# EUI WORKING PAPERS

ECO No. 2006/5

Cointegration in Panel Data with Breaks and Cross-Section Dependence

ANINDYA BANERJEE
AND JOSEP LLUÍS CARRION-I-SILVESTRE

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

# *Cointegration in Panel Data with Breaks and Cross-Section Dependence*

ANINDYA BANERJEE AND JOSEP LLUÍS CARRION-I-SILVESTRE

EUI Working Paper **ECO** No. 2006/5

ISSN 1725-6704



Printed in Italy
European University Institute
Badia Fiesolana
I – 50016 San Domenico di Fiesole (FI)
Italy

http://www.iue.it/
http://cadmus.iue.it/dspace/index.jsp

# Cointegration in panel data with breaks and cross-section dependence*

Anindya Banerjee
Department of Economics
European University Institute

Josep Lluís Carrion-i-Silvestre
Department of Econometrics, Statistics and Spanish Economy
University of Barcelona

February 10, 2006

**Abstract**

The power of standard panel cointegration statistics may be affected by misspecification errors if proper account is not taken of the presence of structural breaks in the data. We propose modifications to allow for one structural break when testing the null hypothesis of no cointegration that retain good properties in terms of empirical size and power. Response surfaces to approximate the finite sample moments that are required to implement the statistics are provided. Since panel cointegration statistics rely on the assumption of cross-section independence, a generalisation of the tests to the common factor framework is carried out in order to allow for dependence among the units of the panel.

**Keywords:** Panel cointegration, structural break, common factors, cross-section dependence
**JEL Codes:** C12, C22

## 1 Introduction

The theory of cointegration establishes that there exist linear combinations of integrated variables that cancel out common stochastic trends. This phenomenon gives rise to equilibrium relationships among integrated variables, which means that in the long-run these variables show co-movement or are cointegrated with each other. Although a large part of the traditional theory has been based upon the assumption of structural stability, the concept of cointegration *per se* does not rule out the possibility that both the cointegrating vector(s) and the deterministic component(s) of the long-run relationship might change during the time period analyzed. In fact, Hansen (1992), and Quintos and Phillips (1993) propose test statistics to assess the stability of the cointegration relationship. More interestingly, it is well known that if no account is taken of changes in the parameters of the model, inference concerning the presence of cointegration can be affected by misspecification errors. This in turn can bias conclusions towards accepting the null hypothesis of no cointegration – *e.g.* see Campos, Ericsson and Hendry (1996), and Gregory and Hansen (1996).

*A. Banerjee thanks the Research Department of the European Central Bank for hosting his visit in July 2005 within the framework of the Research Visitor programme. J. Ll. Carrion-i-Silvestre acknowledges financial support from the Generalitat de Catalunya, under grant 2003BEAI00308 of the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR). We also thank an anonymous referee for comments.

All these considerations have motivated the search for design procedures to test for cointegration allowing for structural breaks. Thus, Gregory and Hansen (1996) generalized the standard cointegration approach in Engle and Granger (1987) to allow for the presence of structural breaks that might affect either the deterministic component or the cointegration vector of the long-run relationship. Hao (1996), Bartley, Lee and Strazicich (2001), and Carrion-i-Silvestre and Sansó (2004) use the multivariate version of the KPSS statistic in Harris and Inder (1994), and Shin (1994) to test for the null of cointegration with one structural break. Finally, Hansen and Johansen (1999), and Busetti (2002) propose methods to estimate the cointegration rank in a multivariate framework.

These proposals are extremely relevant for the imperatives that arise in empirical modelling where structural breaks are very common. Gregory and Hansen (1996) and Gabriel, Da Silva and Nunes (2002) investigate the long-run money demand for the U.S. and Portugal, respectively. Busetti (2002) conducts two illustrations using road casualties in Great Britain, and macroeconomic data for the UK. Finally, Clemente, Marcuello, Montañés and Pueyo (2004) focus on health care expenditure demand functions. The main conclusion that arises from these applications is that inference on cointegration analysis can be affected by the presence of structural breaks. Other applications that may be envisaged for this methodology include looking at models of convergence, real exchange rates, exchange rate pass through and the issue of the solvency of the current account and its relation to the budget deficit, the so-called Feldstein-Horioka puzzle.

The literature on panel data econometrics with integrated data has experienced rapid development since the 1990s. The driving force behind the popularity of the use of the panel data techniques is the idea that the power of tests for unit roots and cointegration might be increased by combining the information that comes from the cross-section ($i = 1, ...N$) and the time ($t = 1, 2, .. > T$) dimensions, especially when the time dimension is restricted by the lack of availability of long series of reliable time-series data. As a result, new statistics to assess the stochastic properties of panel data sets have appeared in the literature – see Banerjee (1999), Baltagi and Kao (2000), and Baltagi (2005) for an overview of the field.

Surprisingly, the issue of instability has not received a great deal of attention in the panel data cointegration framework. In this regard, Kao and Chiang (2000) analyze instability in cointegration relationships assuming that cointegration is present, with a homogeneous cointegrating vector in all the units of the panel – although it is possible to split the panel into two sub-panels using a bootstrap scheme – and a common break point. Breitung (2005) proposes a VAR-based panel data cointegration procedure that permits the introduction of dummy variables outside the long-run relationship. Finally, Westerlund (2004) extends the LM statistic in McCoskey and Kao (1998) by allowing for structural breaks.

As may be seen, the scope of the literature that addresses the panel data cointegration hypothesis testing allowing for structural breaks is fairly limited. The first contribution of our paper is therefore to generalize the approach in Pedroni (1999, 2004) to account for one structural break that may affect the long-run relationship in a number of different ways. Our proposal applies more generally to the class of static-equation-based panel tests for cointegration but does not extend to cointegrated vector error correction models (VECM) for panels with integrated data for which more work is needed in order to develop feasible procedures.

Pedroni proposes seven statistics depending on the way that the individual information is combined to define the panel tests. The statistics can be grouped into either parametric or non-parametric statistics, depending on the way that autocorrelation and endogeneity bias are treated. In this paper we focus only on the parametric statistics, since these are at least asymptotically equivalent to their non-parametric counterparts. A Monte Carlo study, which could be constructed straightforwardly, would reveal the behaviour of non-parametric tests in finite samples when compared to the parametric tests, but is not included here solely for the sake of concision.

One important feature to consider in these tests is cross-section dependence. Most panel data statistics

– including those due to Pedroni – assume cross-section independence, except for common time effects. This is in many contexts a highly restrictive assumption to make. As our second contribution, we address this concern by using a factor model approach due to Bai and Ng (2004) to generalize the degree of permissible cross-section dependency to allow for idiosyncratic responses to multiple common factors.

Taken together we thereby generalize the class of panel cointegration tests to allow for both structural breaks and cross-section dependence. The limiting distributions of the statistics are derived and new sets of critical values are computed wherever required.

Our paper takes the following shape. In section 2 the interest of our proposal is motivated through Monte Carlo simulations. Section 3 presents the models and statistics for the null hypothesis of no cointegration with power against the alternative of broken cointegration. The moments that are required for the computation of the panel data statistics are computed in this section. In this regard, we estimate response surfaces to approximate these moments for whichever sample size. Section 4 extends the approach to the common factor framework. Section 5 focuses on the finite sample properties of the statistics. Finally, section 6 concludes with some remarks. Proofs are collected in the Appendix.

# 2 Motivation

Pedroni (1999, 2004) proposes seven statistics to test the null hypothesis of no cointegration using single-equation methods based on the estimation of static regressions. Since the statistics are based on single-equation methods the cointegrating rank for each unit is either 0 or 1, with a heterogeneous cointegrating vector for each unit. After estimating individual static regressions for each unit, the cointegrating residuals are used to compute each of the statistics. The seven statistics are classified into two different groups depending on whether they are within-dimension-based statistics – homogeneity is assumed when computing the cointegration test statistics – or between-dimension-based statistics where heterogeneous behaviour (across the units of the panel) is allowed. As mentioned in the introduction, we are concerned only with the parametric version of the statistics, i.e. the normalized bias and the pseudo $t$-ratio statistics.

To motivate our proposal we analyze the effects of structural breaks on the parametric group of Pedroni statistics through Monte Carlo simulations. First, we focus on the case where there is cointegration but the deterministic component changes at a point in time. Subsequently we also consider the case of an unstable cointegrating vector.

The Data Generating Process (DGP) is given by:

$$
\begin{aligned}
y_{i,t} &= f_i(t) + x'_{i,t}\delta_{i,t} + e_{i,t} \\
\Delta x_{i,t} &= v_{i,t} \\
e_{i,t} &= \rho_i e_{i,t-1} + \varepsilon_{i,t} \\
\zeta_{i,t} &= (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2),
\end{aligned}
$$

where $f_i(t)$ denotes the deterministic component.

Four different cases are considered. Firstly, we have $f_i(t) = \mu_i + \theta_i DU_{i,t}$ with $DU_{i,t} = 1$ for $t > T_{bi}$ and 0 otherwise, where $T_{bi} = \lambda_i T$ denotes the date of the break with $\lambda_i \in \Lambda$, where $\Lambda$ is a specified closed subset of $(0, 1)$.[1] The parameter set is given by $\mu_i = 1$, $\theta_i = \{0, 1, 3, 5, 10\}$, $\delta_{i,t} = \delta_i = 1$, and $\lambda_i = \{0.25, 0.5, 0.75\}$. The autoregressive parameter comes from the set $\rho_i = \{0, 0.5\}$. The sample size is $T = \{100, 200\}$, the number of units is $N = \{20, 40\}$ and the results are based on $1,000$ replications. For simplicity but without loss of generality, we have specified a common break point for all units in all

[1] We follow the suggestion in Zivot and Andrews (1992) and define $\Lambda = [2/T, (T-1)/T]$.

the simulations. The model that has been estimated to compute the pseudo $t$-ratio Pedroni panel data cointegration test statistics includes a constant term (individual effects) as deterministic component.

Secondly, we have also analyzed the case where the structural break changes both the level and the slope of the time trend. The deterministic function is given by $f_i(t) = \mu_i + \theta_i DU_{i,t} + \beta_i t + \gamma_i DT^*_{i,t}$, where $\mu_i = 1$, $\theta_i = 3$, $\beta_i = 0.3$ and $DT^*_{i,t}$ is the dummy variable defined above. Note that in this case the pseudo $t$-ratio statistic has been computed using a time trend as the deterministic component.

The third case studies the effects of a change both in the level and in the cointegrating vector. As before, the deterministic component is $f_i(t) = \mu_i + \theta_i DU_{i,t}$, with $\mu_i = 1$ and $\theta_i = \{0, 3\}$. Now we focus on the change in the cointegrating vector specifying $\delta_{i,t} = \delta_{i,1} = 1$ for $t \leq T_{bi}$ and $\delta_{i,t} = \delta_{i,2} = \{0, 2, 3, 4, 5, 10\}$ for $t > T_{bi}$. The model estimated to compute the (pseudo $t$-ratio) Pedroni panel data cointegration statistic includes a constant term as deterministic component.

Finally, the fourth case considers a change in the level and time trend, that defines the deterministic component, together with a change in the cointegrating vector. In this case $f_i(t) = \mu_i + \theta_i DU_{i,t} + \beta_i t + \gamma_i DT^*_{i,t}$, with $\mu_i = 1$, $\theta_i = 3$, $\beta_i = 0.3$, $\gamma_i = 0.5$, and $\delta_{i,t} = \delta_{i,1} = 1$ for $t \leq T_{bi}$ and $\delta_{i,t} = \delta_{i,2} = \{0, 2, 3, 4, 5, 10\}$ for $t > T_{bi}$. The model estimated to compute the pseudo $t$-ratio Pedroni panel data cointegration statistic includes individual and time effects.

Detailed results of the simulations for all four cases are available in Tables 1 to 3. In the first case, results in Table 1 show that the effect of a change in level only matters in those situations where the magnitude of the change is large and the break point is located at the end of the time period. Therefore, we can conclude that for small and moderate changes in level the misspecification error of the deterministic component does not damage the power of Pedroni statistic. However, in the second case the consequences of the misspecification error are more serious, since the empirical power approaches zero as the magnitude of the change in trend ($\gamma_i$) increases when the break point is placed either in the middle ($\lambda_i = 0.5$) or at the end ($\lambda_i = 0.75$) of the period. In the third case, Table 2 shows that for the empirical power to diminish the change in the cointegrating vector has to be either moderate or large, and be located in the middle ($\lambda_i = 0.5$) or at the end ($\lambda_i = 0.75$) of the period. Notice that this conclusion is reached irrespective of the change in level that affects the constant term.

Finally, when the level, time trend and the cointegrating vector change, and a model estimated to compute the pseudo $t$-ratio Pedroni panel data cointegration statistic includes individual and time effects, the change in the trend implies further reductions on the empirical power of the statistic when the break point is located in the middle and at the end of the period – see Table 3.

In summary, we may conclude that misspecification errors due to the lack of accounting for a structural break can reduce the power of the panel data cointegration test in Pedroni (2004) in those cases where the break point is placed in the middle or at the end of the time period. Therefore, we observe a bias towards the spurious non-rejection of the null hypothesis of no cointegration. A relevant feature is that the power distortions seem to appear only when the break changes either the slope of the time trend or the cointegrating vector, but no effects are seen when the break only affects the constant term.

# 3 Models and test statistics

In order to consider the issues described above more formally, let $\{Y_{i,t}\}$ be a $(m \times 1)$-vector of non-stationary stochastic process with the following representation

$$\Delta x_{i,t} = v_{i,t}$$
$$y_{i,t} = f_i(t) + x'_{i,t}\delta_{i,t} + e_{i,t}; \quad e_{i,t} = \rho_i e_{i,t} + \varepsilon_{i,t},$$

where $Y_{i,t} = \left(y_{i,t}, x'_{i,t}\right)'$ is conveniently partitioned into a scalar $y_{i,t}$ and the $((m-1)\times 1)$-vector $x_{i,t}$, $i = 1, \dots, N$, $t = 1, \dots, T$. Let $\xi_{i,t} = \left(\varepsilon_{i,t}, v'_{i,t}\right)'$ be a random sequence assumed to be strictly stationary and ergodic, with mean zero and finite variance. In addition, the partial sum process constructed from $\left\{\xi_{i,t}\right\}$ satisfy the multivariate invariance principle defined in Phillips and Durlauf (1986). At this stage and in order to set the analysis in a simplified framework, let us assume that $\{v_{i,t}\}$ and $\{\varepsilon_{i,t}\}$ are independent.

The general functional form for the deterministic term $f(t)$ is given by

$$f_i(t) = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT^*_{i,t}, \tag{1}$$

where

$$DU_{i,t} = \begin{cases} 0 & t \leq T_{bi} \\ 1 & t > T_{bi} \end{cases}; DT^*_{i,t} = \begin{cases} 0 & t \leq T_{bi} \\ (t - T_{bi}) & t > T_{bi} \end{cases},$$

with $T_{bi} = \lambda_i T$, $\lambda_i \in \Lambda$, denoting the time of the break for the $i$-th unit, $i = 1, \dots, N$. Note also that the cointegrating vector is specified as a function of time so that

$$\delta_{i,t} = \begin{cases} \delta_{i,1} & t \leq T_{bi} \\ \delta_{i,2} & t > T_{bi} \end{cases}.$$

Using these elements, we propose up to six different model specifications:

- Model 1. Constant term with a change in level but stable cointegrating vector:

$$y_{i,t} = \mu_i + \theta_i DU_{i,t} + x'_{i,t}\delta_i + e_{i,t} \tag{2}$$

- Model 2. Time trend with a change in level but stable cointegrating vector:

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + x'_{i,t}\delta_i + e_{i,t} \tag{3}$$

- Model 3. Time trend with change in both level and trend but stable cointegrating vector:

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_i + e_{i,t} \tag{4}$$

- Model 4. Constant term with change in both level and cointegrating vector:

$$y_{i,t} = \mu_i + \theta_i DU_{i,t} + x'_{i,t}\delta_{i,t} + e_{i,t} \tag{5}$$

- Model 5. Time trend with change in both level and cointegrating vector (the slope of trend does not change):

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + x'_{i,t}\delta_{i,t} + e_{i,t} \tag{6}$$

- Model 6. The time trend and the cointegrating vector change:

$$y_{i,t} = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_{i,t} + e_{i,t} \tag{7}$$

Using any one of these specifications we propose testing the null hypothesis of no cointegration against the alternative hypothesis of cointegration (with break) using the ADF test statistic applied to the residuals of the cointegration regression as in Engle and Granger (1987) and Gregory and Hansen (1996) but in the panel data framework developed in Pedroni (1999, 2004). In fact, Gregory and Hansen

(1996) propose the specifications given by models 1, 2 and 4 above, so that the specifications in models 3, 5 and 6 allow us to extend their approach.

Our proposal can be described in the following steps. First and following Gregory and Hansen (1996), we proceed to the OLS estimation of one of the models given in (2) to (7) and run the following ADF type-regression equation on the estimated residuals ($\hat{e}_{i,t}(\lambda_i)$):

$$\Delta\hat{e}_{i,t}(\lambda_i) = \rho_i \hat{e}_{i,t-1}(\lambda_i) + \sum_{j=1}^{k} \phi_{i,j} \Delta\hat{e}_{i,t-j}(\lambda_i) + \varepsilon_{i,t}. \tag{8}$$

The notation used refers to the break fraction ($\lambda_i$) parameter, which (if it exists) is in most cases unknown. In order to get rid of the dependence of the statistics on the break fraction parameter, Gregory and Hansen (1996) suggest estimating the models given in (2) to (7) for all possible break dates, subject to trimming, obtaining the estimated OLS residuals and computing the corresponding ADF statistic. With the sequence of ADF statistics in hand, we can also estimate the break point for each unit as the date that minimizes the sequence of individual ADF test statistics – either the $t$-ratio, $t_{\hat{\rho}_i}(\lambda_i)$, or the normalized bias, computed as $T\hat{\rho}_i(\lambda_i) = T\hat{\rho}_i\left(1 - \hat{\phi}_{i,1} - \cdots - \hat{\phi}_{i,k}\right)^{-1}$ – see Hamilton (1994), pp. 523. Gregory and Hansen (1996) derive the limiting distribution of $t_{\hat{\rho}_i}\left(\hat{\lambda}_i\right) = \inf_{\lambda_i\in\Lambda} t_{\rho_i}(\lambda_i)$ and $T\hat{\rho}_i\left(\hat{\lambda}_i\right) = \inf_{\lambda_i\in\Lambda} T\hat{\rho}_i(\lambda_i)$, which are shown not to depend on the break fraction parameter. Specifically, Gregory and Hansen (1996) show that $T\hat{\rho}_i\left(\hat{\lambda}_i\right) \Rightarrow \inf_{\lambda_i\in\Lambda} \int_0^1 Q(\lambda_i,s)\,dQ(\lambda_i,s) \Big/ \int_0^1 Q(\lambda_i,s)^2\,ds$, and $t_{\hat{\rho}_i}\left(\hat{\lambda}_i\right) \Rightarrow \inf_{\lambda_i\in\Lambda} \int_0^1 Q(\lambda_i,s)\,dQ(\lambda_i,s) \Big/ \left[\int_0^1 Q(\lambda_i,s)^2\,dr\left(1 + \varrho(\lambda_i)' D(\lambda_i)\varrho(\lambda_i)\right)\right]^{1/2}$, where $\Rightarrow$ denotes weak convergence, $Q(\lambda_i,s)$ and $\varrho(\lambda_i)$ are functions of Brownian motions and the deterministic component, and $D(\lambda_i)$ depends on the model – see the Theorem in Gregory and Hansen (1996) for further details. As mentioned above Gregory and Hansen (1996) deal only with some of the specifications in this paper, although their developments can be easily extended and similar limiting distributions obtained for the statistics. Note that the estimation of the break point $\hat{T}_{bi}$ is conducted as

$$\hat{T}_{bi} = \arg\min_{\lambda_i\in\Lambda} t_{\hat{\rho}_i}(\lambda_i)\,; \quad \hat{T}_{bi} = \arg\min_{\lambda_i\in\Lambda} T\hat{\rho}_i(\lambda_i)\,,$$

$\forall i = 1, \ldots, N$. At this point we could either follow Gregory and Hansen (1996) and test the null hypothesis for each unit or decide to combine the unit-specific information in a panel data statistic.

The panel statistics on which we focus in order to test the null hypothesis are given by the $Z_{\hat{\rho}_{NT}}$ and $Z_{\hat{t}_{NT}}$ tests in Pedroni (1999, 2004), which can be thought as analogous to the residual-based tests in Engle and Granger (1987). These test statistics are defined by pooling the individual ADF tests, so that they belong to the class of between-dimension test statistics. Specifically, they are computed as:

$$N^{-1/2} Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right) = N^{-1/2} \sum_{i=1}^{N} T\hat{\rho}_i\left(\hat{\lambda}_i\right) \tag{9}$$

$$N^{-1/2} Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right) = N^{-1/2} \sum_{i=1}^{N} t_{\hat{\rho}_i}\left(\hat{\lambda}_i\right). \tag{10}$$

where $\hat{\rho}_i\left(\hat{\lambda}_i\right)$ and $t_{\hat{\rho}_i}\left(\hat{\lambda}_i\right)$ are the estimated coefficient and associated $t$-ratio from (8) and

$$\hat{\lambda} = \left(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_i, \ldots, \hat{\lambda}_N\right)'$$

is the vector of estimated break fractions.

Note that in this framework we allow for a high degree of heterogeneity since the cointegrating vector, the short run dynamics and the break point estimate might differ among units. The use of the panel data cointegration test aims to increase the power of the statistical inference when testing the null hypothesis of no cointegration, but some heterogeneity is preserved when conducting the estimation of the parameters individually.

Following Pedroni (1999, 2004), the panel test statistics are shown to converge to standard Normal distributions once they have been properly standardized.

**Theorem 1** *Let* $\Theta$ *and* $\Psi$ *denote the mean and variance for the vector Brownian motion funcional* $\Upsilon' \equiv \left(\inf_{\lambda_i\in\Lambda} \int_0^1 Q(\lambda_i, s)\, dQ(\lambda_i, s) \left[\int_0^1 Q(\lambda_i, s)^2\, ds\right]^{-1}, \inf_{\lambda_i\in\Lambda} \int_0^1 Q(\lambda_i, s)\, dQ(\lambda_i, s) \times \left[\int_0^1 Q(\lambda_i, s)^2\, ds \;\; \left(1+\varrho(\lambda_i)' D(\lambda_i)\, \varrho(\lambda_i)\right)\right]^{-1/2}\right)$*. Then, under the null hypothesis of no cointegration the asymptotic distribution of the statistics* $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ *and* $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ *defined in (9) and (10), respectively, are given by*

$$\begin{aligned} N^{-1/2} Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right) - \Theta_1\sqrt{N} &\Rightarrow N(0, \Psi_1) \\ N^{-1/2} Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right) - \Theta_2\sqrt{N} &\Rightarrow N(0, \Psi_2), \end{aligned}$$

*as* $(T, N \to \infty)_{\text{seq}}$*, where* $\Rightarrow$ *denotes weak convergence.*

As in Pedroni (2004), in order to prove Theorem 1 we require only the assumption of finite second moments of the random variables characterized as Brownian motion functionals, which will allow to apply the Lindberg-Levy Central Limit Theorem as $N \to \infty$.

The moments of the limiting distributions, $\Theta_1, \Psi_1, \Theta_2$ and $\Psi_2$, are approximated by Monte Carlo simulation for the different specifications and allowing up to seven stochastic regressors in the cointegrating relationship – i.e. the dimension of the $Y_{i,t}$ $(m \times 1)$-vector goes from $(2 \times 1)$ to $(8 \times 1)$. Table 4 presents the moments of the limit distributions based on $T = 1,000$. As can be seen, the moments of the distribution depends both on the specification and the number of stochastic regressors.

Since the limiting distribution of the tests can provide a poor approximation in finite samples, we have approximated the moments of the test statistics for different values of the sample size, specifically $T = \{30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 400, 500, 1,000\}$. In addition, the finite sample distributions depend on the procedure that is applied when selecting the order $(k)$ of the parametric correction in (8). The results reported in Tables 5 to 8 are fixed lag lenght at for $k = 0,\ 2$ and 5, and lag length selection using the $t$-sig criterion in Ng and Perron (1995) with a $k_{\max} = 5$ as the maximum order of lags, respectively. Since reporting the moments of the finite sample distribution for the different values of $T$ and the number of stochastic regressors $p = (m-1)$. The general functional form that has been estimated is

$$g(T, p) = \sum_{l=0}^{3} \left(\beta_{0,l} + \beta_{1,l}\frac{1}{T} + \beta_{2,l}\frac{1}{T^2} + \beta_{3,l}\frac{1}{T^3}\right) p^l,$$

where $g(T, p)$ in the relevant columns of Tables 5 to 8 refer to $\Theta_1, \Psi_1, \Theta_2$ and $\Psi_2$, for the different model specifications. These functions have been estimated by OLS using the Newey-West robust covariance disturbance matrix to assess the individual significance of the regressors – the level of significance is 10%. A GAUSS code is available from the authors to compute the statistics and corresponding moments. In all simulations 10,000 replications were used to simulate the moments.

# 4 Common factors in panel cointegration

In the sections above, we have generalized static-regression-based tests for cointegration to include structural breaks in the deterministic components of the processes. These derivations are valid only under the assumption that the units are cross-sectionally independent. However, this requirement is rarely likely to be satisfied in empirical economic applications where countries or regions depend each other. Therefore, in order to generalize the framework and applicability of the paper further, we have extended our approach to allow for cross-section dependence. We model such dependence by using common factors as in Bai and Ng (2004). In addition to dependence, our tests also can accommodate the presence of structural breaks.[2] We deal first with the case where the break date is known and then proceed to the more realistic scenario of an unknown break date.

## 4.1 Break point known

In this framework the model is given in structural form as:

$$y_{i,t} = f_i(t) + x_{i,t}'\delta_{i,t} + u_{i,t} \tag{11}$$

$$u_{i,t} = F_t'\pi_i + e_{i,t} \tag{12}$$

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

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

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

$t = 1, \ldots, T$, $i = 1, \ldots, N$, where $C(L) = \sum_{j=0}^{\infty} C_j L^j$, and $f_i(t)$ denotes the deterministic component (which may be broken as in 1 above), $F_t$ denotes a $(r \times 1)$-vector containing the common factors, with $\pi_i$ the vector of loadings. Despite the operator $(1 - L)$ in equation (13), $F_t$ does not have to be I(1). In fact, $F_t$ can be I(0), I(1), or a combination of both, depending on the rank of $C(1)$. If $C(1) = 0$, then $F_t$ is I(0). If $C(1)$ is of full rank, then each component of $F_t$ is I(1). If $C(1) \neq 0$, but not full rank, then some components of $F_t$ are I(1) and some are I(0). Our analysis is based on the same set of assumptions in Bai and Ng (2004), and Bai and Carrion-i-Silvestre (2005). Let $M < \infty$ be a generic positive number, not depending on $T$ and $N$:

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

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

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

*Assumption D:* The errors $\varepsilon_{i,t}$, $w_t$, and the loadings $\pi_i$ are three mutually independent groups.

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

*Assumption F:* (i) $v_{i,t} \sim iid(0, \Sigma_v)$, $E\|v_{i,t}\|^4 \leq M$, and (ii) $Var\left(\Delta x_{i,t}'\right) = \sum_{j=0}^{\infty} G_{i,j} \Sigma_v G_{i,j}' > 0$, (iii) $\sum_{j=0}^{\infty} j\|G_{i,j}\| < M$; and (iv) $G(1)$ has full rank.

*Assumption G:* (i) $E(e_{i,t}|v_{i,t}) = 0$ when stochastic regressors are assumed to be strictly exogenous or (ii) $E(e_{i,t}|v_{i,t}) = \Delta x_{i,t}' A_i(L) + \xi_{i,t}$, with $A_i(L)$ being a $(k \times 1)$-vector of lags and leads polynomials of finite orders and $\xi_{i,t} \sim iid(0, \Sigma_\xi)$, when stochastic regressors are non-strictly exogenous.

Assumption A ensures that the factor loadings are identifiable. Assumption B establishes the conditions on the short and long-run variance of $\Delta F_t$ – i.e. the short-run variance matrix is positive definite and the long-run variance matrix may have reduced rank in order to accommodate stationary linear combinations of $I(1)$ factors. Assumption C(i) allows for some weak serial correlation in $(1 - \rho_i L) e_{i,t}$,

[2] An alternative approach to dealing with cross-sectional dependence is proposed by Chang (2005) using a non-linear IV technique.

whereas C(ii) and C(iii) allow for weak cross-section correlation. Assumption E defines the initial conditions. Assumption F establishes conditions on the first differences of the stochastic regressors. Finally, Assumption G defines two situations depending on whether the stochastic regressors are strictly exogenous regressors or non-strictly exogenous. This distinction is important here, because in the common factor framework the limiting distributions of the statistics do not depend on the number of stochastic regressors if strict exogeneity holds. However, this is no longer true when correlation between $e_{i,t}$ and $v_{i,s}$ is allowed and modifications need to be introduced to account for endogenous regressors. Here we suggest using the DOLS estimation method in Stock and Watson (1993) to account for endogeneity, where we assume that the number of leads and lags is fixed as in Stock and Watson (1993), although they can be chosen using a BIC information criterion.[3]

For ease of exposition, we assume strictly exogenous stochastic regressors, although the Appendix contains a discussion of the more general case. The estimation of the common factors is done as in Bai and Ng (2004). We compute the first differences:

$$\Delta y_{i,t} = \Delta f_i(t) + \Delta x'_{i,t}\delta_{i,t} + \Delta F_t \pi_i + \Delta e_{i,t},$$

and take the orthogonal projections:

$$\begin{aligned} M_i \Delta y_i &= M_i \Delta F \pi_i + M_i \Delta e_i \\ &= f\pi_i + z_i, \end{aligned} \tag{16}$$

with $M_i = I - \Delta x_i^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime}$ being the idempotent matrix, and $f = M_i \Delta F$ and $z_i = M_i \Delta e_i$. The superscript $d$ in $\Delta x_i^d$ indicates that there are deterministic elements. The estimation of the common factors and factor loadings can be done as in Bai and Ng (2004) using principal components. Specifically, 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 $y^* y^{*\prime}$, where $y_i^* = M_i \Delta y_i$. Under the normalization $\tilde{f}\tilde{f}'/(T-1) = I_r$, the estimated loading matrix is $\tilde{\Pi} = \tilde{f}' y^* / (T-1)$. Therefore, the estimated residuals are defined as

$$\tilde{z}_{i,t} = y^*_{i,t} - \tilde{f}_t \tilde{\pi}_i. \tag{17}$$

We can recover the idiosyncratic disturbance terms through cumulation, i.e. $\tilde{e}_{i,t} = \sum_{j=2}^{t} \tilde{z}_{i,j}$, and test the unit root hypothesis ($\alpha_{i,0} = 0$) using the ADF regression equation

$$\Delta \tilde{e}_{i,t}\left(\hat{\lambda}_i\right) = \alpha_{i,0} \tilde{e}_{i,t-1}\left(\hat{\lambda}_i\right) + \sum_{j=1}^{k} \alpha_{i,j} \Delta \tilde{e}_{i,t-j}\left(\hat{\lambda}_i\right) + \varepsilon_{i,t}. \tag{18}$$

We denote by $ADF_{\tilde{e}}^{c}(i)$, $ADF_{\tilde{e}}^{\tau}(i)$ and $ADF_{\tilde{e}}^{\gamma}(i)$ the pseudo $t$-ratio ADF statistics for testing $\alpha_{i,0} = 0$ in (18), for the model that includes a constant, a linear time trend, and a time trend with a change in trend, respectively. When $r = 1$ we can use an ADF-type equation to analyze the order of integration of $F_t$ as well. However, in this case we need to proceed in two steps. In the first step we regress $\tilde{F}_t$ on the deterministic specification and the stochastic regressors. In the second step we estimate the ADF

[3] In the more standard panel cointegration framework without common factors, as discussed above, the distributions of the test statistics have been computed without making any assumption about strict exogeneity. The distributions depend on the number of regressors in the model and this is reflected in Table 4 for the asymptotic moments and the response surfaces in Tables 5 to 8, both computed for varying $m$.

regression equation using the detrended common factor $\left(\tilde{F}_t^d\right)$, i.e. the residuals of the first step:

$$\Delta \tilde{F}_t^d = \delta_0 \tilde{F}_{t-1}^d + \sum_{j=1}^{k} \delta_j \Delta \tilde{F}_{t-j}^d + u_t, \tag{19}$$

and test if $\delta_0 = 0$ – $ADF_{\tilde{F}}^d(\lambda)$ denotes the pseudo $t$-ratio ADF statistic for testing $\delta_0 = 0$ in (19).

Finally, if $r > 1$ we should use one of the two statistics proposed in Bai and Ng (2004) to fix the number of common stochastic trends ($q$). As before, let $\tilde{F}_t^d$ denote the detrended common factors. Start with $q = r$ and proceed in three stages – we reproduce these steps here for completeness:

1. Let $\tilde{\beta}_\perp$ be the $q$ eigenvectors associated with the $q$ largest eigenvalues of $T^{-2}\sum_{t=2}^{T} \tilde{F}_t^d \tilde{F}_t^{d\prime}$.
2. Let $\tilde{Y}_t^d = \tilde{\beta}_\perp \tilde{F}_t^d$, from which we can define two statistics:
   (a) Let $K(j) = 1 - j/(J+1)$, $j = 0, 1, 2, \ldots, J$:
      i. Let $\tilde{\xi}_t^d$ be the residuals from estimating a first-order VAR in $\tilde{Y}_t^d$, and let

$$\tilde{\Sigma}_1^d = \sum_{j=1}^{J} K(j) \left( T^{-1} \sum_{t=2}^{T} \tilde{\xi}_t^d \tilde{\xi}_t^{d\prime} \right).$$

      ii. Let $\tilde{v}_c^d(q) = \frac{1}{2}\left[\sum_{t=2}^{T}\left(\tilde{Y}_t^d \tilde{Y}_{t-1}^{d\prime} + \tilde{Y}_{t-1}^d \tilde{Y}_t^{d\prime}\right) - T\left(\tilde{\Sigma}_1^d + \tilde{\Sigma}_1^{d\prime}\right)\right]\left(T^{-1}\sum_{t=2}^{T} \tilde{Y}_{t-1}^d \tilde{Y}_{t-1}^{d\prime}\right)^{-1}$.
      iii. Define $MQ_c^d(q) = T\left[\tilde{v}_c^d(q) - 1\right]$ for the case of no change in the trend and $MQ_c^d(q,\lambda) = T\left[\tilde{v}_c^d(q,\lambda) - 1\right]$ for the case of a change in the trend.
   (b) For $p$ fixed that does not depend on $N$ and $T$:
      i. Estimate a VAR of order $p$ in $\Delta \tilde{Y}_t^d$ to obtain $\tilde{\Pi}(L) = I_q - \tilde{\Pi}_1 L - \ldots - \tilde{\Pi}_p L^p$. Filter $\tilde{Y}_t^d$ by $\tilde{\Pi}(L)$ to get $\tilde{y}_t^d = \tilde{\Pi}(L)\tilde{Y}_t^d$.
      ii. Let $\tilde{v}_f^d(q)$ be the smallest eigenvalue of

$$\Phi_f^d = \frac{1}{2}\left[\sum_{t=2}^{T}\left(\tilde{Y}_t^d \tilde{Y}_{t-1}^{d\prime} + \tilde{Y}_{t-1}^d \tilde{Y}_t^{d\prime}\right)\right]\left(T^{-1}\sum_{t=2}^{T} \tilde{Y}_{t-1}^d \tilde{Y}_{t-1}^{d\prime}\right)^{-1}.$$

      iii. Define the statistic $MQ_f^d(q) = T\left[\tilde{v}_f^d(q) - 1\right]$ for the case of no change in the trend and $MQ_f^d(q,\lambda) = T\left[\tilde{v}_f^d(q,\lambda) - 1\right]$ for the case of a change in the trend.
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 following Theorem offers the main results concerning these statistics.

**Theorem 2** *Let $\{y_{i,t}\}$ be the stochastic process with DGP given by (11) to (15). The following results hold as $N, T \to \infty$. Let $k$ be the order of autoregression chosen such that $k \to \infty$ and $k^3/\min[N,T] \to 0$.*
*(1) Under the null hypothesis that $\rho_i = 1$ in (14),*
*(1.a) for the specification that does not include a time trend, with or without change in level:*

$$ADF_{\tilde{e}}^c(i) \Rightarrow \frac{\frac{1}{2}\left(W_i(1)^2 - 1\right)}{\left(\int_0^1 W_i(s)^2 ds\right)^{1/2}},$$

*(1.b) for those specifications including a time trend with or without change in level:*

$$ADF_{\tilde{e}}^{\tau}(i) \Rightarrow -\frac{1}{2}\left(\int_0^1 V_i(s)^2\, ds\right)^{-1/2},$$

*where* $V_i(s) = W_i(s) - sW_i(1)$.
*(1.c) for those specifications including a time trend with change in trend:*

$$ADF_{\tilde{e}}^{\gamma}(i) \Rightarrow -\frac{1}{2}\left(\lambda^2 \int_0^1 V_i(b_1)^2\, dr + (1-\lambda)^2 \int_0^1 V_i(b_2)^2\, dr\right)^{-1/2},$$

*where* $V_i(b_j) = W_i(b_j) - b_j W_i(1)$, $j = 1, 2$, *are two independent detrended Brownian processes.*
*(2) When* $r = 1$, *under the null hypothesis that* $F_t$ *has a unit root and no change in trend:*

$$ADF_{\tilde{F}}^{d} \Rightarrow \frac{\int_0^1 W_w^d(s)\, dW_w^d(s)}{\left(\int_0^1 W_w^d(s)^2\, ds\right)^{1/2}},$$

*where* $W_w^d(s)$ *denotes the detrended Brownian motion, while when we allow for change in trend:*

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

*where* $W_w^d(s,\lambda)$ *is the detrended Brownian motion and* $\lambda$ *denotes the break fraction parameter.*
*(3) When* $r > 1$, *let* $W_q$ *be a q-vector of standard Brownian motion and* $W_q^d$ *the detrended counterpart. Let* $v_*^d(q)$ *be the smallest eigenvalues of the statistic computed for a model that does not include change in trend. Then:*

$$\Phi_*^d = \frac{1}{2}\left[W_q^d(1)\, W_q^d(1)' - I_p\right]\left[\int_0^1 W_q^d(s)\, W_q^d(s)'\, ds\right]^{-1},$$

*and letting* $v_*^d(q,\lambda)$ *be the smallest eigenvalues of the statistic computed for the model that includes change in trend:*

$$\Phi_*^d(\lambda) = \frac{1}{2}\left[W_q^d(1,\lambda)\, W_q^d(1,\lambda)' - I_p\right]\left[\int_0^1 W_q^d(s,\lambda)\, W_q^d(s,\lambda)'\, ds\right]^{-1},$$

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

The proof of the Theorem is outlined in the Appendix. Some remarks are in order. First, note that the definition of the common factors framework implies that the matrix of projections $M_i$ that is used above cannot depend on $i$, which means that all elements that are defined in $\Delta x_i^d$ should be the same across $i$. There are two different kind of elements in $\Delta x_i^d$: (i) the deterministic regressors and (ii) the stochastic regressors. Regarding the latter, we have shown in the Appendix that the limiting distribution of the statistics do not depend on the presence of stochastic regressors, so that we can ignore the effect of these elements when defining $M_i$. Unfortunately, this is not true for the deterministic regressors. Thus, to warrant that $M_i$ does not (asymptotically) depend on $i$ we have to assume common break dates, i.e. we assume that the break points are the same for all units. This restriction can be seen as a limitation of our analysis, but in fact it is due to the definition of the common factors framework. Thus, (16) specifies a common factor structure for all units, so that $f_t$ cannot depend on $i$. If we look at the definition of $f_t = M_i \Delta F_t$ we can see that the specification of heterogeneous structural breaks implies that the idempotent matrix $M_i$ depends on $i$. The only way to overcome this situation is to impose $M_i = M \ \forall i$ so that the structural breaks are the same for all units. This is the reason why in Theorem 1 we have not included any subscript on $\lambda$ for the units.

Second, the limiting distribution of the ADF statistic for the idiosyncratic disturbance term does not depend on the presence of stochastic regressors. Moreover, the presence of changes in level does not affect the limiting distribution of the ADF statistic that is computed using the idiosyncratic disturbance term.

Third, the distributions of the statistics that focus on the common factors depend on some elements that define the deterministic component although, surprisingly, they do not depend on the number of stochastic regressors. Specifically, the presence of changes in level does not affect the limiting distribution of the ADF and $\Phi_*^d$ statistics, although this is not true when there are changes in trend. For the latter, the test statistics depends on the number and location of the structural breaks. Moreover, in this case we have to assume that these structural breaks are common to all units.

Finally, some remarks should be made concerning the limiting distributions of the statistics derived in Theorem 1. The limiting distributions for $ADF_{\tilde{e}}^{c}(i)$ and $ADF_{\tilde{F}}^{d}$ derived in *(1.a)* and *(2)* are the standard Dickey-Fuller distributions for constant and constant and trend respectively. The moments for $ADF_{\tilde{e}}^{c}(i)$, $ADF_{\tilde{e}}^{\tau}(i)$ and $ADF_{\tilde{e}}^{\gamma}(i)$ for different sample sizes are reported in Table 9 which are used to compute the pooled test given by (20) below.

The ADF statistic when there is one structural break given by $ADF_{\tilde{F}}^{d}(\lambda)$ derived in *(2)* can be found in Perron (1989) for the specification denoted as Model C. The limiting distributions of the MQ tests without break stated in *(3)* may be found in Bai and Ng (2004), while the corresponding distributions for a single known break point, $\Phi_*^d(\lambda)$, have been simulated by us and are reported in Table 10. The asymptotic critical values reported in this table depend both on the number of stochastic common trends and on the break fraction. Note however that these critical values correspond to the case of only one structural break, though our approach can be easily extended to multiple changes in trend.

The individual ADF statistics for the idiosyncratic disturbance terms can be pooled to define a panel data cointegration test. Thus, following the steps given in the previous section we can define

$$N^{-1/2} Z_{\hat{t}_{NT}}^{e}(\lambda) - \Theta_2^e(\lambda)\sqrt{N} \Rightarrow N\left(0, \Psi_2^e(\lambda)\right), \tag{20}$$

where the superscript $e$ denotes the idiosyncratic disturbance term using our results in *(1)* of Theorem 1 above. The moments $\Theta_2^e(\lambda)$ and $\Psi_2^e(\lambda)$ depend on the deterministic specification used and, except for the case of changes in trend, are the same as the ones for the statistics in Bai and Ng (2004) (where these do not depend on the break fraction $\lambda$).[4] Table 9 reports finite sample moments $\Theta_2^e(\lambda)$ and $\Psi_2^e(\lambda)$ for the different statistics and different values of $T$.

## 4.2 Break point unknown

Up to now developments in this section have been based on the implicit assumption of known break point. When the break point is unknown we can proceed to estimate it using the infimum functional as described above. However in contrast with case where factors were not present, we have to constrain

[4] Note that Bai and Ng prefer to combine individual p-values instead of using these moments.

the (unknown) break point to be common to all units in the panel data set and to estimate both the subspace spanned by the common factors and the idiosyncratic disturbance terms for all possible break points. We then compute the $Z^e_{\hat{t}_{NT}}(\lambda) = N^{-1}\sum_{i=1}^{N} t_{\hat{\rho}_i}(\lambda)$ statistic for each break point using the idiosyncratic disturbance terms and estimate the break point as the argument that minimizes the sequence of standardized $Z^e_{\hat{t}_{NT}}(\lambda)$ statistics. Thus, the test statistic that is used to test the null hypothesis of non-cointegration for the idiosyncratic disturbance term is given by

$$Z^e_{\hat{t}_{NT}}\left(\hat{\lambda}\right) = \inf_{\lambda\in\Lambda}\left(\frac{N^{-1/2}Z^e_{\hat{t}_{NT}}(\lambda) - \Theta^e_2(\lambda)\sqrt{N}}{\sqrt{\Psi^e_2(\lambda)}}\right), \tag{21}$$

where the moments again depend on the specification of the deterministic term.

The estimated break date denoted $\hat{T}_b$ is given by

$$\hat{T}_b = \arg\min_{\lambda\in\Lambda}\left(\frac{N^{-1/2}Z^e_{\hat{t}_{NT}}(\lambda) - \Theta^e_2(\lambda)\sqrt{N}}{\sqrt{\Psi^e_2(\lambda)}}\right).$$

The limiting distribution of $Z^e_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ is given in the following Theorem.

**Theorem 3** *Let $\{y_{i,t}\}$ the stochastic process with DGP given by (11) to (15). Then, as $N, T \to \infty$ the $Z^e_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ test in (21) converges to*

$$Z^e_{\hat{t}_{NT}}\left(\hat{\lambda}\right) \Rightarrow \inf_{\lambda\in\Lambda}\kappa(\lambda),$$

*where $\kappa(\lambda)$ denotes a standard Normal distribution for a given $\lambda$.*

The proof follows from the Continuous Mapping Theorem (CMT). Theorem 3 establishes the limiting distribution of $Z^e_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ as the infimum of a sequence of correlated standard Normal variables. It has been shown that when the break point is known, the panel data statistics derived above converge to standard Normal distributions. When the test statistic is computed for all possible break points we obtain a correlated sequence of statistics, each of which is standard Normally distributed. The correlation comes from the fact that the statistics in the sequence are all computed from the same time series information. Critical values for (21) are obtained by simulation for different values of $T$ and for $N = 100$ – see panel A of Table 11.[5]

It is worth mentioning here that we need to consider finally the case of testing for unit roots in the common factors when the break is not known. As shown above, this matters only when there is a change in trend. Our procedure would then involve estimating the break date by using the statistic given in (21). This break date is then used to compute the ADF and the MQ tests for the common factors. Critical values are reported in panel B of Table 11.

# 5 Monte Carlo simulation

In this section we analyze by conducting simulation experiments the finite sample performance of the statistics that have been proposed in the paper. We begin by considering a DGP where the units are not cross-section dependent, so that our results in section 3 can be used. We then consider a DGP with cross-section dependence which uses our results in section 4.

[5] As is usual in the literature, we introduce trimming at the end points of the sample so that $\lambda$ varies between $[0.15, 0.85]$.

## 5.1 Cross-section independent

The empirical size of the tests is studied regressing two independent random walks, which have been generated as the cumulated sum of $iid\ N(0,1)$ processes. The sample size has been set equal to $T = \{50, 100, 250\}$ and the number of units at $N = \{20, 40\}$. The results reported in Table 12 are obtained from 5,000 replications, assuming that the break point is unknown and using the estimated response surfaces of the previous section. As can be seen, the empirical size of both the normalized bias and the pseudo $t$-ratio statistics is close to the nominal size irrespective of $T$ and $N$.

The empirical power of the statistics is assessed using the DGP given by:

$$\begin{aligned} y_{i,t} &= \mu_i + \theta_i DU_{i,t} + \beta_i t + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_{i,t} + e_{i,t} \\ e_{i,t} &= \rho_i e_{i,t-1} + \varepsilon_{i,t}, \end{aligned}$$

where $\varepsilon_{i,t} \sim iid\ N(0,1)\ \forall i,\ i = 1, \ldots, N$. The specification of the values of the parameters depends on the model under consideration. In general, the constant and, when required, the slope of the trend are set equal to $\mu_i = 1$ and $\beta_i = 0.3$, respectively. When there is a change in the level the magnitude is set equal to $\theta_i = 3$, while for the change in trend we consider $\gamma_i = 0.5$. The change in the cointegrating vector is given by $\delta_{i,t} = \delta_{i,1} = 1$ for $t \leq T_{bi}$ and $\delta_{i,t} = \delta_{i,1} = 3$ for $t > T_{bi}$, for a break point randomly located at $\lambda_i \sim U(0.15, 0.85)$, $\forall i,\ i = 1, \ldots, N$, where $U$ denotes the uniform distribution – the same results are obtained when break fraction is fixed either at $\lambda_i = 0.25$, $\lambda_i = 0.5$ or $\lambda_i = 0.75\ \forall i$. Simulations were performed for two autoregressive coefficients $\rho_i = \{0.5, 0.8\}$, although we only report results for $\rho_i = 0.8$ to save space. The computation of the statistics controls the autocorrelation in the disturbance term including up to $k_{\max} = 5$ lags using the $t$-sig criterion to select the order of the autoregressive correction. Results in Tables 13 and 14 show the empirical power of both statistics, respectively, for different combinations of DGP's and estimated models when $\rho_i = 0.8$. Thus, we can assess the empirical power of the statistics when DGP does not coincide with the model that is estimated. When the DGP and estimated model coincide both statistics show good power, which increases with $T$ and $N$ – see bold-typed columns in Tables 13 and 14. However, $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ outperforms $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ since for the former statistic the power equals one in all cases. In general, when the estimated model is misspecified and misspecification involves the cointegrating vector, the empirical power of $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ decreases. For instance, when DGP is given by Model 1 and we estimate Model 4, the power of the $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ statistic is reduced – note that the converse is also true. The same is found when either the DGP is given by Model 2 and we estimate Model 5, or when the DGP is given by Model 3 and we estimate Model 6. However, this feature is not found for the $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic, which does not lose any power when this sort of misspecification occurs. Finally, misspecification due to lack of accounting for time trend – i.e. DGP given by Models 2, 3, 5 and 6, and estimation of specifications given by Models 1 and 4 – reduces the power of both statistics as $T$ increases, although for the $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic the specifications that allow for a change in the cointegrating vector always show higher power.

In all, simulations lead us to conclude that $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic outperforms $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ in all situations that have been considered. Furthermore, overparameterisation of the estimated model does not cause loss of power for the $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic. These features indicate that $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ should be preferred to $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ in empirical applications.

## 5.2 Cross-section dependent

In order to deal with the situation with common factors, to mimic the impact of cross-sectional dependence, consider the DGP given by a bivariate system:

$$
\begin{aligned}
y_{i,t} &= f_i(t) + x'_{i,t}\delta_{i,t} + u_{i,t} \\
u_{i,t} &= F_t\pi_i + e_{i,t} \\
F_t &= \phi F_{t-1} + \sigma_F w_t \\
e_{i,t} &= \rho_i e_{i,t-1} + \varepsilon_{i,t} \\
\Delta x_{i,t} &= v_{i,t},
\end{aligned}
$$

where $(w_t, \varepsilon_{i,t}, v_{i,t})'$ follow a mutually *iid* standard multivariate Normal distribution for $\forall i, j\ i \neq j$ and $\forall t, s\ t \neq s$. In this paper we consider two different situations depending on the number of common factors, i.e. $r = \{1, 3\}$, and specify three values for the autoregressive parameters $\phi = \{0.8, 0.9, 1\}$ and $\rho_i = \{0.8, 0.9, 1\}\ \forall i$. Note that these values allow us to analyze both the empirical size and power of the statistics. The importance of the common factors is controlled through the specification of $\sigma_F^2 = \{0.5, 1, 10\}$. 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 $N = 40$ units and $T = \{50, 100, 250\}$ time observations.

The simulation results for size and power for the case with no breaks (with one or more factors) reported in Tables 15, 16 and 17 are close to those results in Bai and Ng (2004) – we only include the set of results for the only constant case, although the ones for the linear time trend are available upon request. From these results it may be seen that the empirical size of the ADF pooled idiosyncratic $t$-ratio statistic $\left(Z^e_{\hat{t}_{NT}}\right)$ and the ADF statistic of the common factor – when there is only one factor in the DGP – is close to the nominal size, which is set at the 5% level of significance. As expected the power of the tests increases as the autoregressive parameter moves away from unity. Moreover, the power of the $Z^e_{\hat{t}_{NT}}$ test is higher or equal to the power shown by the $ADF^d_{\hat{F}}$ test.

These results do not change when specifying three common factors – see Tables 16 and 17. Thus, the $Z^e_{\hat{t}_{NT}}$ test shows correct empirical size and good power. The $MQ^d_c(q)$ test also shows correct empirical size, while as expected the test has low power for large values of the autoregressive parameter – the bandwidth for the Bartlett spectral window is set as $J = 4ceil\left[\min\left[N, T\right]/100\right]^{1/4}$.

Turning now to the results for the case where there is one structural break, we start by assuming that the break point is known and located at $\lambda_i = \{0.25, 0.5, 0.75\}\ \forall i$. Table 18 reports results for the empirical size and power for the model that allows for one change in level with $\lambda_i = 0.5$ and one common factor. It should be noted that the results are not altered substantially either for other values of $\lambda_i$ or for a model that also includes a change in trend – these results are available upon request. On the one hand, the panel data unit root test on the idiosyncratic disturbance terms show good properties in terms of empirical size and power. On the other hand, the ADF statistic for the common factor shows the right size although, as expected, it has low power when the autoregressive parameter is close to unity and the sample size is small. Our results for three factors reported in Tables 19 and 20 confirm those for the one-factor case. Finally, Table 21 reports results for one common factor with one unknown break, which show that the statistics retain their good finite sample properties when the common break point has to be estimated.

# 6 Conclusions

This paper has shown that inference based on parametric Pedroni panel cointegration test statistics can be affected by the presence of structural breaks. Monte Carlo evidence indicates that in some situations the power of the tests drops as the magnitude of the structural break increases. Specifically, when the structural break affects either the slope of the time trend or the cointegrating vector the power approaches zero as $T$, $N$ and the magnitude of the break increases. In contrast, the power of the standard parametric Pedroni panel cointegration statistics is affected to a much lesser extent when the structural break only changes the level – we require a large magnitude of structural breaks located at the end of the time period to reduce the power of the statistics.

These features have motivated our proposal, and have led us to design statistical procedures to account for the presence of structural breaks when testing for cointegration. Six different specifications have been introduced depending on the effect of structural breaks on the long-run relationship. Finite sample and asymptotic moments have been computed that allow us to define panel cointegration statistics for the specifications considered.

The issue of cross-section dependence is addressed in the paper by assuming an approximate common factor structure. We derive the limiting distributions of statistics in two situations of interest, i.e. (i) for the case of no structural break, and (ii) when there are changes in level and trend. The performance of the approach is investigated through Monte Carlo simulations, from which we conclude that the statistics show good performance once the procedures have accounted for structural breaks.

## References

[1] Bai, J. and J. L. Carrion-i-Silvestre (2005): "Testing Panel Cointegration with Unobservable Dynamic Common Factors", *Mimeo*, Departament d'Econometria, Estadística i Economia Espanyola. Universitat de Barcelona.

[2] Bai, J. and Ng, S. (2002): "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, 70, 191-221.

[3] Bai, J. and S. Ng (2004): "A PANIC Attack on Unit Roots and Cointegration", *Econometrica*, 72, 1127-1177.

[4] Baltagi, B. H. (2005): *Econometric Analysis of Panel Data*, 3rd edition, John Wiley & Sons, New York.

[5] Baltagi, B. H., and C. Kao (2000): Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey. In Baltagi (Ed.) *Nonstationary Panels, Cointegration in Panels and Dynamic Panels*, Advances in Econometrics, 15. North-Holland, Elsevier, Amsterdam.

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

[7] Benerjee, A., Marcellino, M. and C. Osbat (2004): "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-Economic Data", *Econometrics Journal*, 7, 322-340.

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

[9] Bartley, W. A., J. Lee and M. C. Strazicich (2001): "Testing the Null of Cointegration in the Presence of a Structural Break", *Economics Letters*, 73, 315–323.

[10] Breitung, J. (2005): "A Parametric Approach to the Estimation of Cointegration Vectors in Panel Data", *Econometric Reviews*, 24, 151-173.

[11] Busetti, F. (2002): "Testing for (Common) Stochastic Trends in the Presence of Structural Breaks", *Journal of Forecasting*, 21, 81–105.

[12] Campos, J., N. Ericsson and D. F. Hendry (1996): "Cointegration Tests in the Presence of Structural Breaks", *Journal of Econometrics*, 70, 187-220.

[13] Carrion-i-Silvestre, J. Ll. and A. Sansó (2004): "Testing the Null of Cointegration with Structural Breaks", *Oxford Bulletin of Economics and Statistics*. Forthcoming.

[14] Chang, Y. (2005): "Residual Based Tests for Cointegration in Dependent Panels", *Mimeo*. Department of Economics, Rice University.

[15] Engle, R. F. and C. W. J. Granger (1987): "Cointegration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55, 251-276.

[16] Gregory, A. W. and B. E. Hansen (1996): "Residual-based Tests for Cointegration in Models with Regime Shifts", *Journal of Econometrics*, 70, 99-126.

[17] Hansen, B. E. (1992): "Tests for Parameter Instability in Regressions with I(1) Processes", *Journal of Business and Economic Statistics*, 10, 321-335.

[18] Hansen, H. and S. Johansen (1999): “Some Tests for Parameter Constancy in Cointegrated-VAR Models”, *Econometrics Journal*, 2, 306-333.

[19] Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press.

[20] Hao, K. (1996): “Testing for Structural Change in Cointegrated Regression Models: some Comparisons and Generalisations”, *Econometric Reviews*, 15, 401–429.

[21] Harris, D. and B. Inder (1994): “A Test of the Null Hypothesis of Cointegration”. In Colin P. Hargreaves (Ed.) *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press, Oxford.

[22] Jansen, W. J. (2000): “International Capital Mobility”, *Journal of International Money and Finance*, 19, 507-511.

[23] Kao, C. and M. H. Chiang (2000): “Testing for Structural Change of a Cointegrated Regression” *Mimeo*. Department of Economics, Syracuse University.

[24] McCoskey, S. and C. Kao (1998): “A Residual-Based Test of the Null of Cointegration in Panel Data”, *Econometric Reviews*, 17, 57-84.

[25] Ng, S. and P. Perron (1995): “Unit Root Test in ARMA models with Data-Dependent Methods for the Selection of the Truncation Lag”, *Journal of the American Statistical Association*, 90, 268-281.

[26] Pedroni, P. (1999): “Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors”, *Oxford Bulletin of Economics and Statistics*, Special Issue, 653-670.

[27] Pedroni, P. (2004): “Panel Cointegration. Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis”, *Econometric Theory*, 20, 3, 597-625.

[28] Perron, P. (1989): “The Great Crash, the Oil Price Shock and the Unit Root Hypothesis”, *Econometrica*, 57, 1361-1401.

[29] Phillips, P. C. B. (1987): “Time Series Regression with a Unit Root”, *Econometrica*, 55, 277-301.

[30] Phillips, P. C. B. (2001): “Bootstrapping Spurious Regressions”. *Cowles Foundation Discussion Paper* num. 1330. Yale.

[31] Phillips, P.C.B. and S. N. Durlauf (1986): “Multiple Time Series Regression with Integrated Processes”, *Review of Economic Studies*, LIII, 473-495.

[32] Phillips, P. C. B. and P. Perron (1988): “Testing for a Unit Root in Time Series Regression”, *Biometrika*, 335-346.

[33] Quintos, C. and P. C. B. Phillips (1993): “Parameter Constancy in Cointegrating Regressions”, In *New Developments in Time Series Econometrics*, J. M. Dufour and B. Raj (Eds.). Studies in Empirical Economics. Physica-Verlag. Springer Verlag.

[34] Shin, Y. (1994): “A Residual-based Test of the Null of Cointegration against the Alternative of no Cointegration”, *Econometric Theory*, 10, 91–115.

[35] Stock, J. H. and M. W. Watson (1993): “A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems”, *Econometrica*, 64, 783-820.

[36] Westerlund, J. (2004): “Testing for Panel Cointegration with Multiple Structural Breaks”, *Oxford Bulletin of Economics and Statistics*. Forthcoming.

[37] Zivot, E. and D. W. K. Andrews (1992): "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit-Root Hypothesis", *Journal of Business & Economic Statistics*, 10, 251–270."

# A Mathematical Appendix

For the sake of simplicity let us first assume that the stochastic regressors are strictly exogenous. Once the main result is derived, we show how these derivations can be extended to account for non-strictly exogenous regressors.

## A.1 Proof of statement *(1.a)* of Theorem 2

Let us assume the model given by (11) and (12). Furthermore, consider the case where there are no structural breaks affecting the model and there are no deterministic elements in the model – note that the presence of a constant term does not change the results since it disappears when taking first differences. Alternatively, the model can be expressed as:

$$y_{i,t} = x_{i,t}'\delta_i + F_t\pi_i + e_{i,t}.$$

As can be seen, the model assumes that residuals from the static regression follow a factor structure as defined in Bai and Ng (2004). Note that if we introduce (16) in (17) we obtain

$$\begin{aligned} \tilde{z}_{i,t} &= z_{i,t} + f_t\pi_i - \tilde{f}_t\tilde{\pi}_i \\ &= z_{i,t} - v_t H^{-1}\pi_i - \tilde{f}_t d_i, \end{aligned} \tag{22}$$

where $v_t = \tilde{f}_t - f_t H$ and $d_i = \tilde{\pi}_i - H^{-1\prime}\pi_i$, where $H$ is an $(r \times r)$ matrix defined as follows $H = V_{NT}^{-1}\left(\hat{f}'f/T\right)(\Pi'\Pi/N)$ with $V_{NT}$ the $(r \times r)$ diagonal matrix of the first $r$ largest eigenvalues of $(NT)^{-1}\, y^* y^{*\prime}$ in decreasing order. The computation of the partial sum processes of (22) gives:

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\sum_{j=2}^{t} z_{i,j} - T^{-1/2}\sum_{j=2}^{t} v_j H^{-1}\pi_i - T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j d_i. \tag{23}$$

Let us analyse each element of (23) separately. The left-hand side of (23) is equal to

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} &= T^{-1/2}\sum_{j=2}^{t}\left[M_i\Delta\tilde{e}_i\right]_j \\ &= T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} - T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{e}_i\right]_j, \end{aligned} \tag{24}$$

where $[\cdot]_j$ denotes the $j$-th element of the vector between parentheses, and $P_i = I_{T-1} - M_i$. The first element on the right of (24) is equal to

$$T^{-1/2}\sum_{j=2}^{t}\Delta\tilde{e}_{i,j} = T^{-1/2}\tilde{e}_{i,t} - T^{-1/2}\tilde{e}_{i,1} = T^{-1/2}\tilde{e}_{i,t} + O_p(1),$$

so that by the invariance principle

$$T^{-1/2}\sum_{j=2}^{[sT]}\Delta\tilde{e}_{i,j} \Rightarrow \sigma_i W_i(s).$$

The second element on the right hand of (24) is

$$T^{-1/2}\sum_{j=2}^{t}\left[P_i\Delta\tilde{e}_i\right]_j = T^{-1/2}\left(x_{i,t}-x_{i,1}\right)'\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{e}_i.$$

Note that $\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{e}_i = \left(T^{-1}\Delta x_i'\Delta x_i\right)^{-1}\left(T^{-1}\Delta x_i'\Delta\tilde{e}_i\right) = o_p\left(1\right)$, since $\left(T^{-1}\Delta x_i'\Delta x_i\right) \to^p Q_{\Delta x_i \Delta x_i}$, the variance and covariance matrix of $\Delta x_i'\Delta x_i$, and $T^{-1}\Delta x_i'\Delta\tilde{e}_i \to^p 0$ since these elements are orthogonal by definition. On the other hand, $T^{-1/2}x_{i,t} \Rightarrow \Omega_{22,i}^{1/2}W_{m-1}\left(s\right)$ and $T^{-1/2}x_{i,1} \to^p 0$ by assumption. These derivations lead us to

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\tilde{e}_{i,t} + o_p\left(1\right),$$

since $T^{-1/2}x_{i,t}\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{e}_i = o_p\left(1\right)$. The same result can be achieved for $T^{-1/2}\sum_{j=2}^{t} z_{i,j}$, i.e.

$$T^{-1/2}\sum_{j=2}^{t} z_{i,j} = T^{-1/2}e_{i,t} + o_p\left(1\right).$$

This indicates that the presence of stochastic regressors does not have any effect on the partial sum processes. Regarding the term involving $\{v_t\}$ we see from Eq. (A.3) in Bai and Ng (2004) that

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

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

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

From all these results it follows that

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

that is, the limiting distribution is the same derived in Bai and Ng (2004) for the constant case –see Bai and Ng (2004) for the proof. The same result is found for the ADF test provided that the order of the autoregressive correction is selected such that $k \to \infty$ and $k^3/\min\left[N, T\right] \to 0$. This implies that the presence of stochastic regressors does not affect the limiting distribution of the statistic.

## A.2 Proof of statement *(1.b)* of Theorem 2

The generalization that includes a time trend can be carried out as well. In this case the model (11) is replaced by

$$y_{i,t} = \mu_i + \beta_i t + x_{i,t}'\beta_i + u_{i,t}.$$

Note that as before we are not dealing with the structural break case since we are defining the benchmark limiting distributions. Contrary to the previous specification, taking first differences does not remove the deterministic elements, since now the trend becomes a constant. This is a relevant feature since the limiting distribution of the ADF-type statistic varies. However, the asymptotic distribution of the statistic is the same as the one derived in Bai and Ng (2004) for the trend case. The proof follows similar steps above. Now the first difference of regressors defines the following idempotent matrix

$$M_i = I_{T-1} - \Delta x_i^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime},$$

where the $\Delta x_i^d$ matrix is defined by the row vectors $\left(1, \Delta x_{i,t}'\right)'$. Note that as before the first element of (24) converges to

$$T^{-1/2} \sum_{j=2}^{[sT]} \Delta \tilde{e}_{i,j} \Rightarrow \sigma_i W_i(s).$$

The limiting expression of the second element in (24) has to be derived in several steps. First, note that $T^{-1} \Delta x_i^{d\prime} \Delta x_i^d$ converges to variance and covariance matrix of $\Delta x_i^d$, so that all these elements are $O_p(1)$. The first element of the vector $T^{-1} \Delta x_i^{d\prime} \Delta \tilde{e}_i$ is given by $T^{-1/2} \left(T^{-1/2} \sum_{t=1}^T \Delta \tilde{e}_{i,t}\right) = T^{-1/2} \left(T^{-1/2} \left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right)$, where $T^{-1/2} \left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right) \Rightarrow \sigma_i W_i(1)$ since $T^{-1/2} \tilde{e}_{i,1} \to^p 0$. Note that the extra rescaling term $T^{-1/2}$ would be used below. The rest of the elements in $T^{-1} \Delta x_i^{d\prime} \Delta \tilde{e}_i$ involve cross-products among the first difference of the stochastic regressors and $\Delta \tilde{e}_i$ that converges to zero since we have assumed independency. Therefore,

$$\left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime} \Delta \tilde{e}_i = \left[ \begin{array}{c} E\, T^{-1/2} \left(T^{-1/2} \left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right) + o_p(1) \\ \left(-D^{-1} C E\right) T^{-1/2} \left(T^{-1/2} \left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right) + o_p(1) \end{array} \right]$$

where $E = \left(A - B D^{-1} C\right)^{-1}$ and $A = 1, B = T^{-1} \iota' \Delta x_i, C = B'$ and $D = T^{-1} \Delta x_i' \Delta x_i$ denote the elements of the partitioned matrix $T^{-1} \Delta x_i^{d\prime} \Delta x_i^d$, with $\iota = (1, \ldots, 1)'$. The partial sum process of $\Delta x_{i,t}^d$ is

$$T^{-1/2} \sum_{j=2}^{t} \Delta x_{i,j}^d = \left[T^{-1/2} t \quad T^{-1/2} \left(x_{i,t} - x_{i,1}\right)'\right],$$

so that

$$T^{-1/2} \sum_{j=2}^{t} \Delta x_{i,j}^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime} \Delta \tilde{e}_i = \frac{t}{T} E \left(T^{-1/2} \left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right) + o_p(1),$$

since $T^{-1} \left(x_{i,t} - x_{i,1}\right)' = o_p(1)$. Moreover, the matrix $E$ can be expressed as

$$\begin{aligned} \left(A - B D^{-1} C\right)^{-1} &= A^{-1} + A^{-1} B \left(D - C A^{-1} B\right)^{-1} C A^{-1} \\ &= 1 + B \left(D - B' B\right)^{-1} B'. \end{aligned}$$

Note that $B = T^{-1} \iota' \Delta x_i \to^p 0$ so that $\left(A - B D^{-1} C\right)^{-1} = 1 + o_p(1)$. Therefore,

$$\begin{aligned} T^{-1/2} \sum_{j=2}^{t} \Delta x_{i,j}^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime} \Delta \tilde{e}_i &= \frac{t}{T} \left(T^{-1/2} \left(\tilde{e}_{i,T} - \tilde{e}_{i,1}\right)\right) + o_p(1) \\ &\Rightarrow r\, \sigma_i W_i(1). \end{aligned}$$

From Bai and Ng (2004), the terms $T^{-1/2} \left\| \sum_{j=2}^t v_j \right\| = O_p\left(C_{NT}^{-1}\right)$, $\|d_i\| = O_p\left(C_{NT}^{-1}\right)$ and $T^{-1/2} \left\| \sum_{j=2}^t \tilde{f}_j \right\| = O_p(1)$. These derivations lead us to

$$\begin{aligned} T^{-1/2} \sum_{j=2}^{[sT]} \tilde{z}_{i,j} &= T^{-1/2} \tilde{e}_{i,t} - \frac{s}{T} T^{-1/2} \tilde{e}_{i,T} + O_p\left(C_{NT}^{-1}\right) \\ &\Rightarrow \sigma_i \left(W_i(s) - s\, W_i(1)\right) \equiv \sigma_i V_i(s). \end{aligned}$$

The DF statistic is

$$DF_{\tilde{e}}^{\tau}(i) = \frac{T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t}}{\left(\tilde{\sigma}_i^2 T^{-2}\sum_{t=2}^{T}\tilde{e}_{i,t-1}^2\right)^{1/2}}.$$

Note that the following identity holds

$$T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t} = \frac{\tilde{e}_{i,T}^2}{2T} - \frac{\tilde{e}_{i,1}^2}{2T} - \frac{1}{2T}\sum_{t=2}^{T}(\Delta\tilde{e}_{i,t})^2,$$

which shows that $T^{-1}\tilde{e}_{i,T}^2 \Rightarrow \sigma_i^2 V_i(1)^2 = 0$, $T^{-1}\tilde{e}_{i,1}^2 = 0$ and $T^{-1}\sum_{t=2}^{T}(\Delta\tilde{e}_{i,t})^2 \to^p \sigma_i^2$, from which it follows that $T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t} \to^p -\sigma_i^2/2$ and $T^{-2}\sum_{t=2}^{[sT]}\tilde{e}_{i,t-1}^2 \Rightarrow \sigma_i^2\int_0^1 V_i(s)^2 ds$ – see Bai and Ng (2004), Lemma G.4. Using these elements it is straightforward to see that

$$DF_{\tilde{e}}^{\tau}(i) \Rightarrow -\frac{1}{2}\left(\int_0^1 V_i(s)^2 ds\right)^{-1/2},$$

where $V_i(s) = W_i(s) - s\, W_i(1)$, *i.e.* the limiting distribution is the same derived in Bai and Ng (2004) for the trend case. Although the proof is more involved, the same result is achieved for the ADF test. As before, this implies that the presence of stochastic regressors does not affect the limiting distribution of the statistic. Note that this result is also achieved when there are level shifts in the model, since the impulse dummies do not affect the limiting distribution of the $ADF_{\tilde{e}}^{\tau}(i)$ statistic.

## A.3 Proof of statement *(1.c)* of Theorem 2

The model is given by the following deterministic specification

$$f_i(t) = \mu_i + \beta_i t + \theta_i DU_{i,t} + \gamma_i DT_{i,t}^*,$$

which implies that $\Delta f_i(t) = \beta_i + \theta_i D\left(T_b^i\right)_t + \gamma_i DU_{i,t}$ and $\Delta x_{i,t}^d = \left(1, D\left(T_b^i\right)_t, DU_{i,t}, \Delta x_{i,t}'\right)$. In order to simplify the steps of the proof, we deal with the equivalent specification that does not include the impulse dummy, *i.e.* $\Delta x_{i,t}^d = \left(1, DU_{i,t}, \Delta x_{i,t}'\right)$. This simplifies derivations, although it does not imply loss of generality. Moreover, note that the subspace spanned by $\left(1, DU_{i,t}, \Delta x_{i,t}'\right)$ is equivalent to the one spanned by $\left(DU_{i,t}^1, DU_{i,t}^2, \Delta x_{i,t}'\right)$ where $DU_{i,t}^1 = 1$ for $t \le T_b$ and 0 otherwise, and $DU_{i,t}^2 = 1$ for $t > T_b$ and 0 otherwise. This redefinition makes $DU_{i,t}^1$ and $DU_{i,t}^2$ to be orthogonal. Note that as before the first element of (24) converges to

$$T^{-1/2}\sum_{j=2}^{[sT]}\Delta\tilde{e}_{i,j} \Rightarrow \sigma_i W_i(s).$$

The limiting expression of the second element in (24) has to be derived in several steps. First, note that $T^{-1}\Delta x_i^{d\prime}\Delta x_i^d$ converges to variance and covariance matrix of $\Delta x_i^d$, so that all these elements are $O_p(1)$. The first element of the vector $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i$ is given by $T^{-1/2}\left(T^{-1/2}\sum_{t=1}^{T_b}\Delta\tilde{e}_{i,t}\right) = T^{-1/2}\left(T^{-1/2}(\tilde{e}_{i,T_b} - \tilde{e}_{i,1})\right)$, where $T^{-1/2}(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}) \Rightarrow \sigma_i W_i(\lambda)$ since $T^{-1/2}\tilde{e}_{i,1} \to^p 0$. The second element is $T^{-1/2}\left(T^{-1/2}\sum_{t=T_b+1}^{T}\Delta\tilde{e}_{i,t}\right) = T^{-1/2}\left(T^{-1/2}(\tilde{e}_{i,T} - \tilde{e}_{i,T_b})\right)$, where $T^{-1/2}(\tilde{e}_{i,T} - \tilde{e}_{i,T_b}) \Rightarrow \sigma_i W_i(1) - \sigma_i W_i(\lambda)$. Note that as before the extra rescaling term $T^{-1/2}$ would be used below. Finally, the third set of elements in the product is $T^{-1}\Delta x_i'\Delta\tilde{e}_i$ that converges to zero since we have assumed independency. Therefore,

$$\left(\Delta x_i^{d\prime}\Delta x_i^d\right)^{-1}\Delta x_i^{d\prime}\Delta\tilde{e}_i = \begin{bmatrix} E\, T^{-1/2}\left(T^{-1/2}(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}), T^{-1/2}(\tilde{e}_{i,T} - \tilde{e}_{i,T_b})\right)' + o_p(1) \\ \left(-D^{-1}CE\right)\, T^{-1/2}\left(T^{-1/2}(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}), T^{-1/2}(\tilde{e}_{i,T} - \tilde{e}_{i,T_b})\right)' + o_p(1) \end{bmatrix}$$

where $E = \left(A - BD^{-1}C\right)^{-1}$ and $A = diag\left(\lambda, 1-\lambda\right), B = T^{-1}\left[DU_i^1, DU_i^2\right]' \Delta x_i, C = B'$ and $D = T^{-1}\Delta x_i' \Delta x_i$ denote the elements of the partitioned matrix $T^{-1}\Delta x_i^{d\prime} \Delta x_i^d$. Moreover, following the steps given above $\left(A - BD^{-1}C\right)^{-1} = A^{-1} + o_p(1)$, since $B \to^p 0$. The partial sum process of $\Delta x_{i,t}^d$ for $t \leq T_b$ is

$$T^{-1/2}\sum_{j=2}^{t} \Delta x_{i,j}^d = \left[T^{-1/2}t \quad 0 \quad T^{-1/2}\left(x_{i,t} - x_{i,1}\right)'\right],$$

while for $t > T_b$ is

$$T^{-1/2}\sum_{j=2}^{[sT]} \Delta x_{i,j}^d = \left[T^{-1/2}T_b \quad T^{-1/2}\left(s - T_b\right) \quad T^{-1/2}\left(x_{i,t} - x_{i,1}\right)'\right],$$

so that for $t \leq T_b$

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{[sT]} \Delta x_{i,j}^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime} \Delta \tilde{e}_i &= \frac{s}{T}\frac{1}{\lambda}\left(T^{-1/2}\left(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}\right)\right) + o_p(1) \\ &\Rightarrow \frac{s}{\lambda}\sigma_i W_i(\lambda), \end{aligned}$$

since $T^{-1}\left(x_{i,t} - x_{i,1}\right)' = o_p(1)$. Therefore, for $t \leq T_b$

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{[sT]} \tilde{z}_{i,j} &= T^{-1/2}\tilde{e}_{i,t} - \frac{s}{T}T^{-1/2}\tilde{e}_{i,T} + O_p\left(C_{NT}^{-1}\right) \\ &\Rightarrow \sigma_i\left(W_i(s) - \frac{s}{\lambda}W_i(\lambda)\right), \end{aligned}$$

since from Bai and Ng (2004), the terms $T^{-1/2}\left\|\sum_{j=2}^{t} v_j\right\| = O_p\left(C_{NT}^{-1}\right)$, $\|d_i\| = O_p\left(C_{NT}^{-1}\right)$ and $T^{-1/2}\left\|\sum_{j=2}^{t} \tilde{f}_j\right\| = O_p(1)$. Note that we can define $b_1 = s/\lambda$ so that $0 < b_1 < 1$, which in turn implies that

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{[sT]} \tilde{z}_{i,j} &\Rightarrow \sigma_i\sqrt{\lambda}W_i(b_1) - \sigma_i b_1\sqrt{\lambda}W_i(1) \\ &= \sigma_i\sqrt{\lambda}\left(W_i(b_1) - b_1 W_i(1)\right) \equiv \sigma_i\sqrt{\lambda}V_i(b_1), \end{aligned}$$

given the properties of Brownian motions. On the other hand, for $t > T_b$

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{[sT]} \Delta x_{i,j}^d \left(\Delta x_i^{d\prime} \Delta x_i^d\right)^{-1} \Delta x_i^{d\prime} \Delta \tilde{e}_i &= \frac{T_b}{T}\frac{1}{\lambda}\left(T^{-1/2}\left(\tilde{e}_{i,T_b} - \tilde{e}_{i,1}\right)\right) \\ &\quad + \frac{s - T_b}{T}\frac{1}{1-\lambda}\left(T^{-1/2}\left(\tilde{e}_{i,T} - \tilde{e}_{i,T_b}\right)\right) + o_p(1) \\ &\Rightarrow \sigma_i\left(W_i(\lambda) + \frac{s-\lambda}{1-\lambda}\left(W_i(1) - W_i(\lambda)\right)\right), \end{aligned}$$

so that

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{[sT]} \tilde{z}_{i,j} &= T^{-1/2}\tilde{e}_{i,t} - \frac{s}{T}T^{-1/2}\tilde{e}_{i,T} + O_p\left(C_{NT}^{-1}\right) \\ &\Rightarrow \sigma_i\left(W_i(s) - W_i(\lambda) - \frac{s-\lambda}{1-\lambda}\left(W_i(1) - W_i(\lambda)\right)\right). \end{aligned}$$

As before, we can define $b_2 = (s-\lambda)/(1-\lambda)$ so that $0 < b_2 < 1$, which in turn implies that

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} \Rightarrow \sigma_i\sqrt{1-\lambda}\left(W_i(b_2) - b_2 W_i(1)\right) \equiv \sigma_i\sqrt{1-\lambda}V_i(b_2).$$

Using similar developments as in the previous proof, the numerator of the DF statistic converges to $T^{-1}\sum_{t=2}^{T}\tilde{e}_{i,t-1}\Delta\tilde{e}_{i,t} \to^p -\sigma^2/2$, while the denominator is

$$\begin{aligned} T^{-2}\sum_{t=2}^{T}\tilde{e}_{i,t-1}^2 &= T^{-2}\sum_{t=2}^{T_b+1}\tilde{e}_{i,t-1}^2 + T^{-2}\sum_{t=T_b+2}^{T}\tilde{e}_{i,t-1}^2 \\ &\Rightarrow \sigma_i^2\left(\lambda^2\int_0^1 V_i(b_1)^2\,db_1 + (1-\lambda)^2\int_0^1 V_i(b_2)^2\,db_2\right), \end{aligned}$$

with $V(b_1)$ and $V(b_2)$ two independent Brownian bridges. Therefore, the limiting distribution of the DF statistic is

$$DF_{\tilde{e}}^{\tau}(i) \Rightarrow -\frac{1}{2}\left(\lambda^2\int_0^1 V_i(b_1)^2\,db_1 + (1-\lambda)^2\int_0^1 V_i(b_2)^2\,db_2\right)^{-1/2}.$$

It can be shown that this limiting distribution is symmetric around $\lambda = 0.5$ since in this case we can interchange $\lambda^2$ and $(1-\lambda)^2$ and obtain the same distribution. As before, the same limiting distribution is found for the ADF statistic.

## A.4 Proof of statement *(2)* of Theorem 2

Let us now deal with the unit root hypothesis testing when there is $r = 1$ common factor and no change in trend. The model in first differences defines an idempotent matrix $M_i$ that is unit-dependent. At first sight this goes against the definition of a common factor since we assume that this element is common to all units and, hence, cannot depend on $i$. Nevertheless, it is shown below that the elements that depend on $i$ vanish asymptotically. Thus, note that

$$\begin{aligned} \sum_{j=2}^{t}\tilde{f}_j &= \sum_{j=2}^{t}\left[M_i\Delta\tilde{F}\right]_j \\ &= \tilde{F}_t - (x_{i,t} - x_{i,1})'\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{F}, \end{aligned} \tag{25}$$

since we define $\tilde{F}_1 = 0$, where $[\cdot]_j$ refers to the $j$-th element of the matrix between parentheses. Note that the first element of (25) is

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

since $\Delta\tilde{F}_t = H\ \Delta F_t + v_t$ and $V_t = \sum_{j=2}^{t} v_j$.

The detrended estimated factor will remove $F_1$:

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

and it can be shown that

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

since $T^{-1/2}V_t^d = O_p\left(C_{NT}^{-1}\right)$ –see Bai and Ng (2004), Lemma B.2. The second term in (25) is $T^{-1/2}(x_{i,t} - x_{i,1})'$ $\left(\Delta x_i'\Delta x_i\right)^{-1}\Delta x_i'\Delta\tilde{F} = o_p(1)$, since $T^{-1}\Delta x_i'\Delta x_i$ converges to the matrix of covariance of $\Delta x_i$ and

$T^{-1}\Delta x_i' \Delta \tilde{F} = o_p(1)$ by assumption. Since

$$\begin{aligned} T^{-1/2}\tilde{F}_t^d &\Rightarrow H\ W_w^d(s) \\ T^{-2}\sum_{t=2}^{T}\tilde{F}_{t-1}^d\tilde{F}_{t-1}^{d\prime} &\Rightarrow H^2\ \sigma_w^2\int_0^1 W_w^d(s)^2\,ds \\ T^{-1}\sum_{t=2}^{T}\tilde{F}_{t-1}^d\Delta\tilde{F}_t &\Rightarrow H^2\ \sigma_w^2\int_0^1 W_w^d(s)\,dW(s)\,, \end{aligned}$$

the DF statistic converges to

$$\begin{aligned} DF_{\tilde{F}}^d &= \frac{T^{-1}\sum_{t=2}^{T}\tilde{F}_{t-1}^d\Delta\tilde{F}_t}{\left(\tilde{\sigma}_u^2 T^{-2}\sum_{t=2}^{T}\left(\tilde{F}_{t-1}^d\right)^2\right)^{1/2}} \qquad (26) \\ &\Rightarrow \frac{\int_0^1 W_w^d(s)\,dW(s)}{\left(\int_0^1 W_w^d(s)^2\,ds\right)^{1/2}}, \end{aligned}$$

where $W_w^d(s)$ denotes the detrended Brownian motion and $\tilde{\sigma}_w^2 \xrightarrow{p} H^2\ \sigma_w^2$. The ADF statistic has the same limiting distribution provided that the order of the autoregressive correction is selected such that $k\to\infty$ and $k^3/\min[N,T]\to 0$.

Following similar steps, it can be shown that when there is a time trend in the model

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j &= H\ T^{-1/2}\left(F_t - F_1 - (F_T - F_1)\frac{t}{T}\right) + O_p\left(C_{NT}^{-1}\right) \\ &= H\ T^{-1/2}F_t^d + O_p\left(C_{NT}^{-1}\right), \end{aligned}$$

where $F_t^d$ denotes the detrended common factor, which is obtained as the residual of a regression on a constant and a time trend. Therefore, DF statistic given by (26) converges to

$$DF_{\tilde{F}}^d \Rightarrow \frac{\int_0^1 W_w^d(s)\,dW(s)}{\left(\int_0^1 W_w^d(s)^2\,ds\right)^{1/2}},$$

where, as before, $W_w^d(s)$ denotes the detrended Brownian motion and $\tilde{\sigma}_w^2 \to^p H^2\ \sigma_w^2$. The ADF statistic has the same limiting distribution provided that the order of the autoregressive correction is selected such that $k\to\infty$ and $k^3/\min[N,T]\to 0$.

Finally, when there is one structural break that affects the time trend, we can see that

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{t}\tilde{f}_j &= H\ T^{-1/2}\left(F_t - F_1 - (F_T - F_1)\frac{t}{T} - (F_T - F_{T_b})\frac{t-T_b}{T}1(t>T_b)\right) + O_p\left(C_{NT}^{-1}\right) \\ &= H\ T^{-1/2}F_t^d + O_p\left(C_{NT}^{-1}\right), \end{aligned}$$

where $1(t>T_b)$ is an indicator function. Now $F_t^d$ is obtained as the residual of a regression on a constant, a time trend and the dummy variable $DT_t^* = (t-T_b)\ 1(t>T_b)$. Using these elements it is straightforward to see that the DF statistic given by (26) converges to

$$DF_{\tilde{F}}^d(\lambda) \Rightarrow \frac{\int_0^1 W_w^d(s,\lambda)\,dW(s,\lambda)}{\left(\int_0^1 W_w^d(s,\lambda)^2\,dr\right)^{1/2}},$$

where, as before, $W_w^d(s,\lambda)$ denotes the detrended Brownian motion, $\lambda$ is the break fraction parameter and $\tilde{\sigma}_w^2 \xrightarrow{p} H^2\ \sigma_w^2$. Note that this limiting distribution has been considered in Perron (1989) for the specification denoted as Model C. Finally, note that these derivations are valid when stochastic regressors are non-strictly exogenous provided the regression equation includes leads and lags of their first difference.

## A.5 Proof of statement *(3)* of Theorem 2

The limiting distributions of the test statistics that are used when there is more than one common factor ($r > 1$) but no break are the same as the ones derived in Bai and Ng (2004). These steps may be followed routinely to derive the distributions given in *(3)* for the case where the break is unknown. As stated in Bai and Ng (2004), pp. 1167, Remark 1, the validity of the $MQ$ tests using detrended estimated factors relies on the closeness of the true detrended factors, which has been shown in previous proofs. Thus, the limiting distribution of the $MQ$ tests is the same as in Bai and Ng (2004), but using properly detrended Brownian motions.

## A.6 Non strictly-exogenous regressors

Following developments in Bai and Carrion-i-Silvestre (2005) we can show that the same results are obtained when stochastic regressors are non-strictly exogenous. Here we only consider the specification without any deterministic component, although derivations extend to all models proposed in the paper. Thus, the model given by (11) and (12) with non-strictly exogenous regressors can be expressed as

$$y_{i,t} = x_{i,t}'\delta_i + \Delta x_{i,t}' A_i(L) + F_t\lambda_i + \xi_{i,t},$$

where $A_i(L)$ denotes the $(k \times 1)$-vector of lead and lag polynomials. Previous derivations concerning idiosyncratic disturbance term still hold but replacing $\Delta\tilde{e}_{i,t}$ with $\Delta\tilde{\xi}_{i,t}$. Now we define $\Delta x_{i,t}^d = \left(\Delta x_{i,t}', \Delta^2 x_{i,t}'\right)'$. Note that $T^{-1/2}(\Delta x_{i,t} - \Delta x_{i,1}) = T^{-1/2}O_p(1) \to^p 0$, $T^{-1}\Delta x_i^{d\prime}\Delta x_i^d \to^p Q_{\Delta x_i^d \Delta x_i^d}$, the covariance matrix of $\Delta x_i^{d\prime}\Delta x_i^d$, and $T^{-1}\Delta x_i^{d\prime}\Delta\tilde{\xi}_i \to^p 0$, so that we can see that $T^{-1/2}\sum_{j=2}^{[sT]}\left[P_i\Delta\tilde{\xi}_i\right]_j \to^p 0$. Then,

$$T^{-1/2}\sum_{j=2}^{t}\tilde{z}_{i,j} = T^{-1/2}\tilde{\xi}_{i,t} + o_p(1),$$

and

$$T^{-1/2}\sum_{j=2}^{t} z_{i,j} = T^{-1/2}\xi_{i,t} + o_p(1),$$

which indicates that the presence of (non-strictly) stochastic regressors does not have any effect on the partial sum processes once endogeneity has been taken into account and, hence, the rest of the proof follows the one above for strictly exogenous regressors.

Table 1: Empirical power of Pedroni pseudo t-ratio cointegration statistic. The structural change affects the deterministic component

| | | $\rho_i = 0$ | | | | $\rho_i = 0.5$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $T = 100$ | | $T = 250$ | | $T = 100$ | | $T = 250$ | |
| $\lambda_i$ | $(\theta_i, \gamma_i)$ | $N = 20$ | $N = 40$ | $N = 20$ | $N = 40$ | $N = 20;$ | $N = 40$ | $N = 20;$ | $N = 40$ |
| 0.25 | $(0, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(1, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(3, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(5, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(10, 0)$ | 1 | 1 | 1 | 1 | 0.49 | 0.88 | 1 | 1 |
| 0.5 | $(1, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(3, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(5, 0)$ | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 |
| | $(10, 0)$ | 0.94 | 1 | 1 | 1 | 0.08 | 0.09 | 0.90 | 0.99 |
| 0.75 | $(1, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(3, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(5, 0)$ | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 |
| | $(10, 0)$ | 0.83 | 0.98 | 1 | 1 | 0.01 | 0.00 | 0.72 | 0.94 |
| 0.25 | $(0, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(3, 0.5)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(3, 0.7)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(3, 1)$ | 1 | 1 | 0.99 | 1 | 1 | 1 | 0.99 | 1 |
| 0.5 | $(3, 0.5)$ | 0.65 | 0.89 | 0.01 | 0 | 0.02 | 0 | 0 | 0 |
| | $(3, 0.7)$ | 0.02 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $(3, 1)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.75 | $(3, 0.5)$ | 0.34 | 0.54 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $(3, 0.7)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $(3, 1)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

DGP: $y_t = \mu_i + \theta_i DU_{i,t} + \beta_i t + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_i + z_{i,t}$; $\Delta x_{i,t} = \varepsilon_{i,t}$ and $z_{i,t} = \rho_i z_{i,t-1} + v_{i,t}$ with $\zeta_{i,t} = (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2)$, $\mu_i = 1$, $\beta_i = 0.3$ and $\delta_i = 1$. The nominal size is set at the 5% level and 1,000 replications are carried out.

Table 2: Empirical power of Pedroni pseudo t-ratio cointegration statistic. The structural change affects both the deterministic component and the cointegrating vector

| | | | $\rho_i = 0$ | | | | $\rho_i = 0.5$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $N$ $(T = 100)$ | | $N$ $(T = 250)$ | | $N$ $(T = 100)$ | | $N$ $(T = 250)$ | |
| $\lambda_i$ | $(\theta_i, \gamma_i)$ | $(\delta_{i,1}, \delta_{i,2})$ | 20 | 40 | 20 | 40 | 20 | 40 | 20 | 40 |
| 0.25 | $(0, 0)$ | $(1, 0)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(0, 0)$ | $(1, 2)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(0, 0)$ | $(1, 3)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(0, 0)$ | $(1, 4)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(0, 0)$ | $(1, 5)$ | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 |
| | $(0, 0)$ | $(1, 10)$ | 0.99 | 1 | 1 | 1 | 0.97 | 1 | 1 | 1 |
| 0.5 | $(0, 0)$ | $(1, 2)$ | 1 | 1 | 1 | 1 | 0.98 | 1 | 1 | 1 |
| | $(0, 0)$ | $(1, 3)$ | 0.98 | 1 | 0.99 | 1 | 0.50 | 0.77 | 0.76 | 0.94 |
| | $(0, 0)$ | $(1, 4)$ | 0.71 | 0.92 | 0.86 | 0.99 | 0.27 | 0.42 | 0.42 | 0.67 |
| | $(0, 0)$ | $(1, 5)$ | 0.45 | 0.68 | 0.62 | 0.853 | 0.17 | 0.31 | 0.32 | 0.50 |
| | $(0, 0)$ | $(1, 10)$ | 0.17 | 0.30 | 0.26 | 0.406 | 0.13 | 0.18 | 0.19 | 0.31 |
| 0.75 | $(0, 0)$ | $(1, 2)$ | 1 | 1 | 1 | 1 | 0.83 | 0.97 | 0.96 | 1 |
| | $(0, 0)$ | $(1, 3)$ | 0.76 | 0.92 | 0.86 | 0.98 | 0.11 | 0.11 | 0.20 | 0.28 |
| | $(0, 0)$ | $(1, 4)$ | 0.26 | 0.32 | 0.33 | 0.48 | 0.02 | 0.01 | 0.04 | 0.03 |
| | $(0, 0)$ | $(1, 5)$ | 0.09 | 0.10 | 0.12 | 0.13 | 0.01 | 0.01 | 0.02 | 0.01 |
| | $(0, 0)$ | $(1, 10)$ | 0.01 | 0 | 0.01 | 0 | 0.01 | 0 | 0.01 | 0 |
| 0.25 | $(3, 0)$ | $(1, 2)$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | $(3, 0)$ | $(1, 3)$ | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 |
| | $(3, 0)$ | $(1, 4)$ | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 |
| | $(3, 0)$ | $(1, 5)$ | 1 | 1 | 1 | 1 | 0.98 | 1 | 1 | 1 |
| | $(3, 0)$ | $(1, 10)$ | 0.98 | 1 | 1 | 1 | 0.97 | 1 | 0.99 | 1 |
| 0.5 | $(3, 0)$ | $(1, 2)$ | 1 | 1 | 1 | 1 | 0.97 | 1 | 1 | 1 |
| | $(3, 0)$ | $(1, 3)$ | 0.97 | 1 | 1 | 1 | 0.51 | 0.74 | 0.72 | 0.92 |
| | $(3, 0)$ | $(1, 4)$ | 0.71 | 0.92 | 0.84 | 0.98 | 0.23 | 0.44 | 0.43 | 0.69 |
| | $(3, 0)$ | $(1, 5)$ | 0.44 | 0.66 | 0.63 | 0.88 | 0.18 | 0.29 | 0.29 | 0.50 |
| | $(3, 0)$ | $(1, 10)$ | 0.18 | 0.28 | 0.26 | 0.42 | 0.12 | 0.18 | 0.19 | 0.32 |
| 0.75 | $(3, 0)$ | $(1, 2)$ | 1 | 1 | 1 | 1 | 0.77 | 0.95 | 0.96 | 1 |
| | $(3, 0)$ | $(1, 3)$ | 0.74 | 0.91 | 0.86 | 0.98 | 0.11 | 0.10 | 0.18 | 0.26 |
| | $(3, 0)$ | $(1, 4)$ | 0.22 | 0.35 | 0.32 | 0.47 | 0.03 | 0.01 | 0.04 | 0.03 |
| | $(3, 0)$ | $(1, 5)$ | 0.09 | 0.09 | 0.10 | 0.14 | 0.01 | 0.00 | 0.02 | 0.01 |
| | $(3, 0)$ | $(1, 10)$ | 0.01 | 0 | 0.01 | 0.01 | 0 | 0 | 0.01 | 0 |

DGP: $y_t = \mu_i + \theta_i DU_{i,t} + \beta_i t + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_{i,t} + z_{i,t}$; $\Delta x_{i,t} = \varepsilon_{i,t}$ and $z_{i,t} = \rho_i z_{i,t-1} + v_{i,t}$ with $\zeta_{i,t} = (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2)$, $\mu_i = 1$, $\beta_i = 0.3$ and $\delta_{i,t} = \delta_{i,1}$ for $t \leq T_{b,i}$ and $\delta_{i,t} = \delta_{i,2}$ for $t > T_{b,i}$. The nominal size is set at the 5% level and 1,000 replications are carried out.

Table 3: Empirical power of Pedroni pseudo t-ratio cointegration statistic. The structural change affects both the deterministic component and the cointegrating vector

| | | | $\rho_i = 0$ | | | | $\rho_i = 0.5$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $N$ $(T = 100)$ | | $N$ $(T = 250)$ | | $N$ $(T = 100)$ | | $N$ $(T = 250)$ | |
| $\lambda_i$ | $(\theta_i, \gamma_i)$ | $(\delta_{i,1}, \delta_{i,2})$ | 20 | 40 | 20 | 40 | 20 | 40 | 20 | 40 |
| 0.25 | $(3, 0.5)$ | $(1, 2)$ | 1 | 1 | 1 | 1 | 0.99 | 1 | 1 | 1 |
| | $(3, 0.5)$ | $(1, 3)$ | 1 | 1 | 1 | 1 | 0.99 | 1 | 0.98 | 1 |
| | $(3, 0.5)$ | $(1, 4)$ | 1 | 1 | 1 | 1 | 0.96 | 1 | 0.95 | 1 |
| | $(3, 0.5)$ | $(1, 5)$ | 0.98 | 1 | 0.98 | 1 | 0.92 | 1 | 0.95 | 1 |
| | $(3, 0.5)$ | $(1, 10)$ | 0.85 | 0.98 | 0.95 | 1 | 0.88 | 0.98 | 0.93 | 1 |
| 0.5 | $(3, 0.5)$ | $(1, 2)$ | 0.43 | 0.72 | 0 | 0 | 0.01 | 0.01 | 0 | 0 |
| | $(3, 0.5)$ | $(1, 3)$ | 0.36 | 0.53 | 0.01 | 0 | 0.05 | 0.04 | 0 | 0 |
| | $(3, 0.5)$ | $(1, 4)$ | 0.28 | 0.41 | 0.03 | 0.01 | 0.08 | 0.09 | 0.01 | 0 |
| | $(3, 0.5)$ | $(1, 5)$ | 0.23 | 0.30 | 0.05 | 0.04 | 0.08 | 0.10 | 0.01 | 0.01 |
| | $(3, 0.5)$ | $(1, 10)$ | 0.14 | 0.21 | 0.08 | 0.13 | 0.12 | 0.19 | 0.09 | 0.10 |
| 0.75 | $(3, 0.5)$ | $(1, 2)$ | 0.71 | 0.89 | 0.04 | 0.02 | 0.04 | 0.02 | 0 | 0 |
| | $(3, 0.5)$ | $(1, 3)$ | 0.52 | 0.68 | 0.11 | 0.08 | 0.08 | 0.08 | 0.01 | 0 |
| | $(3, 0.5)$ | $(1, 4)$ | 0.28 | 0.34 | 0.09 | 0.08 | 0.08 | 0.05 | 0.01 | 0 |
| | $(3, 0.5)$ | $(1, 5)$ | 0.15 | 0.16 | 0.06 | 0.04 | 0.05 | 0.05 | 0.01 | 0.01 |
| | $(3, 0.5)$ | $(1, 10)$ | 0.04 | 0.03 | 0.03 | 0.01 | 0.05 | 0.03 | 0.03 | 0.01 |

DGP: $y_t = \mu_i + \theta_i DU_{i,t} + \beta_i t + \gamma_i DT^*_{i,t} + x'_{i,t}\delta_{i,t} + z_{i,t}$; $\Delta x_{i,t} = \varepsilon_{i,t}$ and $z_{i,t} = \rho_i z_{i,t-1} + v_{i,t}$ with $\zeta_{i,t} = (\varepsilon_{i,t}, v_{i,t})' \sim iid\ N(0, I_2)$, $\mu_i = 1$, $\beta_i = 0.3$ and $\delta_{i,t} = \delta_{i,1}$ for $t \leq T_{b,i}$ and $\delta_{i,t} = \delta_{i,2}$ for $t > T_{b,i}$. The nominal size is set at the 5% level and 1,000 replications are carried out.

Table 4: Asymptotic moments for the test statistics

| | Model 1 | | | | Model 2 | | | | Model 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| $m-1$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| 1 | -25.124 | 73.605 | -3.558 | 0.388 | -31.702 | 80.102 | -4.003 | 0.341 | -36.102 | 98.290 | -4.276 | 0.366 |
| 2 | -30.807 | 89.178 | -3.943 | 0.392 | -37.262 | 97.782 | -4.343 | 0.355 | -41.353 | 113.560 | -4.581 | 0.374 |
| 3 | -36.241 | 99.942 | -4.285 | 0.373 | -42.352 | 112.792 | -4.637 | 0.369 | -46.254 | 124.446 | -4.853 | 0.364 |
| 4 | -41.323 | 113.847 | -4.580 | 0.373 | -47.420 | 127.582 | -4.912 | 0.368 | -51.393 | 136.173 | -5.124 | 0.364 |
| 5 | -46.457 | 121.902 | -4.865 | 0.365 | -51.847 | 136.375 | -5.145 | 0.362 | -56.221 | 148.416 | -5.366 | 0.366 |
| 6 | -51.609 | 142.541 | -5.131 | 0.384 | -56.491 | 152.524 | -5.375 | 0.378 | -60.893 | 159.531 | -5.593 | 0.365 |
| 7 | -56.732 | 151.879 | -5.389 | 0.372 | -61.259 | 163.744 | -5.606 | 0.375 | -65.777 | 172.601 | -5.820 | 0.369 |
| | Model 4 | | | | Model 5 | | | | Model 6 | | | |
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| $m-1$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| 1 | -28.682 | 91.014 | -3.798 | 0.431 | -36.915 | 106.592 | -4.324 | 0.393 | -45.094 | 139.700 | -4.783 | 0.418 |
| 2 | -38.757 | 123.284 | -4.427 | 0.436 | -45.797 | 136.480 | -4.821 | 0.408 | -58.158 | 175.030 | -5.453 | 0.415 |
| 3 | -48.118 | 149.200 | -4.944 | 0.431 | -54.411 | 161.488 | -5.271 | 0.415 | -70.768 | 217.036 | -6.037 | 0.432 |
| 4 | -56.713 | 173.081 | -5.380 | 0.430 | -63.063 | 184.648 | -5.687 | 0.410 | -83.254 | 256.429 | -6.573 | 0.441 |
| 5 | -65.513 | 206.886 | -5.798 | 0.447 | -71.671 | 210.886 | -6.081 | 0.417 | -95.459 | 284.133 | -7.065 | 0.435 |
| 6 | -73.589 | 221.307 | -6.163 | 0.427 | -79.723 | 240.506 | -6.425 | 0.434 | -106.892 | 318.951 | -7.498 | 0.443 |
| 7 | -81.754 | 240.575 | -6.513 | 0.423 | -88.079 | 251.068 | -6.771 | 0.417 | -118.597 | 357.847 | -7.923 | 0.455 |

Table 5: Response surfaces for $(k=0)$

| | Model 1 | | | | Model 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.39 | 60.648 | -3.127 | -19.196 | 0.339 | 67.8 | -3.684 | -26.679 |
| $\hat{\beta}_{0,1}$ | 5.064 | -1226.67 | -8.833 | 121.763 | 6.104 | -1885.589 | -9.439 | 144.172 |
| $\hat{\beta}_{0,2}$ | | | 179.334 | -2571.386 | | 16698.79 | 28.308 | 3575.522 |
| $\hat{\beta}_{0,3}$ | | 196196.7 | -1990.403 | 58983.27 | 1029.447 | | | -72734.42 |
| $\hat{\beta}_{1,0}$ | -0.005 | 16.530 | -0.429 | -6.238 | 0.003 | 17.645 | -0.341 | -5.625 |
| $\hat{\beta}_{1,1}$ | 1 | -1325.654 | | 124.468 | 0.902 | -1543.665 | | 180.54 |
| $\hat{\beta}_{1,2}$ | 34.590 | 42679.5 | -60.807 | -1312.53 | 39.629 | 53149.58 | -51.393 | -4444.318 |
| $\hat{\beta}_{1,3}$ | | -532567.3 | | | | -663605.3 | | 48906.87 |
| $\hat{\beta}_{2,0}$ | | -0.362 | 0.016 | 0.112 | | -0.39 | 0.01 | 0.067 |
| $\hat{\beta}_{2,1}$ | | | -0.236 | 3.084 | | 6.859 | -0.228 | 1.208 |
| $\hat{\beta}_{2,2}$ | | 225.078 | 5.935 | -51.736 | | | 4.325 | |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 3 | | | | Model 4 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.359 | 91.108 | -3.971 | -31.767 | 0.43 | 60.884 | -3.221 | -19.845 |
| $\hat{\beta}_{0,1}$ | 7.472 | -3645.426 | -8.979 | 442.209 | 3.046 | | | 318.553 |
| $\hat{\beta}_{0,2}$ | 59.681 | 75512.06 | -49.326 | -10829.74 | 102.433 | -87968.11 | -110.06 | -16385.62 |
| $\hat{\beta}_{0,3}$ | | -777252.4 | | 164392.7 | | 1874059 | | 307418.8 |
| $\hat{\beta}_{1,0}$ | | 14.514 | -0.314 | -5.334 | | 35.776 | -0.628 | -10.047 |
| $\hat{\beta}_{1,1}$ | 0.852 | -1361.209 | -2.06 | 124.516 | 3.307 | -3225.963 | -2.236 | 219.694 |
| $\hat{\beta}_{1,2}$ | 42.03 | 47092.270 | | -1139.025 | | 121345.6 | | -1980.416 |
| $\hat{\beta}_{1,3}$ | | -562391.3 | | | | -1725484 | | |
| $\hat{\beta}_{2,0}$ | | -0.216 | 0.008 | 0.039 | 0.001 | -1.033 | 0.023 | 0.136 |
| $\hat{\beta}_{2,1}$ | | | 0.038 | 5.521 | -0.165 | | -0.188 | 12.356 |
| $\hat{\beta}_{2,2}$ | | | -3.393 | -128.867 | 11.955 | 797.530 | -3.325 | -290.951 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 5 | | | | Model 6 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.364 | 74.286 | -3.78 | -27.851 | 0.366 | 87.342 | -3.968 | -30.483 |
| $\hat{\beta}_{0,1}$ | 6.564 | -2146.293 | -6.974 | 242.942 | 10.855 | -2699.1 | -8.23 | 191.322 |
| $\hat{\beta}_{0,2}$ | | 20055.64 | | | -266.021 | 28358.7 | -83.457 | 4931.966 |
| $\hat{\beta}_{0,3}$ | | | | -42063.93 | 7384.621 | | | -123591.5 |
| $\hat{\beta}_{1,0}$ | 0.008 | 34.679 | -0.544 | -9.615 | 0.007 | 33.827 | -0.505 | -9.373 |
| $\hat{\beta}_{1,1}$ | 2.617 | -3212.648 | -0.868 | 322.608 | 3.982 | -3213.574 | -1.767 | 357.392 |
| $\hat{\beta}_{1,2}$ | 41.638 | 115262.4 | -43.717 | -8330.38 | | 118816.1 | | -9875.101 |
| $\hat{\beta}_{1,3}$ | | -1488387 | | 113929.5 | | -1614095 | | 131909.7 |
| $\hat{\beta}_{2,0}$ | | -1.053 | 0.018 | 0.097 | | -0.888 | 0.014 | 0.072 |
| $\hat{\beta}_{2,1}$ | -0.161 | 24.408 | -0.306 | 11.106 | -0.325 | | -0.189 | 9.476 |
| $\hat{\beta}_{2,2}$ | 10.166 | | | -273.637 | 15.916 | 730.025 | -5.392 | -222.194 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

Table 6: Response surfaces for $(k = 2)$

| | Model 1 | | | | Model 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.415 | 62.309 | -3.213 | -19.672 | 0.336 | 69.482 | -3.735 | -26.724 |
| $\hat{\beta}_{0,1}$ | 0.967 | -104.685 | 4.102 | 91.646 | 2.873 | | 1.953 | -42.725 |
| $\hat{\beta}_{0,2}$ | 85.478 | | -428.601 | -6704.974 | 58.52 | -18212.26 | -286.032 | 421.283 |
| $\hat{\beta}_{0,3}$ | | | 6757.605 | 103243.4 | | 275990.6 | 5567.945 | |
| $\hat{\beta}_{1,0}$ | -0.018 | 15.196 | -0.414 | -5.961 | 0.005 | 15.915 | -0.334 | -5.411 |
| $\hat{\beta}_{1,1}$ | 1.579 | -172.849 | 4.4 | 17.452 | 1.368 | -236.01 | 6.006 | 56.212 |
| $\hat{\beta}_{1,2}$ | | | -26.560 | 438.162 | -30.879 | | -59.458 | -245.555 |
| $\hat{\beta}_{1,3}$ | | | | | 612.861 | | | |
| $\hat{\beta}_{2,0}$ | 0.002 | -0.147 | 0.015 | 0.08 | -0.001 | -0.195 | 0.010 | 0.05 |
| $\hat{\beta}_{2,1}$ | -0.085 | -9.382 | | 4.173 | | | -0.138 | 0.614 |
| $\hat{\beta}_{2,2}$ | | | -2.239 | -71.787 | | | | |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 3 | | | | Model 4 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.353 | 89.831 | -4.011 | -31.141 | 0.429 | 66.591 | -3.235 | -19.246 |
| $\hat{\beta}_{0,1}$ | 6.456 | -173.345 | 5.695 | 25.550 | 1.626 | -1025.367 | | 66.531 |
| $\hat{\beta}_{0,2}$ | | -6455.393 | -543.11 | -4224.155 | 100.548 | 30787.53 | -76.879 | -6527.479 |
| $\hat{\beta}_{0,3}$ | | | 8627.961 | 84886.75 | | | | 120101.1 |
| $\hat{\beta}_{1,0}$ | 0.006 | 14.775 | -0.317 | -5.476 | -0.002 | 29.482 | -0.624 | -9.880 |
| $\hat{\beta}_{1,1}$ | -1.009 | -274.989 | 5.92 | 63.485 | 2.983 | 438.987 | 13.891 | 135.547 |
| $\hat{\beta}_{1,2}$ | 81.566 | | -53.692 | -245.299 | -45.199 | -24349.6 | -374.707 | -2851.772 |
| $\hat{\beta}_{1,3}$ | -631.881 | | | | | | 4741.349 | 32282.1 |
| $\hat{\beta}_{2,0}$ | -0.001 | -0.155 | 0.009 | 0.054 | | | 0.024 | 0.104 |
| $\hat{\beta}_{2,1}$ | 0.181 | | -0.147 | | -0.199 | -138.416 | -0.304 | 2.734 |
| $\hat{\beta}_{2,2}$ | -5.953 | | | | 6.356 | 3434.739 | | |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 5 | | | | Model 6 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.380 | 78.16 | -3.825 | -27.922 | 0.383 | 91.354 | -4.016 | -31.322 |
| $\hat{\beta}_{0,1}$ | | -1049.361 | 4.26 | 97.299 | 2.626 | | 6.241 | 156.004 |
| $\hat{\beta}_{0,2}$ | 94.123 | 11495.16 | -98.231 | -5411.362 | 90.144 | -40668.1 | -493.482 | -11876.56 |
| $\hat{\beta}_{0,3}$ | | | | 88658.21 | | 521643.9 | 7199.83 | 218847.8 |
| $\hat{\beta}_{1,0}$ | 0.004 | 29.349 | -0.524 | -9.171 | 0.011 | 29.639 | -0.488 | -8.640 |
| $\hat{\beta}_{1,1}$ | 2.825 | | 14.206 | 143.512 | 2.271 | -281.767 | 13.469 | 106.736 |
| $\hat{\beta}_{1,2}$ | | -7443.609 | -434.678 | -3425.206 | | 4060.874 | -326.613 | -744.496 |
| $\hat{\beta}_{1,3}$ | | | 5633.642 | 49471.7 | | | 3497.908 | |
| $\hat{\beta}_{2,0}$ | | | 0.017 | 0.047 | -0.001 | | 0.014 | |
| $\hat{\beta}_{2,1}$ | -0.136 | -86.63 | -0.236 | 5.337 | -0.121 | -58.02 | -0.272 | 6.34 |
| $\hat{\beta}_{2,2}$ | | 1363.246 | | -74.929 | 1.486 | | | -89.481 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

Table 7: Response surfaces for $(k=5)$

| | Model 1 | | | | Model 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.411 | 61.076 | -3.196 | -19.09 | 0.327 | 70.537 | -3.758 | -26.36 |
| $\hat{\beta}_{0,1}$ | | 2333.282 | -2.138 | -251.033 | 1.926 | 3688.82 | 9.947 | -577.95 |
| $\hat{\beta}_{0,2}$ | 89.8 | 14804.35 | | -2084.949 | | -111525 | -102.888 | 8408.465 |
| $\hat{\beta}_{0,3}$ | 2785.821 | | -2085.401 | | 5474.382 | 2935591 | | -70550.62 |
| $\hat{\beta}_{1,0}$ | -0.018 | 14.491 | -0.419 | -5.96 | 0.008 | 14.356 | -0.324 | -5.371 |
| $\hat{\beta}_{1,1}$ | 1.282 | 1468.171 | 15.196 | -102.32 | 0.596 | 1834.07 | 14.124 | -70.904 |
| $\hat{\beta}_{1,2}$ | | -14669.95 | -348.385 | 2139.019 | | -25876.53 | -368.115 | 1714.215 |
| $\hat{\beta}_{1,3}$ | | | 4192.348 | | | | 4228.759 | |
| $\hat{\beta}_{2,0}$ | 0.001 | | 0.016 | 0.068 | -0.001 | | 0.010 | 0.029 |
| $\hat{\beta}_{2,1}$ | -0.069 | | -0.289 | 1.362 | | | -0.172 | |
| $\hat{\beta}_{2,2}$ | | | | | | | | |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 3 | | | | Model 4 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.367 | 89.609 | -4.013 | -30.645 | 0.435 | 58.969 | -3.269 | -19.333 |
| $\hat{\beta}_{0,1}$ | | 6021.446 | 10.749 | -722.651 | | 3904.387 | 8.328 | -180.318 |
| $\hat{\beta}_{0,2}$ | 139.06 | -119796.1 | -322.281 | 10502.56 | | -154668.9 | -527.974 | -3147.907 |
| $\hat{\beta}_{0,3}$ | 4467.837 | 2779171 | | -173970.8 | 4215.836 | 3727902 | 6371.796 | |
| $\hat{\beta}_{1,0}$ | -0.004 | 13.944 | -0.307 | -5.325 | -0.003 | 27.297 | -0.59 | -9.213 |
| $\hat{\beta}_{1,1}$ | 1.052 | 1272.166 | 15.51 | -75.522 | 1.582 | 3537.474 | 24.265 | -136.207 |
| $\hat{\beta}_{1,2}$ | -11.117 | | -394.685 | 1780.818 | 9.137 | -47364.92 | -666.318 | 2432.973 |
| $\hat{\beta}_{1,3}$ | | | 4719.719 | | | | 8318.551 | |
| $\hat{\beta}_{2,0}$ | | | 0.008 | 0.014 | | 0.826 | 0.023 | |
| $\hat{\beta}_{2,1}$ | | 70.693 | -0.27 | | -0.092 | | -0.479 | |
| $\hat{\beta}_{2,2}$ | | -2726.643 | | | | | | 81.938 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 5 | | | | Model 6 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.343 | 60.513 | -3.828 | -27.858 | 0.378 | 71.175 | -4.026 | -31.256 |
| $\hat{\beta}_{0,1}$ | 1.636 | 6899.273 | 12.998 | -253.539 | 0.867 | 11057.17 | 13.951 | -408.183 |
| $\hat{\beta}_{0,2}$ | -88.332 | -322583.2 | -182.882 | -13556.49 | | -499199.5 | -403.515 | -10959.25 |
| $\hat{\beta}_{0,3}$ | 7587.446 | 5934887 | | 359446.7 | 6664.312 | 10140177 | | 228140.6 |
| $\hat{\beta}_{1,0}$ | 0.016 | 32.949 | -0.496 | -8.767 | 0.011 | 29.816 | -0.451 | -8.12 |
| $\hat{\beta}_{1,1}$ | 1.581 | 2404.486 | 25.834 | -154.416 | 1.509 | 3570.033 | 24.749 | -194.826 |
| $\hat{\beta}_{1,2}$ | | | -842.643 | 6727.948 | | -40673.95 | -789.654 | 8114.972 |
| $\hat{\beta}_{1,3}$ | | | 9999.87 | -124097.4 | | | 10535.88 | -121465.5 |
| $\hat{\beta}_{2,0}$ | -0.001 | | 0.017 | -5.136 | -0.001 | 0.754 | 0.012 | -0.078 |
| $\hat{\beta}_{2,1}$ | -0.083 | 187.542 | -0.574 | 247.073 | -0.074 | | -0.331 | |
| $\hat{\beta}_{2,2}$ | | -7067.742 | 9.204 | | | | | 120.421 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

Table 8: Response surfaces for the automatic lag length selection method ($k_{max} = 5$)

| | Model 1 | | | | Model 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{t}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{\rho}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{t}_{NT}}(\hat{\lambda})$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.41 | 56.823 | -3.218 | -19.638 | 0.372 | 71.034 | -3.778 | -26.654 |
| $\hat{\beta}_{0,1}$ | 10.777 | 2079.863 | -34.87 | -97.193 | 1.676 | 1730.194 | -42.359 | -392.725 |
| $\hat{\beta}_{0,2}$ | -284.429 | | 737.622 | -3103.602 | 97.645 | 40207.55 | 1018.228 | 4124.832 |
| $\hat{\beta}_{0,3}$ | 4332.145 | | -11377.84 | | | | -13147.4 | |
| $\hat{\beta}_{1,0}$ | -0.004 | 18.14 | -0.442 | -6.027 | 0.005 | 13.145 | -0.351 | -5.628 |
| $\hat{\beta}_{1,1}$ | -2.036 | | 1.628 | -68.79 | | 1293.969 | 3.225 | |
| $\hat{\beta}_{1,2}$ | 55.887 | 28710.63 | | 1511.876 | | -18644.32 | -36.265 | |
| $\hat{\beta}_{1,3}$ | | | | | | | | |
| $\hat{\beta}_{2,0}$ | | -0.748 | 0.017 | 0.081 | -0.001 | | 0.01 | 0.064 |
| $\hat{\beta}_{2,1}$ | 0.165 | 205.976 | -0.114 | 0.967 | | 48.061 | -0.161 | -5.218 |
| $\hat{\beta}_{2,2}$ | | -5962.404 | | | 5.98 | | | 140.98 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 3 | | | | Model 4 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{t}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{\rho}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{t}_{NT}}(\hat{\lambda})$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.389 | 72.251 | -4.061 | -31.033 | 0.517 | 70.453 | -3.286 | -19.519 |
| $\hat{\beta}_{0,1}$ | 5.779 | 7427.681 | -43.941 | -465.591 | 1.919 | | -26.176 | -101.832 |
| $\hat{\beta}_{0,2}$ | -225.895 | -177465.4 | 921.364 | 1330.639 | | 44801.21 | 166.728 | -2334.409 |
| $\hat{\beta}_{0,3}$ | 5584.734 | 2808044 | -13082.02 | | | | | |
| $\hat{\beta}_{1,0}$ | | 19.721 | -0.335 | -5.637 | -0.02 | 26.003 | -0.649 | -9.78 |
| $\hat{\beta}_{1,1}$ | -0.798 | | 3.616 | | | 2162.096 | 3.806 | -26.883 |
| $\hat{\beta}_{1,2}$ | 56.865 | 37740.3 | -30.174 | | 72.559 | | -45.143 | |
| $\hat{\beta}_{1,3}$ | | | | | | | | |
| $\hat{\beta}_{2,0}$ | 0.001 | -0.737 | 0.009 | 0.059 | 0.001 | | 0.025 | 0.086 |
| $\hat{\beta}_{2,1}$ | 0.093 | 190.91 | -0.194 | -6.067 | 0.176 | 275.749 | -0.227 | -8.473 |
| $\hat{\beta}_{2,2}$ | | -5491.499 | | 146.45 | | -8513.873 | | 292.56 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

| | Model 5 | | | | Model 6 | | | |
|---|---|---|---|---|---|---|---|---|
| | $Z_{\hat{\rho}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{t}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{\rho}_{NT}}(\hat{\lambda})$ | | $Z_{\hat{t}_{NT}}(\hat{\lambda})$ | |
| | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ | $\Theta_1$ | $\Psi_1$ | $\Theta_2$ | $\Psi_2$ |
| $\hat{\beta}_{0,0}$ | 0.399 | 109.977 | -3.875 | -27.694 | 0.424 | 87.103 | -4.071 | -31.407 |
| $\hat{\beta}_{0,1}$ | | -8193.521 | -35.047 | -296.345 | | 5713.206 | -41.846 | -243.518 |
| $\hat{\beta}_{0,2}$ | 119.632 | 607421 | 681.665 | 1996.116 | 147.021 | -286832.2 | 938.021 | -12550.07 |
| $\hat{\beta}_{0,3}$ | | -9915995 | -8374.721 | | | 7973939 | -14997.98 | 240521.8 |
| $\hat{\beta}_{1,0}$ | 0.011 | 7.772 | -0.549 | -9.262 | 0.011 | 19.269 | -0.509 | -8.675 |
| $\hat{\beta}_{1,1}$ | 0.937 | 6841.871 | 4.83 | -9.079 | 0.858 | 4385.407 | 5.806 | -66.601 |
| $\hat{\beta}_{1,2}$ | | -256154.3 | -74.577 | -540.416 | | -58727.81 | -111.692 | 3590.481 |
| $\hat{\beta}_{1,3}$ | | 3915350 | | | | | 1563.76 | -69517.66 |
| $\hat{\beta}_{2,0}$ | -0.001 | 1.661 | 0.018 | 0.048 | -0.001 | 1.239 | 0.014 | |
| $\hat{\beta}_{2,1}$ | | | -0.235 | -8.318 | | | -0.178 | -7.49 |
| $\hat{\beta}_{2,2}$ | 13.846 | | | 283.209 | 15.245 | | -3.999 | 271.886 |
| $\hat{\beta}_{2,3}$ | | | | | | | | |

Table 9: Mean and variance for the $ADF_{\tilde{e}}^{c}$, $ADF_{\tilde{e}}^{\tau}$ and $ADF_{\tilde{e}}^{\gamma}$ statistics

| $ADF_{\tilde{e}}^{c}(i)$ statistic | | | $ADF_{\tilde{e}}^{\tau}(i)$ statistic | | |
|---|---|---|---|---|---|
| $T$ | $\Theta_2^e$ | $\Psi_2^e$ | $T$ | $\Theta_2^e$ | $\Psi_2^e$ |
| 50 | -0.418 | 0.991 | 50 | -1.549 | 0.367 |
| 100 | -0.419 | 0.980 | 100 | -1.541 | 0.353 |
| 250 | -0.424 | 0.955 | 250 | -1.538 | 0.346 |
| 500 | -0.418 | 0.959 | 500 | -1.536 | 0.346 |
| 1000 | -0.424 | 0.964 | 1000 | -1.535 | 0.341 |

$ADF_{\tilde{e}}^{\gamma}(i)$ statistic

| $T$ | $\lambda$ | $\Theta_2^e$ | $\Psi_2^e$ | $T$ | $\lambda$ | $\Theta_2^e$ | $\Psi_2^e$ | $T$ | $\lambda$ | $\Theta_2^e$ | $\Psi_2^e$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.1 | -1.684 | 0.423 | 100 | 0.1 | -1.680 | 0.405 | 250 | 0.1 | -1.682 | 0.399 |
| | 0.2 | -1.829 | 0.450 | | 0.2 | -1.816 | 0.415 | | 0.2 | -1.810 | 0.394 |
| | 0.3 | -1.932 | 0.414 | | 0.3 | -1.920 | 0.402 | | 0.3 | -1.904 | 0.378 |
| | 0.4 | -2.013 | 0.398 | | 0.4 | -1.981 | 0.368 | | 0.4 | -1.957 | 0.354 |
| | 0.5 | -2.022 | 0.383 | | 0.5 | -1.998 | 0.358 | | 0.5 | -1.967 | 0.330 |
| | 0.6 | -2.011 | 0.404 | | 0.6 | -1.975 | 0.368 | | 0.6 | -1.961 | 0.349 |
| | 0.7 | -1.940 | 0.425 | | 0.7 | -1.913 | 0.390 | | 0.7 | -1.913 | 0.385 |
| | 0.8 | -1.834 | 0.447 | | 0.8 | -1.817 | 0.423 | | 0.8 | -1.808 | 0.402 |
| | 0.9 | -1.681 | 0.423 | | 0.9 | -1.676 | 0.397 | | 0.9 | -1.682 | 0.400 |
| 500 | 0.1 | -1.688 | 0.395 | 1000 | 0.1 | -1.676 | 0.389 | | | | |
| | 0.2 | -1.812 | 0.395 | | 0.2 | -1.809 | 0.392 | | | | |
| | 0.3 | -1.900 | 0.369 | | 0.3 | -1.902 | 0.371 | | | | |
| | 0.4 | -1.954 | 0.343 | | 0.4 | -1.950 | 0.338 | | | | |
| | 0.5 | -1.967 | 0.330 | | 0.5 | -1.972 | 0.339 | | | | |
| | 0.6 | -1.955 | 0.344 | | 0.6 | -1.953 | 0.346 | | | | |
| | 0.7 | -1.898 | 0.369 | | 0.7 | -1.900 | 0.365 | | | | |
| | 0.8 | -1.800 | 0.396 | | 0.8 | -1.799 | 0.390 | | | | |
| | 0.9 | -1.678 | 0.392 | | 0.9 | -1.691 | 0.392 | | | | |

Table 10: Asymptotic critical values for the MQ tests

| | $\lambda = 0.1$ | | | $\lambda = 0.2$ | | | $\lambda = 0.3$ | | |
|---|---|---|---|---|---|---|---|---|---|
| $r$ | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| 1 | -32.163 | -23.629 | -19.865 | -34.858 | -26.091 | -22.144 | -36.123 | -27.562 | -23.619 |
| 2 | -43.372 | -34.321 | -30.056 | -46.436 | -37.139 | -32.688 | -46.773 | -37.778 | -33.492 |
| 3 | -53.648 | -44.378 | -39.748 | -55.828 | -46.232 | -41.766 | -57.136 | -47.511 | -42.775 |
| 4 | -63.359 | -53.470 | -48.595 | -65.206 | -55.582 | -50.645 | -65.570 | -55.883 | -51.370 |
| 5 | -73.691 | -62.796 | -57.434 | -74.601 | -64.165 | -59.199 | -75.573 | -64.731 | -59.919 |
| 6 | -81.346 | -71.238 | -65.663 | -83.575 | -72.562 | -67.309 | -83.921 | -73.247 | -67.908 |
| | $\lambda = 0.4$ | | | $\lambda = 0.5$ | | | $\lambda = 0.6$ | | |
| r | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| 1 | -36.635 | -28.147 | -24.140 | -36.775 | -28.226 | -24.419 | -36.805 | -28.178 | -24.176 |
| 2 | -47.134 | -38.391 | -34.282 | -48.148 | -38.907 | -34.553 | -47.611 | -38.587 | -34.246 |
| 3 | -57.176 | -47.642 | -43.088 | -56.753 | -47.715 | -43.333 | -57.230 | -47.865 | -43.200 |
| 4 | -67.481 | -56.958 | -52.039 | -65.752 | -56.418 | -51.708 | -67.094 | -56.599 | -51.785 |
| 5 | -75.603 | -65.386 | -60.204 | -75.378 | -65.302 | -60.251 | -75.182 | -64.986 | -60.057 |
| 6 | -84.718 | -73.703 | -68.372 | -83.902 | -73.746 | -68.222 | -84.059 | -73.136 | -67.973 |
| | $\lambda = 0.7$ | | | $\lambda = 0.8$ | | | $\lambda = 0.9$ | | |
| r | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| 1 | -36.302 | -27.751 | -23.890 | -35.249 | -26.722 | -22.713 | -32.918 | -24.712 | -20.896 |
| 2 | -47.383 | -38.223 | -34.045 | -46.572 | -37.227 | -33.085 | -43.959 | -35.248 | -31.190 |
| 3 | -56.908 | -47.282 | -42.693 | -55.960 | -46.442 | -41.998 | -54.568 | -45.183 | -40.623 |
| 4 | -66.869 | -56.270 | -51.337 | -65.833 | -55.750 | -50.890 | -63.920 | -53.985 | -49.399 |
| 5 | -75.074 | -64.828 | -59.867 | -74.046 | -64.430 | -59.290 | -74.177 | -63.063 | -57.839 |
| 6 | -85.434 | -73.646 | -68.332 | -83.244 | -72.857 | -67.721 | -82.664 | -71.518 | -66.449 |

Table 11: Critical values for the $Z^{e}_{\hat{t}_{NT}}(\hat{\lambda})$, $ADF_{\tilde{F}}(\hat{\lambda})$ and $MQ(q,\hat{\lambda})$ statistics

**Panel A: $Z^{e}_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic**

Constant with or without change in level

| $T$ | 1% | 2.5% | 5% | 10% |
|---|---|---|---|---|
| 50 | -2.926 | -2.517 | -2.219 | -1.901 |
| 100 | -2.824 | -2.402 | -2.113 | -1.759 |
| 250 | -2.560 | -2.250 | -1.985 | -1.619 |

Time trend with or without change in level

| $T$ | 1% | 2.5% | 5% | 10% |
|---|---|---|---|---|
| 50 | -2.389 | -2.042 | -1.670 | -1.273 |
| 100 | -2.441 | -2.040 | -1.708 | -1.357 |
| 250 | -2.296 | -1.953 | -1.619 | -1.260 |

Time trend with one change in trend

| $T$ | 1% | 2.5% | 5% | 10% |
|---|---|---|---|---|
| 50 | -3.679 | -3.389 | -3.097 | -2.714 |
| 100 | -3.826 | -3.467 | -3.147 | -2.804 |
| 250 | -3.740 | -3.373 | -3.134 | -2.794 |

**Panel B: Common factor statistics**

$ADF_{\tilde{F}}\left(\hat{\lambda}\right)$: Time trend with one change in trend

| $T$ | 1% | 2.5% | 5% | 10% |
|---|---|---|---|---|
| 50 | -4.779 | -4.306 | -4.008 | -3.679 |
| 100 | -4.549 | -4.243 | -3.930 | -3.602 |
| 250 | -4.474 | -4.136 | -3.873 | -3.594 |

$MQ\left(q,\hat{\lambda}\right)$

| $T$ | $r$ | 1% | 2.5% | 5% | 10% |
|---|---|---|---|---|---|
| 50 | 1 | -31.046 | -27.569 | -24.828 | -21.669 |
| | 2 | -38.827 | -35.362 | -32.792 | -29.925 |
| | 3 | -44.744 | -42.436 | -39.703 | -36.641 |
| | 4 | -47.752 | -46.476 | -44.865 | -42.381 |
| | 5 | -48.756 | -48.305 | -47.472 | -46.119 |
| | 6 | -48.890 | -48.746 | -48.444 | -47.879 |
| 100 | 1 | -34.474 | -30.234 | -26.833 | -23.102 |
| | 2 | -44.748 | -40.147 | -36.464 | -32.729 |
| | 3 | -53.423 | -49.142 | -45.879 | -41.862 |
| | 4 | -61.972 | -57.307 | -53.251 | -49.284 |
| | 5 | -69.033 | -64.937 | -61.099 | -56.747 |
| | 6 | -74.663 | -70.434 | -67.183 | -63.437 |
| 250 | 1 | -32.985 | -28.983 | -25.697 | -22.843 |
| | 2 | -46.953 | -41.768 | -38.103 | -33.778 |
| | 3 | -52.827 | -48.542 | -45.066 | -41.136 |
| | 4 | -59.494 | -56.474 | -53.392 | -49.240 |
| | 5 | -70.495 | -66.474 | -62.404 | -57.440 |
| | 6 | -78.589 | -73.456 | -68.748 | -64.459 |

Table 12: Empirical size of the tests (nominal size = 5%)

$Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ statistic

| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|---|
| 20 | 50 | 0.039 | 0.046 | 0.043 | 0.033 | 0.054 | 0.045 |
| | 100 | 0.055 | 0.049 | 0.053 | 0.059 | 0.048 | 0.050 |
| | 250 | 0.050 | 0.053 | 0.046 | 0.052 | 0.056 | 0.059 |
| 40 | 50 | 0.040 | 0.049 | 0.046 | 0.030 | 0.044 | 0.056 |
| | 100 | 0.047 | 0.047 | 0.057 | 0.066 | 0.051 | 0.047 |
| | 250 | 0.056 | 0.061 | 0.047 | 0.044 | 0.046 | 0.055 |

$Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic

| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|---|
| 20 | 50 | 0.044 | 0.045 | 0.049 | 0.047 | 0.050 | 0.045 |
| | 100 | 0.050 | 0.050 | 0.045 | 0.046 | 0.043 | 0.053 |
| | 250 | 0.043 | 0.047 | 0.043 | 0.040 | 0.049 | 0.053 |
| 40 | 50 | 0.045 | 0.051 | 0.055 | 0.048 | 0.041 | 0.052 |
| | 100 | 0.041 | 0.047 | 0.047 | 0.044 | 0.046 | 0.043 |
| | 250 | 0.048 | 0.053 | 0.046 | 0.032 | 0.045 | 0.048 |

Simulation results based on 5,000 replications.

Table 13: Empirical power of the normalised bias statistic (nominal size = 5%)

| | | | | $Z_{\hat{\rho}_{NT}}\left(\hat{\lambda}\right)$ statistic | | | |
|---|---|---|---|---|---|---|---|
| | | | | **DGP: Model 1** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | **0.455** | 0.312 | 0.216 | 0.361 | 0.223 | 0.183 |
| | 100 | **1** | 0.998 | 0.989 | 1 | 0.931 | 0.980 |
| | 250 | **1** | 1 | 1 | 1 | 1 | 1 |
| 40 | 50 | **0.676** | 0.467 | 0.320 | 0.577 | 0.310 | 0.269 |
| | 100 | **1** | 1 | 1 | 1 | 0.998 | 1 |
| | 250 | **1** | 1 | 1 | 1 | 1 | 1 |
| | | | | **DGP: Model 2** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0.001 | **0.306** | 0.219 | 0.004 | 0.211 | 0.185 |
| | 100 | 0 | **1** | 0.988 | 0.001 | 0.935 | 0.983 |
| | 250 | 0 | **1** | 1 | 0.000 | 1 | 1 |
| 40 | 50 | 0 | **1** | 0.334 | 0.001 | 0.309 | 0.261 |
| | 100 | 0 | **1** | 1 | 0.000 | 0.998 | 1 |
| | 250 | 0 | **1** | 1 | 0 | 1 | 1 |
| | | | | **DGP: Model 3** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0 | 0.016 | **0.110** | 0 | 0.015 | 0.088 |
| | 100 | 0 | 0.089 | **0.907** | 0.001 | 0.121 | 0.861 |
| | 250 | 0 | 0.932 | **1** | 0 | 0.998 | 1 |
| 40 | 50 | 0 | 0.010 | **0.125** | 0 | 0.005 | 0.129 |
| | 100 | 0 | 0.085 | **0.995** | 0 | 0.159 | 0.992 |
| | 250 | 0 | 0.787 | **1** | 0 | 0.997 | 1 |
| | | | | **DGP: Model 4** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0.864 | 0.389 | 0.527 | **0.987** | 0.478 | 0.413 |
| | 100 | 1 | 1 | 1 | **1** | 1 | 1 |
| | 250 | 1 | 1 | 1 | **1** | 1 | 1 |
| 40 | 50 | 0.996 | 0.754 | 0.671 | **1** | 0.781 | 0.687 |
| | 100 | 1 | 1 | 1 | **1** | 1 | 1 |
| | 250 | 1 | 1 | 1 | **1** | 1 | 1 |
| | | | | **DGP: Model 5** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0.093 | 0.356 | 0.330 | 0.187 | **0.485** | 0.578 |
| | 100 | 0.236 | 0.999 | 1 | 0.283 | **1** | 1 |
| | 250 | 0.044 | 1 | 1 | 0.113 | **1** | 1 |
| 40 | 50 | 0.089 | 0.657 | 0.667 | 0.233 | **0.714** | 0.743 |
| | 100 | 0.305 | 1 | 1 | 0.515 | **1** | 1 |
| | 250 | 0.037 | 1 | 1 | 0.105 | **1** | 1 |
| | | | | **DGP: Model 6** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0.002 | 0.151 | 0.311 | 0.018 | 0.077 | **0.546** |
| | 100 | 0.009 | 0.990 | 1 | 0.054 | 0.914 | **1** |
| | 250 | 0.001 | 0.997 | 1 | 0.022 | 0.998 | **1** |
| 40 | 50 | 0 | 0.328 | 0.606 | 0.005 | 0.244 | **0.785** |
| | 100 | 0 | 1 | 1 | 0.021 | 0.994 | **1** |
| | 250 | 0 | 1 | 1 | 0.003 | 1 | **1** |

Simulation results based on 5,000 replications.

Table 14: Empirical power of the pseudo $t$-ratio statistic (nominal size = 5%)

| | | | | $Z_{\hat{t}_{NT}}\left(\hat{\lambda}\right)$ statistic | | | |
|---|---|---|---|---|---|---|---|
| | | | | **DGP: Model 1** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | **1** | 1 | 1 | 1 | 1 | 1 |
| | 100 | **1** | 1 | 1 | 1 | 1 | 1 |
| | 250 | **1** | 1 | 1 | 1 | 1 | 1 |
| 40 | 50 | **1** | 1 | 1 | 1 | 1 | 1 |
| | 100 | **1** | 1 | 1 | 1 | 1 | 1 |
| | 250 | **1** | 1 | 1 | 1 | 1 | 1 |
| | | | | **DGP: Model 2** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0.957 | **1** | 1 | 0.995 | 1 | 1 |
| | 100 | 0.403 | **1** | 1 | 0.675 | 1 | 1 |
| | 250 | 0.034 | **1** | 1 | 0.073 | 1 | 1 |
| 40 | 50 | 1 | **1** | 1 | 1 | 1 | 1 |
| | 100 | 0.647 | **1** | 1 | 0.908 | 1 | 1 |
| | 250 | 0.026 | **1** | 1 | 0.070 | 1 | 1 |
| | | | | **DGP: Model 3** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0.303 | 1 | **1** | 0.765 | 1 | 1 |
| | 100 | 0.018 | 1 | **1** | 0.221 | 1 | 1 |
| | 250 | 0.001 | 1 | **1** | 0.009 | 1 | 1 |
| 40 | 50 | 0.497 | 1 | **1** | 0.958 | 1 | 1 |
| | 100 | 0.014 | 1 | **1** | 0.324 | 1 | 1 |
| | 250 | 0 | 1 | **1** | 0.003 | 1 | 1 |
| | | | | **DGP: Model 4** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 1 | 1 | 1 | **1** | 1 | 1 |
| | 100 | 1 | 1 | 1 | **1** | 1 | 1 |
| | 250 | 1 | 1 | 1 | **1** | 1 | 1 |
| 40 | 50 | 1 | 1 | 1 | **1** | 1 | 1 |
| | 100 | 1 | 1 | 1 | **1** | 1 | 1 |
| | 250 | 1 | 1 | 1 | **1** | 1 | 1 |
| | | | | **DGP: Model 5** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 1 | 1 | 1 | 1 | **1** | 1 |
| | 100 | 0.917 | 1 | 1 | 0.981 | **1** | 1 |
| | 250 | 0.219 | 1 | 1 | 0.413 | **1** | 1 |
| 40 | 50 | 1 | 1 | 1 | 1 | **1** | 1 |
| | 100 | 0.992 | 1 | 1 | 1 | **1** | 1 |
| | 250 | 0.308 | 1 | 1 | 0.594 | **1** | 1 |
| | | | | **DGP: Model 6** | | | |
| $N$ | $T$ | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| 20 | 50 | 0.976 | 1 | 1 | 0.999 | 1 | **1** |
| | 100 | 0.287 | 1 | 1 | 0.741 | 1 | **1** |
| | 250 | 0.009 | 1 | 1 | 0.083 | 1 | **1** |
| 40 | 50 | 0.999 | 1 | 1 | 1 | 1 | **1** |
| | 100 | 0.229 | 1 | 1 | 0.901 | 1 | **1** |
| | 250 | 0.003 | 1 | 1 | 0.060 | 1 | **1** |

Simulation results based on 5,000 replications.

Table 15: Empirical size and power when there is one common factor ($N = 40$)

| | | | Constant case: $f_i(t) = \mu_i$ | | | | | | Trend case: $f_i(t) = \mu_i + \beta_i t$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\sigma_F^2 = 0.5$ | | $\sigma_F^2 = 1$ | | $\sigma_F^2 = 10$ | | $\sigma_F^2 = 0.5$ | | $\sigma_F^2 = 1$ | | $\sigma_F^2 = 10$ | |
| $T$ | $\rho_i$ | $\phi$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ |
| 50 | 1 | 1 | 0.049 | 0.067 | 0.039 | 0.045 | 0.052 | 0.065 | 0.054 | 0.077 | 0.054 | 0.058 | 0.058 | 0.062 |
| 100 | 1 | 1 | 0.056 | 0.047 | 0.054 | 0.058 | 0.058 | 0.048 | 0.069 | 0.056 | 0.062 | 0.042 | 0.063 | 0.052 |
| 250 | 1 | 1 | 0.053 | 0.047 | 0.050 | 0.052 | 0.055 | 0.055 | 0.059 | 0.048 | 0.048 | 0.049 | 0.060 | 0.052 |
| 50 | 1 | 0.9 | 0.039 | 0.141 | 0.041 | 0.110 | 0.032 | 0.121 | 0.053 | 0.108 | 0.033 | 0.116 | 0.052 | 0.104 |
| 100 | 1 | 0.9 | 0.049 | 0.288 | 0.042 | 0.305 | 0.053 | 0.345 | 0.054 | 0.227 | 0.038 | 0.200 | 0.053 | 0.191 |
| 250 | 1 | 0.9 | 0.039 | 0.853 | 0.047 | 0.928 | 0.052 | 0.959 | 0.052 | 0.773 | 0.050 | 0.829 | 0.042 | 0.838 |
| 50 | 1 | 0.8 | 0.043 | 0.294 | 0.038 | 0.366 | 0.048 | 0.337 | 0.042 | 0.226 | 0.061 | 0.212 | 0.045 | 0.220 |
| 100 | 1 | 0.8 | 0.057 | 0.763 | 0.042 | 0.817 | 0.041 | 0.854 | 0.070 | 0.611 | 0.059 | 0.642 | 0.052 | 0.661 |
| 250 | 1 | 0.8 | 0.054 | 0.995 | 0.046 | 1 | 0.054 | 1 | 0.053 | 1 | 0.043 | 1 | 0.044 | 1 |
| 50 | 0.9 | 1 | 1 | 0.047 | 1 | 0.049 | 1 | 0.060 | 0.965 | 0.068 | 0.971 | 0.084 | 0.875 | 0.071 |
| 100 | 0.9 | 1 | 1 | 0.054 | 1 | 0.061 | 1 | 0.048 | 1 | 0.071 | 1 | 0.053 | 1 | 0.061 |
| 250 | 0.9 | 1 | 1 | 0.056 | 1 | 0.060 | 1 | 0.038 | 1 | 0.056 | 1 | 0.054 | 1 | 0.074 |
| 50 | 0.9 | 0.9 | 1 | 0.145 | 1 | 0.139 | 1 | 0.145 | 0.96 | 0.123 | 0.945 | 0.099 | 0.861 | 0.093 |
| 100 | 0.9 | 0.9 | 1 | 0.338 | 1 | 0.300 | 1 | 0.310 | 1 | 0.208 | 1 | 0.211 | 1 | 0.203 |
| 250 | 0.9 | 0.9 | 1 | 0.968 | 1 | 0.971 | 1 | 0.978 | 1 | 0.840 | 1 | 0.829 | 1 | 0.833 |
| 50 | 0.9 | 0.8 | 1 | 0.367 | 1 | 0.342 | 1 | 0.338 | 0.961 | 0.238 | 0.948 | 0.224 | 0.834 | 0.231 |
| 100 | 0.9 | 0.8 | 1 | 0.852 | 1 | 0.861 | 1 | 0.883 | 1 | 0.67 | 1 | 0.675 | 1 | 0.671 |
| 250 | 0.9 | 0.8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 50 | 0.8 | 1 | 1 | 0.061 | 1 | 0.063 | 1 | 0.049 | 1 | 0.063 | 1 | 0.058 | 1 | 0.057 |
| 100 | 0.8 | 1 | 1 | 0.058 | 1 | 0.067 | 1 | 0.050 | 1 | 0.053 | 1 | 0.061 | 1 | 0.058 |
| 250 | 0.8 | 1 | 1 | 0.048 | 1 | 0.042 | 1 | 0.047 | 1 | 0.065 | 1 | 0.053 | 1 | 0.057 |
| 50 | 0.8 | 0.9 | 1 | 0.126 | 1 | 0.128 | 1 | 0.127 | 1 | 0.100 | 1 | 0.107 | 1 | 0.115 |
| 100 | 0.8 | 0.9 | 1 | 0.366 | 1 | 0.324 | 1 | 0.327 | 1 | 0.193 | 1 | 0.221 | 1 | 0.217 |
| 250 | 0.8 | 0.9 | 1 | 0.974 | 1 | 0.959 | 1 | 0.963 | 1 | 0.858 | 1 | 0.847 | 1 | 0.83 |
| 50 | 0.8 | 0.8 | 1 | 0.386 | 1 | 0.337 | 1 | 0.352 | 1 | 0.227 | 1 | 0.246 | 1 | 0.232 |
| 100 | 0.8 | 0.8 | 1 | 0.861 | 1 | 0.882 | 1 | 0.867 | 1 | 0.68 | 1 | 0.668 | 1 | 0.669 |
| 250 | 0.8 | 0.8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 16: Empirical size and power. Constant case with three common factors ($N = 40$)

| $T$ | $\rho_i$ | $\alpha$ | $\sigma_F^2$ | $Z^e_{\hat{t}_{NT}}$ | $MQ(0)$ | $MQ(1)$ | $MQ(2)$ | $MQ(3)$ |
|---|---|---|---|---|---|---|---|---|
| 50 | 1 | 1 | 0.5 | 0.082 | 0.006 | 0.167 | 0.340 | 0.484 |
| 100 | 1 | 1 | 0.5 | 0.057 | 0.003 | 0.024 | 0.186 | 0.784 |
| 250 | 1 | 1 | 0.5 | 0.050 | 0.001 | 0.02 | 0.128 | 0.848 |
| 50 | 1 | 0.9 | 0.5 | 0.117 | 0.021 | 0.139 | 0.312 | 0.525 |
| 100 | 1 | 0.9 | 0.5 | 0.061 | 0.086 | 0.053 | 0.206 | 0.652 |
| 250 | 1 | 0.9 | 0.5 | 0.051 | 0.771 | 0.017 | 0.075 | 0.134 |
| 50 | 1 | 0.8 | 0.5 | 0.121 | 0.066 | 0.090 | 0.302 | 0.539 |
| 100 | 1 | 0.8 | 0.5 | 0.051 | 0.509 | 0.041 | 0.122 | 0.325 |
| 250 | 1 | 0.8 | 0.5 | 0.061 | 0.986 | 0.007 | 0.003 | 0.001 |
| 50 | 1 | 1 | 1 | 0.061 | 0 | 0.003 | 0.030 | 0.967 |
| 100 | 1 | 1 | 1 | 0.052 | 0 | 0.013 | 0.063 | 0.921 |
| 250 | 1 | 1 | 1 | 0.050 | 0 | 0.010 | 0.078 | 0.909 |
| 50 | 1 | 0.9 | 1 | 0.030 | 0.001 | 0.006 | 0.045 | 0.945 |
| 100 | 1 | 0.9 | 1 | 0.036 | 0.093 | 0.033 | 0.134 | 0.737 |
| 250 | 1 | 0.9 | 1 | 0.034 | 0.844 | 0.008 | 0.041 | 0.104 |
| 50 | 1 | 0.8 | 1 | 0.033 | 0.039 | 0.010 | 0.062 | 0.886 |
| 100 | 1 | 0.8 | 1 | 0.048 | 0.56 | 0.025 | 0.095 | 0.317 |
| 250 | 1 | 0.8 | 1 | 0.052 | 0.994 | 0.001 | 0.001 | 0.001 |
| 50 | 1 | 1 | 10 | 0.060 | 0 | 0.002 | 0.015 | 0.979 |
| 100 | 1 | 1 | 10 | 0.049 | 0.001 | 0.006 | 0.059 | 0.931 |
| 250 | 1 | 1 | 10 | 0.060 | 0.004 | 0.009 | 0.084 | 0.900 |
| 50 | 1 | 0.9 | 10 | 0.044 | 0.008 | 0.001 | 0.027 | 0.957 |
| 100 | 1 | 0.9 | 10 | 0.053 | 0.116 | 0.030 | 0.133 | 0.717 |
| 250 | 1 | 0.9 | 10 | 0.042 | 0.904 | 0.006 | 0.022 | 0.065 |
| 50 | 1 | 0.8 | 10 | 0.030 | 0.034 | 0.012 | 0.059 | 0.886 |
| 100 | 1 | 0.8 | 10 | 0.049 | 0.651 | 0.014 | 0.076 | 0.256 |
| 250 | 1 | 0.8 | 10 | 0.043 | 0.994 | 0.001 | 0.001 | 0.001 |

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 17: Empirical size and power. Constant case with three common factors ($N = 40$)

| $T$ | $\rho_i$ | $\alpha$ | $\sigma_F^2$ | $Z^e_{\hat{t}_{NT}}$ | $MQ(0)$ | $MQ(1)$ | $MQ(2)$ | $MQ(3)$ | $\rho_i$ | $Z^e_{\hat{t}_{NT}}$ | $MQ(0)$ | $MQ(1)$ | $MQ(2)$ | $MQ(3)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.9 | 1 | 0.5 | 0.918 | 0.014 | 0.224 | 0.378 | 0.384 | 0.8 | 0.929 | 0.018 | 0.299 | 0.372 | 0.311 |
| 100 | 0.9 | 1 | 0.5 | 0.988 | 0.004 | 0.049 | 0.249 | 0.698 | 0.8 | 0.989 | 0.007 | 0.071 | 0.282 | 0.640 |
| 250 | 0.9 | 1 | 0.5 | 0.998 | 0.002 | 0.020 | 0.191 | 0.787 | 0.8 | 0.998 | 0.003 | 0.029 | 0.223 | 0.745 |
| 50 | 0.9 | 0.9 | 0.5 | 1 | 0.053 | 0.168 | 0.357 | 0.422 | 0.8 | 1 | 0.058 | 0.224 | 0.351 | 0.367 |
| 100 | 0.9 | 0.9 | 0.5 | 1 | 0.111 | 0.082 | 0.276 | 0.531 | 0.8 | 1 | 0.134 | 0.095 | 0.292 | 0.479 |
| 250 | 0.9 | 0.9 | 0.5 | 1 | 0.946 | 0.008 | 0.028 | 0.018 | 0.8 | 1 | 0.954 | 0.006 | 0.022 | 0.018 |
| 50 | 0.9 | 0.8 | 0.5 | 1 | 0.134 | 0.121 | 0.299 | 0.446 | 0.8 | 1 | 0.155 | 0.145 | 0.305 | 0.395 |
| 100 | 0.9 | 0.8 | 0.5 | 1 | 0.690 | 0.031 | 0.120 | 0.159 | 0.8 | 1 | 0.742 | 0.031 | 0.110 | 0.117 |
| 250 | 0.9 | 0.8 | 0.5 | 1 | 1 | 0 | 0 | 0 | 0.8 | 1 | 1 | 0 | 0 | 0 |
| 50 | 0.9 | 1 | 1 | 1 | 0 | 0.004 | 0.056 | 0.94 | 0.8 | 0.996 | 0 | 0.004 | 0.064 | 0.932 |
| 100 | 0.9 | 1 | 1 | 1 | 0 | 0.007 | 0.098 | 0.895 | 0.8 | 1 | 0 | 0.007 | 0.105 | 0.888 |
| 250 | 0.9 | 1 | 1 | 1 | 0.001 | 0.010 | 0.127 | 0.862 | 0.8 | 1 | 0 | 0.01 | 0.127 | 0.863 |
| 50 | 0.9 | 0.9 | 1 | 1 | 0.004 | 0.008 | 0.071 | 0.917 | 0.8 | 1 | 0.005 | 0.013 | 0.083 | 0.899 |
| 100 | 0.9 | 0.9 | 1 | 1 | 0.087 | 0.057 | 0.226 | 0.630 | 0.8 | 1 | 0.098 | 0.057 | 0.223 | 0.622 |
| 250 | 0.9 | 0.9 | 1 | 1 | 0.935 | 0.007 | 0.032 | 0.026 | 0.8 | 1 | 0.943 | 0.006 | 0.031 | 0.020 |
| 50 | 0.9 | 0.8 | 1 | 1 | 0.036 | 0.025 | 0.123 | 0.816 | 0.8 | 1 | 0.039 | 0.026 | 0.119 | 0.816 |
| 100 | 0.9 | 0.8 | 1 | 1 | 0.693 | 0.031 | 0.109 | 0.167 | 0.8 | 1 | 0.708 | 0.033 | 0.113 | 0.146 |
| 250 | 0.9 | 0.8 | 1 | 1 | 1 | 0 | 0 | 0 | 0.8 | 1 | 1 | 0 | 0 | 0 |
| 50 | 0.9 | 1 | 10 | 0.937 | 0.003 | 0.002 | 0.032 | 0.963 | 0.8 | 0.985 | 0.004 | 0.002 | 0.033 | 0.961 |
| 100 | 0.9 | 1 | 10 | 1 | 0.001 | 0.007 | 0.095 | 0.897 | 0.8 | 1 | 0.001 | 0.008 | 0.098 | 0.893 |
| 250 | 0.9 | 1 | 10 | 1 | 0.001 | 0.009 | 0.116 | 0.874 | 0.8 | 1 | 0 | 0.008 | 0.125 | 0.867 |
| 50 | 0.9 | 0.9 | 10 | 0.936 | 0.008 | 0.008 | 0.058 | 0.926 | 0.8 | 0.983 | 0.005 | 0.011 | 0.06 | 0.924 |
| 100 | 0.9 | 0.9 | 10 | 1 | 0.082 | 0.058 | 0.230 | 0.630 | 0.8 | 1 | 0.091 | 0.055 | 0.225 | 0.629 |
| 250 | 0.9 | 0.9 | 10 | 1 | 0.938 | 0.006 | 0.032 | 0.024 | 0.8 | 1 | 0.942 | 0.007 | 0.031 | 0.020 |
| 50 | 0.9 | 0.8 | 10 | 0.929 | 0.041 | 0.021 | 0.105 | 0.833 | 0.8 | 0.979 | 0.041 | 0.023 | 0.108 | 0.828 |
| 100 | 0.9 | 0.8 | 10 | 1 | 0.698 | 0.031 | 0.117 | 0.154 | 0.8 | 1 | 0.699 | 0.028 | 0.113 | 0.160 |
| 250 | 0.9 | 0.8 | 10 | 1 | 1 | 0 | 0 | 0 | 0.8 | 1 | 1 | 0 | 0 | 0 |

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 18: Empirical size and power. One level shift, known break point ($\lambda_i = 0.5$) and one common factor ($N = 40$)

| $T$ | $\rho_i$ | $\alpha$ | $\sigma_F^2$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $\rho_i$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $\rho_i$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 1 | 1 | 0.5 | 0.050 | 0.058 | 0.9 | 1 | 0.059 | 0.8 | 1 | 0.060 |
| 100 | 1 | 1 | 0.5 | 0.053 | 0.053 | 0.9 | 1 | 0.058 | 0.8 | 1 | 0.055 |
| 250 | 1 | 1 | 0.5 | 0.046 | 0.051 | 0.9 | 1 | 0.051 | 0.8 | 1 | 0.053 |
| 50 | 1 | 0.9 | 0.5 | 0.042 | 0.121 | 0.9 | 1 | 0.128 | 0.8 | 1 | 0.138 |
| 100 | 1 | 0.9 | 0.5 | 0.049 | 0.275 | 0.9 | 1 | 0.324 | 0.8 | 1 | 0.316 |
| 250 | 1 | 0.9 | 0.5 | 0.047 | 0.837 | 0.9 | 1 | 0.948 | 0.8 | 1 | 0.948 |
| 50 | 1 | 0.8 | 0.5 | 0.042 | 0.282 | 0.9 | 1 | 0.303 | 0.8 | 1 | 0.319 |
| 100 | 1 | 0.8 | 0.5 | 0.049 | 0.695 | 0.9 | 1 | 0.782 | 0.8 | 1 | 0.803 |
| 250 | 1 | 0.8 | 0.5 | 0.050 | 0.981 | 0.9 | 1 | 1 | 0.8 | 1 | 1 |
| 50 | 1 | 1 | 1 | 0.041 | 0.057 | 0.9 | 1 | 0.059 | 0.8 | 1 | 0.060 |
| 100 | 1 | 1 | 1 | 0.050 | 0.058 | 0.9 | 1 | 0.053 | 0.8 | 1 | 0.056 |
| 250 | 1 | 1 | 1 | 0.050 | 0.049 | 0.9 | 1 | 0.048 | 0.8 | 1 | 0.053 |
| 50 | 1 | 0.9 | 1 | 0.041 | 0.119 | 0.9 | 1 | 0.137 | 0.8 | 1 | 0.128 |
| 100 | 1 | 0.9 | 1 | 0.054 | 0.292 | 0.9 | 1 | 0.307 | 0.8 | 1 | 0.308 |
| 250 | 1 | 0.9 | 1 | 0.042 | 0.889 | 0.9 | 1 | 0.949 | 0.8 | 1 | 0.953 |
| 50 | 1 | 0.8 | 1 | 0.039 | 0.304 | 0.9 | 1 | 0.310 | 0.8 | 1 | 0.316 |
| 100 | 1 | 0.8 | 1 | 0.048 | 0.748 | 0.9 | 1 | 0.797 | 0.8 | 1 | 0.798 |
| 250 | 1 | 0.8 | 1 | 0.053 | 0.994 | 0.9 | 1 | 1 | 0.8 | 1 | 1 |
| 50 | 1 | 1 | 10 | 0.048 | 0.058 | 0.9 | 1 | 0.060 | 0.8 | 1 | 0.057 |
| 100 | 1 | 1 | 10 | 0.054 | 0.057 | 0.9 | 1 | 0.054 | 0.8 | 1 | 0.052 |
| 250 | 1 | 1 | 10 | 0.053 | 0.045 | 0.9 | 1 | 0.049 | 0.8 | 1 | 0.052 |
| 50 | 1 | 0.9 | 10 | 0.038 | 0.113 | 0.9 | 1 | 0.122 | 0.8 | 1 | 0.130 |
| 100 | 1 | 0.9 | 10 | 0.046 | 0.288 | 0.9 | 1 | 0.287 | 0.8 | 1 | 0.296 |
| 250 | 1 | 0.9 | 10 | 0.049 | 0.941 | 0.9 | 1 | 0.944 | 0.8 | 1 | 0.951 |
| 50 | 1 | 0.8 | 10 | 0.038 | 0.289 | 0.9 | 1 | 0.291 | 0.8 | 1 | 0.290 |
| 100 | 1 | 0.8 | 10 | 0.045 | 0.791 | 0.9 | 1 | 0.790 | 0.8 | 1 | 0.793 |
| 250 | 1 | 0.8 | 10 | 0.044 | 1 | 0.9 | 1 | 0.999 | 0.8 | 1 | 1 |

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 19: Empirical size and power with three common factors. One level shift, known break point ($\lambda = 0.5$, $N = 40$)

| $T$ | $\rho_i$ | $\alpha$ | $\sigma_F^2$ | $Z^e_{\hat{t}_{NT}}$ | $MQ(0)$ | $MQ(1)$ | $MQ(2)$ | $MQ(3)$ |
|---|---|---|---|---|---|---|---|---|
| 50 | 1 | 1 | 0.5 | 0.082 | 0.011 | 0.179 | 0.349 | 0.461 |
| 100 | 1 | 1 | 0.5 | 0.064 | 0.002 | 0.039 | 0.196 | 0.763 |
| 250 | 1 | 1 | 0.5 | 0.063 | 0.001 | 0.013 | 0.130 | 0.856 |
| 50 | 1 | 0.9 | 0.5 | 0.117 | 0.032 | 0.137 | 0.332 | 0.499 |
| 100 | 1 | 0.9 | 0.5 | 0.070 | 0.047 | 0.061 | 0.206 | 0.686 |
| 250 | 1 | 0.9 | 0.5 | 0.052 | 0.653 | 0.079 | 0.117 | 0.151 |
| 50 | 1 | 0.8 | 0.5 | 0.126 | 0.077 | 0.104 | 0.274 | 0.545 |
| 100 | 1 | 0.8 | 0.5 | 0.055 | 0.361 | 0.104 | 0.184 | 0.351 |
| 250 | 1 | 0.8 | 0.5 | 0.054 | 0.930 | 0.066 | 0.004 | 0 |
| 50 | 1 | 1 | 1 | 0.050 | 0 | 0.001 | 0.034 | 0.965 |
| 100 | 1 | 1 | 1 | 0.056 | 0.001 | 0.004 | 0.066 | 0.929 |
| 250 | 1 | 1 | 1 | 0.051 | 0.001 | 0.009 | 0.092 | 0.898 |
| 50 | 1 | 0.9 | 1 | 0.039 | 0.002 | 0.006 | 0.052 | 0.940 |
| 100 | 1 | 0.9 | 1 | 0.042 | 0.039 | 0.042 | 0.157 | 0.762 |
| 250 | 1 | 0.9 | 1 | 0.042 | 0.770 | 0.038 | 0.089 | 0.103 |
| 50 | 1 | 0.8 | 1 | 0.034 | 0.014 | 0.015 | 0.071 | 0.900 |
| 100 | 1 | 0.8 | 1 | 0.036 | 0.408 | 0.080 | 0.179 | 0.333 |
| 250 | 1 | 0.8 | 1 | 0.047 | 0.989 | 0.011 | 0 | 0 |
| 50 | 1 | 1 | 10 | 0.054 | 0.001 | 0.001 | 0.020 | 0.976 |
| 100 | 1 | 1 | 10 | 0.054 | 0 | 0.004 | 0.060 | 0.935 |
| 250 | 1 | 1 | 10 | 0.051 | 0 | 0.009 | 0.093 | 0.898 |
| 50 | 1 | 0.9 | 10 | 0.038 | 0.003 | 0.005 | 0.038 | 0.950 |
| 100 | 1 | 0.9 | 10 | 0.046 | 0.047 | 0.046 | 0.166 | 0.74 |
| 250 | 1 | 0.9 | 10 | 0.050 | 0.855 | 0.019 | 0.055 | 0.071 |
| 50 | 1 | 0.8 | 10 | 0.032 | 0.013 | 0.013 | 0.071 | 0.896 |
| 100 | 1 | 0.8 | 10 | 0.044 | 0.486 | 0.070 | 0.152 | 0.291 |
| 250 | 1 | 0.8 | 10 | 0.048 | 1 | 0 | 0 | 0 |

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 20: Empirical size and power with three common factors. One level shift, known break point ($\lambda = 0.5$, $N = 40$)

| $T$ | $\rho_i$ | $\alpha$ | $\sigma_F^2$ | $Z^e_{\hat{t}_{NT}}$ | $MQ(0)$ | $MQ(1)$ | $MQ(2)$ | $MQ(3)$ | $\rho_i$ | $Z^e_{\hat{t}_{NT}}$ | $MQ(0)$ | $MQ(1)$ | $MQ(2)$ | $MQ(3)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.9 | 1 | 0.5 | 0.908 | 0.013 | 0.229 | 0.361 | 0.397 | 0.8 | 0.926 | 0.017 | 0.274 | 0.388 | 0.321 |
| 100 | 0.9 | 1 | 0.5 | 0.989 | 0.003 | 0.039 | 0.209 | 0.749 | 0.8 | 0.991 | 0.004 | 0.066 | 0.266 | 0.664 |
| 250 | 0.9 | 1 | 0.5 | 0.998 | 0 | 0.014 | 0.119 | 0.867 | 0.8 | 0.997 | 0.001 | 0.021 | 0.166 | 0.812 |
| 50 | 0.9 | 0.9 | 0.5 | 1 | 0.041 | 0.184 | 0.333 | 0.442 | 0.8 | 1 | 0.048 | 0.218 | 0.367 | 0.367 |
| 100 | 0.9 | 0.9 | 0.5 | 1 | 0.060 | 0.062 | 0.245 | 0.633 | 0.8 | 1 | 0.073 | 0.071 | 0.261 | 0.595 |
| 250 | 0.9 | 0.9 | 0.5 | 1 | 0.729 | 0.026 | 0.084 | 0.161 | 0.8 | 1 | 0.108 | 0.163 | 0.325 | 0.404 |
| 50 | 0.9 | 0.8 | 0.5 | 1 | 0.088 | 0.13 | 0.306 | 0.476 | 0.8 | 1 | 0.509 | 0.058 | 0.175 | 0.258 |
| 100 | 0.9 | 0.8 | 0.5 | 1 | 0.467 | 0.06 | 0.169 | 0.304 | 0.8 | 1 | 0.996 | 0.004 | 0 | 0 |
| 250 | 0.9 | 0.8 | 0.5 | 1 | 0.996 | 0.004 | 0 | 0 | 0.8 | 0.996 | 0.001 | 0.004 | 0.057 | 0.938 |
| 50 | 0.9 | 1 | 1 | 0.996 | 0 | 0.005 | 0.051 | 0.944 | 0.8 | 1 | 0.001 | 0.003 | 0.058 | 0.938 |
| 100 | 0.9 | 1 | 1 | 1 | 0 | 0.004 | 0.056 | 0.94 | 0.8 | 1 | 0 | 0.003 | 0.058 | 0.939 |
| 250 | 0.9 | 1 | 1 | 1 | 0.001 | 0.002 | 0.052 | 0.945 | 0.8 | 1 | 0.003 | 0.007 | 0.074 | 0.916 |
| 50 | 0.9 | 0.9 | 1 | 1 | 0.002 | 0.008 | 0.063 | 0.927 | 0.8 | 1 | 0.043 | 0.031 | 0.151 | 0.775 |
| 100 | 0.9 | 0.9 | 1 | 1 | 0.038 | 0.03 | 0.155 | 0.777 | 0.8 | 1 | 0.727 | 0.033 | 0.087 | 0.153 |
| 250 | 0.9 | 0.9 | 1 | 1 | 0.722 | 0.031 | 0.09 | 0.157 | 0.8 | 1 | 0.021 | 0.018 | 0.099 | 0.862 |
| 50 | 0.9 | 0.8 | 1 | 1 | 0.020 | 0.019 | 0.097 | 0.864 | 0.8 | 1 | 0.466 | 0.054 | 0.168 | 0.312 |
| 100 | 0.9 | 0.8 | 1 | 1 | 0.437 | 0.055 | 0.159 | 0.349 | 0.8 | 1 | 0.997 | 0.003 | 0 | 0 |
| 250 | 0.9 | 0.8 | 1 | 1 | 0.996 | 0.004 | 0 | 0 | 0.8 | 0.977 | 0 | 0.002 | 0.024 | 0.971 |
| 50 | 0.9 | 1 | 10 | 0.911 | 0 | 0.001 | 0.025 | 0.971 | 0.8 | 1 | 0 | 0.005 | 0.050 | 0.944 |
| 100 | 0.9 | 1 | 10 | 0.999 | 0 | 0.003 | 0.056 | 0.94 | 0.8 | 1 | 0 | 0.005 | 0.051 | 0.944 |
| 250 | 0.9 | 1 | 10 | 1 | 0.001 | 0.005 | 0.048 | 0.946 | 0.8 | 0.976 | 0.002 | 0.005 | 0.053 | 0.936 |
| 50 | 0.9 | 0.9 | 10 | 0.912 | 0.002 | 0.006 | 0.05 | 0.936 | 0.8 | 1 | 0.039 | 0.034 | 0.153 | 0.774 |
| 100 | 0.9 | 0.9 | 10 | 1 | 0.035 | 0.029 | 0.155 | 0.78 | 0.8 | 1 | 0.724 | 0.024 | 0.085 | 0.167 |
| 250 | 0.9 | 0.9 | 10 | 1 | 0.718 | 0.026 | 0.093 | 0.163 | 0.8 | 0.974 | 0.021 | 0.015 | 0.086 | 0.875 |
| 50 | 0.9 | 0.8 | 10 | 0.898 | 0.018 | 0.015 | 0.096 | 0.865 | 0.8 | 1 | 0.460 | 0.055 | 0.161 | 0.323 |
| 100 | 0.9 | 0.8 | 10 | 0.999 | 0.461 | 0.054 | 0.16 | 0.324 | 0.8 | 1 | 0.998 | 0.002 | 0 | 0 |
| 250 | 0.9 | 0.8 | 10 | 1 | 0.995 | 0.005 | 0 | 0 | 0.8 | 1 | 0.998 | 0.002 | 0 | 0 |

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.

Table 21: Empirical size and power. One level shift, unknown break point ($\lambda_i = 0.5$) and one common factor ($N = 40$)

| $T$ | $\rho_i$ | $\alpha$ | $\sigma_F^2$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $\rho_i$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ | $\rho_i$ | $Z^e_{\hat{t}_{NT}}$ | $ADF^d_{\hat{F}}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | 1 | 1 | 0.5 | 0.039 | 0.045 | 0.9 | 1 | 0.030 | 0.8 | 1 | 0.044 |
| 100 | 1 | 1 | 0.5 | 0.047 | 0.052 | 0.9 | 1 | 0.038 | 0.8 | 1 | 0.044 |
| 250 | 1 | 1 | 0.5 | 0.049 | 0.062 | 0.9 | 1 | 0.043 | 0.8 | 1 | 0.051 |
| 50 | 1 | 0.9 | 0.5 | 0.050 | 0.111 | 0.9 | 1 | 0.118 | 0.8 | 1 | 0.206 |
| 100 | 1 | 0.9 | 0.5 | 0.041 | 0.252 | 0.9 | 1 | 0.479 | 0.8 | 1 | 0.608 |
| 250 | 1 | 0.9 | 0.5 | 0.041 | 0.843 | 0.9 | 1 | 0.998 | 0.8 | 1 | 0.999 |
| 50 | 1 | 0.8 | 0.5 | 0.048 | 0.277 | 0.9 | 1 | 0.058 | 0.8 | 1 | 0.097 |
| 100 | 1 | 0.8 | 0.5 | 0.046 | 0.710 | 0.9 | 1 | 0.215 | 0.8 | 1 | 0.283 |
| 250 | 1 | 0.8 | 0.5 | 0.054 | 0.979 | 0.9 | 1 | 0.960 | 0.8 | 1 | 0.937 |
| 50 | 1 | 1 | 1 | 0.055 | 0.064 | 0.9 | 1 | 0.021 | 0.8 | 1 | 0.035 |
| 100 | 1 | 1 | 1 | 0.056 | 0.057 | 0.9 | 1 | 0.024 | 0.8 | 1 | 0.034 |
| 250 | 1 | 1 | 1 | 0.052 | 0.053 | 0.9 | 1 | 0.056 | 0.8 | 1 | 0.055 |
| 50 | 1 | 0.9 | 1 | 0.040 | 0.132 | 0.9 | 1 | 0.154 | 0.8 | 1 | 0.195 |
| 100 | 1 | 0.9 | 1 | 0.053 | 0.279 | 0.9 | 1 | 0.307 | 0.8 | 1 | 0.450 |
| 250 | 1 | 0.9 | 1 | 0.045 | 0.898 | 0.9 | 1 | 0.987 | 0.8 | 1 | 0.990 |
| 50 | 1 | 0.8 | 1 | 0.035 | 0.271 | 0.9 | 1 | 0.054 | 0.8 | 1 | 0.082 |
| 100 | 1 | 0.8 | 1 | 0.051 | 0.745 | 0.9 | 1 | 0.141 | 0.8 | 1 | 0.194 |
| 250 | 1 | 0.8 | 1 | 0.030 | 0.994 | 0.9 | 1 | 0.942 | 0.8 | 1 | 0.936 |
| 50 | 1 | 1 | 10 | 0.069 | 0.060 | 0.9 | 1 | 0.016 | 0.8 | 1 | 0.022 |
| 100 | 1 | 1 | 10 | 0.053 | 0.047 | 0.9 | 1 | 0.006 | 0.8 | 1 | 0.023 |
| 250 | 1 | 1 | 10 | 0.054 | 0.057 | 0.9 | 1 | 0.051 | 0.8 | 1 | 0.061 |
| 50 | 1 | 0.9 | 10 | 0.046 | 0.134 | 0.9 | 1 | 0.076 | 0.8 | 1 | 0.091 |
| 100 | 1 | 0.9 | 10 | 0.056 | 0.278 | 0.9 | 1 | 0.217 | 0.8 | 1 | 0.302 |
| 250 | 1 | 0.9 | 10 | 0.051 | 0.936 | 0.9 | 1 | 0.981 | 0.8 | 1 | 0.995 |
| 50 | 1 | 0.8 | 10 | 0.043 | 0.266 | 0.9 | 1 | 0.032 | 0.8 | 1 | 0.049 |
| 100 | 1 | 0.8 | 10 | 0.054 | 0.750 | 0.9 | 1 | 0.096 | 0.8 | 1 | 0.122 |
| 250 | 1 | 0.8 | 10 | 0.056 | 0.999 | 0.9 | 1 | 0.917 | 0.8 | 1 | 0.936 |

The nominal size is set at the 5% level. Simulation results based on 5,000 replications.