Testing Panel Cointegration with Unobservable Dynamic Common Factors

Bai, Jushan and Carrion-i-Silvestre, Josep Lluis

July 2009

Testing Panel Cointegration with Unobservable Dynamic Common Factors*

Jushan Bai[†] Josep Lluís Carrion-i-Silvestre[‡]

July 18, 2009

Abstract

The paper proposes statistics to test the null hypothesis of no cointegration in panel data when common factors drive the cross-section dependence. We consider both the case in which regressors are independent of the common factors and the case in which regressors are affected by the common factors. The test statistics are shown to have limiting distributions independent of the common factors, making it possible to pool the individual statistics. Simulations indicates that the proposed procedures have good finite sample performance.

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

JEL codes: C12, C22

---

*This paper was presented at the 13th International Conference on Panel Data (Cambridge, 2006), 61th European Meeting of the Econometric Society (Vienna, 2006), XXXI Simposio de Análisis Económico (Oviedo, 2006), Factor Models in Theory and Practice (Florence, 2006), New Developments in Macroeconomic Modelling and Growth Dynamics conference (Faro, 2006), and 20 Years of Cointegration: Theory and Practice in Prospect and Retrospect (Rotterdam, 2007). Financial support for this research is provided through NSF grants SES0424540 and SES0551275 (Bai) and through the Spanish Ministerio de Ciencia y Tecnología grant ECO2008-06241/ECON (Carrion-i-Silvestre).

[†]Economics Department, Columbia University. 420 West 118th Street, New York. E-mail: jb3064@columbia.edu

[‡]AQR-IREA Research Group. Department of Econometrics, Statistics and Spanish Economy. University of Barcelona. Av. Diagonal, 690. 08034-Barcelona. E-mail: carrion@ub.edu

1 Introduction

The literature on panel cointegration has experienced a huge development since the 90's. Earlier analysis assumed cross-section independence when designing the inference procedures.[1] This assumption is convenient because it allows the application of the central limit theorem over the cross sections to achieve asymptotic normality for the underlying statistics. A key feature of cointegration is co-movement of economic variables, or existence of common stochastic trends. While cross-section independence allows within-unit common stochastic trends, it cannot capture the cross-section (global) common stochastic trends, thereby limiting the model's applicability. To tackle this problem, we follow a similar framework as in Bai and Ng (2004) and Bai (2009), who use the approximate common factor model to characterize common shocks and common stochastic trends; also see Moon and Perron (2004). We consider a model of the form:

$$\begin{aligned} Y_{i,t} &= \mu_i + \gamma_i\, t + X_{i,t}'\beta_i + F_t'\lambda_i + e_{i,t} \\ i &= 1, 2, ..., N; \quad t = 1, 2, ..., T \end{aligned}$$

where $\mu_i + \gamma_i\, t$ describes the deterministic component, $X_{i,t}$ is a vector of observable I(1) regressors, F_t is a vector of unobservable common shocks whose impact varies over cross sections via λ_i. The $e_{i,t}$ are the idiosyncratic errors.

We refer to F_t, when it is I(1), as unobservable cross-section common stochastic trend. When $e_{i,t}$ are I(0), then $Y_{i,t}$, $X_{i,t}$, F_t are cointegrated, even though $Y_{i,t}$ and $X_{i,t}$ are not cointegrated. So this paper considers cointegration between Y_{it} and $X_{i,t}$ up to a small number of unobservable common stochastic trends. When both $e_{i,t}$ and F_t are I(0), $Y_{i,t}$ and $X_{i,t}$ are cointegrated. In this case, we may regard F_t as common shocks, which capture the cross-section correlations.

A similar framework has been adopted by a number of recent panel cointegration studies. Banerjee and Carrion-i-Silvestre (2006), Gengenbach, Palm and Urbain (2006), Westerlund (2008), and Westerlund and Edgerton (2008) extend the residual-based Engle-Granger approach to panel data with common factors. Gengenbach, Urbain and Westerlund (2008) focus on the error correction model with common factors. Groen and Kleinberger (2003) and Breitung (2005) use the vector error correction specification to test the presence of cointegration, where dependence is considered through the residual covariance matrix. Finally, Carrion-i-Silvestre and Surdeanu (2009) propose a panel cointegration rank test with global stochastic trends. A recent survey of the field is provided by Baltagi (2008), Breitung and Pesaran (2008), and Banerjee and Wagner (2009).

Panel cointegration with cross-section dependence has important empirical applications. Gengenbach, Palm and Urbain (2005) test the PPP hypothesis using panel cointegration techniques that allow for common factors. Banerjee and Carrion-i-Silvestre (2006) analyze the long-run exchange rate pass-through for the euro area. Constantini and Lupi (2006) estimate the long-run relationship between Italian regional unemployment rates. Westerlund (2008) analyzes the Fisher effect, while Gengenbach, Urbain and Westerlund (2009) examine both the Fisher effect and the monetary exchange rates. Moverover, Banerjee and Wagner (2009) study the environmental Kuznets curve; Holly, Pesaran and Yamagata (2009) examine the long-run relationship between housing prices and incomes, and Carrion-i-Silvestre and Surdeanu (2009) focus on money demand.[2]

[1] See, e.g., McCoskey and Kao (1998), Kao (1999), Pedroni (2000, 2004) and Larsson, Lyhagen and Löthgren (2001).

[2] There is also a related literature using common factors when estimating panel cointegration relationships. For instance, Pedroni (2007) estimates an augmented neoclassical Solow growth model, and Tosetti and Moscone (2007) for a health-care demand model, using the approach in Pesaran (2006). Westerlund (2007) estimates a panel model based on the forward rate unbiasedness hypothesis and Costantini and Destefanis (2009) estimate the Italian regional production functions, using the approach in Bai and Kao (2006).

Few of the above studies consider the case where the common factors are allowed to be correlated with stochastic regressors. Correlation between the common factors and I(1) regressors arises in practice since common factors that affect the endogenous variable, in general, also affect the stochastic regressors. Not only do we want to control for cross-sectional correlation, but also we want to determine if the unobserved component F_t is integrated. If F_t is integrated, then y_t and x_{it} are not cointegrated directly, but may be cointegrated up to a small number of cross-sectional unobserved stochastic trends. Our analysis permits F_t to contain both I(1) and I(0) components. We do not regard cross-section dependence as nuisance or a burden on inference, but rather a structure that is potentially informative about the way in which the panel data are linked. A further difference between our framework from the previous panel cointegration studies is the use of the modified Sargan-Bhargava (MSB) statistic. The MSB statistic possesses some optimality properties within the class of tests that are invariant to heterogeneous trends, as is shown by Ploberger and Phillips (2004). Our analysis complements the analysis in Bai and Kao (2006), and Bai, Kao and Ng (2009), who assume the existence of cointegration.[3]

Under the null hypothesis of no cointegration, the disturbances e_{it} are I(1). To consistently estimate the factors and residuals, we follow Bai and Ng (2004) by taking the first order difference of the data. After estimating the factors and residuals from the differenced data, we re-cumulate them and construct test statistics based on these estimated quantities. This procedure has notable advantages. The individual statistics do not depend on the dimension of the stochastic regressors. Therefore, there is no need for many tables of critical values. Nor do the individual statistics depend on the common factors. This implies that the individual statistics are cross-sectionally independent as long as the idiosyncratic errors are cross-sectionally independent. This allows pooled statistics to be constructed.

We find it useful to distinguish two setups: one having $X_{i,t}$ and F_t to be independent, and the other having $X_{i,t}$ and F_t to be correlated. The first setup permits a simpler procedure when constructing the test statistics. For the second setup, an iterated procedure is needed to consistently estimate the slope parameters and the common factors in order to construct the test statistics.

The paper is organized as follows. Section 2 describes the model and the underlying assumptions. We distinguish two situations depending on whether the stochastic regressors are strictly exogenous or non-strictly exogenous with respect to the idiosyncratic errors. Limiting distributions of the test statistics are derived in this section. Section 3 considers the case in which regressors are correlated with the unobservable common factors. Section 4 studies pooled test statistics. Section 5 conducts Monte Carlo simulations to investigate the finite sample properties of proposed statistics. Section 6 concludes. All proofs are collected in the appendix.

2 Heterogeneous panel cointegration

Let $\{Y_{i,t}\}$ be a stochastic process with DGP expressed as:

$$Y_{i,t} = \mu_i + \gamma_i\, t + X_{i,t}'\beta_i + u_{i,t} \tag{1}$$

$t = 1, \ldots, T$, $i = 1, \ldots, N$, where $X_{i,t}$ is a $p \times 1$ vector of I(1) regressors such that

$$(I - L)\, X_{i,t} = G_i\,(L)\, v_{i,t} \tag{2}$$

[3] Related approaches can be found in Pesaran (2006) and Kapetanios, Pesaran and Yamagata (2006), who approximate the common factors using cross-section means of the variables in the model.

and the disturbances $u_{i,t}$ have a factor structure such that

$$\begin{aligned} u_{i,t} &= F_t'\lambda_i + e_{i,t}, & (3)\\ (I-L)\,F_t &= C\,(L)\,w_t & (4)\\ (1-\rho_i L)\,e_{i,t} &= d_i\,(L)\,\varepsilon_{i,t}; & (5) \end{aligned}$$

with F_t a vector of $(r \times 1)$ unobservable dynamic factors and λ_i the vector of loadings. We assume $C\,(L) = \sum_{j=0}^{\infty} C_j L^j$. Despite the operator $(1-L)$ in equation (4), 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 of full rank, then some components of F_t are I(1) and some are I(0). Regarding the deterministic component $\mu_i + \gamma_i\,t$, we consider two specifications: (1) the intercept only model ($\gamma_i = 0$ for all i) and (2) the general linear trend model (without imposing $\gamma_i = 0$). These two cases are separately considered as the resulting test statistics have different limiting distributions. Our analysis is based on similar assumptions introduced in Bai and Ng (2004). Let $S < \infty$ be a generic positive number, not depending on T and N:

Assumption A: (i) $E\,\|\lambda_i\|^4 \leq S$, (ii) $\frac{1}{N}\sum_{i=1}^{N} \lambda_i\lambda_i' \xrightarrow{p} \Sigma_\Lambda$, a $(r \times r)$ positive definite matrix.

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

Assumption C: (i) for each i, $\varepsilon_{i,t} \sim iid\left(0, \sigma_i^2\right)$, $E\,|\varepsilon_{i,t}|^8 \leq S$, $\sum_{j=0}^{\infty} j\,|d_{i,j}| < S$, $\omega_i^2 = d_i\,(1)^2\,\sigma_i^2 > 0$; (ii) ε_{it} are independent across i.

Assumption D: (i) For each i, $v_{i,t} \sim iid\,(0, \Sigma_v)$, $E\,\|v_{i,t}\|^4 \leq S$, and (ii) $Var\,(\Delta X_{i,t}) = \sum_{j=0}^{\infty} G_{i,j}\Sigma_v G_{i,j}' > 0$, (iii) $\sum_{j=0}^{\infty} j\,\|G_{i,j}\| < S$; and (iv) $G_i\,(1)$ has full rank.

Assumption E: The errors $\{\varepsilon_{i,t}\}$, $\{w_t\}$, and the loadings $\{\lambda_i\}$ are mutually independent.

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

Assumptions A and B imply r factors, they are necessary for consistent estimation of factor loadings and the factors (up to a rotation). Assumption B specifies the short-run and long-run variances of ΔF_t. The short-run variance is positive definite (implying r factors), but the long-run variance can be of reduced rank in order to accommodate linear combinations of $I\,(1)$ factors to be stationary. Assumption C(i) allows for some weak serial correlation in $(1-\rho_i L)\,e_{i,t}$, whereas C(ii) assumes cross-section independence, a useful assumption when pooling individual test statistics. Assumption D gives conditions on the first order difference of the stochastic regressors. Assumption E assumes the unobservable common factors are independent of the regression errors, and of the factor loadings, a standard assumption for factor models. Assumption F is for initial conditions.

In the next two subsections, we consider two situations depending on whether stochastic regressors are strictly exogenous regressors or non-strictly exogenous regressors. The first case is quite simple, it is shown that the limiting distribution of statistics does not depend on the stochastic regressors $X_{i,t}$ nor on F_t. With non-strictly exogenous regressors, the procedure needs to be modified in order to achieve the same result.

2.1 Strictly exogenous regressors

In this section, we assume that $X_{i,t}$ is independent of $u_{i,t} = F_t'\lambda_i + e_{i,t}$. This assumption will be relaxed in the next section. Under this assumption, a simpler estimation procedure (without iteration) is sufficient. The proof requires, for the case of intercept model ($\gamma_i = 0$ for all i),

$$\frac{1}{T}\sum_{t=1}^{T} \Delta X_{i,t}\Delta e_{i,t} = O_p(T^{-1/2}), \quad \frac{1}{T}\sum_{t=1}^{T} \Delta X_{i,t}\Delta F_t' = O_p(T^{-1/2}). \tag{6}$$

For the case of linear trends, the requirements become

$$\frac{1}{T}\sum_{t=1}^{T} (\Delta X_{i,t} - \overline{\Delta X}_i)(\Delta F_t - \overline{\Delta F})' = O_p(T^{-1/2}), \tag{7}$$

and a similar expression with $e_{i,t}$ in place of F_t. These requirements are met for strictly exogenous regressors $X_{i,t}$, as explained below. We make this assumption explicit:

Assumption G: $X_{i,t}$ is independent of $(e_{i,s}, F_s)$ for all t and s.

The intercept only case and the linear trend case will be studied separately. The former requires that the I(1) regressors $X_{i,t}$ and the common trends F_t have no drifts. The latter allows drift in $X_{i,t}$ and in F_t. The reason is that for the intercept case, we need $T^{-1/2}X_{i,t} = O_p(1)$ and $T^{-1/2}F_t = O_p(1)$. This cannot be true if drifts exist. When linear trend is included in the estimation, the model is invariant to whether the I(1) regressors have drifts. In this case, the proof of our results needs $T^{-1/2}(X_{i,t} - \frac{t}{T}X_{i,T}) = O_p(1)$ and $T^{-1/2}(F_t - \frac{t}{T}F_T) = O_p(1)$, but these are true even if drifts exist in $X_{i,t}$ and F_t.

2.1.1 Intercept only case

This case assumes no linear trend in the model so that $\gamma_i = 0$ for all i

$$Y_{i,t} = \mu_i + X_{i,t}'\beta_i + F_t'\lambda_i + e_{i,t}. \tag{8}$$

We also assume $X_{i,t}$ and F_t have no drifts. If these series do exhibit drifts, test statistics in the next subsection should be used as they are invariant to drifts. Differencing the above model, we have

$$\Delta Y_{i,t} = \Delta X_{i,t}'\beta_i + \Delta F_t\lambda_i + \Delta e_{i,t}.$$

By the driftless assumption for X_{it} and F_t, $E(\Delta X_{it}) = 0$ and $E(\Delta F_t) = 0$. Since they are also independent, it follows that (6) holds. The above equation can be written as, in vector notation,

$$\Delta Y_i = \Delta X_i\beta_i + \Delta F\lambda_i + \Delta e_i,$$

where

$$\Delta Y_i = \begin{bmatrix} \Delta Y_{i,2} \\ \Delta Y_{i,3} \\ \vdots \\ \Delta Y_{i,T} \end{bmatrix}, \Delta X_i = \begin{bmatrix} \Delta X_{i,2}' \\ \Delta X_{i,3}' \\ \vdots \\ \Delta X_{i,T}' \end{bmatrix}, \Delta F = \begin{bmatrix} \Delta F_2' \\ \Delta F_3' \\ \vdots \\ \Delta F_T' \end{bmatrix},$$

and Δe_i is defined similarly as ΔY_i. We further introduce

$$y_{it} = \Delta Y_{i,t}, \quad x_{it} = \Delta X_{i,t}, \quad f_t = \Delta F_t.$$

The differenced model can be rewritten as

$$y_i = x_i\beta_i + f\lambda_i + \Delta e_i. \tag{9}$$

Define $(T-1)\times(T-1)$ projection matrix as

$$M_i = I_{T-1} - x_i(x_i'x_i)^{-1}x_i' = I_{T-1} - P_i.$$

Left multiplying M_i on each side of (9)

$$\begin{aligned} M_i y_i &= M_i f \lambda_i + M_i \Delta e_i \\ &= f \lambda_i - P_i f \lambda_i + M_i \Delta e_i, \end{aligned}$$

which can be rewritten as

$$y_i^* = f \lambda_i + z_i, \tag{10}$$

where

$$y_i^* = M_i y_i, \quad z_i = M_i \Delta e_i - P_i f \lambda_i. \tag{11}$$

Therefore, (10) is a factor model with new observable variables y_i^*. In the appendix, we show that

$$z_{i,t} = \Delta e_{i,t} + \Delta X_{i,t} O_p(T^{-1/2}),$$

and furthermore,

$$T^{-1/2} \sum_{s=1}^{t} z_{i,s} = T^{-1/2} e_{i,t} + O_p(T^{-1/2}).$$

Thus under the null hypothesis of no cointegration, we have

$$T^{-1/2} \sum_{s=1}^{t} z_{i,s} \to \sigma_i W_i(r),$$

where $W_i(r)$ denotes a standard Brownian motion.

In order to use $z_{i,t}$ to form test statistics, we must have an estimate for $z_{i,t}$. This requires an estimate for f and $\Lambda = (\lambda_1, ..., \lambda_N)'$. The estimation of the common factors and factor loadings can be done as in Bai and Ng (2004) using principal components. Let

$$y^* = (y_1^*, y_2^*, ..., y_N^*)$$

be the $(T-1) \times N$ data matrix. The estimated principal component of $f = (f_2, f_3, \ldots, f_T)$, denoted as $\tilde{f}$, is $\sqrt{T-1}$ times the r eigenvectors corresponding to the first r largest eigenvalues of the $(T-1) \times (T-1)$ matrix $y^* y^{*\prime}$, under the normalization $\tilde{f}\tilde{f}' / (T-1) = I_r$. The estimated loading matrix is $\tilde{\Lambda} = y^{*\prime} \tilde{f} / (T-1)$. Therefore, the estimated residuals are defined as

$$\tilde{z}_{i,t} = y_{i,t}^* - \tilde{f}_t' \tilde{\lambda}_i. \tag{12}$$

We can estimate the idiosyncratic disturbance terms through cumulation, *i.e.*

$$\tilde{e}_{i,t} = \sum_{s=2}^{t} \tilde{z}_{i,s}.$$

The null hypothesis of no cointegration is based on $\tilde{e}_{i,t}$ in place of unobservable $e_{i,t}$.

We use the modified Sargan-Bhargava (MSB) statistic proposed in Stock (1999) to test the null hypothesis. As mentioned in the introduction, this statistic possesses some optimality properties within the class of tests that are invariant to heterogeneous trends as shown by Ploberger and Phillips (2004). The MSB statistic on the idiosyncratic disturbance terms is given by

$$MSB_{\tilde{e}}(i) = \frac{T^{-2} \sum_{t=1}^{T} \tilde{e}_{i,t-1}^2}{\tilde{\sigma}_i^2}, \tag{13}$$

where $\tilde{\sigma}_i^2$ is an estimation of the long-run variance of $\{\Delta e_{i,t}\}$. Here we suggest estimating the long-run variance as in Ng and Perron (2001)

$$\tilde{\sigma}_i^2 = \frac{\tilde{\sigma}_{k,i}^2}{\left(1 - \tilde{\phi}(1)\right)^2}, \tag{14}$$

with $\tilde{\phi}(1) = \sum_{j=1}^{k} \tilde{\phi}_j$ and $\tilde{\sigma}_{k,i}^2 = (T-k)^{-1} \sum_{t=k+1}^{T} \tilde{v}_{i,t}^2$, where $\tilde{\phi}_j$ and $\{\tilde{v}_{i,t}\}$ are obtained from the OLS estimation of

$$\Delta \tilde{e}_{i,t} = \phi_0 \tilde{e}_{i,t-1} + \sum_{j=1}^{k} \phi_j \Delta \tilde{e}_{i,t-j} + v_{i,t} \tag{15}$$

where the lag order k is specified in the theorem below. An alternative estimator for σ_i^2 is that of Newey-West based on the residuals $\tilde{e}_{it} - \hat{\rho}_i \tilde{e}_{i,t-1}$, where $\hat{\rho}_i$ is obtained from regressing $\tilde{e}_{it}$ on $\tilde{e}_{i,t-1}$.

We can also test whether the common factor F_t is I(1). Define

$$\tilde{F}_t = \sum_{s=2}^{t} \tilde{f}_s.$$

When there is one common factor, *i.e.* $r = 1$, we construct the unit root test statistic as in (13), using $\tilde{F}_t$ instead of $\tilde{e}_{i,t}$, that is,

$$MSB_{\tilde{F}} = \frac{T^{-2} \sum_{t=1}^{T} \tilde{F}_{t-1}^2}{\tilde{\sigma}_f^2}, \tag{16}$$

where the long run variance $\left(\tilde{\sigma}_f^2\right)$ can be estimated as described above.

When the number of common factors is $r > 1$ we suggest to use the modified Q statistic – hereafter MQ statistic – in Bai and Ng (2004). Let $\tilde{F}_t^c = \tilde{F}_t - \bar{\tilde{F}}$ denote the demeaned common factors. Start with $q = r$ and proceed in three stages:

1. Let $\tilde{\alpha}_\perp$ be the q eigenvectors associated with the q largest eigenvalues of $T^{-2} \sum_{t=2}^{T} \tilde{F}_t^c \tilde{F}_t^{c\prime}$.
2. Let $\tilde{Y}_t^c = \tilde{\alpha}_\perp \tilde{F}_t^c$, from which we can define two statistics – the first one $(MQ_c^c(q))$ accounts for autocorrelation in a non-parametric way, while the second one $\left(MQ_f^c(q)\right)$ in a parametric way:
   (a) Let $K(j) = 1 - j/(J+1)$, $j = 0, 1, 2, \ldots, J$:
      i. Let $\tilde{\xi}_t^c$ be the residuals from estimating a first-order VAR in $\tilde{Y}_t^c$, and let
      $$\tilde{\Sigma}_1^c = \sum_{j=1}^{J} K(j) \left(T^{-1} \sum_{t=2}^{T} \tilde{\xi}_t^c \tilde{\xi}_t^{c\prime} \right).$$
      ii. Let $\tilde{v}_c^c(q) = \frac{1}{2} \left[\sum_{t=2}^{T} \left(\tilde{Y}_t^c \tilde{Y}_{t-1}^{c\prime} + \tilde{Y}_{t-1}^c \tilde{Y}_t^{c\prime} \right) - T \left(\tilde{\Sigma}_1^c + \tilde{\Sigma}_1^{c\prime} \right) \right] \left(T^{-1} \sum_{t=2}^{T} \tilde{Y}_{t-1}^c \tilde{Y}_{t-1}^{c\prime} \right)^{-1}$.
      iii. Define $MQ_c^c(q) = T\left[\tilde{v}_c^c(q) - 1\right]$.
   (b) For p fixed that does not depend on N and T:
      i. Estimate a VAR of order p in $\Delta \tilde{Y}_t^c$ to obtain $\tilde{\Pi}(L) = I_q - \tilde{\Pi}_1 L - \ldots - \tilde{\Pi}_p L^p$. Filter $\tilde{Y}_t^c$ by $\tilde{\Pi}(L)$ to get $\tilde{y}_t^c = \tilde{\Pi}(L) \tilde{Y}_t^c$.
      ii. Let $\tilde{v}_f^c(q)$ be the smallest eigenvalue of
      $$\Phi_f^c = \frac{1}{2} \left[\sum_{t=2}^{T} \left(\tilde{Y}_t^c \tilde{Y}_{t-1}^{c\prime} + \tilde{Y}_{t-1}^c \tilde{Y}_t^{c\prime} \right) \right] \left(T^{-1} \sum_{t=2}^{T} \tilde{Y}_{t-1}^c \tilde{Y}_{t-1}^{c\prime} \right)^{-1}.$$
      iii. Define the statistic $MQ_f^c(q) = T\left[\tilde{v}_f^c(q) - 1\right]$.

3. If $H_0 : r_1 = q$ is rejected, set $q = q - 1$ and return to the first step. Otherwise, $\tilde{r}_1 = q$ and stop.

The limiting distribution of these statistics are given in the following Theorem.

Theorem 1 *Let $\{Y_{i,t}\}$ be the stochastic process with DGP given by (1) to (5), with $\gamma_i = 0$ in (1). Under Assumptions A-G, the following results hold as $N, T \to \infty$. Let k be the order of autoregression in (15) chosen such that $k \to \infty$ and $k^3 / \min[N, T] \to 0$.*

(i) Under the null hypothesis that $\rho_i = 1$ in (5),

$$MSB_{\tilde{e}}(i) \Rightarrow \int_0^1 W_i(r)^2 \, dr,$$

where $W_i(r)$ denotes a standard Brownian motion.

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

$$MSB_{\tilde{F}} \Rightarrow \int_0^1 W_w(r)^2 \, dr,$$

where $W_w(r)$ denotes a standard Brownian motion.

(iii) When $r > 1$, let W_q^c be a vector of demeaned Brownian motions. Let $v_*^c(q)$ be the smallest eigenvalues of the statistic

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

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

It is interesting to note that the limiting distribution in part (i) does not depend on the stochastic regressors $X_{i,t}$, nor on the unobservable common stochastic trend F_t. This is a very useful property as it does not require many tables for critical values. Furthermore, since the limit is free from the common shocks, the individual test statistics can be pooled if $e_{i,t}$ are cross-sectionally uncorrelated. As is shown in the next section, the limiting distribution is different, however, when linear trends are entertained in the model.

To sum up, the statistics that have been proposed in this section can be constructed following these steps:

1. Take the first order difference for the dependent and the explanatory variables, and label them as y_i, which is $(T-1) \times 1$, and x_i, which is $(T-1) \times p$, for $i = 1, 2, ..., N$.

2. Construct the projection matrix M_i, and define $y_i^* = M_i y_i$ $i = 1, 2, ..., N$, and let $y^* = (y_1^*, y_2^*, ..., y_N^*)$.

3. Estimate f and Λ from the $(T-1) \times (T-1)$ matrix $y^* y^{*\prime}$ via singular value decomposition. Define

$$\tilde{z}_{i,t} = y_{i,t}^* - \tilde{f}_t' \tilde{\lambda}_i.$$

4. For each i, construct the cumulative sum $\tilde{e}_{i,t} = \sum_{s=1}^{t} \tilde{z}_{i,s}$, estimate the long-run variance $\tilde{\sigma}_i^2$ using (14) and (15), and construct the MSB test given in (13) based on $\tilde{e}_{i,t}$. Response surfaces to approximate finite sample p-values are provided in Bai and Carrion-i-Silvestre (2009).

5. If there is only one common factor ($r = 1$), construct the cummulative sum $\tilde{F}_t = \sum_{s=2}^{t} \tilde{f}_s$. Estimate the long-run variance $\tilde{\sigma}^2$ using (14) and (15), but with $\tilde{F}_t$ instead of $\tilde{e}_{i,t}$, and construct the MSB test given in (16) based on $\tilde{F}_t$. Response surfaces to approximate finite sample p-values are provided in Bai and Carrion-i-Silvestre (2009).

6. If there are more than one common factor ($r > 1$), define the cummulative sum $\tilde{F}_t = \sum_{s=2}^{t} \tilde{f}_s$, and compute the demeaned $\tilde{F}_t^c = \tilde{F}_t - \overline{\tilde{F}}$ series. Start with $q = r$ and proceed to test the number of stochastic trends following the three stages described earlier. This requires the computation of either the $MQ_c^c(q)$ or the $MQ_f^c(q)$ statistics. Asymptotic critical values are provided in Bai and Ng (2004), Table I.

2.1.2 Linear trend case

In the previous section we assume $\gamma_i = 0$ for all i. We now relax this assumption to allow heterogeneous linear trends as in (1)

$$Y_{i,t} = \mu_i + \gamma_i t + X_{i,t}'\beta_i + F_t'\lambda_i + e_{i,t}. \tag{17}$$

The estimation starts with model transformation that purges the deterministic component $\mu_i + \gamma_i t$. By doing so, the analysis also allows drifts in $X_{i,t}$ and in F_t. In fact, the analysis is invariant to drifts, as explained in details in the appendix. Purging the deterministic part requires differencing and then demeaning. Differencing (17) yields,

$$\Delta Y_{i,t} = \gamma_i + \Delta X_{i,t}'\beta_i + \Delta F_t'\lambda_i + \Delta e_{i,t}.$$

The first difference does not remove the deterministic elements as the trend becomes an intercept for the differenced data. This is a relevant feature, leading to a different limiting distribution of the MSB statistic. Further demeaning yields

$$\Delta Y_{i,t} - \overline{\Delta Y_i} = (\Delta X_{i,t} - \overline{\Delta X_i})\beta_i + (\Delta F_t - \overline{\Delta F})\lambda_i + \Delta e_{i,t} - \overline{\Delta e_i},$$

where $\overline{\Delta Y}_i = \frac{1}{T-1}\sum_{t=2}^{T} \Delta Y_{i,t}$ with $\overline{\Delta X}_i$ and $\overline{\Delta F}$ defined similarly. Rewrite the above as

$$y_i = x_i\beta_i + f\lambda_i + \Delta e_i - \iota\,\overline{\Delta e}_i, \tag{18}$$

where

$$y_i = \Delta Y_i - \iota\,\overline{\Delta Y}_i, \quad x_i = \Delta X_i - \iota\overline{\Delta X}_i, \quad f = \Delta F - \iota\,\overline{\Delta F},$$

these are, respectively, $(T-1)\times 1$, $(T-1)\times p$, and $(T-1)\times r$ matrices. Introduce the projection matrix,

$$M_i = I_{T-1} - x_i(x_i'x_i)^{-1}x_i',$$

which has the same form as in the previous section, but x_i is defined differently. Left multiply M_i on each side of (18), we have

$$\begin{aligned} M_i y_i &= M_i f\lambda_i + M_i(\Delta e_i - \iota\,\overline{\Delta e}_i) \\ &= f\lambda_i + \Delta e_i - \iota\,\overline{\Delta e}_i - P_i f\lambda_i - P_i(\Delta e_i - \iota\,\overline{\Delta e}_i), \end{aligned}$$

or

$$y_i^* = f\lambda_i + z_i, \tag{19}$$

where

$$y_i^* = M_i y_i, \quad z_i = \Delta e_i - \iota\,\overline{\Delta e_i} - P_i f \lambda_i - P_i \Delta e_i, \tag{20}$$

note $P_i\,\iota\,\overline{\Delta e_i} = 0$ as $P_i\,\iota = 0$.

To estimate f and $\Lambda = (\lambda_1, ..., \lambda_N)'$, we introduce,

$$y^* = (y_1^*, y_2^*, ..., y_N^*),$$

a $(T-1) \times N$ matrix. Let $\tilde{f}$ and $\tilde{\lambda}$ be computed the same way as in the previous subsection. Define

$$\tilde{z}_{i,t} = y_{i,t}^* - \tilde{f}_t' \tilde{\lambda}_i.$$

Finally,

$$\tilde{e}_{i,t} = \sum_{s=2}^{t} \tilde{z}_{i,s}$$

$$\tilde{F}_t = \sum_{s=2}^{t} \tilde{f}_s.$$

Let $MSB_{\tilde{e}}$ and $MSB_{\tilde{F}}$ be constructed exactly the same way as before. When $r > 1$ we can compute the MQ statistics defined in the previous subsection where now $\tilde{F}_t^c$ is replaced by $\tilde{F}_t^\tau$, $\tilde{F}_t^\tau$ being the residuals from a regression of $\tilde{F}_t$ on a constant and a time trend. Then, testing the number of common stochastic trends proceeds exactly in the same way using either the $MQ_c^\tau(q)$ or the $MQ_f^\tau(q)$ statistics, with $\tilde{v}_c^\tau(q)$ and $\tilde{v}_f^\tau(q)$ computed as $\tilde{v}_c^c(q)$ and $\tilde{v}_f^c(q)$ in the previous subsection, respectively, but using detrended common factors.

Theorem 2 *Let $\{Y_{i,t}\}$ be the stochastic process with DGP given by (1) to (5), with linear trends allowed in (1). Under Assumptions A-G, the following results hold as $N, T \to \infty$. Let k be the order of autoregression chosen such that $k \to \infty$ and $k^3/\min[N,T] \to 0$.*

(i) Under the null hypothesis that $\rho_i = 1$ in (5)

$$MSB_{\tilde{e}}(i) \Rightarrow \int_0^1 V_i(r)^2\,dr,$$

where $V_i(r) = W_i(r) - rW_i(1)$, $i = 1, \ldots, N$, denotes a standard Brownian bridge.

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

$$MSB_{\tilde{F}} \Rightarrow \int_0^1 V_w(r)^2\,dr,$$

where $V_w(r) = W_w(r) - rW_w(1)$ denotes a standard Brownian bridge.

(iii) When $r > 1$, let W_q^τ a vector of detrended Brownian motions. Let $v_*^\tau(q)$ be the smallest eigenvalues of the statistic

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

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

The proof is provided in the appendix. As expected, the limiting distribution of these statistics depend on the deterministic specification, but it does not depend on the stochastic regressors in the cointegrating relationship. This is quite convenient since it reduces the amount of tables needed to carry out the statistical inference.

To sum up, the statistics that have been proposed in this section for the linear trend case can be constructed as follows:

1. Differencing and demeaning both the dependent and the explanatory variables, and label them as y_i, which is $(T-1)\times 1$, and x_i, which is $(T-1)\times p$, for $i=1,2,...,N$.

2. Construct the projection matrix M_i, and define $y_i^* = M_i y_i$ $i=1,2,...,N$, and let $y^* = (y_1^*, y_2^*, ..., y_N^*)$.

3. The computation of the $MSB_{\tilde{e}}$ and $MSB_{\tilde{F}}$ statistics is identical to the previous section. Response surfaces to approximate finite sample p-values are provided in Bai and Carrion-i-Silvestre (2009).

4. If $r > 1$, define the cummulative sum $\tilde{F}_t = \sum_{s=2}^{t} \tilde{f}_s$, and compute the detrended $\tilde{F}_t^\tau$ factors, where $\tilde{F}_t^\tau$ denotes the residuals from a regression of $\tilde{F}_t$ on a constant and a linear time trend. Start with $q=r$ and proceed to test the number of stochastic trends following the three stages described earlier, computing the $MQ_c^\tau(q)$ or the $MQ_f^\tau(q)$ statistics. Asymptotic critical values are provided in Bai and Ng (2004), Table I.

2.2 Non-strictly exogenous regressors

In this section we allow $X_{i,t}$ to be correlated with the disturbances $e_{i,t}$ but maintain the assumption that $X_{i,t}$ and the factors F_t are independent. The case of dependence between $X_{i,t}$ and F_t is considered in the next section. Using idea from dynamic least squares method, by adding leads and lags of $\Delta X_{i,t}$ to control for endogeneity, we assume the model can be written as

$$Y_{i,t} = \mu_i + \gamma_i t + X_{i,t}'\beta_i + \Delta X_{i,t}' A_i(L) + F_t \lambda_i + \xi_{i,t}, \tag{21}$$

where $A_i(L)$ is a vector of polynomials of lead and lag operators with m_1 lags and m_2 leads. Let $m = m_1 + m_2$. For simplicity, we assume m_1 and m_2 are finite. The regressors $X_{i,t}$ and $\Delta X_{i,s}$ are strictly exogenous relative to $\xi_{i,t}$. In addition, the error term $\xi_{i,t}$ is I(0) when $e_{i,t}$ is I(0), and $\xi_{i,t}$ is I(1) when $e_{i,t}$ is I(1).

Equation (21) follows from the projection argument. If $e_{i,t}$ is I(0), we can directly project $e_{i,t}$ on leads and lags of $\Delta X_{i,t}$ such that $e_{i,t} = \Delta X_{i,t}' A_i(L) + \xi_{i,t}$ with $\xi_{i,t}$ being I(0), and (21) follows immediately. When $e_{i,t}$ is I(1), we can project $\Delta e_{i,t}$ onto $\Delta X_{i,t}$ such that $\Delta e_{i,t} = \Delta X_{i,t}' B_i(L) + \eta_{i,t}$. This implies that $e_{i,t} = X_{i,t}' B_i(L) + \xi_{i,t}$ with $\xi_{i,t} = \sum_{s=0}^{t} \eta_{i,s} \sim I(1)$. But by the Beveridge-Nelson decomposition, we can write $X_{i,t}'\beta_i + X_{i,t}' B_i(L)$ as $X_{i,t}'\tau_i + \Delta X_{i,t}' A_i(L)$ for some τ_i and $A_i(L)$. Then (21) follows upon renaming τ_i as β_i. The idea is that $\xi_{i,t}$ has the same order of integration as $e_{i,t}$.

The intercept only specification imposes $\gamma_i = 0$ in (21), while for the time trend specification $\gamma_i \neq 0$. Differencing (21) gives

$$\Delta Y_{i,t} = \gamma_i + \Delta X_{i,t}'\beta_i + \Delta^2 X_{i,t}' A_i(L) + \Delta F_t' \lambda_i + \Delta\xi_{i,t}. \tag{22}$$

As in section 2.1, introduce the following notation for the intercept only case. Let y_i be the $(T-m-1)\times 1$ vector consisting of $\Delta Y_{i,t}$ $(t = m_1+2, ...T-m_2)$, and let x_i be the $(T-m-1)\times(m+2)p$ matrix with each row of the form $(\Delta X_{i,t}', \Delta^2 X_{i,t-m_1}', ..., \Delta^2 X_{i,t+m_2}')$. Similarly, let f be $(T-m-1)\times r$ matrix with row elements $\Delta F_t'$ and let $\Delta\xi_i$ be $(T-m-1)\times 1$ vector with elements $\Delta\xi_{i,t}$ $(t = m_1+2, ..., T-m_2)$. We can rewrite (22) with $\gamma_i = 0$ as

$$y_i = x_i\delta_i + f\lambda_i + \Delta\xi_i, \tag{23}$$

where δ_i is a vector of parameters consisting of β_i and the coefficients in $A_i(L)$. Let us define the $(T-m-1)\times(T-m-1)$ projection matrix

$$M_i = I_{T-m-1} - x_i(x_i'x_i)^{-1}x_i' = I_{T-m-1} - P_i.$$

Left multiplying M_i each side of (23), we obtain (10) with $y_i^* = M_i y_i$ and $z_i = M_i\Delta\xi_i - P_i f\lambda_i$ as in (11). The whole analysis in Section 2.1.1 goes through. The requirement $\frac{1}{T}\sum_{t=1}^{T}\Delta X_{i,t}\Delta e_{i,t} = O_p(T^{-1/2})$ is now replaced by $\frac{1}{T}\sum_{t=m_1+2}^{T-m_2} x_{i,t}\Delta\xi_{i,t} = O_p(T^{-1/2})$, which holds since $\Delta\xi_{i,t}$ is uncorrelated with $x_{i,t}$.

In the presence of linear trends, we define y_i and x_i as the above but with their time series sample means (columnwise means) removed. Similarly, f and $\Delta\xi_i$ are defined with their sample means removed as well. The analysis is the same as that of section 2.1.2. We summarize the result in the following theorem.

Theorem 3 *Let $\{Y_{i,t}\}$ be the stochastic process with DGP given by (1) to (5). Suppose that Assumptions A-F hold. Let $MSB_{\tilde{e}}(i)$ and $MSB_{\tilde{F}}$ be the test statistics based on newly defined y_i and x_i, then Theorem 1 and Theorem 2 still hold.*

3 Regressors correlated with common factors

Previous derivations rely on the assumption that stochastic regressors are not correlated with the common factors. In this section, we relax this assumption by allowing correlations between $X_{i,t}$ and F_t. In fact, $X_{i,t}$ can be correlated with F_t, or with λ_i or both. The idea is that, similar to the left hand side variable $Y_{i,t}$, the regressors $X_{i,t}$ are likely to be impacted by the common shocks F_t. For example, $X_{i,t}$ may take on the form

$$X_{i,t} = A_t\lambda_i + B_i F_t + \sum_{k=1}^{r} C_{i,k}(F_{k,t}\lambda_{k,t}) + \Pi_i G_t + \eta_{i,t},$$

where A_t, B_i, $C_{i,k}$ are matrices or vectors, and G_t is vector of another common factors not influencing $Y_{i,t}$, and $\eta_{i,t}$ are iid, say. As a result, the following condition used earlier

$$\frac{1}{T}\sum_{t=1}^{T}\Delta X_{i,t}\Delta F_t = O_p(T^{-1/2}),$$

(for the intercept only case), or

$$\frac{1}{T}\sum_{t=1}^{T}(\Delta X_{i,t} - \overline{\Delta X}_i)(\Delta F_t - \overline{\Delta F}) = O_p(T^{-1/2}),$$

(for the linear trend case) may not hold. The above limit is nonzero in general when $X_{i,t}$ and F_t are correlated. To tackle the problem, we estimate β_i and F jointly. This will permit consistent estimation of both the regression parameters and factors, and thus the residuals.

We reproduce model (17) here

$$Y_{i,t} = \mu_i + \gamma_i t + X_{i,t}'\beta_i + F_t'\lambda_i + e_{i,t}. \tag{24}$$

In the context of stationary regressors and stationary disturbances, Bai (2009) considers the estimation of the above model, allowing for correlation between $X_{i,t}$ and F_t. Bai, Kao and Ng (2009) estimate the model with I(1) regressors and I(1) factors, taking cointegration as given. Our purpose here is to test for cointegration.

In the present setting, the null hypothesis implies $e_{i,t}$ to be I(1). We therefore need to difference the data to achieve stationarity. As in the previous sections, an added advantage of differencing is that the limit of the test statistic, $MSB_{\tilde{e}}(i)$, does not depend on $X_{i,t}$ and F_t. Without differencing, the resulting test statistic would have a limit involving residual Brownian motion, which is obtained as a projection residual by projecting the Brownian motion associated with $e_{i,t}$ onto those associated with $X_{i,t}$ and F_t. The resulting test statistics cannot be pooled due to cross correlations induced by the common trend F_t.

Differencing gives

$$\Delta Y_{i,t} = \gamma_i + \Delta X_{i,t}'\beta_i + \Delta F_t'\lambda_i + \Delta e_{i,t}.$$

In vector notation,

$$\Delta Y_i = \gamma_i \,\iota + \Delta X_i \,\beta_i + \Delta F \,\lambda_i + \Delta e_i$$

where ι is a vector of ones. The discussion in this section assumes $X_{i,t}$ is strictly exogenous with respect to the idiosyncratic errors, otherwise, we need to add leads and lags of $\Delta X_{i,t}$ in equation (24), as in Section 2.2.

If no linear trend is assumed ($\gamma_i = 0$ for all i), we define the projection matrix to be an identical matrix, i.e.,

$$M = I_{T-1}.$$

If linear trend is allowed, we define

$$M = I_{T-1} - T^{-1}\iota\,\iota',$$

(a demean operator). Multiply M on each side of the model equation we have

$$M\Delta Y_i = M\Delta X_i\,\beta_i + M\Delta F\,\lambda_i + M\Delta e_i,$$

or

$$y_i = x_i\beta_i + f\lambda_i + z_i, \tag{25}$$

where

$$y_i = M\Delta Y_i, \quad x_i = M\Delta X_i, \quad f = M\Delta F, \quad z_i = M\Delta e_i.$$

Note that M does not depend on i.

We use the least squares method to estimate (β_i, f, Λ). They are estimated jointly. The least squares objective function is defined as:

$$SSR\,(\beta_i, f, \Lambda) = \sum_{i=1}^{N} (y_i - x_i\beta_i - f\lambda_i)'\,(y_i - x_i\beta_i - f\lambda_i)\,, \tag{26}$$

subject to the constraint $f'f/\,(T-1) = I_r$ and $\Lambda'\Lambda$ being diagonal. Concentrating out Λ, the least squares estimator $(\tilde{\beta}_1, ...\tilde{\beta}_N, \tilde{f})$ must satisfy, see Bai (2009), the following system of nonlinear equations:[4]

$$\tilde{\beta}_i = \left(x_i'x_i\right)^{-1} x_i'\left(y_i - \tilde{f}\tilde{\lambda}_i\right), \quad (i = 1, 2, ..., N) \tag{27}$$

$$\left[\frac{1}{NT}\sum_{i=1}^{N}\left(y_i - x_i\tilde{\beta}_i\right)'\left(y_i - x_i\tilde{\beta}_i\right)\right]\tilde{f} = \tilde{f}V_{NT}, \tag{28}$$

where V_{NT} is the diagonal matrix containing the r largest eigenvalues of the matrix in the squared brackets. Note that $\tilde{\beta}_i$ and $\tilde{f}$ can be obtained iteratively. Given β_i, we can estimate

[4]If common slope coefficient $\beta_i = \beta$ is assumed, equation (27) becomes

$$\tilde{\beta} = \left(\sum_{i=1}^{N} x_i'x_i\right)^{-1}\sum_{i=1}^{N} x_i'\left(y_i - \tilde{f}\tilde{\lambda}_i\right)$$

and equation (28) remains the same with $\tilde{\beta}_i$ replaced by $\tilde{\beta}$.

f, and given f we can estimate β_i. This process is iterated until convergence. Once $\left(\tilde{\beta}_i, \tilde{f}\right)$ is available we can obtain the loading matrix as $\tilde{\lambda}_i = (T-1)^{-1}\tilde{f}'\left(y_i - x_i\tilde{\beta}\right)$. Finally, define

$$\tilde{z}_i = y_i - x_i\tilde{\beta}_i - \tilde{f}\tilde{\lambda}_i.$$

Bai (2009) shows that this iterated approach gives consistent estimation of β_i, f and λ_i (for each i). Because the differenced data are I(0), the rate of convergence for β_i is $\sqrt{T}$. But this rate is sufficient for our purpose. In addition, the estimated $\tilde{f}$ and $\tilde{\Lambda}$ possess properties similar to a pure factor model, despite correlations between $\Delta X_{i,t}$ and ΔF_t. In particular, we have

$$T^{-1/2}\sum_{s=2}^{t} v_s = T^{-1/2}\sum_{s=2}^{t}(\tilde{f}_s - Hf_s) = O_p(C_{NT}^{-1}),$$

and

$$d_i = \tilde{\lambda}_i - H^{-1\prime}\lambda_i = O_p(C_{NT}^{-1}).$$

Exactly as before, estimate $e_{i,t}$ again by

$$\tilde{e}_{i,t} = \sum_{s=2}^{t} \tilde{z}_{i,s},$$

and estimate F_t by

$$\tilde{F}_t = \sum_{s=2}^{t} \tilde{f}_s.$$

Let $MSB_{\tilde{e}}$, $MSB_{\tilde{F}}$ and MQ test statistics be defined as in Section 2. The limiting distributions of these statistics are given in the following Theorem.

Theorem 4 *Let the DGP for the stochastic process $\{Y_{i,t}\}$ be given by (24) together with (2) to (5). Suppose that Assumptions A-F hold and the slope coefficients and the factors are estimated jointly. Then the limiting distributions in Theorem 1 and Theorem 2 still hold.*

In summary, in spite of correlations between $X_{i,t}$ with F_t or with λ_i, the results in previous sections continue to hold. Simulations show this approach indeed works quite well in terms of size and power properties.

4 Pooled test statistics

Using results of previous sections, we can define panel cointegration statistics that combine individual statistics for each cross-section. Pooling individual statistics can yield more powerful tests. We consider several approaches to combining. Each of those approaches assumes asymptotic independence of individual statistics. Assuming idiosyncratic errors e_{it} are cross-sectionally independent, then all cross-section correlations are captured by the common factors F_t. In view that the individual test statistics $MSB_{\tilde{e}}(i)$ do not depend on the common factors in the limit, they are asymptotically independent. Thus pooling is permitted.

The first approach of combining standardizes the sample average of individual statistics so that

$$MSB_{\tilde{e}} = \sqrt{N}\frac{\overline{MSB_{\tilde{e}}\left(i\right)} - \bar{\xi}}{\bar{\varsigma}} \to N\left(0,1\right),$$

where $\overline{MSB_{\tilde{e}}\left(i\right)} = N^{-1}\sum_{i=1}^{N} MSB_{\tilde{e}}\left(i\right)$, $\bar{\xi} = N^{-1}\sum_{i=1}^{N}\xi_i$ and $\bar{\varsigma}^2 = N^{-1}\sum_{i=1}^{N}\varsigma_i^2$, where ξ_i and ς_i^2 denotes the mean and variance of $MSB_{\tilde{e}}\left(i\right)$ respectively. The following Lemma provides these moments.

Lemma 1 *Let $MSB_{\tilde{e}}(i) = \tilde{\sigma}_i^{-2} T^{-2} \sum_{t=1}^{T} \tilde{e}_{i,t-1}^2$ be the test statistic with limit distribution given in Theorems 1 to 4. Let ξ_i and ς_i^2 denote the mean and variance, respectively, of the limiting random variable of $MSB_{\tilde{e}}(i)$, then*

(1) The only constant case: $\xi_i = \frac{1}{2}$ *and* $\varsigma_i^2 = \frac{1}{3}$

(2) The time trend case: $\xi_i = \frac{1}{6}$ *and* $\varsigma_i^2 = \frac{1}{45}$.

It is possible to define panel statistics through the combination of individual p-values. Thus, under the assumption of cross-section independence of $e_{i,t}$, $-2\ln p_i \sim \chi_2^2$, a result that was used in Maddala and Wu (1999) to define the Fisher-type test statistic:

$$P = -2\sum_{i=1}^{N} \ln p_i \sim \chi_{2N}^2,$$

where p_i denotes the p-value of the $MSB_{\tilde{e}}(i)$ statistic for the i-th unit. Choi (2001) proposes the following test when $N \to \infty$:

$$P_m = \frac{-2\sum_{i=1}^{N} \ln p_i - 2N}{\sqrt{4N}} \to N(0,1),$$

as $N \to \infty$.

The computation of these statistics requires the corresponding p-values. Bai and Carrion-i-Silvestre (2009) provide response surfaces that can be used to approximate these p-values for the MSB statistic. In summary, we have three different ways to combine the individual statistics. Monte Carlo simulations are conducted in the next section to evaluate the performance of those aggregated statistics.

5 Monte Carlo simulation

5.1 Regressors independent of the common factors

Finite sample properties of our procedure are investigated through the specification of the following bivariate DGP:

$$\begin{aligned} Y_{i,t} &= \mu_i + \gamma_i t + X_{i,t}\beta_i + u_{i,t} \\ u_{i,t} &= F_t'\lambda_i + e_{i,t} \\ F_t &= \alpha F_{t-1} + \sigma_F w_t \\ e_{i,t} &= \rho_i e_{i,t-1} + \varepsilon_{i,t} \\ \Delta X_{i,t} &= v_{i,t}, \end{aligned}$$

where $(w_t, \varepsilon_{i,t}, v_{i,t})'$ consists of iid standard normal random variables for all i and t. We consider various combinations for the number of factors r and the value of AR parameters (α, ρ_i). More specifically, $r = \{1, 3\}$, $\alpha = \{0.9, 0.95, 1\}$ and $\rho_i = \{0.95, 0.99, 1\}$ for all i. These values allow analyzing both the empirical size and power of the statistics. The relative importance of the common factors is controlled through the value of $\sigma_F^2 = \{0.5, 1, 10\}$. Note that the test statistics are invariant to the values of μ_i and γ_i, therefore they are set to zero. The test statistics only depend on whether trends are allowed or not in the estimation procedure. In addition, we set $\beta_i = 1$ for all i. The heterogenous slope coefficients will be considered later. Throughout the simulation experiments the number of common factors is estimated using the panel BIC information criterion in Bai and Ng (2002) with $r_{\max} = 6$ as the maximum number of factors. We consider $N = 40$ individuals and $T = \{50, 100, 250\}$ time observations. The number of replications in all cases is set at 5,000 and the nominal size is set at the 5% level. In order to save space, we only report the results for the time trend deterministic specification – the results for the intercept-only case are similar in all cases.

Table 1 reports the empirical size and power for the time trend case. As can be seen, the $MSB_{\tilde{e}}$ statistic is undersized since the empirical size is mildly below the nominal size. The panel statistics that are based on the combination of individual p-values, P and P_m statistics, show good size. All three panel data statistics present high power, even for $\rho_i = 0.99$. In most cases the empirical power is almost one for $\rho_i = 0.95$. The $MSB_{\tilde{F}}$ statistic has the correct size and, as expected, the power increases as the autoregressive coefficient moves away from unity. It is worth noticing that these features are also found for the constant deterministic specification.

Similar conclusions are obtained for the case of three common factors. As above, Table 2 suggests that statistics using individual p-values have better empirical size and power. One reason for the mild oversize shown by $MSB_{\tilde{e}}$ could be the fact that the limiting distribution of this statistic is not symmetric. Regarding the MQ tests, we observe that when $\alpha = 1$ and large T the parametric MQ statistic has the correct empirical size, while the non-parametric one shows some size distortion. Note that, in Table 2, MQ(3) denotes the frequency that the MQ statistics have detected three common stochastic trends, MQ(2) the frequency of two stochastic trends, MQ(1) the frequency of one stochastic trend, and finally, MQ(0) denotes the frequency that the statistics detect no stochastic trend. Regarding the empirical power, we see that the MQ tests do not show high power unless T is large and α moves away from unity, which is expected even if F_t is observable, and is due to the non-panel nature of F_t. This is in contrast with evidence for the panel statistics, which show good power.

5.2 Regressors correlated with common factors

The DGP that is used to assess the performance of the statistics when stochastic regressors are correlated with either the common factors or the loadings is given by

$$Y_{i,t} = X_{1i,t}\beta_{i1} + X_{2i,t}\beta_{i2} + F_t'\lambda_i + e_{i,t}, \tag{29}$$

$i = 1, \ldots, N$, $t = 1, \ldots, T$, where the stochastic regressors are generated according to

$$\begin{gathered} X_{1i,t} = \mu_1 + c_1 F_t'\lambda_i + \iota'\lambda_i + \iota' F_t + \eta_{1i,t} \\ X_{2i,t} = \mu_2 + c_2 F_t'\lambda_i + \iota'\lambda_i + \iota' F_t + \eta_{2i,t} \\ \eta_{1i,t} = \eta_{1i,t-1} + v_{1i,t}; \quad \eta_{2i,t} = \eta_{2i,t-1} + v_{2i,t} \\ (v_{1i,t}, v_{2i,t}) \sim iid\ N(0, I_2). \end{gathered}$$

Common factors and idiosyncratic disturbance terms are given by

$$\begin{aligned} e_{i,t} &= \rho_i e_{i,t-1} + \varepsilon_{i,t} \\ F_t &= \alpha F_{t-1} + \sigma_F w_t, \end{aligned} \tag{30}$$

with $(w_t, \varepsilon_{i,t})'$ consists of iid standard normal random variables for all i and t. We set $\mu_1 = \mu_2 = c_1 = c_2 = 1$. Empirical size and power are investigated for all possible pairs of $\rho_i = \{0.95, 0.99, 1\}$ and $\alpha = \{0.9, 0.95, 1\}$. As above, $r = \{1, 3\}$, and the importance of the common factors is controlled through the value of $\sigma_F^2 = \{0.5, 1, 10\}$. Simulations are performed for $T = \{50, 100, 250\}$ observations and $N = 40$ individuals. Computational cost due to the iterative estimation procedure has led us to base the results on 1,000 replications. For the slope parameters, we consider two cases. The first case is for common slope parameters, so that β_{i1} and β_{i2} do not depend on i. The second case considers heterogenous slope parameters.

Common slope parameters. The true parameter values are $(\beta_1, \beta_2) = (1, 3)$. Table 3 offers results when $r = 1$, which shows similar conclusions as the case in Section 5.1. Regarding the statistics for the idiosyncratic disturbance terms, we see that the statistics based on pooling the p-values show better performance in terms of the empirical size. In all cases the tests present non-trivial power, even for $\rho_i = 0.99$. The MSB statistic computed for the common factor shows good empirical size and power as well. Note that these results are obtained irrespective of the deterministic specifications.

Results under three common factors are reported in Tables 4 to 6. As before, the panel data statistics using the p-values have empirical size close to the nominal one. Their empirical power is quite good even for large autoregressive coefficient. Note that these results are obtained regardless of the deterministic specification. Regarding the MQ statistics, we are able to detect the existence of three common factors. The MQ tests show the correct empirical size. However, we require large T and large σ_F^2 for the statistics to have good empirical power. As mentioned above, this feature is due to the fixed dimension of F_t.

Heterogeneous parameters. The set-up of the simulation experiment in this case is the same as for the homogeneous case, except that the slope parameters β_1 and β_2 in (29) are randomly distributed as $\beta_1 \sim N(1, 1)$ and $\beta_2 \sim N(3, 1)$. The results in Tables 7 to 9 are similar to those in the previous analysis. In general, panel data unit root tests based on p-value combination have an empirical size that is closer to the nominal one, while the panel test that combines the statistics is mildly under-sized. As for the empirical power, the statistics show higher power for the constant only case than for the time trend case. This is in accordance with the findings in Moon, Perron and Phillips (2004) and in our another paper, where it was shown that the more complicated the deterministic component the lower power of the statistics around the null hypothesis. Finally, the performance of the MQ tests is not altered when considering the heterogeneous parameters case; the non-parametric version of the MQ test IS more powerful than the parametric one.

6 Conclusions

This paper contributes to the literature on panel data cointegration analysis by considering cross-section dependence. The framework used is the approximate factor models. We distinguish two important aspects of the model. First, stochastic regressors are assumed to be independent of the unobservable common factors and factor loadings. Second and more important is the allowance of correlation amongst regressors and common factors and factor loadings. In both cases, the paper proposes statistics to test the presence of cointegration, whether or not the stochastic regressors are strictly or non-strictly exogenous. It is shown that the limiting distribution of these statistics depend on the deterministic specification but not on the number of stochastic regressors.

References

[1] Bai, J., 2009. Panel data models with interactive fixed effects. Econometrica, forthcoming.

[2] Bai, J., Carrion-i-Silvestre, J. Ll., 2009. Structural changes, common stochastic trends, and unit roots in panel data. Review of Economic Studies, 76, 2, 471-501.

[3] Bai, J., Kao, C., 2006. On the estimation and inference of a panel cointegration model with cross-sectional dependence. In Baltagi, Badi (Ed.), Contributions to economic analysis. Elsevier, 3-30.

[4] Bai, J., Kao, C., Ng S., 2009. Panel cointegration with global stochastic trends. Journal of Econometrics, 149, 1, 82-99.

[5] Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica, 70, 191-221.

[6] Bai, J., Ng, S., 2004. A PANIC attack on unit roots and cointegration. Econometrica, 72, 4, 1127-1177.

[7] Baltagi, B. H., 2008. Econometric analysis of panel data. Fourth Edition. John Wiley & Sons. New York.

[8] Banerjee, A., Carrion-i-Silvestre, J.Ll., 2006. Cointegration in panel data with breaks and cross-Section dependence. Working paper 591, European Central Bank.

[9] Banerjee, A., Wagner, M., 2009. Testing economic hypotheses using macro-panels of data. Forthcoming in The Palgrave Handbook of Econometrics, Volume II: Applied Econometrics, edited by T. C. Mills and K. Patterson, Palgrave Macmillan, Basingstoke.

[10] Breitung, J., 2005. A parametric approach to the estimation of cointegration vectors in panel data. Econometric Reviews 24, 151-173.

[11] Breitung, J., Pesaran, M. H., 2008. Unit roots and cointegration in panels. In Matyas, L. and Sevestre, P., The econometrics of panel data (Third Edition), Ch 9, pp. 279-322, Kluwer Academic Publishers.

[12] Carrion-i-Silvestre, J.Ll., Surdeanu, L., 2009. Panel cointegration rank testing with cross-section dependence. Mimeo, Department of Econometrics, Statistics and Spanish Economy, University of Barcelona.

[13] Chang, Y., 2002. Nonlinear IV unit root tests in panels with cross-sectional dependency. Journal of Econometrics, 110, 261-292.

[14] Choi, I., 2001. Unit root tests for panel data. Journal of Econometrics, 20, 249-272.

[15] Costantini, M., Destefanis, S., 2009. Cointegration analysis for cross-sectionally dependent panels: The case of regional production functions. Economic Modelling, 26, 320-327.

[16] Costantini, M., Lupi, C., 2006. Divergence and long-run equilibria in Italian regional unemployment. Applied Economic Letters, 13, 899-904.

[17] Gengenbach, C., Palm, F.C., Urbain, J.P., 2005. Cointegration testing in panels with common factors, METEOR Research Memorandum, Universiteit Maastricht.

[18] Gengenbach, C., Palm, F.C., Urbain, J.P., 2006. Cointegration testing in panels with common factors, Oxford Bulletin of Economics and Statistics 68 (Supplement), 683-719.

[19] Gengenbach, C., Urbain, J.P., Westerlund, J., 2009. Panel error cointegration testing with global stochastic trends. Working paper Meteor RM/08/051, Maastricht University.

[20] Groen, J.J.J., Kleibergen, F., 2003. Likelihood-based cointegration analysis in panels of vector error-correction models. Journal of Business & Economic Statistics, 21, 2, 295-318.

[21] Holly, S., Pesaran, M. H., Yamagata, T., 2009. A spatio-temporal model of house prices in the US. Journal of Econometrics, forthcoming.

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

[23] Kapetanios, G., Pesaran, M. H., Yamagata, T., 2006. Panels with nonstationary multifactor error structures. Working paper, Cambridge University.

[24] Larsson, R., Lyhagen, J., Löthgren, M., 2001. Likelihood-based cointegration tests in heterogeneous panels. Econometrics Journal, 4, 109-142.

[25] Maddala, G. S., 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, 631-652.

[26] McCoskey, S., Kao, C., 1998. A residual based test of the null of cointegration in panel data. Econometric Reviews, 17, 57-84.

[27] Moon, R. H., Perron, B., 2004. Testing for unit root in panels with dynamic factors. Journal of Econometrics, 122, 81–126.

[28] Moon, R. H., Perron, B., Phillips, P. C. B., 2007. Incidental trends and the power of panel unit root tests. Journal of Econometrics, 141, 416-459.

[29] Ng, S., Perron, P., 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, 1519-1554.

[30] Pedroni, P., 2000. Fully modified OLS for heterogeneous cointegrated panels. Advances in Econometrics, 15, 93-130.

[31] Pedroni, P., 2004. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory, 3, 579-625.

[32] Pedroni, P., 2007. Social capital, barriers to production and capital shares: Implications for the importance of parameter heterogeneity from a nonstationary panel approach. Journal of Applied Econometrics, 22, 429-451.

[33] Pesaran, M. H., 2006. Estimation and inference in large heterogeneous panels with multifactor error structure. Econometrica 74, 4, 967-1012.

[34] Ploberger, W., Phillips, P. C. B., 2004). Optimal testing for unit roots in panel data. Mimeo. University of Rochester.

[35] Stock, J. H., 1999. A class of tests for integration and cointegration. In Engle, R. F. and H. White (ed.), Cointegration, causality and forecasting. A Festchrift in Honour of Clive W. F. Granger. Oxford University Press.

[36] Tosetti, E., Moscone, F., 2007. Health Expenditure and Income in the United States. Working paper 07/14, Deparment of Economics. University of Leicester.

[37] Westerlund, J., 2007. Estimating cointegrated panels with common factors and the forward rate unbiasedness hypothesis. Journal of Financial Econometrics, 5, 491-522.

[38] Westerlund, J., 2008. Panel cointegration tests of the Fisher effect. Journal of Applied Econometrics, 23, 193-233.

[39] Westerlund, J., Edgerton, D., 2008. A simple test for cointegration in dependent panels with structural breaks. Oxford Bulletin of Economics and Statistics, 70, 5, 665-704.

A Mathematical Appendix

A.1 Common factors with strictly exogenous regressors

Proof of Theorem 1. From $y_i^* = f\lambda_i + z_i$ and $y_i^* = \tilde{f}\tilde{\lambda}_i + \tilde{z}_i$, we have

$$\tilde{z}_i = z_i + f\lambda_i - \tilde{f}\tilde{\lambda}_i.$$

That is,

$$\begin{aligned} \tilde{z}_{i,t} &= z_{i,t} + f_t'\lambda_i - \tilde{f}_t'\tilde{\lambda}_i \\ &= z_{i,t} - v_t H^{-1\prime}\lambda_i - \tilde{f}_t' d_i, \end{aligned} \tag{31}$$

where $v_t = \tilde{f}_t - Hf_t$ and $d_i = \tilde{\lambda}_i - H^{-1\prime}\lambda_i$. The computation of the partial sum processes of (31) gives:

$$T^{-1/2}\sum_{s=2}^{t}\tilde{z}_{i,s} = T^{-1/2}\sum_{s=2}^{t} z_{i,s} - T^{-1/2}\sum_{s=2}^{t} v_s H^{-1\prime}\lambda_i - T^{-1/2}\sum_{s=2}^{t}\tilde{f}_s' d_i. \tag{32}$$

We next analyze each term on the right hand side of (32). For the first term, recall

$$z_i = \Delta e_i - P_i[\Delta e_i + f\lambda_i],$$

or

$$\begin{aligned} z_{i,t} &= \Delta e_{i,t} - \Delta X_{i,t}'(x_i'x_i)^{-1}\Big[x_i'\Delta e_i + x_i' f\lambda_i\Big] \\ &= \Delta e_{i,t} - \Delta X_{i,t}'\Big(T^{-1}x_i'x_i\Big)^{-1}\Big[T^{-1}x_i'\Delta e_i + T^{-1}x_i' f\lambda_i\Big]. \end{aligned}$$

Note that

$$T^{-1}x_i'\Delta e_i = T^{-1}\sum_{t=2}^{T}\Delta X_{i,t}'\Delta e_{i,t} = O_p(T^{-1/2}), \quad T^{-1}x_i' f = T^{-1}\sum_{t=2}^{T}\Delta X_{i,t}'\Delta F_t = O_p(T^{-1/2}).$$

Thus

$$z_{i,t} = \Delta e_{i,t} + \Delta X_{i,t}' O_p(T^{-1/2}).$$

The cumulative sum of $z_{i,t}$ after dividing by $\sqrt{T}$ is,

$$\begin{aligned} T^{-1/2}\sum_{s=1}^{t} z_{i,s} &= T^{-1/2}e_{i,t} - (T^{-1/2}X_{i,t}')O_p(T^{-1/2}) \\ &= T^{-1/2}e_{i,t} + O_p(T^{-1/2}), \end{aligned}$$

we have assumed $e_{i,1} = 0$ and $X_{i,1} = 0$ for notational simplicity, without loss of generality.

Regarding the term involving $\{v_t\}$ we see from Eq. (A.3) in Bai and Ng (2004) that

$$T^{-1/2}\sum_{s=2}^{t} v_s = O_p\left(C_{NT}^{-1}\right),$$

where $C_{NT} = \min\left[\sqrt{N}, \sqrt{T}\right]$. Moreover and as shown in Bai and Ng (2004), the term $d_i = O_p\left(C_{NT}^{-1}\right)$ and $T^{-1/2}\sum_{s=2}^{t}\tilde{f}_s = O_p(1)$, so that

$$T^{-1/2}\sum_{s=2}^{t}\tilde{z}_{i,s} = T^{-1/2}\sum_{s=2}^{t} z_{i,s} + O_p\left(C_{NT}^{-1}\right) = T^{-1/2}e_{i,t} + O_p\left(C_{NT}^{-1}\right).$$

Since $T^{-1/2}e_{i,t} \Rightarrow \sigma_i W_i(r)$, it follows that

$$MSB_{\tilde{e}}(i) \Rightarrow \int_0^1 W_i(r)^2 \, dr,$$

that is, the limiting distribution is the same as derived in Stock (1999) for the constant case. This implies that the presence of stochastic regressors does not affect the limiting distribution of the statistic.

Next consider unit root test for F_t, the case of $r = 1$. From

$$T^{-1/2}\sum_{s=2}^{t} v_s = T^{-1/2}\sum_{s=2}^{t}(\tilde{f}_s - Hf_s) = O_p(C_{NT}^{-1}),$$

from $f_s = \Delta F_s$ and the definition of $\tilde{F}_t$, we have

$$T^{-1/2}[\tilde{F}_t - H(F_t - F_1)] = O_p(C_{NT}^{-1}),$$

and from $T^{-1/2}F_t \Rightarrow \sigma_F W_w(r)$, where $W_w(r)$ is standard Brownian motion, we have

$$T^{-1/2}H^{-1}\tilde{F}_t \Rightarrow \sigma_F W_w(r),$$

note that H^{-1} is scalar for $r = 1$. The MSB test is scale invariant, the scalar H^{-1} is cancelled out from the numerator and the denominator. This implies that $MSB_{\tilde{F}} \xrightarrow{d} \int_0^1 W_w(r)^2 dr$. The case of $r > 1$ is similar to Bai and Ng (2004), and is thus omitted.

To prove Theorem 2, we need the following lemma. Recall in the linear trend case, $x_{i,t} = \Delta X_{i,t} - \overline{\Delta X}_i$ and $f_t = \Delta F_t - \overline{\Delta F}$.

Lemma 2 *For drifted or driftless $X_{i,t}$ and F_t, we have*

(i) $T^{-1/2}\sum_{s=2}^{t} x_{i,s} = O_p(1)$

(ii) $T^{-1}x_i' f = O_p(T^{-1/2})$

(iii) $T^{-1}x_i'\Delta e_i = O_p(T^{-1/2})$

(iv) Let $z_{i,t} = \Delta e_{i,t} - \overline{\Delta e}_i - x_{i,t}'(x_i'x_i)^{-1}\Big[x_i' f\lambda_i + x_i'(\Delta e_i - \iota\,\overline{\Delta e}_i)\Big]$. *Then* $T^{-1/2}\sum_{s=2}^{t} z_{i,s} = T^{-1/2}[e_{i,t} - (\frac{t}{T})e_{i,T}] + o_p(1) \Rightarrow \sigma_i[W_i(r) - rW_i(1)] = \sigma_i V_i(r)$

Proof of (i). Note $x_{i,t} = \Delta X_{i,t} - \overline{\Delta X}_i$ is invariant to drift of $X_{i,t}$ (i.e., does not depend on the drift of $X_{i,t}$, if any). Without loss of generality, one may assume $X_{i,t}$ has no drift. In addition, $x_{i,t} = \Delta X_{i,t} - \frac{t}{T}(X_{i,T} - X_{i,1})$. Thus

$$T^{-1/2}\sum_{s=2}^{t} x_{i,s} = T^{-1/2}\Big[X_{i,t} - \frac{t}{T}X_{i,T}] - T^{-1/2}X_{i,1}(1 - \frac{t}{T}) = O_p(1).$$

Proof of (ii). We have

$$T^{-1}x_i' f = T^{-1}\sum_{t=2}^{T}(\Delta X_{i,t} - \overline{\Delta X_i})'(\Delta F_t - \overline{\Delta F}),$$

but $(\Delta X_{i,t} - \overline{\Delta X_i})$ and $(\Delta F_t - \overline{\Delta F})$ are invariant to drift, so without loss of generality, we can assume they are driftless. Then the above is $O_p(T^{-1/2})$ due to the independence of $X_{i,t}$ and F_t.

Proof of (iii). Same as (ii).

Proof of (iv). Combining (i),(ii), and (iii), we have

$$T^{-1/2}\sum_{s=2}^{t} x_{i,s}'\Big[(T^{-1}x_i'x_i)^{-1}[T^{-1}x_i'f + T^{-1}x_i'(\Delta e_i - \iota\,\overline{\Delta e_i})\Big] = O_p(T^{-1/2}),$$

it follows that

$$T^{-1/2}\sum_{s=2}^{t} z_{i,s} = T^{-1/2}\sum_{s=2}^{t}(\Delta e_{i,t} - \overline{\Delta e_i}) + O_p(T^{-1/2})$$
$$= T^{-1/2}\Big(e_{i,t} - \frac{t}{T}e_{i,T}\Big) + O_p(T^{-1/2}) \Rightarrow \sigma_i V_i(r).$$

Q.E.D.

Proof of Theorem 2. As in the proof of Theorem 1,

$$\tilde{z}_i = z_i + f\lambda_i - \tilde{f}\tilde{\lambda}_i$$

$$\tilde{z}_{i,t} = z_{i,t} - v_t H^{-1\prime}\lambda_i - \tilde{f}_t' d_i,$$

where $v_t = \tilde{f}_t - Hf_t$ and $d_i = \tilde{\lambda}_i - H^{-1\prime}\lambda_i$. Again, as before, cumulative sum leads to

$$T^{-1/2}\sum_{s=2}^{t}\tilde{z}_{i,s} = T^{-1/2}\sum_{s=2}^{t} z_{i,s} - T^{-1/2}\sum_{s=2}^{t} v_s H^{-1\prime}\lambda_i - T^{-1/2}\sum_{s=2}^{t}\tilde{f}_s' d_i$$

where, from (20),

$$z_{i,t} = \Delta e_{i,t} - \overline{\Delta e_i} - x_{i,t}'(x_i'x_i)^{-1}\Big[x_i'f\lambda_i + x_i'\Delta e_i - \iota\overline{\Delta e_i})\Big].$$

By Lemma 2(iv),

$$T^{-1/2}\sum_{s=2}^{t} z_{i,s} \Rightarrow \sigma_i V_i(r).$$

From Bai and Ng (2004),

$$T^{-1/2}\sum_{s=2}^{t} v_s = O_p(C_{NT}^{-1}) = o_p(1),$$

and

$$d_i = \tilde{\lambda}_i - H^{-1\prime}\lambda_i = O_p(C_{NT}^{-1}),$$

we have

$$T^{-1/2}\sum_{s=2}^{t}\tilde{z}_{i,s} = T^{-1/2}\sum_{s=2}^{t} z_{i,s} + o_p(1) \Rightarrow \sigma_i V_i(r).$$

It follows that

$$MSB_{\tilde{e}}\,(i) \Rightarrow \int_0^1 V_i(r)^2 dr.$$

Consider testing the stationarity of F_t with $r = 1$. From Bai and Ng (2004),

$$T^{-1/2}\sum_{s=2}^{t}(\tilde{f}_s - Hf_s) = O_p(C_{NT}^{-1})$$

where $f_t = \Delta F_t - \overline{\Delta F}$ with $\overline{\Delta F} = \frac{1}{T-1}(F_T - F_1)$. Cumulative sum of the true f_t

$$T^{-1/2}\sum_{s=2}^{t} f_s = T^{-1/2}\Big(F_t - F_1 - \frac{t-1}{T-1}(F_T - F_1)\Big) \Rightarrow \sigma_w V_w(r),$$

where $V_w(r)$ is a Brownian bridge. Next,

$$T^{-1/2}\tilde{F}_t = T^{-1/2}\sum_{s=2}^{t}\tilde{f}_s = H\ T^{-1/2}\left(F_t - F_1 - \frac{t-1}{T-1}(F_T - F_1)\right) + O_p\left(C_{NT}^{-1}\right).$$

It follows that

$$T^{-1/2}H^{-1}\tilde{F}_t \Rightarrow \sigma_w V_w(r).$$

By the definition of MSB test,

$$MSB_{\tilde{f}} \xrightarrow{d} \int_0^1 V_w(r)^2 dr.$$

The proof of $r > 1$ is the same as in Bai and Ng (2004), thus omitted.

A.2 Stochastic regressors correlated with common factors

Proof of Theorem 4. From $y_i = x_i\beta_i + f\lambda_i + z_i$ and $y_i = x_i\hat{\beta}_i + \hat{f}\hat{\lambda}_i + \hat{z}_i$, we have

$$\begin{aligned}\hat{z}_i &= z_i - x_i(\hat{\beta}_i - \beta_i) + f\lambda_i - \tilde{f}\tilde{\lambda}_i \\ &= z_i - x_i(\tilde{\beta}_i - \beta_i) - (\hat{f} - fH)H^{-1\prime}\lambda_i - \tilde{f}(\hat{\lambda}_i - H^{-1\prime}\lambda_i),\end{aligned}$$

or

$$\tilde{z}_{i,t} = z_{i,t} - x_{i,t}'(\tilde{\beta}_i - \beta_i) - v_t H^{-1\prime}\lambda_i - \tilde{f}_t d_i,$$

where $v_t = \tilde{f}_t - Hf_t$ and $d_i = \tilde{\lambda}_i - H^{-1\prime}\lambda_i$. Thus,

$$\frac{1}{\sqrt{T}}\sum_{s=2}^{t}\tilde{z}_{i,s} = \frac{1}{\sqrt{T}}\sum_{s=2}^{t}z_{i,s} - (\frac{1}{\sqrt{T}}\sum_{s=2}^{t}x_{i,s}')(\tilde{\beta}_i - \beta_i) - (\frac{1}{\sqrt{T}}\sum_{s=2}^{t}v_s)H^{-1\prime}\lambda_i - (\frac{1}{\sqrt{T}}\sum_{s=2}^{t}\tilde{f}_s)d_i. \quad (33)$$

The remaining proof focuses on the linear trend model, as the intercept only model is simpler. In this case, $x_i = M\Delta X_i = \Delta X_i - \iota\,\overline{\Delta X}_i$ and $f = M\Delta F = \Delta F - \iota\,\overline{\Delta F}$ and $z_i = M\Delta e_i = \Delta e_i - \iota\,\overline{\Delta e}_i$. Consider the first term on the right hand side of (33),

$$\begin{aligned}T^{-1/2}\sum_{s=2}^{t}z_{i,s} &= T^{-1/2}\Big(e_{i,t} - e_{i,1} - \frac{t-1}{T-1}[e_{i,T} - e_{i,1}]\Big) \\ &= T^{-1/2}\Big(e_{i,t} - \frac{t-1}{T-1}e_{i,T}\Big) + O_p(T^{-1/2}) \Rightarrow \sigma_i V_i(r)\end{aligned}$$

where V_i is a Brownian bridge, and σ_i^2 is the long run variance of Δe_{it}. Next,

$$T^{-1/2}\sum_{s=2}^{t}x_{i,t} = T^{-1/2}\Big(X_{i,t} - X_{i,1} - \frac{t-1}{T-1}(X_{i,T} - X_{i,1})\Big) = O_p(1).$$

The above being $O_p(1)$ holds even if X_{it} is a drifted random walk (containing a linear trend component). Thus the second term on the right hand side of (33) is $O_p(1)(\tilde{\beta}_i - \beta_i) = O_p(T^{-1/2})$. As in Bai and Ng (2004), we have

$$T^{-1/2}\sum_{s=2}^{t}v_s = T^{-1/2}\sum_{s=2}^{t}(\tilde{f}_s - Hf_s) = O_p(C_{NT}^{-1}),$$

and

$$d_i = \tilde{\lambda}_i - H^{-1\prime}\lambda_i = O_p(C_{NT}^{-1}).$$

Combining these results, we have

$$T^{-1/2}\tilde{e}_{i,t} = T^{-1/2}\sum_{s=2}^{t}\tilde{z}_{i,s} \Rightarrow \sigma_i V_i(r),$$

it follows that

$$MSB_{\tilde{e}} \xrightarrow{d} \int_0^1 V_i(r)^2 dr.$$

Next consider testing unit root in F_t for the case of $r = 1$. By definition, $\tilde{F}_t = \sum_{s=2}^{t}\tilde{f}_s$. Adding and subtracting,

$$T^{-1/2}\tilde{F}_t = T^{-1/2}H\sum_{s=2}^{t} f_s + T^{-1/2}\sum_{s=2}^{t}(\tilde{f}_s - Hf_s) = T^{-1/2}H\sum_{s=2}^{t} f_s + O_p(C_{NT}^{-1}).$$

But

$$T^{-1/2}\sum_{s=2}^{t} f_s = T^{-1/2}\Big[F_t - F_1 - \frac{t-1}{T-1}(F_T - F_1)\Big] \Rightarrow \sigma_w V_w(r).$$

It follows that

$$H^{-1}T^{-1/2}\tilde{F}_t \Rightarrow \sigma_w V_w(r),$$

and

$$MSB_{\tilde{F}} \xrightarrow{d} \int_0^1 V_w(r)^2 dr.$$

The proof of $r > 1$ is the similar to that of Bai and Ng (2004), thus omitted.

Table 1: Empirical size and power for the time trend case, when regressors are independent of the common factor. One common factor and $N = 40$

			$\rho_i = 1$				$\rho_i = 0.99$				$\rho_i = 0.95$			
T	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MSB_{\tilde{F}}$	$MSB_{\tilde{e}}$	P_m	P	$MSB_{\tilde{F}}$	$MSB_{\tilde{e}}$	P_m	P	$MSB_{\tilde{F}}$
50	1	0.5	0.014	0.042	0.035	0.051	0.019	0.047	0.040	0.050	0.252	0.239	0.217	0.058
100	1	0.5	0.028	0.051	0.043	0.052	0.052	0.075	0.064	0.051	0.990	0.923	0.913	0.049
250	1	0.5	0.038	0.050	0.043	0.050	0.432	0.312	0.289	0.050	1	1	1	0.053
50	0.95	0.5	0.015	0.036	0.030	0.061	0.019	0.051	0.044	0.073	0.251	0.226	0.207	0.066
100	0.95	0.5	0.025	0.049	0.040	0.107	0.048	0.076	0.067	0.105	0.991	0.937	0.924	0.116
250	0.95	0.5	0.031	0.060	0.053	0.377	0.441	0.331	0.301	0.382	1	1	1	0.411
50	0.9	0.5	0.012	0.035	0.029	0.110	0.016	0.040	0.033	0.108	0.242	0.224	0.201	0.102
100	0.9	0.5	0.021	0.040	0.032	0.258	0.044	0.071	0.061	0.261	0.993	0.927	0.916	0.259
250	0.9	0.5	0.031	0.053	0.045	0.788	0.438	0.316	0.291	0.775	1	1	1	0.835
50	1	1	0.019	0.042	0.036	0.050	0.019	0.050	0.043	0.048	0.253	0.225	0.203	0.045
100	1	1	0.024	0.048	0.041	0.046	0.056	0.085	0.074	0.049	0.992	0.923	0.913	0.045
250	1	1	0.033	0.059	0.050	0.051	0.446	0.316	0.288	0.054	1	1	1	0.052
50	0.95	1	0.014	0.039	0.032	0.065	0.023	0.046	0.040	0.061	0.238	0.221	0.201	0.063
100	0.95	1	0.025	0.046	0.040	0.107	0.054	0.072	0.063	0.113	0.991	0.933	0.921	0.117
250	0.95	1	0.027	0.049	0.042	0.392	0.434	0.317	0.292	0.402	1	1	1	0.407
50	0.9	1	0.013	0.034	0.029	0.099	0.018	0.041	0.036	0.113	0.237	0.221	0.2	0.104
100	0.9	1	0.025	0.048	0.042	0.270	0.046	0.068	0.059	0.277	0.993	0.928	0.914	0.281
250	0.9	1	0.027	0.055	0.048	0.823	0.441	0.311	0.282	0.839	1	1	1	0.853
50	1	10	0.017	0.039	0.033	0.044	0.020	0.043	0.035	0.031	0.180	0.182	0.161	0.049
100	1	10	0.026	0.048	0.044	0.048	0.045	0.073	0.064	0.038	0.959	0.869	0.853	0.044
250	1	10	0.030	0.052	0.044	0.050	0.417	0.302	0.277	0.051	1	1	1	0.050
50	0.95	10	0.014	0.038	0.032	0.058	0.017	0.043	0.037	0.041	0.179	0.176	0.157	0.069
100	0.95	10	0.023	0.051	0.044	0.113	0.051	0.075	0.065	0.086	0.947	0.847	0.828	0.112
250	0.95	10	0.027	0.052	0.045	0.433	0.411	0.305	0.279	0.437	1	1	1	0.434
50	0.9	10	0.016	0.039	0.035	0.104	0.016	0.039	0.034	0.076	0.168	0.182	0.162	0.116
100	0.9	10	0.025	0.045	0.039	0.305	0.042	0.069	0.059	0.295	0.946	0.844	0.825	0.298
250	0.9	10	0.030	0.054	0.046	0.873	0.395	0.283	0.259	0.880	1	1	1	0.891

Nominal size is set at the 5% level of significance. Results based on 5,000 replications

Table 2: Empirical size and power for the time trend case, when regressors are independent of the common factor. Three common factors and $N = 40$

							Non-parametric test				Parametric test			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	1	1	0.5	0.028	0.049	0.043	0.001	0.047	0.207	0.745	0.002	0.047	0.210	0.741
100	1	1	0.5	0.048	0.067	0.058	0.010	0.079	0.591	0.320	0.005	0.073	0.598	0.324
250	1	1	0.5	0.034	0.059	0.051	0.002	0.009	0.080	0.909	0	0.004	0.037	0.959
50	1	0.95	0.5	0.065	0.086	0.077	0.017	0.356	0.345	0.282	0.016	0.359	0.345	0.280
100	1	0.95	0.5	0.070	0.104	0.093	0.018	0.070	0.860	0.052	0.011	0.056	0.881	0.052
250	1	0.95	0.5	0.032	0.053	0.045	0.045	0.048	0.205	0.702	0.023	0.036	0.167	0.774
50	1	0.9	0.5	0.031	0.061	0.055	0.008	0.090	0.195	0.707	0.007	0.094	0.196	0.703
100	1	0.9	0.5	0.083	0.114	0.104	0.064	0.127	0.344	0.465	0.054	0.121	0.350	0.475
250	1	0.9	0.5	0.044	0.067	0.058	0.556	0.059	0.160	0.225	0.314	0.063	0.177	0.446
50	1	1	1	0.018	0.039	0.033	0	0.001	0.011	0.989	0	0.001	0.018	0.981
100	1	1	1	0.025	0.051	0.045	0.001	0.006	0.054	0.939	0.001	0.004	0.042	0.953
250	1	1	1	0.039	0.058	0.050	0.001	0.005	0.079	0.915	0	0.002	0.034	0.964
50	1	0.95	1	0.014	0.036	0.032	0	0.001	0.014	0.985	0	0.002	0.018	0.980
100	1	0.95	1	0.021	0.044	0.037	0.004	0.008	0.060	0.928	0.003	0.007	0.051	0.939
250	1	0.95	1	0.034	0.054	0.048	0.047	0.053	0.211	0.689	0.023	0.041	0.173	0.763
50	1	0.9	1	0.008	0.032	0.027	0.001	0.001	0.013	0.985	0.001	0.001	0.019	0.979
100	1	0.9	1	0.021	0.039	0.033	0.012	0.018	0.098	0.872	0.008	0.014	0.094	0.884
250	1	0.9	1	0.031	0.046	0.038	0.590	0.056	0.151	0.203	0.336	0.060	0.169	0.435
50	1	1	10	0.026	0.049	0.043	0	0.001	0.008	0.987	0	0.001	0.011	0.984
100	1	1	10	0.031	0.059	0.049	0	0.002	0.040	0.958	0	0.002	0.034	0.964
250	1	1	10	0.037	0.060	0.051	0	0.005	0.073	0.922	0	0.002	0.038	0.961
50	1	0.95	10	0.016	0.041	0.035	0	0.001	0.009	0.972	0.001	0.001	0.014	0.966
100	1	0.95	10	0.026	0.045	0.039	0.003	0.010	0.082	0.905	0.003	0.008	0.069	0.920
250	1	0.95	10	0.030	0.050	0.042	0.092	0.056	0.214	0.638	0.057	0.052	0.184	0.707
50	1	0.9	10	0.011	0.034	0.028	0.001	0.001	0.012	0.981	0.001	0.001	0.016	0.977
100	1	0.9	10	0.022	0.044	0.038	0.027	0.028	0.122	0.823	0.021	0.022	0.114	0.843
250	1	0.9	10	0.028	0.047	0.041	0.659	0.051	0.136	0.154	0.385	0.058	0.179	0.378

Nominal size is set at the 5% level of significance. Results based on 5,000 replications

Table 2 (cont): Empirical size and power for the time trend case, when regressors are independent of the common factor. Three common factors and $N = 40$

							Non-parametric test				Parametric test			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.99	1	0.5	0.041	0.064	0.057	0.005	0.077	0.220	0.698	0.004	0.078	0.222	0.696
100	0.99	1	0.5	0.058	0.075	0.066	0	0.003	0.057	0.940	0	0.003	0.050	0.948
250	0.99	1	0.5	0.440	0.315	0.290	0	0.008	0.080	0.912	0	0.005	0.058	0.937
50	0.99	0.95	0.5	0.048	0.072	0.065	0.008	0.157	0.477	0.358	0.010	0.160	0.475	0.355
100	0.99	0.95	0.5	0.047	0.070	0.060	0.001	0.008	0.093	0.898	0.001	0.007	0.087	0.905
250	0.99	0.95	0.5	0.417	0.307	0.283	0.042	0.048	0.205	0.705	0.023	0.040	0.166	0.771
50	0.99	0.9	0.5	0.035	0.067	0.061	0.004	0.055	0.237	0.704	0.005	0.056	0.238	0.701
100	0.99	0.9	0.5	0.049	0.069	0.062	0.013	0.022	0.113	0.852	0.009	0.017	0.108	0.866
250	0.99	0.9	0.5	0.414	0.300	0.276	0.548	0.046	0.177	0.229	0.447	0.082	0.186	0.285
50	0.99	1	1	0.024	0.047	0.040	0	0.004	0.041	0.955	0	0.004	0.044	0.952
100	0.99	1	1	0.056	0.08	0.070	0	0.003	0.041	0.956	0	0.003	0.032	0.965
250	0.99	1	1	0.432	0.314	0.289	0.001	0.007	0.080	0.912	0	0.004	0.060	0.936
50	0.99	0.95	1	0.022	0.044	0.038	0	0.002	0.085	0.913	0	0.004	0.089	0.907
100	0.99	0.95	1	0.047	0.074	0.065	0.002	0.008	0.056	0.934	0.001	0.007	0.049	0.943
250	0.99	0.95	1	0.404	0.297	0.276	0.071	0.051	0.206	0.672	0.04	0.048	0.173	0.739
50	0.99	0.9	1	0.019	0.041	0.034	0.001	0.002	0.012	0.985	0.001	0.002	0.018	0.979
100	0.99	0.9	1	0.041	0.064	0.057	0.017	0.02	0.106	0.857	0.014	0.016	0.099	0.871
250	0.99	0.9	1	0.407	0.290	0.266	0.606	0.054	0.156	0.184	0.498	0.099	0.173	0.230
50	0.99	1	10	0.026	0.050	0.044	0	0.001	0.005	0.992	0	0.001	0.010	0.987
100	0.99	1	10	0.048	0.074	0.063	0	0.004	0.053	0.943	0	0.002	0.044	0.954
250	0.99	1	10	0.340	0.261	0.236	0.001	0.010	0.098	0.891	0.001	0.008	0.073	0.918
50	0.99	0.95	10	0.020	0.044	0.038	0	0.001	0.008	0.985	0	0.001	0.012	0.980
100	0.99	0.95	10	0.041	0.07	0.061	0.003	0.012	0.064	0.919	0.001	0.010	0.060	0.926
250	0.99	0.95	10	0.300	0.229	0.206	0.056	0.062	0.194	0.688	0.03	0.050	0.158	0.762
50	0.99	0.9	10	0.014	0.035	0.028	0	0.001	0.013	0.983	0	0.001	0.020	0.976
100	0.99	0.9	10	0.042	0.058	0.050	0.013	0.021	0.104	0.861	0.01	0.017	0.095	0.877
250	0.99	0.9	10	0.311	0.235	0.216	0.670	0.057	0.123	0.150	0.559	0.107	0.135	0.199

Nominal size is set at the 5% level of significance. Results based on 5,000 replications

Table 2 (cont): Empirical size and power for the time trend case, when regressors are independent of the common factor. Three common factors and $N = 40$

							Non-parametric test				Parametric test			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.95	1	0.5	0.222	0.211	0.191	0	0.016	0.077	0.907	0	0.017	0.081	0.902
100	0.95	1	0.5	0.910	0.809	0.794	0.003	0.026	0.396	0.575	0.001	0.019	0.402	0.578
250	0.95	1	0.5	1	1	1	0	0.008	0.080	0.912	0.001	0.004	0.059	0.936
50	0.95	0.95	0.5	0.290	0.249	0.230	0.004	0.094	0.710	0.192	0.004	0.097	0.708	0.191
100	0.95	0.95	0.5	0.976	0.902	0.885	0.003	0.011	0.110	0.876	0.002	0.009	0.101	0.888
250	0.95	0.95	0.5	1	1	1	0.208	0.255	0.469	0.068	0.145	0.276	0.502	0.077
50	0.95	0.9	0.5	0.246	0.228	0.209	0.011	0.114	0.217	0.658	0.008	0.117	0.220	0.655
100	0.95	0.9	0.5	0.983	0.899	0.887	0.018	0.021	0.121	0.840	0.012	0.018	0.114	0.856
250	0.95	0.9	0.5	1	1	1	0.612	0.058	0.149	0.181	0.493	0.106	0.163	0.238
50	0.95	1	1	0.206	0.199	0.179	0	0.001	0.013	0.987	0.001	0.001	0.015	0.983
100	0.95	1	1	0.976	0.879	0.86	0	0.004	0.041	0.955	0	0.004	0.033	0.963
250	0.95	1	1	1	1	1	0.002	0.011	0.087	0.900	0.001	0.007	0.064	0.928
50	0.95	0.95	1	0.174	0.186	0.169	0	0.001	0.007	0.993	0	0	0.013	0.987
100	0.95	0.95	1	0.967	0.868	0.851	0.003	0.007	0.063	0.927	0.001	0.007	0.058	0.934
250	0.95	0.95	1	1	1	1	0.076	0.052	0.218	0.654	0.049	0.048	0.188	0.715
50	0.95	0.9	1	0.16	0.173	0.157	0	0.002	0.019	0.979	0.001	0.002	0.027	0.970
100	0.95	0.9	1	0.969	0.870	0.848	0.021	0.026	0.117	0.836	0.016	0.019	0.112	0.853
250	0.95	0.9	1	1	1	1	0.626	0.058	0.140	0.176	0.523	0.103	0.152	0.222
50	0.95	1	10	0.108	0.133	0.116	0	0.001	0.009	0.988	0.001	0	0.014	0.983
100	0.95	1	10	0.784	0.681	0.652	0.001	0.003	0.054	0.942	0.001	0.002	0.046	0.951
250	0.95	1	10	1	1	1	0.002	0.012	0.086	0.900	0.002	0.008	0.062	0.928
50	0.95	0.95	10	0.096	0.109	0.094	0	0.002	0.008	0.989	0	0.002	0.012	0.984
100	0.95	0.95	10	0.736	0.640	0.613	0.003	0.009	0.072	0.916	0.002	0.007	0.062	0.928
250	0.95	0.95	10	1	1	1	0.039	0.054	0.200	0.707	0.022	0.041	0.172	0.765
50	0.95	0.9	10	0.079	0.099	0.085	0	0.002	0.011	0.986	0	0.003	0.019	0.977
100	0.95	0.9	10	0.756	0.649	0.621	0.02	0.021	0.103	0.855	0.017	0.02	0.096	0.866
250	0.95	0.9	10	1	1	1	0.658	0.044	0.156	0.142	0.543	0.094	0.176	0.187

Nominal size is set at the 5% level of significance. Results based on 5,000 replications

Table 3: Empirical size and power for the linear time trend case, when regressors are correlated with the common factor. One common factor, panel BIC ($r_{max} = 6$) and $N = 40$

			$\rho_i = 1$				$\rho_i = 0.99$				$\rho_i = 0.95$				$\rho_i = 0.9$			
T	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MSB_{\tilde{F}}$	$MSB_{\tilde{e}}$	P_m	P	$MSB_{\tilde{F}}$	$MSB_{\tilde{e}}$	P_m	P	$MSB_{\tilde{F}}$	$MSB_{\tilde{e}}$	P_m	P	$MSB_{\tilde{F}}$
50	1	0.5	0.004	0.052	0.043	0.050	0.010	0.068	0.062	0.051	0.089	0.231	0.202	0.056	0.311	0.783	0.764	0.047
100	1	0.5	0.014	0.046	0.039	0.040	0.029	0.082	0.076	0.052	0.613	0.877	0.868	0.041	0.609	0.985	0.985	0.048
250	1	0.5	0.020	0.047	0.042	0.050	0.355	0.328	0.296	0.054	0.854	0.996	0.996	0.054	0.948	0.999	0.999	0.068
50	0.95	0.5	0.008	0.058	0.052	0.061	0.006	0.075	0.065	0.081	0.111	0.283	0.263	0.060	0.334	0.828	0.821	0.082
100	0.95	0.5	0.018	0.050	0.045	0.096	0.029	0.082	0.073	0.106	0.655	0.939	0.929	0.091	0.682	0.998	0.998	0.149
250	0.95	0.5	0.028	0.081	0.068	0.367	0.408	0.372	0.340	0.356	0.967	0.999	0.999	0.425	0.999	0.999	0.999	0.477
50	0.9	0.5	0.005	0.069	0.063	0.089	0.004	0.080	0.077	0.092	0.117	0.338	0.319	0.113	0.399	0.899	0.889	0.119
100	0.9	0.5	0.008	0.070	0.062	0.260	0.041	0.109	0.097	0.254	0.727	0.950	0.940	0.294	0.741	0.999	0.999	0.335
250	0.9	0.5	0.020	0.082	0.072	0.726	0.398	0.392	0.358	0.786	0.987	0.999	0.999	0.848	0.999	0.999	0.999	0.883
50	1	1	0.014	0.046	0.037	0.056	0.019	0.049	0.043	0.027	0.294	0.266	0.239	0.054	0.936	0.926	0.915	0.062
100	1	1	0.022	0.059	0.048	0.045	0.069	0.075	0.058	0.049	0.986	0.941	0.936	0.063	0.997	0.999	0.999	0.044
250	1	1	0.029	0.052	0.046	0.053	0.434	0.329	0.310	0.061	0.999	0.999	0.999	0.041	0.999	0.999	0.999	0.045
50	0.95	1	0.017	0.034	0.030	0.066	0.018	0.044	0.034	0.072	0.267	0.266	0.245	0.064	0.940	0.910	0.898	0.071
100	0.95	1	0.020	0.038	0.034	0.117	0.047	0.076	0.068	0.115	0.993	0.958	0.951	0.120	0.996	0.999	0.999	0.101
250	0.95	1	0.035	0.063	0.062	0.369	0.484	0.340	0.319	0.383	0.999	0.999	0.999	0.445	0.999	0.999	0.999	0.445
50	0.9	1	0.008	0.038	0.032	0.103	0.012	0.040	0.036	0.103	0.279	0.262	0.232	0.105	0.948	0.935	0.921	0.117
100	0.9	1	0.025	0.052	0.041	0.267	0.053	0.080	0.065	0.265	0.993	0.958	0.942	0.261	0.996	0.999	0.999	0.292
250	0.9	1	0.041	0.054	0.049	0.792	0.455	0.342	0.318	0.810	0.999	0.999	0.999	0.861	0.999	0.999	0.999	0.862
50	1	10	0.012	0.031	0.027	0.039	0.018	0.047	0.043	0.044	0.296	0.278	0.254	0.035	0.992	0.931	0.920	0.042
100	1	10	0.026	0.037	0.032	0.067	0.048	0.074	0.067	0.047	0.994	0.948	0.939	0.046	0.999	0.999	0.999	0.052
250	1	10	0.029	0.051	0.042	0.055	0.461	0.327	0.302	0.048	0.999	0.999	0.999	0.054	0.999	0.999	0.999	0.042
50	0.95	10	0.011	0.035	0.032	0.062	0.023	0.038	0.032	0.071	0.276	0.262	0.237	0.058	0.981	0.926	0.914	0.051
100	0.95	10	0.020	0.052	0.046	0.124	0.051	0.080	0.068	0.095	0.996	0.950	0.939	0.105	0.999	0.999	0.999	0.097
250	0.95	10	0.028	0.048	0.037	0.427	0.436	0.321	0.296	0.402	0.999	0.999	0.999	0.414	0.999	0.999	0.999	0.397
50	0.9	10	0.017	0.047	0.041	0.118	0.009	0.037	0.029	0.101	0.250	0.253	0.228	0.094	0.988	0.929	0.920	0.096
100	0.9	10	0.026	0.038	0.034	0.297	0.041	0.080	0.071	0.301	0.996	0.941	0.934	0.296	0.999	0.999	0.999	0.292
250	0.9	10	0.035	0.050	0.044	0.857	0.456	0.356	0.329	0.864	0.999	0.999	0.999	0.845	0.999	0.999	0.999	0.863

Nominal size is set at the 5% level of significance. Results based on 1,000 replications

Table 4: Time trend case, homogeneous parameters, three common factors correlated with stochastic regressors, panel BIC ($r_{max} = 6$) and $N = 40$

							Non-parametric				Parametric			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	1	1	0.5	0.023	0.047	0.038	0.000	0.001	0.007	0.993	0.000	0.001	0.010	0.990
100	1	1	0.5	0.020	0.054	0.045	0.000	0.002	0.028	0.970	0.000	0.002	0.023	0.975
250	1	1	0.5	0.040	0.052	0.047	0.000	0.002	0.077	0.921	0.000	0.002	0.030	0.968
50	1	0.95	0.5	0.012	0.029	0.025	0.000	0.001	0.009	0.990	0.000	0.001	0.014	0.986
100	1	0.95	0.5	0.031	0.050	0.042	0.000	0.002	0.064	0.934	0.000	0.001	0.055	0.944
250	1	0.95	0.5	0.032	0.060	0.050	0.003	0.041	0.189	0.764	0.000	0.012	0.105	0.882
50	1	0.9	0.5	0.009	0.030	0.026	0.000	0.001	0.008	0.992	0.000	0.001	0.025	0.975
100	1	0.9	0.5	0.011	0.038	0.034	0.000	0.012	0.094	0.894	0.000	0.010	0.091	0.899
250	1	0.9	0.5	0.027	0.040	0.035	0.242	0.179	0.287	0.289	0.103	0.118	0.241	0.535
50	1	1	1	0.018	0.043	0.035	0.000	0.001	0.004	0.996	0.000	0.001	0.004	0.995
100	1	1	1	0.021	0.057	0.051	0.000	0.001	0.032	0.968	0.000	0.001	0.024	0.976
250	1	1	1	0.040	0.054	0.047	0.000	0.002	0.071	0.927	0.000	0.001	0.026	0.974
50	1	0.95	1	0.016	0.033	0.030	0.000	0.001	0.004	0.996	0.000	0.001	0.008	0.992
100	1	0.95	1	0.033	0.044	0.037	0.000	0.001	0.051	0.948	0.000	0.001	0.042	0.957
250	1	0.95	1	0.026	0.053	0.045	0.004	0.029	0.207	0.757	0.000	0.011	0.105	0.884
50	1	0.9	1	0.011	0.035	0.029	0.000	0.001	0.005	0.995	0.000	0.001	0.009	0.991
100	1	0.9	1	0.019	0.036	0.034	0.000	0.015	0.079	0.906	0.000	0.009	0.080	0.910
250	1	0.9	1	0.025	0.042	0.038	0.332	0.166	0.253	0.246	0.145	0.103	0.235	0.514
50	1	1	10	0.014	0.036	0.032	0.000	0.001	0.007	0.993	0.000	0.001	0.009	0.991
100	1	1	10	0.035	0.049	0.039	0.000	0.001	0.028	0.971	0.000	0.001	0.023	0.976
250	1	1	10	0.025	0.053	0.047	0.000	0.002	0.063	0.935	0.000	0.001	0.029	0.971
50	1	0.95	10	0.012	0.035	0.031	0.000	0.001	0.007	0.992	0.000	0.001	0.018	0.981
100	1	0.95	10	0.021	0.037	0.029	0.000	0.003	0.042	0.955	0.000	0.001	0.038	0.962
250	1	0.95	10	0.041	0.051	0.047	0.008	0.036	0.214	0.739	0.000	0.013	0.086	0.898
50	1	0.9	10	0.014	0.029	0.026	0.000	0.001	0.006	0.994	0.000	0.001	0.013	0.987
100	1	0.9	10	0.024	0.049	0.045	0.000	0.014	0.106	0.879	0.000	0.010	0.095	0.894
250	1	0.9	10	0.025	0.044	0.039	0.467	0.147	0.181	0.202	0.243	0.107	0.198	0.449

Nominal size is set at the 5% level of significance. Results based on 1,000 replications

Table 5: Time trend case, homogeneous parameters, three common factors correlated with stochastic regressors, panel BIC ($r_{max} = 6$) and $N = 40$

							Non-parametric				Parametric			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.99	1	0.5	0.026	0.058	0.053	0.000	0.001	0.008	0.992	0.000	0.001	0.013	0.986
100	0.99	1	0.5	0.074	0.083	0.072	0.000	0.001	0.044	0.956	0.000	0.001	0.037	0.963
250	0.99	1	0.5	0.447	0.319	0.290	0.000	0.002	0.049	0.949	0.000	0.001	0.018	0.982
50	0.99	0.95	0.5	0.024	0.058	0.055	0.000	0.001	0.008	0.991	0.000	0.001	0.014	0.985
100	0.99	0.95	0.5	0.049	0.078	0.070	0.000	0.002	0.032	0.966	0.000	0.001	0.030	0.969
250	0.99	0.95	0.5	0.444	0.313	0.288	0.005	0.031	0.182	0.779	0.000	0.014	0.083	0.900
50	0.99	0.9	0.5	0.013	0.041	0.034	0.000	0.001	0.003	0.997	0.000	0.001	0.009	0.991
100	0.99	0.9	0.5	0.042	0.071	0.060	0.000	0.012	0.086	0.901	0.000	0.007	0.076	0.916
250	0.99	0.9	0.5	0.431	0.336	0.310	0.255	0.188	0.265	0.289	0.104	0.099	0.233	0.561
50	0.99	1	1	0.020	0.047	0.044	0.000	0.001	0.009	0.991	0.000	0.001	0.015	0.985
100	0.99	1	1	0.052	0.092	0.081	0.000	0.001	0.032	0.968	0.000	0.001	0.023	0.977
250	0.99	1	1	0.48	0.325	0.297	0.000	0.001	0.073	0.927	0.000	0.001	0.024	0.976
50	0.99	0.95	1	0.027	0.037	0.033	0.000	0.001	0.006	0.994	0.000	0.001	0.009	0.990
100	0.99	0.95	1	0.037	0.07	0.057	0.000	0.003	0.059	0.938	0.000	0.001	0.051	0.948
250	0.99	0.95	1	0.453	0.334	0.308	0.011	0.045	0.197	0.744	0.000	0.013	0.106	0.878
50	0.99	0.9	1	0.019	0.029	0.025	0.000	0.001	0.006	0.994	0.000	0.001	0.014	0.985
100	0.99	0.9	1	0.048	0.069	0.059	0.000	0.008	0.098	0.893	0.000	0.008	0.090	0.901
250	0.99	0.9	1	0.449	0.281	0.258	0.376	0.179	0.219	0.223	0.168	0.136	0.226	0.467
50	0.99	1	10	0.028	0.05	0.049	0.000	0.001	0.005	0.995	0.000	0.001	0.008	0.992
100	0.99	1	10	0.066	0.092	0.079	0.000	0.001	0.038	0.961	0.000	0.001	0.036	0.963
250	0.99	1	10	0.465	0.323	0.29	0.000	0.005	0.061	0.934	0.000	0.002	0.029	0.969
50	0.99	0.95	10	0.015	0.064	0.054	0.000	0.001	0.007	0.993	0.000	0.001	0.008	0.992
100	0.99	0.95	10	0.053	0.074	0.059	0.000	0.003	0.052	0.945	0.000	0.003	0.048	0.949
250	0.99	0.95	10	0.424	0.300	0.272	0.007	0.058	0.205	0.727	0.000	0.018	0.111	0.868
50	0.99	0.9	10	0.016	0.033	0.027	0.000	0.001	0.010	0.99	0.000	0.001	0.015	0.984
100	0.99	0.9	10	0.043	0.051	0.041	0.000	0.013	0.111	0.874	0.000	0.012	0.099	0.887
250	0.99	0.9	10	0.430	0.311	0.285	0.448	0.154	0.200	0.195	0.190	0.127	0.236	0.444

Nominal size is set at the 5% level of significance. Results based on 1,000 replications

Table 6: Time trend case, homogeneous parameters, three common factors correlated with stochastic regressors, panel BIC ($r_{max} = 6$) and $N = 40$

							Non-parametric				Parametric			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.95	1	0.5	0.278	0.268	0.244	0.000	0.001	0.003	0.997	0.000	0.001	0.006	0.994
100	0.95	1	0.5	0.991	0.929	0.915	0.000	0.001	0.037	0.962	0.000	0.001	0.034	0.965
250	0.95	1	0.5	1	1	1	0.000	0.001	0.072	0.927	0.000	0.001	0.026	0.973
50	0.95	0.95	0.5	0.249	0.219	0.199	0.000	0.001	0.011	0.989	0.000	0.001	0.015	0.985
100	0.95	0.95	0.5	0.991	0.929	0.917	0.000	0.003	0.043	0.954	0.000	0.003	0.035	0.962
250	0.95	0.95	0.5	1	1	1	0.006	0.055	0.236	0.700	0.000	0.026	0.108	0.864
50	0.95	0.9	0.5	0.224	0.204	0.187	0.000	0.001	0.010	0.989	0.000	0.001	0.016	0.983
100	0.95	0.9	0.5	0.996	0.95	0.939	0.000	0.012	0.101	0.886	0.000	0.006	0.099	0.894
250	0.95	0.9	0.5	1	1	1	0.429	0.150	0.209	0.209	0.195	0.115	0.222	0.465
50	0.95	1	1	0.273	0.248	0.217	0.000	0.001	0.006	0.994	0.000	0.001	0.008	0.992
100	0.95	1	1	0.995	0.94	0.934	0.000	0.001	0.031	0.968	0.000	0.001	0.028	0.971
250	0.95	1	1	1	1	1	0.000	0.004	0.081	0.915	0.000	0.002	0.029	0.969
50	0.95	0.95	1	0.242	0.236	0.211	0.000	0.003	0.004	0.993	0.000	0.002	0.005	0.993
100	0.95	0.95	1	0.993	0.947	0.936	0.000	0.006	0.054	0.940	0.000	0.002	0.051	0.947
250	0.95	0.95	1	1	1	1	0.006	0.054	0.225	0.712	0.000	0.014	0.132	0.851
50	0.95	0.9	1	0.252	0.232	0.207	0.000	0.001	0.009	0.991	0.000	0.001	0.019	0.981
100	0.95	0.9	1	0.992	0.931	0.92	0.000	0.011	0.090	0.898	0.000	0.007	0.084	0.908
250	0.95	0.9	1	1	1	1	0.411	0.149	0.229	0.208	0.189	0.111	0.234	0.463
50	0.95	1	10	0.278	0.25	0.231	0.000	0.001	0.005	0.995	0.000	0.001	0.008	0.992
100	0.95	1	10	0.994	0.938	0.926	0.000	0.001	0.040	0.96	0.000	0.001	0.030	0.970
250	0.95	1	10	1	1	1	0.000	0.001	0.072	0.927	0.000	0.001	0.027	0.973
50	0.95	0.95	10	0.26	0.252	0.221	0.000	0.001	0.005	0.995	0.000	0.001	0.009	0.990
100	0.95	0.95	10	0.987	0.941	0.936	0.000	0.004	0.052	0.944	0.000	0.004	0.045	0.951
250	0.95	0.95	10	1	1	1.000	0.017	0.060	0.222	0.698	0.000	0.028	0.116	0.854
50	0.95	0.9	10	0.24	0.232	0.204	0.000	0.001	0.011	0.989	0.000	0.001	0.017	0.983
100	0.95	0.9	10	0.991	0.924	0.918	0.000	0.013	0.100	0.884	0.000	0.011	0.088	0.900
250	0.95	0.9	10	1	1	1.000	0.473	0.129	0.214	0.181	0.236	0.114	0.228	0.419

Nominal size is set at the 5% level of significance. Results based on 1,000 replications

Table 7: Time trend case, heterogeneous parameters, three common factors correlated with stochastic regressors, panel BIC ($r_{max} = 6$) and $N = 40$

							Non-parametric				Parametric			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	1	1	0.5	0.026	0.044	0.040	0.001	0.018	0.063	0.917	0.001	0.018	0.065	0.915
100	1	1	0.5	0.025	0.036	0.028	0.000	0.000	0.041	0.959	0.000	0.001	0.033	0.966
250	1	1	0.5	0.033	0.054	0.048	0.000	0.002	0.061	0.937	0.000	0.001	0.026	0.973
50	1	0.95	0.5	0.015	0.035	0.030	0.003	0.011	0.054	0.932	0.002	0.012	0.056	0.930
100	1	0.95	0.5	0.015	0.036	0.034	0.000	0.001	0.057	0.942	0.000	0.001	0.053	0.946
250	1	0.95	0.5	0.031	0.044	0.037	0.003	0.034	0.199	0.764	0.000	0.007	0.101	0.892
50	1	0.9	0.5	0.013	0.033	0.029	0.003	0.009	0.052	0.936	0.003	0.011	0.052	0.934
100	1	0.9	0.5	0.012	0.037	0.030	0.000	0.005	0.095	0.900	0.000	0.003	0.096	0.901
250	1	0.9	0.5	0.029	0.039	0.036	0.194	0.163	0.285	0.358	0.074	0.098	0.228	0.600
50	1	1	1	0.024	0.044	0.038	0.000	0.000	0.006	0.993	0.000	0.000	0.010	0.989
100	1	1	1	0.024	0.033	0.029	0.000	0.000	0.033	0.967	0.000	0.001	0.027	0.972
250	1	1	1	0.032	0.052	0.047	0.000	0.002	0.060	0.938	0.000	0.003	0.022	0.975
50	1	0.95	1	0.016	0.039	0.027	0.000	0.000	0.007	0.993	0.000	0.000	0.011	0.989
100	1	0.95	1	0.015	0.034	0.032	0.000	0.001	0.057	0.942	0.000	0.001	0.047	0.952
250	1	0.95	1	0.031	0.045	0.037	0.003	0.041	0.219	0.737	0.002	0.010	0.101	0.887
50	1	0.9	1	0.013	0.028	0.022	0.000	0.001	0.010	0.989	0.000	0.001	0.014	0.985
100	1	0.9	1	0.013	0.036	0.026	0.000	0.009	0.096	0.895	0.000	0.005	0.089	0.906
250	1	0.9	1	0.030	0.038	0.036	0.301	0.155	0.262	0.282	0.112	0.100	0.246	0.542
50	1	1	10	0.025	0.042	0.038	0.000	0.001	0.005	0.993	0.000	0.000	0.007	0.992
100	1	1	10	0.023	0.034	0.026	0.000	0.002	0.031	0.967	0.000	0.001	0.029	0.970
250	1	1	10	0.034	0.058	0.047	0.000	0.003	0.062	0.935	0.000	0.003	0.019	0.978
50	1	0.95	10	0.015	0.036	0.029	0.000	0.001	0.005	0.993	0.000	0.001	0.007	0.991
100	1	0.95	10	0.014	0.031	0.028	0.000	0.002	0.060	0.938	0.000	0.001	0.050	0.949
250	1	0.95	10	0.031	0.045	0.040	0.006	0.040	0.226	0.728	0.002	0.010	0.109	0.879
50	1	0.9	10	0.012	0.025	0.021	0.000	0.001	0.010	0.988	0.000	0.001	0.013	0.985
100	1	0.9	10	0.012	0.032	0.027	0.000	0.015	0.102	0.883	0.000	0.010	0.091	0.899
250	1	0.9	10	0.030	0.045	0.037	0.447	0.132	0.202	0.219	0.218	0.115	0.208	0.459

Nominal size is set at the 5% level of significance. Results based on 1,000 replications

Table 8: Time trend case, heterogeneous parameters, three common factors correlated with stochastic regressors, panel BIC ($r_{max} = 6$) and $N = 40$

							Non-parametric				Parametric			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.99	1	0.5	0.026	0.045	0.039	0.001	0.019	0.069	0.910	0.001	0.019	0.073	0.906
100	0.99	1	0.5	0.047	0.060	0.054	0.000	0.001	0.039	0.960	0.000	0.001	0.034	0.965
250	0.99	1	0.5	0.414	0.306	0.267	0.000	0.003	0.062	0.935	0.000	0.000	0.024	0.976
50	0.99	0.95	0.5	0.019	0.045	0.035	0.003	0.014	0.058	0.925	0.003	0.016	0.057	0.924
100	0.99	0.95	0.5	0.038	0.054	0.048	0.000	0.001	0.058	0.941	0.000	0.001	0.058	0.941
250	0.99	0.95	0.5	0.395	0.285	0.264	0.003	0.027	0.206	0.764	0.002	0.008	0.101	0.889
50	0.99	0.9	0.5	0.014	0.035	0.027	0.003	0.013	0.054	0.930	0.003	0.015	0.055	0.927
100	0.99	0.9	0.5	0.033	0.053	0.048	0.000	0.004	0.095	0.901	0.000	0.003	0.097	0.900
250	0.99	0.9	0.5	0.391	0.282	0.258	0.209	0.163	0.285	0.343	0.083	0.105	0.236	0.576
50	0.99	1	1	0.028	0.047	0.042	0.000	0.000	0.005	0.994	0.000	0.000	0.011	0.988
100	0.99	1	1	0.045	0.058	0.054	0.000	0.001	0.034	0.965	0.000	0.001	0.030	0.969
250	0.99	1	1	0.415	0.305	0.266	0.000	0.002	0.058	0.940	0.000	0.001	0.024	0.975
50	0.99	0.95	1	0.020	0.037	0.033	0.000	0.000	0.007	0.993	0.000	0.000	0.010	0.990
100	0.99	0.95	1	0.036	0.054	0.044	0.000	0.001	0.057	0.942	0.000	0.001	0.050	0.949
250	0.99	0.95	1	0.396	0.285	0.264	0.004	0.037	0.213	0.746	0.003	0.011	0.108	0.878
50	0.99	0.9	1	0.016	0.032	0.026	0.000	0.001	0.008	0.991	0.000	0.001	0.014	0.985
100	0.99	0.9	1	0.032	0.051	0.042	0.000	0.010	0.096	0.894	0.000	0.005	0.090	0.905
250	0.99	0.9	1	0.390	0.285	0.253	0.319	0.149	0.239	0.293	0.128	0.112	0.230	0.530
50	0.99	1	10	0.028	0.052	0.040	0.000	0.001	0.006	0.992	0.000	0.000	0.007	0.992
100	0.99	1	10	0.048	0.061	0.051	0.000	0.002	0.030	0.968	0.000	0.001	0.027	0.972
250	0.99	1	10	0.418	0.293	0.258	0.000	0.004	0.061	0.935	0.000	0.003	0.019	0.978
50	0.99	0.95	10	0.024	0.041	0.034	0.000	0.001	0.005	0.993	0.000	0.001	0.007	0.991
100	0.99	0.95	10	0.035	0.055	0.046	0.000	0.002	0.058	0.940	0.000	0.001	0.048	0.951
250	0.99	0.95	10	0.400	0.279	0.254	0.006	0.043	0.220	0.731	0.002	0.010	0.110	0.878
50	0.99	0.9	10	0.014	0.033	0.023	0.000	0.001	0.009	0.989	0.000	0.001	0.012	0.986
100	0.99	0.9	10	0.035	0.048	0.041	0.000	0.013	0.103	0.884	0.000	0.010	0.089	0.901
250	0.99	0.9	10	0.394	0.278	0.255	0.444	0.132	0.202	0.222	0.218	0.114	0.209	0.459

Nominal size is set at the 5% level of significance. Results based on 1,000 replications

Table 9: Time trend case, heterogeneous parameters, three common factors correlated with stochastic regressors, panel BIC ($r_{max} = 6$) and $N = 40$

							Non-parametric				Parametric			
T	ρ_i	α	σ_F^2	$MSB_{\tilde{e}}$	P_m	P	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$	$MQ(0)$	$MQ(1)$	$MQ(2)$	$MQ(3)$
50	0.95	1	0.5	0.232	0.207	0.180	0.001	0.031	0.079	0.889	0.001	0.031	0.080	0.888
100	0.95	1	0.5	0.985	0.909	0.897	0.001	0.003	0.043	0.953	0.000	0.004	0.038	0.958
250	0.95	1	0.5	1	1	1	0.000	0.001	0.069	0.930	0.000	0.002	0.025	0.973
50	0.95	0.95	0.5	0.223	0.198	0.177	0.004	0.020	0.070	0.906	0.004	0.021	0.073	0.902
100	0.95	0.95	0.5	0.988	0.917	0.903	0.000	0.003	0.068	0.929	0.000	0.003	0.063	0.934
250	0.95	0.95	0.5	1	1	1	0.010	0.053	0.227	0.710	0.004	0.013	0.108	0.875
50	0.95	0.9	0.5	0.199	0.187	0.173	0.003	0.017	0.061	0.919	0.004	0.017	0.065	0.914
100	0.95	0.9	0.5	0.987	0.912	0.894	0.000	0.004	0.111	0.885	0.001	0.003	0.099	0.897
250	0.95	0.9	0.5	1	1	1	0.397	0.131	0.220	0.252	0.194	0.102	0.215	0.489
50	0.95	1	1	0.238	0.216	0.199	0.000	0.002	0.006	0.992	0.000	0.002	0.011	0.987
100	0.95	1	1	0.987	0.916	0.897	0.000	0.001	0.031	0.968	0.000	0.001	0.029	0.970
250	0.95	1	1	1	1	1	0.000	0.002	0.066	0.932	0.000	0.001	0.024	0.975
50	0.95	0.95	1	0.228	0.2	0.183	0.000	0.000	0.009	0.991	0.000	0.000	0.012	0.988
100	0.95	0.95	1	0.989	0.915	0.907	0.000	0.001	0.056	0.943	0.000	0.001	0.048	0.951
250	0.95	0.95	1	1	1	1	0.011	0.050	0.225	0.714	0.003	0.012	0.099	0.886
50	0.95	0.9	1	0.197	0.191	0.173	0.000	0.001	0.010	0.989	0.000	0.002	0.012	0.986
100	0.95	0.9	1	0.988	0.914	0.897	0.000	0.009	0.103	0.888	0.000	0.004	0.094	0.902
250	0.95	0.9	1	1	1	1	0.431	0.121	0.213	0.235	0.225	0.101	0.218	0.456
50	0.95	1	10	0.242	0.224	0.202	0.000	0.001	0.006	0.992	0.000	0.000	0.007	0.992
100	0.95	1	10	0.988	0.914	0.9	0.000	0.000	0.034	0.966	0.000	0.001	0.028	0.971
250	0.95	1	10	1	1	1	0.000	0.003	0.060	0.937	0.000	0.001	0.022	0.977
50	0.95	0.95	10	0.22	0.206	0.185	0.000	0.000	0.006	0.993	0.000	0.000	0.006	0.993
100	0.95	0.95	10	0.988	0.922	0.906	0.000	0.003	0.058	0.939	0.000	0.002	0.043	0.955
250	0.95	0.95	10	1	1	1	0.009	0.043	0.221	0.727	0.004	0.012	0.102	0.882
50	0.95	0.9	10	0.196	0.187	0.172	0.000	0.001	0.010	0.988	0.000	0.000	0.013	0.986
100	0.95	0.9	10	0.988	0.917	0.902	0.001	0.013	0.101	0.885	0.000	0.011	0.088	0.901
250	0.95	0.9	10	1	1	1	0.461	0.131	0.198	0.210	0.232	0.117	0.203	0.448

Nominal size is set at the 5% level of significance. Results based on 1,000 replications