Testing for Breaks in Cointegrated Panels*

Chihwa Kao
Syracuse University (USA)

Lorenzo Trapani
Cass Business School, London (UK)

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

March 30, 2015

Abstract

We investigate the issue of testing for structural breaks in large cointegrated panels with common and idiosyncratic regressors. We prove a panel Functional Central Limit Theorem. We show that the estimated coefficients of the common regressors have a mixed normal distribution, whilst the estimated coefficients of the idiosyncratic regressors have a normal distribution. We consider strong dependence across the idiosyncratic regressors by allowing for the presence of (stationary and nonstationary) common factors. We show that tests based on transformations of Wald-type statistics have power versus alternatives of order $n^{-1/2}T^{-1}$.

JEL codes: C23.

Keywords: Structural change, Panel cointegration, Common trends.

*This is a revised version of an earlier article entitled "Modelling and Testing for Structural Breaks in Panel Cointegration Models with Common and Idiosyncratic Stochastic Trends". We wish to thank two anonymous Referees for very constructive comments. 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), "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 the "17th International Panel Data Conference" (Montreal, 8-10 July 2011), 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, Yupin Hu, Lynda Kalaf, Hashem Pesaran, Marcel Voia and Jean-Pierre Urbain. The usual disclaimer applies.

†Cass Business School, Faculty of Finance, 106 Bunhill Row, London EC1Y 8TZ. Tel.: +44 (0) 207 8698; email: G.Urga@city.ac.uk

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. Hansen (1992) was the first to consider the issue of testing for breaks in a cointegrating regression. We also 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.

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. 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 (2010, 2011) investigates the estimation of change points in panel time trend models with cross-sectional dependence; Breitung and Eickmeier (2011), Chen, Dolado and Gonzalo (2011) and Han and Inoue (2011) investigate testing for changes in the loadings of a panel factor model.

This paper proposes an estimation and testing framework for slope parameter instability in cointegrated panel regression. Strong cross-sectional dependence among the dependent variables and among regressors is allowed for through the presence of common factors. Our model contains a set of variables that are common across all units (common shocks), and unit-specific variables (idiosyncratic shocks):

$$y_{it} = \alpha_i + \beta' f_t + \gamma' x_{it} + e_{it}, \tag{1}$$

where $i = 1, ..., n$ and $t = 1, ..., T$ and β and γ are $k \times 1$ and $p \times 1$ respectively. We assume that f_t and x_{it} are $I(1)$ and e_{it} is $I(0)$, so that (1) is a cointegrating regression for all units

i. Cross-sectional dependence among the y_{it}s arises directly from f_t. Although here we focus on the case of observable f_t, in a working paper version of this article (Kao, Trapani and Urga, 2011a), we consider the case of f_t being unobservable and estimated from a set of exogenous variables. When using the estimated f_ts, (1) becomes a factor augmented panel regression, and, as expected, β loses its structural interpretation since f_t is identified only up to a rotation. In spite of β being not identifiable due to the well-known issue of rotational indeterminacy, inference on breaks on β is not affected by this: Wald-type test statistics have the same distribution as in the case of observable β.

Main results of this paper

This paper makes three contributions to the existing literature.

First, we derive a panel functional central limit theorem (FCLT), which is the building block to study the asymptotics of the partial sample estimates. The FCLT is derived under the basic framework that the regressors x_{it} in (1) are independent across i; similarly, we assume that the error term e_{it} also are cross sectionally independent.

Second, we investigate the impact of dependence in the x_{it}s across i, showing under which conditions the asymptotics remains unchanged. Specifically, we allow for both stationary and nonstationary common factors in the DGP of the x_{it}s. The former case is tantamount to assuming that the innovations of x_{it} are (strongly) cross correlated across i. We show that, in this case, cross dependence has no impact on the asymptotics: stationary common factors are washed out and do not provide a contribution to the asymptotics, which is driven by the nonstationary components. When nonstationary common factors are present in the DGP of the x_{it}s, we show that the asymptotics becomes non-standard (and, therefore, not analytically tractable), since the impact of the common factors is not washed out. In this case, we study the restrictions on the common factors that allow for the asymptotics to be the same as when no cross dependence is present; we show that this is the case only if the common factors are not pervasive (i.e. if they have "small" loadings).

Finally, we show that tests based on Wald-type statistics (Hansen, 1992; Andrews, 1993; Andrews and Ploberger, 1994) have nontrivial power versus local alternatives shrinking at a rate $O_p\left(1/\sqrt{n}T\right)$, which provides further justification for using panels to enhance the power of tests.

Although the tests are constructed under the alternative of an abrupt change which is common to all units, we show that they have power versus other classes of alternatives, e.g. smooth parameter changes. Also, whilst our tests are designed for a common changepoint alternative, they also have power when breaks occur at different points in time across units. Finally, 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 discusses estimation under the basic setting of no cross dependence across the x_{it}s; the section contains the FCLT for the partial sample Fully Modified OLS (henceforth, FM-OLS) estimators. Section 3 extends the basic framework to the case of cross dependent x_{it}s, presenting the asymptotic theory for the cases of stationary and nonstationary common factors in the DGP of the x_{it}s (Sections 3.1 and 3.2 respectively). The test statistic and its null distribution are reported in Section 4, where we also analyze the power versus local alternatives. In Section 5 we report the finite sample properties, i.e. size and power, of our proposed tests. Section 6 provides concluding remarks. Technical Lemmas and proofs are in the appendices.

NOTATION. We define $A^{1/2}$ to be any matrix such that $A = \left(A^{1/2}\right)\left(A^{1/2}\right)'$. We use $a \wedge b$ to represent $\min\{a, b\}$; $\|v\|$ to denote the Euclidean norm of a vector v, and $\|A\|$ the Frobenius norm of a matrix A, i.e. $\|A\| = \sqrt{tr\,(A'A)}$; "$\to$" to denote the ordinary limit; "$\Rightarrow$" to denote weak convergence for sequences of random elements with respect to the Skorohod metric (see e.g. Pollard, 1984, pp. 64-66); "$\xrightarrow{d}$" to denote convergence in distribution; "$\xrightarrow{p}$" to denote convergence in probability, $\lfloor x \rfloor$ to denote the integer part, W_m to denote an m-dimensional Brownian motion, and $\bar{W}_m = W_m - \int W_m$ to denote the demeaned version of W_m. More generally, the symbol "$\bar{\cdot}$" above a series, say y_t, denotes demeaning, i.e. $\bar{y}_t = y_t - T^{-1}\sum_{t=1}^{T} y_t$. We let $M, M', M''... < \infty$ be generic positive constants that do not depend on n or T.

2 Estimation and asymptotic theory

Consider the alternative formulation of (1):

$$y_{it} = \alpha_i + \theta' z_{it} + e_{it},$$

where $\theta = \left[\beta', \gamma'\right]'$ and $z_{it} = [f_t', x_{it}']'$.

Henceforth, we define the long run variance of e_{it} as Ω_e, assuming it to be homogeneous across i for simplicity. Similarly, we define the long run covariance and one sided long run covariance of z_{it} as Ω_z and Λ_z respectively; also, we define the long run covariance and one sided long run covariance between z_{it} and e_{it} as Ω_{ze} and Λ_{ze} respectively:

$$\begin{array}{ll} \Omega_z = \lim_{T\to\infty} E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \Delta z_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \Delta z_{it}\right)'\right]; & \Lambda_z = \lim_{T\to\infty}\sum_{t=0}^T E\left[\Delta z_{i0}\Delta z_{it}'\right]; \\ \Omega_{ze} = \lim_{T\to\infty} E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \Delta z_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T e_{it}\right)\right]; & \Lambda_{ze} = \lim_{T\to\infty}\sum_{t=0}^T E\left[\Delta z_{i0} e_{it}\right]; \\ \Omega_e = \lim_{T\to\infty} E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T e_{it}\right)^2\right]. & \end{array} \tag{2}$$

In order to implement the (infeasible) FM-OLS estimator, let

$$y_{it}^{+} = y_{it} - \Delta z_{it}^{\dagger\prime}\Omega_z^{-1}\Omega_{ze}, \tag{3}$$

$$\Lambda_{ze}^{+} = \Lambda_{ze}^{\dagger} - \Lambda_z^{\dagger}\Omega_z^{-1}\Omega_{ze}, \tag{4}$$

with $\Delta z_{it}^{\dagger} = \left[n^{-1/2}\Delta f_t', \Delta x_{it}'\right]'$, $\Lambda_{ze}^{\dagger} = \left[n^{-1/2}\Lambda_{fe}', \Lambda_{xe}'\right]'$, and $\Lambda_z^{\dagger} = diag\left\{n^{-1/2}\Lambda_f, \Lambda_x\right\}$.

Let $r \in (0,1)$. The partial sample estimators are given by

$$\hat{\theta}_{1,\lfloor Tr\rfloor}^{FM*} = \left[\sum_{i=1}^n\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\sum_{i=1}^n\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}\bar{y}_{it}^{+} - \Lambda_{ze}^{+}\right)\right], \tag{5}$$

$$\hat{\theta}_{2,\lfloor Tr\rfloor}^{FM*} = \left[\sum_{i=1}^n\sum_{t=\lfloor Tr\rfloor+1}^{T}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\sum_{i=1}^n\sum_{t=\lfloor Tr\rfloor+1}^{T}\left(\bar{z}_{it}\bar{y}_{it}^{+} - \Lambda_{ze}^{+}\right)\right]. \tag{6}$$

Equations (3)-(6) define the FM-OLS estimates. The main difference with standard panel FM-OLS estimation is that, in the definition of $\Delta z_{it}^{\dagger}$, Δf_t is normalized by $n^{-1/2}$. To illustrate the rationale of this, consider, as an example, the numerator of (5). This contains the term $\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{f}_t\Delta f_t'$. In order for it not to diverge, it needs to be normalized by $\frac{1}{nT}$. Given that the numerator of $\hat{\theta}_{1,\lfloor Tr\rfloor}^{FM*}-\theta$ is normalised by $\frac{1}{\sqrt{nT}}$, the extra $\frac{1}{\sqrt{n}}$ must be applied.

Let e_t^f and e_{it}^x be the innovations of f_t and x_{it} respectively, i.e. $f_t = f_{t-1} + e_t^f$ and $x_{it} = x_{it-1} + e_{it}^x$. Define the m-dimensional vector of innovations $E_{it}^z = \left[e_{it}, e_t^{f\prime}, e_{it}^{x\prime}\right]'$, with $m = p+k+1$. In order to study the asymptotics of $\hat{\theta}_{1,\lfloor Tr\rfloor}^{FM*}$ and $\hat{\theta}_{2,\lfloor Tr\rfloor}^{FM*}$, we need the following assumption.

Assumption 1: *(a)* $E_{it}^z = \sum_{s=0}^{\infty} G_s V_{it-s}$, where V_{it} is i.i.d. across t with $E(V_{it}) = 0$, $E[V_{it}V_{it}'] = I_m$, $E\|V_{it}\|^{4+\delta} < \infty$ for some $\delta > 0$ and

$$G_s = \begin{bmatrix} -a'\tilde{C}_s \\ \Theta C_s \end{bmatrix},$$

with $a = \left[1, -\beta', -\gamma'\right]'$, $\Theta = [0_{(p+k)\times 1}\ I_{p+k}]$ and $\tilde{C}_s = \sum_{j=s+1}^{\infty} C_j$; *(b)* $\sum_{s=0}^{\infty} s^3\|C_s\| < \infty$; *(c)* E_{i0} is i.i.d. across i with $E\|E_{i0}\|^4 < \infty$ for all i; *(d)* letting $G(1) \equiv \sum_{s=0}^{\infty} G_s$, $G^f(1)$ has rank k, where $G^f(1)$ is the block on the main diagonal of $G(1)$ corresponding to f_t; also, $G^x(1)$ has rank p, where $G^x(1)$ is the block on the main diagonal of $G(1)$ corresponding to x_{it}; *(e)* e_{it} and e_{it}^x are independent across i; also, f_t and x_{it} are mutually independent.

Assumption 1 states that E_{it}^z is a linear process across time. The assumption requires that the errors e_{it} are cross-sectionally i.i.d. - similarly, the innovations e_{it}^x are assumed to be i.i.d. across i. In both cases, this assumption is made only for simplicity. However, all results in the paper hold under milder conditions on cross sectional dependence and heterogeneity for both sequences. For example, Lemma A.2 in Appendix stipulates the asymptotic independence between the estimated β and γ. However, it only requires that $n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\tau_{ij}^x\right| = o(1)$, where, letting $E\left(x_{it}x_{jt}'\right) = \tau_{ij,t}^x$ it holds that $\left\|\tau_{ij,t}^x\right\| < \left|\tau_{ij}^x\right|$ for all t. Similarly, Lemma A.4 contains the panel FCLT. The proof requires that $(nT)^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}|\tau_{ij,ts}| = O(1)$, where $\tau_{ij,ts} = E(e_{it}e_{jt})$. As well as this assumption, which is very similar to the ones in the panel factor model literature (e.g. Assumption C in Bai, 2003), Lemma A.4 also requires results such as $E\left|n^{-1/2}\sum_{i=1}^{n}e_{it}\right|^{2+\delta}<\infty$ for $\delta>0$, which we prove by applying Burkholder's inequality (see e.g. Davidson, 2002, p. 242). Again, it would be possible to replace the assumption of cross-sectional independence with some milder requirement, as long as $E\left|n^{-1/2}\sum_{i=1}^{n}e_{it}\right|^{2+\delta}$ $<\infty$ still holds. However, the presence of cross dependence is bound to cause problems when covariances such as e.g. $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\lim_{T\to\infty}E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{jt}\right)\right]$ need to be estimated. Unless the order of cross sectional correlation is known, HAC-type estimators are not feasible. Part *(d)* of the assumption entails that $G_i(1)G_i(1)'$ is positive definite for each i, which is a standard requirement in the context of cointegration analysis. Requiring independence between f_t and x_{it} (part *(e)* of the assumption) is used in the proof of Lemma A.2 - indeed, this requirement is stronger than necessary, and some dependence between f_t and x_{it} could in principle be allowed.

Define the p-dimensional, zero mean, variance transformed Gaussian process $C_p(u;v)$, with covariance structure given by

$$E\left[C_p(b;a)\,C_p(d;c)\right]=\left[\frac{(d-a)(b-c)}{(b-a)(d-c)}(b-c)^2\right]I_p, \tag{7}$$

for $a<c<b<d$. From (7), it holds that, e.g. $E\left[C_p(r;0)\,C_p(s;0)\right]=\left(\min\{r,s\}\right)^2 I_p$. Note that $C_p(r;0)$ has the same covariance structure as $\int WdB-r^{-1}B(1)\int W$, where W and B are independent standard Brownian motions of dimensions p and 1 respectively. However, $C_p(r;0)$ is normal, whilst $\int WdB-r^{-1}B(1)\int W$ is mixed normal. Indeed, $C_p(r;0)$ is the limit of $\frac{1}{\sqrt{n}}\left[\sum_{i=1}^{n}\int W_i dB_i-r^{-1}B_i(1)\int W_i\right]$, under independence across i.

It holds that

Theorem 1 *Under Assumption 1, as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$*

$$\sqrt{n}T\begin{bmatrix}\hat{\beta}_{1,\lfloor Tr\rfloor}^{FM*}-\beta\\ \hat{\gamma}_{1,\lfloor Tr\rfloor}^{FM*}-\gamma\end{bmatrix}\Rightarrow\begin{bmatrix}\int_0^r\bar{W}_f\bar{W}_f' & 0\\ 0 & \frac{1}{6}\Omega_x r^2\end{bmatrix}^{-1}\begin{bmatrix}\int_0^r\bar{W}_f dW_{e+}\\ \frac{1}{\sqrt{6}}\Omega_x^{1/2}\Omega_{e+}^{1/2}C_p(r;0)\end{bmatrix},$$

$$\sqrt{n}T\begin{bmatrix}\hat{\beta}_{2,\lfloor Tr\rfloor}^{FM*}-\beta\\ \hat{\gamma}_{2,\lfloor Tr\rfloor}^{FM*}-\gamma\end{bmatrix}\Rightarrow\begin{bmatrix}\int_r^1\bar{W}_f\bar{W}_f' & 0\\ 0 & \frac{1}{6}\Omega_x(1-r)^2\end{bmatrix}^{-1}\begin{bmatrix}\int_r^1\bar{W}_f dW_{e+}\\ \frac{1}{\sqrt{6}}\Omega_x^{1/2}\Omega_{e+}^{1/2}C_p(1;r)\end{bmatrix},$$

where $\bar{W}_f$ is a k-dimensional demeaned Brownian motion with covariance matrix Ω_f, Ω_x is the long run covariance matrix of x_{it}, and W_{e+} is a Brownian motion with variance $\Omega_{e+} = \Omega_e - \Omega_{ez}\Omega_z^{-1}\Omega_{ze}$.

Remarks

T1.1 Theorem 1 contains the FCLT for the partial sample estimator of θ. The finite dimensional distributions of $\hat{\gamma}_{1,\lfloor Tr \rfloor}^{FM*} - \gamma$ are normal, which is a difference with respect to the case of a single cointegrating equation, where the partial sample estimates have a mixed normal asymptotic distribution. This is a consequence of cross-sectional averaging (Kao, 1999; Phillips and Moon, 1999), and of having stationary common factors only in the DGP of x_{it}. Conversely, the finite dimensional distributions of $\hat{\beta}_{1,\lfloor Tr \rfloor}^{FM*} - \beta$ are mixed normal, i.e., the same as in the single equation case. As pointed out by Kao, Trapani and Urga (2011b), this is due to the strong cross dependence arising from the common factors f_t, which is not washed out by cross sectional averaging.

T1.2 The distributions of the two estimates, $\hat{\beta}_{1,\lfloor Tr \rfloor}^{FM*}$ and $\hat{\gamma}_{1,\lfloor Tr \rfloor}^{FM*}$, are asymptotically independent. This is a consequence of Lemma A.2, which stipulates that the asymptotic covariance matrix of $\sqrt{n}T\left[\hat{\theta}_{1,\lfloor Tr \rfloor}^{FM*} - \theta\right]$ is diagonal. Thus, inference on β and γ can be carried out separately.

Theorem 1 contains the asymptotics of the infeasible FM-OLS estimator. However, tests require $\sqrt{n}$-consistent estimates of the long run covariance matrices, i.e., letting $\hat{\Omega}_z$ be the estimator of Ω_z, it must hold that $\sqrt{n}\left\|\hat{\Omega}_z - \Omega_z\right\| = o_p(1)$ and similarly for the other estimated long run covariance matrices. Estimation can be carried out along the same lines as in Phillips and Moon (1999, pp. 1084-1085). Define $\hat{e}_{it} = \bar{y}_{it} - \hat{\theta}'\bar{z}_{it}$, where

$$\hat{\theta} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}\bar{y}_{it}\right] \tag{8}$$

$\hat{\theta}$ is a standard, full sample least squares dummy variable estimator. Under endogeneity, Assumption 1 entails that $\hat{\theta} - \theta = O_p\left(T^{-1}\right)$. The estimator is T-consistent, and not $\sqrt{n}T$-consistent, due to the presence of the one sided long run covariance matrix Λ_{ze} which does not vanish in cross sectional averages; see e.g., the discussion in Phillips and Moon (1999, pp. 1083-1084). We consider a restricted, full sample estimation of θ and of the long run covariance matrices. Alternatively, unrestricted, partial sample estimators could be used. It is expected that both estimators converge to the same limit under the null and under sequences of local alternatives, although this is not necessarily so under fixed alternatives (see also the discussion in Andrews, 1993, p. 833).

Let Ω_z and Λ_z be diagonal for simplicity: $\Omega_z = diag\{\Omega_f, \Omega_x\}$ and $\Lambda_z = diag\{\Lambda_f, \Lambda_z\}$. Define the autocovariance estimators

$$
\begin{aligned}
&\hat{\Phi}_{xi,j} = \tfrac{1}{T}\textstyle\sum_{t=j+1}^{T} \Delta x_{it}\Delta x'_{it-j}; && \hat{\Phi}_{f,j} = \tfrac{1}{T}\textstyle\sum_{t=j+1}^{T} \Delta f_t \Delta f'_{t-j};\\
&\hat{\Phi}_{xei,j} = \tfrac{1}{T}\textstyle\sum_{t=j+1}^{T} \Delta x_{it}\Delta x'_{it-j}\hat{e}_{it}\hat{e}_{it-j}; && \hat{\Phi}_{fei,j} = \tfrac{1}{T}\textstyle\sum_{t=j+1}^{T} \Delta f_t \Delta f'_{t-j}\hat{e}_{it}\hat{e}_{it-j};\\
&\hat{\Phi}_{ei,j} = \tfrac{1}{T}\textstyle\sum_{t=j+1}^{T} \hat{e}_{it}\hat{e}_{it-j}.
\end{aligned}
$$

Hence, $\hat{\Omega}_{x,i} = \hat{\Phi}_{xi,0} + 2\sum_{j=1}^{l} \kappa\left(\frac{j}{l}\right)\hat{\Phi}_{xi,j}$, $\hat{\Lambda}_{x,i} = \sum_{j=0}^{l} \kappa\left(\frac{j}{l}\right)\hat{\Phi}_{xi,j}$, and similarly the other estimators, using some kernel $\kappa(\cdot)$ with bandwidth l. Finally, we compute

$$
\begin{aligned}
&\hat{\Omega}_x = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \hat{\Omega}_{x,i}; && \hat{\Omega}_e = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \hat{\Omega}_{e,i};\\
&\hat{\Lambda}_x = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \hat{\Lambda}_{x,i};\\
&\hat{\Omega}_{xe} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \hat{\Omega}_{xe,i}; && \hat{\Omega}_{fe} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \hat{\Omega}_{fe,i};\\
&\hat{\Lambda}_{xe} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \hat{\Lambda}_{xe,i}; && \hat{\Lambda}_{fe} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \hat{\Lambda}_{fe,i}.
\end{aligned} \tag{9}
$$

In Section 3, we present $\sqrt{n}$-consistency results for the long run covariance matrices estimators.

3 Cross-correlated regressors

In this section, we analyse the impact of the presence of cross dependence across the regressors, which we model through the presence of common factors, viz.

$$x_{it} = w_{it} + \Gamma_i g_t, \tag{10}$$

where w_{it} is a unit specific nonstationary shock, with Γ_i a $p \times q$ matrix and g_t is a set of common factors that could be $I(0)$ or $I(1)$, or a mixture of the two. We firstly consider the case of g_t being $I(0)$, showing that the presence of common factors in the DGP of x_{it} does not alter any of the results that hold when the x_{it}s are assumed to be cross sectionally independent. We then consider the case of g_t being $I(1)$; in such case, we show that it is still possible to conduct standard inference as long as the common factors are not pervasive, i.e. if Γ_i is $n^{-1}\sum_{i=1}^{n}\Gamma_i\Gamma_i' = o(1)$.

The presence of the g_ts in (10) also affects y_{it} indirectly. Equation (10) can be further interpreted by considering the reduced form of (1)

$$y_{it} = \alpha_i + \gamma' w_{it} + \beta' f_t + \phi_i' g_t + e_{it}, \tag{11}$$

where $\phi_i = \gamma'\Gamma_i$. This is a model with common factors, some of which with a homogeneous set of coefficients (the f_ts) and some with a heterogeneous set of coefficients (the g_ts).

Recalling (10), we define the DGP of w_{it} as

$$w_{it} = w_{it-1} + e_{it}^{w}.$$

In Section 3.1, we derive the asymptotics of the partial sample estimator of θ for the case $g_t \sim I(0)$; the case of $g_t \sim I(1)$ is discussed in Section 3.2.

3.1 The case of stationary g_t

Let $h_{it} \equiv [f_t', w_{it}']'$. Long run covariance matrices are expected to be the same as in Section 2, since g_t is stationary and thus its impact is expected to be negligible. However, in this section, and in the next one, we use a slightly different notation: the long run covariance and one sided long run covariance of z_{it} are denoted as Ω_h and Λ_h respectively. Also, the long run covariance and one sided long run covariance between h_{it} and e_{it} are Ω_{he} and Λ_{he} respectively; all long run covariance matrices are defined in the same way as in (2). We use the subscript "h" since, in view of the stationarity of g_t, the only contribution to long run covariance matrices comes from h_{it}. Estimates are denoted as $\hat{\Omega}_e$, $\hat{\Omega}_h$, $\hat{\Lambda}_h$, $\hat{\Omega}_{he}$ and $\hat{\Lambda}_{he}$.

The (feasible) FM-OLS estimator is based on

$$y_{it}^{+} = y_{it} - \Delta z_{it}^{\dagger\prime} \hat{\Omega}_h^{-1} \hat{\Omega}_{he}, \tag{12}$$

$$\hat{\Lambda}_{he}^{+} = \hat{\Lambda}_{he}^{\dagger} - \hat{\Lambda}_h^{\dagger} \hat{\Omega}_h^{-1} \hat{\Omega}_{he}, \tag{13}$$

with $\hat{\Lambda}_h^{\dagger} = diag\left\{n^{-1/2}\hat{\Lambda}_f, \hat{\Lambda}_w\right\}$ and $\hat{\Lambda}_{he}^{\dagger} = \left[n^{-1/2}\hat{\Lambda}_{fe}^{\prime}, \hat{\Lambda}_{we}^{\prime}\right]^{\prime}$.

The feasible partial sample FM-OLS estimators are given, similarly to (5)-(6), by

$$\hat{\theta}_{1,\lfloor Tr\rfloor}^{FM} = \left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor} \bar{z}_{it}\bar{z}_{it}^{\prime}\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}\bar{y}_{it}^{+} - \hat{\Lambda}_{he}^{+}\right)\right], \tag{14}$$

$$\hat{\theta}_{2,\lfloor Tr\rfloor}^{FM} = \left[\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T} \bar{z}_{it}\bar{z}_{it}^{\prime}\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\left(\bar{z}_{it}\bar{y}_{it}^{+} - \hat{\Lambda}_{he}^{+}\right)\right]. \tag{15}$$

Let $E_{it} = \left[e_{it}, e_t^{f\prime}, e_{it}^{w\prime}\right]^{\prime}$, and consider the following assumptions.

Assumption 1*: Assumption 1 holds for E_{it}.

Assumption 2: *(a)* g_t is i.i.d. across t with zero mean, independent of E_{it}, and such that $E\left\|g_t\right\|^{4+\delta} < \infty$ for all t and $\delta > 0$; *(b)* it holds that $T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor} \bar{f}_t \bar{g}_t^{\prime} = O_p(1)$, $T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor} \bar{g}_t \bar{g}_t^{\prime} = O_p(1)$, and $T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor} \bar{g}_t \bar{w}_{it}^{\prime} = O_p(1)$ uniformly in r for all i; *(c)* it holds that $E\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Delta g_t \Delta w_{it}^{\prime}\right\|^2 = O(1)$ for all i; *(d)* Γ_i is non random with $\left\|\Gamma_i\right\| < \infty$ for all i.

Assumption 1* states that e_{it}^{w} is cross sectionally independent (as well as e_{it}). Similarly to Assumption 1, this is imposed only for simplicity, and cross sectional independence could be relaxed. Independence between $\{f_t\}$ and $\{w_{it}\}$ also could be relaxed, as it would suffice to have $E\left(w_{it}\middle|\{f_t\}_{t=1}^{T}\right) = 0$. Under Assumption 2, the impact of g_t on the asymptotics is negligible. Assuming that g_t is independent over time is made only for convenience. Indeed, g_t could be assumed to be a martingale mifference sequence (MDS), and the proofs would not require any changes; this is also true if g_t is assumed to be generated by a linear process, as long as it admits an MDS approximation.

When $g_t \sim I(0)$, the amount of cross dependence among the x_{it}s is given by $E\left(x_{it}x'_{jt}\right) = \Gamma_i E\left(g_t g'_t\right)\Gamma'_j$, which is $O(1)$ for all t under Assumption 2. In view of the x_{it}s being $I(1)$, $\max_t E\left(x_{it}x'_t\right) = O(T)$; thus, having stationary common factors introduces some "weak" cross dependence of relative magnitude $O\left(\frac{1}{T}\right)$.

Estimates of the long run covariance matrices are computed according to (9). All estimators are $\sqrt{n}$-consistent under Assumptions 1* and 2. This is shown in Lemma B.3 in Appendix, where we also show that $\sqrt{n}$-consistency requires that

$$\frac{l}{\sqrt{T}} + l\frac{n}{T} + \frac{n}{l^{2\psi}} \to 0, \tag{16}$$

where ψ is the Parzen exponent of $\kappa(\cdot)$ (see Andrews, 1991). This, in turn, entails that $\frac{n^{1+\varepsilon}}{T} \to 0$ for some $\varepsilon > \frac{1}{2\psi}$, which is a stronger restriction on the relative speed of divergence between n and T than $\frac{n}{T} \to 0$, assumed in Theorem 1. The "optimal" bandwidth l^* with $l^* \to \infty$ can be chosen as $l^* = \arg\min\left[nl\left(\frac{1}{n\sqrt{T}} + \frac{1}{T}\right) + \frac{n}{l^{2\psi}}\right]$. This entails that $l^* = \left(\frac{nT}{n+\sqrt{T}}\right)^{\frac{1}{1+2\psi}}$; noting that $\frac{n^{1+\varepsilon}}{T} \to 0$ entails $\frac{n}{\sqrt{T}} \to 0$, this yields $l^* = O\left[\left(n\sqrt{T}\right)^{\frac{1}{1+2\psi}}\right]$.

It holds that:

Proposition 1 *Under Assumptions 1* and 2 and equation (16), as* $(n, T) \to \infty$ *with* $\frac{n^{1+\varepsilon}}{T} \to 0$, $\sqrt{n}T\left[\hat{\theta}^{FM}_{1,\lfloor Tr\rfloor} - \theta\right]$ *has the same distribution as in Theorem 1.*

Proposition 1 states that, when g_t is stationary, its impact on the limiting distribution of $\hat{\theta}^{FM}_{1,\lfloor Tr\rfloor}$ is negligible. Indeed, the asymptotic distribution of $\hat{\theta}^{FM}_{1,\lfloor Tr\rfloor}$ is the same as in Theorem 1. Thus, even though cross dependence among the x_{it}s is present, inference can be carried out as if no common factors were present in (10). However, the presence of g_t has an impact on the asymptotics: the relative speed of divergence of n and T has to be restricted according to equation (16).

3.2 The case of nonstationary g_t

We turn to discussing the case of nonstationary g_t, and their impact on the asymptotics of $\hat{\theta}^{FM}_{1,\lfloor Tr\rfloor}$ and $\hat{\theta}^{FM}_{2,\lfloor Tr\rfloor}$ defined in (12)-(13). Prior to reporting the main results, we discuss the impact of nonstationary g_t on the asymptotics of the FM-OLS estimator.

In general, in presence of nonstationary g_t and no further restrictions, the asymptotics of $\hat{\theta}^{FM}_{1,\lfloor Tr \rfloor} - \theta$ is nonstandard. Although θ can still be estimated consistently, its limiting distribution depends on several nuisance parameters, which makes testing fraught with difficulties. In order to illustrate the main argument, consider the estimation error $\hat{\theta}^{FM}_{1,\lfloor Tr \rfloor} - \theta = \left[\sum_{i=1}^{n} \sum_{t=1}^{\lfloor Tr \rfloor} \bar{z}_{it}\bar{z}_{it}'\right]^{-1} \left[\sum_{i=1}^{n} \sum_{t=1}^{\lfloor Tr \rfloor} \left(\bar{z}_{it}e_{it}^{+} - \Lambda_{he}^{+}\right)\right]$, where we use population quantities for the long run and one-sided covariance matrices, and assume Γ_i is non random. To illustrate the main problems, we focus on the denominator:

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{z}_{it}\bar{z}_{it}' = \begin{bmatrix} \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{f}_t\bar{f}_t' & \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{f}_t\bar{x}_{it}' \\ \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{x}_{it}\bar{f}_t' & \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{x}_{it}\bar{x}_{it}' \end{bmatrix} = \begin{bmatrix} a & b \\ b' & c \end{bmatrix}.$$

Consider $b = \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{f}_t\bar{w}_{it}' + \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{f}_t\bar{g}_t'\Gamma_i' = b_1 + b_2$. We have $b_1 = o_p(1)$, by virtue of Lemma A.1 in Appendix. As far as b_2 is concerned, its order of magnitude is expected to be $O_p(1)$ unless some restrictions on Γ_i are introduced. A sufficient condition for b_2 to be negligible is that $\max_i \|\Gamma_i\| = o(1)$ as $(n,T) \to \infty$. If this holds, independence between $\hat{\beta}^{FM}_{1,\lfloor Tr \rfloor} - \beta$ and $\hat{\gamma}^{FM}_{1,\lfloor Tr \rfloor} - \gamma$ is preserved, and inference on the stability of each vector of parameters can be carried out separately. Also, consider c

$$\begin{aligned} c &= \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{w}_{it}\bar{w}_{it}' + \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \Gamma_i\bar{g}_t\bar{w}_{it}' \\ &\quad + \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{w}_{it}\bar{g}_t'\Gamma_i' + \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \Gamma_i\bar{g}_t\bar{g}_t'\Gamma_i' \\ &= c_1 + c_2 + c_3 + c_4. \end{aligned}$$

In view of similar arguments as in Lemma A.2, it can be expected that c_2 and c_3 are $o_p(1)$. Also, Lemma A.3 stipulates that $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{w}_{it}\bar{w}_{it}' \xrightarrow{p} \frac{1}{6}\Omega_w r^2$. Turning to c_4, its order of magnitude is given by $\left(\frac{1}{n}\sum_{i=1}^{n}\|\Gamma_i\|^2\right) E\left\|\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{g}_t\bar{g}_t'\right\|^2$, which in general is not negligible unless some restrictions on Γ_i are imposed. Again, $\max_i \|\Gamma_i\| = o(1)$ would serve the purpose. If this were not the case, then the asymptotic covariance of $\hat{\gamma}^{FM}_{1,\lfloor Tr \rfloor} - \gamma$ would be the sum of a non random term, $\frac{1}{6}\Omega_w r^2$, and of a random variable, corresponding to the limit of c_4. As a consequence, the limiting distribution of $\hat{\gamma}^{FM}_{1,\lfloor Tr \rfloor} - \gamma$ would depend on several nuisance parameters, such as Ω_w and the Γ_is (and, thus, on p and q also).

In order to study under which conditions Theorem 1 still holds, consider the following assumption.

Assumption 3: *(a)* $g_t = g_{t-1} + e_t^g$, with e_t^g a zero mean, i.i.d. process independent of E_{it}; *(b)* it holds that $E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{g}_t e_t^g\right\|^2 < \infty$, $\max_i E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{g}_t e_{it}\right\|^2 < \infty$, $\max_i E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{g}_t e_{it}^w\right\|^2 < \infty$ and $\max_i E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{w}_{it} e_t^g\right\|^2 < \infty$ uniformly in r; *(c)* Γ_i is non random with $\max_i \|\Gamma_i\| = o\left(n^{-1/4-\omega/2}\right)$ for some $\omega > 0$.

Assumption 3 states that g_t is nonstationary (part *(a)*), and it imposes some integrability conditions (part *(b)*) and some restrictions on the pervasiveness of the factors (part *(c)*). In principle, part *(a)* could be made more general, by allowing for serial dependence in e_t^g. Similarly, the high-level assumptions in part *(b)* could be replaced by more primitive conditions on the moments of e_t^g (and of e_{it} and e_{it}^w). The important condition is part *(c)*, which requires that the Γ_i are of smaller order than $n^{-1/4}$. This is needed in order to remove the effect of g_t from the denominator of $\hat{\theta}_{1,\lfloor Tr \rfloor}^{FM} - \theta$, which requires $\max_i \|\Gamma_i\| = o(1)$. In addition to the discussion above, we need the even stronger condition that $\max_i \|\Gamma_i\| = o\left(n^{-1/4}\right)$. This is needed in order to wipe out the impact of the g_ts on the numerator of $\hat{\theta}_{1,\lfloor Tr \rfloor}^{FM} - \theta$. The quantity ω plays a role in determining the rate of convergence of long run covariance matrices estimators (see Lemma B.3).

When $g_t \sim I(1)$, it holds that $\max_t E\left(x_{it}x_{jt}'\right) = \Gamma_i \max_t E\left(g_t g_t'\right)\Gamma_j'$, which is $O\left(Tn^{-1/2-\omega}\right)$ under Assumption 3. In view of the x_{it}s being $I(1)$, $\max_t E\left(x_{it}x_t'\right) = O(T)$; thus, having stationary common factors introduces some "weak" cross dependence of relative magnitude $O\left(n^{-1/2-\omega}\right)$.

Similarly to Proposition 1, the estimated long run covariance matrices should be $\sqrt{n}$-consistent. In Lemma B.3, we show that $\hat{\Omega}_w$, $\hat{\Omega}_e$, $\hat{\Lambda}_w$, $\hat{\Omega}_{we}$, $\hat{\Omega}_{fe}$, $\hat{\Lambda}_{we}$ and $\hat{\Lambda}_{fe}$ computed according to (9) are $\sqrt{n}$-consistent under Assumptions 1* and 3, as long as

$$l\frac{\sqrt{n}}{T} + \frac{l}{n^{\omega}} \to 0. \tag{17}$$

From equation (17), a rule to choose the bandwidth can be derived, in a similar fashion as in the case of stationary g_t. In this case, the "optimal" bandwidth should satisfy $l^* \to \infty$ and $l^* = o\left(\min\left\{\frac{T}{\sqrt{n}}, n^{\omega}\right\}\right)$.

It holds that:

Proposition 2 *Under Assumptions 1* and 3 and equation (17), as* $(n,T) \to \infty$ *with* $\frac{n}{T} \to 0$, $\sqrt{n}T\left[\hat{\theta}^{FM}_{1,\lfloor Tr \rfloor} - \theta\right]$ *has the same distribution as in Theorem 1.*

Proposition 2 states that when g_t is nonstationary, it has a negligible impact as long as its loadings are small. This result mirrors Proposition 1. In order to make $\Gamma_i g_t$ negligible, a restriction on its time series properties was imposed, which allowed to recover the same asymptotics as in Theorem 1. In the context of Proposition 2, in order to have the same asymptotics as in Theorem 1, we need to impose restrictions on the cross-sectional properties of $\Gamma_i g_t$. When these hold, inference on breaks can be conducted exactly in the same way as if g_t were stationary, or in absence of g_t.

4 Testing

In this section we present two results. First, we derive the limiting distribution of Wald-type statistics for the null of no changes in θ. This is a direct application of the asymptotic theory derived above. In particular, test statistics for the null of no change in γ are independent of statistics for the null of no change in β. As a consequence of normality, we show that the limiting distribution of Wald-type statistics for no change in γ is the same as in the literature (e.g. Andrews, 1993; Andrews and Ploberger, 1994). As a consequence of mixed normality, we show that the limiting distribution of Wald-type statistics for no change in β is the same as found by Hansen (1992) in the context of a single cointegrating equation.

We also analyse the power versus local alternatives. We show that, as a consequence of the panel approach, tests have power versus alternatives shrinking at a rate $\frac{1}{\sqrt{n}T}$. Tests also have power versus changes of different magnitude and timing across units.

For each r, we define the test statistic

$$W_{nT}(r) = \frac{1}{\hat{\Omega}_{e+}} nT^2 \left[\hat{\theta}^{FM}_{1,\lfloor Tr \rfloor} - \hat{\theta}^{FM}_{2,\lfloor Tr \rfloor}\right]' \left[H_{nT}(r)\right]^{-1} \left[\hat{\theta}^{FM}_{1,\lfloor Tr \rfloor} - \hat{\theta}^{FM}_{2,\lfloor Tr \rfloor}\right], \tag{18}$$

where $\hat{\Omega}_{e+} = \hat{\Omega}_e - \hat{\Omega}_{eh}\hat{\Omega}_h^{-1}\hat{\Omega}_{he}$ and

$$H_{nT}(r) = \left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor} \bar{z}_{it}\bar{z}_{it}'\right)^{-1} + \left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=\lfloor Tr \rfloor+1}^{T} \bar{z}_{it}\bar{z}_{it}'\right)^{-1}.$$

4.1 Null distribution

We report the asymptotic distribution of (18) under the null that θ is constant over time.

Define $s(r) = \left(\int_0^r \bar{W}_k\bar{W}_k'\right)^{-1/2} Z_1 - \left(\int_r^1 \bar{W}_k\bar{W}_k'\right)^{-1/2} Z_2$, where W_k is a k-dimensional standard Brownian motion and Z_1 and Z_2 are two k-dimensional standard normals independent of each other and of W_k. Define also $M(r) = \left[\left(\int_0^r \bar{W}_k\bar{W}_k'\right)^{-1} + \left(\int_r^1 \bar{W}_k\bar{W}_k'\right)^{-1}\right]$. Finally, define

$$J(r) = \begin{bmatrix} \frac{\frac{1-r}{r}C_p(r;0) - \frac{r}{1-r}C_p(1;r)}{\sqrt{r^2+(1-r)^2}} \\ M^{-1/2}(r)\, s(r) \end{bmatrix}. \tag{19}$$

The null distribution is given in the following Theorem.

Theorem 2 *Let Assumption 1 hold. As $(n,T) \to \infty$ with $\frac{n}{T} \to 0$, it holds that*

$$W_{nT}(r) \Rightarrow J'(r)\, J(r) = Q_k(r) + Q_p(r), \tag{20}$$

under the null of no break, where $Q_k(r)$ and $Q_p(r)$ are independent and defined as

$$Q_k(r) = s(r)'\, M^{-1}(r)\, s(r), \tag{21}$$

$$Q_p(r) = \frac{\left\|(1-r)^2 C_p(r;0) - r^2 C_p(1;r)\right\|^2}{r^4(1-r)^2 + r^2(1-r)^4}. \tag{22}$$

For fixed r, $Q_k(r) \xrightarrow{d} \chi^2_k$, $Q_p(r) \xrightarrow{d} \chi^2_p$ and $Q_k(r) + Q_p(r) \xrightarrow{d} \chi^2_{p+k}$. Further, if $g_t \sim I(0)$ and Assumptions 1 and 2, equation (16) and $\frac{n^{1+\varepsilon}}{T} \to 0$ hold; or, alternatively, if $g_t \sim I(1)$,* *and Assumptions 1* and 3, equation (17) and* $\frac{n}{T} \to 0$ *hold; then (20)-(22) still hold.*

Remarks

T2.1 Theorem 2 is an application of Theorem 1 (and Propositions 1 and 2) and the Continuous Mapping Theorem (CMT). The distribution of $Q_k(r)$ is the same as in Hansen (1992), when testing for breaks in the context of a single cointegrating equation; this is a consequence of $\hat{\beta}^{FM*}_{1,\lfloor Tr \rfloor} - \beta$ (and $\hat{\beta}^{FM}_{1,\lfloor Tr \rfloor} - \beta$) having a mixed normal asymptotic distribution. Conversely, the distribution of $Q_p(r)$ is the same as in Andrews (1993). This is a consequence of the limiting distribution of $\hat{\gamma}^{FM*}_{1,\lfloor Tr \rfloor} - \gamma$ (and $\hat{\gamma}^{FM}_{1,\lfloor Tr \rfloor} - \gamma$) being normal - see also Remark T1.1. The Theorem also states that, for fixed r, the limiting distribution of the Wald-type statistics is the usual chi-squared one.

T2.2 Note that $Q_k(r)$ and $Q_p\left(r\right)$ are independent. This is a consequence of the asymptotic independence between $\hat{\beta}^{FM}_{1,\lfloor Tr \rfloor} - \beta$ and $\hat{\gamma}^{FM}_{1,\lfloor Tr \rfloor} - \gamma$ stipulated in Theorem 1; see Remark T1.2. As Propositions 1 and 2 show, the same holds for $\hat{\beta}^{FM}_{1,\lfloor Tr \rfloor} - \beta$ and $\hat{\gamma}^{FM}_{1,\lfloor Tr \rfloor} - \gamma$. Thus, it is possible to conduct separate inference for the constancy of β and γ.

T2.3 Tests for the null of no breaks can be based on $SupW \equiv \max_{[Tr^*] \leq \lfloor Tr \rfloor \leq T - [Tr^*]} W_{nT}\left(r\right)$, where r^* represents the fraction of the sample trimmed away from the beginning and the end of the sample. The CMT ensures that $SupW \xrightarrow{d} \sup_{r^* \leq r \leq 1 - r^*} \left[J'\left(r\right) J\left(r\right)\right]$. Alternatively, following Andrews and Ploberger (1994), one can use $AveW \equiv T^{-1} \sum_{\lfloor Tr \rfloor = [Tr^*]}^{T-[Tr^*]} W_{nT}\left(r\right)$ or $ExpW \equiv \ln\left\{T^{-1} \sum_{\lfloor Tr \rfloor = [Tr^*]}^{T-[Tr^*]} \exp\left[\frac{1}{2} W_{nT}\left(r\right)\right]\right\}$; using the CMT, $AveW \xrightarrow{d} \int_{r^*}^{1-r^*} J'\left(r\right) J\left(r\right) dr$ and $ExpW \xrightarrow{d} \ln\left\{\int_{r^*}^{1-r^*} \exp\left[\frac{1}{2} J'\left(r\right) J\left(r\right)\right] dr\right\}$.

4.2 Consistency

We analyse the power of tests based on Wald-type statistics versus a general class of local-to-null alternatives shrinking as $O_p\left(1/\sqrt{n}T\right)$. We show that, as far as the time series properties of the power are concerned, the test has nontrivial power versus alternatives of order $O_p\left(1/T\right)$, and versus both abrupt and smooth changes. As far as cross-sectional properties are concerned, the test has nontrivial power versus alternatives of order $O_p\left(1/\sqrt{n}\right)$ (a consequence of the panel approach). Nontrivial power is also attained when different units undergo changes at different points in time. Finally, we discuss the power when only some units (possibly a finite number) have a break. We show that when only a finite number of units have a break, the test has power versus local alternatives of order $O_p(\sqrt{n}/T)$.

We consider the following sequence of local alternatives:

$$H_a : \theta_{it} = \theta + \frac{\sqrt{n}}{m_n T}\psi_i\left(\frac{t}{T}\right), \tag{23}$$

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

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

Assumption 4: *(a)* the function $\psi_i\left(\frac{t}{T}\right)$ is nonconstant and Riemann integrable; *(b)* $(n,T) \to \infty$, uniformly in r, $\frac{1}{m_n T^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor} \bar{z}_{it}\bar{z}_{it}'\psi_i\left(\frac{t}{T}\right) = O_p(1)$.

As discussed in Andrews (1993), possible alternative functional forms for $\psi_i(\cdot)$ include: a single step function, i.e. $\psi_i(s) = 0$ if $s < r$ and $\psi_i(s) = \triangle\theta$ (finite) if $s \geq r$, which represents a one-time change on θ at $\lfloor Tr\rfloor$; multiple steps functions that represent multiple changes; time trending functions, i.e. $\psi_i(\cdot) = t/T$.

Equation (23) encompasses various possible cases of cross-sectional behaviour. As mentioned above, having different $\psi_i(\cdot)$ across i entails having breaks of possibly different magnitude. Also, the timing of the breaks (and the presence of breaks itself) is not restricted to be the same across units. Finally, 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.

The consistency of tests based on $W_{nT}(r)$ is in the following Theorem. The expression of the noncentrality parameter $d(r)$ is in equation (37) in Appendix.

Theorem 3 *Let Assumption 1 hold. As $(n,T) \to \infty$ with $\frac{n}{T} \to 0$, it holds that, under H_a, $W_{nT}(r) \Rightarrow [J(r) + d(r)]' [J(r) + d(r)]$. The same result holds if: either $g_t \sim I(0)$ and Assumptions 1* and 2, equation (16) and $\frac{n^{1+\varepsilon}}{T} \to 0$ hold; or if $g_t \sim I(1)$ and Assumptions 1* and 3 and equation (17) hold.*

Remarks

T3.1 Consider $m_n = n$, so that (23) states that $\theta_{it} = \theta + \frac{1}{\sqrt{nT}} \psi_i \left(\frac{t}{T}\right)$. Theorem 3 shows that 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 and Propositions 1 and 2. This finding is directly related to 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 cross sectional units increases.

T3.2 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).

T3.3 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 or inertia across units are encompassed by the test.

T3.4 Of course, when $m_n < n$, the test is bound to have less power. It could be interesting to 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 shrinking at a rate $O_p\left(\sqrt{n}/T\right)$. This is worse than in the univariate case where nontrivial power is attained versus alternatives shrinking as $O_p(1/T)$.

5 Monte Carlo simulations

In this section, we report some evidence based on synthetic data on the size and power properties of the tests proposed above; in particular, we study the properties of tests based on $SupW$, $AveW$ and $ExpW$ as defined in Remark T2.3; critical values at 5% level are taken from Andrews (1993) and Andrews and Ploberger (1994), based on $r^* = 0.15$. This means that the critical values used for $SupW$, $AveW$ and $ExpW$ are 2.88, 8.85 and 2.06 respectively.

The design of the Monte Carlo exercise is based on (1)

$$y_{it} = \alpha_i + \beta f_t + \gamma x_{it} + e_{it},$$

where we set $p = k = 1$, $\beta = \gamma = 1$ and α_i is simulated as i.i.d. $N(0,1)$ across i. Also, we generate f_t and w_{it} in (10) as

$$\begin{aligned} f_t &= f_{t-1} + e_t^f, \\ w_{it} &= w_{it-1} + e_{it}^w, \end{aligned}$$

with e_t^f and e_{it}^w both i.i.d. across t and $N(0,1)$; further, e_{it}^w is generated as i.i.d. across i. We only report simulations for the case $g_t \sim I(0)$. Unreported experiments for the case $g_t \sim I(1)$ with $\Gamma_i \propto n^{-1/2}$ were equal (to the second decimal) to the ones reported here. In (10), we set $q = 1$ and generate Γ_i as i.i.d. $N(1,1)$ across i; g_t is i.i.d. $N(0,1)$.

In order to consider serial correlation and endogeneity, we firstly generate $\dot{E}_{it} = \left[\dot{e}_{it}, \dot{e}_t^f, \dot{e}_{it}^w\right]'$ as i.i.d. Gaussian with identity covariance matrix, so as to ensure that $\dot{e}_{it}$ is independent of $\left[\dot{e}_t^f, \dot{e}_{it}^w\right]'$. Contemporaneous correlation is imposed by premultiplying $\dot{E}_{it}$ by the Choleski factor of

$$\Pi = \begin{bmatrix} I_{\lfloor Tr \rfloor} & 0 & \rho^w 1_{\lfloor Tr \rfloor, n} \\ 0 & I_n & \rho^f I_n \\ \rho^w 1_{n, \lfloor Tr \rfloor} & \rho^f I_n & I_n \end{bmatrix},$$

so that ρ^f and ρ^w represent, respectively, the correlation between $\ddot{e}_{it}$ and $\ddot{e}_t^f$, and $\ddot{e}_{it}$ and $\ddot{e}_{it}^w$ in the vector $\ddot{E}_{it} = \left[\ddot{e}_{it}, \ddot{e}_t^f, \ddot{e}_{it}^w\right]'$; also, $I_{\lfloor Tr \rfloor}$ denotes an identity matrix of dimension $\lfloor Tr \rfloor$, $1_{n,\lfloor Tr \rfloor}$ is a matrix of ones of dimension $n \times \lfloor Tr \rfloor$, and similarly the other quantities. Serial correlation is induced by creating $E_{it} = \left[e_{it}, e_t^f, e_{it}^w\right]'$ according to an ARMA(1,1) specification as

$$E_{it} = \rho E_{it-1} + \ddot{E}_{it} + \vartheta \ddot{E}_{it-1}.$$

We conduct our simulations under $\left(\rho^f, \rho^w\right) = (0.5, 0.5)$, and we consider the following combinations of (ρ, ϑ): $(0,0)$, $(0.7,0)$, $(0,0.7)$, $(0.7,0.7)$.

We consider combinations of (n, T) with $n = (20, 50, 100)$ and $T = (50, 100, 250)$. When estimating long run covariance matrices, we use the HAC-type estimators discussed in Lemma B.3. As far as the kernel is concerned, we use Bartlett kernel (whose Parzen exponent is $\psi = 1$). As far as the selection of the bandwidth l is concerned, we select it using $l^* = O\left[\left(n\sqrt{T}\right)^{\frac{1}{1+2\psi}}\right]$, as discussed in Section 3.1. This entails that $l^* = \lfloor T^{1/3} \rfloor$ in our case. Also, tests should be carried out under the restriction that $\frac{n^{3/2}}{T} \to 0$. It can be expected that tests do not have the correct size when this is not the case; however, l^* still is a minimizer for (16), which should ameliorate this case too.

When computing the power, we consider the following alternative:

$$\gamma_t = \begin{cases} 1 \text{ for } t = 1, ... \left\lfloor \frac{T}{2} \right\rfloor \\ 1 + c_i \text{ for } t = \left\lfloor \frac{T}{2} \right\rfloor + 1, ..., T \end{cases}, \tag{24}$$

where $c_i \sim i.i.d.N(c, 1)$ and c is set equal to 0.25 and 0.5 in two sets of experiments.

Finally, all simulated data have been computed with 1000 replications.

Table 1 reports empirical rejection frequencies at a 5% level. Given the number of simulations, a 95% confidence interval for the empirical size is $0.05 \pm 2\sqrt{\frac{0.05(1-0.95)}{1000}} \simeq [0.036, 0.064]$.

[Insert Table 1 somewhere here]

As a general finding, tests based on $SupW$ have the correct size across almost all experiments; there are only a few cases (corresponding to the presence of MA roots in E_{it}) in which the test has a tendency to be slightly conservative. On the other hand, tests based on $AveW$ and $ExpW$ have a tendency to be oversized; this is particularly true for tests based on $AveW$. This entails that, for small samples, tests based on $SupW$ are preferable, at least as far as size properties are concerned. This, however, vanishes as T increases; there are few exceptions, all for $AveW$, whereas $ExpW$ attains the correct size in all cases considered for $T = 200$. As predicted by the theory, as T increases the tests improve, having the correct size. There are few exceptions, in presence of AR roots in the DGP of E_{it}. The impact of n on the empirical sizes seems to be quite limited: in spite of the theoretical need for the restriction $\frac{n^{3/2}}{T} \to 0$, as n increases results are relatively stable as long as T is large. When T is small relatively to n, the empirical sizes are, especially for the $AveW$ statistic, grossly overstated. Finally, the dynamics of E_{it} has a quite clear impact on the size. In presence of AR roots, the tests all have a tendency to over-reject, in finite samples; conversely, tests become more conservative when MA roots are present. This is attenuated as T increases.

As far as the power is concerned, we report the empirical rejection frequencies, adjusted for size distortion, in Tables 2 and 3. Two sets of experiments were carried out, setting c in (24) as 0.25 and 0.5 respectively.

[**Insert Tables 2 and 3 somewhere here**]

The power of all the three tests has a behaviour which is in line with the theory: as n and T increase, the power increases monotonically across all experiments, in both cases of $c = 0.25$ and $c = 0.5$. Indeed, the impact of n is even stronger than predicted by the theory: adapting the results in Theorem 3, the power versus fixed alternatives should grow as $\sqrt{n}T$. It can be noted that, when n is smaller than 200, the tests based on $SupW$ have a smaller power. For sufficiently large (n, T), the power properties of the three tests are very similar. As far as the dynamics of E_{it} is concerned, when AR roots are present, tests are more powerful, ceteris paribus; this is strictly related to the behaviour under H_0, where tests, as noted above, have a (slight) tendency to over-reject. Conversely, it is interesting to note that, in presence of MA roots tests, which become more conservative under H_0, are more powerful than in the white noise case.

6 Conclusions

In this paper, we derived an inferential theory for testing for an unknown common change point in a cointegrated large panel regression. Cross sectional dependence is allowed for across y_{it}

directly (through f_t) and in the x_{it}s (through g_t). We show that the two cases of g_t being stationary and nonstationary have different impacts, and discuss the conditions under which the asymptotics is the same.

We consider Wald-type statistics, showing that under the null the limiting distributions are nuisance parameters free and depend only on the number of regressors. 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 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.

As well as the issue of nonstationary common factors in x_{it} being an open issue, it would be natural to extend the framework developed here 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.

Appendix: proofs and preliminary Lemmas

Throughout this whole section, we use the notation $\delta_{nT} \equiv \min\left\{\sqrt{n}, \sqrt{T}\right\}$, and we also define $h_{it} \equiv [f_t', w_{it}']'$.

The Appendix is organised as follows. We start with a preliminary Lemma of general interest. We then report proofs and preliminary Lemmas for Sections 2-3 and 4 in separate Appendices (Appendix A and B respectively).

Lemma A.1 *Let $\varphi_i = \varphi_{i-1} + u_i^x$, with u_i^x a v-dimensional, zero mean, MDS with $E\left(u_i^x u_i^{x\prime}\right) = I_v$, and let u_i be an MDS independent of u_i^x, with unit variance. Then, as $T \to \infty$ for $a < c < b < d$*

$$
\begin{aligned}
& E\left\{\frac{1}{T}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\left[\varphi_{i-1} - \frac{1}{T(b-a)}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\varphi_{i-1}\right]u_i \times \frac{1}{T}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}\left[\varphi_{i-1}' - \frac{1}{T(d-c)}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}\varphi_{i-1}'\right]u_i\right\} \\
= & \left[\frac{1}{6}\frac{(d-a)(b-c)}{(b-a)(d-c)}(b-c)^2\right]I_v. \qquad (25)
\end{aligned}
$$

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

$$
\begin{aligned}
& E\left[\frac{1}{T^2}\left(\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\bar{\varphi}_{i-1}u_i\right)\left(\sum_{i=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}\bar{\varphi}_{i-1}u_i\right)'\right] = E\left[\frac{1}{T^2}\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Td\rfloor}\bar{\varphi}_{i-1}\bar{\varphi}_{j-1}'u_iu_j\right] \\
= \; & E\left[\frac{1}{T^2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\bar{\varphi}_{i-1}\bar{\varphi}_{i-1}'u_i^2\right] = E\left(\frac{1}{T^2}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\bar{\varphi}_{i-1}\bar{\varphi}_{i-1}'\right)
\end{aligned}
$$

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

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

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

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

Similar passages yield

$$
\begin{aligned}
III &= \frac{1}{b-a}\frac{1}{T^3}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Ta\rfloor}^{\lfloor Tb\rfloor} E\left(\varphi_{i-1}\varphi_{j-1}'\right) \\
&= \frac{1}{b-a}\frac{1}{T^3}\left[\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tc\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor} E\left(\varphi_{i-1}\varphi_{j-1}'\right)+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor} E\left(\varphi_{i-1}\varphi_{j-1}'\right)\right] \\
&= \frac{1}{b-a}\frac{1}{T^3}\left[T\left(b-c\right)\sum_{i=\lfloor Ta\rfloor}^{\lfloor Tc\rfloor}\left(i-1\right)+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(i\wedge j-1\right)\right]\times I_v \\
&\to \frac{1}{b-a}\left[\int_a^c udu+\int_c^b\int_c^b\left(u\wedge v\right)dudv\right]\times I_v \\
&= \left[\frac{1}{2}\frac{b-c}{b-a}\left(c^2-a^2\right)+\frac{1}{3}\frac{b^3-c^3}{b-a}\right]\times I_v,
\end{aligned}
$$

$$
\begin{aligned}
IV &= \frac{1}{d-c}\frac{1}{T^3}\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Td\rfloor} E\left(\varphi_{i-1}\varphi_{j-1}'\right) \\
&= \frac{1}{d-c}\frac{1}{T^3}\left[\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor} E\left(\varphi_{i-1}\varphi_{j-1}'\right)+\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tb\rfloor}^{\lfloor Td\rfloor} E\left(\varphi_{i-1}\varphi_{j-1}'\right)\right] \\
&= \frac{1}{d-c}\frac{1}{T^3}\left[\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\sum_{j=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(i\wedge j-1\right)+T\left(d-b\right)\sum_{i=\lfloor Tc\rfloor}^{\lfloor Tb\rfloor}\left(i-1\right)\right]\times I_v \\
&\to \frac{1}{d-c}\left[\int_c^b\int_c^b\left(u\wedge v\right)dudv+\left(d-b\right)\int_c^b udu\right]\times I_v \\
&= \left[\frac{1}{3}\frac{b^3-c^3}{d-c}+\frac{1}{2}\frac{d-b}{d-c}\left(b^2-c^2\right)\right]\times I_v.
\end{aligned}
$$

Putting everything together, the Lemma follows. QED

Appendix A: proofs of results in Sections 2 and 3

The following three Lemmas are needed to prove Theorem 1 and Proposition 1. The focus of the three Lemmas is on the setup in Section 3.1, of which Section 2 is a special case.

Lemma A.2 *Under Assumptions 1* and 2, as $(n, T) \to \infty$, it holds that, uniformly in r*

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t\bar{x}_{it}' = O_p\left(\frac{1}{\delta_{nT}}\right).$$

Proof. We have

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t\bar{x}_{it}' = \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t'\bar{w}_{it}' + \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t\bar{g}_t'\Gamma_i' = I + II.$$

Consider I, and let $\varsigma_{i\lfloor Tr \rfloor} \equiv T^{-2}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t\bar{w}_{it}'$. The Beveridge-Nelson (BN henceforth) decomposition entailed by Assumption 1* allows us to write

$$\varsigma_{i\lfloor Tr \rfloor} = \frac{1}{T^2}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t^*\bar{w}_{it}^{*\prime} + R_{i\lfloor Tr \rfloor} = \varsigma_{i\lfloor Tr \rfloor}^* + R_{i\lfloor Tr \rfloor},$$

where $\bar{f}_t^*$ and $\bar{w}_{it}^*$ are such that Δf_t^* and Δw_{it}^* are i.i.d. across t with variances Ω_f and Ω_w respectively. As far as the remainder $R_{i\lfloor Tr \rfloor}^{\varsigma}$ is concerned, Assumption 1* entails that $R_{i\lfloor Tr \rfloor}^{\varsigma} = O_p\left(T^{-1/2}\right)$. This can be proved following the same steps as in Phillips and Moon (1999, p. 1101-1102).

We turn to show that $\frac{1}{n}\sum_{i=1}^{n}\varsigma_{i\lfloor Tr \rfloor}^*$ is $O_p\left(\frac{1}{\sqrt{n}}\right)$. Let C be the σ-field associated with $\{f_t\}_{t=1}^{T}$. The sequence $\varsigma_{i\lfloor Tr \rfloor}^*$ has mean zero: $E\left[\varsigma_{i\lfloor Tr \rfloor}^*\middle| C\right] = T^{-2}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t^* E\left(\bar{w}_{it}^{*\prime}\right) = 0$, since $\bar{w}_{it}^*$ has mean zero. Also, define the σ-field $I_i \equiv \{\varsigma_{1T}^*, ..., \varsigma_{iT}^*\} \cup C$; in view of Assumption 1*, $E\left[\varsigma_{i\lfloor Tr \rfloor}^*\middle| I_{i-1}\right] = E\left[\varsigma_{i\lfloor Tr \rfloor}^*\middle| C\right] = 0$. This entails that $\left\{\varsigma_{i\lfloor Tr \rfloor}^*, I_i\right\}$ is an MDS. Thus, $n^{-1}\sum_{i=1}^{n}\varsigma_{i\lfloor Tr \rfloor}^*$ is bounded by the square root of

$$E\left[\left(\frac{1}{n}\sum_{i=1}^{n}\varsigma_{i\lfloor Tr \rfloor}^*\right)^2\right] = \frac{1}{n^2}\sum_{i=1}^{n}E\left(\varsigma_{i\lfloor Tr \rfloor}^{*2}\right) = \frac{1}{n^2}\sum_{i=1}^{n}E\left[\left(\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t^*\bar{w}_{it}^{*\prime}\right)^2\right].$$

This entails that $n^{-1}\sum_{i=1}^{n}\varsigma_{i\lfloor Tr \rfloor}^*$ is bounded by $n^{-1}E\left[\left(\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr \rfloor}\bar{f}_t^*\bar{w}_{it}^{*\prime}\right)^2\right]$, which yields $E\left[\left(\frac{1}{n}\sum_{i=1}^{n}\varsigma_{i\lfloor Tr \rfloor}^*\right)^2\right] = O\left(\frac{1}{n}\right)$. Thus, $\frac{1}{n}\sum_{i=1}^{n}\varsigma_{i\lfloor Tr \rfloor}^* = O_p\left(\frac{1}{\sqrt{n}}\right)$. Putting all together

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

whence $I = O_p\left(\delta_{nT}^{-1}\right)$. Turning to II, Assumption 2*(b)* yields $II = O_p\left(\frac{1}{T}\right)$. QED

Lemma A.3 *Under Assumptions 1* and 2, as* $(n,T) \to \infty$

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}_{it}' \Rightarrow \begin{bmatrix} \int_0^r \bar{W}_f\bar{W}_f' & 0_{k\times p} \\ 0_{p\times k} & \frac{1}{6}\Omega_w r^2 \end{bmatrix}.$$

Proof. We have

$$\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}_{it}' = \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\begin{bmatrix} \bar{f}_t\bar{f}_t' & \bar{f}_t\bar{x}_{it}' \\ \bar{x}_{it}\bar{f}_t' & \bar{x}_{it}\bar{x}_{it}' \end{bmatrix}.$$

Consider $\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_t\bar{f}_t'$. Assumption 1* entails that an FCLT holds, whereby $\frac{1}{T^2}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_t\bar{f}_t' \Rightarrow \int_0^r \bar{W}_f\bar{W}_f'$. The off-diagonal blocks converge to zero in view of Lemma A.2. Finally, consider $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{x}_{it}\bar{x}_{it}' = \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_t\bar{g}_t'\Gamma_i' + \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_t\bar{w}_{it}' + \frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it}\bar{g}_t'\Gamma_i'$ $+\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it}\bar{w}_{it}' = I + II + III + IV$. It holds that $I = O_p\left(T^{-1}\right)$, using Assumption 2*(b)*. As far as II and III are concerned, they are both bounded by $\frac{1}{T}\left\|\Gamma_i\right\| E\left[\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{g}_t\bar{w}_{it}'\right\|\right]$, which is $O_p\left(\frac{1}{T}\right)$ based on Assumptions 2*(b)* and 2*(c)*; this is not the sharpest bound, but it suffices for our purposes. As far as IV is concerned, Kao and Chiang (2000) prove that $\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it}\bar{w}_{it}' \xrightarrow{p} \frac{1}{6}\Omega_w r^2$. QED

Lemma A.4 *Let* $e_{it}^{+} = e_{it} - \Delta z_{it}^{\dagger\prime}\Omega_h^{-1}\Omega_{he}$. *Under Assumptions 1* and 2, as* $(n,T) \to \infty$ *with* $\frac{n}{T} \to 0$

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}e_{it}^{+} - \Lambda_{he}^{+}\right) \Rightarrow \begin{bmatrix} \int_0^r \bar{W}_f dW_{e+} \\ \frac{1}{\sqrt{6}}\Omega_w^{1/2}\Omega_{e+}^{1/2}C_p\left(r;0\right) \end{bmatrix}.$$

Proof. We have

$$
\begin{aligned}
& \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}e_{it}^{+}-\Lambda_{he}^{+}\right) \qquad (26)\\
= & \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{h}_{it}e_{it}^{h+}-\Lambda_{he}^{+}\right)+\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left[\begin{array}{c}0\\ \left(\Gamma_{i}\bar{g}_{t}e_{it}^{+}-\Lambda_{g,i}\Omega_{w}^{-1}\Omega_{we}\right)\end{array}\right]\\
& -\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{h}_{it}\Delta\bar{g}_{t}'\Gamma_{i}'\Omega_{w}^{-1}\Omega_{we}\\
= & I+II+III,
\end{aligned}
$$

where $e_{it}^{h+}\equiv e_{it}-\Delta h_{it}^{\dagger\prime}\Omega_{z}^{-1}\Omega_{ze}$ and $h_{it}^{\dagger}=\left[n^{-1/2}f_{t}',w_{it}'\right]'$.

We start by showing that II and III are asymptotically negligible. Consider II

$$
\begin{aligned}
& \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\Gamma_{i}\bar{g}_{t}e_{it}^{+}-\Lambda_{g,i}\Omega_{z}^{-1}\Omega_{ze}\right)\\
= & \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_{i}\bar{g}_{t}e_{it}-\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_{i}\bar{g}_{t}\Delta w_{it}'\Omega_{w}^{-1}\Omega_{we}\\
& -\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\Gamma_{i}\bar{g}_{t}\bar{g}_{t}'+\Lambda_{g,i}\right)\Omega_{w}^{-1}\Omega_{we}+\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_{i}\bar{g}_{t}\bar{g}_{t-1}'\Omega_{w}^{-1}\Omega_{we}\\
= & II_{a}+II_{b}+II_{c}+II_{d}.
\end{aligned}
$$

We have $II_{a}=\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Gamma_{i}e_{it}\right)g_{t}+\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Gamma_{i}e_{it}\right)\left(\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}g_{t}\right)=II_{a,1}$ $+\ II_{a,2}$. Consider $II_{a,1}$, and let $\upsilon_{nt}=\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Gamma_{i}e_{it}\right)g_{t}$; υ_{nt} is an MDS across t for every n, in view of Assumption 2. Thus, the order of magnitude of $T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\upsilon_{nt}$ is given by the square root of $E\left\Vert T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\upsilon_{nt}\right\Vert^{2}$. In view of the MDS property, $E\left\Vert T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\upsilon_{nt}\right\Vert^{2}$ $=T^{-2}\sum_{t=1}^{\lfloor Tr\rfloor}E\left\Vert\upsilon_{nt}\right\Vert^{2}$; this is bounded by $T^{-1}\max_{1\leq t\leq T}E\left\Vert\upsilon_{nt}\right\Vert^{2}$, uniformly in r. We have $E\left\Vert\upsilon_{nt}\right\Vert^{2}\leq E\left\Vert g_{t}\right\Vert^{2}\left(n^{-1}\sum_{i=1}^{n}\left\Vert\Gamma_{i}\right\Vert^{2}E\left\vert e_{it}\right\vert^{2}\right)$, using the fact that g_{t} is independent of e_{it}. By Assumptions 1*(a) and 2(a), $E\left\Vert\upsilon_{nt}\right\Vert^{2}<\infty$, whence $II_{a}=O_{p}\left(T^{-1/2}\right)$ uniformly in r. Turning to $II_{a,2}$, Assumption 2 entails that $\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}g_{t}=O_{p}\left(T^{-1/2}\right)$. Also, $\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Gamma_{i}e_{it}\right)$ is bounded by the square root of $\frac{1}{nT^{2}}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\sum_{s=1}^{\lfloor Tr\rfloor}\Gamma_{i}\Gamma_{i}'E\left(e_{it}e_{is}\right)$, due to the assumption of cross sectional independence. This is bounded by $T^{-1}\sup_{i}\left\Vert\Gamma_{i}\right\Vert^{2}\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\sum_{s=1}^{\lfloor Tr\rfloor}\sup_{i}\left\vert E\left(e_{it}e_{is}\right)\right\vert$, which is of order $O\left(T^{-1}\right)$ due to the summability assumed in Assumption 1. Thus, $II_{a,2}=$

$O_p\left(T^{-1}\right)$. The same logic applies to II_b, which is also $O_p\left(T^{-1/2}\right)$. Consider II_c. Note that $\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_t\bar{g}_t'=\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i g_t g_t'-\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\left(\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}g_t\right)\left(\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}g_t\right)'$; also, $\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}g_t=O_p\left(T^{-1/2}\right)$. Thus, $II_c\leq\sqrt{n}\left\|\Omega_w^{-1}\Omega_{we}\right\|\frac{1}{n}\sum_{i=1}^{n}\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i g_t g_t'+\Lambda_{g,i}\right\|+O_p\left(T^{-1}\right)$. By definition, $E\left(\Gamma_i g_t g_t'\right)=-\Lambda_{g,i}$; also, $\Gamma_i g_t g_t'$ is i.i.d. with finite second moment in view of Assumption 2*(a)*. Thus, $T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\Gamma_i g_t g_t'-\Lambda_{g,i}\right)=O_p\left(T^{-1/2}\right)$ uniformly in r. Hence, $II_c=O_p\left(\sqrt{\frac{n}{T}}\right)$. Finally, consider II_d. Again $\leq\sqrt{n}\,\sup_i\left\|\Gamma_i\right\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}g_t g_{t-1}'+O_p\left(T^{-1}\right)$ due to similar arguments. The sequence $g_t g_{t-1}'$ is an MDS, since $E\left[g_t g_{t-1}'\middle|g_{t-1},g_{t-2},...\right]=E\left[g_t\right]g_{t-1}'=0$; also, $E\left\|g_t g_{t-1}'\right\|^2$ is finite in view of Assumption 2*(a)*. Thus, $T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}g_t g_{t-1}'=O_p\left(T^{-1/2}\right)$, whence $II_d=O_p\left(\sqrt{\frac{n}{T}}\right)$. Putting all together, $II=O_p\left(\sqrt{\frac{n}{T}}\right)$ uniformly in r.

Consider now III in (26). We have

$$\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{h}_{it}\Delta\bar{g}_t'=\frac{1}{T}\bar{h}_{i\lfloor Tr\rfloor}\bar{g}_{\lfloor Tr\rfloor}'-\frac{1}{T}\bar{h}_{i0}\bar{g}_0'-\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Delta\bar{h}_{it}\bar{g}_{t-1}'=III_a+III_b+III_c.$$

Given that $E\left\|\bar{h}_{i\lfloor Tr\rfloor}\bar{g}_{\lfloor Tr\rfloor}'\right\|\leq\left(E\left\|\bar{h}_{i\lfloor Tr\rfloor}\right\|^2\right)^{1/2}\left(E\left\|\bar{g}_{\lfloor Tr\rfloor}\right\|^2\right)^{1/2}=O_p\left(\sqrt{T}\right)O_p\left(1\right)$ uniformly in r, $III_a=O_p\left(\frac{1}{\sqrt{T}}\right)$; similarly, in view of Assumptions 1**(c)* and 2*(a)*, $E\left\|\bar{h}_{i0}\bar{g}_0'\right\|=O_p\left(1\right)$, whence $III_b=O_p\left(\frac{1}{T}\right)$. Finally, consider III_c. We have $\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Delta\bar{h}_{it}\bar{g}_{t-1}'=\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Delta h_{it}g_{t-1}'-\left(\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Delta h_{it}\right)\left(\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{g}_{t-1}'\right)=III_{c,1}+III_{c,2}$. Consider $III_{c,1}$; $\Delta h_{it}g_{t-1}'$ is an MDS across t, with $E\left\|\Delta\bar{h}_{it}\bar{g}_{t-1}'\right\|^2\leq\left(E\left\|\Delta\bar{h}_{it}\right\|^4\right)^{1/2}\left(E\left\|\bar{g}_{t-1}\right\|^4\right)^{1/2}<\infty$ from Assumptions 1**(a)* and 2*(a)*, so that $III_c=O_p\left(\frac{1}{\sqrt{T}}\right)$ uniformly in r. As far as $III_{c,2}$ is concerned, similar considerations to above yield $III_{c,2}=O_p\left(T^{-1}\right)$. Given that III is bounded by $\sqrt{n}\,\sup_i\left(E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{h}_{it}\Delta\bar{g}_t'\right\|\right)$ $\sup_i\left\|\Gamma_i'\right\|\left\|\Omega_w^{-1}\Omega_{we}\right\|$, this yields $III=O_p\left(\sqrt{\frac{n}{T}}\right)$ uniformly in r.

In view of the above, consider (26):

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}e_{it}^{+}-\Lambda_{he}^{+}\right)=\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{h}_{it}e_{it}^{h+}-\Lambda_{he}^{+}\right)+O_p\left(\sqrt{\frac{n}{T}}\right),$$

uniformly in r. As $(n,T)\to\infty$ under $\frac{n}{T}\to 0$, the asymptotics of (26) is driven by $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{h}_{it}e_{it}^{h+}-\Lambda_{he}^{+}\right)$. Assumption 1* entails that a BN decomposition holds for both $\bar{h}_{it}$ and e_{it}^{h+}, whereby

$$\bar{h}_{it} = \bar{h}_{it}^{*} + R_{h,it}, \tag{27}$$

$$e_{it}^{h+} = e_{it}^{h+*} + R_{e,it}^{h}, \tag{28}$$

where $\bar{h}_{it}^{*}$ and e_{it}^{h+*} denote the MDS approximations and $R_{h,it}$ and $R_{e,it}^{h}$ are squared integrable remainders by Assumption 1* - see Phillips and Moon (1999) for details. Based on (27)-(28), we may write

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{h}_{it}e_{it}^{h+} - \Lambda_{he}^{+}\right) = \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{h}_{it-1}^{*}e_{it}^{h+*} + R_{eh,n\lfloor Tr\rfloor}.$$

The same passages as in the proof of Lemma 16 in Phillips and Moon (1999, p. 1105) yield that $R_{eh,n\lfloor Tr\rfloor} = O_p\left(\sqrt{\frac{n}{T}}\right)$ uniformly in r, so that as $(n,T)\to\infty$ with $\frac{n}{T}\to 0$, the asymptotics is driven by

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{h}_{it-1}^{*}e_{it}^{h+*} = \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\begin{array}{c}\bar{f}_{t-1}^{*}e_{it}^{h+*}\\ \bar{w}_{it-1}^{*}e_{it}^{h+*}\end{array}\right).$$

In view of Lemma A.2, $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_{t-1}^{*}e_{it}^{h+*}$ and $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it-1}^{*}e_{it}^{h+*}$ have a diagonal covariance matrix; therefore, they can be studied separately.

Consider first

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_{t-1}^{*}e_{it}^{h+*} = \frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_{t-1}^{*}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}e_{it}^{h+*}\right) = \frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_{t-1}^{*}e_{nt}.$$

By construction, e_{nt} is an MDS, it has variance Ω_{e+}, and it is independent of $\bar{f}_{t-1}$. Also

$$\begin{aligned} E\left|e_{nt}\right|^{2+\delta} &\leq M\left(E\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}e_{it}^{*}\right|^{2+\delta} + E\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}e_{it}^{w*\prime}\Omega_{w}^{-1}\Lambda_{we}\right|^{2+\delta} + E\left|e_{t}^{f*}\Omega_{f}^{-1}\frac{1}{n}\sum_{i=1}^{n}\Lambda_{we}\right|^{2+\delta}\right) \\ &\leq M\sup_{i}E\left|e_{it}^{*}\right|^{2+\delta} + M'\sup_{i}E\left|e_{it}^{w*}\right|^{2+\delta} + M''E\left|e_{t}^{f*}\right|^{2+\delta}. \end{aligned}$$

To show the last inequality, consider $n^{-\frac{2+\delta}{2}}E\left|\sum_{i=1}^{n}e_{it}^{*}\right|^{2+\delta}$. By Burkholder's inequality, this is bounded by $n^{-\frac{2+\delta}{2}}E\left|\sum_{i=1}^{n}(e_{it}^{*})^{2}\right|^{\frac{2+\delta}{2}}$; Holder's inequality ensures that this, in turn, is bounded by $n^{-\frac{2+\delta}{2}} E \left| \left(\sum_{i=1}^n |e_{it}^*|^{2+\delta} \right)^{\frac{2+\delta}{2}} n^{1-\frac{2}{2+\delta}} \right|^{2+\delta}$. Finally, applying the C_r-inequality and Jensen's inequality, $n^{-\frac{2+\delta}{2}} E \left| \sum_{i=1}^n e_{it}^* \right|^{2+\delta} \leq M \frac{1}{n} \sum_{i=1}^n E |e_{it}^*|^{2+\delta} \leq M \sup_i |e_{it}^*|^{2+\delta}$. The same applies to $E \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n e_{it}^{w*\prime} \Omega_w^{-1} \Lambda_{we} \right|^{2+\delta}$. Thus, by Assumption 1**(a)*, $E |e_{nt}|^{2+\delta}$ is bounded uniformly in n. An FCLT for MDS (see e.g. Hall and Heyde, 1980, Theorem 4.1) holds with $T^{-1/2} \sum_{t=1}^{\lfloor Tr \rfloor} e_{nt} \Rightarrow W_{e+}(r)$. Standard arguments in the theory of weak convergence to stochastic integrals (Phillips, 1988) yield

$$\frac{1}{T} \sum_{t=1}^{\lfloor Tr \rfloor} \bar{f}_{t-1}^* \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n e_{it}^{h+*} \right) \Rightarrow \int_0^r \bar{W}_f dW_{e+}.$$

As far as $\frac{1}{\sqrt{nT}} \sum_{i=1}^n \sum_{t=1}^{\lfloor Tr \rfloor} \bar{w}_{it-1}^* e_{it}^{h+*}$ is concerned, let $\xi_{nt} = n^{-1/2} \sum_{i=1}^n \bar{w}_{it-1}^* e_{it}^{h+*}$, and let $p = 1$ for the sake of notational simplicity. Let $C_{\xi nt-1} \equiv \left\{ \xi_{n1}, ..., \xi_{nt-1} \right\}$; since e_{it}^{h+*} is i.i.d. across time, $E\left(\xi_{nt} | C_{\xi nt-1} \right) = 0$. Thus, ξ_{nt} is an MDS. Finally, we show that a Lindeberg condition holds uniformly in n. Indeed, let $S_{n\lfloor Tr \rfloor}^2 \equiv E\left(\sum_{t=1}^{\lfloor Tr \rfloor} \xi_{nt}^2 \right)$; it holds that

$$\begin{aligned}
\sum_{t=1}^T E \left| \frac{\xi_{nt}}{S_{nT}} \right|^{2+\delta} &\leq M \frac{1}{T^{2+\delta}} \sum_{t=1}^T E \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \bar{w}_{it-1}^* e_{it}^{h+*} \right|^{2+\delta} \\
&\leq M' \frac{1}{T^{2+\delta}} n^{-\frac{2+\delta}{2}} \sum_{t=1}^T E \left| \sum_{i=1}^n \left(\bar{w}_{it-1}^* e_{it}^{h+*} \right)^2 \right|^{\frac{2+\delta}{2}} \\
&\leq M'' \frac{1}{T^{2+\delta}} n^{-\frac{2+\delta}{2}} \sum_{t=1}^T E \left| \left(\sum_{i=1}^n \left| \bar{w}_{it-1}^* e_{it}^{h+*} \right|^{2+\delta} \right)^{\frac{2}{2+\delta}} n^{1-\frac{2}{2+\delta}} \right|^{\frac{2+\delta}{2}} \\
&\leq M''' \frac{1}{T^{2+\delta}} n^{-\frac{2+\delta}{2}} n^{\frac{2+\delta}{2}\left(1-\frac{2}{2+\delta}\right)} n \sum_{t=1}^T \left(\frac{1}{n} E \sum_{i=1}^n \left| \bar{w}_{it-1}^* e_{it}^{h+*} \right|^{2+\delta} \right) \\
&\leq M'''' \frac{1}{T^{2+\delta}} \sum_{t=1}^T \left(\sup_i E \left| \bar{w}_{it-1}^* \right|^{4+2\delta} \right)^{1/2} \left(\sup_i E \left| e_{it}^{h+*} \right|^{4+2\delta} \right)^{1/2},
\end{aligned}$$

where we have used, respectively, Burkholder's inequality (second line), Holder's inequality (third line), the C_r-inequality and convexity (fourth line), the Cauchy-Schwartz inequality (fifth line). By Assumption 1*, $E \left| e_{it}^{h+*} \right|^{4+\delta} < \infty$ for all i; also, $\bar{w}_{it-1}^* = \sum_{s=0}^{t-1} \Delta \bar{w}_{is}^*$ with

$\sup_{i,s} E\left|\Delta\bar{w}_{is}^{*}\right|^{4+2\delta} < \infty$, so that

$$\begin{aligned}\sum_{t=1}^{T} E\left|\frac{\xi_{nt}}{S_{nT}}\right|^{2+\delta} &\leq M\frac{1}{T^{2+\delta}}\sum_{t=1}^{T}\left(\sup_{i} E\left|\bar{w}_{it-1}^{*}\right|^{4+2\delta}\right)^{1/2} \\ &\leq M'\frac{1}{T^{2+\delta}}\sum_{t=1}^{T} t^{1+\delta/2}\left(\sup_{i,s} E\left|\frac{1}{\sqrt{t}}\sum_{s=0}^{t-1}\Delta\bar{w}_{is}^{*}\right|^{4+2\delta}\right)^{1/2} \leq M''\frac{1}{T^{2+\delta}}\sum_{t=1}^{T} t^{1+\delta/2} = O\left(\frac{1}{T^{\delta/2}}\right).\end{aligned}$$

Thus

$$\lim_{T\to\infty}\limsup_{n\to\infty}\sum_{t=1}^{T} E\left|\frac{\xi_{nt}}{S_{nT}}\right|^{2+\delta} = 0.$$

An FCLT for MDS can be applied (see e.g. Hall and Heyde, 1980, Theorem 4.1) uniformly in n: that $n\to\infty$ is only ancillary to the main argument of the proof. The covariance kernel of the process is given by $\lim_{n,T\to\infty} S_{n\lfloor Tr\rfloor}^{2}$, which follows from Lemma A.1. QED

We are now ready to report the proof of the main results. The proof of Theorem 1 is a special case of Proposition 1, and thus it is omitted. We present directly the proof of Proposition 1.

Proof of Proposition 1 It holds that

$$\sqrt{n}T\begin{bmatrix}\hat{\beta}_{1,\lfloor Tr\rfloor}^{FM}-\beta \\ \hat{\gamma}_{1,\lfloor Tr\rfloor}^{FM}-\gamma\end{bmatrix} = \left[\frac{1}{nT^{2}}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}\hat{e}_{it}^{+}-\tilde{\Lambda}_{he}^{+}\right)\right].$$

Since the covariance matrix is diagonal in view of Lemma A.2, we can study $\sqrt{n}T\left(\tilde{\beta}_{1,\lfloor Tr\rfloor}^{FM}-\beta\right)$ and $\sqrt{n}T\left(\tilde{\gamma}_{1,\lfloor Tr\rfloor}^{FM}-\gamma\right)$ separately. We define $\hat{e}_{it}^{+} = e_{it} - \Delta z_{it}^{\dagger\prime}\hat{\Omega}_{h}^{-1}\hat{\Omega}_{he}$.

Consider the numerator of $\sqrt{n}T\left(\tilde{\beta}^{FM}_{1,\lfloor Tr\rfloor}-\beta\right)$:

$$\begin{aligned}\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{f}_t\hat{e}^{+}_{it}-\hat{\Lambda}^{+}_{fe}\right) &= \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{f}_t e^{+}_{it}-\Lambda^{+}_{fe}\right)+\frac{\lfloor Tr\rfloor}{T}\sqrt{n}\left(\hat{\Lambda}_{fe}-\Lambda_{fe}\right)\\ &\quad -\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{f}_t\Delta f_t'-\Lambda_f\right)\left(\hat{\Omega}_f^{-1}\hat{\Omega}_{fe}-\Omega_f^{-1}\Omega_{fe}\right)\\ &\quad -\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_t\Delta x_{it}'\left(\hat{\Omega}_w^{-1}\hat{\Omega}_{we}-\Omega_w^{-1}\Omega_{we}\right)\\ &\quad +\frac{\lfloor Tr\rfloor}{T}\left(\hat{\Lambda}_f-\Lambda_f\right)\Omega_f^{-1}\Omega_{fe}\\ &= I+II+III+IV+V.\end{aligned}$$

Term II is bounded by $\sqrt{n}E\left\|\hat{\Lambda}_{fe}-\Lambda_{fe}\right\|$, which is $o_p(1)$ as shown by Phillips and Moon (1999). Turning to III, it is $o_p(1)$ due to $T^{-1}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{f}_t\Delta f_t'-\Lambda_f\right)$ being $O_p(1)$ and to $\left(\hat{\Omega}_f^{-1}\hat{\Omega}_{fe}-\Omega_f^{-1}\Omega_{fe}\right)$ $= o_p(1)$. As far as IV is concerned, $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{f}_t\Delta x_{it}'$ is at most $O_p(\sqrt{n})$; in view of $\hat{\Omega}_w^{-1}\hat{\Omega}_{we}-\Omega_w^{-1}\Omega_{we}$ being $o_p\left(n^{-1/2}\right)$, $IV=o_p(1)$. Finally, V is $o_p(1)$ in light of $\hat{\Lambda}_f-\Lambda_f$ being $o_p(1)$. The term that dominates is I, and Lemma A.4 gives its asymptotic distribution. The limiting distribution of the denominator is given by Lemma A.2; applying the CMT, the final result follows.

We now turn to $\sqrt{n}T\left(\tilde{\gamma}^{FM}_{1,\lfloor Tr\rfloor}-\gamma\right)$. Similar passages as above ensure that the asymptotics of the numerator is driven by $\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{x}_{it}e^{+}_{it}-\Lambda^{+}_{we}\right)$. Lemmas A.2 and A.4, and the CMT, yield the final result. QED

Proof of Proposition 2. Consider the estimation error

$$\sqrt{n}T\left[\hat{\theta}^{FM}_{1,\lfloor Tr\rfloor}-\theta\right]=\left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}\hat{e}^{+}_{it}-\hat{\Lambda}^{+}_{he}\right)\right].$$

Under Assumptions 1* and 3, the denominator has the same block-diagonal structure as in Lemma A.3, as discussed in the main body of the text. This entails that the asymptotics of $\sqrt{n}T\left[\hat{\beta}^{FM}_{1,\lfloor Tr\rfloor}-\beta\right]$ is the same as in Theorem 1. As far as $\sqrt{n}T\left[\hat{\gamma}^{FM}_{1,\lfloor Tr\rfloor}-\gamma\right]$ is concerned, its denominator converges to $\frac{1}{6}\Omega_w r^2$. Turning to the numerator:

$$\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{x}_{it}\hat{e}_{it}^{+}-\hat{\Lambda}_{we}^{+}\right)=\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{w}_{it}e_{it}^{+}-\Lambda_{we}^{+}\right)+o_p\left(1\right),$$

where the $o_p(1)$ term comes from the $\sqrt{n}$-consistency of the long run covariance matrices estimates, in a similar fashion as in the proof of Proposition 1; we use the notation Λ_{we}^{+} because, in view of Assumption 3*(a)*, g_t is expected not to contribute to the one-sided long run covariance. It holds that

$$\begin{aligned}\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{w}_{it}e_{it}^{+}-\Lambda_{we}^{+}\right) &= \frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{w}_{it}e_{it}^{h+}-\Lambda_{we}^{+}\right)+\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_te_{it}^{+}\\ &\quad+\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it}\Delta\bar{g}_t'\Gamma_i'\Omega_w^{-1}\Omega_{we}\\ &= I+II+III.\end{aligned}$$

As far as I is concerned, we know from the proof of Lemma A.4 that it converges to $\frac{1}{\sqrt{6}}\Omega_w^{1/2}\Omega_{e+}^{1/2}C_p\left(r;0\right)$ uniformly in r as $(n,T)\to\infty$ and $\frac{n}{T}\to 0$. We now show that II and III are negligible. Consider II, and note

$$II=\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_te_{it}^{h+}+\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_t\Delta g_t'\Gamma_i'\Omega_w^{-1}\Omega_{we}=II_a+II_b.$$

As far as II_a is concerned, note that $\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_te_{it}^{h+}$ has mean zero by virtue of Assumption 3*(a)*; also, conditional on g_t it is a cross-sectionally independent sequence in view of Assumption 1*. Thus, II_a is bounded by the square root of $\frac{1}{n}\sum_{i=1}^{n}E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_te_{it}^{h+}\right\|^2\leq\max_i\|\Gamma_i\|^2$ $\max_i E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{g}_te_{it}^{h+}\right\|^2$, which is $o(1)$ in view of Assumption 3. Turning to II_b, it is bounded by

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\Gamma_i\bar{g}_t\Delta g_t'\Gamma_i'\right\|\leq\sqrt{n}\max_i\|\Gamma_i\|^2\max_i E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{g}_t\Delta g_t'\right\|,$$

which is $o(1)$ if $\sqrt{n}\max_i\|\Gamma_i\|^2=o(1)$. This entails that $\max_i\|\Gamma_i\|=o\left(n^{-1/4}\right)$, which follows from part *(c)* of Assumption 3.

Finally, we consider III. The sequence $\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it}\Delta\bar{g}_t'\Gamma_i'$ has mean zero and it is, conditionally on g_t, cross-sectionally independent across i by Assumption 1*. Thus, III is bounded by the square root of $\frac{1}{n}\sum_{i=1}^{n} E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it}\Delta\bar{g}_t'\Gamma_i'\right\|^2 \leq \max_i \|\Gamma_i\|^2 \max_i E\left\|\frac{1}{T}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{w}_{it}\Delta\bar{g}_t'\right\|^2$, which is $o(1)$ in view of Assumption 3. QED

Appendix B: proofs of results in Section 4

In this Appendix, we report the proof of Theorems 2 and 3. Prior to reporting the latter, we show two preliminary lemmas (Lemmas B.1 and B.2), and Lemma B.3, where we discuss the $\sqrt{n}$-consistency of the long run covariance matrices estimators $\hat{\Omega}_w$, $\hat{\Omega}_e$, $\hat{\Lambda}_w$, $\hat{\Omega}_{we}$, $\hat{\Omega}_{fe}$, $\hat{\Lambda}_{we}$ and $\hat{\Lambda}_{fe}$ under the null and the alternative, for both cases of stationary and nonstationary g_t.

Proof of Theorem 2. The proof of the Theorem is an application of Theorem 1 and of Propositions 1 and 2. We show the main passages that lead to show (21)-(22), and then show that $Q_k(r)$ and $Q_p(r)$ have a chi-squared distribution for fixed r; passages are reported for the setup of Section 3.

Consider (21)-(22). Theorem 1 and Propositions 1 and 2 state that the limiting distribution of $\hat{\beta}$ and $\hat{\gamma}$ are independent. This is a consequence of Lemma A.2, which states that matrix $H_{nT}(r)$ in (18) is diagonal. Indeed, by Lemma A.2

$$H_{nT}(r) \Rightarrow \begin{bmatrix} \Omega_f^{-1/2}\left[\left(\int_0^r \bar{W}_k\bar{W}_k'\right)^{-1} + \left(\int_r^1 \bar{W}_k\bar{W}_k'\right)^{-1}\right]\Omega_f^{-1/2} & 0 \\ 0 & \frac{1}{6}\Omega_w\left[\frac{1}{r^2}+\frac{1}{(1-r)^2}\right] \end{bmatrix}, \tag{29}$$

where $\bar{W}_k$ is a k-dimensional demeaned standard Brownian motion, and Ω_f is the long run covariance matrix of f_t. Thus, (21) and (22) can be shown separately, using the CMT.

In order to show that $Q_k(r)$ and $Q_p(r)$ have a chi-squared distribution pointwise in r, consider $Q_k(r)$ first. We use the fact that $\int_0^r \bar{W}_k dW_u \stackrel{d}{=} \left(\int_0^r \bar{W}_k\bar{W}_k'\right)^{1/2} Z$, where W_u is a standard Brownian motion that is independent of W_k and Z a k-dimensional standard normal independent of W_k. It holds that $\sqrt{n}T\left(\hat{\beta}_{1,\lfloor Tr\rfloor}^{FM}-\hat{\beta}_{2,\lfloor Tr\rfloor}^{FM}\right) = \sqrt{n}T\left(\hat{\beta}_{1,\lfloor Tr\rfloor}^{FM}-\beta\right) - \sqrt{n}T\left(\hat{\beta}_{2,\lfloor Tr\rfloor}^{FM}-\beta\right)$; using

Theorem 1

$$\sqrt{n}T\left(\hat{\beta}^{FM}_{1,\lfloor Tr\rfloor}-\hat{\beta}^{FM}_{2,\lfloor Tr\rfloor}\right)\Rightarrow\Omega^{1/2}_{e+}\Omega^{-1/2}_{f}\left[\left(\int_0^r\bar{W}_k\bar{W}_k'\right)^{-1}\int_0^r\bar{W}_kdW_u-\left(\int_r^1\bar{W}_k\bar{W}_k'\right)^{-1}\int_r^1\bar{W}_kdW_u\right]$$
$$\stackrel{d}{=}\Omega^{1/2}_{e+}\Omega^{-1/2}_{f}\left[\left(\int_0^r\bar{W}_k\bar{W}_k'\right)^{-1/2}Z_1-\left(\int_r^1\bar{W}_k\bar{W}_k'\right)^{-1/2}Z_2\right].$$

Since Z_1 and Z_2 are independent, conditioning on the σ-field associated with $\bar{W}_k$, we have

$$\Omega^{1/2}_{e+}\Omega^{-1/2}_{f}\left[\left(\int_0^r\bar{W}_k\bar{W}_k'\right)^{-1/2}Z_1-\left(\int_r^1\bar{W}_k\bar{W}_k'\right)^{-1/2}Z_2\right]$$
$$\stackrel{d}{=}\Omega^{1/2}_{e+}\Omega^{-1/2}_{f}\left[\left(\int_0^r\bar{W}_k\bar{W}_k'\right)^{-1}+\left(\int_r^1\bar{W}_k\bar{W}_k'\right)^{-1}\right]^{1/2}Z,$$

where Z is a k-dimensional standard normal independent of W_k. Using (29), $s(r)'M^{-1}(r)s(r)$ $=Z'Z\sim\chi^2_k$. This is true for all possible elements of the σ-field associated with $\bar{W}_k$. Turning to $Q_p(r)$, we have $\sqrt{n}T\left(\hat{\gamma}^{FM}_{1,\lfloor Tr\rfloor}-\hat{\gamma}^{FM}_{2,\lfloor Tr\rfloor}\right)=\sqrt{n}T\left(\hat{\gamma}^{FM}_{1,\lfloor Tr\rfloor}-\gamma\right)-\sqrt{n}T\left(\hat{\gamma}^{FM}_{2,\lfloor Tr\rfloor}-\gamma\right)$; thus, using Theorem 1 and Propositions 1 and 2

$$\sqrt{n}T\left(\hat{\gamma}^{FM}_{1,\lfloor Tr\rfloor}-\hat{\gamma}^{FM}_{2,\lfloor Tr\rfloor}\right)\Rightarrow\sqrt{6}\Omega^{1/2}_{e+}\Omega^{-1/2}_{w}\left[\frac{1}{r^2}C_p(r;0)-\frac{1}{(1-r)^2}C_p(1;r)\right];$$

in view of Lemma A.1, this means that, for fixed r, $\sqrt{n}T\left(\hat{\gamma}^{FM}_{1,\lfloor Tr\rfloor}-\hat{\gamma}^{FM}_{2,\lfloor Tr\rfloor}\right)\Rightarrow\sqrt{6}\Omega^{1/2}_{e+}\Omega^{-1/2}_{w}$ $N\left[0,\frac{r^2+(1-r)^2}{r^2(1-r)^2}I_k\right]$. Using (29), $Q_p(r)$ is, for every r

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

Lemma B.1 *Consider the full sample estimator $\hat{\theta}$ defined in (8). Let Assumptions 1*, 2* *and 4 hold, and let* $\Omega_{\Delta g,j} = E\left[\Delta g_t \Delta g_{t-j}'\right]$. *Under both* H_0 *and* H_a*, it holds that, for all* j

$$
\begin{aligned}
\hat{\theta} - \theta &= O_p\left(\frac{1}{T}\right); & (30)\\
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{e}_{it}\hat{e}_{it-j} - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}e_{it-j} &= O_p\left(\frac{1}{T}\right); & (31)\\
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it-j}' - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta w_{it-j}' &= \frac{1}{n}\sum_{i=1}^{n}\Gamma_i\Omega_{\Delta g,j}\Gamma_i' & (32)\\
&\quad + O_p\left(T^{-1/2}\right) + O_p\left(n^{-1/2}T^{-1/2}\right); \\
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it-j}'\hat{e}_{it}\hat{e}_{it-j} - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it-j}'e_{it}e_{it-j} &= O_p\left(T^{-1}\right); & (33)\\
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it-j}'e_{it}e_{it-j} - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta w_{it-j}'e_{it}e_{it-j} &= O_p\left(T^{-1}\right). & (34)
\end{aligned}
$$

Proof. Let d_i be a dummy variable that is equal to 1 if unit i has a break and zero otherwise. Consider (30), and recall equation (8); under H_a we have

$$
\begin{aligned}
\hat{\theta} - \theta &= \left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}e_{it}\right] \\
&\quad + \left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\frac{1}{nT^2}\frac{\sqrt{n}}{m_nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}\bar{z}_{it}'\psi_i\left(\frac{t}{T}\right)d_i\right];
\end{aligned}
$$

as far as the first term is concerned, it is present under H_0 also. From equation (5.12) in Phillips and Moon (1999, p. 1083), it follows that it is $O_p\left(T^{-1}\right)$. As far as the second term is concerned, it is present only under H_a; using Assumption 4*(b)*, it holds that it is of magnitude $O_p\left(\frac{1}{\sqrt{n}T}\right)$. Hence, (30) follows immediately.

Consider (31); we report the proof for the case $j = 0$. It holds that

$$
\begin{aligned}
\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{e}_{it}^2 - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}e_{it}^2 &= T\left(\hat{\theta} - \theta\right)'\left(\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}\bar{z}_{it}'\right)\left(\hat{\theta} - \theta\right) \\
&\quad -2\left(\hat{\theta} - \theta\right)'\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}e_{it}\right) = I + II.
\end{aligned}
$$

Term I has the same order of magnitude as $T\left\|\hat{\theta} - \theta\right\|^2\left\|\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}\bar{z}_{it}'\right\|^2$, which, in view of (30) and Lemma A.3, is $O_p\left(\frac{1}{T}\right)$. Also, $II \leq \left\|\hat{\theta}-\theta\right\| \left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{z}_{it}e_{it}\right\| = O_p\left(\frac{1}{T}\right)$ again using (30) and some passages in its proof. Putting all together, (31) follows.

We turn to (32). It holds that

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it-j}' - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta w_{it-j}' \\
= \; & \frac{1}{n}\sum_{i=1}^{n}\Gamma_i\left(\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_{t-j}'\right)\Gamma_i' + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Gamma_i\Delta g_t\Delta w_{it-j}' \\
& + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta g_{t-j}'\Gamma_i' \\
= \; & I + II + III.
\end{aligned} \tag{35}$$

Consider I. Under Assumptions 1* and 2, $\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_{t-j}'$ converges to $\Omega_{\Delta g,j}$; also, $\frac{1}{T}\sum_{t=1}^{T}\left(\Delta g_t\Delta g_{t-j}' - \Omega_{\Delta g,j}\right) = O_p\left(T^{-1/2}\right)$ since $\Delta g_t\Delta g_{t-j}' - \Omega_{\Delta g,j}$ has mean zero by construction and $\Delta g_t\Delta g_{t-j}'$ has finite second moment (in view of Assumption 2(a)). Thus, $\frac{1}{n}\sum_{i=1}^{n}\Gamma_i\left(\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_{t-j}'\right)\Gamma_i'$ $= \frac{1}{n}\sum_{i=1}^{n}\Gamma_i\Omega_{\Delta g,j}\Gamma_i' + O_p\left(T^{-1/2}\right)$. Turning to II, this is bounded by the square root of $\max_i\left\|\Gamma_i\right\|^2 E\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Delta g_t\Delta w_{it-j}'\right\|^2$, so that $II = O_p\left(\frac{1}{\sqrt{nT}}\right)$ in view of Assumption 2*(c)*. Term III is the transpose of II. Thus, (32) follows putting everything together.

As far as (33) is concerned, we focus on $j=0$. It holds that

$$\begin{aligned}
&\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'\hat{e}_{it}^2 - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'e_{it}^2 \\
= \; & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'\left(\hat{\theta}-\theta\right)'\bar{z}_{it}\bar{z}_{it}'\left(\hat{\theta}-\theta\right) + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}'\left(\hat{\theta}-\theta\right)'\bar{z}_{it}e_{it} \\
= \; & I + II.
\end{aligned}$$

Consider I; by the Cauchy-Schwartz inequality, $I \leq T\left\|\hat{\theta}-\theta\right\|^2\left[T^{-1}\sum_{t=1}^{T}\left\|\Delta x_{it}\right\|^4\right]^{1/2}\left[T^{-3}\sum_{t=1}^{T}\left\|\bar{z}_{it}\right\|^4\right]^{1/2}$ $= O_p\left(T^{-1}\right)$. This follows from Theorem 5.3 in Park and Phillips (1999), whereby $\sum_{t=1}^{T}\left\|\bar{z}_{it}\right\|^4 = O_p\left(T^3\right)$. As far as II is concerned, using cross-sectional independence, it is bounded by the square root of $n^{-1}T^{-2}\left\|\hat{\theta}-\theta\right\|^2\sum_{i=1}^{n}\sum_{t=1}^{T}E\left[\left\|\Delta x_{it}\right\|^4\left\|\bar{z}_{it}\right\|^2 e_{it}^2\right]$, which is of order $O_p\left(T^{-1}\right)$.

Finally, consider (34), focusing on $j = 0$; we have

$$\begin{aligned}
& \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it}' e_{it}^2 - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta w_{it}' e_{it}^2 \\
= & \frac{1}{n}\sum_{i=1}^{n}\Gamma_i\left(\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_t' e_{it}^2\right)\Gamma_i' + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Gamma_i\Delta g_t\Delta w_{it}' e_{it}^2 + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta g_t'\Gamma_i' e_{it}^2 \\
= & I + II + III.
\end{aligned}$$

Consider I. Assumption 2 states that g_t is independent of e_{it}. Thus, the same arguments as for the proof of (33) hold. Similar considerations apply to II and III which is the transpose of II.

Lemma B.2 *Let Assumptions 1*, 3 and 4 hold. Under both H_0 and H_a, it holds that, for all j, (31), (33) and (34) hold. Also*

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it-j}' - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta w_{it-j}' = o_p\left(\frac{1}{n^{1/2+\omega}}\right) + O_p\left(\frac{1}{n^{3/4+\omega/2}\sqrt{T}}\right), \quad (36)$$

where $\omega > 0$ is defined in Assumption 3(c).

Proof. Equations (31), (33) and (34) are not affected by the stationarity or non-stationarity of g_t, as it emerges from the proofs of Lemma B.1. Considering (36), we use

$$\begin{aligned}
& \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta x_{it}\Delta x_{it-j}' - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta w_{it-j}' \\
= & \frac{1}{n}\sum_{i=1}^{n}\Gamma_i\left(\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_{t-j}'\right)\Gamma_i' + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Gamma_i\Delta g_t\Delta w_{it-j}' + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Delta w_{it}\Delta g_{t-j}'\Gamma_i' \\
= & I + II + III.
\end{aligned}$$

Term I is bounded by $\max_i \|\Gamma_i\|^2 \left\|\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_t'\right\|$, which is $o_p\left(n^{-1/2-\omega}\right)$ by Assumption 3*(c)*. Considering II, using the same bound as for the case of stationary g_t, part *(c)* of Assumption 3 yields $II = O_p\left(\frac{1}{n^{3/4+\omega/2}T^{1/2}}\right)$. The same applies to III.

Lemma B.3 *Let $\kappa\left(\frac{j}{l}\right)$ be a kernel with Parzen exponent $\psi > 1/2$ and bandwidth $l \to \infty$.*

Let Assumptions 1, 2 and 4 hold, and assume that equation (16) holds. As $(n, T) \to \infty$, all estimates in (9) are $\sqrt{n}$-consistent under both H_0 and H_a.*

Let Assumptions 1, 3 and 4 hold and assume that equation (17) holds. As $(n, T) \to \infty$, all* *estimates in (9) are $\sqrt{n}$-consistent under both H_0 and H_a.*

Proof. We start by proving the $\sqrt{n}$-consistency of $\hat{\Omega}_e$; the proof is valid under both cases of stationary and nonstationary g_t. Let $\tilde{\Phi}_{ei,j} = T^{-1}\sum_{t=j+1}^{T} e_{it}e_{it-j}$; define $\tilde{\Omega}_e = \tilde{\Phi}_{ei,0} + 2\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\tilde{\Phi}_{ei,j}$. We have

$$\hat{\Omega}_e - \Omega_e = \tilde{\Omega}_e - \Omega_e + \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\Phi}_{ei,0} - \tilde{\Phi}_{ei,0}\right) + 2\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\left(\hat{\Phi}_{ei,j} - \tilde{\Phi}_{ei,j}\right).$$

Phillips and Moon (1999, p. 1109) show that $\sqrt{n}\left\|\tilde{\Omega}_e - \Omega_e\right\| = o_p(1)$. As far as $\hat{\Phi}_{ei,j} - \tilde{\Phi}_{ei,j}$ is concerned, (31) states that it is of order $O_p\left(T^{-1}\right)$; hence, $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\left(\hat{\Phi}_{ei,j} - \tilde{\Phi}_{ei,j}\right)$ $= O_p(l/T)$, so that, putting all together, $\sqrt{n}\left\|\hat{\Omega}_e - \Omega_e\right\| = o_p(1) + O_p(l\sqrt{n}/T)$, which is $o_p(1)$.

We now turn to $\hat{\Omega}_w$; the proof for the $\sqrt{n}$-consistency of $\hat{\Lambda}_w$ is similar, and thus omitted. We consider the case of stationary g_t first. Let $\tilde{\Phi}_{wi,j} = T^{-1}\sum_{t=j+1}^{T}\Delta w_{it}\Delta w_{it-j}'$; define $\tilde{\Omega}_w = \tilde{\Phi}_{wi,0} + 2\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\tilde{\Phi}_{wi,j}$. We have

$$\hat{\Omega}_w - \Omega_w = \tilde{\Omega}_w - \Omega_w + \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\Phi}_{wi,0} - \tilde{\Phi}_{wi,0}\right) + 2\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\left(\hat{\Phi}_{wi,j} - \tilde{\Phi}_{wi,j}\right).$$

As before, $\sqrt{n}\left\|\tilde{\Omega}_w - \Omega_w\right\| = o_p(1)$. Also

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\Phi}_{wi,0} - \tilde{\Phi}_{wi,0}\right) + 2\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\left(\hat{\Phi}_{wi,j} - \tilde{\Phi}_{wi,j}\right) \\
= & \frac{1}{n}\sum_{i=1}^{n}\Gamma_i\left(\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_t' - \Omega_{\Delta g}\right)\Gamma_i' + 2\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\Gamma_i\left(\frac{1}{T}\sum_{t=1}^{T}\Delta g_t\Delta g_{t-j}' - \Omega_{\Delta g,j}\right)\Gamma_i' + II \\
= & I + II,
\end{aligned}$$

where the remainder II comes from (35). Indeed, II is bounded by $O_p\left(\frac{l}{\sqrt{n}T}\right)$, so that $\sqrt{n}II$ $= O_p\left(\frac{l}{\sqrt{T}}\right) = o_p(1)$ under (16). Considering I, Assumption 2 entails that Δg_t is a stationary process with finite fourth order cumulant; thus, Proposition 1 in Andrews (1991) yields $I =$

$O\left(\sqrt{l/T}\right)+O\left(l^{-\psi}\right)$. Therefore

$$\sqrt{n}\left\|\hat{\Omega}_w-\Omega_w\right\|=o_p(1)+O_p\left(\frac{l}{\sqrt{T}}\right)+O_p\left(\sqrt{\frac{nl}{T}}\right)+O_p\left(\frac{\sqrt{n}}{l^{\psi}}\right),$$

which is $o_p(1)$ under (16). The restriction $\frac{n^{1+\varepsilon}}{T}\to 0$ comes from noting that $n=o\left(l^{2\psi}\right)$, so that in order for $\frac{nl}{T}\to 0$, it is necessary that $\frac{n^{1+1/(2\psi)+\delta}}{T}\to 0$ for some $\delta>0$.

When g_t is nonstationary, using (36), we have

$$\sqrt{n}\left\|\hat{\Omega}_w-\Omega_w\right\|=o_p(1)+o_p\left(ln^{-\omega}\right)+O_p\left(\frac{l}{n^{1/4+\omega/2}\sqrt{T}}\right),$$

which is $o_p(1)$ under (17).

Finally, consider $\hat{\Omega}_{we}$ (and, similarly, $\hat{\Lambda}_{we}$). Let $\tilde{\Phi}_{wei,j}=T^{-1}\sum_{t=j+1}^{T}\Delta w_{it}\Delta w'_{it-j}e_{it}e_{it-j}$; define $\tilde{\Omega}_{we}=\tilde{\Phi}_{wei,0}+2\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\tilde{\Phi}_{wei,j}$. We have $\hat{\Omega}_{we}=\tilde{\Omega}_{we}-\Omega_{we}+\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\Phi}_{wei,0}-\tilde{\Phi}_{wei,0}\right)$ $+2\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{l}\kappa\left(\frac{j}{l}\right)\left(\hat{\Phi}_{wei,j}-\tilde{\Phi}_{wei,j}\right)$. Using (33) and (34), the proof is the same as for $\hat{\Omega}_w$ in both cases of stationary and nonstationary g_t.

Proof of Theorem 3 We consider the setup of Section 3. The proof requires showing: that, under H_a, $\sqrt{n}T\left[\hat{\theta}^{FM}_{1\lfloor Tr\rfloor}-\hat{\theta}^{FM}_{2\lfloor Tr\rfloor}\right]$ converges to $\frac{\sqrt{6}}{r^2}\,\Omega_w^{-1/2}\,\Omega_{e+}^{1/2}\,\left[C_p(r;0)-C_p(1;r)\right]$ plus a noncentrality parameter; and that estimated long run covariance matrices are consistent under H_a. The latter requirement follows from Lemma B.3.

Consider $\sqrt{n}T\left[\hat{\theta}^{FM}_{1\lfloor Tr\rfloor}-\hat{\theta}^{FM}_{2\lfloor Tr\rfloor}\right]=\sqrt{n}T\left[\hat{\theta}^{FM}_{1\lfloor Tr\rfloor}-\theta\right]-\sqrt{n}T\left[\hat{\theta}^{FM}_{2\lfloor Tr\rfloor}-\theta\right]$, and focus on $\sqrt{n}T$ $\left[\hat{\theta}^{FM}_{1\lfloor Tr\rfloor}-\theta\right]$. We have

$$\begin{aligned}\sqrt{n}T\left(\hat{\theta}^{FM}_{1\lfloor Tr\rfloor}-\theta\right) &= \left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}'_{it}\right]^{-1}\left[\frac{1}{\sqrt{n}T}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\left(\bar{z}_{it}\hat{e}^{+}_{it}-\tilde{\Lambda}^{+}_{he}\right)\right]\\ &\quad+\left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}'_{it}\right]^{-1}\left[\frac{1}{m_nT}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}'_{it}\psi_i\left(\frac{t}{T}\right)\right]\\ &= I+II.\end{aligned}$$

Term I is not affected by H_a. As far as II is concerned, it is $O_p(1)$ by Assumption 4*(b)*. This entails that $\sqrt{n}T\left[\hat{\theta}^{FM}_{1\lfloor Tr\rfloor}-\hat{\theta}^{FM}_{2\lfloor Tr\rfloor}\right]$ has noncentrality parameter given by

$$\begin{aligned}
&\left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\frac{1}{m_nT^2}\sum_{i=1}^{m_n}\sum_{t=1}^{\lfloor Tr\rfloor}\bar{z}_{it}\bar{z}_{it}'\psi_i\left(\frac{t}{T}\right)\right] \\
&-\left[\frac{1}{nT^2}\sum_{i=1}^{n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\bar{z}_{it}\bar{z}_{it}'\right]^{-1}\left[\frac{1}{m_nT^2}\sum_{i=1}^{m_n}\sum_{t=\lfloor Tr\rfloor+1}^{T}\bar{z}_{it}\bar{z}_{it}'\psi_i\left(\frac{t}{T}\right)\right] \\
\equiv\ & D_{nT}(r),
\end{aligned}$$

from whence we define

$$d(r)=\lim_{n,T\to\infty}\left[H_{nT}(r)\right]^{-1/2}D_{nT}(r). \tag{37}$$

Putting all together, the Theorem follows. QED

References

Andrews, D. W. K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817-858.

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

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

Bai, J., 2003. Inferential theory for structural models of large dimensions. Econometrica 71, 135-171.

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

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.

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

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

Breitung, J., Eickmeier, S., 2011. Testing for structural breaks in dynamic factor models. Journal of Econometrics 163, 71-84.

Chen, L., Dolado, J.J., Gonzalo, J., 2011. Detecting big structural breaks in large factor models. Mimeo.

Csorgo, M., Horvath, L., 1997. Limit theorems in change-point analysis. Wiley, Chichester.

Davidson, J., 2002. Stochastic Limit Theory. Oxford University Press, Oxford.

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

Hall, P., Heyde, C.C., 1980. Martingale limit theory and its applications. New York, University Press.

Han, X., Inoue, A., 2011. Tests for parameter instability in dynamic factor models. Mimeo.

Hansen, B.H., 1992. Testing for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics 10, 321-335.

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

Kao, C., Chiang, M-H., 2000. Testing for structural change of a cointegrated regression in panel data. Mimeo

Kao, C., Trapani, L., Urga, G., 2011a. Testing for breaks in cointegrated panels with common and idiosyncratic stochastic trends. Working Paper N. WP-CEA-02-2011, Centre for Econometric Analysis, Cass Business School.

Kao, C., Trapani, L., Urga, G., 2011b. Asymptotics for panel models with common shocks. Econometric Reviews (forthcoming).

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.

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

Kim, D., 2010. Common local breaks in time trends for large panel data. Mimeo. Department of Economics, University of Virginia.

Kim, D., 2011. Estimating a common deterministic time trend break in large panels with cross sectional dependence. Journal of Econometrics 164, 310-330.

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

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

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.

Phillips, P.C.B., 1988. Weak convergence of sample covariance matrices to stochastic integrals via martingale approximations. Econometric Theory 4, 528-533.

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

Pollard, D., 1984. Convergence of probability measures. Springer-Verlag, New York.

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

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

n	T	(ρ, ϑ)	(0, 0)			(0.7, 0)			(0, 0.7)			(0.7, 0.7)		
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
20	50		0.102	0.053	0.081	0.135	0.060	0.104	0.079	0.041	0.062	0.128	0.049	0.100
	100		0.089	0.057	0.071	0.119	0.047	0.094	0.058	0.033	0.053	0.111	0.043	0.076
	200		0.060	0.040	0.042	0.062	0.042	0.058	0.045	0.037	0.045	0.062	0.049	0.046
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
50	50		0.112	0.050	0.094	0.159	0.069	0.122	0.100	0.035	0.065	0.162	0.056	0.117
	100		0.097	0.055	0.067	0.101	0.042	0.070	0.065	0.027	0.046	0.093	0.038	0.061
	200		0.056	0.038	0.044	0.080	0.050	0.052	0.043	0.035	0.043	0.070	0.046	0.048
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
100	50		0.108	0.066	0.095	0.145	0.064	0.116	0.087	0.036	0.066	0.138	0.052	0.108
	100		0.095	0.055	0.075	0.117	0.057	0.090	0.064	0.033	0.052	0.107	0.052	0.082
	200		0.057	0.037	0.043	0.065	0.036	0.056	0.040	0.035	0.036	0.065	0.035	0.053

Table 1.Empirical rejection frequencies at 5% level for the null of no change in γ.

n	T	(ρ, ϑ)	$(0, 0)$			$(0.7, 0)$			$(0, 0.7)$			$(0.7, 0.7)$		
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
20	50		0.031	0.040	0.037	0.042	0.034	0.042	0.022	0.026	0.033	0.044	0.034	0.042
	100		0.052	0.045	0.054	0.133	0.061	0.105	0.048	0.032	0.033	0.150	0.074	0.124
	200		0.181	0.097	0.151	0.458	0.337	0.440	0.206	0.094	0.166	0.493	0.380	0.477
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
50	50		0.035	0.050	0.045	0.024	0.024	0.026	0.057	0.055	0.058	0.060	0.048	0.056
	100		0.101	0.075	0.079	0.295	0.170	0.251	0.124	0.074	0.092	0.326	0.203	0.294
	200		0.471	0.347	0.439	0.880	0.869	0.885	0.673	0.560	0.649	0.902	0.892	0.920
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
100	50		0.065	0.086	0.080	0.067	0.047	0.060	0.033	0.037	0.032	0.082	0.049	0.061
	100		0.228	0.137	0.204	0.539	0.454	0.539	0.284	0.148	0.240	0.593	0.511	0.586
	200		0.779	0.762	0.797	0.995	0.997	0.999	0.946	0.931	0.949	1.000	1.000	1.000

Table 2. Rejection frequencies under the alternative of a change in γ. The alternative is specified as in (24), setting $c = 0.25$.

n	T	(ρ, ϑ)	$(0, 0)$			$(0.7, 0)$			$(0, 0.7)$			$(0.7, 0.7)$		
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
20	50		0.045	0.060	0.060	0.070	0.052	0.078	0.024	0.024	0.024	0.073	0.058	0.077
	100		0.167	0.106	0.136	0.447	0.331	0.434	0.236	0.115	0.197	0.498	0.386	0.487
	200		0.617	0.523	0.610	0.950	0.936	0.953	0.818	0.741	0.819	0.960	0.948	0.963
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
50	50		0.127	0.172	0.160	0.168	0.135	0.166	0.083	0.085	0.094	0.196	0.160	0.200
	100		0.486	0.396	0.469	0.853	0.814	0.863	0.618	0.515	0.602	0.907	0.885	0.917
	200		0.920	0.919	0.927	0.999	1.000	1.000	0.990	0.992	0.992	1.000	1.000	1.000
			$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$	$AveW$	$SupW$	$ExpW$
100	50		0.263	0.300	0.296	0.370	0.320	0.370	0.199	0.184	0.199	0.435	0.393	0.440
	100		0.740	0.742	0.765	0.993	0.988	0.992	0.868	0.862	0.877	0.996	0.996	0.997
	200		0.973	0.982	0.980	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000

Table 3. Rejection frequencies under the alternative of a change in γ. The alternative is specified as in (24), setting $c = 0.5$.