CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Testing for Breaks in Cointegrated Panels with Common and Idiosyncratic Stochastic Trends

Chihwa Kao, Lorenzo Trapani and Giovanni Urga

CEA@Cass Working Paper Series

WP–CEA–02-2011

Testing for Breaks in Cointegrated Panels with Common and Idiosyncratic Stochastic Trends

Chihwa Kao
Syracuse University (USA)

Lorenzo Trapani
Cass Business School, London (UK)

Giovanni Urga*
Cass Business School, London (UK) and Bergamo University (Italy)

February 8, 2011

Abstract

In this paper, we develop tests for structural change in cointegrated panel regressions with common and idiosyncratic trends. We consider both the cases of observable and non-observable common trends, deriving a Functional Central Limit Theorem for the partial sample estimators under the null of no break. We show that test based on sup-Wald statistics are powerful versus breaks of size $O_p\left(1/\sqrt{n}T\right)$, also proving that power is present when the time of change differs across units and when only some units have a break. Our framework is extended to the case of cross correlated regressors and endogeneity. Monte Carlo evidence shows that the tests have the correct size and good power properties.

JEL codes: C23.

Keywords: Structural change, Panel cointegration, Common stochastic trends, Functional Central Limit Theorem.

*Corresponding author: g.urga@city.ac.uk. Centre for Econometric Analysis, Faculty of Finance, Cass Business School, 106, Bunhill Row, London, EC1Y 8TZ (U.K.) Tel. 00.44.(0)20.70408698; Fax. 00.44.(0)20.70408881; http://www.cass.city.ac.uk/faculty/g.urga http://www.cass.city.ac.uk/cea/index.html

1 Introduction

Since the seminal contributions by Perron (1989) and Rappoport and Reichlin (1989), the literature has produced a comprehensive set of results on the changepoint problem in a time series framework - we refer, inter alia, to the articles by Andrews (1993), Andrews and Ploberger (1994), Bai and Perron (1998), and Kejriwal and Perron (2008, 2010). Useful surveys can be found in Banerjee and Urga (2005) and Perron (2006). When extending the framework to a multivariate setting, the literature has shown that the cross sectional dimension can lead to better inference; for example, Bai, Lumsdaine and Stock (1998) show that the estimation of the changepoint in a VAR improves with the dimension of the VAR, due to the presence of cross sectional information. As pointed out by Qu and Perron (2007), a crucial condition is having nonzero correlations across equations, even when including equations without breaks.

Thus, a natural development to enhance the power of tests for structural breaks is to use panel data models, especially when cross sectional dependence is present. Despite the potential usefulness, the inferential theory on structural changes in panels is still underdeveloped. There are a few exceptions: Feng, Kao and Lazarova (2008) and Bai (2010) propose procedures for dating breaks in simple settings with no cross sectional dependence amongst units; Kim (2010a, 2010b) investigates the estimation of change points in panel time trend models with cross-sectional dependence; Breitung and Eickmeier (2010) propose a test for changes in the loadings of a panel factor model.

This paper fills the gap in the literature by proposing an estimation and testing framework for slope parameter instability in cointegrated panel regression; strong cross-sectional dependence is allowed for through the presence of common stochastic trends.

Basic model and extensions

We study a cointegrated panel with unit-specific variables (idiosyncratic shocks) and a set of possibly unobservable variables that are common across all units (common shocks):

$$y_{it} = \alpha_i + \beta_t' F_t + \gamma_t' x_{it} + u_{it}, \tag{1}$$

where $i = 1, ..., n$ and $t = 1, ..., T$ and β_t and γ_t are $R \times 1$ and $p \times 1$ respectively. We assume that (1) is a cointegrating regression for all units i, so that the vector $[F_t', x_{it}']'$ is $I(1)$ and the error term u_{it} is stationary for all is. The presence of (strong) cross-sectional dependence arises directly from the common shocks F_t; due to the slope β being homogeneous across units, the term $\beta' F_t$ represents an R-dimensional stochastic time effect. As well as having cross dependence due to the common shock F_t, and as well as having homogeneous response to common shocks, we also consider dependence and heterogeneity in the x_{it}s as

$$x_{it} = \Gamma_i G_t + w_{it}, \tag{2}$$

where G_t is a set of common factors that can be independent of the regressors F_t or overlap with them, and w_{it} is a unit specific (stationary or nonstationary) shock - see also Pesaran (2006) and Kapetanios, Pesaran and Yamagata (2011) for a similar (in spirit) framework. From (2), cross-dependence is accounted for directly (via F_t) and indirectly (via the factor structure in x_{it}). Heterogeneity in the response to common shocks is also allowed for through the possibly heterogeneous loadings Γ_i; also, the response to F_t is allowed to be (indirectly) heterogeneous across individuals if G_t contains F_t.

Model (1) encompasses a wide set of models in economics and finance which may be subject to breaks. Such a model may represent a situation whereby the decision y_{it} of microeconomic agent i is influenced by macroeconomic factors F_t and by a set of individual specific characteristics, α_i and x_{it}. Examples that have been studied in the literature include, inter alia: demand for household food consumption (see e.g., Dynarski and Sheffrin, 1985, where households are assumed to have the same elasticity to food price, which is the common shock, and to permanent income, which is the idiosyncratic variable); firm size evolving according to a random walk, a case known in the literature as Gibrat's law (see Sutton, 1997; Geroski et al, 2002); other examples can also be found in micro demand for investment, consumption, labour demand. Moreover, the forward rate unbiasedness hypothesis postulates that the forward rate is an unbiased predictor of the corresponding future spot rate. This hypothesis has been extensively tested for exchange rates (Baillie and Bollerslev, 1989; Liu and Maynard, 2005; Westerlund, 2007). Another example in finance are models for default intensity for firm i at time t expressed as function of common factors (such as U.S. 3-month T-bill and the trailing 1-year returns) and idiosyncratic covariates such as distance to default and trailing 1-year stock return of the firm i (see Das et al., 2007). Relevant is also the literature on output convergence where output for country i at time t depends on a set of common, to all n countries, technological shocks/knowledge and heterogenous degrees of access to the technological knowledge (Pesaran, 2007; Phillips and Sul, 2007).

Finally, most of the results in this paper are derived assuming zero long run correlation between $[F_t', x_{it}']'$ and u_{it}. However, we show that our framework can be accommodated to allow for endogeneity. Whilst this involves modifying the estimation technique, i.e., from ordinary least squares (OLS) to fully-modified OLS (FMOLS), the limiting distribution of the test and the power versus local alternatives remain unaltered. Other estimation techniques have been proposed in similar contexts - see e.g. Bai, Kao and Ng (2009).

Main results of this paper

We focus our attention on testing for the constancy over time of $\theta_t = \left(\beta_t', \gamma_t'\right)'$, thus developing tests for changes in the cointegration relationship between y_{it} and $(x_{it}', F_t')'$. Considering, for simplicity, the alternative of only one abrupt change at (unknown) time $\lfloor Tr \rfloor$, the null is H_0 :

$\theta_t = \theta$ for all t, whereas the alternative could be defined as

$$H_a : \theta_t = \begin{cases} \theta_1 & \text{for } t = 1, ..., \lfloor Tr \rfloor \\ \theta_2 & \text{for } t = \lfloor Tr \rfloor + 1, ..., T \end{cases},$$

with $\theta_1 \neq \theta_2$.

This paper makes two contributions to the existing literature. First, we develop a Functional Central Limit Theorem (FCLT) for the partial sample estimators of θ, considering both the cases of observed and unobserved common shocks. Results are extended to the case of endogeneity by proving an FCLT for the partial sample FMOLS estimators. These results are of independent interest: ordinary large panels asymptotic theory (Phillips and Moon, 1999; Kao, 1999) cannot be applied in our framework due to the strong cross-sectional dependence introduced by the common shocks. Second, we show that tests based on Wald-type statistics (Andrews, 1993; Andrews and Ploberger, 1994) have nontrivial power versus local alternatives of order $O_p\left(1/\sqrt{n}T\right)$, which provides further justification towards the use of panel models in order to enhance the power of tests. Although the tests are constructed under the alternative H_a of an abrupt and common change, we show that they have power versus other classes of alternatives, e.g. smooth parameter changes. Also, we prove that our tests, albeit designed for the common changepoint alternative H_a, have nontrivial power versus alternatives where series have a break at potentially different points in time. This is a desirable property, since a break could be induced by a change common to all units, but each unit could have different levels of hysteresis and therefore respond with different lag. Also, we study the presence of power when only some units (say m_n) are subject to a change. We show that, in the extreme case of m_n finite, tests have power versus local alternatives shrinking as $O_p\left(\frac{\sqrt{n}}{T}\right)$: when the panel contains many units that do not have a break, there is a loss of power with respect to the case of testing for one unit at a time.

The remainder of the paper is organized as follows. Section 2 introduces the model. Section 3 discusses the asymptotics; test statistics and their null distribution are in Section 4 defines the test statistic and it discusses the local power; extensions (including FMOLS) are in Section 5. In Section 6 we report the finite sample properties, i.e., size and power, of our proposed tests. Section 7 provides concluding remarks. Appendix A contains some preliminary lemmas; the proofs of the main results (test distribution under the null and under local-to-null alternatives) are in Appendix B.

NOTATION. We write integrals involving Brownian motions such as, e.g., $\int_0^1 B(s)ds$ as $\int B$ when there is no ambiguity over limits. We define $\Omega^{1/2}$ to be any matrix such that $\Omega = \left(\Omega^{1/2}\right)\left(\Omega^{1/2}\right)'$. We use $\|\cdot\|$ to denote the Euclidean norm of a vector, $\overset{a.s.}{=}$ to denote almost surely equality, $\rightarrow$ to denote the ordinary limit, $\xrightarrow{d}$ to denote convergence in distribution, $\xrightarrow{p}$ to denote convergence in probability, $\lfloor x \rfloor$ to denote the integer part, $B = B(\Omega)$ to denote Brownian motion with covariance matrix Ω, and $\bar{B} = B - \int B$ to denote the demeaned version of B. We let $M < \infty$ be a generic positive number which does not depend on n or T.

2 Model and assumptions

Consider the panel model with common and idiosyncratic shocks

$$y_{it} = \alpha_i + \beta_t' F_t + \gamma_t' x_{it} + u_{it}, \tag{3}$$

with $i = 1, ..., n$ and $t = 1, ..., T$. We let $F_t = (F_{1t}, ..., F_{Rt})'$ be a $R \times 1$ vector of common stochastic trends

$$F_t = F_{t-1} + \varepsilon_t, \tag{4}$$

x_{it} a $p \times 1$ vector of observable individual-specific regressors,

$$x_{it} = x_{it-1} + \epsilon_{it}, \tag{5}$$

and $[u_{it}, \varepsilon_t', \epsilon_{it}']'$ the error terms.

When common shocks F_t are not observable in (3), F_t can be estimated by a set of observable exogenous variables, z_{it}, defined up to a factor-loading specification as

$$z_{it} = \lambda_i' F_t + e_{it}, \tag{6}$$

where λ_i is a vector of factor loadings and e_{it} is the error term. Whilst results are presented under the simplifying assumption that R is known, the number of common shocks R could be estimated using e.g. the criteria discussed in Bai (2004). Also, since panel z_{it} is employed solely for the estimation of F_t, whilst panel y_{it} is used to estimate β, the number of cross sectional units in the two panels need not be the same. Similarly, the cross sectional index i needs not refer to the same units: for example, in panel y_{it}, i could index individuals, but as far as z_{it} is concerned, i could index different macro variables such as in Stock and Watson (1999, 2002, 2005).

Consider the following assumptions:

Assumption 1: Let $\omega_{it} = [u_{it}, \varepsilon_t', \epsilon_{it}', e_{it}]'$. We assume that *(a)* ω_{it} is a linear process across t with $E\|\omega_{it}\|^{4+\delta} < \infty$ for some $\delta > 0$, and a Beveridge-Nelson decomposition exists such that

$$w_t \overset{a.s.}{=} w_t^* + R_{wT}, \tag{7}$$

$$\widetilde{x}_{it} \overset{a.s.}{=} \widetilde{x}_{it}^* + R_{xiT}, \tag{8}$$

$$u_{it} \overset{a.s.}{=} u_{it}^* + R_{uiT}, \tag{9}$$

$$e_{it} \overset{a.s.}{=} e_{it}^* + R_{eiT}, \tag{10}$$

where w_t^* and $\widetilde{x}_{it}^*$ are two random walks with long run covariance matrices Ω_ε and Ω_ϵ respectively, and u_{it}^* and e_{it}^* are *i.i.d.* processes with variances σ_u^2, and σ_e^2 respectively. Also,

$$\sup_i |R_{jiT}| = O_p\left(\frac{1}{\sqrt{T}}\right), \tag{11}$$

for $j \in \{w, x, u, e\}$; *(b)* for a given t, $\{u_{it}\}$, $\{\epsilon_{it}\}$, and $\{e_{it}\}$ are mutually independent across i and independent of $\{\varepsilon_t\}$; *(c)* $\{x_{it}, F_t\}$ are not cointegrated and Ω_ε and Ω_ϵ are non singular; *(d)* the eigenvalues of Ω_ε and the random matrix $\int B_\varepsilon B_\varepsilon'$ are distinct with probability 1.

Assumption 2: $\|\lambda_i\| \leq M$ and $\frac{1}{n}\sum_{i=1}^n \lambda_i \lambda_i' \to \Sigma_\Lambda$ as $n \to \infty$, where Σ_Λ is non singular.

Assumption 1(a) is a standard requirement on the amount of serial dependence allowed for, and it enables the asymptotic theory developed by Phillips and Solo (1992) and Phillips and Moon (1999) to hold. The only requirement which is nonstandard is (11), which is needed as a sufficient condition in order to prove a panel functional central limit theorem (FCLT). Assumption 1(b) considers a framework where: *(a)* regressors are strictly exogenous and *(b)* no cross sectional dependence is allowed other than the one determined by the presence of the common regressors F_t. Particularly, 1(b) also rules out the presence of cross-sectional dependence among the idiosyncratic shocks x_{it}s. As discussed in the introduction, these restrictions are considered only for the purpose of simplicity of the exposition. We show that the main results of the paper (null distribution and local power of the test) still hold after relaxing each of these assumptions. We discuss the presence of cross-dependence among the idiosyncratic explanatory variables x_{it}s and the presence of endogeneity in Section 5. The requirement that $\{e_{it}\}$ is independent of the other innovations considered in 1(b) is needed for the inference on the F_ts when they are not observable and it is a standard assumption - see Assumption D in Bai (2004, p. 141). Assumption 1(c) rules out cointegration among regressors. This too is a standard requirement in cointegration analysis - see e.g., Park and Phillips (1988) for discussion. Assumption 1(d) is a standard requirement in large panel factor literature, and it is needed in order to identify the factors F_ts in (6) when they are not observable.

Assumption 2 is also standard and it ensures that each factor has a nontrivial contribution towards y_{it}.

Let $\widetilde{x}_{it} = x_{it} - \frac{1}{T}\sum_{t=1}^T x_{it}$ and $w_t = F_t - \frac{1}{T}\sum_{t=1}^T F_t$. The following proposition is important for developing the asymptotics in this paper.

Proposition 1 *Let Assumption 1 hold. As* $(n, T) \to \infty$, *for all* $r \in (0, 1)$

$$\frac{1}{\sqrt{n}T^2}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} w_t \widetilde{x}_{it}' = O_p(1).$$

Proposition 1 states that the asymptotic magnitude of the cross term $\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} w_t \widetilde{x}_{it}'$ is $O_p\left(\sqrt{n}T^2\right)$. The $\sqrt{n}$-convergence is achieved since a CLT holds for the cross-sectional average

$\sum_{i=1}^{n} \widetilde{x}_{it}'$, and this holds since the $\widetilde{x}_{it}$s are *i.i.d.* among i by Assumption 1(b) and because they have zero mean by construction.

We now turn to partial sample estimation of θ, studying its asymptotics under the null of no structural change.

3 Partial sample estimation

This section contains the FCLT for the partial sample estimator (PSE) of θ in (3).

When F_t is unobservable, we propose a two step approach. First, we derive an estimate of F_t, say $\widehat{F}_t$, using (6). Second, we estimate θ using OLS in (3) replacing F_t with $\widehat{F}_t$. The estimator $\widehat{F}_t$ can be obtained applying principal components to the z_{it}; we refer to Bai (2004). It is well known that F_t is identifiable only up to a transformation, say $H'F_t$ where H is an $R \times R$ matrix. Thus, using $\widehat{F}_t$ in (3) allows to estimate $H^{-1}\beta$ rather than β. However, as far as testing is concerned, knowledge of $H'F_t$ and of $H^{-1}\beta$ is the same as estimating F_t and β. Hence, for simplicity, we assume H being a $R \times R$ identity matrix in this paper.

Let $\hat{w}_{t,r} = \widehat{F}_t - \frac{1}{Tr}\sum_{t=1}^{\lfloor Tr \rfloor} \widehat{F}_t$ and $\widehat{W}_{it,r} = \left[\hat{w}_{t,r}', \widetilde{x}_{it}'\right]'$; define, similarly, $\hat{w}_{t,1-r} = \widehat{F}_t - \frac{1}{T(1-r)}\sum_{t=\lfloor Tr \rfloor+1}^{T} \widehat{F}_t$ and $\widehat{W}_{it,1-r} = \left[\hat{w}_{t,1-r}', \widetilde{x}_{it}'\right]'$. Also, consider

$$y_{it} = \alpha_i + \beta_t'\widehat{F}_t + \gamma_t' x_{it} + v_{it}, \tag{12}$$

where $v_{it} = u_{it} + \beta'\left(F_t - \widehat{F}_t\right)$. We define the PSE as

$$\hat{\theta}_{1\lfloor Tr \rfloor} = \left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widehat{W}_{it,r}\widehat{W}_{it,r}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widehat{W}_{it,r}y_{it}\right], \tag{13}$$

$$\hat{\theta}_{2\lfloor Tr \rfloor} = \left[\sum_{i=1}^{n}\sum_{t=\lfloor Tr \rfloor+1}^{T} \widehat{W}_{it,1-r}\widehat{W}_{it,1-r}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=\lfloor Tr \rfloor+1}^{T} \widehat{W}_{it,1-r}y_{it}\right], \tag{14}$$

for some $r \in (0,1)$. Under the null

$$\begin{aligned}
\hat{\theta}_{1\lfloor Tr \rfloor} - \theta &= \begin{bmatrix} \hat{\beta}_{1\lfloor Tr \rfloor} - \beta \\ \hat{\gamma}_{1\lfloor Tr \rfloor} - \gamma \end{bmatrix} \\
&= \left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widehat{W}_{it,r}\widehat{W}_{it,r}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widehat{W}_{it,r}v_{it}\right] \\
&= \begin{bmatrix} \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_{t,r}\hat{w}_{t,r}' & \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widetilde{x}_{it}\hat{w}_{t,r}' \\ \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_{t,r}\widetilde{x}_{it}' & \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widetilde{x}_{it}\widetilde{x}_{it}' \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_{t,r}v_{it} \\ \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widetilde{x}_{it}v_{it} \end{bmatrix},
\end{aligned}$$

and a similar expression can be derived for $\hat{\theta}_{2\lfloor Tr \rfloor} - \theta$.

Before presenting the main result for the asymptotics of $\hat{\theta}_{1\lfloor Tr\rfloor}$ and $\hat{\theta}_{2\lfloor Tr\rfloor}$, we introduce some notation; let

$$\sigma_\zeta^2 = \sigma_u^2 + \sigma_\Pi^2, \tag{15}$$

$$\sigma_\Pi^2 = \sigma_e^2 \left(\beta' \tilde{Q}_B \Sigma_\Lambda \tilde{Q}_B' \beta\right), \tag{16}$$

and $\frac{1}{T^2}\sum_{t=1}^T \hat{w}_t w_t' \xrightarrow{d} \tilde{Q}_B$.

It holds that

Theorem 1 *Let Assumptions 1-2 hold. Then, as $(n,T) \longrightarrow \infty$ with $\frac{n}{T} \rightarrow 0$*

$$\begin{aligned}
\sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr\rfloor} - \theta\right] &= \sqrt{n}T\begin{bmatrix} \hat{\beta}_{1\lfloor Tr\rfloor} - \beta \\ \hat{\gamma}_{1\lfloor Tr\rfloor} - \gamma \end{bmatrix} \xrightarrow{d} \\
&\begin{bmatrix} \sigma_\zeta\left(\int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} & 0_{R\times p} \\ 0_{p\times R} & \frac{6}{r^2}\sigma_u \Omega_\epsilon^{-1/2} \end{bmatrix} \times \begin{bmatrix} Z_1 \\ \frac{1}{\sqrt{6}} C(r;0) \end{bmatrix},
\end{aligned} \tag{17}$$

$$\begin{aligned}
\sqrt{n}T\left[\hat{\theta}_{2\lfloor Tr\rfloor} - \theta\right] &= \sqrt{n}T\begin{bmatrix} \hat{\beta}_{2\lfloor Tr\rfloor} - \beta \\ \hat{\gamma}_{2\lfloor Tr\rfloor} - \gamma \end{bmatrix} \xrightarrow{d} \\
&\begin{bmatrix} \sigma_\zeta\left(\int_r^1 \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} & 0_{R\times p} \\ 0_{p\times R} & \frac{6}{(1-r)^2}\sigma_u \Omega_\epsilon^{-1/2} \end{bmatrix} \times \begin{bmatrix} Z_2 \\ \frac{1}{\sqrt{6}} C(1;r) \end{bmatrix},
\end{aligned} \tag{18}$$

uniformly in r, where $\bar{B}_\varepsilon(r)$ is a standard, demeaned R-dimensional Brownian motion, Z_1 and Z_2 are two independent R-dimensional Gaussian random variables and the stochastic process $\frac{1}{\sqrt{6}}C(r;s)$ is Gaussian, independent of $\bar{B}_\varepsilon(r)$, with independent increments, mean zero and covariance kernel given in (36).

Remarks

1.1 Theorem 1 is the building block to implement sequential testing for breaks. The process $C(r;0)$ is, in essence, a variance transformed Brownian motion. As far as $\hat{\beta}_{1\lfloor Tr\rfloor} - \beta$ is concerned, it is $\sqrt{n}T$ consistent, as in Phillips and Moon (1999) and Kao (1999). However, the limiting distribution is not normal, contrary to what typically found in cointegrated panels, but mixed normal. This is due to the shocks w_t being nonstationary and common across units, which results in the covariance matrix $\left(nT^2\right)^{-1}\sum_{i=1}^n\sum_{t=1}^T w_t w_t'$ converging to the random matrix $\int \bar{B}_\varepsilon \bar{B}_\varepsilon'$ rather than to a matrix of constants as in standard panel cointegration - see the proof of Lemma A.3.

1.2 Equation (36) in Lemma A.3 in Appendix describes the covariance structure of $C(r;s)$; note that $E\left[C(r;0)\right]^2 = \frac{1}{6}r^2$, and $E\left[C(1;r)\right]^2 = \frac{1}{6}(1-r)^2$, which is essentially the same result as Lemma A.4 in Chiang et al. (2002). The covariance structure of the process is $E\left[C(r;0)\,C(s;0)\right] = \frac{1}{6}(r \vee s)^2$.

1.3 The estimates $\hat{\beta}_{1\lfloor Tr\rfloor} - \beta$ and $\hat{\gamma}_{1\lfloor Tr\rfloor} - \gamma$ are asymptotically independent. This result is a consequence of Proposition 1, whereby

$$\frac{1}{nT^2}\left[\begin{array}{cc} \sum_{i=1}^{n}\sum_{t=1}^{T} w_t w_t' & \sum_{i=1}^{n}\sum_{t=1}^{T} \widetilde{x}_{it} w_t' \\ \sum_{i=1}^{n}\sum_{t=1}^{T} w_t \widetilde{x}_{it}' & \sum_{i=1}^{n}\sum_{t=1}^{T} \widetilde{x}_{it}\widetilde{x}_{it}' \end{array}\right] = \left[\begin{array}{cc} O_p(1) & O_p\left(\frac{1}{\sqrt{n}}\right) \\ O_p\left(\frac{1}{\sqrt{n}}\right) & O_p(1) \end{array}\right]. \quad (19)$$

Also, $\hat{\theta}_{1\lfloor Tr\rfloor} - \theta$ and $\hat{\theta}_{2\lfloor Tr\rfloor} - \theta$ are independent, as a consequence of both $B(r)$ and $C(r;0)$ having independent increments.

1.4 A technical note on Theorem 1. The FCLT is shown using a slightly different approach than ordinary panel CLT (Phillips and Moon, 1999). Here, the proof is based on showing that, as $T \to \infty$, the FCLT can hold and on working out the variance for each r. That $n \to \infty$ is only incidental to the proof. The proof essentially shows that (17) and (18) hold for all r, thereby proving convergence of the finite dimensional distributions; tightness is shown using conventional tightness arguments, based on bounding the remainder terms in the Beveridge-Nelson decomposition of w_t, $\widetilde{x}_{it}$ and u_{it} - see also the discussion in Phillips and Solo (1992).

As a consequence of Theorem 1, a similar result can also be shown to hold for the case of observable common shocks.

Proposition 2 *Let Assumption 1 hold. Then, as* $(n,T) \longrightarrow \infty$ *with* $\frac{n}{T} \to 0$, $\sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr\rfloor} - \theta\right]$ *and* $\sqrt{n}T\left[\hat{\theta}_{2\lfloor Tr\rfloor} - \theta\right]$ *converge to the same distributions as in (17) and (18) respectively, with* σ_ζ *replaced by* σ_u.

The FCLT derived in Theorem 1 can now be used to derive the distribution of test statistics under the null of no break.

4 Testing

In this section, we consider three statistics: the supremum of the Wald statistic, $SupW$, the average Wald statistic, $AveW$, and the logarithm of the Andrews-Ploberger exponential Wald statistic, $ExpW$ - see Andrews (1993), Andrews and Ploberger (1994).

Assumption 3: $r \in (0,1)$.

Assumption 3 states that the fraction of T at which the change point occurs, r, is bounded away from zero and one. Thus, our tests are designed to have power versus mid-sample alternatives, as it is typical in this literature (see Andrews, 1993, p.838).

Consider $\hat{\theta}_{1\lfloor Tr\rfloor}$ and $\hat{\theta}_{2\lfloor Tr\rfloor}$ defined in (13) and (14) respectively, and define, for $j = 1, 2$

$$\hat{\theta}^{*}_{j\lfloor Tr\rfloor} = \begin{bmatrix} \hat{\sigma}_{\zeta,j} I_R & 0 \\ 0 & \hat{\sigma}_{u,j} I_p \end{bmatrix}^{-1} \hat{\theta}_{j\lfloor Tr\rfloor},$$

where $\hat{\sigma}_{\zeta,j}$ and $\hat{\sigma}_{u,j}$ are consistent for σ^2_u and σ^2_ζ respectively under H_0. Compute the Wald-type statistic $W(\lfloor Tr\rfloor)$ as

$$W(\lfloor Tr\rfloor) = \left[\hat{\theta}^{*}_{1\lfloor Tr\rfloor} - \hat{\theta}^{*}_{2\lfloor Tr\rfloor}\right]' \begin{bmatrix} \left(\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor} \widehat{W}_{it}\widehat{W}'_{it}\right)^{-1} \\ + \left(\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T} \widehat{W}_{it}\widehat{W}'_{it}\right)^{-1} \end{bmatrix}^{-1} \left[\hat{\theta}^{*}_{1\lfloor Tr\rfloor} - \hat{\theta}^{*}_{2\lfloor Tr\rfloor}\right]. \quad (20)$$

Before showing the limiting distribution for $W(\lfloor Tr\rfloor)$ as n and T pass to infinity, some preliminary notation is necessary. Let $\mathbf{s}(r) = \left(\int_0^r \bar{B}_\varepsilon \bar{B}'_\varepsilon\right)^{-1/2} Z_1 - \left(\int_r^1 \bar{B}_\varepsilon \bar{B}'_\varepsilon\right)^{-1/2} Z_2$ and $\mathbf{M}(r) = \left[\left(\int_0^r \bar{B}_\varepsilon \bar{B}'_\varepsilon\right)^{-1} + \left(\int_r^1 \bar{B}_\varepsilon \bar{B}'_\varepsilon\right)^{-1}\right]$. Define

$$J(r) = \begin{bmatrix} \frac{[C(1;r)-C(r;0)]}{\sqrt{r^2+(1-r)^2}} \\ [\mathbf{M}(r)]^{-1/2}\mathbf{s}(r) \end{bmatrix}. \quad (21)$$

It holds that

Theorem 2 *Let Assumptions 1–3 hold. Then, under H_0, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$*

$$W(\lfloor Tr\rfloor) \xrightarrow{d} J(r)' J(r) = Q_R(r) + Q_p(r) \quad (22)$$

where $Q_R(r)$ and $Q_p(r)$ are independent and defined as

$$Q_R(r) = \mathbf{s}(r)' [\mathbf{M}(r)]^{-1} \mathbf{s}(r) \quad (23)$$

and

$$Q_p(r) = \frac{\left\| (1-r)^2 C(r;0) - r^2 C(1;r) \right\|^2}{r^4(1-r)^2 + r^2(1-r)^4}. \quad (24)$$

Remarks

2.1 Theorem 2 is an application of Theorem 1 and of the Continuous Mapping Theorem. For a given r, $Q_R(r) \sim \chi^2_R$ and $Q_p(r) \sim \chi^2_p$; thus, when suitably normalized, the difference of the partial sample estimates has a chi-squared distribution for fixed r with $R+p$ degrees of freedom.

2.2 Note that $J(r)' J(r)$ is a variance transformed, tied-down Bessel process; when suitably normalized by its covariance kernel, as in (20), it has the same distribution as in Andrews (1993).

2.3 In principle, one can construct tests separately for β and γ using $Q_R(r)$ and $Q_p(r)$ since $Q_R(r)$ and $Q_p(r)$ are independent. Theorem 2 states that if one wants to test only for the constancy of β it holds that $W(\lfloor Tr \rfloor) \xrightarrow{d} Q_R(r)$; if one is interested in testing merely for the constancy of γ it holds that $W(\lfloor Tr \rfloor) \xrightarrow{d} Q_p(r)$.

2.4 Theorem 2 holds under more general conditions than Assumption 1. For example, the theorem still holds if one allows for the presence of (strong) cross-dependence among the x_{it}s. As we show in greater detail in Section 5, this is because the limiting distribution of $W(\lfloor Tr \rfloor)$ follows from the asymptotic normality of $\left[\hat{\theta}'_{1\lfloor Tr \rfloor}, \hat{\theta}'_{2\lfloor Tr \rfloor}\right]'$, which still holds for cross correlated x_{it}s.

Theorem 2 is valid for any consistent estimators of σ_u^2 and σ_ζ^2 (of course the choice of these estimators will affect the finite sample performance). Although here we propose to use estimators based on the pre- and post- break subsamples, alternatively one could use estimators based on the full sample - see also the discussion in Andrews (1993, p. 833). To estimate σ_u^2, one could compute

$$\hat{\sigma}_{u,1}^2 = \frac{1}{n(Tr)} \sum_{i=1}^{n} \sum_{t=1}^{\lfloor Tr \rfloor} \left[y_{it} - \bar{y}_i - \hat{\theta}'_{1\lfloor Tr \rfloor} \widehat{W}_{it} \right]^2, \tag{25}$$

which is consistent under H_0 (a similar definition would apply for $\hat{\sigma}_{u,2}^2$).

To find a consistent estimator of σ_ζ^2, consider the case of e_{it} being *i.i.d.*. From (15) a possible choice is

$$\hat{\sigma}_{\zeta,1}^2 = \hat{\sigma}_{u,1}^2 + \hat{\sigma}_{\Pi,1}^2,$$

and similarly for $\hat{\sigma}_{\zeta,2}^2$. From (16), we have $\hat{\sigma}_{\Pi,1}^2 = \hat{\beta}'_{1\lfloor Tr \rfloor} \hat{\sigma}_{\pi,1}^2 \hat{\beta}_{1\lfloor Tr \rfloor}$ with

$$\hat{\sigma}_{\pi,1}^2 = \left[\frac{1}{(Tr)^2} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t \hat{w}'_t \right] \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{Tr} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{e}_{it}^2 \right) \hat{\lambda}_i \hat{\lambda}'_i \right] \left[\frac{1}{(Tr)^2} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t \hat{w}'_t \right], \tag{26}$$

where $\hat{\lambda}_i$ is a consistent estimate of λ_i and $\hat{e}_{it}$ can be computed as $\hat{e}_{it} = z_{it} - \hat{\lambda}'_i \widehat{F}_t$. Therefore, we can provide an estimate for σ_ζ^2 for each subsample as

$$\hat{\sigma}_{\zeta,1}^2 = \hat{\sigma}_{u,1}^2 + \hat{\beta}'_{1\lfloor Tr \rfloor} \hat{\sigma}_{\pi,1}^2 \hat{\beta}_{1\lfloor Tr \rfloor}. \tag{27}$$

If e_{it} is serially correlated, a different formula should be used in (26), replacing $\frac{1}{Tr}\sum_{t=1}^{\lfloor Tr \rfloor} \hat{e}_{it}^2$ with e.g. some weighted sum-of-covariances estimator.

Proposition 3 *Suppose Assumptions 1-3 hold. Then, as $(n,T) \to \infty$, under H_0, it holds that $\hat{\sigma}_{u,1}^2 \xrightarrow{p} \sigma_u^2$ and $\hat{\sigma}_{\zeta,1}^2 \xrightarrow{p} \sigma_\zeta^2$, and similarly for $\hat{\sigma}_{u,2}^2$ and $\hat{\sigma}_{\zeta,2}^2$.*

Following Andrews (1993) and Andrews and Ploberger (1994), we consider three test statistics for the null of no break:

$$
\begin{aligned}
SupW &\equiv \max_{[Tr^*]\leq\lfloor Tr\rfloor\leq T-[Tr^*]} W\left(\lfloor Tr\rfloor\right), \\
AveW &\equiv \frac{1}{T}\sum_{\lfloor Tr\rfloor=[Tr^*]}^{T-[Tr^*]} W\left(\lfloor Tr\rfloor\right), \\
ExpW &\equiv \ln\left\{\frac{1}{T}\sum_{\lfloor Tr\rfloor=[Tr^*]}^{T-[Tr^*]} \exp\left[\frac{1}{2}W\left(\lfloor Tr\rfloor\right)\right]\right\},
\end{aligned}
$$

where r^* represents the fraction of the sample trimmed away from the beginning and the end of the sample. Using the continuous mapping theorem (CMT) we have the following result:

Corollary 1 *Suppose Assumptions 1-3 hold. Under* H_0 *as* $(n,T)\to\infty$ *with* $\frac{n}{T}\to 0$

$$
\begin{aligned}
SupW &\xrightarrow{d} \underset{r^*\leq r\leq 1-r^*}{sup}\left[J(r)'J(r)\right], \\
AveW &\xrightarrow{d} \int_{r^*}^{1-r^*}\left[J(r)'J(r)\right]dr, \\
ExpW &\xrightarrow{d} \ln\left\{\int_{r^*}^{1-r^*}\exp\left[\frac{1}{2}J(r)'J(r)\right]dr\right\},
\end{aligned}
$$

uniformly in r.

Critical values for $SupW$, $AveW$, and $ExpW$ can be taken from Andrews (1993) and Andrews and Ploberger (1994) in light of Remark 2.2. For example, when $r^*=0.15$ and $R=p=1$, the critical values of the 5% level for $SupW$, $AveW$, and $ExpW$ are 11.79, 4.61, and 3.22 respectively.

4.1 Consistency of the test

In this section, we show that Wald-type tests have non-trivial power versus a general class of local-to-null alternatives of order $O_p\left(1/\sqrt{n}T\right)$. Such alternatives include the case of abrupt change around one common changepoint, but they also include smooth transitions from one regime to another and the possibility of different changepoints for different units in the panel. The time series properties of the local alternatives considered here are in line with the findings in the literature, and particularly the results that the test has nontrivial power versus alternatives of order $1/T$ and the presence of power versus smooth transition changes. The cross-sectional properties of the local alternatives are found to be more general than those the test is designed for. The test is shown to have nontrivial power versus alternatives of order $1/\sqrt{n}$ (a consequence of the panel approach) and under the case whereby different units may undergo changes at different points in time (if any). We also investigate the power when only some units (possibly a finite number) have a break, showing that when a finite number of units have a break, test have power versus local alternatives of order $O_p\left(\sqrt{n}/T\right)$.

We assume the following sequence of local alternatives:

$$H_a^{(nT)} : \theta_{it}^{(nT)} = \theta + \frac{\sqrt{n}}{m_n T} g_i\left(\frac{t}{T}\right), \tag{28}$$

where $g_i(\cdot)$ is a $(R+p)\times 1$ arbitrary, finite and non-zero function defined on the unit interval, and m_n is the number of units for which $g_i(\cdot) \neq 0$ (i.e. the units that have a nontrivial break); m_n can be finite or pass to infinity as $n \to \infty$.

The properties of $g_i\left(\frac{t}{T}\right)$ are specified in the following assumption.

Assumption 4: The function $g_i\left(\frac{t}{T}\right)$ is nonconstant and it belongs to the class of Riemann integrable functions and as $(n,T)\to\infty$, for all r: *(a)* $\frac{1}{m_n T}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|g_i\left(\frac{t}{T}\right)\right\|^2 = O_p(1)$; *(b)* $\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor} W_{it}W_{it}' g_i\left(\frac{t}{T}\right) = O_p(1)$; *(c)* $\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor} g_i'\left(\frac{t}{T}\right) W_{it}W_{it}' g_i\left(\frac{t}{T}\right) = O_p(1)$; and *(d)* $\frac{1}{m_n T}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor} W_{it}' g_i\left(\frac{t}{T}\right) u_{it} = O_p(1)$.

Possible alternative functional forms for $g_i(\cdot)$ include: a single step function, i.e., $g_i(s) = 0$ if $s < r$ and $g_i(s) = \triangle\theta$ (finite) if $s \geq r$, which represents a one-time change on θ at $\lfloor Tr\rfloor = \lfloor Tr\rfloor$; multiple steps functions that represent multiple changes; time trending function $g_i(\cdot) = t/T$.

Under the alternative, we assume that θ can differ across individual. This has a twofold implication. First, having different $g_i(\cdot)$ across i entails having, under the alternative, breaks of possibly different magnitude. The case whereby some units have a zero size break (i.e., no change) is taken into account also. Second, allowing for $g_i(\cdot)$ to differ across i also allows to consider a specification, under the alternative, where the time of the break (and the presence of breaks itself) is not restricted to be the same across units. This case could be envisaged to take into account the presence of a common source of break but differently timed reactions due to different hysteresis across units. Of course, some units may not have any breaks at all, which is taken into account by allowing for m_n to be strictly smaller than n.

We now turn to studying the asymptotic distribution of the Wald statistic under the sequence of local alternatives (28). Model (3) can be rewritten as $y_{it}^{(nT)} = \alpha_i + W_{it}'\theta_t^{(nT)} + u_{it}$. Similarly, when common shocks are replaced by their estimates $\widehat{W}_{it}$ we have $y_{it}^{(nT)} = \alpha_i + \widehat{W}_{it}'\theta_t^{(nT)} + v_{it}$, with $v_{it} = u_{it} + \left(F_t - \widehat{F}_t\right)'\beta_t^{(nT)}$. Let $\hat{\theta}_{1\lfloor Tr\rfloor}^{(nT)}$ and $\hat{\theta}_{2\lfloor Tr\rfloor}^{(nT)}$ be the OLS estimators under the local alternative (28), and let $\tilde{\sigma}_{u,j}^2$ and $\tilde{\sigma}_{\zeta,j}^2$ be consistent estimators for σ_u^2 and σ_ζ^2 respectively under $H_a^{(nT)}$. Define

$$\hat{\theta}_{j\lfloor Tr\rfloor}^{*(nT)} = \begin{bmatrix} \tilde{\sigma}_{\zeta,j} I_R & 0 \\ 0 & \tilde{\sigma}_{u,j} I_p \end{bmatrix}^{-1} \hat{\theta}_{j\lfloor Tr\rfloor}^{(nT)}$$

for $j=1,2$, the Wald statistics under the local alternative can be computed as

$$W^{(nT)}\left(\lfloor Tr\rfloor\right)=\left[\hat{\theta}^{*(nT)}_{1\lfloor Tr\rfloor}-\hat{\theta}^{*(nT)}_{2\lfloor Tr\rfloor}\right]'\left[\begin{array}{c}\left(\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right)^{-1}\\+\left(\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right)^{-1}\end{array}\right]^{-1}\left[\hat{\theta}^{*(nT)}_{1\lfloor Tr\rfloor}-\hat{\theta}^{*(nT)}_{2\lfloor Tr\rfloor}\right]. \tag{29}$$

The local asymptotic power for the Wald statistics is given in the following theorem:

Theorem 3 *Suppose Assumptions 1-4 hold. Then under $H_a^{(nT)}$, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$, it holds that $W^{(nT)}\left(\lfloor Tr\rfloor\right)\xrightarrow{d}\left[J(r)+d(r)\right]'\left[J(r)+d(r)\right]$, where $J(r)$ is defined in (21) and $d(r)$ is defined in Appendix.*

Remarks

3.1 Consider the case $m_n=n$. The test has power versus alternatives shrinking as $O_p\left(1/\sqrt{n}T\right)$; this is a direct consequence of the $\sqrt{n}T$ rate of convergence in Theorem 1. This finding is consistent with the analysis in Bai, Lumsdaine and Stock (1998), Qu and Perron (2007) and Bai (2010), where it is shown that the quality of the breakpoint estimates improves as the number of time series employed increases.

3.2 Theorem 3 shows that the test has some well-known time series properties, e.g., the presence of power versus "smooth" changes as opposed to abrupt changes for which it is designed for; this is consistent with the findings in Andrews (1993). The test also has some cross-sectional properties: albeit designed for the detection of common changepoints, the test exhibits nontrivial power versus breakpoints located at different times for different time series. Thus, cases whereby a common shock introduces breaks in all units but at different points in time due, e.g., to different levels of hysteresis of inertia across units are encompassed by the test.

3.3 When m_n is $o(n)$, the test has a loss of power. This is not surprising, since only some of the cross sectional information (the one from the units which have a break) is actually relevant. As an extreme, consider the case when the number of units that do have a break is finite, i.e. $m_n=O(1)$. Theorem 3 shows that in this case the test has power versus alternatives of order $O_p\left(\sqrt{n}/T\right)$, thus being worse than in the univariate case where nontrivial power is attained versus local to null alternatives of order $O_p(1/T)$. Including cross sectional information is beneficial when there are breaks in the units, whereas it worsens the performance of the tests when many units do not have breaks.

Theorem 3 holds for any choice of the estimators $\tilde{\sigma}_u^2$ and $\tilde{\sigma}_\zeta^2$, as long as they are consistent under $H_a^{(nT)}$. A possible estimator for σ_u^2 is $\tilde{\sigma}_{u,1}^2=\frac{1}{n(Tr)}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left[y_{it}-\bar{y}_i-\hat{\theta}^{(nT)\prime}_{1\lfloor Tr\rfloor}\widehat{W}_{it}\right]^2$. To estimate σ_ζ^2 we propose $\tilde{\sigma}_{\zeta,1}^2=\tilde{\sigma}_{u,1}^2+\hat{\beta}^{(nT)\prime}_{1\lfloor Tr\rfloor}\hat{\sigma}_{\pi,1}^2\hat{\beta}^{(nT)}_{1\lfloor Tr\rfloor}$, where $\hat{\sigma}_{\pi,1}^2$ is defined in equation (26) and $\hat{\beta}_{1\lfloor Tr\rfloor}^{(nT)}$ is the partial sample OLS estimator for β under $H_a^{(nT)}$; similar definitions can be computed for $\tilde{\sigma}_{u,2}^2$ and $\tilde{\sigma}_{\zeta,2}^2$.

The following proposition establishes consistency for $\tilde{\sigma}_{u,1}^2$ and $\tilde{\sigma}_{\zeta,1}^2$ under $H_a^{(nT)}$:

Proposition 4 *Suppose Assumptions 1-4 hold. Then under the local alternative hypotheses $H_a^{(nT)}$ defined in equation (28), it holds that as $(n,T)\to\infty$, $\tilde{\sigma}_{u,1}^2 \xrightarrow{p} \sigma_u^2$ and $\tilde{\sigma}_{\zeta,1}^2 \xrightarrow{p} \sigma_\zeta^2$. The same holds for $\tilde{\sigma}_{u,2}^2$ and $\tilde{\sigma}_{\zeta,2}^2$.*

5 Extensions

In this section, we consider two extensions: *(a)* the presence of cross-sectional correlation among the idiosyncratic shocks x_{it}, and *(b)* the presence of endogeneity in the cointegration relationship. First, we prove that even though the asymptotic law of the OLS estimator changes when cross-dependence is considered among the x_{it}s, however the asymptotic distribution (and local power) of Wald-type statistics is unaltered. Second, as far as the case of endogeneity is concerned, we develop an FMOLS estimator to accommodate for it, showing that, when using this estimator, results concerning the test remain unchanged.

Cross-sectional dependence among the x_{it}s

Theorem 1 shows an FCLT for $\hat{\theta}_{1\lfloor Tr\rfloor}$ and $\hat{\theta}_{2\lfloor Tr\rfloor}$, from which the limiting distribution of the Wald statistic $W(\lfloor Tr\rfloor)$ can be inferred. We show that asymptotic mixed normality is preserved also when we allow for cross-dependence among the x_{it}s, even though asymptotic orthogonality between $\hat{\beta}-\beta$ and $\hat{\gamma}-\gamma$ does not hold any more. Consider a simpler version of (2), viz. $x_{it}=\Gamma_i G_t$, where G_t is a set of $I(1)$ shocks that could contain some of the F_ts, and let the following assumptions hold.

Assumption 5: $G_t = G_{t-1}+\varepsilon_t^G$ with *(a)* $\{u_{it}\}$, $\{\varepsilon_t, \varepsilon_t^G\}$, and $\{e_{it}\}$ mutually independent across i for all t; *(b)* letting $\omega_{2it}=\left[u_{it}, \varepsilon_t', \varepsilon_t^{G\prime}, e_{it}\right]'$, ω_{2it} is a linear process across t with mean zero and finite 4th moment and an FCLT holds so that uniformly in r for all i

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tr\rfloor}\omega_{2it}\xrightarrow{d} B_{2\omega}(r)=\begin{bmatrix} B_u(r)\\ B_\varepsilon(r)\\ B_G(r)\\ B_e(r)\end{bmatrix},$$

where $B_\omega(r)$ is a multivariate Brownian motion, whose elements have covariance matrices σ_u^2, Ω_ε, Ω_G and σ_e^2; *(c)* a Beveridge-Nelson decomposition holds for ω_{2it} with the remainder bounded as in equation (11).

Assumption 6: *(a)* $\{\Gamma_i\}_{i=1}^n$ is a random sequence independent of $\left[u_{it}, \varepsilon_t', \varepsilon_t^{G\prime}, e_{it}\right]'$ with *(i)* $n^{-1}\sum_{i=1}^n \Gamma_i\Gamma_i' \xrightarrow{p} \Sigma_\Gamma$ where Σ_Γ is non singular and *(ii)* $n^{-1}\sum_{i=1}^n\Gamma_i \xrightarrow{p} \bar{\Gamma}$; *(b)* as $(n,T)\to\infty$ it holds that

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{G}_t\bar{G}_t'\Gamma_i' \xrightarrow{d} \Sigma_\Gamma^{1/2}\left(\int_0^r \bar{B}_G\bar{B}_G'\right)\Sigma_\Gamma^{1/2}$$

and (under $\frac{n}{T}\to 0$)

$$\frac{1}{\sqrt{n}T^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{G}_t w_{t,r}' \xrightarrow{d} \bar{\Gamma}\int_0^r \bar{B}_G\bar{B}_\varepsilon'$$

uniformly in r.

Assumption 5 extends Assumption 1 to take account of the factor structure in the x_{it}s; for the sake of simplicity, the idiosyncratic component in (2), w_{it}, is omitted here. Assumption 6*(a)(ii)* considers two possible, alternative cases, whereby the loadings Γ_i have either zero or nonzero mean. This will be shown to play an important role in the asymptotic variance of $\hat{\theta}_{\lfloor Tr\rfloor}$.

The following result, which is the equivalent to Theorem 1, holds:

Proposition 5 *Suppose Assumptions 2, 5 and 6 hold. Under H_0, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$*

$$\sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr\rfloor}-\theta\right]\xrightarrow{d}\sigma_\zeta\left[\begin{array}{cc}\int_0^r \bar{B}_\varepsilon\bar{B}_\varepsilon' & \int_0^r \bar{B}_\varepsilon\bar{B}_G'\bar{\Gamma}' \\ \bar{\Gamma}\int_0^r \bar{B}_G\bar{B}_\varepsilon' & \Sigma_\Gamma^{1/2}\left(\int_0^r \bar{B}_G\bar{B}_G'\right)\Sigma_\Gamma^{1/2}\end{array}\right]^{-1/2}\times Z,$$

uniformly in r, with Z an $(R+p)$-dimensional Gaussian random variable.

Proposition 5 shows, in essence, that the FCLT for $\hat{\theta}_{1\lfloor Tr\rfloor}-\theta$ is the same as $\hat{\beta}_{1\lfloor Tr\rfloor}-\beta$ for in Theorem 1 (up to adjusting the dimension from R to $p+R$). This entails that $W(\lfloor Tr\rfloor)$ has, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$, the same distribution as $Q_R(r)$ in (23), although the dimension changes to $R+p$. Therefore, tests based on $W(\lfloor Tr\rfloor)$ have the same properties under Assumptions 5 and 6 as under Assumptions 1; this includes power versus local alternatives and consistency of the estimator of σ_ζ.

Proposition 5 also has some important differences with Theorem 1, mainly in terms of the correlation between $\hat{\beta}-\beta$ and $\hat{\gamma}-\gamma$ and the presence of σ_ζ, instead of σ_u, in the expression of $\hat{\gamma}-\gamma$. The former is in general not zero, thus not having asymptotic independence, unless $\bar{\Gamma}=0$; this, in essence, is because as $T\to\infty$, $T^{-2}\sum_{t=1}^{T}G_tF_t'$ does not converge to zero even if F_t and G_t are independent.

Endogeneity and FMOLS estimation

This section considers extending the framework to incorporate the case of endogeneity; we develop a FMOLS estimator and we show that the null distribution of the Wald-type test statistic is the same as in Theorem 2. This is based on showing that an FCLT holds for the partial sample FMOLS estimator, and thus all results derived above still hold in the case of endogeneity.

Assumption 7: *(a)* Assumption 1(a) holds; *(b)* a multivariate invariance principle for ω_{it} holds, such that, as $T \to \infty$ for all i, $\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Tr \rfloor} \omega_{it} \xrightarrow{d} B_i(r) = \left[B_{u,i}(r), B'_{\varepsilon}(r), B'_{\epsilon,i}(r), B_{e,i}(r) \right]'$, with long run covariance matrix given by $diag\left\{\Omega, \sigma_e^2\right\}$, where $diag\{\cdot\}$ represents a block diagonal matrix and

$$\Omega = \begin{bmatrix} \Omega_u & \Omega_{u\varepsilon} & \Omega_{u\epsilon} \\ \Omega_{\varepsilon u} & \Omega_{\varepsilon} & \Omega_{\varepsilon\epsilon} \\ \Omega_{\epsilon u} & \Omega_{\epsilon\varepsilon} & \Omega_{\epsilon} \end{bmatrix},$$

partitioned conformably with ω_{it}, where we assume $\Omega_{\epsilon\varepsilon}$ and $\Omega_{\varepsilon\epsilon}$ both equal to zero for simplicity; *(c)* $\{u_{it}, \varepsilon'_t, \epsilon'_{it}\}$ and $\{e_{it}\}$ are two independent groups.

Assumption 7 substitutes Assumption 1, and it considers the presence of endogeneity, which is taken into account through the terms $\Omega_{\varepsilon u}$ and $\Omega_{\epsilon u}$ in the long run covariance matrix: both the common and the idiosyncratic shocks can be correlated with the error term u_{it}. For simplicity, the long run covariance matrix of the $B_i(r)$ is the same for all i. Independence between $\{e_{it}\}$ and $\{u_{it}, \varepsilon_t, \epsilon_{it}\}$ is still required in order for the asymptotic theory for $\widehat{F}_t$ to hold.

Define also

$$\Lambda = \lim_{T\to\infty} \frac{1}{T} E \begin{bmatrix} \sum_t u_{it}u_{i0} & \sum_t u_{it}\Delta F'_0 & \sum_t u_{it}\Delta x'_{i0} \\ \sum_t \Delta F_t u_{i0} & \sum_t \Delta F_t \Delta F'_0 & \sum_t \Delta F_t \Delta x'_{i0} \\ \sum_t \Delta x_{it} u_{i0} & \sum_t \Delta x_{it}\Delta F'_0 & \sum_t \Delta x_{it}\Delta x'_{i0} \end{bmatrix} = \begin{bmatrix} \Lambda_u & \Lambda_{u\varepsilon} & \Lambda_{u\epsilon} \\ \Lambda_{\varepsilon u} & \Lambda_{\varepsilon} & \Lambda_{\varepsilon\epsilon} \\ \Lambda_{\epsilon u} & \Lambda_{\epsilon\varepsilon} & \Lambda_{\epsilon} \end{bmatrix},$$

and as before, for simplicity we assume $\Lambda_{\epsilon\varepsilon}$ and $\Lambda_{\varepsilon\epsilon}$ equal to zero. Assumption 7 entails (see e.g., Phillips and Durlauf, 1986)

$$T\left[\sum_{t=1}^{T} \widehat{W}_{it}\widehat{W}'_{it}\right]^{-1} \left\{\sum_{t=1}^{T} \widehat{W}_{it}\left[u_{it} + \beta'(w_t - \hat{w}_t)\right]\right\}$$
$$\xrightarrow{d} \begin{bmatrix} \left(\int \bar{B}_{\varepsilon}\bar{B}'_{\varepsilon}\right)^{-1/2} \sqrt{\Omega_u + \sigma_e^2\left(\beta'\tilde{Q}_B\Sigma_{\Lambda}\tilde{Q}'_B\beta\right)} \times Z + \Lambda_{\varepsilon u} \\ \int B_{\epsilon,i} dB_{u,i} + \Lambda_{\epsilon u} \end{bmatrix},$$

for all i. Therefore, the OLS estimator of θ is no longer $\sqrt{n}T$ consistent due to the persistence of the asymptotic bias terms $\Lambda_{\varepsilon u}$ and $\Lambda_{\epsilon u}$ across units - see also the discussion in Phillips and Moon (1999, p. 1084).

Let $\Omega_{(\varepsilon\epsilon),u} = \begin{bmatrix} \Omega'_{\varepsilon u} & \Omega'_{\epsilon u} \end{bmatrix}'$, $\Omega_{(\varepsilon\epsilon)} = diag\{\Omega_{\varepsilon}, \Omega_{\epsilon}\}$, $\Lambda_{(\varepsilon\epsilon),u} = \begin{bmatrix} \Lambda'_{\varepsilon u} & \Lambda'_{\epsilon u} \end{bmatrix}'$ and $\Lambda_{(\varepsilon\epsilon)} = diag\{\Lambda_{\varepsilon}, \Lambda_{\epsilon}\}$, where $diag\{\Omega_{\varepsilon}, \Omega_{\epsilon}\}$ is understood as a block diagonal matrix with Ω_{ε} and Ω_{ϵ} on the main diagonal and similarly for $\Lambda_{(\varepsilon\epsilon)}$. Estimates of Ω and Λ are based on $\hat{\Gamma}_i(j) = \frac{1}{T}\sum_{t=1}^{T} W^+_{i,t+j} W^{+\prime}_{i,t}$, with $W^+_{i,t} = \left[\hat{u}_{it}, \Delta\widehat{F}_t, \Delta x'_{it}\right]'$. Estimates of Ω_i and Λ_i can be calculated as $\hat{\Omega}_i = \sum_{j=-T+1}^{T-1} w\left(\frac{j}{h}\right)\hat{\Gamma}_i(j)$ and $\hat{\Lambda}_i = \sum_{j=0}^{T-1} w\left(\frac{j}{h}\right)\hat{\Gamma}_i(j)$ respectively, where $w(x)$ is a kernel and h is the bandwidth.

The averaged kernel estimators of Ω and Λ are calculated as $\hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n}\hat{\Omega}_i$ and $\hat{\Lambda} =$

$\frac{1}{n}\sum_{i=1}^{n}\hat{\Lambda}_i$. The following assumption characterizes the kernel $w(x)$.

Assumption 8: *(a)* $w(0)=1$, $w(x)=w(-x)$, $\int_{-\infty}^{\infty} w^2(x)\,dx<\infty$ and $\lim_{x\to\infty}|x|^{-q}[1-w(x)]<\infty$, for some $q\in(0,\infty)$; *(b)* as n, T and $h\to\infty$, it holds that $\frac{h}{T}\to 0$ and $\frac{h^{2q}}{T}\to\varepsilon>0$.

Assumption 8 is the same as Assumption 11 in Phillips and Moon (1999, p. 1085), and it ensures that

$$\begin{aligned}\sqrt{n}\left(\hat{\Omega}-\Omega\right) &= o_p(1), & (30)\\ \sqrt{n}\left(\hat{\Lambda}-\Lambda\right) &= o_p(1). & (31)\end{aligned}$$

These are stronger than simple consistency and they are needed to prove the asymptotics of the FMOLS estimator (see Phillips and Moon, 1999, p. 1109). Note that n does not play a role in determining the value of the bandwidth h.

Let

$$\hat{\Lambda}^{+}_{(\varepsilon\epsilon),u}=\hat{\Lambda}_{(\varepsilon\epsilon),u}-\frac{1}{\sqrt{n}}\begin{pmatrix}\hat{\Lambda}_{\varepsilon} & 0\\ 0 & \sqrt{n}\hat{\Lambda}_{\epsilon}\end{pmatrix}\hat{\Omega}^{-1}_{(\varepsilon\epsilon)}\hat{\Omega}_{(\varepsilon\epsilon),u};$$

the partial sample estimates for θ are defined as

$$\begin{aligned}\hat{\theta}^{FM}_{1\lfloor Tr\rfloor} &= \left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\widehat{W}_{it}\widehat{W}'_{it}\right]^{-1}\times\\ &\quad\left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\widehat{W}_{it}y_{it}-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\begin{pmatrix}\hat{w}_t\Delta\widehat{F}'_t & \hat{w}_t\Delta x'_{it}\\ x_{it}\Delta\widehat{F}'_t & \sqrt{n}x_{it}\Delta x'_{it}\end{pmatrix}\hat{\Omega}^{-1}_{(\varepsilon\epsilon)}\hat{\Omega}_{(\varepsilon\epsilon),u}-nk\hat{\Lambda}^{+}_{(\varepsilon\epsilon),u}\right],\end{aligned}$$

and

$$\begin{aligned}\hat{\theta}^{FM}_{2\lfloor Tr\rfloor} &= \left[\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\widehat{W}_{it}\widehat{W}'_{it}\right]^{-1}\times\\ &\quad\left[\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\widehat{W}_{it}y_{it}-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\begin{pmatrix}\hat{w}_t\Delta\widehat{F}'_t & \hat{w}_t\Delta x'_{it}\\ x_{it}\Delta\widehat{F}'_t & \sqrt{n}x_{it}\Delta x'_{it}\end{pmatrix}\hat{\Omega}^{-1}_{(\varepsilon\epsilon)}\hat{\Omega}_{(\varepsilon\epsilon),u}\right.\\ &\quad\left.-n\left(T-\lfloor Tr\rfloor\right)\hat{\Lambda}^{+}_{(\varepsilon\epsilon),u}\right].\end{aligned}$$

This estimation procedure follows similar lines as e.g., in Phillips and Moon (1999, p. 1085); the only difference here is the presence of the normalization term $\sqrt{n}$, due to the presence of common regressors across units.

The Wald statistic for the null of no structural change is defined as

$$W^{FM}(\lfloor Tr \rfloor) = \left[\hat{\theta}^{FM*}_{1\lfloor Tr \rfloor} - \hat{\theta}^{FM*}_{2\lfloor Tr \rfloor}\right]' \left[\begin{array}{c} \left(\sum_{i=1}^{n} \sum_{t=1}^{\lfloor Tr \rfloor} \widehat{W}_{it}\widehat{W}'_{it}\right)^{-1} \\ + \left(\sum_{i=1}^{n} \sum_{t=\lfloor Tr \rfloor+1}^{T} \widehat{W}_{it}\widehat{W}'_{it}\right)^{-1} \end{array} \right]^{-1} \left[\hat{\theta}^{FM*}_{1\lfloor Tr \rfloor} - \hat{\theta}^{FM*}_{2\lfloor Tr \rfloor}\right], \tag{32}$$

where

$$\hat{\theta}^{FM*}_{j\lfloor Tr \rfloor} = \left[\begin{array}{cc} \hat{\Omega}^2_{\zeta} I_R & 0 \\ 0 & \left(\hat{\Omega}_u - \hat{\Omega}_{u\epsilon}\hat{\Omega}^{-1}_{\varepsilon}\hat{\Omega}_{\varepsilon u}\right) I_p \end{array} \right]^{-1} \hat{\theta}^{FM}_{j\lfloor Tr \rfloor},$$

with $\hat{\Omega}_{\zeta}$ being an estimator of

$$\Omega_{\zeta} = \left(\Omega_u - \Omega_{u\varepsilon}\Omega^{-1}_{\varepsilon}\Omega_{\varepsilon u}\right) + \sigma^2_e \left(\beta' \tilde{Q}_B \Sigma_{\Lambda} \tilde{Q}'_B \beta\right), \tag{33}$$

constructed as $\hat{\Omega}_{\zeta} = \left(\hat{\Omega}_u - \hat{\Omega}_{u\varepsilon}\hat{\Omega}^{-1}_{\varepsilon}\hat{\Omega}_{\varepsilon u}\right) + \hat{\sigma}^2_e \left(\hat{\beta}' \hat{\sigma}^2_{\pi} \hat{\beta}\right)$, where $\hat{\beta}' \hat{\sigma}^2_{\pi} \hat{\beta}$ is defined in (26).

It holds that

Proposition 6 *Let Assumptions 2, 7 and 8 hold. Under H_0, as $(n, T) \to \infty$ with $\frac{n}{T} \to 0$, it holds that $W^{FM}(\lfloor Tr \rfloor) \xrightarrow{d} J(r)' J(r)$.*

Proposition states that the limiting distribution of $W^{FM}(\lfloor Tr \rfloor)$ is the same as the one in Theorem 2; thus, none of the results obtained so far, including Corollary 1 and the critical values for the statistics $SupW^{FM}(\lfloor Tr \rfloor)$, $AveW^{FM}(\lfloor Tr \rfloor)$, and $ExpW^{FM}(\lfloor Tr \rfloor)$, change.

6 Monte Carlo Simulations

In this section, we use synthetic data to assess the null rejection probabilities and the power properties of $SupW(\lfloor Tr \rfloor)$, $AveW(\lfloor Tr \rfloor)$, and $ExpW(\lfloor Tr \rfloor)$. We consider combinations of $n = \{20, 40, 60, 120, 240, 480\}$ and $T = \{20, 40, 60, 120, 240, 480\}$. The Monte Carlo experiments are based on the following design

$$\begin{aligned} y_{it} &= \alpha_i + \beta'_t F_t + \gamma'_t x_{it} + u_{it}, \\ F_t &= F_{t-1} + \varepsilon_t, \\ x_{it} &= x_{it-1} + \epsilon_{it}, \\ z_{it} &= \lambda'_i F_t + e_{it}. \end{aligned}$$

We assume a single factor and one single idiosyncratic component, such that $R = p = 1$. Under the null hypothesis of no structural change, we set the values of the parameters $\beta = 1$ and $\gamma = 1$. Also, we generate α_i and λ_i from *i.i.d.* $N(0, 1)$ and $N(2, 1)$ respectively.

We run a first set of experiments with *i.i.d.* data, reported in Tables 1a and 2a

[Insert Tables 1a and 2a here]

In order to assess power and size under serial correlation and endogeneity, we generate $[u_{it}, \varepsilon'_{it}, \epsilon'_{it}, e'_{it}]'$ as follows. We firstly create a Gaussian *i.i.d.* sequence $[\dot{u}_{it}, \dot{\varepsilon}'_{it}, \dot{\epsilon}'_{it}, e'_{it}]'$, which ensures that the error term e_{it} is independent of $\dot{u}_{it}$, $\dot{\varepsilon}'_{it}$ and $\dot{\epsilon}_{it}$. Contemporaneous correlation between $\dot{u}_{it}$, $\dot{\varepsilon}'_{it}$ and $\dot{\epsilon}_{it}$ is imposed by premultiplying $\dot{e}_t = [\dot{u}_{it}, \dot{\varepsilon}'_{it}, \dot{\epsilon}'_{it}]'$ by the Choleski factor of

$$\Pi = \begin{bmatrix} I_{\lfloor Tr \rfloor} & 0 & \rho^F 1_{\lfloor Tr \rfloor, n} \\ 0 & I_{np} & \rho^X 1^X_{np,n} \\ \rho^F 1_{n, \lfloor Tr \rfloor} & \rho^X 1^X_{n,np} & I_n \end{bmatrix}.$$

The coefficients ρ^F and ρ^X represent the correlation coefficients between $\ddot{u}_{it}$ and $\ddot{\varepsilon}_{it}$ and $\ddot{u}_{it}$ and $\ddot{\epsilon}'_{it}$ respectively in the new vector $\ddot{e}_t = [\ddot{u}_{it}, \ddot{\varepsilon}'_{it}, \ddot{\epsilon}'_{it}]'$; also, $I_{\lfloor Tr \rfloor}$ denotes an identity matrix of dimension $\lfloor Tr \rfloor$ and e.g., $1_{\lfloor Tr \rfloor, n}$ is a $\lfloor Tr \rfloor \times n$ matrix of ones. Serial correlation is induced by creating $e_t = [u_{it}, \varepsilon'_{it}, \epsilon'_{it}]'$ according to an ARMA(1,1) specification defined by

$$e_t = \rho_e e_{t-1} + \ddot{e}_t + \theta_e \ddot{e}_{t-1}. \tag{34}$$

We report results for $\{\rho^F, \rho^X, \rho_e, \theta_e\} = \{-0.4, -0.4, 0, 0.4\}$ in Tables 1b and 2b

[Insert Tables 1b and 2b here]

Size and power are evaluated at 5% level. For the purpose of size distortion assessment, we note that the critical values of the 5% level for $SupW$, $AveW$, and $ExpW$ are 11.79, 4.61, and 3.22 respectively, as derived by Andrews (1993) and Andrews and Ploberger (1994). Power assessment is conducted under the alternative hypothesis of structural change in $\theta = [\beta', \gamma']'$, where the break is located at the 40% of the sample. To control for the break magnitude, we set under H_a

$$\theta_t = \begin{cases} \theta \text{ for } t < \lfloor Tr \rfloor \\ (1+c)\,\theta \text{ for } t \geq \lfloor Tr \rfloor \end{cases}$$

where c is a scalar that defines the percentage change in the parameter values. We set $c = 0.5$.

All our results have been obtained using $10,000$ iterations; when generating DGPs, the first $1,000$ observations are discarded to avoid dependence on the initial conditions. All routines have been written using Gauss 6.0.

Table 1a contains empirical rejection frequencies of the test statistics, $SupW$, $AveW$, and $ExpW$, under the null that β and γ are stable over time. Overall, all three test statistics show good size as n and T increase. The three test statistics are undersized if n and T are small, though size distortion is significantly reduced as T increases. Table 1b illustrates the null rejection frequencies for tests constructed upon the FMOLS estimator. The test statistics tend to be slightly undersized for small samples, and using the FMOLS estimator leads to a more evident improvement in the sizes for all statistics for small samples, starting from $T = 20$. As samples get larger, results are substantially equivalent as those found when tests are conducted using the OLS estimator, with the only exception being the cases of $(n, T) = (20, 120)$ and $(n, T) = (20, 240)$ where the three statistics become substantially oversized.

Table 2a gives the power of the test statistics. All tests show very good power properties. The power gain is substantial as T increases and more moderate for increasing sizes of n. This result is consistent with the $\sqrt{n}T$ asymptotics of the three tests, as reported in the paper. When using the FMOLS estimator (Table 2b), the power improves even for moderate samples ($n \geq 40$), which is consistent with the efficiency gain with respect to OLS. We note that the power improvement is more pronounced as n increases rather than as T increases. Even with $T = 20$, the power improves substantially as n increases.

7 Conclusions

In this paper, we derived an inferential theory for testing for an unknown common change point in a cointegrated large panel regression. The model we considered contains unit specific regressors but it can also accommodate for the presence of common shocks, thereby allowing for strong cross-sectional dependence. We analyze both the cases of observable and unobservable common shocks, and we prove an FCLT for the partial sample estimators under various assumptions - *i.i.d.* data, strong cross dependence among idiosyncratic regressors, endogeneity.

For the purpose of testing, we study various transformations of Wald-type statistics, showing that under the null the limiting distributions are nuisance parameters free and depend only on the number of regressors. The distributions of the Wald-type statistics are functionals of the tied-down Bessel process, as found in the univariate stationary regression framework in Andrews (1993). The proposed tests are shown to have nontrivial power versus sequences of local alternatives of order $O_p\left(1/\sqrt{n}T\right)$; the term $1/\sqrt{n}$ shows the usefulness of the panel approach. Although the tests are designed for the case of one abrupt change common to all units, we show that the tests have power versus smooth, transition-type alternatives (similarly to Andrews, 1993) and also versus heterogeneous changepoints. Monte Carlo evidence shows that tests have the correct size and good power properties, the power gain being substantial as T increases and more moderate for increasing sizes of n, consistent with the $\sqrt{n}T$ asymptotics. However, when only some units have a break, our results show that the performance of tests actually becomes worse than in the one-unit-at-a-time case, as tests have power versus local alternatives shrinking at a rate $O_p\left(\sqrt{n}/T\right)$ in the extreme case of a finite number of units having a break.

An interesting development of the framework studied in this paper could be the extension to the multiple breaks case, following a similar approach as Kejriwal and Perron (2008, 2010). Also, our test statistics are based on taking the supremum of the Wald-type statistics over a trimmed interval. Alternatively, the Wald-type statistics could be normalised to take the supremum over the whole sample. This approach is discussed in various contributions (we refer to Csorgo and Horvath, 1997, for a comprehensive review) in a time series setting; it would be interesting to extend it to a panel setting, analysing the role of $n \to \infty$. This is an exciting research agenda for future work.

Acknowledgement

This is a revised version of an earlier article entitle "Modelling and Testing for Structural Breaks in Panel Cointegration Models with Common and Idiosyncratic Stochastic Trends". We would like to thank participants at the International Conferences "Common Features in London" (Cass Business School, 16-17 December 2004), "New York Camp Econometrics I" (Saratoga Springs, February 4-5, 2006) and "Breaks and Persistence in Econometrics" (Cass Business School, 11-12 December 2006), "Factor Models in Economics and Finance" (Cass Business School, 4-5 December 2009) , and econometrics seminars at Ohio State University, Academia Sinica, Bangor Business School, and Syracuse University for helpful comments. We benefited from discussions with Badi Baltagi, Robert De Jong, Long-Fei Lee, Zongwu Cai and Yupin Hu. The usual disclaimer applies.

References

[1] Andrews, D. W. K., 1993. Tests for parameter instability and structural changes with unknown change point. Econometrica 61, 821-856.

[2] Andrews, D. W. K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383-1414.

[3] Bai, J., 2004. Estimating cross-section common stochastic trends in nonstationary panel data. Journal of Econometrics 122, 137-183.

[4] Bai, J., 2010. Common breaks in means and variances for panel data. Journal of Econometrics 157, 78-92.

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

[6] Bai, J., Lumsdaine, R. L., Stock, J. H., 1998. Testing for and dating common breaks in multivariate time series. Review of Economic Studies 65, 395-432.

[7] Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47-78.

[8] Baillie, R. T., Bollerslev, T., 1989. Common stochastic trends in a system of exchange rates. The Journal of Finance 44, 167-181.

[9] Banerjee, A., Urga, G., 2005. Modelling structural breaks, long memory and stock market volatility: an overview. Journal of Econometrics 129, 1-34.

[10] Breitung, J., Eickmeier, S., 2010. Testing for structural breaks in dynamic factor models. Journal of Econometrics (article in press).

[11] Chiang, M-H., McCoskey, S., Kao, C., 2002. Panel data tests for a common break point in a cointegrated regression: asymptotic and small sample results. Mimeo, Center for Policy Research, Syracuse University.

[12] Csorgo, M., Horvath, L., 1997. Limit Theorems in Change-Point Analysis. Wiley, Chichester.

[13] Das, S. R., Duffie, D., Kapadia, N., Saita, L., 2007. Common failings: how corporate defaults are correlated. The Journal of Finance 57, 93-117.

[14] Dynarski, M., Sheffrin, S. M., 1985. Housing purchases and transitory income: a study with panel data. The Review of Economics and Statistics 67, 195-204.

[15] Feng, Q., Kao, C., Lazarova, S., 2008. Estimation and identification of change points in panel models. Mimeo, Center for Policy Research, Syracuse University.

[16] Geroski, P., Lazarova, S., Urga, G., Walters, C. F., 2002. Are differences in firm size transitory or permanent? Journal of Applied Econometrics 18, 47-59.

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

[18] Kapetanios, G., Pesaran, M. H., Yamagata, T., 2011. Panels with nonstationary multifactor error structures. Journal of Econometrics 160, 326-348.

[19] Kejriwal, M., Perron, P., 2008. The limiting distribution of the estimates in cointegrated regressions models with multiple structural changes. Journal of Econometrics 146, 59-73.

[20] Kejriwal, M., Perron, P., 2010. Testing for multiple structural changes in cointegrated regression models. Journal of Business and Economic Statistics 28, 503-522.

[21] Kim, K., 2010a. Estimating a common deterministic time trend break in large panels with cross sectional dependence. Mimeo. Department of Economics, University of Virginia.

[22] Kim, K., 2010b. Common local breaks in time trends for large panel data. Mimeo. Department of Economics, University of Virginia,

[23] Liu, W., Maynard, A., 2005. Testing forward rate unbiasedness allowing for persistent regressors. Journal of Empirical Finance 12, 613-628.

[24] Park, J. Y., Phillips, P. C. B., 1988. Statistical inference in regressions with integrated processes: part 1. Econometric Theory 4, 468-497.

[25] Park, J. Y., Phillips, P. C. B., 1999. Asymptotics for nonlinear transformations of integrated time series. Econometric Theory 15, 269, 298.

[26] Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis. Econometrica 57, 1361-1401

[27] Perron, P., 2006. Dealing with breaks," Chapter 8 in the Palgrave Handbook of Econometrics: Vol. 1 Econometric Theory edited by T.C. Mills and K. Patterson. Palgrave Macmillan: Basingstoke.

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

[29] Pesaran, M. H., 2007. A pair-wise Approach to Testing for Output and Growth Convergence," *Journal of Econometrics*, 138, 312-355.

[30] Phillips, P. C. B., Durlauf, S. N., 1986. Multiple time series regression with integrated processes," *Review of Economic Studies*, 53, 473-95.

[31] Phillips, P. C. B., Moon, H. R., 1999. Linear regression limit theory for nonstationary panel data. Econometrica 67, 1057-1112.

[32] Phillips, P. C. B., Solo, V., 1992. Asymptotics for linear processes. Annals of Statistics 20, 971–1001.

[33] Phillips, P. C. B., and Sul, D. (2007), "Transition Modeling and Econometric Convergence Tests," *Econometrica*, 75, 1771-1855.

[34] Qu, Z., Perron, P., 2007. Estimating and testing structural changes in multivariate regressions. Econometrica 75, 459-502.

[35] Rappoport, P., Reichlin, L., 1989. Segmented trends and non-stationary time series. Economic Journal 99, 168-77

[36] Stock, J. H., Watson, M. W., 1999. Forecasting inflation. Journal of Monetary Economics 44, 293-335.

[37] Stock, J. H., Watson, M. W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167-1179.

[38] Stock, J. H., Watson, M. W., 2005. Implications of dynamic factor models for VAR analysis. NBER Working Papers 11467, National Bureau of Economic Research, Inc.

[39] Sutton, J., 1997.Gibrat's legacy. Journal of Economic Literature 35, 40-59.

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

Panel A: Size for $SupW$

$n\backslash T$	20	40	60	120	240	480
20	0.0175	0.0375	0.0372	0.0694	0.0587	0.0519
40	0.0145	0.0236	0.0248	0.0462	0.0514	0.0604
60	0.0149	0.0260	0.0340	0.0337	0.0397	0.0550
120	0.0151	0.0287	0.0346	0.0373	0.0470	0.0561
240	0.0172	0.0309	0.0306	0.0360	0.0480	0.0454
480	0.0212	0.0285	0.0351	0.0349	0.0501	0.0560

Panel B: Size for $AveW$

$n\backslash T$	20	40	60	120	240	480
20	0.0267	0.0375	0.0350	0.0534	0.0407	0.0350
40	0.0258	0.0267	0.0242	0.0342	0.0349	0.0403
60	0.0220	0.0273	0.0312	0.0299	0.0265	0.0339
120	0.0238	0.0298	0.0315	0.0325	0.0306	0.0354
240	0.0241	0.0333	0.0330	0.0311	0.0367	0.0314
480	0.0312	0.0300	0.0307	0.0298	0.0375	0.0349

Panel C: Size for $ExpW$

$n\backslash T$	20	40	60	120	240	480
20	0.0306	0.0472	0.0455	0.0724	0.0537	0.0458
40	0.0272	0.0325	0.0534	0.0475	0.0461	0.0525
60	0.0241	0.0353	0.0392	0.0337	0.0352	0.0455
120	0.0269	0.0388	0.0428	0.0396	0.0411	0.0473
240	0.0289	0.0427	0.0403	0.0362	0.0453	0.0405
480	0.0370	0.0403	0.0415	0.0377	0.0489	0.0460

Table 1a. Empirical null rejection frequencies at 5% nominal level for tests based on $SupW$, $AveW$ and $ExpW$. Data are generated as i.i.d.. The estimation method for θ is OLS.

Panel A. Power for $SupW$

$n\backslash T$	20	40	60	120	240	480
20	0.0715	0.2085	0.4723	0.7962	0.9972	1.0000
40	0.0850	0.2332	0.5129	0.9467	1.0000	1.0000
60	0.0932	0.3281	0.6545	0.9837	1.0000	1.0000
120	0.1340	0.5551	0.8512	0.9999	1.0000	1.0000
240	0.2545	0.9640	0.9697	1.0000	1.0000	1.0000
480	0.4195	0.9327	0.9996	1.0000	1.0000	1.0000

Panel B. Power for $AveW$

$n\backslash T$	20	40	60	120	240	480
20	0.0859	0.2170	0.4477	0.8364	0.9983	1.0000
40	0.1055	0.2699	0.5523	0.9655	1.0000	1.0000
60	0.1172	0.3693	0.6917	0.9927	1.0000	1.0000
120	0.1686	0.6003	0.8921	1.0000	1.0000	1.0000
240	0.2991	0.8142	0.9877	1.0000	1.0000	1.0000
480	0.4700	0.9637	1.0000	1.0000	1.0000	1.0000

Panel C. Power for $ExeW$

$n\backslash T$	20	40	60	120	240	480
20	0.1009	0.2448	0.4993	0.8321	0.9980	1.0000
40	0.1224	0.2838	0.5628	0.9620	1.0000	1.0000
60	0.1329	0.3874	0.7006	0.9910	1.0000	1.0000
120	0.1806	0.6151	0.8899	1.0000	1.0000	1.0000
240	0.3168	0.8131	0.9850	1.0000	1.0000	1.0000
480	0.4894	0.9599	0.9999	1.0000	1.0000	1.0000

Table 2a. Empirical power at 5% nominal level for tests based on $SupW$, $AveW$ and $ExpW$. Data are generated as i.i.d.. The estimation method for θ is OLS.

Panel A: Size for $SupW$

$n \backslash T$	20	40	60	120	240	480
20	0.021	0.030	0.054	0.113	0.180	0.043
40	0.028	0.057	0.042	0.037	0.052	0.066
60	0.061	0.038	0.035	0.080	0.043	0.043
120	0.031	0.038	0.037	0.040	0.054	0.055
240	0.046	0.032	0.042	0.038	0.053	0.043
480	0.032	0.034	0.043	0.043	0.042	0.037

Panel B: Size for $AveW$

$n \backslash T$	20	40	60	120	240	480
20	0.030	0.033	0.045	0.070	0.101	0.024
40	0.039	0.056	0.039	0.027	0.033	0.038
60	0.068	0.037	0.035	0.056	0.028	0.026
120	0.042	0.039	0.035	0.031	0.032	0.033
240	0.056	0.037	0.039	0.033	0.035	0.026
480	0.044	0.037	0.041	0.035	0.029	0.026

Panel C: Size for $ExpW$

$n \backslash T$	20	40	60	120	240	480
20	0.035	0.041	0.066	0.114	0.168	0.039
40	0.048	0.071	0.054	0.039	0.057	0.058
60	0.088	0.049	0.043	0.084	0.041	0.038
120	0.052	0.050	0.045	0.041	0.050	0.048
240	0.071	0.047	0.052	0.043	0.052	0.039
480	0.055	0.048	0.051	0.047	0.040	0.034

Table 1b. Empirical null rejection frequencies at 5% nominal level for tests based on $SupW$, $AveW$ and $ExpW$. Data are generated according to (34). The estimation method for θ is FMOLS.

Panel A. Power for $SupW$

$n\backslash T$	20	40	60	120	240	480
20	0.064	0.144	0.324	0.763	0.974	1.000
40	0.086	0.319	0.555	0.933	0.999	1.000
60	0.156	0.405	0.692	0.983	1.000	1.000
120	0.207	0.618	0.886	1.000	1.000	1.000
240	0.365	0.809	0.987	1.000	1.000	1.000
480	0.536	0.961	1.000	1.000	1.000	1.000

Panel B. Power for $AveW$

$n\backslash T$	20	40	60	120	240	480
20	0.081	0.182	0.372	0.804	0.985	1.000
40	0.107	0.342	0.604	0.959	1.000	1.000
60	0.173	0.442	0.742	0.992	1.000	1.000
120	0.234	0.662	0.925	1.000	1.000	1.000
240	0.398	0.864	0.994	1.000	1.000	1.000
480	0.569	0.982	1.000	1.000	1.000	1.000

Panel C. Power for $ExeW$

$n\backslash T$	20	40	60	120	240	480
20	0.097	0.193	0.383	0.807	0.985	1.000
40	0.126	0.373	0.614	0.955	1.000	1.000
60	0.204	0.467	0.745	0.992	1.000	1.000
120	0.265	0.678	0.922	1.000	1.000	1.000
240	0.435	0.864	0.993	1.000	1.000	1.000
480	0.606	0.980	1.000	1.000	1.000	1.000

Table 2b. Empirical power at 5% nominal level for tests based on $SupW$, $AveW$ and $ExpW$. Data are generated according to (34). The estimation method for θ is FMOLS.

Appendix A: Preliminary Lemmas

Henceforth, we use the following notation: $C_{nT} = \min\{\sqrt{n}, T\}$, $\bar{F}^0 = T^{-1}\sum_{t=1}^{T} F_t$, and $\bar{F} = T^{-1}\sum_{t=1}^{T} \widehat{F}_t$.

Lemma A.1 *Under Assumptions 1 and 2, as $(n, T) \to \infty$ and for all $r \in (0, 1)$*

(a) $T^{-1}\sum_{t=1}^{\lfloor Tr \rfloor} \left\| \widehat{F}_t - F_t \right\|^2 = O_p\left(C_{nT}^{-2}\right)$,

(b) $T^{-1}\sum_{t=1}^{\lfloor Tr \rfloor} \left\| \hat{w}_t - w_t \right\|^2 = O_p\left(C_{nT}^{-2}\right)$,

(c) $T^{-1}\sum_{t=1}^{\lfloor Tr \rfloor} w_t'\left(\hat{w}_t - w_t\right) = O_p\left(C_{nT}^{-1}\right)$.

Proof. Without loss of generality, we prove the Lemma for $r = 1$. Part (a) is taken from Lemma 1 in Bai (2004). Consider part (b):

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} \left\| \hat{w}_t - w_t \right\|^2 &= \frac{1}{T}\sum_{t=1}^{T} \left\| \left(\widehat{F}_t - F_t\right) + \left(\bar{F} - \bar{F}^0\right) \right\|^2 \\
&\leq \frac{2}{T}\sum_{t=1}^{T} \left[\left\| \widehat{F}_t - F_t \right\|^2 + \left\| \bar{F} - \bar{F}^0 \right\|^2 \right] = I + II.
\end{aligned}$$

Part (a) yields $I = O_p\left(C_{nT}^{-2}\right)$; as far as II is concerned, $\left\| \bar{F} - \bar{F}^0 \right\|^2 = \left\| T^{-1}\sum_{t=1}^{T}\left(\widehat{F}_t - F_t\right) \right\|^2 \leq \left(T^{-1}\sum_{t=1}^{T} \left\| \widehat{F}_t - F_t \right\|^2 \right) = O_p\left(C_{nT}^{-2}\right)$. Therefore $II = O_p\left(C_{nT}^{-2}\right)$, and thus $T^{-1}\sum_{t=1}^{T} \left\| \hat{w}_t - w_t \right\|^2 = O_p\left(C_{nT}^{-2}\right)$. To prove part (c), note

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} w_t'\left(\hat{w}_t - w_t\right) &= \frac{1}{T}\sum_{t=1}^{T} \left(F_t - \bar{F}\right)' \left[\left(\widehat{F}_t - F_t\right) + \left(\bar{F} - \bar{F}^0\right) \right] \\
&= \frac{1}{T}\sum_{t=1}^{T} F_t\left(\widehat{F}_t - F_t\right) + \left(\frac{1}{T}\sum_{t=1}^{T} w_t'\right)\left(\bar{F} - \bar{F}^0\right) \\
&\quad -\bar{F}'\frac{1}{T}\sum_{t=1}^{T}\left(\widehat{F}_t - F_t\right) \\
&= I + II + III.
\end{aligned}$$

Lemma B.4(i) in Bai (2004) ensures that $I = O_p\left(C_{nT}^{-1}\right)$. As far as II is concerned, we have $II = \left[T^{-1}\sum_{t=1}^{T} w_t\right]' \left[T^{-1}\sum_{t=1}^{T}\left(\widehat{F}_t - F_t\right)\right] \leq \left[T^{-1}\sum_{t=1}^{T} w_t\right]' \left[T^{-1}\sum_{t=1}^{T}\left\|\widehat{F}_t - F_t\right\|\right]$; Assumption 1 entails $T^{-1}\sum_{t=1}^{T} w_t = O_p\left(\sqrt{T}\right)$. From Lemma B.4(iii) in Bai (2004), $T^{-1}\sum_{t=1}^{T}\left\|\widehat{F}_t - F_t\right\| = O_p\left(T^{-1/2}C_{nT}^{-1}\right)$; thus, $II = O_p\left(C_{nT}^{-1}\right)$. As far as III is concerned, note $III = \left[T^{-1}\sum_{t=1}^{T} F_t\right]' \left[T^{-1}\sum_{t=1}^{T}\left(\widehat{F}_t - F_t\right)\right]$, and therefore $III = O_p\left(C_{nT}^{-1}\right)$ similarly to II. ■

Lemma A.2 *Under Assumptions 1 and 2,* as $(n, T) \longrightarrow \infty$ *and for all* $r \in (0, 1)$.

(a) $T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_t\hat{w}_t' = T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor}w_tw_t' + O_p\left(T^{-1/2}C_{nT}^{-1}\right)$, *with* $T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor}w_tw_t' = O_p(1)$;

(b) $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_tu_{it} = n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}w_tu_{it} + O_p\left(C_{nT}^{-1}\right)$;

(c) $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_t\left(F_t-\widehat{F}_t\right) = \sqrt{n}T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}w_t'\left(F_t-\widehat{F}_t\right) + O_p\left(\sqrt{n}C_{nT}^{-2}\right)$, *with* $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}w_t'\left(F_t-\widehat{F}_t\right) = O_p\left(\sqrt{n}C_{nT}^{-2}\right)$;

(d) $n^{-1/2}T^{-2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}x_{it}\hat{w}_t' = n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}x_{it}w_t' + n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}x_{it}\left(\hat{w}_t-w_t\right)' = O_p(1) + O_p\left(T^{-1}C_{nT}^{-1}\right)$.

Proof. We prove the Lemma for $r=1$. For part (a), note that

$$\begin{aligned}
\frac{1}{T^2}\sum_{t=1}^{T}\hat{w}_t\hat{w}_t' &= \frac{1}{T^2}\sum_{t=1}^{T}\left(w_t+\hat{w}_t-w_t\right)\left(w_t+\hat{w}_t-w_t\right)' \\
&= \frac{1}{T^2}\sum_{t=1}^{T}w_tw_t' + \frac{1}{T^2}\sum_{t=1}^{T}w_t\left(\hat{w}_t-w_t\right)' \\
&\quad + \frac{1}{T^2}\sum_{t=1}^{T}\left(\hat{w}_t-w_t\right)w_t' + \frac{1}{T^2}\sum_{t=1}^{T}\left(\hat{w}_t-w_t\right)\left(\hat{w}_t-w_t\right)' \\
&= I + II + III + IV.
\end{aligned}$$

Assumption 1 ensures that $I = O_p(1)$. As far as II and III are concerned, application of the Cauchy-Schwartz inequality and of Lemma A.1(a) ensures that they are bounded by

$$\begin{aligned}
II &\leq \frac{1}{T^2}\left(\sum_{t=1}^{T}\|w_t\|^2\right)^{1/2}\left(\sum_{t=1}^{T}\|\hat{w}_t-w_t\|^2\right)^{1/2} \\
&= \frac{1}{T^2}O_p(T)\,O_p\left(\frac{\sqrt{T}}{C_{nT}}\right) = O_p\left(\frac{1}{\sqrt{T}C_{nT}}\right).
\end{aligned}$$

As far as IV is concerned, using Lemma A.1(b) we have $\left\|T^{-2}\sum_{t=1}^{T}\left(\hat{w}_t-w_t\right)\left(\hat{w}_t-w_t\right)'\right\| \leq T^{-2}\sum_{t=1}^{T}\|\hat{w}_t-w_t\|^2 = O_p\left(T^{-1}C_{nT}^{-2}\right)$. Hence, $T^{-2}\sum_{t=1}^{T}\hat{w}_t\hat{w}_t' = T^{-2}\sum_{t=1}^{T}w_tw_t' + O_p\left(T^{-1/2}C_{nT}^{-1}\right) + O_p\left(T^{-1}C_{nT}^{-2}\right)$.

Turning to part (b), $n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{w}_tu_{it} = n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}w_tu_{it} + n^{-1/2}T^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{w}_t-w_t\right)u_{it}$. The Cauchy-Schwartz inequality and Lemma A.1(b) entail

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\hat{w}_t-w_t\right)u_{it} \leq \left(\frac{1}{T}\sum_{t=1}^{T}\|\hat{w}_t-w_t\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{i=1}^{n}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_{it}\right\|^2\right)^{1/2} = O_p\left(\frac{1}{C_{nT}}\right). \tag{35}$$

To prove (c) consider

$$\frac{\sqrt{n}}{T}\sum_{t=1}^{T}\hat{w}_t'\left(F_t-\tilde{F}_t\right)=\frac{\sqrt{n}}{T}\sum_{t=1}^{T}w_t'\left(F_t-\tilde{F}_t\right)+\frac{\sqrt{n}}{T}\sum_{t=1}^{T}\left(\hat{w}_t-w_t\right)'\left(F_t-\widehat{F}_t\right)=I+II.$$

From Lemma A.1(c) it follows that $I=O_p\left(\sqrt{n}C_{nT}^{-1}\right)$. As far as II is concerned,

$$\begin{aligned} II &\leq \sqrt{n}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{w}_t-w_t\right\|^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|F_t-\widehat{F}_t\right\|^2\right)^{1/2} \\ &= \sqrt{n}O_p\left(\frac{1}{C_{nT}}\right)O_p\left(\frac{1}{C_{nT}}\right)=O_p\left(\frac{\sqrt{n}}{C_{nT}^2}\right). \end{aligned}$$

To prove part (d), let $n^{-1/2}T^{-2}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{x}_{it}w_t'+n^{-1/2}T^{-2}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{x}_{it}\left(\hat{w}_t-w_t\right)'=I+II$. Proposition 1 (proved below) ensures that $I=O_p(1)$. As far as II is concerned, note that the sequence $\widetilde{x}_{it}\left(\hat{w}_t-w_t\right)'$ has zero mean and is conditionally *i.i.d.* across i; thus, a CLT yields $\sum_{i=1}^{n}\widetilde{x}_{it}\left(\hat{w}_t-w_t\right)'=O_p\left(\sqrt{n}\right)$; the rest of the proof, to show that $\sum_{t=1}^{T}\widetilde{x}_{it}\left(\hat{w}_t-w_t\right)'=O_p\left(TC_{nT}^{-1}\right)$, is similar to that of Lemma B.4(i) in Bai (2004), and thus passages are omitted. ■

Lemma A.3 *Let $x_i=x_{i-1}+u_i^x$, with u_i^x a zero mean, unit variance MDS, and let u_i be an MDS independent of u_i^x, with unit variance. Then, as $T\to\infty$ for $a<c<b<d$*

$$\begin{aligned} &E\left\{\frac{1}{T}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\left[x_i-\frac{1}{T(b-a)}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}x_i\right]u_i\times\frac{1}{T}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}\left[x_i-\frac{1}{T(d-c)}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}x_i\right]u_i\right\} \\ &= \frac{(d-a)(b-c)}{(b-a)(d-c)}\left[\frac{1}{6}(b-c)^2\right]. \end{aligned} \tag{36}$$

Proof. Let $\bar{x}_i$ be the demeaned version of x_i and note

$$\begin{aligned} &E\left[\frac{1}{T^2}\left(\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\bar{x}_iu_i\right)\left(\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}\bar{x}_iu_i\right)\right]=E\left[\frac{1}{T^2}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}\bar{x}_i\bar{x}_ju_iu_j\right] \\ &= E\left[\frac{1}{T^2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(\bar{x}_iu_i\right)^2\right]=E\left(\frac{1}{T^2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\bar{x}_i^2\right) \end{aligned}$$

due to the independence of u_i and to $E\left(u_i^2\right)=1$. We have

$$
\begin{aligned}
& E\left(\frac{1}{T^2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\bar{x}_i^2\right)=E\left(\frac{1}{T^2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}x_i^2\right)\\
& +E\left[\frac{1}{T^2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(\frac{1}{T\left(b-a\right)}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}x_i\right)\left(\frac{1}{T\left(d-c\right)}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}x_i\right)\right]\\
& -E\left[\frac{1}{T^2}\frac{1}{T\left(b-a\right)}\left(\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}x_i\right)\left(\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}x_i\right)\right]\\
& -E\left[\frac{1}{T^2}\frac{1}{T\left(d-c\right)}\left(\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}x_i\right)\left(\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}x_i\right)\right]\\
= & \; I+II-III-IV.
\end{aligned}
$$

As far as I is concerned, it holds that, as $T\to\infty$, $I=T^{-2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}E\left(x_i^2\right)=T^{-2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}i\to\int_c^b udu=\frac{1}{2}\left(b^2-c^2\right)$. Also

$$
\begin{aligned}
II &= \frac{b-c}{\left(b-a\right)\left(d-c\right)}\frac{1}{T^3}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}E\left(x_ix_j\right)\\
&= \frac{b-c}{\left(b-a\right)\left(d-c\right)}\frac{1}{T^3}\left[\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tc\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}E\left(x_ix_j\right)\right.\\
&\quad \left.+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}E\left(x_ix_j\right)+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tb\rfloor}^{\lfloor Td\rfloor}E\left(x_ix_j\right)\right]\\
&= \frac{b-c}{\left(b-a\right)\left(d-c\right)}\frac{1}{T^3}\left[T\left(d-c\right)\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tc\rfloor}i\right.\\
&\quad \left.+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(i\wedge j\right)+T\left(d-b\right)\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}i\right]\\
&\to \frac{b-c}{\left(b-a\right)\left(d-c\right)}\left(d-c\right)\int_a^c udu+\int_c^b\int_c^b\left(u\wedge v\right)dudv+\left(d-b\right)\int_c^b udu\\
&= \frac{1}{2}\frac{\left(b-c\right)\left(d-c\right)}{\left(b-a\right)\left(d-c\right)}\left(c^2-a^2\right)+\frac{1}{3}\frac{\left(b-c\right)}{\left(b-a\right)\left(d-c\right)}\left(b^3-c^3\right)\\
&\quad +\frac{1}{2}\frac{\left(b-c\right)\left(d-b\right)}{\left(b-a\right)\left(d-c\right)}\left(b^2-a^2\right).
\end{aligned}
$$

Similar passages yield

$$\begin{aligned}
III &= \frac{1}{b-a}\frac{1}{T^3}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}E\left(x_i x_j\right) \\
&= \frac{1}{b-a}\frac{1}{T^3}\left[\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tc\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}E\left(x_i x_j\right)+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}E\left(x_i x_j\right)\right] \\
&= \frac{1}{b-a}\frac{1}{T^3}\left[T\left(b-c\right)\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tc\rfloor}i+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(i\wedge j\right)\right] \\
&\to \frac{1}{b-a}\left[\int_a^c udu+\int_c^b\int_c^b\left(u\wedge v\right)dudv\right] \\
&= \frac{1}{2}\frac{b-c}{b-a}\left(c^2-a^2\right)+\frac{1}{3}\frac{b^3-c^3}{b-a},
\end{aligned}$$

$$\begin{aligned}
IV &= \frac{1}{d-c}\frac{1}{T^3}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}E\left(x_i x_j\right) \\
&= \frac{1}{d-c}\frac{1}{T^3}\left[\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}E\left(x_i x_j\right)+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tb\rfloor}^{\lfloor Td\rfloor}E\left(x_i x_j\right)\right] \\
&= \frac{1}{d-c}\frac{1}{T^3}\left[\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(i\wedge j\right)+T\left(d-b\right)\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}i\right] \\
&\to \frac{1}{d-c}\left[\int_c^b\int_c^b\left(u\wedge v\right)dudv+\left(d-b\right)\int_c^b udu\right] \\
&= \frac{1}{3}\frac{b^3-c^3}{d-c}+\frac{1}{2}\frac{d-b}{d-c}\left(b^2-c^2\right).
\end{aligned}$$

Putting everything together, (36) follows.

Appendix B: Proofs and Lemmas

Proof of Proposition 1. Define $\varsigma_{i\lfloor Tr\rfloor} = T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor} w_t \widetilde{x}_{it}'$. Then, the BN decompositions (7) and (8) in Assumption 1 yield

$$\varsigma_{i\lfloor Tr\rfloor} = \frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor} w_t^* \widetilde{x}_{it}^{*\prime} + R_{i\lfloor Tr\rfloor} = \varsigma_{i\lfloor Tr\rfloor}^* + R_{i\lfloor Tr\rfloor},$$

with $R_{i\lfloor Tr\rfloor} = O_p\left(T^{-1/2}\right)$ due to (11). Let C be the σ-field associated with $\{w_t\}_{t=1}^T$ and consider $\varsigma_{i\lfloor Tr\rfloor}^*$; conditioning on C, Assumption 1 ensures that $\varsigma_{i\lfloor Tr\rfloor}^*$ is *i.i.d.* across i. Also, $E\left[\varsigma_{i\lfloor Tr\rfloor}^* \middle| C\right] = T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor} w_t^* E\left(\widetilde{x}_{it}^{*\prime}\right) = 0$ for all r, since $\widetilde{x}_{it}^{*\prime}$ has zero mean by construction. Letting I_i denote the σ-field defined as the union of C and $\{\varsigma_{1T}^*, ..., \varsigma_{iT}^*\}$, we have $E\left[\varsigma_{i\lfloor Tr\rfloor}^* \middle| I_{i-1}\right] = E\left[\varsigma_{i\lfloor Tr\rfloor}^* \middle| C\right] = 0$. Therefore $\left\{\varsigma_{i\lfloor Tr\rfloor}^*, I_i\right\}$ is a zero mean MDS for all r. A Liapunov condition can be proved such that for some $\delta > 0$, $E\left[\left\|\varsigma_{i\lfloor Tr\rfloor}^*\right\|^{2+\delta} \middle| C\right] < \infty$. This holds because

$$\begin{aligned}
\left\|\varsigma_{i\lfloor Tr\rfloor}^*\right\|^{2+\delta} &= \left\|\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor} w_t^* \widetilde{x}_{it}^{*\prime}\right\|^{2+\delta} \leq M_\delta \frac{1}{T^{2(2+\delta)}}\sum_{t=1}^{\lfloor Tr\rfloor} \|w_t^*\|^{2+\delta} \left\|\widetilde{x}_{it}^{*\prime}\right\|^{2+\delta} \\
&\leq M_\delta \left[\frac{1}{T^{2(2+\delta)}}\sum_{t=1}^{\lfloor Tr\rfloor} \|w_t^*\|^{2(2+\delta)}\right]^{1/2} \left[\frac{1}{T^{2(2+\delta)}}\sum_{t=1}^{\lfloor Tr\rfloor} \left\|\widetilde{x}_{it}^{*\prime}\right\|^{2(2+\delta)}\right]^{1/2} = O_p(1),
\end{aligned}$$

where the last equality holds in light of Assumption 1(a) and Theorem 5.2 in Park and Phillips (1999). Thus, $\left\|\varsigma_{i\lfloor Tr\rfloor}^*\right\|^{2+\delta}$ is bounded by $O_p(1)$ for all i as $T \to \infty$, and therefore an MDS CLT can be applied such that, for all r, $n^{-1/2}\sum_{i=1}^n \varsigma_{i\lfloor Tr\rfloor}^* = O_p(1)$ as $(n, T) \to \infty$. Thus, as far as $\varsigma_{i\lfloor Tr\rfloor}$ is concerned, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \varsigma_{i\lfloor Tr\rfloor} = \frac{1}{\sqrt{n}}\sum_{i=1}^n \varsigma_{i\lfloor Tr\rfloor}^* + \frac{1}{\sqrt{n}}\sum_{i=1}^n R_{i\lfloor Tr\rfloor} = O_p(1) + O_p\left(\sqrt{\frac{n}{T}}\right).$$

Thus, as $(n, T) \to \infty$ under $n/T \to 0$, $n^{-1/2}\sum_{i=1}^n\sum_{t=1}^T w_t \widetilde{x}_{it}' = O_p(1)$.

Lemma B.1 *Under Assumptions 1-2, as $(n, T) \longrightarrow \infty$ it holds that, for all $r \in (0, 1)$*

(a)

$$\frac{1}{nT^2}\sum_{i=1}^n\sum_{t=1}^{\lfloor Tr\rfloor} W_{it}W_{it}' \xrightarrow{d} \begin{bmatrix} \int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon' & 0 \\ 0 & \frac{1}{6}r^2\Omega_\epsilon \end{bmatrix},$$

(b)

$$\frac{1}{nT^2}\sum_{i=1}^n\sum_{t=1}^{\lfloor Tr\rfloor} \widehat{W}_{it}\widehat{W}_{it}' \xrightarrow{d} \begin{bmatrix} \int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon' & 0 \\ 0 & \frac{1}{6}r^2\Omega_\epsilon \end{bmatrix},$$

Proof. Consider part (a). Proposition 1 implies that the off-diagonal terms of the matrix converge in probability to zero. Also, $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor} w_t w_t' = \frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor} w_t w_t'$, and a standard FCLT yields the result. Finally, standard joint LLN - see Phillips and Moon (1999) - yields $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor} \tilde{x}_{it}\tilde{x}_{it}' \xrightarrow{p} \frac{1}{6}r^2\Omega_\epsilon$. To prove part (b), note $n^{-1}T^{-2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor} \hat{w}_t\hat{w}_t' = T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor} w_t w_t' + O_p\left(T^{-1/2}C_{nT}^{-1}\right)$, where the last equality holds from Lemma A2.(a). The desired result holds as long as $(n,T)\to\infty$.

Lemma B.2 *Under Assumptions 1-2, as* $(n,T)\longrightarrow\infty$ *with* $\frac{n}{T}\to 0$ *it holds that, for all* $r\in(0,1)$

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\widehat{W}_{it}u_{it}\xrightarrow{d}\begin{bmatrix}\sigma_\zeta\left(\int_0^r \bar{B}_\varepsilon\bar{B}_\varepsilon'\right)^{1/2} & 0_{R\times p}\\ 0_{p\times R} & \sigma_u\Omega_\epsilon^{1/2}\end{bmatrix}\times\begin{bmatrix}B(r)\\ C(r;0)\end{bmatrix},$$

where $B(r)$ *and* $C(r;0)$ *are defined in Theorem 1.*

Proof. The Lemma is an FCLT for cointegrated panel data. To prove it, note first that the covariance matrix is diagonal in light of Proposition 1; thus, we show separately that

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_t v_{it}\xrightarrow{d}\sigma_\zeta\left(\int_0^r\bar{B}_\varepsilon\bar{B}_\varepsilon'\right)^{1/2}\times B(r), \tag{37}$$

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\tilde{x}_{it}v_{it}\xrightarrow{d}\sigma_u\Omega_\epsilon^{1/2}\times C(r;0). \tag{38}$$

Consider first (37), rewritten as

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_t v_{it}=\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_t\left[\sum_{i=1}^{n}\frac{1}{\sqrt{n}}\left(u_{it}+\beta'\tilde{Q}_{B,T}\lambda_i e_{it}\right)\right]+o_p(1);$$

using the BN decomposition, this is equivalent to

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_t\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(u_{it}+\beta'\tilde{Q}_{B,T}\lambda_i e_{it}\right)\right] &= \frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_t^*\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(u_{it}^*+\beta'\tilde{Q}_{B,T}\lambda_i e_{it}^*\right)\right]\\ &\quad+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}R_{wvi,\lfloor Tr\rfloor}+o_p(1).\end{aligned}$$

Let $u_{it}^* + \beta' \tilde{Q}_{B,T} \lambda_i e_{it}^* = v_{it}^*$; for all n, conditional on C, $\hat{w}_t^* \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n v_{it}^* \right)$ is a zero mean MDS with covariance matrix

$$
\begin{aligned}
& E \left[\frac{1}{T} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t^* \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} v_{it}^* \right) \middle| C \right]^2 = \frac{1}{T^2} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t^* \hat{w}_t^{*\prime} E \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} v_{it}^* \right)^2 \\
= & \ \sigma_\zeta^2 \frac{1}{T^2} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t^* \hat{w}_t^{*\prime},
\end{aligned}
$$

so that, as $(n,T) \to \infty$, $E \left[T^{-1} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t^* \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n v_{it}^* \right) \middle| C \right]^2 = \sigma_\zeta^2 \int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon'$, in light of Lemma B.1. Thus, for all r, $T^{-1} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t^* \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n v_{it}^* \right) \xrightarrow{d} \sigma_\zeta \left(\int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon' \right)^{1/2} \times Z$, where Z is a standard normal of dimension R. Tightness follows if

$$
\sup_r \left| \frac{1}{T} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} v_{it} \right) - \frac{1}{T} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t^* \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} v_{it}^* \right) \right| = o_p(1);
$$

however, this is equivalent to $\sup_r \left| T^{-1} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n v_{it} \right) - T^{-1} \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_t^* \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n v_{it}^* \right) \right| + o_p(1) = \sup_r \left| n^{-1/2} \sum_{i=1}^n R_{wvi,\lfloor Tr \rfloor} \right|$. Neglecting the $o_p(1)$ term and using (11)

$$
\sup_r \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_{wvi,\lfloor Tr \rfloor} \right| \leq \sup_r \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left| R_{wvi,\lfloor Tr \rfloor} \right| \leq \sqrt{n} \sup_{i,r} \left| R_{wvi,\lfloor Tr \rfloor} \right| = \sqrt{n} O_p \left(\frac{1}{\sqrt{T}} \right),
$$

which vanishes as $(n,T) \to \infty$ under $\frac{n}{T} \to 0$.

As far as (38) is concerned, we have, applying Lemma A.2(d)

$$
\frac{1}{\sqrt{n}T} \sum_{i=1}^{n} \sum_{t=1}^{\lfloor Tr \rfloor} \tilde{x}_{it} v_{it} = \frac{1}{\sqrt{n}T} \sum_{i=1}^{n} \sum_{t=1}^{\lfloor Tr \rfloor} \tilde{x}_{it}^* u_{it}^* + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_{xui,\lfloor Tr \rfloor} + o_p(1).
$$

Let $\xi_{nt}^* = \frac{1}{\sqrt{n}} \sum_{i=1}^n \tilde{x}_{it}^* u_{it}^*$; ξ_{nt}^* is an MDS across t. It holds that

$$
\begin{aligned}
& E \left(\frac{1}{T} \sum_{t=1}^{\lfloor Tr \rfloor} \xi_{nt}^* \right)^2 = \frac{1}{T^2} \sum_{t=1}^{\lfloor Tr \rfloor} \sum_{s=1}^{\lfloor Tr \rfloor} E \left(\xi_{nt}^* \xi_{ns}^* \right) \\
= & \ \frac{1}{T^2} \sum_{t=1}^{\lfloor Tr \rfloor} E \left(\xi_{nt}^{*2} \right) = \sigma_u^2 \frac{1}{T^2} \sum_{t=1}^{\lfloor Tr \rfloor} \frac{1}{n} \sum_{i=1}^{n} E \left(x_{it}^* - \frac{1}{Tr} \sum_{t=1}^{\lfloor Tr \rfloor} x_{it}^* \right)^2 \\
= & \ \sigma_u^2 \frac{1}{T^2} \sum_{t=1}^{\lfloor Tr \rfloor} E \left(x_{it}^* - \frac{1}{Tr} \sum_{t=1}^{\lfloor Tr \rfloor} x_{it}^* \right)^2.
\end{aligned}
$$

Applying Lemma A.3 with $a = c = 0$ and $b = d = r$, it holds that $\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr \rfloor} E\left(x_{it}^* - \frac{1}{Tr}\sum_{t=1}^{\lfloor Tr \rfloor} x_{it}^*\right)^2$ $\to \Omega_\epsilon \frac{1}{6} r^2$. That $n \to \infty$ is only incidental to the main argument of the proof; thus, as $(n, T) \to \infty$, $E\left(\frac{1}{T}\sum_{t=1}^{\lfloor Tr \rfloor} \xi_{nt}^*\right)^2 = \frac{1}{6}\sigma_u^2 \Omega_\epsilon r^2$. In order to prove tightness, similar arguments as above entail

$$\begin{aligned}
&\sup_r \left| \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \tilde{x}_{it} u_{it} - \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \tilde{x}_{it}^* u_{it}^* \right| \\
= &\sup_r \left| \frac{1}{\sqrt{n}}\sum_{i=1}^{n} R_{xui,\lfloor Tr \rfloor} \right| + o_p(1) \leq \sqrt{n}\sup_{i,r}\left| R_{xui,\lfloor Tr \rfloor}\right| + o_p(1) = O_p\left(\sqrt{\frac{n}{T}}\right) + o_p(1);
\end{aligned}$$

thus, an FCLT holds as $(n, T) \to \infty$ under $\frac{n}{T} \to 0$. The covariance structure of $C(\cdot;\cdot)$ can be calculated using (36) in Lemma A.3, recalling that the long run variances of x_{it} and u_{it} are, respectively, Ω_ϵ and σ_u^2.

Proof of Theorem 1. The proof follows from Lemmas B.1 and B.2.

Proof of Theorem 2. To prove the theorem, note $\left(\hat{\theta}_{1\lfloor Tr \rfloor}^* - \hat{\theta}_{2\lfloor Tr \rfloor}^*\right) = \left(\hat{\theta}_{1\lfloor Tr \rfloor}^* - \theta\right)$ $-\left(\hat{\theta}_{2\lfloor Tr \rfloor}^* - \theta\right)$, and therefore

$$\sqrt{n}T\left(\hat{\theta}_{1\lfloor Tr \rfloor}^* - \hat{\theta}_{2\lfloor Tr \rfloor}^*\right) = \begin{bmatrix} I & -I \end{bmatrix} \begin{bmatrix} \sqrt{n}T\left(\hat{\theta}_{1\lfloor Tr \rfloor} - \theta\right) \\ \sqrt{n}T\left(\hat{\theta}_{2\lfloor Tr \rfloor} - \theta\right) \end{bmatrix},$$

where I is $(R+p) \times (R+p)$ identity matrix. Theorem 1 and the continuous mapping theorem entail, for any consistent estimators $\hat{\sigma}_u^2$ and $\hat{\sigma}_\zeta^2$

$$\begin{aligned}
\sqrt{n}T\left(\hat{\theta}_{1\lfloor Tr \rfloor}^* - \hat{\theta}_{2\lfloor Tr \rfloor}^*\right) &\xrightarrow{d} \begin{bmatrix} I & -I \end{bmatrix} \begin{bmatrix} \begin{pmatrix} \left(\int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} Z_1 \\ \frac{\sqrt{6}}{r^2}\Omega_\epsilon^{-1/2} C(r;0) \end{pmatrix} \\ \begin{pmatrix} \left(\int_r^1 \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} Z_2 \\ \frac{\sqrt{6}}{(1-r)^2}\Omega_\epsilon^{-1/2} C(1;r) \end{pmatrix} \end{bmatrix} \\
= &\begin{bmatrix} \left(\int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} Z_1 - \left(\int_r^1 \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1/2} Z_2 \\ \sqrt{6}\Omega_\epsilon^{-1/2}\left[\frac{C(r;0)}{r^2} - \frac{C(1;r)}{(1-r)^2}\right] \end{bmatrix} = \mathbf{v}(r).
\end{aligned}$$

Using Lemma B.1(b) yields

$$\begin{aligned}
&\begin{bmatrix} \left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \widehat{W}_{it}\widehat{W}_{it}'\right)^{-1} \\ +\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=\lfloor Tr \rfloor+1}^{T} \widehat{W}_{it}\widehat{W}_{it}'\right)^{-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \left(\int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1} + \left(\int_r^1 \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1} & 0 \\ 0 & \frac{6}{r^2}\Omega_\epsilon^{-1} + \frac{6}{(1-r)^2}\Omega_\epsilon^{-1} \end{bmatrix} \\
= &\ G(r).
\end{aligned}$$

Therefore, by the CMT and uniformly in r we have

$$\begin{aligned} W(\lfloor Tr \rfloor) &= \sqrt{n}T\left(\hat{\theta}^*_{1k} - \hat{\theta}^*_{2k}\right)' \left[\begin{array}{c} \left(\frac{1}{nT^2}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} \widehat{W}_{it}\widehat{W}_{it}'\right)^{-1} \\ +\left(\frac{1}{nT^2}\sum_{i=1}^n \sum_{t=\lfloor Tr \rfloor+1}^{T} \widehat{W}_{it}\widehat{W}_{it}'\right)^{-1} \end{array} \right]^{-1} \sqrt{n}T\left(\hat{\theta}^*_{1k} - \hat{\theta}^*_{2k}\right) \\ &\xrightarrow{d} \mathbf{v}(r)' [G(r)]^{-1} \mathbf{v}(r) \\ &= \mathbf{s}(r)' [\mathbf{M}(r)]^{-1} \mathbf{s}(r) \\ &\quad + \left[\frac{1}{r^2} + \frac{1}{(1-r)^2}\right]^{-1} \times \left[\frac{(1-r)^2 C(r;0) - r^2 C(1;r)}{r^2(1-r)^2}\right]' \left[\frac{(1-r)^2 C(r;0) - r^2 C(1;r)}{r^2(1-r)^2}\right] \\ &= I + II. \end{aligned}$$

Consider I. Conditioning on C, Theorem 1 yields, for fixed r, $\mathbf{s}(r)| C \sim \left[\left(\int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1} + \left(\int_r^1 \bar{B}_\varepsilon \bar{B}_\varepsilon'\right)^{-1}\right]^{1/2} \times N[0, I_R]$, so that, conditioning on C, $\mathbf{s}(r)' [\mathbf{M}(r)]^{-1} \mathbf{s}(r)\Big| C \sim \chi^2_R$. This result does not depend on C - i.e. it holds true for all the possible elements in the sigma-field C. Thus $\mathbf{s}(r)' [\mathbf{M}(r)]^{-1} \mathbf{s}(r) = Q_R(r) \sim \chi^2_R$. As far as II is concerned, normality and independence of the increments of $C(r;0)$ entail

$$(1-r)^2 C(r;0) - r^2 C(1;r) \sim N\left\{0, \left[(1-r)^4 r^2 + r^4 (1-r)^2\right] I_p\right\}.$$

Therefore, for every r

$$II \sim \frac{r^2(1-r)^2}{r^2 + (1-r)^2} \frac{\left\| N\left\{0, \left[(1-r)^4 r^2 + r^4 (1-r)^2\right] I_p\right\}\right\|}{r^4(1-r)^4} \sim \chi^2_p.$$

Hence, for fixed r, $W(\lfloor Tr \rfloor) \xrightarrow{d} Q_R(r) + Q_p(r)$; this proves equation (2). Independence of $Q_R(r)$ and $Q_p(r)$ follows from the fact that $\widehat{\beta}$ and $\widehat{\gamma}$ are asymptotically independent. This also proves that both $Q_p(r)$ and $Q_R(r)$ are nuisance parameters free.

Proof of Proposition 3. Consider $\hat{\sigma}^2_u$, defined as $\hat{\sigma}^2_u = \frac{1}{nT}\sum_{i=1}^n \sum_{t=1}^T \left(y_{it} - \bar{y}_i - \hat{\theta}'\widehat{W}_{it}\right)^2$. Consistency of $\hat{\theta}$ under H_0 entails $\hat{\sigma}^2_u \xrightarrow{p} \sigma^2_u$. As far as $\hat{\sigma}^2_\zeta$ is concerned, we have $\hat{\sigma}^2_\zeta = \hat{\sigma}^2_u + \hat{\beta}'\hat{\sigma}^2_\pi \hat{\beta}$; under H_0, $\hat{\sigma}^2_u = \sigma^2_u + o_p(1)$ and likewise $\hat{\beta} = \beta + o_p(1)$. Also, it holds that $\hat{\sigma}^2_\pi = \left(\frac{1}{T^2}\sum_{t=1}^T \hat{w}_t \hat{w}_t'\right) \left[\frac{1}{n}\sum_{i=1}^n \left(\frac{1}{T}\sum_{t=1}^T \hat{e}^2_{it}\right)\hat{\lambda}_i \hat{\lambda}_i'\right] \left(\frac{1}{T^2}\sum_{t=1}^T \hat{w}_t \hat{w}_t'\right)$. From Lemma 2.(a) it holds that $\frac{1}{T^2}\sum_{t=1}^T \hat{w}_t \hat{w}_t' = \frac{1}{T^2}\sum_{t=1}^T w_t w_t' + o_p(1)$. Since $\hat{\lambda}_i = \lambda_i + O_p\left(T^{-1}\right)$ (see Bai, 2004), $T^{-1}\sum_{t=1}^T \hat{e}^2_{it} \xrightarrow{p} \sigma^2_e$. Therefore, $\hat{\sigma}^2_\pi = \sigma^2_e \left(\tilde{Q}_B \Sigma_\Lambda \tilde{Q}_B'\right) + o_p(1)$, whence $\hat{\sigma}^2_\zeta \xrightarrow{p} \sigma^2_\zeta$.

Proof of Theorem 3. Under the local alternative $H_a^{(nT)}$ the model can be rewritten as

$$
\begin{aligned}
y_{it}^{(nT)} &= \alpha_i + \widehat{W}_{it}'\theta_t^{(nT)} + v_{it} \\
&= \alpha_i + \widehat{W}_{it}'\theta + \frac{\sqrt{n}}{m_n T}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) + v_{it}.
\end{aligned}
$$

The partial sample OLS estimate for θ is defined as $\hat{\theta}_{1k}^{(nT)} = \left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}y_{it}^{(nT)}$, and therefore

$$
\begin{aligned}
\hat{\theta}_{1k}^{(nT)} &= \left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\left[\alpha_i + \widehat{W}_{it}'\theta + \frac{\sqrt{n}}{m_n T}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) + v_{it}\right] \\
&= \theta + \left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\left\{\frac{\sqrt{n}}{m_n T}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) + v_{it}\right\}\right].
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\sqrt{n}T\left[\hat{\theta}_{1k}^{(nT)} - \theta\right] \\
&= \left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}v_{it} + \frac{1}{\sqrt{n}T}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\frac{\sqrt{n}}{m_n T}\widehat{W}_{it}'g\left(\frac{t}{T}\right)\right],
\end{aligned}
$$

and similarly for the second partial sample, $\hat{\theta}_{2k}^{(nT)}$

$$
\begin{aligned}
&\sqrt{n}T\left[\hat{\theta}_{2k}^{(nT)} - \theta\right] \\
&= \left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=\lfloor Tr \rfloor+1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=\lfloor Tr \rfloor+1}^{T}\widehat{W}_{it}v_{it} + \frac{1}{\sqrt{n}T}\sum_{i=1}^{m_n}\sum_{t=\lfloor Tr \rfloor+1}^{T}\widehat{W}_{it}\frac{\sqrt{n}}{m_n T}\widehat{W}_{it}g_i\left(\frac{t}{T}\right)\right].
\end{aligned}
$$

Combining these two results, it can be shown that

$$
\begin{aligned}
&\sqrt{n}T\left[\hat{\theta}_{1k}^{(nT)} - \hat{\theta}_{2k}^{(nT)}\right] \\
&= \sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr \rfloor} - \hat{\theta}_{2\lfloor Tr \rfloor}\right] + \left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right)^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\frac{\sqrt{n}}{m_n T}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right)\right] \\
&\quad - \left[\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=\lfloor Tr \rfloor+1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right)\right]^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{m_n}\sum_{t=\lfloor Tr \rfloor+1}^{T}\widehat{W}_{it}\frac{\sqrt{n}}{m_n T}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right)\right] \qquad (39) \\
&= I + II.
\end{aligned}
$$

Consider I and II. It holds that

$$
\begin{aligned}
& \frac{1}{\sqrt{n}T}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\widehat{W}_{it}\frac{\sqrt{n}}{m_n T}\widehat{W}_{it}' g_i\left(\frac{t}{T}\right) \\
= & \frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left[W_{it}+\left(\widehat{W}_{it}-W_{it}\right)\right]\left[W_{it}+\left(\widehat{W}_{it}-W_{it}\right)\right]' g_i\left(\frac{t}{T}\right) \\
= & \frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}W_{it}W_{it}' g_i\left(\frac{t}{T}\right)+\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}W_{it}\left(\widehat{W}_{it}-W_{it}\right)' g_i\left(\frac{t}{T}\right) \\
& +\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\widehat{W}_{it}-W_{it}\right)W_{it}' g_i\left(\frac{t}{T}\right)+\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\widehat{W}_{it}-W_{it}\right)\left(\widehat{W}_{it}-W_{it}\right)' g_i\left(\frac{t}{T}\right) \\
= & a+b+c+d.
\end{aligned}
$$

Assumption 4(b) states that $a=O_p(1)$. Also, as far as b and c are concerned, it holds that

$$
\begin{aligned}
\|b\| & \leq \left(\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|\widehat{W}_{it}-W_{it}\right\|^2\right)^{1/2}\left(\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|W_{it}' g_i\left(\frac{t}{T}\right)\right\|^2\right)^{1/2} \\
& = \left(\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|\hat{w}_t-w_t\right\|^2\right)^{1/2}\left(\frac{1}{m_n T}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|\frac{W_{it}'}{\sqrt{T}} g_i\left(\frac{t}{T}\right)\right\|^2\right)^{1/2}.
\end{aligned}
$$

As $(n,T)\to\infty$, Lemma A.1(b) ensures that $\left(\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor}\|\hat{w}_t-w_t\|^2\right)^{1/2}=\left[\frac{1}{T}O_p\left(\frac{1}{C_{nT}^2}\right)\right]^{1/2}=o_p(1)$; also, by Assumption 4(c), $\frac{1}{\sqrt{n}T}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|\frac{W_{it}'}{\sqrt{T}}g_i\left(\frac{t}{T}\right)\right\|^2=O_p(1)$. Consequently, b and c are both $o_p(1)$. Finally as far as d is concerned

$$
\begin{aligned}
\|d\| & \leq \left(\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|\widehat{W}_{it}-W_{it}\right\|^2\right)\left(\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|g_i\left(\frac{t}{T}\right)\right\|^2\right)^{1/2} \\
& \leq \left(\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|\hat{w}_t-w_t\right\|^2\right)\left(\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\left\|g_i\left(\frac{t}{T}\right)\right\|^2\right)^{1/2} \\
& = O\left(\frac{1}{T}\right)O_p\left(\frac{1}{C_{nT}^2}\right)O\left(\frac{1}{T}\right)O_p(1)=o_p(1),
\end{aligned}
$$

using Assumption 4(a). Combining these results we get

$$\begin{aligned}
& \sqrt{n}T\left[\hat{\theta}_{1k}^{(nT)}-\hat{\theta}_{2k}^{(nT)}\right] \\
= \; & \sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr\rfloor}-\hat{\theta}_{2\lfloor Tr\rfloor}\right]+\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}W_{it}W_{it}'\right)^{-1}\left[\frac{1}{m_nT^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}W_{it}W_{it}'g_i\left(\frac{t}{T}\right)\right] \\
& -\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}W_{it}W_{it}'\right)^{-1}\left[\frac{1}{m_nT^2}\sum_{i=1}^{m_n}\sum_{t=\lfloor Tr\rfloor+1}^{T}W_{it}W_{it}'g_i\left(\frac{t}{T}\right)\right]+o_p(1) \\
= \; & \sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr\rfloor}-\hat{\theta}_{2\lfloor Tr\rfloor}\right]+D_{n,T}+o_p(1).
\end{aligned}$$

Letting $d(r)$ be the limit of $D_{n,T}$ as $(n,T)\to\infty$, $\sqrt{n}T\left[\hat{\theta}_{1k}^{(nT)}-\hat{\theta}_{2k}^{(nT)}\right]\xrightarrow{d}[J(r)+d(r)]\,\Xi^{-1/2}$ uniformly in r, with

$$\Xi=\lim_{n,T\to\infty}\left[\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right)^{-1}+\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right)^{-1}\right].$$

Upon assuming that both $\tilde{\sigma}_\zeta$ and $\tilde{\sigma}_u$ are consistent for σ_ζ and σ_u, Theorem 3 follows from the Continuous Mapping Theorem.

Proof of Proposition 4. We first consider the consistency of $\tilde{\sigma}_u^2$. Let $\widetilde{y}_{it}=y_{it}-\bar{y}_i$; it holds that

$$\begin{aligned}
\tilde{\sigma}_u^2 &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\widetilde{y}_{it}-\widehat{W}_{it}'\hat{\theta}_t^{(nT)}\right]^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{\widetilde{y}_{it}-\widehat{W}_{it}'\left[\hat{\theta}+\frac{1}{\sqrt{n}T}g_i\left(\frac{t}{T}\right)\right]\right\}^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\widetilde{y}_{it}-\widehat{W}_{it}'\hat{\theta}\right)^2+\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{n}{m_n^2T^2}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) \\
&\quad -2\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\left(\widetilde{y}_{it}-\hat{\theta}_t'\widehat{W}_{it}\right)\frac{\sqrt{n}}{m_nT}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) \\
&= I+II+III.
\end{aligned}$$

As far as term I is concerned, we have

$$\begin{aligned} I &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\widetilde{y}_{it} - \widehat{W}_{it}'\theta + \widehat{W}_{it}'\left(\hat{\theta}-\theta\right)\right]^2 \\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left\{v_{it} + \widehat{W}_{it}'\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it}\right\}^2 \\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}^2 - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}\widehat{W}_{it}'\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it}. \end{aligned}$$

As far as III is concerned, it holds that

$$\begin{aligned} &\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\left(\widetilde{y}_{it} - \widehat{W}_{it}'\hat{\theta}\right)\frac{\sqrt{n}}{m_nT}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) \\ = &\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\left[v_{it} + \widehat{W}_{it}'\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it}\right]\frac{\sqrt{n}}{m_nT}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) \\ = &\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{\sqrt{n}}{m_nT}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}v_{it} \\ &+\frac{1}{nT}\left[\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{\sqrt{n}}{m_nT}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}\widehat{W}_{it}'\right]\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it}\right]. \end{aligned}$$

Combing I and III with II we get

$$\begin{aligned} \tilde{\sigma}_u^2 = &\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}^2 \\ &-\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}\widehat{W}_{it}'\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it} \\ &+\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{n}{m_n^2T^2}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) \\ &-\frac{2}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{\sqrt{n}}{m_nT}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}v_{it} \\ &-\frac{2}{nT}\left[\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{\sqrt{n}}{m_nT}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}\widehat{W}_{it}'\right]\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it}\right]. \end{aligned}$$

Lemmas B.1 and B.2 ensure that

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}\widehat{W}_{it}'\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it} = O_p\left(\frac{1}{nT}\right);$$

from Lemma B.1 and Assumption 4(c) and 4(d) we have

$$\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{n}{m_n^2T^2}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}\widehat{W}_{it}'g_i\left(\frac{t}{T}\right) &= O_p\left(\frac{1}{m_nT}\right), \\
\frac{1}{nT}\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{\sqrt{n}}{m_nT}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}v_{it} &= O_p\left(\frac{1}{\sqrt{n}T}\right);
\end{aligned}$$

finally, Lemmas B.1 and B.2 and Assumption 4(b) entail

$$\frac{1}{nT}\left[\sum_{i=1}^{m_n}\sum_{t=1}^{T}\frac{\sqrt{n}}{m_nT}g_i\left(\frac{t}{T}\right)'\widehat{W}_{it}\widehat{W}_{it}'\right]\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{W}_{it}v_{it}\right] = O_p\left(\frac{1}{nT}\right).$$

Hence, $\tilde{\sigma}_u^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}^2 + o_p(1)$. It holds that

$$\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}v_{it}^2 &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\theta'\left(W_{it}-\widehat{W}_{it}\right)+u_{it}\right]^2 \\
&= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2 + \frac{1}{T}\sum_{t=1}^{T}\left[\theta'\left(w_t-\hat{w}_t\right)\right]^2 + \frac{2}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\theta'\left(w_t-\hat{w}_t\right)u_{it} \\
&= I + II + III.
\end{aligned}$$

Assumption 1 and the LLN yield $I = \sigma_u^2 + o_p(1)$. As far as II and III are concerned, we have $\|II\| \leq \|\theta\|^2 \frac{1}{T}\sum_{t=1}^{T}\|(w_t-\hat{w}_t)\|^2 = o_p(1)$, using Lemma A.1(b). Also, $\|III\| \leq \|\theta\|^2$ $\frac{1}{nT}\left(\sum_{t=1}^{T}\|(w_t-\hat{w}_t)\|^2\right)^{1/2}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}u_{it}^2\right)^{1/2} = o_p(1)$, after equation (35). Then under the local alternatives $H_a^{(nT)}$ it holds that $\tilde{\sigma}_u^2 \xrightarrow{p} \sigma_u^2$. We are now ready also to prove consistency of $\tilde{\sigma}_\zeta^2$. By definition, $\tilde{\sigma}_\zeta^2 = \tilde{\sigma}_u^2 + \hat{\beta}^{(nT)\prime}\hat{\sigma}_\pi^2\hat{\beta}^{(nT)}$, and since $\hat{\sigma}_\pi^2 \xrightarrow{p} \sigma_\pi^2$ under $H_a^{(nT)}$ (as it does not depend on $H_a^{(nT)}$ being true or not), consistency of $\hat{\beta}^{(nT)}$ entails $\tilde{\sigma}_\zeta^2 \xrightarrow{p} \sigma_u^2 + \beta^{(nT)\prime}\sigma_\pi^2\beta^{(nT)\prime}$ $= \sigma_\zeta^2$.

Proof of Proposition 5. Consider $\hat{\theta}_{1\lfloor Tr\rfloor} - \theta$; the estimation error is

$$\begin{bmatrix} \hat{\beta}_{1\lfloor Tr\rfloor}-\beta \\ \hat{\gamma}_{1\lfloor Tr\rfloor}-\gamma \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_{t,r}\hat{w}_{t,r}' & \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{G}_t\hat{w}_{t,r}' \\ \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_{t,r}\bar{G}_t'\Gamma_i' & \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{G}_t\bar{G}_t'\Gamma_i' \end{bmatrix}^{-1}\begin{bmatrix} \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{w}_{t,r}v_{it} \\ \sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{G}_tv_{it} \end{bmatrix}.$$

Consider the denominator. The Cauchy-Schwartz inequality entails, for all r, $\sum_{t=1}^{\lfloor Tr\rfloor}\bar{G}_t\left(\hat{w}_{t,r}-w_{t,r}\right)' \leq$ $\left(\sum_{t=1}^{\lfloor Tr\rfloor}\|\bar{G}_t\|^2\right)^{1/2}\left(\sum_{t=1}^{\lfloor Tr\rfloor}\|\hat{w}_{t,r}-w_{t,r}\|^2\right)^{1/2} = O_p(T)\,O_p\left(\sqrt{T}C_{nT}^{-1}\right)$. Thus, using Assumption 6, as $(n,T)\to\infty$ it holds that $\left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{G}_t\hat{w}_{t,r}' = \left(nT^2\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{G}_tw_{t,r}' +$

$O_p\left(T^{-1/2}C_{nT}^{-1}\right) \xrightarrow{d} \bar{\Gamma}\int_0^r \bar{B}_G \bar{B}_\varepsilon'$. Assumptions 5 and 6 and Lemma B.1(a) entail

$$\left[\begin{array}{cc} \sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_{t,r}\hat{w}_{t,r}' & \sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} \Gamma_i \bar{G}_t \hat{w}_{t,r}' \\ \sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} \hat{w}_{t,r}\bar{G}_t'\Gamma_i' & \sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} \Gamma_i \bar{G}_t \bar{G}_t' \Gamma_i' \end{array}\right] \xrightarrow{d} \left[\begin{array}{cc} \int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon' & \bar{\Gamma}\int_0^r \bar{B}_G \bar{B}_\varepsilon' \\ \int_0^r \bar{B}_\varepsilon \bar{B}_G' \bar{\Gamma}' & \Sigma_\Gamma^{1/2}\left(\int_0^r \bar{B}_G \bar{B}_G'\right)\Sigma_\Gamma^{1/2} \end{array}\right].$$

As far as the numerator is concerned, define two non-zero $R-$ and $p-$dimensional vectors ϕ_1 and ϕ_2. Applying Lemma A.1(b), it holds that

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\left(\phi_1'\hat{w}_{t,r} + \phi_2'\Gamma_i\bar{G}_t\right)v_{it} = \frac{1}{\sqrt{n}T}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\left(\phi_1' w_{t,r} + \phi_2'\Gamma_i\bar{G}_t\right)v_{it} + o_p(1).$$

An FCLT for $\frac{1}{\sqrt{n}T}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\left(\phi_1' w_{t,r} + \phi_2'\Gamma_i\bar{G}_t\right)v_{it}$ follows from the proof of Theorem 1. By definition, the covariance matrix is

$$\begin{aligned} & \lim_{n,T\to\infty}\frac{1}{nT^2}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\left[\phi_1' w_{t,r}w_{t,r}'\phi_1 + \phi_2'\Gamma_i\bar{G}_t\bar{G}_t'\Gamma_i'\phi_2 + 2\phi_1' w_{t,r}\bar{G}_t'\Gamma_i'\phi_2\right]E\left(v_{it}^{*2}\right) \\ = \; & \sigma_\zeta^2\left[\phi_1'\int_0^r \bar{B}_\varepsilon\bar{B}_\varepsilon'\phi_1 + \phi_2'\Sigma_\Gamma^{1/2}\left(\int_0^r \bar{B}_G\bar{B}_G'\right)\Sigma_\Gamma^{1/2}\phi_2 + 2\phi_1'\int_0^r \bar{B}_\varepsilon\bar{B}_G'\bar{\Gamma}'\phi_2\right], \end{aligned}$$

using Lemma B1.(b) and Assumption 6. Thus, the finite-dimensional distributions are

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\left[\begin{array}{c} \hat{w}_{t,r}v_{it} \\ \Gamma_i\bar{G}_t v_{it} \end{array}\right] \xrightarrow{d} \sigma_\zeta \left[\begin{array}{cc} \int_0^r \bar{B}_\varepsilon \bar{B}_\varepsilon' & \bar{\Gamma}\int_0^r \bar{B}_G \bar{B}_\varepsilon' \\ \int_0^r \bar{B}_\varepsilon \bar{B}_G' \bar{\Gamma}' & \Sigma_\Gamma^{1/2}\left(\int_0^r \bar{B}_G \bar{B}_G'\right)\Sigma_\Gamma^{1/2} \end{array}\right]^{1/2} \times Z.$$

Tightness follows from Assumption 5*(c)*, along the same lines as in the proof of Theorem 1.

Proof of Proposition 6. The proof follows similar lines to the proof of Theorem 9 in Phillips and Moon (1999) and is based on two steps: firstly, we show an FCLT for $\sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr \rfloor}^{FM} - \theta\right]$; secondly, we prove the consistency of the rescaling matrix

$$\left[\begin{array}{cc} \hat{\Omega}_\zeta I_R & 0 \\ 0 & \left(\hat{\Omega}_u - \hat{\Omega}_{u\epsilon}\hat{\Omega}_\varepsilon^{-1}\hat{\Omega}_{\varepsilon u}\right)I_p \end{array}\right].$$

Consider $\hat{\theta}_{1\lfloor Tr \rfloor}^{FM} - \theta$. We have

$$\begin{aligned} \sqrt{n}T\left[\hat{\theta}_{1\lfloor Tr \rfloor}^{FM} - \theta\right] = \; & \left[\frac{1}{nT^2}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\widehat{W}_{it}'\right]^{-1} \times \\ & \left[\frac{1}{\sqrt{n}T}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}u_{it} - \frac{1}{nT}\sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor}\left(\begin{array}{cc} \hat{w}_t\Delta\widehat{F}_t' & \hat{w}_t\Delta x_{it}' \\ x_{it}\Delta\widehat{F}_t' & \sqrt{n}x_{it}\Delta x_{it}' \end{array}\right)\hat{\Omega}_{(\varepsilon\epsilon)}^{-1}\hat{\Omega}_{(\varepsilon\epsilon),u}\right. \\ & \left. -\sqrt{n}\hat{\Lambda}_{(\varepsilon\epsilon),u}^{+}\right]. \end{aligned}$$

Consider the denominator of $\hat{\theta}^{FM}_{1\lfloor Tr \rfloor} - \theta$; Assumption 7 and the application of Lemma B.1(b) yield

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\widehat{W}_{it}\widehat{W}_{it}' \xrightarrow{d} \left[\begin{array}{cc} \int_0^r \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}' & 0 \\ 0 & \frac{1}{6}r^2\Omega_{\epsilon} \end{array}\right].$$

As far as the numerator of $\sqrt{n}T\left[\hat{\theta}^{FM}_{1\lfloor Tr \rfloor} - \theta\right]$ is concerned, we can write it as

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\begin{array}{c} \hat{w}_t^*\left(u_{it}^* - n^{-1/2}\Omega_{\varepsilon,u}'\Omega_{\varepsilon}^{-1}\Delta\widehat{F}_t^*\right) - \Lambda_{\varepsilon u}^{+} \\ x_{it}^*\left(u_{it}^* - \Omega_{\epsilon,u}'\Omega_{\epsilon}^{-1}\Delta x_{it}^*\right) - \Lambda_{\epsilon u}^{+} \end{array}\right] + o_p(1) + O_p\left(\sqrt{\frac{n}{T}}\right), \tag{40}$$

where we used the fact that $\Omega_{(\varepsilon\epsilon)}$ is assumed diagonal. In (40), the remainder of order $o_p(1)$ comes from (30) and (31); the term of order $O_p\left(\sqrt{\frac{n}{T}}\right)$ comes from the BN decomposition assumed in Assumption 7*(a)*, following similar lines as in the proof of Proposition 1 and Lemma B.2. From Assumption 7 and Assumption 1*(a)*, $\hat{w}_t^*\left(u_{it}^* - n^{-1/2}\Omega_{\varepsilon,u}'\Omega_{\varepsilon}^{-1}\Delta\widehat{F}_t^*\right) - \Lambda_{\varepsilon u}^{+}$ and $x_{it}^*\left(u_{it}^* - \Omega_{\epsilon,u}'\Omega_{\epsilon}^{-1}\Delta x_{it}^*\right) - \Lambda_{\epsilon u}^{+}$ are (by construction) zero mean MDS with finite second moments. Using the same approach as in the proof of Lemma B.2, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$

$$\begin{aligned} &\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\begin{array}{c} \hat{w}_t^*\left(u_{it}^* - \frac{1}{\sqrt{n}}\Omega_{\varepsilon,u}'\Omega_{\varepsilon}^{-1}\Delta\widehat{F}_t^*\right) \\ x_{it}^*\left(u_{it}^* - \Omega_{\epsilon,u}'\Omega_{\epsilon}^{-1}\Delta x_{it}^*\right) \end{array}\right] \\ &\xrightarrow{d} \left[\begin{array}{cc} \Omega_{\zeta}\left(\int_0^r \bar{B}_{\varepsilon}\bar{B}_{\varepsilon}'\right)^{1/2} & 0_{R\times p} \\ 0_{p\times R} & \frac{1}{6}\left(\Omega_u - \Omega_{u,\epsilon}\Omega_{\epsilon}^{-1}\Omega_{\epsilon,u}\right)^{1/2} \end{array}\right] \times \left[\begin{array}{c} B(r) \\ C(r;0) \end{array}\right], \end{aligned}$$

uniformly in $r\in[0,1]$. Combining this result with the denominator, the FCLT follows.

As far as the rescaling matrix is concerned, equation (30) and Proposition 3 entail $\hat{\Omega}_{\zeta} \xrightarrow{p} \Omega_{\zeta}$ and $\left(\hat{\Omega}_u - \hat{\Omega}_{u,\epsilon}\hat{\Omega}_{\epsilon}^{-1}\hat{\Omega}_{\epsilon,u}\right) \xrightarrow{p} \left(\Omega_u - \Omega_{u,\epsilon}\Omega_{\epsilon}^{-1}\Omega_{\epsilon,u}\right)$ - see also Phillips and Moon (1999; p. 1109). The rest of the proof can be derived following the same lines as the proof of Theorem 2, and is therefore omitted.