Q\right)^{-1}\right)\bar{H}^{\prime}Y_i\right\| \\ &\quad + T^{-2}\left\|Y_i^{\prime}Q\left(Q^{\prime}Q\right)^{-1}\left(\bar{H}^{\prime}Y_i - Q^{\prime}Y_i\right)\right\| \\ &= I + II + III. \end{aligned}$$

Consider part I and recall that $\bar{H} = Q + \bar{U}$.

We then have

$$T^{-2}\left\|\left(Y_i^{\prime}\bar{H} - Y_i^{\prime}Q\right)\left(\bar{H}^{\prime}\bar{H}\right)^{-1}\bar{H}^{\prime}Y_i\right\| \leq \left\|\frac{Y_i^{\prime}\bar{U}}{T^2}\right\|\left\|\left(\frac{\bar{H}^{\prime}\bar{H}}{T^2}\right)^{-1}\frac{\bar{H}^{\prime}Y_i}{T^2}\right\|,$$

where

$$\left\|\frac{Y_i^{\prime}\bar{U}}{T^2}\right\| = \left\|\frac{1}{N}\sum_{j=1}^{N}\frac{Y_i^{\prime}U_j}{T^2}\right\| = O_p\left(\frac{1}{\sqrt{N}}\right),$$

given that $T^{-2}Y_i'U_j = O_p(1)$ and $N^{-1/2}\sum_{j=1}^{N} T^{-2}Y_i'U_j = O_p(1)$. Assuming that the rank condition $rank(\bar{\pi}) = r \leq (1+k)$ for all N as $T, N \to \infty$ holds, and provided that $T^{-2}\bar{H}'\bar{H} = O_p(1)$ and $T^{-2}\bar{H}'Y_i = O_p(1)$, we have that $I = O_p\left(N^{-1/2}\right)$.

For part II we have

$$II \leq \left\| \frac{\bar{U}'\bar{U}}{T^2} + \frac{Q'\bar{U}}{T^2} + \frac{\bar{U}'Q}{T^2} \right\| \left\| \frac{Y_i'Q}{T^2} \left(\frac{\bar{H}'\bar{H}}{T^2} \right)^{-1} \right\| \left\| \left(\frac{Q'Q}{T^2} \right)^{-1} \frac{\bar{H}'Y_i}{T^2} \right\|,$$

where $T^{-2}Y_i'Q = O_p(1)$ and $T^{-2}Q'Q = O_p(1)$. Note that

$$\left\| \frac{\bar{U}'\bar{U}}{T^2} \right\| = \left\| \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{U_i'U_j}{T^2} \right\| = O_p\left(\frac{1}{\sqrt{N}} \right),$$

given that $T^{-2}U_i'U_j = O_p(1)$ and $N^{-1/2}\sum_{j=1}^{N} T^{-2}U_i'U_j = O_p(1)$. Similarly, $\left\|T^{-2}Q'\bar{U}\right\| = O_p\left(N^{-1/2}\right)$, so that $II = O_p\left(N^{-1/2}\right)$.

Part III is given by

$$III \leq \left\| \frac{Y_i'Q}{T^2} \left(\frac{Q'Q}{T^2} \right)^{-1} \right\| \left\| \frac{Y_i'\bar{U}}{T^2} \right\|.$$

Using the elements above, it may be easily shown that $III = O_p\left(N^{-1/2}\right)$.

Consequently, when the rank condition $rank(\bar{\pi}) = r \leq (1+k)$ holds for all N as $T, N \to \infty$, $M_g = M_q$, so that, conditional on K, we have

$$T^{-2}\left\|Y_i'\bar{M}Y_i - Y_i'M_gY_i\right\| \leq O_p\left(\frac{1}{\sqrt{N}} \right),$$

uniformly over i. Therefore, conditional on K, we have

$$\hat{\beta} = \left[\frac{1}{N} \sum_{i=1}^{N} T^{-2}\left(x_i'M_gx_i\right) + O_p\left(\frac{1}{\sqrt{N}} \right) \right]^{-1} \left[\frac{1}{N} \sum_{i=1}^{N} T^{-2}\left(x_i'M_gy_i\right) + O_p\left(\frac{1}{\sqrt{N}} \right) \right],$$

so that as $T, N \to \infty$

$$\hat{\beta} \xrightarrow{p} \beta = \Omega_{U_xU_x}^{-1}\Omega_{U_xU_y},$$

a result that was already established in Theorem 1 and confirmed in the Monte Carlo experiment that has been conducted above.

Table A.1: Mean, median and root mean square error of the $\hat{\beta}_{PCCE}$

			$N = 10$			$N = 20$			$N = 50$			$N = 100$		
ϕ_i	ρ	T	Mean	Median	RMSE	Mean	Median	RMSE	Mean	Median	RMSE	Mean	Median	RMSE
1	1	50	0.996	0.996	0.144	0.999	0.999	0.092	1.000	1.000	0.059	1.000	1.000	0.041
1	1	100	1.000	0.998	0.140	1.000	1.000	0.092	1.002	1.001	0.059	1.000	0.999	0.040
1	1	250	1.001	1.001	0.146	0.999	1.001	0.092	0.999	0.999	0.059	1.001	1.001	0.042
1	0.99	50	0.996	0.996	0.144	0.999	1.000	0.092	1.000	1.000	0.059	1.000	1.000	0.041
1	0.99	100	1.000	0.997	0.140	1.000	1.001	0.092	1.002	1.001	0.059	1.000	0.999	0.040
1	0.99	250	1.000	1.000	0.147	0.999	1.000	0.093	0.999	1.000	0.059	1.001	1.001	0.042
1	0.95	50	0.996	0.995	0.145	0.999	0.999	0.092	1.000	1.000	0.060	1.000	1.000	0.041
1	0.95	100	1.000	0.998	0.143	1.000	1.001	0.094	1.002	1.003	0.059	1.000	0.999	0.041
1	0.95	250	1.000	1.003	0.148	0.999	1.001	0.095	0.999	1.000	0.060	1.001	1.001	0.043
1	0.9	50	0.997	0.996	0.145	0.999	0.999	0.094	0.999	0.999	0.060	1.000	1.000	0.041
1	0.9	100	1.000	0.999	0.144	1.000	1.001	0.095	1.002	1.003	0.060	1.000	0.999	0.042
1	0.9	250	1.001	1.003	0.147	0.999	1.000	0.096	0.999	1.000	0.060	1.001	1.001	0.043
0.9	1	50	0.996	0.996	0.117	0.998	0.998	0.076	1.000	1.000	0.048	1.000	1.000	0.033
0.9	1	100	1.001	1.001	0.087	1.000	1.000	0.057	1.001	1.001	0.036	1.000	1.000	0.025
0.9	1	250	1.000	0.999	0.049	1.000	1.000	0.032	1.000	1.000	0.020	1.000	1.000	0.014
0.9	0.99	50	0.996	0.996	0.117	0.998	0.999	0.076	1.000	1.000	0.048	1.000	1.000	0.033
0.9	0.99	100	1.001	1.000	0.087	1.000	1.000	0.057	1.001	1.001	0.036	1.000	1.000	0.025
0.9	0.99	250	1.000	1.000	0.049	1.000	1.000	0.032	1.000	1.000	0.020	1.000	1.000	0.014
0.9	0.95	50	0.996	0.996	0.118	0.998	0.999	0.076	1.000	1.000	0.048	1.000	1.000	0.033
0.9	0.95	100	1.000	1.000	0.087	1.000	1.000	0.057	1.001	1.001	0.036	1.000	1.000	0.025
0.9	0.95	250	1.000	1.000	0.048	1.000	1.000	0.032	1.000	1.000	0.019	1.000	1.000	0.014
0.9	0.9	50	0.996	0.996	0.117	0.998	0.999	0.076	1.000	0.999	0.048	1.000	1.000	0.034
0.9	0.9	100	1.000	1.000	0.087	1.000	1.000	0.057	1.001	1.001	0.035	1.000	1.000	0.025
0.9	0.9	250	1.000	1.000	0.048	1.000	1.000	0.031	1.000	1.000	0.019	1.000	1.000	0.014

Table A.2: Empirical size and power of the panel cointegration tests with normalized spatial dependence, $N = 10$

		SAR				SMA			
		$\vartheta = 0.4$		$\vartheta = 0.8$		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.131	0.064	0.135	0.126	0.134	0.062	0.153	0.096
1	100	0.142	0.069	0.140	0.123	0.140	0.061	0.161	0.090
1	250	0.128	0.060	0.131	0.126	0.130	0.061	0.170	0.095
0.99	50	0.136	0.066	0.139	0.126	0.132	0.062	0.155	0.098
0.99	100	0.152	0.074	0.155	0.137	0.158	0.063	0.180	0.097
0.99	250	0.217	0.096	0.211	0.173	0.225	0.099	0.254	0.136
0.95	50	0.178	0.101	0.173	0.168	0.171	0.094	0.197	0.134
0.95	100	0.384	0.220	0.360	0.286	0.372	0.203	0.376	0.246
0.95	250	0.956	0.923	0.931	0.845	0.953	0.913	0.894	0.880
0.9	50	0.296	0.203	0.272	0.281	0.291	0.207	0.298	0.237
0.9	100	0.788	0.728	0.725	0.673	0.791	0.718	0.720	0.685
0.9	250	1.000	1.000	0.999	1.000	1.000	1.000	0.997	1.000

		SEC1				SEC2			
		$\vartheta = 0.4$		$\vartheta = 0.8$		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.091	0.050	0.095	0.050	0.097	0.049	0.115	0.052
1	100	0.061	0.050	0.060	0.048	0.061	0.047	0.066	0.047
1	250	0.054	0.047	0.054	0.046	0.052	0.044	0.056	0.049
0.99	50	0.092	0.050	0.096	0.050	0.102	0.052	0.118	0.052
0.99	100	0.073	0.053	0.076	0.053	0.075	0.051	0.080	0.052
0.99	250	0.159	0.077	0.155	0.077	0.152	0.081	0.154	0.078
0.95	50	0.152	0.080	0.153	0.079	0.158	0.078	0.163	0.078
0.95	100	0.450	0.177	0.455	0.178	0.458	0.176	0.452	0.177
0.95	250	0.997	0.938	0.997	0.941	0.998	0.942	0.998	0.940
0.9	50	0.337	0.176	0.331	0.177	0.335	0.182	0.308	0.181
0.9	100	0.972	0.734	0.972	0.736	0.972	0.738	0.959	0.742
0.9	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

		SEC3			
		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.138	0.056	0.128	0.070
1	100	0.123	0.048	0.135	0.059
1	250	0.095	0.052	0.141	0.064
0.99	50	0.143	0.057	0.134	0.070
0.99	100	0.130	0.052	0.157	0.063
0.99	250	0.188	0.082	0.236	0.097
0.95	50	0.174	0.083	0.180	0.096
0.95	100	0.380	0.180	0.391	0.205
0.95	250	0.973	0.934	0.961	0.920
0.9	50	0.282	0.192	0.308	0.199
0.9	100	0.815	0.739	0.820	0.721
0.9	250	1.000	1.000	1.000	1.000

Table A.3: Empirical size and power of the panel cointegration tests with normalized spatial dependence, $N = 50$

		SAR				SMA			
		$\vartheta = 0.4$		$\vartheta = 0.8$		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.073	0.064	0.131	0.126	0.064	0.061	0.083	0.075
1	100	0.083	0.069	0.156	0.139	0.076	0.064	0.105	0.084
1	250	0.060	0.062	0.138	0.134	0.065	0.065	0.091	0.085
0.99	50	0.083	0.070	0.138	0.128	0.078	0.062	0.093	0.080
0.99	100	0.137	0.078	0.195	0.157	0.120	0.077	0.162	0.100
0.99	250	0.404	0.153	0.391	0.235	0.423	0.141	0.428	0.169
0.95	50	0.346	0.133	0.266	0.196	0.344	0.120	0.359	0.145
0.95	100	0.952	0.496	0.795	0.501	0.961	0.488	0.938	0.488
0.95	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	50	0.936	0.432	0.630	0.451	0.942	0.417	0.918	0.427
0.9	100	1.000	1.000	0.999	0.982	1.000	1.000	1.000	0.998
0.9	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

		SEC1				SEC2			
		$\vartheta = 0.4$		$\vartheta = 0.8$		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.054	0.048	0.053	0.046	0.051	0.044	0.054	0.048
1	100	0.061	0.054	0.059	0.053	0.061	0.056	0.063	0.053
1	250	0.049	0.057	0.048	0.056	0.047	0.057	0.048	0.052
0.99	50	0.060	0.049	0.057	0.049	0.057	0.047	0.062	0.052
0.99	100	0.104	0.064	0.104	0.065	0.101	0.066	0.111	0.064
0.99	250	0.398	0.137	0.399	0.135	0.397	0.132	0.395	0.130
0.95	50	0.334	0.107	0.333	0.108	0.338	0.105	0.335	0.105
0.95	100	0.974	0.493	0.973	0.493	0.974	0.493	0.969	0.490
0.95	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	50	0.960	0.433	0.960	0.429	0.959	0.429	0.955	0.433
0.9	100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	250	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

		SEC3			
		$\vartheta = 0.4$		$\vartheta = 0.8$	
ϕ_i	T	Z_τ	$CADF_P$	Z_τ	$CADF_P$
1	50	0.059	0.054	0.071	0.061
1	100	0.062	0.048	0.079	0.052
1	250	0.053	0.055	0.059	0.066
0.99	50	0.071	0.055	0.081	0.063
0.99	100	0.111	0.061	0.122	0.064
0.99	250	0.397	0.133	0.399	0.144
0.95	50	0.352	0.114	0.349	0.127
0.95	100	0.968	0.488	0.959	0.482
0.95	250	1.000	1.000	1.000	1.000
0.9	50	0.957	0.432	0.938	0.428
0.9	100	1.000	0.999	1.000	1.000
0.9	250	1.000	1.000	1.000	1.000