A CUSUM TEST OF COMMON TRENDS IN LARGE HETEROGENEOUS PANELS

JAVIER HIDALGO AND JUNGYOON LEE

Abstract. This paper examines a nonparametric CUSUM-type test for common trends in large panel data sets with individual fixed effects. We consider, as in Zhang, Su and Phillips (2012), a partial linear regression model with unknown functional form for the trend component, although our test does not involve local smoothings. This conveniently forgoes the need to choose a bandwidth parameter, which due to a lack of a clear and sensible information criteria it is difficult for testing purposes. We are able to do so after making use that the number of individuals increases with no limit. After removing the parametric component of the model, when the errors are homoscedastic, our test statistic converges to a Gaussian process whose critical values are easily tabulated. We also examine the consequences of having heteroscedasticity as well as discussing the problem of how to compute valid critical values due to the very complicated covariance structure of the limiting process. Finally, we present a small Monte-Carlo experiment to shed some light on the finite sample performance of the test.

1. INTRODUCTION

One of the oldest issues and main concerns in time series is perhaps the study and modelling of their trend behaviour. The literature is large and vast, see for instance an overview by Phillips (2001), or more recently by White and Granger (2011) on some practicalities regarding trends for the purpose of, say, prediction. A huge amount of the literature has focused on the topic of stochastic vs. deterministic trends. On the other hand, more recently there has been an interest on how analysis of macroeconomic data may benefit from using cross-sectional information. One of the earliest works is given by Phillips and Moon (1999) on testing for unit roots in panel data, see also the recent work on macroeconomic convergence by Phillips and Sul (2007), and Breitung and Pesaran (2008) for a survey on stochastic trends in panel data models. Recently there has also been a surge of interest on examining deterministic trending regression models. Among others and

Date: **26 June 2014**.

Key words and phrases. **Common Trends, large data set, Partial linear models, Bootstrap algorithms.**

JEL Classification: C12, C13, C23.

We thanks the comments of a referee which have led to a clearer version of the manuscript. Of course, any remaining errors are our sole responsibility.

in a context with spatial dependence, see Robinson (2011*a*) who considers nonparametric trending models or Robinson (2011*b*) for a general parametric model where the trending regressors are of a non polynomial type t^{γ} with $\gamma > -1/2$. On the other hand, semiparametric models have been considered, see Gao and Hawthorne (2006) and in panel data models by Atak, Linton and Xiao (2011) or Zhang, Su, Xu and Phillips (2012) or Degras, Xu, Zhang and Wu (2012). The work by Atak et al. (2011) assumes that the number of cross-section units is finite although the trending regression functions are common among the different units. On the contrary, the latter two references allow for the number of cross-section units to increase without limit and they consider a test for common trends, or, in the language of panel data analysis for homogeneity/poolability.

This paper belongs to the latter framework. More specifically, we are concerned with testing for common trends in the partial panel linear regression model

$$\begin{aligned} y_{it} &= \beta' x_{it} + m_i(t/T) + \upsilon_{it}, \quad i = 1, ..., n, \quad t = 1, ..., T, \\ \upsilon_{it} &= \mu_i + \varepsilon_{it}, \end{aligned} \tag{1.1}$$

where $\{x_{it}\}_{t\geq 1}$, $i \in \mathbb{N}^+$, are $k-$vector of exogenous variables, $m_i(t/T)$ is left to be of unknown functional form, μ_i are the individual fixed effects and the error terms $\{\varepsilon_{it}\}_{t\geq 1}$, $i \in \mathbb{N}^+$, are sequences of zero mean random variables, possibly heteroscedastic. We shall remark that the reason not to introduce fixed time effects in (1.1), is because the tests are invariant to its presence, see (2.5) or (2.6) below. In addition, model (1.1) is a generalization of that discussed in Degras, et al. (2012), where we have now some covariates x_{it} in the model. We are interested, to be more precise the next section, on testing the hypothesis

$$H_0: \quad m_i(\tau) = m(\tau) \quad \text{for all } i \geq 1 \ \text{ and } \tau \in [0,1].$$

This hypothesis testing can be put into a much broader context of testing for equality among a set of curves, such as the conditional expectation and/or the dynamic structure in time series. More generally, our hypothesis can arise when we wish to decide how similar two or more groups of individuals are. When the number of curves is finite, there is a substantial literature. Relevant examples are the classical poolability test in panel data, see Hsiao (1986) and references therein, for parametric models, and Baltagi, Hidalgo and Li (1996) in a nonparametric scenario. In a time series context, we can mention the work by Detter and Paparoditis (2008) and references there, where the interest is to decide whether the covariance structure is the same across the sequences.

Nowadays due to the amount of available data, it is not unrealistic to consider the number of series large or that they increase without limit. Among the latter framework, we can cite Jin and Su (2010) and Degras et al. (2012) and in a partially linear regression model the work by Zhang et al. (2012).

Also it worth mentioning the recent work by Hidalgo and Souza (2013) who tested if the dynamic structure across sequences is the same.

It is worth noting that we can generalize our model (1.1) as

$$y_{it} = \beta' x_{it} + g_i(d_t) + \upsilon_{it},$$

where, as in Pesaran and Tosetti (2011) or more recently Bryan and Jenkins (2013), d_t has the interpretation of being a sequence of observed common effects/factors across individuals. See also Galvao and Kato (2013), where d_t has the meaning of being fixed effects. With this interpretation, if $g_i(d_t) = g_i d_t$, say, we can regard the latest display model as a panel data with interactive fixed effects, so that our hypothesis becomes $g_i = g$, i.e. we are testing the existence of interactive effects versus a more standard additive fixed effects type of model. Bearing this in mind, we focus nevertheless on the case where $d_t = t/T$ merely for historical reasons, although our main results seem to hold true for the last displayed model.

We finish this section relating the results of the paper with the problem of classification with functional data sets, which is a topic of active research. The reason being that this paper tackles the problem of whether a set of curves, that is the trending functions, are the same or not among a set of individual units. Within the functional data analysis framework, this question translates into whether there is some common structure or if we can split the set (of curves) into several classes or groups. See classical examples in Ferraty and Vieu (2006), although their approach uses nonparametric techniques which we try to avoid so that the issues of how to choose a bandwidth parameter and/or the "metric" to decide closeness are avoided. With this in mind, we believe that our approach can be used for a classification scheme. For instance, in economics are the trend components across different industries the same? In the language of functional data analysis, this is a problem of supervised classification which is nothing more than a modern name to one of the oldest statistical problems: namely to decide if an individual belongs to a particular population. The term supervised refers to the case where we have a "training" sample which has been classified without error. Moreover, we can envisage that our methodology can be extended to problems dealing with functional data in a framework similar to those examined by Chang and Ogden (2009).

The remainder of the paper is organized as follows. In the next section, we introduce the test statistics and their limiting distributions in a model without the regressors x_{it}. Section 3 considers the model (1.1), and we explicitly examine the consequences when we allow for heteroscedasticity, that is $E(\varepsilon_{it}) = \sigma_i^2$ for $i \in \mathbb{N}^+$. We show that the limiting process has quite a messy covariance structure to be of any use with real data sets. Because of this, we then examine a transformation to overcome the latter concerned. Section 4 describes the local alternatives for which the tests have non trivial power and their consistency. We also describe and show the validity of a bootstrap version of the test, as our simulation results illustrate that tests based on a subset of proposed statistics yield a very poor finite sample performance, see Table 2. Section 5 presents a Monte Carlo experiment to get some idea of the finite sample performance of our tests. Section 6 concludes and finally we confine the proofs to the Appendix.

2. TESTING FOR HOMOGENEITY

As a starting point we shall first provide a test for common trends in the absence of exogenous regressors x_{it}. Our motivation comes from the fact that the results of Section 3 are then much easier to follow after those given in Theorems 1 and 2 below. So, we begin with the model

$$\begin{aligned} y_{it} &= m_i(t/T) + \upsilon_{it}, \quad i = 1, ..., n, \quad t = 1, ..., T, \\ \upsilon_{it} &= \mu_i + \varepsilon_{it} \end{aligned} \tag{2.1}$$

with both n and T increasing to infinity.

The null hypothesis H_0 is

$$H_0: \; m_i(\tau) = m(\tau) \quad \text{for all } i \geq 1; \quad \tau \in [0, 1] \tag{2.2}$$

and the alternative hypothesis H_1 becomes

$$H_1: \; 0 < \iota(n) = |\Lambda| / n < 1, \tag{2.3}$$

where $\Lambda = \{i : \mu(\Upsilon_i) > 0\}$ with $\Upsilon_i = \{\tau \in [0, 1] : m_i(\tau) \neq m(\tau)\}$ and "$|\mathcal{A}|$" denotes the cardinality of the set $\mathcal{A}$ and $\lambda(\Upsilon)$ denotes the Lebesgue measure of the set Υ. Λ denotes the set of individuals $i \geq 1$ for which $m_i(\tau)$ is different from the "common" trend $m(\tau)$, and thus $\iota =: \iota(n)$ represents the proportion of sequences $i \geq 1$ for which $m_i(\tau) \neq m(\tau)$. One feature of (2.3) is that ι can be (asymptotically) negligible. More specifically, as we shall show in Section 4, the test has nontrivial power under local alternatives such that $\iota = O(n^{-1/2})$. This situation when $\iota \searrow 0$ can be of interest if we want to decide whether a new set of sequences share the same trending behaviour as the existing ones, or for "supervised classification" purposes.

Before we state our results, let's introduce some regularity conditions.

Condition C1: $\{\varepsilon_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, are sequences of zero mean independent random variables with $E\varepsilon_{it}^2 = \sigma^2$ and finite fourth moments. In addition, the sequences $\{\varepsilon_{it}\}_{t\geq 1}$ and $\{\varepsilon_{jt}\}_{t\geq 1}$ are mutually independent for all $i \neq j$.

Condition C2: The functions $m_i(\tau)$ are bounded for all $i \in \mathbb{N}^+$.

Condition $C1$ is standard, although stronger than we really need for our results to follow. An inspection of our proofs indicate that the results would follow provided some type of "weak" dependence in both time and/or cross-sectional dimensions. The only main difference, being that the asymptotic covariance structure of our statistics will be affected by its dependence. So, for simplicity and to ease arguments, we have decided to keep $C1$ as it stands. It is worth emphasizing that we do not assume that the sequences $\{\varepsilon_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$ are identically distributed, although their first two moments are constant, although we shall discuss the consequences of allowing for heteroscedasticity in Section 3 below.

Regarding Condition $C2$, we notice that we do not need any type of smoothness condition on the trend functions $m_i(\tau)$. We require only that they are bounded, although strictly speaking this is only needed when examining the consistency and local power of the tests.

Let's introduce now some notation used in the sequel of the paper. For arbitrary sequences $\{\varsigma_{it}\}_{t=1}^{T}$, $i \geq 1$, we define

$$\bar{\varsigma}_{\cdot t} = \frac{1}{n}\sum_{j=1}^{n}\varsigma_{jt}, \quad \bar{\varsigma}_{i\cdot} = \frac{1}{T}\sum_{s=1}^{T}\varsigma_{is}, \quad \bar{\varsigma}_{\cdot\cdot} = \frac{1}{Tn}\sum_{j=1}^{n}\sum_{s=1}^{T}\varsigma_{js}. \tag{2.4}$$

For presentation reasons we shall consider separately the cases $\mu_i = 0$ and $\mu_i \neq 0$. The arguments needed in the latter case are lengthier, but they are based on those given when $\mu_i = 0$.

2.1. The case without fixed individual effects μ_i.

Define

$$\widetilde{m}_i(t/T) := y_{it} - \overline{y}_{\cdot t}, \quad i = 1, \cdots n, \quad t = 1, ..., T \tag{2.5}$$

$$\widetilde{m}'_i(t/T) := y_{it} - \frac{1}{i-1}\sum_{j=1}^{i-1} y_{jt}, \quad i = 1, \cdots n, \quad t = 1, ..., T. \tag{2.6}$$

The motivation to introduce the quantities (2.5) and (2.6) is that, under the null hypothesis H_0, it holds that $\overline{y}_{\cdot t} = m(t/T) + \overline{\varepsilon}_{\cdot t}$. So,

$$\widetilde{m}_i\left(\frac{t}{T}\right) := \varepsilon_{it} - \overline{\varepsilon}_{\cdot t}; \qquad \widetilde{m}'_i\left(\frac{t}{T}\right) := \varepsilon_{it} - \frac{1}{i-1}\sum_{1=2}^{i-1}\varepsilon_{jt},$$

which do not depend on the trend function $m(\cdot)$ and they have zero mean. However, under the alternative hypothesis, $\widetilde{m}_i(t/T)$ and $\widetilde{m}'_i(t/T)$ will develop a mean different than zero, suggesting that tests based on (2.5) and (2.6) will have the usual desirable statistical properties.

Hence, following ideas in Brown, Durbin and Evans (1975), we base our test for homogeneity on

$$\begin{aligned} Q_{T,n,r,u} &= \frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\left(y_{it} - \overline{y}_{\cdot t}\right), \\ Q'_{T,n,r,u} &= \frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]}\left(y_{it} - \frac{1}{i-1}\sum_{j=1}^{i-1} y_{jt}\right), \qquad r, u \in [0,1], \end{aligned}$$

where $[c]$ denotes the integer part of a number c.

Let's introduce the following notation. We denote "$\Rightarrow_{D\left([0,1]^2\right)}$" weak convergence in the space $D\left([0,1]^2\right)$, the space of continuous functions from the right and equipped with the Skorohod metric, given in p. 1662 of Bickel and

Wichura (1971). Recall that a two-parameter process $W(r,u)$ is said to be a Brownian sheet if it is a Gaussian process, satisfying $W(0,0)=0$ and

$$Cov\left(W\left(r_1,u_1\right),W\left(r_2,u_2\right)\right)=\left(r_1\wedge r_2\right)\left(u_1\wedge u_2\right),$$

where $(a_1\wedge a_2)=\min(a_1,a_2)$. We also introduce two-parameter process

$$B(r,u)=W(r,u)-uW(r,1).$$

Theorem 1. *Assuming $C1$, under H_0, we have that as $T,n\to\infty$,*

$$(\mathbf{a})\quad \sqrt{Tn}\,Q_{T,n,r,u}\Rightarrow_{D([0,1]^2)}\sigma B(r,u),\quad r,u\in[0,1],\tag{2.7}$$

$$(\mathbf{b})\quad \sqrt{Tn}\,Q'_{T,n,r,u}\Rightarrow_{D([0,1]^2)}\sigma W(r,u),\quad r,u\in[0,1].\tag{2.8}$$

2.2. **The case with individual fixed effects μ_i.**

When the individual fixed effects μ_i's are present in (2.1), we need to modify both $Q_{T,n,r,u}$ and $Q'_{T,n,r,u}$, as we need to remove them from the model. To that end, for all $r,u\in[0,1]$, consider

$$U_{T,n,r,u}=\frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\left\{y_{it}-\overline{y}_{\cdot t}-\overline{y}_{i\cdot}+\overline{y}_{\cdot\cdot}\right\}$$

$$U'_{T,n,r,u}=\frac{1}{Tn}\sum_{t=2}^{[Tr]}\sum_{i=2}^{[nu]}\left\{y_{it}-\frac{1}{i-1}\sum_{j=1}^{i-1}y_{jt}-\frac{1}{t-1}\sum_{s=1}^{t-1}y_{is}+\frac{1}{(t-1)(i-1)}\sum_{j=1}^{i-1}\sum_{s=1}^{t-1}y_{js}\right\}.$$

Let's introduce the two-parameter Gaussian process

$$B'(r,u)=W(r,u)-uW(r,1)-rW(1,u)+ruW(1,1),$$

which satisfies $B'(0,0)=0, B'(r,1)=0, B'(1,u)=0\quad\forall r,u\in[0,1]$ and

$$Cov\left(B'(r_1,u_1),B'(r_2,u_2)\right)=(r\wedge r_1)(1-r_1\vee r_2)(u_1\wedge u_2)(1-u_1\vee u_2),$$

where $a_1\vee a_2=\max(a_1,a_2)$. The process $B'(r,u)$ is known as the *Brownian pillow.*

Theorem 2. *Assuming $C1$, under H_0, we have that as $T,n\to\infty$*

$$(\mathbf{a})\quad \sqrt{Tn}U_{T,n,r,u}\Rightarrow_{D([0,1]^2)}\sigma B'(r,u),\quad r,u\in[0,1],\tag{2.9}$$

$$(\mathbf{b})\quad \sqrt{Tn}U'_{T,n,r,u}\Rightarrow_{D([0,1]^2)}\sigma W(r,u),\quad r,u\in[0,1].\tag{2.10}$$

We now comment briefly on the results of Theorems 1 and 2. The first point to notice is that, as we mentioned in the abstract, we are able to avoid local smoothing in the estimation of the functions $m_i(\cdot)$, being the key that both n and T increase to infinity. We notice that both $\sqrt{Tn}U'_{T,n,r,u}$ and $\sqrt{Tn}\,Q'_{T,n,r,u}$ have the same asymptotic behaviour, which is not the case for $\sqrt{Tn}U_{T,n,r,u}$ and $\sqrt{Tn}\,Q_{Tn,r,u}$. This is somehow expected after we notice that both $y_{it}-\frac{1}{i-1}\sum_{j=1}^{i-1}y_{jt}$ and the term inside the braces in the definition of $U'_{T,n,r,u}$ have the same structure, more specifically they behave as martingale difference sequences. We just remark that $y_{it}-\frac{1}{i-1}\sum_{j=1}^{i-1}y_{jt}$ is nothing but the martingale transformation of $y_{it}-\overline{y}_{\cdot t}$.

2.3. Tests for homogeneity.

We put together results of Sections 2.1-2.2 to provide test statistics for homogeneity. To that end we denote by $g(\cdot)$ a continuous functional: $[0,1]^2 \to \mathbb{R}^+$. To implement the test(s) we need to estimate the variance σ^2. For that purpose, denote residuals by

$$e_{it}^{(1)} := y_{it} - \bar{y}_{\cdot t}; \quad e_{it}^{(2)} := y_{it} - \bar{y}_{\cdot t} - \bar{y}_{i\cdot} + \bar{y}_{\cdot\cdot},$$

depending on whether the model does not have or has individual fixed effects. As we mentioned earlier, under the H_0, we have that

$$e_{it}^{(1)} := \varepsilon_{it} - \bar{\varepsilon}_{\cdot t}; \quad e_{it}^{(2)} := \varepsilon_{it} - \bar{\varepsilon}_{\cdot t} - \bar{\varepsilon}_{i\cdot} + \bar{\varepsilon}_{\cdot\cdot}.$$

In addition we estimate σ^2 by

$$\widehat{\sigma}_{(\ell)}^2 := \frac{1}{Tn} \sum_{t=1}^{T} \sum_{i=1}^{n} e_{it}^{(\ell)2}, \quad \text{for } \ell = 1, 2.$$

Proposition 1. *Assuming $C1$, under H_0, as $n, T \to \infty$, we have that*

$$\widehat{\sigma}_{(\ell)}^2 \to_p \sigma^2, \quad \textit{for} \quad \ell = 1, 2.$$

Proof. The proof is standard and so it is omitted. □

Corollary 1. *Assuming $C1 - C2$, under H_0, for any continuous functional $g(\cdot)$, as $T, n \to \infty$,*

(**a**)

$$g\left(\sqrt{Tn}\frac{Q_{T,n,r,u}}{\widehat{\sigma}_{(1)}}\right) \overset{d}{\Rightarrow} g(B(r,u)); \quad g\left(\sqrt{Tn}\frac{Q'_{T,n,r,u}}{\widehat{\sigma}_{(1)}}\right) \overset{d}{\Rightarrow} g(W(r,u))$$

(**b**)

$$g\left(\sqrt{Tn}\frac{U_{T,n,r,u}}{\widehat{\sigma}_{(2)}}\right) \overset{d}{\Rightarrow} g\left(B'(r,u)\right); \quad g\left(\sqrt{Tn}\frac{U'_{T,n,r,u}}{\widehat{\sigma}_{(2)}}\right) \overset{d}{\Rightarrow} g(W(r,u)).$$

Proof. The proof follows by standard arguments using the continuous mapping theorem, Theorems 1 and 2 and Proposition 1, so it is omitted. □

One functional often employed to implement the test is the "sup", that is

$$\sqrt{Tn} \sup_{r,u\in[0,1]} \left|\frac{U_{T,n,r,u}}{\widehat{\sigma}_{(2)}}\right| \overset{d}{\Rightarrow} \sup_{r,u\in[0,1]} |B'(r,u)|,$$

$$\sqrt{Tn} \sup_{r,u\in[0,1]} \left|\frac{U'_{T,n,r,u}}{\widehat{\sigma}_{(2)}}\right| \overset{d}{\Rightarrow} \sup_{r,u\in[0,1]} |W(r,u)|,$$

which corresponds to a Kolmogorov-Smirnov type of test.

Critical values for the supremum of absolute value of the three limiting processes are provided in Table 1 for various significance levels, which have been simulated with 20000 iterations.

TABLE 1. Critical Values for α-significance level, $n, T = 10000$

α	0.1	0.075	0.05	0.025	0.01	0.005
$\sup_{r,u\in[0,1]} \lvert W(r,u)\rvert$	2.2187	2.3308	2.4781	2.7227	3.0282	3.2
$\sup_{r,u\in[0,1]} \lvert B(r,u)\rvert$	1.3419	1.3951	1.4677	1.5782	1.7106	1.8091
$\sup_{r,u\in[0,1]} \lvert B'(r,u)\rvert$	0.7777	0.8039	0.8372	0.8849	0.9515	0.9975

3. TEST OF COMMON TRENDS

We now revisit our model of interest (1.1). Recall that under the null hypothesis we can write (1.1) as

$$y_{it} = \beta' x_{it} + m\left(t/T\right) + \upsilon_{it}, \quad \upsilon_{it} = \mu_i + \varepsilon_{it}. \tag{3.1}$$

It is obvious that if β were known, the problem of testing for homogeneity would be no different from that examined in Section 2. That is, we would test homogeneity in the model

$$\begin{aligned} y_{it} - x_{it}'\beta &= :\upsilon_{it} \\ &= m_i(t/T) + \mu_i + \varepsilon_{it} \end{aligned}$$

replacing y_{it} there by υ_{it}. However β is unknown, so given an estimator $\widehat{\beta}$, we might then consider instead

$$\widehat{\upsilon}_{it} =: y_{it} - x_{it}'\widehat{\beta} = x_{it}'\left(\widehat{\beta} - \beta\right) + m(t/T) + \mu_i + \varepsilon_{it}.$$

Thus, it is sensible to consider $U_{Tn,r,u}$ and $U'_{Tn,r,u}$ but with $\widehat{\upsilon}_{it}$ replacing y_{it} there. More specifically, for $r, u \in [0,1]$, we define

$$\begin{aligned} \widehat{U}_{T,n,r,u} &= \frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\left(\widehat{\upsilon}_{it} - \overline{\widehat{\upsilon}}_{\cdot t} - \overline{\widehat{\upsilon}}_{i\cdot} + \overline{\widehat{\upsilon}}_{\cdot\cdot}\right) \qquad (3.2) \\ \widehat{U}'_{T,n,r,u} &= \frac{1}{Tn}\sum_{t=2}^{[Tr]}\sum_{i=2}^{[nu]}\left(\widehat{\upsilon}_{it} - \frac{\sum_{j=1}^{i-1}\widehat{\upsilon}_{js}}{i-1} - \frac{\sum_{s=1}^{t-1}\widehat{\upsilon}_{is}}{t-1} + \frac{\sum_{j=1}^{i-1}\sum_{s=1}^{t-1}\widehat{\upsilon}_{js}}{(i-1)(t-1)}\right). \end{aligned}$$

So, we now look at how to estimate the slope parameters β in (3.1). For that purpose, given a generic sequence $\{\varsigma_{it}\}_{i\geq 1, t\geq 1}$ and denoting by $\check{\varsigma}_{it} := \varsigma_{it} - \bar{\varsigma}_{\cdot t} - \bar{\varsigma}_{i\cdot} + \bar{\varsigma}_{\cdot\cdot}$, we have that under H_0, (3.1) becomes

$$\check{y}_{it} = \beta'\check{x}_{it} + \check{\varepsilon}_{it}; \quad i = 1, ..., n; \quad t = 1, ..., T,$$

so that we estimate β by least squares methods, that is

$$\widehat{\beta} = \left(\sum_{t=1}^{T}\sum_{i=1}^{n}\check{x}_{it}\check{x}_{it}'\right)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{n}\check{x}_{it}\check{y}_{it}.$$

Observe that $\check{\varsigma}_{it}$ removes the common trend and the fixed effects.

Condition C3: For all $i \in \mathbb{N}^+$, $\{x_{it}\}_{t\in\mathbb{Z}}$ are k-dimensional sequences of independent random variables mutually uncorrelated with the error sequences $\{\varepsilon_{it}\}_{t\in\mathbb{Z}}$, where $E(x_{it}) = \varkappa_i$, $\Sigma = E\left\{(x_{it} - \varkappa_i)(x_{it} - \varkappa_i)'\right\}$ and finite fourth moments.

Condition C4: $T, n \to \infty$ such that $T/n^2 \to 0$.

We have the following two results.

Theorem 3. *Assuming $C1 - C4$, under H_0, as $T, n \to \infty$,*

$$(Tn)^{1/2}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\Sigma^{-1}\right).$$

Theorem 4. *Assuming $C1$ to $C4$, under H_0, as $T, n \to \infty$ we have that for any continuous functional $g(\cdot)$,*

$$(\mathbf{a}) \quad g\left(\sqrt{Tn}\frac{\widehat{U}_{T,n,r,u}}{\widehat{\sigma}}\right) \overset{d}{\Rightarrow} g\left(B'(r,u)\right); \qquad (\mathbf{b}) \quad g\left(\sqrt{Tn}\frac{\widehat{U}'_{T,n,r,u}}{\widehat{\sigma}}\right) \overset{d}{\Rightarrow} g\left(W(r,u)\right),$$

where

$$\widehat{\sigma}^2 := \frac{1}{Tn}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\widehat{\upsilon}_{it} - \overline{\widehat{\upsilon}}_{\cdot t} - \overline{\widehat{\upsilon}}_{i\cdot} + \overline{\widehat{\upsilon}}_{\cdot\cdot}\right)^2.$$

A typical continuous functional $g(\cdot)$ employed is the supremum, that is

$$(Tn)^{1/2}\sup_{r,u\in[0,1]}\left|\widehat{\sigma}^{-1}\widehat{U}'_{T,n,r,u}\right| \xrightarrow{d} \sup_{r,u\in[0,1]}\left|W(r,u)\right|,$$

which corresponds to a Kolmogorov-Smirnov type of statistic.

Remark 1. *The results of the previous theorem indicates that knowledge of β does not affect the asymptotic behaviour of our test.*

3.1. **Unequal variance σ_i^2.**

We now look at the problem of testing for homogeneity with possible heterogeneous second moments in the error term. More specifically, we are concerned with testing for homogeneity when

Condition C5: For all $i \in \mathbb{N}^+$, $\{\varepsilon_{it}\}_{t\in\mathbb{Z}}$ are sequences of random variables as in $C1$ but such that $E\varepsilon_{it}^2 = \sigma_i^2$ satisfying

$$\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 \le C < \infty.$$

We can be tempted to write Condition $C5$ in a more traditional format as saying that $E\left(\varepsilon_{it}^2 \mid x_{it}\right) = \sigma^2(x_{it})$. However, Condition $C3$ implies that the latter conditional variance would not play any relevant role on the asymptotic behaviour of any continuous functional of either $\widehat{U}_{T,n,r,u}$ or $\widehat{U}'_{T,n,r,u}$, in the sense that the asymptotic behaviour is that given in Theorem 4 but with a different scaling σ. On the other hand, this is not the case though, if the unconditional variance of $\{\varepsilon_{it}\}_{t\in\mathbb{Z}}$ is as stated in Condition $C5$ as Theorem 5 below shows. It is also worth mentioning that it is obvious how to handle the situation where we allow for trending variances, say $\sigma_i^2 = \alpha_1 + \alpha_2 i^\gamma$, as long as $\gamma < 1$. This relaxation of $C5$ is trivial and it only complicates the notation and arguments. Finally, similar caveats appear if $C5$ we modified to say that $E\varepsilon_{it}^2 = \sigma_t^2$. However, to save space we decided to keep $C5$ as it stands.

Let's introduce the following notation

$$\sigma^2(u) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{[nu]} \sigma_i^2 .$$

Theorem 5. *Assuming $C3$ and $C5$, under H_0 we have that as $n, T \to \infty$,*

$$\begin{array}{lrcl} (\mathbf{a}) & (Tn)^{1/2}\,\widehat{U}_{T,n,r,u} & \Longrightarrow & \mathcal{G}(r,u) \\ (\mathbf{b}) & (Tn)^{1/2}\,\widehat{U}'_{T,n,r,u} & \Longrightarrow & \mathcal{G}'(r,u), \end{array}$$

where $\mathcal{G}'(r,u)$ and $\mathcal{G}(r,u)$ are Gaussian processes with covariance structures $\Phi(r_1,u_1;r_2,u_2)$ given in (7.16) below and

$$\begin{aligned} Cov(\mathcal{G}(r_1,u_1),\mathcal{G}(r_2,u_2)) \;=\; & (r_1\wedge r_2)\left(\sigma^2(u_1\wedge u_2) - u_2\sigma^2(u_1)\right) \\ & - r_1r_2u_2\sigma^2(u_1) - r_1r_2u_2\sigma^2(u_1\wedge u_2) \\ & + u_1\sigma^2(u_2)(r_1r_2 - (r_1\wedge r_2)), \end{aligned}$$

respectively.

The results of Theorem 5 indicate that to obtain asymptotic critical values on functionals based on $(Tn)^{1/2}\,\widehat{U}'_{T,n,r,u}$ and $(Tn)^{1/2}\,\widehat{U}_{T,n,r,u}$ appear to be a daunting task. This is in view of the very complicated and complex covariance structure of the latter two processes. To overcome this type of problem, we envisage a procedure based on a transformation of $\widehat{U}'_{T,n,r,u}$ or $\widehat{U}_{T,n,r,u}$ in such a way that the transformed statistics will have a known (or at least simpler) covariance structure.

Denote the estimator of σ_i^2 by

$$\widehat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{v}_{it} - \overline{\widehat{v}}_{\cdot t} - \overline{\widehat{v}}_{i\cdot} - \overline{\widehat{v}}_{\cdot\cdot}\right)^2, \;\; i = 1, ..., n. \tag{3.3}$$

It is quite standard in view of $C3$ and $C5$ to show that $\widehat{\sigma}_i^2$ is a (pointwise) consistent estimator for σ_i^2 for all $i = 1, ..., n$, and under $C3$ and $C5$ and under sufficient moment conditions, it is easily shown that

$$E\left(\widehat{\sigma}_i^2 - \sigma_i^2\right)^{2p} = O\left(T^{-p}\right), \;\; i = 1, ..., n. \tag{3.4}$$

So, for instance we modify $\widehat{U}'_{T,n,r,u}$ to

$$\widehat{U}^{\sharp}_{T,n,r,u} = \frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]}\left(\frac{1}{\widehat{\sigma}_i}\widehat{v}_{it} - \frac{\sum_{j=1}^{i-1}\frac{1}{\widehat{\sigma}_j}\widehat{v}_{js}}{i-1} - \frac{\sum_{s=1}^{t-1}\frac{1}{\widehat{\sigma}_i}\widehat{v}_{is}}{t-1} + \frac{\sum_{j=1}^{i-1}\sum_{s=1}^{t-1}\frac{1}{\widehat{\sigma}_j}\widehat{v}_{js}}{(i-1)(t-1)}\right). \tag{3.5}$$

A similar transformation is obtained for $\widehat{U}_{T,n,r,u}$. Now, Theorem 3 implies that the right side of (3.5) is

$$\frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]}\left(\frac{1}{\widehat{\sigma}_i}\varepsilon_{it}-\frac{\sum_{j=1}^{i-1}\frac{1}{\widehat{\sigma}_j}\varepsilon_{js}}{i-1}-\frac{\sum_{s=1}^{t-1}\frac{1}{\widehat{\sigma}_i}\varepsilon_{is}}{t-1}+\frac{\sum_{j=1}^{i-1}\sum_{s=1}^{t-1}\frac{1}{\widehat{\sigma}_j}\varepsilon_{js}}{(i-1)(t-1)}\right)$$
$$+o_p\left((Tn)^{-1/2}\right)$$

uniformly in $r,u\in[0,1]$. Now, if we would have σ_i instead of $\widehat{\sigma}_i$, the right hand side of the last displayed equation would be then what we had in Theorems 2 or 4, but with σ^2 replaced by 1 there, as we had that $\varepsilon_{it}/\sigma_i$ satisfies $C1$ but with unit variance. So, if

$$\frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]}\left(\frac{1}{\widehat{\sigma}_i}-\frac{1}{\sigma_i}\right)\varepsilon_{it}=o_p\left((Tn)^{-1/2}\right) \tag{3.6}$$

uniformly in $r,u\in[0,1]$, we then conclude that the modification given by $\widehat{U}^{\sharp}_{Tn,r,u}$ leads to a statistic with a known covariance structure.

To that end, we first observe that by (3.4) and that $E\sup_{\ell=1,\dots,q}|a_\ell|\le\left(\sum_{\ell=1}^{q}E|a_\ell|^k\right)^{1/k}$, we have that

$$\sup_{i=1,\dots,n}\left|\widehat{\sigma}_i^2-\sigma_i^2\right|=O_p\left(n^{1/2p}T^{-1/2}\right)=o(1).$$

The latter uniform convergence, as opposed to pointwise consistency. But (3.6) is the case, because by Taylor's expansion,

$$\frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]}\left\{\sum_{\ell=1}^{2p}(-1)^\ell\left(\frac{\widehat{\sigma}_i-\sigma_i}{\sigma_i}\right)\right\}\varepsilon_{it}/\sigma_i$$
$$+\frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]}\left(\frac{\widehat{\sigma}_i-\sigma_i}{\sigma_i}\right)^{2p+1}\frac{1}{\widehat{\sigma}_i}\varepsilon_{it}. \tag{3.7}$$

From here, suitable regularity conditions on how n and T increase to infinity and somewhat lengthy but routine algebra lead us to the claim in (3.6) and conclude that the transformation of $\widehat{U}'_{Tn,\tau,\iota}$ given in $\widehat{U}^{\sharp}_{Tn,r,u}$ will overcome the problematic issue of how to tabulate valid asymptotic critical values for our test.

4. ASYMPTOTIC PROPERTIES UNDER LOCAL ALTERNATIVE HYPOTHESIS. BOOTSTRAP

As usual, it is important to examine the behaviour of any test under alternative hypothesis and ultimately under local alternatives. We begin with the discussion under local alternatives, the consistency of the tests being a corollary. To that end, consider

$$H_l: \quad \begin{cases} m_i(\tau) = m(\tau) + \frac{1}{T^{1/\xi}} \varrho_i(\tau) & \text{for all } i \leq \mathbf{I}(n), \\ m_i(\tau) = m(\tau) & \text{for all } \mathbf{I}(n) < i \leq n, \end{cases} \tag{4.1}$$

where $\varrho_i(\tau)$ is different than zero in $\Upsilon_i = \Upsilon$, a set of positive Lebesgue measure, $\xi \leq 1/2$ and $\mathbf{I}(n) \leq Cn^{1/2}$, where C is a finite positive constant. It is worth remarking that we have assumed that Υ is the same for all $i \geq 1$ for notational simplicity. However in the Monte Carlo simulations we shall examine what happens with the power of the test when, given a particular alternative hypothesis, we permute the order of the cross-section units, see Table 8.

We begin with the test based on the statistic $U_{T,n,r,u}$ leaving the case with regressors x_{it} for later.

Proposition 2. *Under H_l with $\xi = 1/2$ and $\mathbf{I}(n) = Cn^{1/2}$, assuming $C1$ and $C2$, we have that*

$$(Tn)^{1/2} U_{T,n,r,u} \Rightarrow_{D([0,1]^2)} B'(r,u) + C^{1/2} \left(\int_{[0,r]\cap\Upsilon} \overline{\varrho}(\tau)\, d\tau - r \int_{\Upsilon} \overline{\varrho}(\tau)\, d\tau \right),$$

where $\overline{\varrho}(\tau) = \lim_{n\to\infty} n^{-1/2} \sum_{i=1}^{n^{1/2}} \varrho_i(\tau)$.

Proof. By definition we have that

$$\begin{aligned} U_{T,n,r,u} &= \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} (\varepsilon_{it} - \overline{\varepsilon}_{\cdot t} - \overline{\varepsilon}_{i\cdot} + \overline{\varepsilon}_{\cdot\cdot}) \\ &\quad + \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{Cn^{1/2}} \left[\mathring{\varrho}_i(t/T) - \frac{1}{n} \sum_{j=1}^{Cn^{1/2}} \mathring{\varrho}_j(t/T) \right]. \end{aligned}$$

The second term on the right of the last displayed equality known as the drift function, is

$$\frac{1}{T^{3/2} n} \sum_{t=1}^{[Tr]} \sum_{i=1}^{Cn^{1/2}} \left(1 - \frac{1}{n}\right) \mathring{\varrho}_i(t/T).$$

So, we have that after we normalize $U_{T,n,r,u}$ by $(Tn)^{1/2}$ the latter displayed expression becomes

$$\frac{1}{T} \sum_{t=1}^{[Tr]} \frac{1}{n^{1/2}} \sum_{i=1}^{Cn^{1/2}} \mathring{\varrho}_i(t/T) \to C^{1/2} \left(\int_{[0,r]\cap\Upsilon} \overline{\mathring{\varrho}}(\tau)\, d\tau - r \int_{\Upsilon} \overline{\mathring{\varrho}}(\tau)\, d\tau \right).$$

Observe that we can interchange the limit in n and the integral using Lebesgue dominated convergence because $|\varrho_i(u)| \leq C$ for all $i \geq 1$. Recall that $C2$ implies that the functions $\varrho_i(\tau)$ are bounded. This concludes the proof of the proposition. □

One immediate conclusion that we draw from Proposition 2 is that our tests detect local alternatives shrinking to the null hypothesis at a parametric rate $O\left(T^{-1/2}n^{-1/2}\right)$. A question of interest can be if test based on $U'_{Tn,r,u}$ can detect local alternatives when $\xi < 1/2$ and/or $\mathbf{I}$ is a finite integer. This type of alternatives can be useful for the purpose of, say, classification or whether we wish to decide if a new set of individuals, $i \leq \mathbf{I}$, shares the "common" trend function $m(\tau)$.

To that end, we notice that the drift function in $(Tn)^{1/2} U_{Tn,r,u}$ becomes

$$\begin{aligned} \frac{1}{T^{\xi+1/2}n^{1/2}} \sum_{t=1}^{[Tr]} \sum_{i=1}^{\mathbf{I}} \left(1 - \frac{1}{n}\right) \mathring{\varrho}_i(t/T) &\simeq \frac{1}{T^{\xi+1/2}n^{1/2}} \sum_{t=1}^{[T\tau]} \sum_{i=1}^{\mathbf{I}} \mathring{\varrho}_i(t/T) \\ &\simeq \frac{T^{1/2-\xi}}{n^{1/2}} \int_{[0,r]\cap\Upsilon} \sum_{i=1}^{\mathbf{I}} \mathring{\varrho}_i(u)\, du. \end{aligned}$$

So, if $n \propto T^{1-2\xi}$, the drift function is nonzero and hence the test based on $U_{Tn,r,u}$, and similarly for $U'_{T,n,r,u}$, have no trivial asymptotic power.

Another easy conclusion from Proposition 2 is that under fixed alternatives, since $\int_{[0,r]\cap\Upsilon} \overline{\mathring{\varrho}}(\tau)\, d\tau - r \int_{\Upsilon} \overline{\mathring{\varrho}}(\tau)\, d\tau$ is nonzero, $(Tn)^{-1/2} \left\| U_{T,n,r,u} \right\|^{-1} \to_P 0$, so that tests based on $U_{T,n,r,u}$ are consistent, so is $U'_{T,n,r,u}$.

We now discuss the local power for the model (3.1). We begin by discussing the consequences on the asymptotic properties of the least squares estimator of β. To that end, recall that (3.1) under local alternatives with $\xi = 1/2$ and $\mathbf{I}(n) = Cn^{1/2}$ becomes

$$y_{it} = \beta' x_{it} + m_i(t/T) + \upsilon_{it}, \quad \upsilon_{it} = \mu_i + \varepsilon_{it}$$

so denoting

$$\mathring{m}_i(t/T) = m_i(t/T) - T^{-1} \sum_{s=1}^{T} m_i(s/T),$$

we have that standard manipulations yields

$$\begin{aligned} \widehat{\beta} - \beta &= \left(\sum_{t=1}^{T} \sum_{i=1}^{n} \check{x}_{it} \check{x}'_{it} \right)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} \check{x}_{it} \left\{ \mathring{m}_i(t/T) - \frac{1}{n} \sum_{j=1}^{n} \mathring{m}_j(t/T) \right\} \\ &\quad + \left(\sum_{t=1}^{T} \sum_{i=1}^{n} \check{x}_{it} \check{x}'_{it} \right)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} \check{x}_{it} \varepsilon_{it}. \end{aligned} \tag{4.2}$$

But as we can take $Ex_{it} = 0$, the only consequence on the asymptotic behaviour of $(Tn)^{1/2}\left(\widehat{\beta} - \beta\right)$ is that the asymptotic variance will have an additional term coming from

$$\frac{1}{(Tn)^{1/2}} \sum_{i=1}^{n} \sum_{t=1}^{T} \check{x}_{it} \left\{ \mathring{m}_i(t/T) - \frac{1}{n} \sum_{j=1}^{n} \mathring{m}_j(t/T) \right\},$$

which is different than zero under fixed alternative hypothesis, whereas under the local alternatives it is clear to be $o_p(1)$ implying that the behaviour of $(Tn)^{1/2}\left(\widehat{\beta}-\beta\right)$ is unchanged. The only condition for the latter is that $\varrho_i(\tau)$ is square integrable. See our comments after Corollary 2 below.

We now discuss the behaviour of $\widehat{U}_{T,n,r,u}$ under the local alternatives in (4.1).

Corollary 2. *Under H_l with $\xi=1/2$ and $\mathbf{I}(n)=Cn^{1/2}$ and assuming $C1-C4$, we have that*

$$(Tn)^{1/2}\widehat{U}_{T,n,r,u}\Rightarrow_{D\left([0,1]^2\right)} B'(r,u)+C^{1/2}\left(\int_{[0,r]\cap\Upsilon}\overline{\varrho}(\tau)\,d\tau-r\int_{\Upsilon}\overline{\varrho}(\tau)\,d\tau\right).$$

Proof. The proof follows easily in view of (4.2), comments that follow and Proposition 2, so it is omitted. □

Corollary 2 then indicates that, as it was the case under the null hypothesis, the behaviour of the test is unaltered by the estimation of the slope parameters β under the local alternatives.

Finally, we need to discuss the consistency of the test(s). But this is a natural consequence of Proposition 2 and Corollary 2. The only difference, but this does not affect the consistency of the test, is that the distribution of $\widehat{\beta}$ is different as it will have an additional term in its asymptotic variance coming from the asymptotic variance of first term on the right of (4.2), which by standard arguments yield to the expression

$$\begin{aligned}&\Sigma^{-1}Var\left(\frac{1}{(Tn)^{1/2}}\sum_{t=1}^{T}\sum_{i=1}^{n}x_{it}\left\{\mathring{m}_i(t/T)-\frac{1}{n}\sum_{j=1}^{n}\mathring{m}_j(t/T)\right\}\right)\Sigma^{-1}\\ =\;&\Sigma^{-1}\lim_{n,T\to\infty}\frac{1}{(Tn)^{1/2}}\sum_{t=1}^{T}\sum_{i=1}^{n}\left\{\mathring{m}_i(t/T)-\frac{1}{n}\sum_{j=1}^{n}\mathring{m}_j(t/T)\right\}^2.\end{aligned}$$

Again regarding the tests based on $(Tn)^{1/2}\widehat{U}'_{T,n,r,u}$, we obtain the same type of conclusion proceeding as with $(Tn)^{1/2}\widehat{U}_{T,n,r,u}$.

4.1. **Bootstrap of the test.**

In this subsection, we provide a valid bootstrap algorithm for our tests. In particular our main concern is to do so for the tests based on $\widehat{U}_{T,n,r,u}$ (and $Q_{T,n,r,u}$) as our Monte-Carlo experiment suggests a poor finite sample performance, especially when compared to those of $\widehat{U}'_{T,n,r,u}$ (and $Q'_{T,n,r,u}$). To that end, because by $C1$ the errors ε_{it} are i.i.d., Efron's naive bootstrap would be appropriate. The bootstrap will have the following *3 STEPS*:

***STEP 1*:** For each $i=1,...,n$, denote by $\{\varepsilon^*_{it}\}_{t=1}^{T}$ a random sample of size T from the empirical distribution of $\{\widehat{\varepsilon}_{it}\}_{t\geq1,i\geq1}$, where

$$\widehat{\varepsilon}_{it}=\widehat{\upsilon}_{it}-\overline{\widehat{\upsilon}}_{i\cdot}-\overline{\widehat{\upsilon}}_{\cdot t}+\overline{\widehat{\upsilon}}_{\cdot\cdot}.$$

Observe that by definition $\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{\varepsilon}_{it}=0$.

***STEP 2*:** We compute our bootstrap panel model as

$$y_{it}^{*}=x_{it}'\widehat{\beta}+\varepsilon_{it}^{*},\ i=1,...,n;\quad t=1,...,T,$$

and the bootstrap least squares estimator as

$$\widehat{\beta}^{*}=\left(\sum_{t=1}^{T}\sum_{i=1}^{n}\check{x}_{it}\check{x}_{it}'\right)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{n}\check{x}_{it}\check{y}_{it}^{*}.$$

***STEP 3*:** Obtain the least squares residuals

$$\widehat{v}_{it}^{*}=\check{y}_{it}^{*}-\check{x}_{it}'\widehat{\beta}^{*},\ \ i=1,...,n;\quad t=1,...,T$$

and then compute the bootstrap analogues of $\widehat{U}_{T,n,r,u}$ and $\widehat{U}'_{T,n,r,u}$ with $\widehat{v}_{it}^{*}$ replacing $\widehat{v}_{it}$ there. That is, for $r,u\in[0,1]$, we define

$$\begin{aligned}\widehat{U}_{T,n,r,u}^{*}&=\frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\left(\widehat{v}_{it}^{*}-\overline{\widehat{v}^{*}}_{\cdot t}-\overline{\widehat{v}^{*}}_{i\cdot}-\overline{\widehat{v}^{*}}_{\cdot\cdot}\right)\\ \widehat{U}_{T,n,r,u}^{*\prime}&=\frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\left(\widehat{v}_{it}^{*}-\frac{\sum_{j=1}^{i-1}\widehat{v}_{js}^{*}}{i-1}-\frac{\sum_{s=1}^{t-1}\widehat{v}_{is}^{*}}{t-1}+\frac{\sum_{j=1}^{i-1}\sum_{s=1}^{t-1}\widehat{v}_{js}^{*}}{(i-1)(t-1)}\right).\end{aligned}$$

Theorem 6. *Assuming $C1$ to $C4$, under the maintained hypothesis, as $T,n\to\infty$,*

$$(Tn)^{1/2}\left(\widehat{\beta}^{*}-\widehat{\beta}\right)\xrightarrow{d}\mathcal{N}\left(0,\sigma^{2}\Sigma^{-1}\right),\ \ \textit{(in probability)}.$$

We then have the following result.

Theorem 7. *Assuming $C1$ to $C4$, under the maintained hypothesis as $T,n\to\infty$ we have that for any continuous functional $g(\cdot)$,*

$$\begin{aligned}&(\mathbf{a})\quad g\left(\widehat{\sigma}^{*-1}\sqrt{Tn}\widehat{U}_{T,n,r,u}^{*}\right)\overset{d}{\Rightarrow}g\left(B'(r,u)\right),\quad \textit{(in probability)}\\ &(\mathbf{b})\quad g\left(\widehat{\sigma}^{*-1}\sqrt{Tn}\widehat{U}_{T,n,r,u}^{*\prime}\right)\overset{d}{\Rightarrow}g\left(W(r,u)\right)\quad \textit{(in probability)},\end{aligned}$$

where

$$\widehat{\sigma}^{*2}:=\frac{1}{Tn}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\widehat{v}_{it}^{*}-\overline{\widehat{v}^{*}}_{\cdot t}-\overline{\widehat{v}^{*}}_{i\cdot}-\overline{\widehat{v}^{*}}_{\cdot\cdot}\right)^{2}.$$

5. MONTE CARLO STUDY

We now present results of a small Monte-Carlo study to shed some light on the finite sample performance of our tests based on the statistics $Q_{T,n,r,u}$, $Q'_{T,n,r,u}$, $U_{T,n,r,u}$ and $U'_{T,n,r,u}$. All throughout the study, the results reported are based on 1000 iterations. The sample sizes used were $T,n=20,30,50$.

We will only present the results for model (1.1), as those for (2.1) are qualitatively the same. So, we consider (1.1) given by

$$\begin{aligned} y_{it} &= \beta' x_{it} + m_i(t/T) + \upsilon_{it}, \quad i = 1, ..., n, \quad t = 1, ..., T, \\ \upsilon_{it} &= \mu_i + \varepsilon_{it}, \end{aligned}$$

where $\mu_i, \sim N(0,1), i = 1, \cdots, n-1$ with $\mu_n = -\sum_{j=1}^{n-1} \mu_j$ and $m(t/T) = 2\log(1+(t/T)^2)$. Recall that under H_0 $m_i(t/T) = m(t/T)$. For ε_{it}, we consider *i.i.d.* standard normal random variable as is the case for the regressors x_{it} and $\beta = 2$. In all the tables $\widehat{Q}$ and $\widehat{Q}'$ stands the test in Section 3 with $\mu_i = 0$ and the sup norm. Similarly, $\widehat{U}$ and $\widehat{U}'$ are those with μ_i as defined above.

Table 2 reports the Monte Carlo size at the 10%, 5% and 1% nominal size, whereas Tables 3 to 8 reports the power of the different tests.

TABLE 2. Monte Carlo Size of 10%, 5% and 1% tests.

nT	test stat	Asymptotic test			Bootstrap test		
400	$\widehat{Q}$	0.026	0.01	0.001	0.118	0.051	0.003
	$\widehat{Q}'$	0.1	0.049	0.009	0.129	0.062	0.01
	$\widehat{U}$	0.028	0.016	0.006	0.155	0.092	0.025
	$\widehat{U}'$	0.109	0.067	0.02	0.132	0.088	0.018
900	$\widehat{Q}$	0.032	0.014	0.002	0.071	0.031	0.006
	$\widehat{Q}'$	0.067	0.033	0.003	0.075	0.033	0.002
	$\widehat{U}$	0.025	0.013	0.004	0.113	0.052	0.007
	$\widehat{U}'$	0.092	0.044	0.012	0.116	0.048	0.012
2500	$\widehat{Q}$	0.069	0.03	0.005	0.116	0.062	0.006
	$\widehat{Q}'$	0.087	0.043	0.005	0.099	0.043	0.006
	$\widehat{U}$	0.049	0.025	0.009	0.111	0.07	0.024
	$\widehat{U}'$	0.101	0.051	0.012	0.118	0.06	0.009

Table 2 suggests that the sizes of the test based on the asymptotic distribution are pretty good even for small samples for those tests based on the statistics $\widehat{Q}'$ and $\widehat{U}'$. However, the finite sample performance of the tests based on Q and U are very poor, in particular when we compare it with the performance obtained with $\widehat{Q}'$ and $\widehat{U}'$. Due to its poor small sample performance, we decided to employ bootstrap methods to see if we can improve it. The table suggests that bootstrap test definitely improves the finite sample performance, being now tests based on, say, $\widehat{Q}$ or $\widehat{Q}'$ very similar on their behaviour.

Our next experiment deals with the power of the tests. We have considered four different settings of heterogeneity in the trend function m_i.

Setting A (single break point). Let b denote the break point, for which values $0.5n$ and $0.75n$ were tried. For $i = 1, \cdots, b$, we set $m(t/T) = 0$ for $t = 1, \cdots, T/2$ and $m(t/T) = -1$ for $t = 1+T/2, \cdots, T$. For $i = b+1, \cdots, n$, we set $m(t/T) = 2\log(1+(t/T)^2)$, which ranges from 0 to around 1.4. Tables 3-4 report the Monte Carlo power of the 10%, 5% and 1% tests under this setting.

TABLE 3. Monte Carlo Power of 10%, 5% and 1% tests, setting A, $b = 0.5n$

nT	test stat	Asymptotic test			Bootstrap test		
400	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	0.996	0.988	0.963	0.998	0.988	0.96
900	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	1	1	1	1	1	0.999
2500	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	1	1	1	1	1	1

TABLE 4. Monte Carlo Power of 10%, 5% and 1% tests, setting A, $b = 0.75n$

nT	test stat	Asymptotic test			Bootstrap test		
400	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	0.982	0.962	0.872	0.99	0.975	0.904
	$\widehat{U}$	1	1	0.999	1	1	1
	$\widehat{U}'$	0.806	0.72	0.52	0.817	0.723	0.547
900	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	0.992	0.98	0.926	0.996	0.985	0.904
2500	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	1	1	1	1	1	1

Setting B (three break points). Let $b_1 = 0.25n, b_2 = 0.5n, b_3 = 0.75n$ be three break points. For $i = 1, \cdots, b_1$ and $i = b_2 + 1, \cdots, b_3$, we set $m(t/T) = 0$ for $t = 1, \cdots, T/2$ and $m(t/T) = -1$ for $t = 1 + T/2, \cdots, T$. For $i = b_1+1, \cdots, b_2$ and $i = b_3+1, \cdots, n$, we let $m(t/T) = 2\log(1+(t/T)^2)$. Table 5 report the Monte Carlo power of the 10%, 5% and 1% tests under this setting.

Table 5. Monte Carlo Power of 10%, 5% and 1% tests, setting B

nT	test stat	Asymptotic test			Bootstrap test		
400	$\widehat{Q}$	1	1	0.979	1	1	0.995
	$\widehat{Q}'$	0.99	0.972	0.875	0.997	0.978	0.839
	$\widehat{U}$	0.993	0.972	0.86	1	1	0.971
	$\widehat{U}'$	0.86	0.772	0.543	0.873	0.789	0.503
900	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	0.99	0.971	0.902	0.993	0.976	0.918
2500	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	1	1	1	1	1	1

Setting C (single break point with linear trend functions). Let b denote the break point, for which values $0.5n$ and $0.75n$ were tried. For $i = 1, \cdots, b$, we set $m(t/T) = 1 + t/T$, while for $i = b + 1, \cdots, n$, we set $m(t/T) = 2 + 2t/T$. Tables 6-7 report the Monte Carlo power of the 10%, 5% and 1% tests under this setting.

In Settings A and B above, the ordering of the i index was such that the break points neatly divide the sample into subgroups of the same trend functions. However, it is plausible in the real data that the ordering of i index is not so informative. This motivated us to consider the following setting.

Setting D (random permutation). For $i = 1, \cdots, n/2$, we set $m(t/T) = 0$ for $t = 1, \cdots, T/2$ and $m(t/T) = -1$ for $t = 1 + T/2, \cdots, T$, while for $i = 1 + n/2, \cdots, n$, we let $m(t/T) = 2\log(1 + (t/T)^2)$. After data has been generated, the i index was shuffled using random permutation. Table 8 report the Monte Carlo power of the 10%, 5% and 1% tests under this setting.

Across all settings, the Monte Carlo power results seem broadly similar between tests based on asymptotic and bootstrap critical values. Also, as

TABLE 6. Monte Carlo Power of 10%, 5% and 1% tests, setting C, $b = 0.5n$

nT	test stat	Asymptotic test			Bootstrap test		
400	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	0.997	0.99	0.972	1	0.999	0.997
	$\widehat{U}'$	0.943	0.894	0.746	0.953	0.912	0.72
900	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	0.999	0.999	0.99	0.999	0.998	0.985
2500	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	1	1	1	1	1	1

TABLE 7. Monte Carlo Power of 10%, 5% and 1% tests, setting C, $b = 0.75n$

nT	test stat	Asymptotic test			Bootstrap test		
400	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	0.945	0.897	0.718	0.991	0.982	0.916
	$\widehat{U}'$	0.582	0.467	0.265	0.59	0.461	0.292
900	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	0.999	1	1	1
	$\widehat{U}'$	0.891	0.822	0.637	0.902	0.834	0.57
2500	$\widehat{Q}$	1	1	1	1	1	1
	$\widehat{Q}'$	1	1	1	1	1	1
	$\widehat{U}$	1	1	1	1	1	1
	$\widehat{U}'$	1	1	0.995	1	1	0.995

with the size results, the presence of regressor and the resulting need for the first stage estimation of the parameter β has not much altered the Monte Carlo power results between the simple and partial linear models. The results improved with increasing sample sizes and the reported Monte Carlo powers were close to 1 at $T, n = 50$ in all settings.

For both settings A and C which use a single break point, power performance was better when the break was at $0.5n$ compared to when it was $0.75n$. This is to be expected, since the former break point represents a

Table 8. Monte Carlo Power of 10%, 5% and 1% tests, setting D

nT	test stat	Asymptotic test			Bootstrap test		
400	$\widehat{Q}$	0.152	0.071	0.012	0.39	0.182	0.045
	$\widehat{Q}'$	0.493	0.323	0.096	0.52	0.362	0.135
	$\widehat{U}$	0.084	0.033	0.008	0.354	0.182	0.024
	$\widehat{U}'$	0.3	0.209	0.055	0.332	0.212	0.045
900	$\widehat{Q}$	0.999	0.985	0.88	1	1	0.951
	$\widehat{Q}'$	0.965	0.921	0.756	0.974	0.936	0.766
	$\widehat{U}$	0.937	0.872	0.7	0.995	0.983	0.894
	$\widehat{U}'$	0.738	0.608	0.371	0.751	0.638	0.412
2500	$\widehat{Q}$	1	1	0.999	1	1	1
	$\widehat{Q}'$	1	1	0.996	1	1	0.999
	$\widehat{U}$	1	0.998	0.947	1	1	0.996
	$\widehat{U}'$	0.98	0.933	0.75	0.986	0.953	0.839

further deviation away from the null of common trends/homogeneity than the latter.

For settings A, B and C in which the ordering of the i index is preserved, the worst power results were obtained by the $\widehat{U}'$ statistic under the presence of individual fixed effects.

In Setting D, where the ordering of the i index was disturbed, the Monte Carlo results differ between particular realisations of the random permutation of i index and the ones reported in the tables appear to be quite standard. Monte Carlo power results are poorer for this setting when compared to others, highlighting the effect of the ordering of i index on our test statistics. It is encouraging that the power results improves considerably as the sample size increases even in this setting.

6. **CONCLUSIONS**

In this paper we have described and examined a simple test for common trends of unspecified functional form as the number of individuals increases. We were able to perform the test without resorting to nonparametric estimation techniques so that we avoid the uneasy issue of how to choose a bandwidth parameter in a testing procedure like ours. We have allowed the panel regression model to have covariates entering in a linear form. We discuss the consequences that the estimation of the slope parameters may have in the asymptotic behaviour of the test, finding that there were no consequences, so that the asymptotic behaviour of the test is not altered by the estimation of the parameters.

There are several interesting issues worth examining as those already mentioned in the introduction. One of them is how we can extend this methodology to the situation where the sample sizes T_i for the different sequences of individual units $i = 1, ..., n$, are not necessarily the same. We believe that the methodology in the paper can be implemented after some smoothing has been put in place, for instance, via splines. A second relevant extension is what happens when the errors exhibit cross-sectional and/or time dependence. We conjecture that, after inspection of our proofs, the main conclusions of our results will still hold true if they are "weakly" dependent. However, the technical details to accomplish this and in particular those to obtain a consistent estimator of the asymptotic variance might be cumbersome and lengthy. We envisage, though, that it is possible to obtain a simple computational estimator of the long run variance of the test via bootstrap methods using results given in Section 3, together with those obtained by Chang and Ogden (2009). However the details are beyond the scope of this paper.

7. APPENDIX

7.1. Proof of Theorem 1.

Proof. We first notice that under H_0, we have that

$$
\begin{aligned}
(\mathbf{a}) \quad Q_{T,n,r,u} &= \frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\left(\varepsilon_{it} - \overline{\varepsilon}_{\cdot t}\right), \\
(\mathbf{b}) \quad Q'_{T,n,r,u} &= \frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]}\left(\varepsilon_{it} - \frac{1}{i-1}\sum_{j=1}^{i-1}\varepsilon_{jt}\right), \qquad r,u \in [0,1],
\end{aligned}
$$

We begin with part (**b**). To show weak convergence in (2.8) we need to establish convergence of finite dimensional distributions and tightness. To that end, we first we establish that as $T, n \to \infty$,

$$
E\left((Tn)Q'_{T,n,r_1,u_1}Q'_{T,n,r_2,u_2}\right) \to \sigma^2(r_1 \wedge r_2)(u_1 \wedge u_2). \tag{7.1}
$$

Denoting

$$
z_{it} =: \varepsilon_{it} - \frac{1}{i-1}\sum_{j=1}^{i-1}\varepsilon_{jt}, \tag{7.2}
$$

we have

$$
\sqrt{Tn}Q'_{T,n,r,u} = \sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]} z_{it}.
$$

Now, $C1$ and standard arguments imply that

$$
E\left\{z_{i_1 t}z_{i_2 s}\right\} = 0 \tag{7.3}
$$

if $i_1 \neq i_2$ and t, s. So, we have that

$$E\left(\sum_{t=1}^{[Tr_1]}\sum_{i=2}^{[nu_1]} z_{it} \sum_{t=1}^{[Tr_2]}\sum_{i=2}^{[nu_2]} z_{it}\right) = \sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[n(u_1\wedge u_2)]} E\left(z_{it}^2\right)$$
$$\to \sigma^2 (r_1\wedge r_2)(u_1\wedge u_2),$$

which completes the proof of (7.1).

Next, we establish that the finite dimensional distributions (fidi) of $Q'_{T,n,r,u}$ converge to those of $\sigma W(r,u)$. As usual because the Cramer-Wold device, we only need to show the convergence in distribution of

$$\sqrt{Tn}\, Q'_{T,n,r,u} \xrightarrow{d} \sigma W\left(r,u\right), \quad r,u \in [0,1]\,,$$

for any fixed r and u. We have already shown in (7.1) that $Var(\sqrt{Tn}Q'_{T,n,r,u}) \to \sigma^2 ru$. On the other hand, it is well known that for any fixed t,

$$Z_{nt} = \frac{1}{\sqrt{n}}\sum_{i=2}^{[nu]} z_{it} \xrightarrow{d} \mathcal{N}\left(0,\sigma^2 u\right),$$

and as Z_{nt} and Z_{ns} are independent by $C1$, we conclude that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} Z_{nt} \xrightarrow{d} \mathcal{N}\left(0,\sigma^2 ur\right).$$

So, we have established the CLT when we take sequentially the limits in n and then in T. However, we need to show that this is the case when they are taken jointly. To that end, it suffices to verify that Z_{nt} satisfy the generalized Lindeberg-Feller condition: for any $\eta > 0$, as $T, n \to \infty$,

$$q_T(r) := \sum_{t=1}^{[Tr]} E\left(Z_{nt}^2 I(|Z_{nt}| \geq \eta)\right) \to 0,$$

see Theorem 2 in Phillips and Moon (1999), being a sufficient condition

$$E\left(Z_{nt}^4\right) = E\left(\frac{1}{\sqrt{n}}\sum_{i=2}^{[nu]} z_{it}\right)^4 \leq C. \tag{7.4}$$

Proceeding as we did to show (7.1),

$$\begin{aligned}
E\left(Z_{nt}^{4}\right) &= \frac{3}{n^{2}}\left(\sum_{i=2}^{[nu]} E z_{it}^{2}\right)^{2}+\frac{1}{n^{2}} \sum_{i_{1}, \ldots, i_{4}=2}^{[nu]} cum\left(z_{i_{1} t}, z_{i_{2} t}, z_{i_{3} t}, z_{i_{4} t}\right) \\
&= 3 \sigma^{4}(1+o(1))+\frac{1}{n^{2}} \sum_{i_{1}, \ldots, i_{4}=2}^{[nu]} cum\left(\varepsilon_{i_{1} t}, \ldots, \varepsilon_{i_{4} t}\right) \\
&\quad +\frac{1}{n^{2}} \sum_{i_{1}, \ldots, i_{4}=2}^{[nu]} \frac{1}{\prod_{\ell=1}^{4} i_{\ell}-1} cum\left(\sum_{j=1}^{i_{1}-1} \varepsilon_{jt}, \ldots, \sum_{j=1}^{i_{4}-1} \varepsilon_{jt}\right) \\
&= 3 \sigma^{4} u^{2}+O\left(\frac{\log ^{3} n}{n}\right).
\end{aligned}$$

So the sufficient condition (7.4) holds true and hence $\sqrt{Tn} Q_{T,n,r,u}^{\prime} \Longrightarrow_{d} \sigma W(r, u)$.

To complete the proof of part (**b**), we need to show tightness. For that purpose, define an *increment* of $Q_{T,n,r,u}^{\prime}$ as follows. For $0 \leq r_{1} \leq r_{2} \leq 1$ and $0 \leq u_{1} \leq u_{2} \leq 1$,

$$\triangle_{r_{1}, u_{1}, r_{2}, u_{2}}^{\prime}:=\frac{1}{Tn} \sum_{t=[Tr_{1}]+1}^{[Tr_{2}]} \sum_{i=[nu_{1}]+1}^{[nu_{2}]} z_{it}.$$

A sufficient condition for tightness, see Bickel and Wichura (1971), is

$$E((Tn)^{2}|\triangle_{r_{1}, u_{1}, r_{2}, u_{2}}^{\prime}|^{4}) \leq C\left(\frac{[Tr_{2}]-[Tr_{1}]}{T}\right)^{2}\left(\frac{[nu_{2}]-[nu_{1}]}{n}\right)^{2}, \quad (7.5)$$

where we can assume that $n^{-1} \leq u_{2}-u_{1}$ and $T^{-1} \leq r_{2}-r_{1}$ since otherwise the inequality in (7.5) would hold trivially. Now, proceeding as we did with (7.1), we have that the left side of (7.5) is

$$\begin{aligned}
&\frac{3}{T^{2} n^{2}}\left(\sum_{t=[Tr_{1}]+1}^{[Tr_{2}]} \sum_{i=[nu_{1}]+1}^{[nu_{2}]} E z_{it}^{2}\right)^{2} \qquad (7.6)\\
&+\frac{1}{T^{2} n^{2}} \sum_{t_{1}, \ldots, t_{4}=[Tr_{1}]+1}^{[Tr_{2}]} \sum_{i_{1}, \ldots, i_{4}=[nu_{1}]+1}^{[nu_{2}]} cum\left(z_{i_{1} t_{1}}, z_{i_{2} t_{2}}, z_{i_{3} t_{3}}, z_{i_{4} t_{4}}\right).
\end{aligned}$$

By $C1$, we have that $cum\left(z_{i_1t}, z_{i_2t}, z_{i_3t}, z_{i_4t}\right) \neq 0$ only if $t_1 = ... = t_4$, so that the second term of (7.6) is

$$\begin{aligned}
&\frac{1}{T^2n^2}\sum_{t=[Tr_1]+1}^{[Tr_2]}\sum_{i_1,...,i_4=[nu_1]+1}^{[nu_2]} cum\left(z_{i_1t}, z_{i_2t}, z_{i_3t}, z_{i_4t}\right) \\
\leq\ & C\left(\frac{[Tr_2]-[Tr_1]}{T}\right)^2\left\{\frac{1}{n^2}\sum_{i_1,...,i_4=2}^{[nu]} cum\left(\varepsilon_{i_1t}, ..., \varepsilon_{i_4t}\right)\right. \\
&\left. + \frac{1}{n^2}\sum_{i_1,...,i_4=2}^{[nu]}\frac{1}{\prod_{\ell=1}^{4} i_\ell - 1} cum\left(\sum_{j=1}^{i_1-1}\varepsilon_{jt}, ..., \sum_{j=1}^{i_4-1}\varepsilon_{jt}\right)\right\} \\
\leq\ & C\left(\frac{[Tr_2]-[Tr_1]}{T}\right)^2\left(\frac{[nu_2]-[nu_1]}{n}\right)^2.
\end{aligned}$$

On the other hand, the first term of (7.6) is also bounded by the far right side of the last displayed inequality. So this concludes the proof of the tightness condition and so part (**b**) of the theorem.

We now look at part (**a**). The proof of part (**a**) proceeds similarly, if it is not easier, to that given in part (**b**) once we observe that

$$\begin{aligned}
\sqrt{Tn}Q_{T,n,r,u} &= \frac{1}{\sqrt{Tn}}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[un]}\left(\varepsilon_{it} - \frac{1}{n}\sum_{j=1}^{n}\varepsilon_{jt}\right) \\
&= \frac{1}{\sqrt{Tn}}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\varepsilon_{it} - u\frac{1}{\sqrt{Tn}}\sum_{t=1}^{[Tr]}\sum_{i=1}^{n}\varepsilon_{it}, \qquad (7.7)
\end{aligned}$$

so that it is omitted. □

7.2. **Proof of Theorem 2.**

Proof. As in the proof of Theorem 1, we shall only handled explicitly part (**b**), the proof of part (**a**) being similarly handled. Using (7.2) and then that

$$w_{it} = z_{it} - \frac{1}{t-1}\sum_{s=1}^{t-1} z_{is}, \qquad (7.8)$$

the notation given in the proof of Theorem 1, we have that

$$U'_{T,n,r,u} = \frac{1}{Tn}\sum_{t=2}^{[Tr]}\sum_{i=2}^{[nu]} w_{it}.$$

Before we look at the structure of the second moments of $U'_{T,n,r,u}$, it is worth emphasizing that if $t \neq s$ for all i,j $E\left(w_{it}w_{js}\right) = 0$. Bearing this in mind together with (7.3), we then have that

$$E\left(\sum_{t=1}^{[Tr_1]}\sum_{i=2}^{[nu_1]}w_{it}\sum_{t=1}^{[Tr_2]}\sum_{i=2}^{[nu_2]}w_{it}\right) = \sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[n(u_1\wedge u_2)]}E\left(w_{it}^2\right)$$
$$\to \sigma^2(r_1\wedge r_2)(u_1\wedge u_2),$$

because by definition of w_{it} and $C1$, $E\left(w_{it}^2\right) = Ez_{it}^2\left(t/\left(t-1\right)\right)$.

Next, we examine the convergence in distribution of

$$\sqrt{Tn}U'_{T,n,r,u} = \frac{1}{\sqrt{Tn}}\sum_{t=2}^{[rT]}\sum_{i=2}^{[un]}w_{it}.$$

As we argued in the proof of Theorem 1, for any fixed t, we have that

$$\frac{1}{\sqrt{n}}\sum_{i=2}^{[nu]}w_{it}\xrightarrow{d}\mathcal{N}\left(0,\sigma^2\left(1+\frac{1}{t-1}\right)u\right).$$

From here we obtain that

$$\frac{1}{\sqrt{T}}\sum_{t=2}^{[Tr]}\frac{1}{\sqrt{n}}\sum_{i=2}^{[nu]}w_{it}\xrightarrow{d}\mathcal{N}\left(0,\sigma^2 ru\right) \tag{7.9}$$

since the sequence $\frac{1}{\sqrt{n}}\sum_{i=2}^{[un]}w_{it}$ are independent in t by $C1$. However, we need to show that (7.9) also holds when the limit in n and T is taken jointly. To that end it suffices to show that

$$E\left(\frac{1}{\sqrt{n}}\sum_{i=2}^{[nu]}w_{it}\right)^4 \le C$$

but the left side of the last displayed expression is

$$\begin{aligned}
& \frac{3}{n^2}\left(\sum_{i=2}^{[nu]}Ew_{it}^2\right)^2 + \frac{1}{n^2}\sum_{i_1,\dots,i_4=2}^{[u_2 n]}cum\left(w_{i_1t},w_{i_2t},w_{i_3t},w_{i_4t}\right) \\
= \ & 3\sigma^4\left(1+o\left(1\right)\right)+\frac{1}{n^2}\sum_{i_1,\dots,i_4=2}^{[nu]}cum\left(z_{i_1t},\dots,z_{i_4t}\right) \\
& +\frac{1}{n^2}\sum_{i_1,\dots,i_4=2}^{[nu]}\frac{1}{\prod_{\ell=1}^{4}i_\ell-1}cum\left(\sum_{j=1}^{i_1-1}z_{jt},\dots,\sum_{j=1}^{i_4-1}z_{jt}\right) \\
= \ & 3\sigma^4u^2+O\left(\frac{\log^3 n}{n}\right).
\end{aligned}$$

So, we conclude that (7.9) also holds when the limits are taking jointly.

To finish the proof it remains to verify the tightness condition of

$$\sqrt{Tn}U'_{T,n,r,u} = \frac{1}{\sqrt{Tn}} \sum_{t=2}^{[rT]} \sum_{i=2}^{[un]} w_{it}.$$

To that end, and proceeding as in the proof of Theorem 1, it is enough to show that for $0 \le r_1 \le r_2 \le 1$ and $0 \le u_1 \le u_2 \le 1$,

$$\triangle'_{r_1,u_1,r_2,u_2} := \frac{1}{Tn} \sum_{t=[r_1 T]+1}^{[r_2 T]} \sum_{i=[u_1 n]+1}^{[u_2 n]} w_{it}$$

satisfies (7.5). The proof is essentially the same as that given for (7.5), i.e. (7.6), but with the obvious changes and so is omitted. This completes the proof of part (**b**). The proof of part (**a**) is essentially the same as, if not easier than, that of part (**b**), once we notice that

$$\begin{aligned}
\sqrt{Tn}U_{T,n,r,u} &= \frac{1}{\sqrt{Tn}} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \left(z_{it} - \frac{1}{T} \sum_{s=1}^{T} z_{jt} \right) \\
&= \frac{1}{\sqrt{Tn}} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \varepsilon_{it} - u \frac{1}{\sqrt{Tn}} \sum_{t=1}^{[Tr]} \sum_{i=1}^{n} \varepsilon_{it} \\
&\quad - r \frac{1}{\sqrt{Tn}} \sum_{t=1}^{T} \sum_{i=1}^{[nu]} \varepsilon_{it} + ru \frac{1}{\sqrt{Tn}} \sum_{t=1}^{T} \sum_{i=1}^{n} \varepsilon_{it},
\end{aligned}$$

and hence is omitted. □

7.3. **Proof of Theorem 3.**

Proof. The proof is quite standard after using Moon and Phillips (1999) results. First standard arguments indicate that

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \check{x}_{it} \check{x}'_{it} \xrightarrow{P} \Sigma.$$

Indeed this is the case because by $C3$, $(x_{it} - \varkappa_i)(x_{it} - \varkappa_i)'$ are i.i.d. sequences of random variables with finite second moments.

Next,

$$\begin{aligned}
\frac{1}{(nT)^{1/2}} \sum_{i=1}^{n} \sum_{t=1}^{T} \check{x}_{it} \varepsilon_{it} &= \frac{1}{(nT)^{1/2}} \sum_{i=1}^{n} \sum_{t=1}^{T} x_{it} \varepsilon_{it} \\
&\quad - \frac{1}{(nT)^{1/2}} \sum_{i=1}^{n} \sum_{t=1}^{T} \left(\frac{1}{T} \sum_{s=1}^{T} x_{is} \right) \varepsilon_{it} \left\{ \overline{x}_{i\cdot} + \overline{x}_{\cdot t} - \overline{x}_{\cdot\cdot} \right\},
\end{aligned}$$

where we have assumed that $\varkappa_i = 0$ as the left side of the last displayed equality is invariant to the value of $\varkappa_i$. It is clear that the asymptotic behaviour of the left side of the last displayed equality is given by that of

$$\frac{1}{(nT)^{1/2}} \sum_{i=1}^{n} \sum_{t=1}^{T} x_{it} \varepsilon_{it}. \tag{7.10}$$

To show that (7.10) converges in distribution to a normal random variable we employ Theorem 2 of Moon and Phillips (1999). But,

$$E \left(\frac{1}{T^{1/2}} \sum_{t=1}^{T} x_{it} \varepsilon_{is} \right)^4 \leq C,$$

by $C1$, $C3$ and standard results. This concludes the proof of the theorem. □

7.4. **Proof of Theorem 4.**

Proof. First by definition, we have that

$$\widehat{U}'_{T,n,r,u} = \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \breve{\varepsilon}_{it} - \left(\widehat{\beta} - \beta \right)' \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \breve{x}_{it}.$$

So compared to what we had in the case of no regressors x_{it}, we have now the extra term $\left(\widehat{\beta} - \beta \right)' \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \breve{x}_{it}$. But, as we can consider the situation where $E(x_{it}) = \varkappa_i \equiv 0$ without loss of generality as $E\breve{x}_{it} = 0$, we can easily conclude that

$$\sup_{r,u \in [0,1]} \left| \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \breve{x}_{it} \right| = o_p(1). \tag{7.11}$$

Indeed this is the case because for each $r, u \in [0, 1]$ we have

$$\breve{\mathbb{X}}_{nT}(r,u) =: \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \breve{x}_{it} = o_p(1),$$

by ergodicity. On the other hand, the tightness condition of $\breve{\mathbb{X}}_{nT}(r,u)$ proceeds as that in Theorem 3. Thus, we conclude that

$$\widehat{U}'_{T,n,r,u} = \frac{1}{Tn} \sum_{t=1}^{[Tr]} \sum_{i=1}^{[nu]} \breve{\varepsilon}_{it} + o_p\left((Tn)^{-1/2} \right)$$

uniformly in $r, u \in [0, 1]$, because by Theorem 4, $\widehat{\beta} - \beta = O_p\left((Tn)^{-1/2} \right)$. From here the proof of the theorem follows in a straightforward fashion. □

7.5. **Proof of Theorem 5.**

Proof. As we have done with the results in Theorem 1 and 3, we shall explicitly deal with the proof of part (**b**), part (**a**) being similarly handled. To that end, we first notice that we still have that $(Tn)^{1/2}\left(\widehat{\beta}-\beta\right) \to_d \mathcal{N}\left(0,\sigma^2\Sigma^{-1}\right)$. Next, proceeding as in the proof of Theorem 4, and using the notation in (7.2) and (7.8), we have that

$$\widehat{U}'_{T,n,r,u} = \frac{1}{Tn}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]} w_{it} + o_p\left((Tn)^{-1/2}\right),$$

uniformly in $r,u \in [0,1]$. However Condition $C5$ will yield to the fact that, say $E\left(z_{it}z_{jt}\right)$ can be different than zero even when $i \neq j$. Indeed for, standard algebra gives that

$$E\left(z_{it}z_{jt}\right) = \begin{cases} \sigma_i^2 + \frac{1}{(i-1)^2}\sum_{\ell=1}^{i-1}\sigma_\ell^2 & i=j \\ -\frac{1}{(ij)_+-1}\left(\sigma_{(ij)_-}^2 - \frac{1}{(ij)_--1}\sum_{\ell=1}^{(ij)_--1}\sigma_\ell^2\right) & i\neq j \end{cases} \tag{7.12}$$

using the notation $(ij)_- = i \wedge j$ and $(ij)_+ = i \vee j$. On the other hand,

$$E\left(w_{it}w_{is}\right) = 0 \quad \text{if } s \neq t. \tag{7.13}$$

So, the latter two displayed expressions indicate that the covariance structure of $(Tn)^{1/2}\,\widehat{U}'_{T,n,r,u}$ differs from when the second moments of ε_{it} are the same. Due to this, we now proceed to characterize the covariance structure of $(Tn)^{1/2}\,\widehat{U}'_{T,n,r,u}$, that is

$$\frac{1}{Tn}E\left(\sum_{t=1}^{[Tr_1]}\sum_{i=2}^{[nu_1]} w_{it}\sum_{t=1}^{[Tr_2]}\sum_{i=2}^{[nu_2]} w_{it}\right) = \frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[nu_1]}\sum_{j=2}^{[nu_2]} E\left(w_{it}w_{jt}\right) \tag{7.14}$$

by (7.13). Now,

$$\begin{aligned} E\left(w_{it}w_{jt}\right) &= E\left\{\left(z_{it} - \frac{1}{t-1}\sum_{s=1}^{t-1} z_{is}\right)\left(z_{jt} - \frac{1}{t-1}\sum_{s=1}^{t-1} z_{js}\right)\right\} \\ &= E\left(z_{it}z_{jt}\right) + \frac{1}{(t-1)^2}\sum_{s=1}^{t-1} E\left(z_{is}z_{js}\right) \end{aligned}$$

because by $C1$, $E\left(z_{it}z_{js}\right)=0$ if $t\neq s$. So, the right side of (7.14) is

$$\begin{aligned}
&\frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[nu_1]}\sum_{j=1}^{[nu_2]}\left\{E\left(z_{it}z_{jt}\right)+\frac{1}{(t-1)^2}\sum_{s=1}^{t-1}E\left(z_{is}z_{js}\right)\right\}\\
=\;&\frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[n(u_1\wedge u_2)]}\left\{Ez_{it}^2+\frac{1}{(t-1)^2}\sum_{s=1}^{t-1}Ez_{is}^2\right\} \qquad (7.15)\\
&+\frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[nu_1]}\sum_{j\neq i}^{[nu_2]}\left\{E\left(z_{it}z_{jt}\right)+\frac{1}{(t-1)^2}\sum_{s=1}^{t-1}E\left(z_{is}z_{js}\right)\right\}.
\end{aligned}$$

Now because (7.12), $\frac{1}{t-1}\sum_{s=1}^{t-1}Ez_{is}^2\leq C$ and standard algebra gives that the first term on the right of (7.15) is

$$\left(r_1\wedge r_2\right)\frac{1}{n}\sum_{i=2}^{[n(u_1\wedge u_2)]}\sigma_i^2\left(1+O\left(\frac{\log n}{n}+\frac{\log T}{T}\right)\right)$$

by (7.12). On the other hand, (7.12) implies that the contribution due to

$$\frac{1}{(t-1)^2}\sum_{s=1}^{t-1}E\left(z_{is}z_{js}\right)$$

into the second term on the right of (7.15) is $O\left(\frac{\log n\log T}{T}\right)$, so that the behaviour of second term on the right of (7.15) is given by that of

$$\begin{aligned}
&\frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[nu_1]}\sum_{j\neq i}^{[nu_2]}E\left(z_{it}z_{jt}\right)\\
=\;&-\left(r_1\wedge r_2\right)\frac{1}{n}\sum_{i=2}^{[nu_1]}\sum_{j\neq i}^{[nu_2]}\left\{\frac{1}{(ij)_+-1}\left(\sigma_{(ij)_-}^2-\frac{1}{(ij)_--1}\sum_{\ell=1}^{(ij)_--1}\sigma_\ell^2\right)\right\}.
\end{aligned}$$

Hence we conclude that the right side of (7.14) is governed by

$$\begin{aligned}
&\left(r_1\wedge r_2\right)\frac{1}{n}E\left(\sum_{i=1}^{[nu_1]}\sum_{j=1}^{[nu_2]}z_{it}z_{jt}\right)\\
=\;&\left(r_1\wedge r_2\right)\frac{1}{n}\sum_{i=2}^{[nu_1]}\sum_{j=2}^{[nu_2]}E\left(\varepsilon_{i1}-\left(\sum_{\ell=i}^{[nu_1]}\frac{1}{\ell}\right)\varepsilon_{(i-1)1}\right)\left(\varepsilon_{j1}-\left(\sum_{\ell=j}^{[nu_2]}\frac{1}{\ell}\right)\varepsilon_{(j-1)1}\right)
\end{aligned}$$

which implies that as $n,T\nearrow\infty$, it is

$$\left(r_1\wedge r_2\right)\lim_{n\to\infty}\frac{1}{n}\sum_{i=2}^{[n(u_1\wedge u_2)]}\sigma_i^2\left(1+\log\left(\frac{i}{nu_1}\right)+\log\left(\frac{i}{nu_2}\right)+\log\left(\frac{i}{nu_1}\right)\log\left(\frac{i}{nu_2}\right)\right). \qquad (7.16)$$

Next that the fidi of $(Tn)^{1/2}\,\widehat{U}'_{T,n,r,u}$ converge in distribution to a Gaussian random variable with mean zero and variance

$$\Phi\left(r,u;r,u\right)=r\lim_{n\to\infty}\frac{1}{n}\sum_{i=2}^{[nu]}\sigma_i^2\left(1+2\log\left(\frac{i}{nu}\right)+\log^2\left(\frac{i}{nu}\right)\right)$$

proceeds as standard and it is nothing different to the proof under $C1$. Finally the tightness condition, for which a sufficient condition is

$$(Tn)^2\,E\left(\sum_{t=[Tr_1]+1}^{[Tr_2]}\sum_{i=[nu_1]+1}^{[nu_2]}w_{it}\right)^4\leq C\left(r_2-r_1\right)^{1+\psi}\left(u_2-u_1\right)^{1+\psi},$$

for some $\psi>0$. Now the proof of (7.16) suggests that it suffices all we need to show that

$$\begin{aligned}&(Tn)^2\,E\left(\sum_{t=[Tr_1]+1}^{[Tr_2]}\sum_{i=[nu_1]+1}^{[nu_2]}\left(\varepsilon_{it}-\log\left(\frac{i}{n_2u}\right)\varepsilon_{(i-1)t}\right)\right)^4\\ \leq\;& C\left(r_2-r_1\right)^{1+\psi}\left(u_2-u_1\right)^{1+\psi}.\end{aligned}$$

However the proof of the last displayed inequality is almost routine by $C3$. This concludes the proof of part (**b**) of the theorem.

The proof of part (**a**) follows similarly after we notice that

$$\begin{aligned}\frac{1}{Tn}E\left(\sum_{t=1}^{[Tr_1]}\sum_{i=1}^{[nu_1]}\varepsilon_{it}\sum_{t=1}^{[Tr_2]}\sum_{i=1}^{[nu_2]}\varepsilon_{it}\right)&=\frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=1}^{[n(u_1\wedge u_2)]}E\left(\varepsilon_{it}\varepsilon_{jt}\right)\\&=\frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=1}^{[n(u_1\wedge u_2)]}\sigma_i^2.\end{aligned}$$

This concludes the proof of the theorem. □

7.6. **Proof of Theorem 6.**

Proof. By definition of $\widehat{\beta}^*$ we have that

$$(Tn)^{1/2}\left(\widehat{\beta}^*-\widehat{\beta}\right)=\left(\frac{1}{Tn}\sum_{t=1}^{T}\sum_{i=1}^{n}\check{x}_{it}\check{x}'_{it}\right)^{-1}\frac{1}{(Tn)^{1/2}}\sum_{t=1}^{T}\sum_{i=1}^{n}\check{x}_{it}\check{\varepsilon}^*_{it}.$$

So it suffices to show that

$$\frac{1}{(Tn)^{1/2}}\sum_{t=1}^{T}\sum_{i=1}^{n}\check{x}_{it}\check{\varepsilon}^*_{it}\overset{d}{\to}\mathcal{N}\left(0,\sigma^2\Sigma\right)\qquad(\textit{in probability}).$$

Arguing as we did in the proof of Theorem 4, say, it is clear that we only need to show that

$$\frac{1}{(Tn)^{1/2}}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(x_{it}-\varkappa_i\right)\varepsilon^*_{it}\overset{d}{\to}\mathcal{N}\left(0,\sigma^2\Sigma\right)\qquad(\textit{in probability}).$$

But this is standard see Shao and Tu (1995). □

7.7. **Proof of Theorem 7.**

Proof. We would just show part (**a**), with part (**b**) being similarly handled. From the continuous mapping theorem it suffices to show that

$$\sqrt{Tn}\widehat{U}^{*\prime}_{T,n,r,u} = \frac{1}{\sqrt{Tn}}\sum_{t=1}^{[Tr]}\sum_{i=1}^{[nu]}\left(\widehat{v}^{*}_{it} - \frac{\sum_{j=1}^{i-1}\widehat{v}^{*}_{js}}{i-1} - \frac{\sum_{s=1}^{t-1}\widehat{v}^{*}_{is}}{t-1} + \frac{\sum_{j=1}^{i-1}\sum_{s=1}^{t-1}\widehat{v}^{*}_{js}}{(i-1)(t-1)}\right)$$

converges to (in probability) to $W(r,u)$. Denoting

$$z^{*}_{it} =: \varepsilon^{*}_{it} - \frac{1}{i-1}\sum_{j=1}^{i-1}\varepsilon^{*}_{jt}, \quad w^{*}_{it} =: z^{*}_{it} - \frac{1}{t-1}\sum_{s=1}^{t-1} z^{*}_{is}$$

we write

$$\sqrt{Tn}\widehat{U}^{*\prime}_{T,n,r,u} = \frac{1}{\sqrt{Tn}}\sum_{t=1}^{[Tr]}\sum_{i=2}^{[nu]} w^{*}_{it}.$$

Now, because $\{\varepsilon^{*}_{it}\}_{t=1}^{T}$, $i \geq 1$, are i.i.d. sequences of random variables, $E^{*}\left\{z^{*}_{i_1 t} z^{*}_{i_2 s}\right\} = 0$; $E^{*}\left\{w^{*}_{i_1 t} w^{*}_{i_2 s}\right\} = 0$, which implies that

$$\begin{aligned}\frac{1}{Tn}E^{*}\left(\sum_{t=1}^{[Tr_1]}\sum_{i=2}^{[nu_1]} w^{*}_{it}\sum_{t=1}^{[Tr_2]}\sum_{i=2}^{[nu_2]} w^{*}_{it}\right) &= \frac{1}{Tn}\sum_{t=1}^{[T(r_1\wedge r_2)]}\sum_{i=2}^{[n(u_1\wedge u_2)]} E^{*}\left(w^{*2}_{it}\right) \quad (7.17)\\ &= (r_1\wedge r_2)(u_1\wedge u_2)\frac{1}{Tn}\sum_{t=1}^{T}\sum_{i=2}^{n} w^{2}_{it}.\end{aligned}$$

But the right side of (7.17) converges in probability to $(r_1\wedge r_2)(u_1\wedge u_2)\sigma^2$ by standard arguments.

Next, we establish that the finite dimensional distributions (fidi) of $\sqrt{Tn}\widehat{U}^{*\prime}_{T,n,r,u}$ converge (in probability) to those of $\sigma W(r,u)$. As usual because the Cramer-Wold device, we only show the convergence in distribution of

$$\sqrt{Tn}\widehat{U}^{*\prime}_{T,n,r,u} \xrightarrow{d} \sigma W(r,u), \quad r,u \in [0,1], \quad (\textit{in probability}) \qquad (7.18)$$

for any fixed r and u. To that end, it is well known that for any fixed t

$$W^{*}_{nt} = \frac{1}{\sqrt{n}}\sum_{i=2}^{[nu]} w^{*}_{it} \xrightarrow{d} \mathcal{N}\left(0,\sigma^2 u\right), \quad (\textit{in probability})$$

and as W^{*}_{nt} is independent of W^{*}_{ns} by construction, we conclude that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} W^{*}_{nt} \xrightarrow{d} \mathcal{N}\left(0,\sigma^2 ur\right), \quad (\textit{in probability}).$$

So, we have established that when taking limits in n and then in T sequentially, it yields that $\sqrt{Tn}\widehat{U}^{*\prime}_{T,n,r,u}$ converges in bootstrap distribution to $\sigma W(r,u)$. However, we need to show that this is the case when the limits in

n and T are taking jointly. To that end, it suffices to verify that W_{nt}^{*} satisfy the generalized Lindeberg-Feller condition: for any $\eta > 0$, as $T, n \to \infty$,

$$\sum_{t=1}^{[Tr]} E^{*}\left(W_{nt}^{*2} I(|W_{nt}^{*}| \geq \eta)\right) \overset{P}{\to} 0,$$

for which a sufficient condition is

$$E^{*}\left(W_{nt}^{*4}\right) = E^{*}\left(\frac{1}{\sqrt{n}} \sum_{i=2}^{[nu]} w_{it}^{*}\right)^{4} = C_{nT}, \tag{7.19}$$

with C_{nT} being an $O_p(1)$ sequence of random variables. By assumption,

$$\begin{aligned} E^{*}\left(W_{nt}^{*4}\right) &= \frac{3}{n^{2}}\left(\sum_{i=2}^{[nu]} E^{*} w_{it}^{*2}\right)^{2} + \frac{1}{n^{2}} \sum_{i_1,\ldots,i_4=2}^{[nu]} cum^{*}\left(w_{i_1 t}^{*}, w_{i_2 t}^{*}, w_{i_3 t}^{*}, w_{i_4 t}^{*}\right) \\ &= 3\left(\frac{1}{n}\sum_{i=2}^{n} w_{it}^{2}\right) + \frac{1}{n}\widehat{cum}\left(\varepsilon_{it}, \ldots, \varepsilon_{it}\right)\left(1 + O_p(1)\right) \\ &\quad + \frac{1}{n^{2}} \sum_{i_1,\ldots,i_4=2}^{[nu]} \frac{1}{\prod_{\ell=1}^{4} i_\ell - 1} \sum_{j=1}^{i_1 - 1} \widehat{cum}\left(\varepsilon_{jt}, \ldots, \varepsilon_{jt}\right) \\ &\leq C_{nT}. \end{aligned}$$

So, (7.19) holds true so is (7.18).

To complete the proof of part (**a**), we need to show tightness. For that purpose, define an increment of $Q_{T,n,r,u}^{\prime}$ as follows. For $0 \leq r_1 \leq r_2 \leq 1$ and $0 \leq u_1 \leq u_2 \leq 1$,

$$\triangle_{r_1,u_1,r_2,u_2}^{*\prime} := \frac{1}{Tn} \sum_{t=[Tr_1]+1}^{[Tr_2]} \sum_{i=[nu_1]+1}^{[nu_2]} w_{it}^{*}.$$

A sufficient condition for tightness given by

$$E^{*}((Tn)^{2} |\triangle_{r_1,u_1,r_2,u_2}^{*\prime}|^{4}) \leq C_{nT}\left(\frac{[Tr_2] - [Tr_1]}{T}\right)^{2}\left(\frac{[nu_2] - [nu_1]}{n}\right)^{2}, \tag{7.20}$$

and where $n^{-1} \leq u_2 - u_1$ and $T^{-1} \leq r_2 - r_1$ since otherwise (7.20) holds trivially. Now, proceeding as with (7.17), the left side of (7.20) is

$$\begin{aligned} &\frac{3}{T^{2} n^{2}}\left(\sum_{t=[Tr_1]+1}^{[Tr_2]} \sum_{i=[nu_1]+1}^{[nu_2]} E w_{it}^{*2}\right)^{2} \qquad (7.21) \\ &+ \frac{1}{T^{2} n^{2}} \sum_{t_1,\ldots,t_4=[Tr_1]+1}^{[Tr_2]} \sum_{i_1,\ldots,i_4=[nu_1]+1}^{[nu_2]} cum^{*}\left(w_{i_1 t_1}^{*}, w_{i_2 t_2}^{*}, w_{i_3 t_3}^{*}, w_{i_4 t_4}^{*}\right). \end{aligned}$$

By construction, $cum^{*}\left(w^{*}_{i_1t_1}, w^{*}_{i_2t_2}, w^{*}_{i_3t_3}, w^{*}_{i_4t_4}\right) \neq 0$ only if $t_1 = ... = t_4$, so that the second term of (7.21) is bounded by

$$\begin{aligned} & C_{nT}\left(\frac{[Tr_2]-[Tr_1]}{T}\right)^2 \frac{1}{n^2} \sum_{i_1,...,i_4=[nu_1]+1}^{[nu_2]} \left\{ \widehat{cum}\left(\varepsilon_{i_1t}, ..., \varepsilon_{i_4t}\right) \right. \\ & \qquad \left. + \left(\prod_{\ell=1}^{4} i_\ell - 1\right)^{-1} \sum_{j=1}^{i_1-1} \widehat{cum}\left(\varepsilon_{jt}, ..., \varepsilon_{jt}\right) \right\} \\ \leq \; & C_{nT}\left(\frac{[Tr_2]-[Tr_1]}{T}\right)^2 \left(\frac{[nu_2]-[nu_1]}{n}\right)^2 . \end{aligned}$$

On the other hand the first term of (7.21) is also bounded by the right side of (7.21). So this concludes the proof of the tightness condition and part (**a**) of the theorem. □

References

[1] Atak, A., Linton, O. and Xiao, Z. (2011): "A semiparametric panel model for unbalanced data with application to climate change in the United Kingdom," *Journal of Econometrics,* **164**, 92-115.

[2] Baltagi, B., Hidalgo, J. and Li, Q. (1996): "A nonparametric test for poolability using panel data," *Journal of Econometrics*, **75**, 345-367.

[3] Bickel, P. J. and Wichura, M. J.(1971): "Convergence criteria for multiparameter stochastic processes and some applications," *Annals of Mathematical Statistics*, **42**, 1656-1670.

[4] Breitung, J. and Pesaran, M.H. (2008). Unit roots and cointegration in panels. In L. Matyas and P. Sevestre (Eds.), *The Econometrics of Panel Data* (3rd ed.), 279–322, Berlin: Springer.

[5] Brown, R. L., Durbin, J. and Evans, J. M. (1975): "Techniques for testing the constancy of regression relationships over time," *Journal of Royal Statistical Society, Ser. B,* **37**, 149-192.

[6] Bryan, M.L. and Jenkins, S.P. (2013). Regression analysis of country effects using multilevel data: A cautionary tale. Mimeo, Institute for social and economic research.

[7] Chan, C. and Ogden, R.D. (2009): "Bootstrapping sums of independent but not identically distributed continuous processes with applications to functional data," *Journal of Multivariate Analysis*, **100**, 1291-1303.

[8] Degas, D. Xu, Z., Zhang, T. and Wu, W.B. (2012): "Testing for parallelism among trends in multiple time series," *IEEE Transactions on Signal Processing*, **60**, 1087-1097.

[9] Detter, H. and Paparoditis, E. (2009): "Bootstrapping frequency domain tests in multivariate time series with an application to testing equality of spectral densities," *Journal of Royal Statistical Society, Ser. B,* **71**, 831-857.

[10] Efron, B. (1979): "Bootstrap methods: another look at the jackknife," *Annals of Statistics*, **7**, 1-26.

[11] Ferraty, F. and Vieu, P. (2006). Nonparametric functional data analysis: Theory and practice. *Springer-Verlag*. Berlin.

[12] Galvao, A.F. and Kato, K. (2013): "Estimation and inference for linear panel data models under misspecification when both n and T are large," *Forthcoming in Journal of Business and Economic Statistics*.

[13] Gao, J. and Hawthorne, K. (2006): "Semiparametric estimation and testing of the trend of temperature series," *The Econometrics Journal,* **9**, 332-355.

[14] Hidalgo, J. and Souza, P. (2013). "Testing for equality of an increasing number of spectral density functions," Forthcoming.

[15] Hsiao, C. (1986): Analysis of Panel Data. *Cambridge University Press*.

[16] Jin, S. and Su, L. (2010): "A nonparametric poolability test for panel data models with cross section dependence," Mimeo. Singapore Management University.

[17] Pesaran, M.H. and Tosetti, E. (2011): "Large panels with common factors and spatial correlation," *Journal of Econometrics*, **161**, 182-202.

[18] Pesaran, M.H. and Yamagata, T. (2008): "Testing slope heterogeneity in large panels," *Journal of Econometrics*, **142**, 50-93.

[19] Phillips, P.C.B. (2001): "Trending time series and macroeconomic activity: Some present and future challenges," *Journal of Econometrics*, **100**, 21-27.

[20] Phillips, P.C.B. and Moon, R. (1999): "Linear regression limit theory for nonstationary panel data," *Econometrica*, **67**, 1057-1111.

[21] Phillips, P. C. B. and Sul, D. (2007): "Transition modeling and econometric convergence tests," *Econometrica*, **75**, 1771–855.

[22] Robinson, P.M. (2011a): "Asymptotic theory for nonparametric regression with spatial data," *Journal of Econometrics*, **165**, 5-19.

[23] Robinson, P.M. (2011b). Inference on power law spatial trends. Preprint. London School of Economics.

[24] Shao, J. and Tu, D. (1995). The jackknife and bootstrap. *Springer-Verlag*. Berlin.

[25] Zhang, Y., Su, L. and Phillips, P.C.B. (2011): "Testing for common trends in semiparametric panel data models with fixed effects," *Econometrics Journal*, **20**, 102-208.

[26] White, H. and Granger, C. (2011): "Consideration of trends in time series," *Journal of Time Series Econometrics*, **3**, 1–40.

Economics Department, London School of Economics, Houghton Street, London WC2A 2AE, UK.

E-mail address: f.j.hidalgo@lse.ac.uk

New College of the Humanities, 19 Bedford Square, London WC1B 3HH, UK

E-mail address: j.lee2@lse.ac.uk