DOCUMENTS DE TREBALL
DE LA FACULTAT DE CIÈNCIES
ECONÒMIQUES I EMPRESARIALS

Col·lecció d'Economia

Testing for multicointegration in panel data with common factors

Vanessa Berenguer-Rico y Josep Lluís Carrion-i-Silvestre*

Adreça correspondència:

Department of Econometrics, Statistics and Spanish Economy
Facultat de Ciències Econòmiques i Empresarials
Universitat de Barcelona
Avda. Diagonal 690
08034 Barcelona
e-mails: vberenguer@ub.edu, carrion@ub.edu

* We would like to thank Anindya Banerjee and two anonymous referees for helpful comments. Corresponding author: Av. Diagonal, 690. 08034 Barcelona. Tel: +34 934021826; Fax: +34 93 4021821; e-mail: carrion@ub.edu. The authors acknowledge the financial support of the Ministerio de Ciencia y Tecnología under grant SEJ2005-08646/ECON.

Abstract: The paper addresses the concept of multicointegration in panel data frame- work. The proposal builds upon the panel data cointegration procedures developed in Pedroni (2004), for which we compute the moments of the parametric statistics. When individuals are either cross-section independent or cross-section dependence can be re- moved by cross-section demeaning, our approach can be applied to the wider framework of mixed I(2) and I(1) stochastic processes analysis. The paper also deals with the issue of cross-section dependence using approximate common factor models. Finite sample performance is investigated through Monte Carlo simulations. Finally, we illustrate the use of the procedure investigating inventories, sales and production relationship for a panel of US industries.

Key words: Multicointegration, panel data, I(2) processes, common factors, cross-multicointegration, cross-section dependence

JEL classification: C12, C22

Resum: Aquest article estèn el concepte de multicointegració a l’entorn de dades de panell. La proposta es basa en els procediments de contrast de cointegració en dades de panell desenvolupats per Pedroni (2004), pels quals es calculen els moments dels estadístics paramètrics. Quan els individus són o bé independents entre si, o bé la dependencia transversal es pot eliminar treient la mitjana del tall transversal, la nostra aproximació es pot aplicar a l’àmbit a on es consideren processos estocàstics I(2) i I(1) de manera conjunta. El treball també considera aquella situació en què la dependència transversal es pot recollir mitjançant models de factors comuns. El comportament en mostra finita dels estadístics de prova és estudiat a través de simulacions de Monte Carlo. Finalment, el treball il·lustra l’ús del procediment analitzant la relació entre existències, vendes i producció per a un panell d’indústries dels Estats Units.

Paraules clau: Multicointegració, dades de panell, processos estocàstics I(2), factors comuns, multicointegració transversal, dependència transversal

Classificació JEL: C12, C22

1 Introduction

Panel data techniques for macroeconomic analysis have experienced huge development in recent years. The increasing availability of statistical information has allowed to conduct studies using data of different countries, regions or cities to get more insights into economic relationships. In addition, the use of panel data statistics allows improvement in the power of statistical inference since it combines the information in both the time and cross-section dimensions. Macroeconomic panels can be characterized as those panel data sets with moderate or large number of observations (T) compared to the number of individuals (N). This feature implies that non-stationarity in variance can be present in the panel data set, so that practitioners have to check whether the estimation of their model relating economic variables results in spurious regression or in cointegration relationship.

Non-stationarity in variance has been profusely addressed in panel data literature. We can find proposals that extend univariate unit root and stationarity tests to panel data framework, and similar developments have been proposed in cointegration analysis. Overviews of the field can be found in Banerjee (1999), and Breitung and Pesaran (2005). Main developments in panel data framework have addressed cointegration relationships. These proposals allow the assessment of the presence of long-run relationships among variables that in most cases are characterized as $I(1)$ processes. However, standard cointegration analysis might be incomplete even in the case that cointegration is found. Thus, it is possible that a deeper level of cointegration, i.e. multicointegration, exists. As noted in Engsted, Gonzalo and Haldrup (1997), multicointegration is also an important property of the data that needs to be considered empirically. The statistical properties of the procedures that are used for estimating and testing cointegrated systems become invalid if multicointegration is not taken into account when it is present. This will have serious consequences, for instance, in forecasting and hypothesis testing. Therefore, consideration of multicointegration when it is present can give us better statistical results when we analyze long-run economic relationships, especially in those cases where stock-flow relationships are involved. Extending multicointegration to panel data framework is not only a matter of theoretical interest, but useful from an empirical point of view.

Empirical applications considering multicointegration have appeared in time series literature, although they can be extended to panel data framework as well – see Granger and Lee (1989), Lee (1992), Leachman (1996), Leachman and Francis (2002), and Siliverstovs (2003). Previous multicointegration analyses are carried out either for one or more individuals, although the stochastic properties are studied individual-by-individual. For the latter, the application of multicointegration in panel data is of interest, provided that we can gain more insight on panel data stochastic properties through the combination of the information in both the cross-section and time series dimensions. Finally, it is worth mentioning that our approach can be applied to the analysis of cointegration with $I(2)$

processes, since multicointegration is a special case of polynomial cointegration. In this case, further applications can be conducted – for instance, the analyses in Juselius (1999, 2004), and Banerjee, Cockerell and Russell (2001) can be extended to panel data framework. Multicointegration has mostly been tested using either the two-step approach in Granger and Lee (1989), the one-step approach in Engsted, Gonzalo and Haldrup (1997), or the Error Correction Model specification in Engsted and Haldrup (1999). In this paper we address this concern and generalize the approach in Pedroni (2004) to tackle panel data multicointegration.

One important feature of non-stationary panel data analysis is cross-section dependence. So far, cross-section independence among individuals has been commonly assumed in all these cases since it allows standard Normal limiting distributions to be obtained. However, this assumption plays an important role in practice. Banerjee, Marcellino and Osbat (2005) analyze the effects of cross-section dependence in panel data unit root tests that assume independence among individuals. They show that important size distortions (over-rejections) appear when cross-section dependence is ignored. Recent developments in the literature aim to weaken this assumption using different approaches to account for cross-section dependence. In this paper we proceed in two stages. First, we derive the limiting distribution of the panel multicointegration test assuming that the individuals are cross-section independent. In a second stage we consider the factor structure in Bai and Ng (2004), and Banerjee and Carrion-i-Silvestre (2006) to account for cross-section dependence whether the first level cointegrating vector is known or unknown. Both the one and the two-step approaches available in the literature for testing multicointegration are useful in conducting our analysis.

The paper is organized as follows. Section 2 defines the concept of multicointegration and presents the model that is used in the paper. Section 3 defines the panel data multicointegration test statistics, for which both finite sample and asymptotic moments are computed. In Section 4 we consider the presence of cross-section dependence when testing for multicointegration through common factors models. Section 5 analyses the finite sample performance. In Section 6 we investigate the presence of multicointegration between sales and production for a panel data set of forty-eight US industries. Finally, Section 7 concludes.

2 Multicointegration in panel data

Cointegration is a necessary condition for the presence of multicointegration as defined in Granger and Lee (1989). Thus, if we consider one dimensional time series $\{y_{i,t}\}_0^\infty$ and m-dimensional time series $\{x_{i,t}\}_0^\infty$ all being $I(1)$ non-stationary stochastic processes, $t = 1, \ldots, T$, $i = 1, \ldots, N$, these variables are assumed to satisfy the following standard cointegration model:

$$y_{i,t} = c_t\alpha_i + x_{i,t}\beta_i + \vartheta_{i,t}, \tag{1}$$

where $\{c_t\}_1^\infty$ is an s_0-dimensional deterministic sequence of general form – typically, $c_t = 0$, $c_t = 1$ and $c_t = (1, t)$ – and where $\vartheta_{i,t}$ is an $I(0)$ series. Suppose that the cumulated cointegration residuals, $S_{i,t} = \sum_{j=1}^{t} \vartheta_{i,j}$, cointegrate with either $\{y_{i,t}\}_0^\infty$ and/or $\{x_{i,t}\}_0^\infty$, then we obtain the standard multicointegration model, that is

$$S_{i,t} = m_t\delta_i + x_{i,t}\gamma_i + u_{i,t}, \tag{2}$$

where $\{m_t\}_1^\infty$ is the s_1-dimensional deterministic sequence and where $u_{i,t}$ is an $I(0)$ series.

Following Engsted, Gonzalo and Haldrup (1997), we can write (2) as:

$$Y_{i,t} = Cm_t\mu_i + X_{i,t}\beta_i + x_{i,t}\gamma_i + u_{i,t}, \tag{3}$$

where $Y_{i,t} = \sum_{j=1}^{t} y_{i,j}$ and $X_{i,t} = \sum_{j=1}^{t} x_{i,j}$ are $I(2)$ variables and $Cm_t = \sum_{j=1}^{t} c_j + m_t$ is the new m_0-deterministic component associated to multicointegration relation (3), with $m_0 = s_0 + s_1$ and $\mu_i = (\alpha_i', \delta_i')'$. The specification given by (3) can be written using the Phillips' (1991) triangular representation. Thus, $\{Y_{i,t}\}_0^\infty$ and $\{X_{i,t}^m\}_0^\infty$ are assumed to be generated according to $Y_{i,t} = Cm_t\mu_0 + Y_{i,t}^0$ and $X_{i,t}^m = (Cm_t, x_{i,t}, X_{i,t})$, with the stochastic regressors defined as $x_{i,t} = Cm_t\mu_{i,1} + x_{i,t}^0$, $\Delta x_{i,t}^0 = \varepsilon_{i,1t}$, $X_{i,t} = Cm_t\mu_{i,2} + X_{i,t}^0$, $\Delta^2 X_{i,t}^0 = \varepsilon_{i,2t}$, where $x_{i,t}^0$ and $X_{i,t}^0$ are the m_1 and m_2-dimensional stochastic processes integrated of order one and two, respectively, and Cm_t denotes the m_0-deterministic component of the different variables. $Y_{i,t}^0$ is generally integrated of order two and linked to $x_{i,t}^0$ and $X_{i,t}^0$ through

$$Y_{i,t}^0 - x_{i,t}^0\gamma_i - X_{i,t}^0\beta_i = u_{i,t}, \tag{4}$$

with $\Delta^d u_{i,t} = v_{i,t}$. The order of integration d can be either $d = 0, 1$ or 2, which is to be discussed below. The processes $x_{i,t}^0$, $X_{i,t}^0$, $Y_{i,t}^0$ are initialized at $t = 1, 0, 0$, respectively – this does not affect the results. The $w_{i,t} = (v_{i,t}, \varepsilon_{i,1t}, \varepsilon_{i,2t})'$ stochastic processes involved in the definition of the model are assumed to be a strong-mixing sequence satisfying the multivariate invariance principle in Phillips and Durlauf (1986). Thus, let $B_T(r) = T^{-1/2}\sum_{t=1}^{[Tr]} w_{i,t}$ be the partial sum process. Then, as $T \to \infty$, $B_T(r) \Rightarrow B(r) \equiv BM(\Omega)$, where $\Rightarrow$ denotes weak convergence of the associated probability measure on the unit interval [0,1], and $B(r)$ denotes a vector Brownian motion process with long-run variance matrix Ω_i. Moreover, we partition Ω_i conformably with $w_{i,t}$, so that

$$\Omega_i = \begin{bmatrix} \omega_{i,00} & \omega_{i,01} & \omega_{i,02} \\ \omega_{i,10} & \Omega_{i,11} & \Omega_{i,12} \\ \omega_{i,20} & \Omega_{i,21} & \Omega_{i,22} \end{bmatrix} = \Sigma_i + \Lambda_i + \Lambda_i', \tag{5}$$

where $\Sigma_i = E(w_{i,0}w'_{i,0})$ and $\Lambda_i = \sum_{k=1}^{\infty} E(w_{i,0}w'_{i,k})$. For subsequent use, we also define $\Delta_i = \Sigma_i + \Lambda_i$, which can be partitioned in conformity with Ω_i. In (5) the diagonal submatrices $\Omega_{i,11}$ and $\Omega_{i,22}$ are assumed to be positive definite such that $x^0_{i,t}$ and $X^0_{i,t}$ are not permitted to be individually cointegrated.

There are in this $I(2)$ system several cointegration possibilities depending on the order of integration of $u_{i,t}$, i.e. $\Delta^d u_{i,t} = v_{i,t}$ with $d = 0, 1, 2$. When $d = 2$ there do not exist either cointegration or multicointegration because there is not any common stochastic trend – i.e. $u_{i,t}$ process is integrated of order two. When $d = 1$ there is only cointegration at the first level. Note that in this case $Y^0_{i,t}, X^0_{i,t} \sim CI(2,1)$ with cointegrating vector $(1, -\beta_i)$ and, hence, $\sum_{j=1}^{t} z_{i,j} = \sum_{j=1}^{t} y_{i,j} - \sum_{j=1}^{t} x_{i,j}\beta_i$ is integrated of order 1. Then, the residuals $z_{i,t}$ must be stationary showing that there is cointegration at the first level. Finally, when $d = 0$ we conclude that the variables $y_{i,t}$ and $x_{i,t}$ are multicointegrated in such a way that all stochastic trends are cancelled in the multicointegration relation. The conditional model (4) can be expressed as:

$$Y_{i,t} = Cm_t\mu_i + X_{i,t}\beta_i + x_{i,t}\gamma_i + u_{i,t} = X^m_{i,t}\alpha_i + u_{i,t}, \tag{6}$$

where $\mu_i = (\mu_{i,0} - \mu_{i,1}\gamma_i - \mu_{i,2}\beta_i)$. Depending upon the integration order of $u_{i,t}$, there may be stochastic cointegration at different levels as well as deterministic co-trending if some elements in μ_i turn out to be zero, although the series individually have nonzero elements in their deterministic part. The specification in (6) nests the multicointegration framework defined in Haldrup (1994) and Engsted, Gonzalo and Haldrup (1997) once we specify either $Cm_t = 0$, $Cm_t = 1$, $Cm_t = (1, t)$ or $Cm_t = (1, t, t^2)$, i.e. zero, constant, trend and quadratic trend respectively.

3 Testing the null of non-multicointegration in panel data

In this section we present the panel data residual based statistic that allows testing the null hypothesis of non-multicointegration. Although our set-up builds upon the multicointegration framework, the proposal can be applied in more general situations in which the presence of cointegration can be tested for mixed $I(2)$ and $I(1)$ variables. This is of great interest provided that, to the best of our knowledge, there are not any proposal in non-stationary panel data analysis that address this concern.

The computation of the statistics proceeds as follows. First, the OLS estimated residuals in (6) are used to specify an augmented Dickey-Fuller type regression,

$$\Delta\hat{u}_{i,t} = \rho_i\hat{u}_{i,t-1} + \sum_{j=1}^{p_i} \phi_{i,j}\Delta\hat{u}_{i,t-j} + \epsilon_{i,t}, \tag{7}$$

from which either the normalized bias – computed as $T\hat{\rho}_i\left(1-\hat{\phi}_{i,1}-\cdots-\hat{\phi}_{i,p_i}\right)^{-1}$, see Hamilton (1994), pp. 523 – or the t-ratio statistic $\left(t_{\hat{\rho}_i}\right)$ can be defined for each individual. Second, the individual information can be combined using the parametric between-dimension panel data statistics as defined in Pedroni (2004), i.e.

$$N^{-1/2}Z_{\hat{\rho}_{NT}}=N^{-1/2}\sum_{i=1}^{N}\frac{T\hat{\rho}_i}{\left(1-\hat{\phi}_{i,1}-\cdots-\hat{\phi}_{i,p_i}\right)};\quad N^{-1/2}Z_{\hat{t}_{NT}}=N^{-1/2}\sum_{i=1}^{N}t_{\hat{\rho}_i}. \tag{8}$$

Haldrup (1994) follows Phillips and Ouliaris (1990) and derives the limiting distribution for the individual $t_{\hat{\rho}_i}$ statistics, which is shown to converge to $t_{\hat{\rho}_i}\Rightarrow\int_0^1 Q_i(s)\,dQ_i(s)$ $\left[\int_0^1 Q_i(s)^2\,ds\right]^{-1/2}$, and, similarly, $\left(1-\hat{\phi}_{i,1}-\cdots-\hat{\phi}_{i,p_i}\right)^{-1}T\hat{\rho}_i\Rightarrow\int_0^1 Q_i(s)\,dQ_i(s)\left[\int_0^1 Q_i(s)^2\,ds\right]^{-1}$, where $Q_i(s)=\omega_{i,00.1}^{1/2}\left(W_{i,0}(s)-W_{i,*}(s)\left(\int_0^1 W_{i,*}'(s)\,W_{i,*}(s)\right)^{-1}\left(\int_0^1 W_{i,0}(s)\,W_{i,*}(s)\right)\right)$, with $\omega_{i,00.1}=\omega_{i,00}-\omega_{i,01}\Omega_{i,11}^{-1}\omega_{i,10}$, $W_{i,*}(s)=\left(f(s),W_{i,1}(s),W_{i,2}(s)\right)'$, $f(s)$ denotes the limit of the deterministic components, $W_{i,1}(s)$ is a vector of m_1 Brownian motions, and $W_{i,2}(s)$ the vector of m_2 integrated Brownian motions.

Note that this framework considers high degree of heterogeneity since both the cointegrating vector and the short-run dynamics vary among individuals. The panel test statistics are shown to converge to standard Normal distributions once they have been properly standardized.

Theorem 1 *Let $\{y_{i,t}\}_0^\infty$ and $\{x_{i,t}\}_0^\infty$ be the $I(1)$ stochastic processes that define the cointegration relationship given in (1), and $Y_{i,t}=\sum_{j=1}^{t}y_{i,j}$ and $X_{i,t}=\sum_{j=1}^{t}x_{i,j}$ be the $I(2)$ stochastic processes that define the multicointegration relationship in (6). Let Θ and Ψ denote the mean and variance for the vector Brownian motion functional $\Upsilon\equiv\left(\int_0^1 Q_i(s)\,dQ_i(s)\left[\int_0^1 Q_i(s)^2\,ds\right]^{-1},\int_0^1 Q_i(s)\,dQ_i(s)\left[\int_0^1 Q_i(s)^2\,ds\right]^{-1/2}\right)'$. Furthermore, let p_i be the order of autoregression chosen such that $p_i\to\infty$ and $p_i^3/T\to 0$. Under the null hypothesis of non-multicointegration that $\rho_i=0\ \forall i, i=1,\ldots,N$, in (7) and assuming that individuals are cross-section independent, the $Z_{\hat{\rho}_{NT}}$ and $Z_{\hat{t}_{NT}}$ statistics given in (8) converge as $T\to\infty$ followed by $N\to\infty$, i.e. $(T,N\to\infty)_{\text{seq}}$, to:*

$$N^{-1/2}Z_{\hat{\rho}_{NT}}-\Theta_1\sqrt{N}\Rightarrow N(0,\Psi_1);\quad N^{-1/2}Z_{\hat{t}_{NT}}-\Theta_2\sqrt{N}\Rightarrow N(0,\Psi_2).$$

where Θ_1,Ψ_1,Θ_2 and Ψ_2 are the mean and variance of the Brownian motion functionals to which the individual normalized bias and t-ratio statistics converge.

As in Pedroni (2004), in order to prove Theorem 1 we require only the assumption of finite second moments of the random variables characterized as Brownian motion functionals, which will allow us to apply the Lindberg-Levy Central Limit Theorem as $N\to\infty$. The moments of the limiting distributions, Θ_1,Ψ_1,Θ_2 and Ψ_2, are approximated by Monte Carlo simulation for the four deterministic specifications – zero, constant, trend and quadratic trend – for different combinations of m_1 $I(1)$ and m_2 $I(2)$ stochastic regressors in the cointegrating relationship. To be specific, we have followed Haldrup (1994), and Engsted, Gonzalo and Haldrup (1997) and define $(m_2 + 1)$ $I(2)$ stochastic processes – one for the endogenous variable and m_2 for the regressors – using partial sum of partial sum of iid $N(0,1)$, whereas for the m_1 $I(1)$ stochastic processes we have used using partial sum of iid $N(0,1)$ with $T = 1,000$ in all cases. Tables 1 and 2 present the moments of the limit distributions.

Since the limit distribution of the tests can provide poor approximation in finite samples, we have approximated the moments of the statistics for $T = \{50, 100, 250\}$ as well. For these sample sizes the moments have been computed selecting the order (p_i) of the parametric correction in (7) with the t-sig criterion in Ng and Perron (1995) with $p_{\max} = 5$ as the maximum number of lags. Other criteria might be followed to select the order of the autoregressive correction – i.e. we could chose p_i by means of information criteria such as AIC or BIC, or fix p_i in exogenous way. Since the t-sig criterion in Ng and Perron (1995) is one of the most widely used strategies in practice, we have preferred to compute finite sample moments following this approach. Obviously, the use of different methods to select p_i in finite samples affects the moments that have to be used to compute the statistic. Thus, practitioners willing to apply other criteria when selecting the order of the autoregressive correction should compute the moments of the statistics in finite samples. Tables 1 and 2 report the finite sample moments for the different deterministic specifications based on the t-sig criterion in Ng and Perron (1995). In all simulations 10,000 replications were done. As can be seen, the moments of the distribution depends both on the specification and the number of stochastic regressors.

4 Panel multicointegration with common factors

Previous sections have assumed that individuals in the panel data set are independent from each other. Notwithstanding, economic models predict that macroeconomic variables such as GDP, consumption, interest rates, exchange rates and investment for different countries are related. Economic models for which multicointegration can be present are based on some of these variables, so dependence among individuals is found. For instance, life-cycle hypothesis involves income, consumption and wealth, which are expected to be related for different countries. Unfulfilment of independence among individuals implies that previous results no longer hold. As mentioned in the introduction, there are different approaches in the literature to account for cross-section dependence. In this section we adopt approximate common factor models to model cross-section dependence among individuals. Our specification follows that in Banerjee and Carrion-i-Silvestre (2006) for the panel cointegration analysis. However, the application of this approach has led us to write the multicointegration testing procedure in terms of the two-steps procedure in Granger and Lee (1989), instead of using the one-step approach in Engsted, Gonzalo and Haldrup (1997).

Let us assume that we have a panel data multicointegrated set given by:

$$y_{i,t} = c_t\alpha_i + x_{i,t}\beta_i + \vartheta_{i,t} \tag{9}$$

$$y_{i,t} = m_t\mu_i + S_{i,t}\gamma_i + u_{i,t} \tag{10}$$

where $S_{i,t} = \sum_{j=1}^{t}\vartheta_{i,j}$. Equation (9) represents the first level of cointegration, while (10) specifies the multicointegrating relationship, i.e. the relationship between $y_{i,t}$ and the cumulated residuals of the first level cointegrating regression. Since multicointegration requires that variables in levels have to be cointegrated, then $S_{i,t} \sim I(1)$ by definition. The OLS estimated residuals in (9) can be written as $\hat{\vartheta}_{i,t} = \vartheta_{i,t} - c_t(\hat{\alpha}_i - \alpha_i) - x_{i,t}\left(\hat{\beta}_i - \beta_i\right)$. Note that $(\hat{\alpha}_i - \alpha_i) = O_p\left(T^{-1/2}\right)$ and $\left(\hat{\beta}_i - \beta_i\right) = O_p\left(T^{-1}\right)$ – see Phillips and Ouliaris (1990) – so that $\hat{\vartheta}_{i,t} = \vartheta_{i,t} + O_p\left(T^{-1/2}\right)$. This feature allows us to use the cumulated estimated residuals of (9) and defines the following set-up:

$$y_{i,t} = m_t\mu_i + S_{i,t}\gamma_i + u_{i,t} \tag{11}$$

$$u_{i,t} = F_t\pi_i + e_{i,t} \tag{12}$$

$$(I - L)F_t = C(L)v_t \tag{13}$$

$$(1 - \rho_i L)e_{i,t} = H_i(L)\xi_{i,t} \tag{14}$$

$$(I - L)x_{i,t} = G_i(L)\varepsilon_{i,t}, \tag{15}$$

where $S_{i,t} = \sum_{j=1}^{t}\vartheta_{i,j}$, which can be estimated using $\hat{\vartheta}_{i,t}$ obtained in the first step without affecting the results. $C(L) = \sum_{j=0}^{\infty}C_jL^j$, F_t denotes a $(1 \times r)$-vector containing the common factors, with π_i the vector of loadings. Despite the operator $(1 - L)$ in equation (13), 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 $I(0)$. Our analysis is based on the same set of assumptions in Bai and Ng (2004), and Banerjee and Carrion-i-Silvestre (2006). As in Banerjee and Carrion-i-Silvestre (2006), we distinguish two situations depending on whether the stochastic regressor $\hat{S}_{i,t}$ is strictly exogenous or non-strictly exogenous regressor. This distinction is important since under strict exogeneity the limiting distribution of statistics does not depend on $\hat{S}_{i,t}$. However, this is not true when correlation between $e_{i,t}$ and $\hat{\vartheta}_{i,t}$ is allowed so some sort of modifications should be introduced to account for the endogeneity. Here we suggest using the Dynamic OLS (DOLS) estimation method in Stock and Watson (1993). Throughout the paper, we assume that the number of leads and lags is fixed as in

Stock and Watson (1993), although they can be chosen using BIC information criterion as suggested in Westerlund (2005b) – see Banerjee and Carrion-i-Silvestre (2006) for further details on non-strictly exogenous regressors.

For ease of exposition, we assume that $\hat{S}_{i,t}$ is strictly exogenous stochastic regressor. The estimation of the common factors is done as in Bai and Ng (2004), and Banerjee and Carrion-i-Silvestre (2006). The procedure to estimate both the idiosyncratic disturbance term and the common factors proceeds as follows. First, we compute the first difference of the model:

$$\Delta y_{i,t} = \Delta m_t \mu_i + \Delta \hat{S}_{i,t} \gamma_i + \Delta F_t \pi_i + \Delta e_{i,t}. \tag{16}$$

Note that if $\hat{S}_{i,t}$ is non-strictly exogenous, we should introduce leads and lags of $\Delta^2 \hat{S}_{i,t}$ in (16). Then, we take the orthogonal projections $y_{i,t}^* = f_t \pi_i + z_{i,t}$, with $y_i^* = M_i \Delta y_i$, being $M_i = I - \Delta x_i^* (\Delta x_i^{*\prime} \Delta x_i^*)^{-1} \Delta x_i^{*\prime}$ the idempotent matrix, $f = M_i \Delta F$, $z_i = M_i \Delta e_i$ and Δx_i^* the matrix that collects the first difference of the deterministic and the $\hat{S}_{i,t}$ stochastic regressor – the superscript $*$ in Δx_i^* indicates that there are deterministic elements. The estimation of the common factors and factor loadings can be done as in Bai and Ng (2004) using principal components. Then, the estimated residuals are defined as $\tilde{z}_{i,t} = y_{i,t}^* - \tilde{f}_t \tilde{\pi}_i$, so that we can recover the idiosyncratic disturbance terms through cumulation, i.e. $\tilde{e}_{i,t} = \sum_{j=2}^{t} \tilde{z}_{i,j}$, and test the unit root hypothesis using the ADF regression equation. When $r = 1$ we can use the ADF type equation to analyze the order of integration of F_t as well. However, we should proceed in two steps. In the first step, we regress $\tilde{F}_t$ on the deterministic specification and the stochastic regressors. In the second step, we estimate the ADF regression equation using the detrended common factor $\left(\tilde{F}_t^*\right)$, i.e. the residuals of the first step. Finally, if $r > 1$ we should use one of the two statistics proposed in Bai and Ng (2004) – denoted as $MQ_c^*(q)$ for the non-parametric statistic and $MQ_f^*(q)$ for the parametric one – to fix the number of common stochastic trends (q). The following Theorem presents the limiting distribution of these statistics.

Theorem 2 *Let $\{Y_{i,t}\}$ the stochastic process with DGP given by (11) to (15). Let p_i be the order of autoregression chosen such that $p_i \to \infty$ and $p_i^3 / \min[N,T] \to 0$. Then, the following results hold as $(T, N \to \infty)_{\text{seq}}$.*
(1) Under the null hypothesis that $\rho_i = 1$ in (14),

$$ADF_{\tilde{e}}^c(i) \Rightarrow \frac{\frac{1}{2}\left(W_i(1)^2 - 1\right)}{\left(\int_0^1 W_i(r)^2 dr\right)^{1/2}}; \quad ADF_{\tilde{e}}^\tau(i) \Rightarrow -\frac{1}{2}\left(\int_0^1 V_i^\tau(r)^2 dr\right)^{-1/2},$$

where $ADF_{\tilde{e}}^c(i)$ and $ADF_{\tilde{e}}^\tau(i)$ denote the statistics for the constant and time trend specifications, respectively, and $V_i^\tau(r) = W_i(r) - rW_i(1)$.

(2) When $r = 1$, under the null hypothesis that F_t has a unit root

$$ADF_{\tilde{F}}^{*} \Rightarrow \frac{\int_0^1 W_w^*(r)\, dW_w^*(r)}{\left(\int_0^1 W_w^*(r)^2\, dr\right)^{1/2}},$$

where $W_w^*(r)$ denotes the detrended – either by a constant or a linear time trend depending on the deterministic specification – Brownian motion.
(3) When $r > 1$, let W_q be a q-vector of standard Brownian motion and W_q^* the detrended counterpart. Let $v^*(q)$ be the smallest eigenvalues of the statistic computed as

$$\Phi^* = \frac{1}{2}\left[W_q^*(1)\, W_q^*(1)' - I_p\right]\left[\int_0^1 W_q^*(r)\, W_q^*(r)'\, dr\right]^{-1},$$

(3.1) Let J be the truncation lag of the Bartlett kernel, chosen such that $J \to \infty$ and $J/\min\left[\sqrt{N}, \sqrt{T}\right] \to 0$. Then, under the null hypothesis that F_t has q stochastic trends, $T\left[\tilde{v}_c^*(q) - 1\right] \xrightarrow{d} v^*(q)$.
(3.2) Under the null hypothesis that F_t has q stochastic trends with a finite VAR($\bar{p}$) representation and a VAR(p) is estimated with $p \geq \bar{p}$, $T\left[\tilde{v}_f^*(q) - 1\right] \xrightarrow{d} v^*(q)$.

The proof of Theorem 2 is entirely analogous to that in Banerjee and Carrion-i-Silvestre (2006) and sketched in the Appendix. Note that the limiting distribution of the statistics is the same as the ones in Bai and Ng (2004), and in Banerjee and Carrion-i-Silvestre (2006). We can define a panel data unit root statistic using the individual ADF statistics computed for the idiosyncratic disturbance term, i.e. $N^{-1/2} Z_j^{\tilde{e}} = N^{-1/2} \sum_{i=1}^{N} ADF_{\tilde{e}}^{j}(i)$, $j = \{c, \tau\}$, which, standardized, it is shown to converge to the standard Normal distribution. Asymptotic and finite sample moments $\Theta_j^{\tilde{e}}$ and $\Psi_j^{\tilde{e}}$, $j = \{c, \tau\}$, of the statistics are reported in Table 3 using 1,000 replications – note that these moments can be also used to compute those statistics in Bai and Ng (2004), and in Banerjee and Carrion-i-Silvestre (2006). As above, the finite sample moments are based on the use of the t-sig criterion in Ng and Perron (1995) with $p_{\max} = 5$ as the maximum number of lags for the autoregressive correction.

The presence of multicointegration depends on the rank of the $C(1)$ matrix in (13) and on the values of ρ_i in (14). Thus, if $\rho_i = 1\ \forall i$ and $C(1)$ is of full rank, multicointegration does not exist. Multicointegration is present when $\rho_i < 1\ \forall i$ and $C(1) = 0$. Finally, multicointegration will be present with up to r_1 ($\leq r$) non-stationary factors if $\rho_i < 1\ \forall i$ and $C(1) \neq 0$, but not full rank, since then some components of F_t are $I(1)$ and some are $I(0)$. This situation can be encountered if cross-multicointegration is present between $y_{i,t}$ and/or $x_{i,t}$ for different individuals.[1] In this case, the non-stationary common factor

[1]Note that this is a natural extension of the concept of cross-cointegration in Banerjee, Marcellino

might be understood as a common stochastic trend relating $y_{i,t}$ (or $x_{i,t}$) series in each panel data set.

5 Monte Carlo simulations

Finite sample performance of the statistics that have been proposed in this paper is investigated using simulations. We present simulation results for the situation in which individuals are assumed to be cross-section independent, and for the case where cross-section dependence is driven by common factors. Note that throughout the section we assume that the the first level cointegrating vector β_i is assumed to be unknown. Throughout the section, the lag order for the autoregressive correction that is required for computing the ADF statistics is selected using the t-sig criterion in Ng and Perron (1995) with $p_{\max} = 5$ lags as the maximum order.

5.1 Cross-section independent individuals

The data generating process (DGP) that has been used is given by

$$\begin{aligned} x_{i,t} &= A_i W_{i,t} + \alpha_{i,1} \Delta W_{i,t} + x_{i,2t} \\ y_{i,t} &= W_{i,t} + \alpha_{i,2} \Delta W_{i,t} + y_{i,2t}; \quad \Delta W_{i,t} = \xi_{i,t}, \end{aligned}$$

where, without loss of generality, $A_i = 2 \ \forall i = 1, \ldots, N$, and $\xi_{i,t} \sim N\left(0, \sigma^2_{\xi,i}\right)$. Note that when $\alpha_{i,1} = \alpha_{i,2} = 0$ and $x_{i,2t}, y_{i,2t} \sim I(0)$ we are under the null hypothesis of non-multicointegration, while when $\alpha_{i,1} \neq \alpha_{i,2} \neq 0$ and $x_{i,2t}, y_{i,2t} \sim I(-1)$ we are under the alternative hypothesis of multicointegration. We have several local alternatives depending upon the values for $\alpha_{i,1}$, $\alpha_{i,2}$ and the different possibilities to obtain $I(-1)$ processes, i.e. overdifferenced stationary processes. The $x_{i,2t}$ and $y_{i,2t}$ stochastic processes have been defined as follows:

$$\begin{aligned} x_{i,2t} &= \Delta w_{i,1t} & y_{i,2t} &= \Delta w_{i,2t} \\ w_{i,1t} &= \rho_{i,1} w_{i,1t-1} + \varepsilon_{i,1t} & w_{i,2t} &= \rho_{i,2} w_{i,2t-1} + \varepsilon_{i,2t}, \end{aligned}$$

where $\varepsilon_{i,1t} \sim N\left(0, \sigma^2_{\varepsilon_{i,1}}\right)$, $\varepsilon_{i,2t} \sim N\left(0, \sigma^2_{\varepsilon_{i,2}}\right)$ with $\sigma^2_{\varepsilon_{i,1}} = \sigma^2_{\varepsilon_{i,2}} = 1$. Note that when $\rho_{i,1} = \rho_{i,2} = 1$ we are under the null hypothesis, while for $\rho_{i,1}, \rho_{i,2} < 1$ we are under the alternative hypothesis.

Several specifications have been adopted in the Monte Carlo simulations for the parameters of interest when analyzing the empirical power – i.e. under the alternative hypothesis. We have imposed $\alpha_{i,1} = \alpha_{i,2} = 0.5$ for all individuals since they do not

and Osbat (2005).

affect the empirical power of the statistics. There are two different sets of parameters that affect the empirical power of the statistics, i.e. the autoregressive parameters $\rho_{i,1}$ and $\rho_{i,2}$, and $\sigma^2_{\xi,i}$, which has the interpretation of a signal-to-noise ratio since we have set $\sigma^2_{\varepsilon_{i,1}} = \sigma^2_{\varepsilon_{i,2}} = 1$. Regarding these parameters, we have followed two approaches. First, we have assumed that they are fixed and common to all individuals setting $\rho_{i,1} = \rho_{i,2} = \{0.99, 0.9\}$ and $\sigma^2_{\xi,i} = \{10, 30, 50\}$. Second, when investigating the empirical power we have allowed heterogeneous values for both the autoregressive parameters specifying $\rho_{i,1} = \rho_{i,2} \sim U\,[0.9, 0.99]$ and for the signal-to-noise ratio $\sigma^2_{\xi,i} \sim U\,[10, 50]$, where U denotes the uniform distribution. In all cases, we have carried out simulations for $T = \{50, 100, 250, 1,000\}$ and $N = \{20, 40\}$, with 1,000 replications. The nominal size is set at the 5% level of significance.

Tables 4 and 5 report the empirical size and power when the relevant parameters of the model are homogeneous for all individuals. In general, the statistics show mild over size distortions for small sample sizes ($T = 50$), though the empirical size tends to be close to nominal one as T increases. In addition, size distortions decrease with the signal-to-noise ratio. Notwithstanding, note that these size distortions only appear for those specifications that include deterministic terms, since for the non-deterministics case the empirical size is around the nominal one in almost all situations. These oversize distortions could be explained by the specification of the DGP in structural form. Note that if we express the DGP in final form we obtain that the disturbance term is a mixture of the disturbance terms in the structural form, which under the null hypothesis include overdifferenced stochastic processes. Simulations not reported here indicate that the empirical size equals the nominal one in all cases when the stochastic $I(2)$ and $I(1)$ processes are generated in an independent way. Regarding the empirical power, the test statistics show good properties with values that equal one in most cases when $\rho_{i,1} = \rho_{i,2} = 0.9\ \forall i$. Furthermore, the statistics have reasonable power values even when $\rho_{i,1} = \rho_{i,2} = 0.99\ \forall i$, i.e. when we are very close to the null hypothesis. Finally, results are very similar if we allow for heterogeneous individuals – see Table 6. We observe size distortions for those specifications that include linear or quadratic time trend, while the statistics have empirical size close to the nominal one for the non-deterministic and constant specifications. In all cases, the empirical power is high with values that equal one in most situations.

5.2 Cross-section dependent individuals

We define the DGP given by

$$
\begin{aligned}
y_{i,t} &= S_{i,t} + u_{i,t} && (17)\\
x_{i,t} &= S_{i,t} + \Delta S_{i,t} + u_{i,t} && (18)\\
\Delta S_{i,t} &= \vartheta_{i,t}; \qquad u_{i,t} = F_t \pi_i + e_{i,t},
\end{aligned}
$$

where $\vartheta_{i,t} \sim iid\ N(0,1)$ and $\pi_i \sim U[2,10]$. The idiosyncratic disturbance terms are generated according to $e_{i,t} = \rho_i e_{i,t-1} + \xi_{i,t}$, with $\xi_{i,t} \sim iid\ N\left(0, \sigma^2_{\xi_i}\right)$, while the common factor term is given by $F_t = \theta F_{t-1} + \upsilon_t$, with $\upsilon_t \sim iid\ N(0, \sigma^2_\upsilon)$. In this section we consider the case of one known common factor ($r = 1$) as well as the case three ($r = 3$) unknown common factors. For the later, we have estimated the number of common factors using the panel BIC information criteria in Bai and Ng (2004) allowing for up to six common factors. In order to save space, we only investigate the empirical size and power of the $Z^{\tilde{e}}_\tau$ statistic, using $\rho_i = \{0.9, 0.99, 1\}$ $\forall i$, $\theta = \{0.9, 0.95, 1\}$, with $\sigma^2_{\xi_i} = \{3, 5\}$ and $\sigma^2_\upsilon = 1$ – note that $\sigma^2_{\xi_i}$ has interpretation in terms of signal-to-noise ratio. Simulations are computed for $T = \{50, 100, 250\}$ and $N = 40$, using 1,000 replications. The nominal size is set at the 5% level of significance.

The DGP given in (17) and (18) implies that $y_{i,t}$ and $x_{i,t}$ are cointegrated with cointegrating vector (1, -1). The presence of multicointegration depends on the values of ρ_i and θ. Thus, if $\rho_i = 1$ $\forall i$ and/or $\theta = 1$, multicointegration does not exist. Multicointegration is present when both ρ_i and θ are less than one $\forall i$. Finally, multicointegration will be present up to r_1 ($\leq r$) non-stationary factors if $\rho_i < 1$ $\forall i$ and $\theta = 1$ for r_1 common factors. We have obtained the OLS estimated residuals from $y_{i,t} = \alpha_i + x_{i,t}\beta_i + \vartheta_{i,t}$, and defined $S_{i,t} = \sum_{j=1}^{t} \hat{\vartheta}_{i,j}$. Then, we have proceeded following the procedure described in Section 4, with $m_t = (1, t)$ in (11). We use the finite sample moments reported in Table 3 to compute the $Z^{\tilde{e}}_\tau$ statistic.

Let us first focus on the results for the one common factor. Results in Table 7 indicate that for large T the empirical size of the $Z^{\tilde{e}}_\tau$ statistic is close to the nominal one, although size distortion appear for $T = 50$. This feature was to be expected since in this case T is similar to N – note that our approach requires T larger than N. The ADF^*_F statistic shows good empirical size, although mild distortions appear for $T = 50$. As expected, the power of both statistics increases as ρ_i and θ moves away from one. Furthermore, the power of the $Z^{\tilde{e}}_\tau$ statistic increases as $\sigma^2_{\xi_i}$ grows.

Table 8 reports empirical size and power for the $Z^{\tilde{e}}_\tau$ and MQ^*_f statistics when $\rho_i = 1$, $\rho_i = 0.99$ and $\rho_i = 0.9$ – the bandwidth for the Bartlett spectral window used in the computation of the MQ^*_f statistic is set as $J = 4ceil\left[\min[N,T]/100\right]^{1/4}$. We only offer results for the non-parametric version of the MQ statistic $\left(MQ^*_f\right)$ since results for the parametric one (MQ_c^*) were almost equivalent. In this Table, $MQ(3)$ denotes the frequency that the MQ_f^* statistic has detected three common stochastic trends, $MQ(2)$ for the frequency corresponding to two stochastic trends, $MQ(1)$ for one stochastic trend and, finally, $MQ(0)$ denotes the times that the statistic has not detected any stochastic trend. Results in Table 8 indicate that the empirical size of the $Z_\tau^{\tilde{e}}$ statistic approaches the nominal one as T increases, regardless of $\sigma^2_{\xi_i}$. As expected, the power of the test increases as ρ_i moves away from one, and as T gets larger. Note that these conclusions are obtained irrespective of θ. The performance of the MQ_f^* statistic is quite good, since it tends to detect the right number of common stochastic trends most times, while it shows non-trivial power when the common factors are stationary. As for the $Z_\tau^{\tilde{e}}$ statistic, these conclusions are robust to the value of ρ_i. In all, simulations that have been conducted in this paper indicate that the test statistics offer good properties in finite samples when testing the null hypothesis of non-multicointegration, regardless of the test statistic that is used.

6 Production and sales multicointegration relationship in US industries

In this section we illustrate the application of the procedures developed in this paper using monthly inventories and sales series as in Granger and Lee (1989). We use seasonally adjusted series in 1996 constant dollars that cover the period January 1967 to December 1996, i.e. $T = 359$ observations, and are drawn from the US Department of Commerce, Bureau of Economic Analysis database – the analysis cannot be extended from January 1997 onwards due to methodological changes in the definition of these time series. The database offers information on inventories ($Inv_{i,t}$) and sales ($sales_{i,t}$) for 47 US industries that belong to manufacturing, wholesale trade and retail trade sectors. The production ($prod_{i,t}$) series are obtained using the identity in Granger and Lee (1989), i.e. $prod_{i,t} = sales_{i,t} + \Delta Inv_{i,t}$, $i = 1, \ldots, 47$ and $t = 1, \ldots, 359$. Granger and Lee (1989) conclude that empirical results generally support the presence of multicointegration relationships between production and sales in many of the US industries and industrial aggregates. The goal of this section is to extend the previous evidence using the panel data techniques that have been proposed in this paper, which allows the power of the analysis to be increased through the combination of the information of the time and cross-section dimensions.

We have applied the panel data unit root tests in Maddala and Wu (1999) – hereafter, MW statistic – and Im, Pesaran and Shin (2003) – henceforth IPS statistic – to analyze the panel data sets of sales, production and inventories. The order of the autoregressive specification that is used to compute the individual statistics is selected using the t-sig criterion in Ng and Perron (1995) with the maximum order of lags set at twelve as in

Granger and Lee (1989). Panel A of Table 9 indicates that for the case where cross-section independence is assumed, the t-ratio IPS statistic concludes that sales, production and inventories are $I(1)$, while the first difference of inventories is stationary in variance. Mixed evidence is obtained when using the MW statistic. The null hypothesis of non-stationarity cannot be rejected at the 5% level of significance for sales and inventories, while it is rejected for production and change in inventories. Cross-independence might be an unrealistic assumption, especially when analyzing sales and production series of industries that belong to the same economy.

We have accounted for the presence of cross-section dependence in three different ways. First, we have followed the approach in Im, Pesaran and Shin (2003), and have proceeded to remove the cross-section mean – which implies assuming that cross-section dependence is driven by one stationary common factor. Results reported in Panel A of Table 9 indicate that sales, production and inventories are $I(1)$, while change in inventories is $I(0)$. The second way to consider cross-section dependence is based on the computation of the bootstrap distribution for the IPS and MW panel data statistics. In this case, we reach the same conclusions as when individuals were assumed to be independent. Finally, we can base the analysis on the common factor approach in Bai and Ng (2004). Panel A of Table 9 presents the estimated number of factors ($\hat{r}$) determined using the panel BIC information criterion allowing for up to six common factors. Both versions of the MQ statistic indicate that there are non-stationary factors driving production, sales and inventories, while the panel ADF statistic applied to the estimated idiosyncratic disturbance term does not reject the null hypothesis of unit root. Therefore, production, sales and inventories can be characterized as non-stationary panels. The opposite situation is found for change in inventories, since both components are stationary in variance.

So far, the analysis reveals that production and sales are cointegrated with vector (1, -1) since all computation that has been carried out shows that change in inventories is $I(0)$. Furthermore, Bai and Ng (2004) methodology indicates that dependence across US industries in the change in inventories is driven by stationary common factors. Panel B of Table 9 presents the panel data statistics for testing the null hypothesis of non-multicointegration. Assuming that individuals are cross-section independent leads to reject the null hypothesis of non-multicointegration using both $Z_{\hat{\rho}_{NT}}$ and $Z_{\hat{t}_{NT}}$ statistics at the 5% level of significance. Therefore, we find evidence that points to the presence of multicointegration. Nevertheless, these results might be wrong if cross-section dependence is present amongst individuals. First, we have accounted for the presence of cross-section dependence including temporal effects. Thus, working with cross-section demeaned data produces inconclusive results. The $Z_{\hat{\rho}_{NT}}$ statistic finds evidence of panel multicointegration, whereas the $Z_{\hat{t}_{NT}}$ does not.

Previous analyses have revealed that common factors might be driving cross-section dependence. In order to account for this feature, we have computed the statistics in

Section 4 allowing for up to six common factors. Results in panel B of Table 9 investigates whether inventories cointegrates with production and/or sales. In both cases the maximum number of common factors is achieved. Let us focus first on sales and inventories relationship. In this case, there are non-stationary common factors – the number depends both on the information criterion and on the version of the MQ statistic that is used – and panel data ADF statistic computed for the idiosyncratic disturbance term does not reject the null hypothesis of unit root at the 5% level of significance. Therefore, results indicate that there is not cointegration between sales and inventories. When the analysis focuses on production and inventories, we find that, regardless of the information criterion and the version of the MQ statistic that is used, there are two non-stationary in variance common factors. The panel ADF statistic based on the estimated idiosyncratic disturbance term shows that the null hypothesis of unit root can be rejected at the 10% level. Taken at a whole, we have found evidence of mild multicointegration up to the presence of two non-stationary common factors. As mentioned above, note that this situation might be encountered if cross-multicointegration is present between production or inventories for different industries, i.e. if series of production (or inventories) of different industries cointegrate each other. Thus, the non-stationary common factors might be understood as common stochastic trends relating production (or inventories) series in each panel data set.

7 Conclusions

We have proposed test statistics that allow us to analyze the presence of multicointegration relationships in panel data. Although the proposal has focused on multicointegration testing, the statistics can be used to study the presence of cointegration in a wider framework, that is, panel cointegration among $I(2)$ processes when individuals are independent or when cross-section dependence can be modelled by including temporal effects – to the best of our knowledge, this has not been previously considered in the literature. However, this is not the case when cross-section dependence is modelled through approximate factor models. In this situation, the analysis has to be carried out in a $I(1)$ set-up. The use of the common factor approach can be used regardless of whether the cointegration vector of the first level is known or unknown. Simulations conducted in the paper reveal that the statistics show good performance in terms of empirical size and power. Finally, we have illustrated the use of the proposal investigating multicointegration relationships between sales and production of US industries.

References

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

Banerjee, A. (1999). "Panel Data Unit Roots and Cointegration: An Overview", Oxford Bulletin of Economics and Statistics, Special issue, pp. 607–629.

Benerjee,. and Carrion-i-Silvestre, J.L. (2006). *Cointegration in Panel Data with Breaks and Cross-section Dependence*, European Central Bank, Working paper series, num. 591.

Banerjee A., Cockerell, L. and Russell, B. (2001). "An I(2) Analysis of Inflation and the Markup", Journal of Applied Econometrics, Vol. 16, pp. 221-240.

Benerjee, A., Marcellino, M. and Osbat, C. (2005). "Testing for PPP: Should We Use Panel Methods?", Empirical Economics, Vol. 30, pp. 77-91.

Breitung, J. and Pesaran, H. M. (2005). *Unit Roots and Cointegration in Panels*, Working Paper, Department of Economics, University of Cambridge.

Engsted, T., Gonzalo, J., Haldrup, N. (1997). "Testing for Multicointegration", Economic Letters, Vol. 56, pp. 259-266.

Engsted, T. and Haldrup, N. (1999). "Multicointegration in Stock-Flow Models", Oxford Bulletin of Economics and Statistics, Vol. 61, pp. 237-254.

Granger, C. W. J. and Lee, T. (1989). "Multicointegration", In Rhodes, G. F. and Fomby, T. B. (eds.). Advances in Econometrics 8, 71-84. JAI Press, Greenwich.

Haldrup, N. (1994). "The asymptotics of single-equation cointegration regressions with I(1) and I(2) variables", Journal of Econometrics, Vol. 63, pp. 151-181.

Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press.

Herrndorf, N., (1984). "A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables". Anals of Probability 12, pp. 141-153.

Im, K.S., Pesaran, M.H. and Shin, Y. (2003). "Testing for Unit Roots in Heterogeneous Panels", Journal of Econometrics, 115, pp. 53-74.

Juselius, K. (1999): "Price Convergence in the Long Run and the Medium Run. An I(2) Analysis of Six Price Indices", in R.Engle and H. White (ed.). Cointegration, Causality, and Forecasting. Festschrift in Honour of C. W. J. Granger. Oxford University Press.

Juselius, K. (2004): "Inflation, Money Growth, and I(2) Analysis". Working paper. Department of Economics, University of Copenhagen.

Leachman, L., (1996). "New evidence on the Ricardian Equivalence Theorem: A Multicointegration Approach", Applied Economics, 28, pp. 695-704.

Leachman,L., and Francis, B. (2002). "Twin Deficits: Apparition or Reality?", Applied Economics, 34, pp. 1121-1132.

Lee, T., (1992). "Stock flow relationships in US housing construction", Oxford Bulletin of Economics and Statistics, 54, pp. 419-430.

Maddala, G. S. and Wu, S. (1999). "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", Oxford Bulletin of Economics and Statistics, Special Issue, 61, pp. 631-652.

Ng, S. and Perron, P. (1995). "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag", Journal of the American Statistical Association, 90, pp. 268-281.

Pedroni, P. (2004): "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests With an Application to the PPP Hypothesis", Econometric Theory, 20, pp. 597–625.

Phillips, P.C.B. (1991): "Optimal Inference in Cointegrated Systems", Econometrica, 59, 283–306.

Phillips, P.C.B. and Durlauf, S. N. (1986). "Multiple Time Series Regression with Integrated Processes". Review of Economic Studies LIII, pp. 473-495.

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

Siliverstovs, B. (2003). "Multicointegration in US consumption data", Discussion Paper, num. 382. DIW Berlin, German Institute for Economic Research.

Stock, J. H. and Watson, M. W. (1993): "A simple estimator of cointegrating vectors in higher order integrated systems", Econometrica, Vol. 61, pp. 783–820.

Westerlund, J. (2005): "Data Dependent Endogeneity Correction in Cointegrated Panels", Oxford Bulletin of Economics and Statistics, 67, pp. 691-705.

A Mathematical Appendix

A.1 Proof of Theorem 2

We proceed to sketch the proof the case were the deterministic term is driven only by a constant ($m_t = 1$), although the results can be generalized for the other specifications. Thus, we show that our framework can be reduced to the one in Bai and Ng (2004), and Banerjee and Carrion-i-Silvestre (2006), so that interested readers can find further details in those references.

Note that the model given by (11) and (12) can be expressed as:

$$Y_{i,t} = m_t\mu_i + \hat{S}_{i,t}\gamma_i + F_t\pi_i + e_{i,t}.$$

Note that we can write

$$\begin{aligned} \tilde{z}_{i,t} &= z_{i,t} + f_t\pi_i - \tilde{f}_t\tilde{\pi}_i \qquad (19) \\ &= z_{i,t} - v_t H^{-1}\pi_i - \tilde{f}_t d_i, \end{aligned}$$

where $v_t = \tilde{f}_t - f_t H$ and $d_i = \tilde{\pi}_i - H^{-1\prime}\pi_i$. The computation of the partial sum processes of (19) gives:

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\sum_{j=2}^{t} z_{i,j} - T^{-1/2}\sum_{j=2}^{t} v_j H^{-1}\pi_i - T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j d_i. \qquad (20)$$

Let us analyze each element of (20) separately. The left-hand side of (20) is equal to

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} &= T^{-1/2}\sum_{j=2}^{t} M_i\Delta\tilde{e}_{i,j} \qquad (21) \\ &= T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} - T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{e}_i\right]_j, \end{aligned}$$

where $\left[P_i\Delta\tilde{e}_i\right]_j$ denotes the j-th element of the matrix $P_i\Delta\tilde{e}_i$, and $P_i = I_{T-1} - M_i$. The first element on the right of (21) is $T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} = T^{-1/2}\tilde{e}_{i,t} + O_p(1) \Rightarrow \sigma W_i(r)$. The second element on the right hand of (21) tends to $T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{e}_i\right]_j \to 0$, provided that $T^{-1}\Delta\hat{S}_i'\Delta\hat{S}_i = T^{-1}\Delta S_i'\Delta S_i - 2T^{-1}\Delta S_i'\left(c\left(\hat{\alpha}_i - \alpha_i\right) + x_i\left(\hat{\beta}_i - \beta_i\right)\right) + O_p\left(T^{-2}\right) = T^{-1}\Delta S_i'\Delta S_i + O_p\left(T^{-1/2}\right) \to^p Q_{\Delta S_i \Delta S_i}$ and $T^{-1}\Delta\hat{S}_i'\Delta\tilde{e}_i = T^{-1}\Delta S_i'\Delta\tilde{e}_i + O_p\left(T^{-1/2}\right) \to 0$ since we assume that stochastic regressor $S_{i,t}$ is strictly exogenous. Henceforth, we use $\Delta S_{i,t}$ instead of $\Delta\hat{S}_{i,t}$ in the derivations provided that it has been shown that $\Delta\hat{S}_{i,t} = \Delta S_{i,t} + O_p\left(T^{-1/2}\right)$.

These derivations lead us to $T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\tilde{e}_{i,t} + o_p(1)$, since $T^{-1/2}S_{i,t}\left(\Delta S_i'\Delta S_i\right)^{-1}$

$\Delta S_i' \Delta \tilde{e}_i = o_p(1)$ – see Banerjee and Carrion-i-Silvestre (2006) for further details. The same result can be achieved for $T^{-1/2} \sum_{j=2}^{t} z_{i,j}$. This indicates that the presence of stochastic regressors does not have any effect on the partial sum processes. Regarding the term involving $\{v_t\}$ we see from Eq. (A.3) in Bai and Ng (2004) that $T^{-1/2} \sum_{j=2}^{t} v_j = O_p\left(C_{NT}^{-1}\right)$, where $C_{NT} = \min\left\{N^{-1/2}, T^{-1/2}\right\}$. Moreover and as shown in Bai and Ng (2004), the term $d_i = O_p\left(C_{NT}^{-1}\right)$ and $T^{-1/2} \sum_{j=2}^{t} \tilde{f}_j = O_p(1)$, so that $T^{-1/2} \sum_{j=2}^{t} \tilde{z}_{i,j} = T^{-1/2} \sum_{j=2}^{t} z_{i,j} + O_p\left(C_{NT}^{-1}\right)$ .From all these results it follows that

$$DF_{\tilde{e}}^{c}(i) \Rightarrow \frac{\frac{1}{2}\left(W_i(1)^2 - 1\right)}{\left(\int_0^1 W_i(r)^2 dr\right)^{1/2}},$$

that is, the limiting distribution is the same derived in Bai and Ng (2004) for the constant case –see Bai and Ng (2004) for the proof. The same result is found for the ADF test. This implies that the presence of stochastic regressors does not affect the limiting distribution of the statistic.

Let us now deal with the unit root hypothesis testing when there is $r = 1$ common factor. The first difference of the model defines an idempotent matrix M_i that depends on the individual, although it is shown below that the elements that depend on i vanish asymptotically. Thus, note that

$$\begin{aligned} \sum_{j=2}^{t} \tilde{f}_j &= \sum_{j=2}^{t} M_i \Delta \tilde{F}_t \\ &= \tilde{F}_t - (S_{i,t} - S_{i,1})' \left(\Delta S_i' \Delta S_i\right)^{-1} \Delta S_i' \Delta \tilde{F}, \end{aligned} \tag{22}$$

since we define $\tilde{F}_1 = 0$. Note that the first element of (22) is

$$\tilde{F}_t = H(F_t - F_1) + V_t,$$

since $\Delta \tilde{F}_t = H\ \Delta F_t + v_t$ and $V_t = \sum_{j=2}^{t} v_j$. The detrended estimated factor will remove F_1:

$$\tilde{F}_t^* = H\ F_t^* + V_t^*,$$

which can be shown that

$$T^{-1/2} \tilde{F}_t^* = H\ T^{-1/2} F_t^* + O_p\left(C_{NT}^{-1}\right),$$

since $T^{-1/2} V_t^d = O_p\left(C_{NT}^{-1}\right)$ –see Bai and Ng (2004), Lemma B.2. The second term in (22) is $T^{-1/2}(S_{i,t} - S_{i,1})' \left(\Delta S_i' \Delta S_i\right)^{-1} \Delta S_i' \Delta \tilde{F} = o_p(1)$, since $T^{-1} \Delta S_i' \Delta S_i$ converges to the matrix of covariance of ΔS_i and $T^{-1} \Delta S_i' \Delta \tilde{F} = o_p(1)$ by assumption. Therefore, under the null hypothesis the DF statistic converges to

$$DF_{\tilde{F}}^{d} = \frac{T^{-1}\sum_{t=2}^{T}\tilde{F}_{t-1}^{*}\Delta\tilde{F}_{t}}{\left(\tilde{\sigma}_{u}^{2}T^{-2}\sum_{t=2}^{T}\left(\tilde{F}_{t-1}^{*}\right)^{2}\right)^{1/2}} \Rightarrow \frac{\int_{0}^{1}W_{w}^{*}\left(r\right)dW\left(r\right)}{\left(\int_{0}^{1}W_{w}^{*}\left(r\right)^{2}dr\right)^{1/2}}, \tag{23}$$

where $W_{w}^{*}(r)$ denotes the detrended Brownian motion and $\tilde{\sigma}_{w}^{2} \xrightarrow{p} H^{2}\ \sigma_{w}^{2}$. The ADF statistic has the same limiting distribution provided that the order of the autoregressive correction is selected such that $p \to \infty$ and $p^{3}/\min[N,T] \to 0$. The limiting distribution of the test statistic that is used when there is more than one common factor $(r > 1)$ is the same as the one derived in Bai and Ng (2004) for the constant case. We address the reader to Bai and Ng (2004) for the proof of this part of the Theorem, since our framework is equivalent to theirs. Finally, it should be mentioned that all these results hold for the case of non-strictly exogenous regressors once the model given by (11) and (12) has been augmented to include leads and lags of $\Delta\hat{S}_{i,t}$. In practice, the number of leads and lags can be selected using the BIC information criterion. Further details can be found in Banerjee and Carrion-i-Silvestre (2006).

Table 1: Mean and variance for the no deterministics and constant cases

Non-deterministics case

			$T=50$		$T=100$		$T=250$		$T=1,000$	
$Z_{\hat{\rho}_{NT}}$	m_1	m_2	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1
	0	1	-5.654	26.739	-5.373	25.178	-5.213	24.894	-5.093	24.789
	1	1	-10.037	45.394	-9.806	43.489	-9.620	42.471	-9.456	41.596
	2	1	-14.189	65.113	-13.946	59.888	-13.849	59.566	-13.624	57.703
	3	1	-18.157	85.567	-17.968	76.561	-17.899	74.861	-17.684	73.107
	4	1	-21.994	108.195	-21.899	93.358	-21.898	90.320	-21.706	88.715
	0	2	-10.290	49.626	-9.616	43.187	-9.185	41.569	-8.823	40.649
	1	2	-14.626	71.889	-13.989	61.523	-13.603	58.983	-13.214	56.987
	2	2	-18.761	94.887	-18.142	79.496	-17.796	75.546	-17.414	72.969
	3	2	-22.745	120.241	-22.166	97.678	-21.847	91.441	-21.473	88.537
	4	2	-26.596	147.699	-26.094	116.281	-25.887	107.664	-25.507	103.787
$Z_{\hat{t}_{NT}}$	m_1	m_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2
	0	1	-1.492	0.925	-1.377	0.875	-1.302	0.870	-1.250	0.908
	1	1	-2.168	0.860	-2.059	0.778	-1.984	0.727	-1.935	0.706
	2	1	-2.673	0.844	-2.554	0.735	-2.479	0.676	-2.421	0.630
	3	1	-3.089	0.840	-2.963	0.730	-2.879	0.650	-2.816	0.594
	4	1	-3.452	0.849	-3.319	0.725	-3.227	0.638	-3.160	0.577
	0	2	-2.248	0.86	-2.070	0.750	-1.942	0.720	-1.846	0.743
	1	2	-2.748	0.866	-2.579	0.740	-2.462	0.675	-2.376	0.649
	2	2	-3.164	0.860	-2.995	0.729	-2.875	0.656	-2.790	0.606
	3	2	-3.524	0.859	-3.353	0.731	-3.227	0.645	-3.140	0.584
	4	2	-3.841	0.858	-3.671	0.726	-3.547	0.640	-3.454	0.571

Constant case

			$T=50$		$T=100$		$T=250$		$T=1,000$	
$Z_{\hat{\rho}_{NT}}$	m_1	m_2	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1
	0	1	-10.381	44.055	-9.819	38.759	-9.453	36.351	-9.276	35.796
	1	1	-14.259	64.275	-13.707	56.068	-13.329	52.427	-13.126	51.462
	2	1	-18.144	85.923	-17.582	72.963	-17.251	68.480	-17.029	67.094
	3	1	-22.060	110.768	-21.496	90.987	-21.185	84.285	-20.953	82.385
	4	1	-25.878	136.910	-25.359	109.777	-25.125	100.824	-24.824	97.521
	0	2	-15.617	76.183	-14.387	59.183	-13.633	53.168	-13.240	51.150
	1	2	-19.499	100.986	-18.286	77.465	-17.550	69.541	-17.101	66.796
	2	2	-23.363	126.88	-22.163	95.559	-21.462	86.355	-20.980	82.336
	3	2	-27.251	155.956	-26.081	116.434	-25.393	102.557	-24.895	97.005
	4	2	-31.048	188.191	-29.933	137.205	-29.326	119.582	-28.764	111.988
$Z_{\hat{t}_{NT}}$	m_1	m_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2
	0	1	-2.363	0.804	-2.218	0.723	-2.124	0.675	-2.071	0.663
	1	1	-2.780	0.816	-2.641	0.722	-2.544	0.650	-2.489	0.617
	2	1	-3.155	0.825	-3.015	0.720	-2.915	0.637	-2.855	0.594
	3	1	-3.498	0.829	-3.356	0.722	-3.250	0.633	-3.184	0.577
	4	1	-3.806	0.833	-3.664	0.724	-3.556	0.634	-3.480	0.568
	0	2	-2.955	0.821	-2.742	0.702	-2.592	0.631	-2.502	0.608
	1	2	-3.298	0.824	-3.100	0.713	-2.955	0.629	-2.863	0.587
	2	2	-3.615	0.825	-3.426	0.713	-3.282	0.631	-3.187	0.578
	3	2	-3.908	0.828	-3.728	0.715	-3.584	0.631	-3.486	0.565
	4	2	-4.175	0.824	-4.007	0.718	-3.864	0.634	-3.759	0.559

Table 2: Mean and variance for the linear and quadratic time trend cases

Linear trend case

			$T=50$		$T=100$		$T=250$		$T=1,000$	
$Z_{\hat{\rho}_{NT}}$	m_1	m_2	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1
	0	1	-15.978	75.296	-14.712	58.648	-14.118	53.217	-13.706	51.180
	1	1	-19.677	98.989	-18.407	76.561	-17.815	69.352	-17.423	67.149
	2	1	-23.383	124.728	-22.166	94.558	-21.630	86.156	-21.222	82.757
	3	1	-27.090	151.184	-25.947	113.290	-25.483	101.903	-25.062	97.387
	4	1	-30.873	181.851	-29.763	133.997	-29.345	118.118	-28.907	112.825
	0	2	-21.468	120.172	-19.259	83.279	-18.185	71.121	-17.472	66.520
	1	2	-25.244	148.714	-23.055	103.193	-22.003	87.481	-21.281	82.437
	2	2	-29.028	180.650	-26.858	122.963	-25.881	104.295	-25.115	97.409
	3	2	-32.800	214.248	-30.689	144.101	-29.756	120.904	-28.988	111.843
	4	2	-36.592	250.352	-34.532	167.071	-33.635	138.068	-32.839	127.237
$Z_{\hat{t}_{NT}}$	m_1	m_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2
	0	1	-2.972	0.791	-2.759	0.674	-2.634	0.609	-2.548	0.573
	1	1	-3.300	0.804	-3.099	0.694	-2.974	0.619	-2.889	0.573
	2	1	-3.604	0.814	-3.416	0.701	-3.291	0.627	-3.204	0.570
	3	1	-3.889	0.816	-3.711	0.708	-3.588	0.628	-3.497	0.562
	4	1	-4.158	0.823	-3.989	0.715	-3.864	0.629	-3.767	0.559
	0	2	-3.471	0.827	-3.195	0.700	-3.019	0.616	-2.895	0.568
	1	2	-3.754	0.821	-3.501	0.708	-3.332	0.621	-3.210	0.567
	2	2	-4.019	0.813	-3.786	0.708	-3.626	0.626	-3.501	0.560
	3	2	-4.269	0.809	-4.055	0.709	-3.899	0.628	-3.774	0.553
	4	2	-4.507	0.811	-4.309	0.711	-4.157	0.631	-4.028	0.552

Quandratic trend case

			$T=50$		$T=100$		$T=250$		$T=1,000$	
$Z_{\hat{\rho}_{NT}}$	m_1	m_2	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1	Θ_1	Ψ_1
	0	1	-21.887	117.198	-19.737	83.195	-18.556	69.827	-17.895	65.924
	1	1	-25.485	145.813	-23.395	102.971	-22.239	86.462	-21.590	81.553
	2	1	-29.169	177.125	-27.123	123.170	-25.992	103.017	-25.336	96.851
	3	1	-32.826	210.241	-30.888	144.739	-29.814	119.575	-29.145	112.164
	4	1	-36.543	248.073	-34.677	168.437	-33.646	136.887	-32.971	126.942
	0	2	-27.853	182.413	-24.477	115.187	-22.641	88.917	-21.549	81.472
	1	2	-31.606	217.124	-28.277	137.932	-26.448	106.112	-25.343	96.868
	2	2	-35.452	256.907	-32.103	160.992	-30.272	123.135	-29.151	111.829
	3	2	-39.185	296.390	-35.962	185.851	-34.141	140.566	-32.996	126.828
	4	2	-42.988	341.124	-39.786	211.252	-38.004	158.116	-36.848	141.614
$Z_{\hat{t}_{NT}}$	m_1	m_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2	Θ_2	Ψ_2
	0	1	-3.511	0.776	-3.238	0.669	-3.052	0.585	-2.934	0.545
	1	1	-3.774	0.785	-3.529	0.686	-3.351	0.600	-3.236	0.549
	2	1	-4.029	0.788	-3.806	0.692	-3.633	0.611	-3.518	0.549
	3	1	-4.271	0.797	-4.068	0.699	-3.903	0.615	-3.785	0.549
	4	1	-4.501	0.804	-4.317	0.702	-4.156	0.619	-4.037	0.547
	0	2	-3.951	0.816	-3.63	0.709	-3.396	0.604	-3.233	0.550
	1	2	-4.186	0.806	-3.898	0.709	-3.678	0.613	-3.519	0.550
	2	2	-4.412	0.796	-4.153	0.703	-3.942	0.618	-3.786	0.548
	3	2	-4.625	0.792	-4.396	0.704	-4.195	0.620	-4.038	0.546
	4	2	-4.828	0.793	-4.622	0.700	-4.434	0.622	-4.278	0.545

Table 3: Asymptotic and finite sample moments for the ADF idiosyncratic statistics

	$T = 50$		$T = 100$		$T = 250$		$T = 1,000$	
	$\Theta_j^{\tilde{e}}$	$\Psi_j^{\tilde{e}}$	$\Theta_j^{\tilde{e}}$	$\Psi_j^{\tilde{e}}$	$\Theta_j^{\tilde{e}}$	$\Psi_j^{\tilde{e}}$	$\Theta_j^{\tilde{e}}$	$\Psi_j^{\tilde{e}}$
$ADF_{\tilde{e}}^{c}(i)$	-0.401	1.167	-0.410	1.054	-0.420	0.996	-0.421	0.970
$ADF_{\tilde{e}}^{\tau}(i)$	-1.563	0.415	-1.554	0.378	-1.540	0.357	-1.529	0.339

Table 4: Empirical size and power. No deterministics and constant cases

			Non-deterministics case						Constant case					
			Size		Power				Size		Power			
			$\rho_1, \rho_2 = 1$		$\rho_1, \rho_2 = 0.99$		$\rho_1, \rho_2 = 0.9$		$\rho_1, \rho_2 = 1$		$\rho_1, \rho_2 = 0.99$		$\rho_1, \rho_2 = 0.9$	
N	T	σ_ξ^2	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$
20	50	10	0.050	0.044	0.827	0.579	1	0.986	0.090	0.086	0.835	0.649	0.996	0.952
	100	10	0.038	0.035	0.956	0.612	1	1	0.087	0.084	0.994	0.739	1	1
	250	10	0.048	0.044	0.996	0.735	1	1	0.074	0.068	1	0.672	1	1
	1,000	10	0.057	0.044	1	1	1	1	0.088	0.074	1	1	1	1
	50	30	0.045	0.040	1	0.998	1	1	0.088	0.087	1	1	1	1
	100	30	0.052	0.041	1	0.999	1	1	0.067	0.062	1	1	1	1
	250	30	0.046	0.032	1	1	1	1	0.072	0.078	1	1	1	1
	1,000	30	0.051	0.040	1	1	1	1	0.079	0.063	1	1	1	1
	50	50	0.039	0.043	1	1	1	1	0.089	0.085	1	1	1	1
	100	50	0.036	0.032	1	1	1	1	0.067	0.051	1	1	1	1
	250	50	0.047	0.038	1	1	1	1	0.089	0.076	1	1	1	1
	1,000	50	0.042	0.029	1	1	1	1	0.072	0.051	1	1	1	1
40	50	10	0.059	0.047	0.961	0.788	1	1	0.117	0.106	0.985	0.865	1	0.998
	100	10	0.033	0.031	1	0.852	1	1	0.059	0.085	1	0.902	1	1
	250	10	0.043	0.041	1	0.958	1	1	0.080	0.067	1	0.917	1	1
	1,000	10	0.045	0.039	1	1	1	1	0.091	0.085	1	1	1	1
	50	30	0.037	0.035	1	1	1	1	0.101	0.101	1	1	1	1
	100	30	0.055	0.037	1	1	1	1	0.067	0.073	1	1	1	1
	250	30	0.036	0.031	1	1	1	1	0.083	0.074	1	1	1	1
	1,000	30	0.059	0.052	1	1	1	1	0.082	0.077	1	1	1	1
	50	50	0.040	0.030	1	1	1	1	0.107	0.096	1	1	1	1
	100	50	0.052	0.043	1	1	1	1	0.090	0.088	1	1	1	1
	250	50	0.041	0.035	1	1	1	1	0.075	0.074	1	1	1	1
	1,000	50	0.044	0.036	1	1	1	1	0.087	0.078	1	1	1	1

Table 5: Empirical size and power. Linear and quadratic time trend cases

			Linear time trend case						**Quadratic time trend case**					
			Size		Power				Size		Power			
			$\rho_1, \rho_2 = 1$		$\rho_1, \rho_2 = 0.99$		$\rho_1, \rho_2 = 0.9$		$\rho_1, \rho_2 = 1$		$\rho_1, \rho_2 = 0.99$		$\rho_1, \rho_2 = 0.9$	
N	T	σ_ξ^2	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$
20	50	10	0.074	0.063	0.533	0.459	0.843	0.760	0.098	0.094	0.310	0.313	0.452	0.469
	100	10	0.083	0.080	0.980	0.732	1	1	0.104	0.107	0.976	0.790	1	0.997
	250	10	0.077	0.071	0.999	0.603	1	1	0.098	0.089	1	0.642	1	1
	1,000	10	0.110	0.103	1	1	1	1	0.092	0.089	1	1	1	1
	50	30	0.103	0.088	1	0.999	1	1	0.115	0.120	0.998	0.998	0.999	0.996
	100	30	0.100	0.094	1	1	1	1	0.086	0.059	1	1	1	1
	250	30	0.082	0.074	1	1	1	1	0.094	0.092	1	1	1	1
	1,000	30	0.073	0.071	1	1	1	1	0.081	0.079	1	1	1	1
	50	50	0.100	0.093	1	1	1	1	0.124	0.112	1	1	1	0.999
	100	50	0.090	0.066	1	1	1	1	0.081	0.074	1	1	1	1
	250	50	0.077	0.072	1	1	1	1	0.090	0.086	1	1	1	1
	1,000	50	0.063	0.059	1	1	1	1	0.070	0.066	1	1	1	1
40	50	10	0.092	0.076	0.805	0.692	0.974	0.946	0.117	0.124	0.526	0.518	0.731	0.729
	100	10	0.124	0.111	1	0.928	1	1	0.114	0.112	1	0.947	1	1
	250	10	0.108	0.102	1	0.832	1	1	0.089	0.098	1	0.858	1	1
	1,000	10	0.111	0.104	1	1	1	1	0.085	0.076	1	1	1	1
	50	30	0.109	0.119	1	1	1	1	0.151	0.146	1	1	1	1
	100	30	0.101	0.094	1	1	1	1	0.113	0.101	1	1	1	1
	250	30	0.100	0.092	1	1	1	1	0.110	0.112	1	1	1	1
	1,000	30	0.088	0.084	1	1	1	1	0.089	0.082	1	1	1	1
	50	50	0.110	0.121	1	1	1	1	0.142	0.117	1	1	1	1
	100	50	0.090	0.091	1	1	1	1	0.127	0.118	1	1	1	1
	250	50	0.082	0.074	1	1	1	1	0.111	0.124	1	1	1	1
	1,000	50	0.089	0.085	1	1	1	1	0.083	0.075	1	1	1	1

Table 6: Empirical size and power with heterogeneous individuals

Panel A: Empirical size

		No deterministics		Constant		Linear trend		Quadratic trend	
N	T	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$
20	50	0.055	0.051	0.086	0.084	0.081	0.088	0.099	0.106
	100	0.029	0.027	0.075	0.067	0.091	0.086	0.081	0.079
	250	0.036	0.026	0.071	0.060	0.073	0.070	0.091	0.082
	1,000	0.056	0.041	0.085	0.077	0.103	0.099	0.094	0.094
N	T	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$
40	50	0.039	0.032	0.112	0.100	0.110	0.102	0.109	0.104
	100	0.045	0.035	0.083	0.074	0.101	0.082	0.091	0.075
	250	0.031	0.030	0.078	0.070	0.104	0.090	0.088	0.086
	1,000	0.053	0.048	0.087	0.083	0.119	0.106	0.098	0.095

Panel B: Empirical power

		No deterministics		Constant		Linear trend		Quadratic trend	
N	T	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$
20	50	1	1	0.175	0.181	1	1	1	0.988
	100	1	1	0.686	0.627	1	1	1	1
	250	1	1	1	1	1	1	1	1
	1,000	1	1	1	1	1	1	1	1
N	T	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$
40	50	0.353	0.396	0.196	0.251	1	1	1	0.985
	100	0.968	0.968	0.878	0.809	1	1	1	1
	250	1	1	1	1	1	1	1	1
	1,000	1	1	1	1	1	1	1	1

Table 7: Empirical size and power for the $Z_{\tau}^{\tilde{e}}$ and ADF_F^* statistics. One known common factor, $N = 40$ individuals

T	$\sigma_{\xi_i}^2$	θ	ρ_i	$Z_{\tau}^{\tilde{e}}$	ADF_F^*	ρ_i	$Z_{\tau}^{\tilde{e}}$	ADF_F^*	ρ_i	$Z_{\tau}^{\tilde{e}}$	ADF_F^*
50	3	1	1	0.114	0.076	0.99	0.123	0.057	0.9	0.747	0.081
100	3	1	1	0.056	0.053	0.99	0.086	0.064	0.9	0.999	0.056
250	3	1	1	0.046	0.055	0.99	0.270	0.065	0.9	1	0.044
50	3	0.95	1	0.127	0.099	0.99	0.130	0.078	0.9	0.738	0.075
100	3	0.95	1	0.062	0.095	0.99	0.084	0.087	0.9	1	0.098
250	3	0.95	1	0.046	0.273	0.99	0.255	0.307	0.9	1	0.278
50	3	0.9	1	0.117	0.115	0.99	0.120	0.094	0.9	0.709	0.097
100	3	0.9	1	0.060	0.223	0.99	0.069	0.186	0.9	1	0.221
250	3	0.9	1	0.061	0.844	0.99	0.262	0.856	0.9	1	0.83
50	5	1	1	0.127	0.074	0.99	0.129	0.064	0.9	0.803	0.076
100	5	1	1	0.071	0.042	0.99	0.101	0.064	0.9	1	0.051
250	5	1	1	0.054	0.051	0.99	0.365	0.061	0.9	1	0.056
50	5	0.95	1	0.132	0.072	0.99	0.157	0.069	0.9	0.756	0.079
100	5	0.95	1	0.070	0.106	0.99	0.108	0.090	0.9	1	0.100
250	5	0.95	1	0.073	0.286	0.99	0.341	0.273	0.9	1	0.266
50	5	0.9	1	0.112	0.109	0.99	0.105	0.111	0.9	0.789	0.106
100	5	0.9	1	0.086	0.211	0.99	0.107	0.194	0.9	1	0.205
250	5	0.9	1	0.060	0.810	0.99	0.337	0.833	0.9	1	0.861

Table 8: Empirical size and power for the $Z_\tau^{\tilde{e}}$ and MQ_f statistics. Three unknown common factors, $N = 40$ individuals

			$\rho_i = 1$					$\rho_i = 0.99$					$\rho_i = 0.9$				
T	$\sigma^2_{\xi_i}$	θ	$Z_\tau^{\tilde{e}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$Z_\tau^{\tilde{e}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$Z_\tau^{\tilde{e}}$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	3	1	0.16	0.00	0.00	0.01	0.98	0.16	0.00	0.00	0.01	0.98	0.68	0.00	0.00	0.01	0.98
100	3	1	0.09	0.01	0.04	0.15	0.80	0.12	0.00	0.05	0.17	0.77	1.00	0.01	0.05	0.19	0.74
250	3	1	0.07	0.01	0.05	0.21	0.73	0.29	0.01	0.05	0.19	0.74	1.00	0.02	0.08	0.28	0.62
50	3	0.95	0.14	0.00	0.00	0.02	0.97	0.16	0.00	0.00	0.02	0.97	0.67	0.00	0.01	0.02	0.97
100	3	0.95	0.09	0.01	0.04	0.19	0.76	0.10	0.02	0.05	0.18	0.75	1.00	0.02	0.03	0.15	0.81
250	3	0.95	0.06	0.17	0.10	0.25	0.47	0.29	0.15	0.09	0.26	0.51	1.00	0.15	0.09	0.29	0.47
50	3	0.9	0.13	0.00	0.01	0.03	0.96	0.14	0.00	0.00	0.02	0.97	0.63	0.00	0.01	0.03	0.95
100	3	0.9	0.06	0.07	0.05	0.20	0.67	0.09	0.07	0.06	0.19	0.68	1.00	0.05	0.05	0.20	0.71
250	3	0.9	0.05	0.70	0.05	0.14	0.11	0.25	0.72	0.04	0.13	0.11	1.00	0.76	0.06	0.10	0.07
50	5	1	0.14	0.00	0.00	0.06	0.94	0.16	0.00	0.00	0.03	0.96	0.75	0.00	0.00	0.09	0.90
100	5	1	0.08	0.01	0.02	0.15	0.83	0.11	0.00	0.03	0.15	0.82	1.00	0.00	0.02	0.12	0.86
250	5	1	0.06	0.01	0.04	0.17	0.79	0.35	0.01	0.03	0.19	0.77	1.00	0.01	0.04	0.23	0.72
50	5	0.95	0.14	0.00	0.00	0.03	0.96	0.15	0.00	0.00	0.02	0.98	0.74	0.00	0.00	0.02	0.98
100	5	0.95	0.05	0.01	0.02	0.14	0.82	0.11	0.02	0.04	0.14	0.81	1.00	0.01	0.04	0.16	0.79
250	5	0.95	0.06	0.14	0.09	0.26	0.51	0.34	0.15	0.08	0.28	0.49	1.00	0.13	0.10	0.26	0.51
50	5	0.9	0.13	0.00	0.00	0.03	0.97	0.12	0.00	0.01	0.04	0.95	0.73	0.00	0.00	0.07	0.93
100	5	0.9	0.05	0.05	0.07	0.23	0.65	0.09	0.05	0.05	0.20	0.71	1.00	0.05	0.05	0.18	0.71
250	5	0.9	0.06	0.66	0.04	0.13	0.16	0.32	0.65	0.03	0.13	0.18	1.00	0.78	0.05	0.09	0.08

Table 9: Panel data unit root test. IPS, MW and Bai-Ng statistics

	Raw data				CS demeaned				Bootstrap IPS test		Bootstrap MW test	
	IPS	p-val	MW	p-val	IPS	p-val	MW	p-val	5%	10%	90%	95%
$Sales$	0.33	0.63	115.96	0.06	1.59	0.94	89.03	0.63	-2.86	-2.55	97.46	102.27
$Prod$	-0.40	0.34	126.47	0.01	0.24	0.59	110.97	0.11	-2.41	-2.00	96.04	101.36
$Inven$	0.38	0.65	83.48	0.77	1.18	0.88	70.89	0.96	-2.81	-2.45	98.92	104.07
$\Delta Inven$	-36.03	0.00	1274.32	0.00	-37.07	0.00	1153.12	0.00	-2.42	-1.70	109.91	116.43

Bai-Ng panel data unit root test

	$\hat{r}$	MQ_f^*	$S.t.$	MQ_c^*	$S.t.$	$ADF-idio$	p-val
$Sales$	6	-12.18	2	-19.83	2	1.20	0.89
$Prod$	6	-18.64	2	-24.56	2	-0.87	0.19
$Inven$	6	-32.13	5	-53.18	6	2.73	1.00
$\Delta Inven$	6	-266.81	0	-116.23	0	-23.43	0.00

Panel B: Panel data multicointegration testing

	$Z_{\hat{\rho}_{NT}}$	$Z_{\hat{t}_{NT}}$
Independent	-14.72	-2.30
CS demeaned	-12.88	-1.31

	$\hat{r}$	MQ_f^*	$S.t.$	MQ_c^*	$S.t.$	$ADF-idio$	p-val
Sales vs Inventories	6	-41.72	4	-40.10	3	1.34	0.91
Production vs Inventories	6	-20.01	2	-23.43	2	-1.38	0.08

MQ_f^* denotes the non-parametric version of the MQ test, while MQ_c^* is the parametric one. $S.t.$ indicates the number of common stochastic trends that are present in the estimated common factors ($\hat{r}$). $ADF-idio$ denotes the ADF panel data unit root tests applied on the estimated idiosyncratic disturbance terms.