INFERENCE AND TESTING BREAKS IN LARGE DYNAMIC PANELS WITH STRONG CROSS SECTIONAL DEPENDENCE

JAVIER HIDALGO AND MARCIA SCHAFGANS

ABSTRACT. This paper is concerned with various issues related to inference in large dynamic panel data models (where both n and T increase without bound) in the presence of, possibly, strong cross-sectional dependence. Our first aim is to provide a Central Limit Theorem for estimators of the slope parameters of the model under mild conditions. To that end, we extend and modify existing results available in the literature. Our second aim is to study two, although similar, tests for breaks/homogeneity in the time dimension. The first test is based on the *CUSUM* principle; whereas the second test is based on a Hausman-Durbin-Wu approach. Some of the key features of the tests are that they have nontrivial power when the number of individuals, for which the slope parameters may differ, is a "negligible" fraction or when the break happens to be towards the end of the sample. Due to the fact that the asymptotic distribution of the tests may not provide a good approximation for their finite sample distribution, we describe a simple bootstrap algorithm to obtain (asymptotic) valid critical values for our statistics. An important and surprising feature of the bootstrap is that there is no need to know the underlying model of the cross-sectional dependence, and hence the bootstrap does not require to select any bandwidth parameter for its implementation, as is the case with moving block bootstrap methods which may not be valid with cross-sectional dependence and may depend on the particular ordering of the individuals. Finally, we present a Monte-Carlo simulation analysis to shed some light on the small sample behaviour of the tests and their bootstrap analogues.

JEL classification: C12, C13, C23

Key words: Large panel data. Dynamic models. Cross-sectional strong-dependence. Central Limit Theorems. Homogeneity. Bootstrap algorithms.

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

Corresponding author: Marcia Schafgans. Tel.: +44 207 955 7512; fax: +44 207 955 6592.
E-mail addresses f.j.hidalgo@lse.ac.uk (J. Hidalgo), m.schafgans@lse.ac.uk.(M. Schafgans)

1. INTRODUCTION

Nowadays it is widely recognized that economic agents are interrelated due to common factors, contagion, spillovers and so on. This dependence has been systematically neglected until quite recently in econometrics, possibly, due to a lack of a clear framework to characterize such a dependence which is exacerbated by the fact that, contrary to time series data, there is an absence of a clear or natural ordering of the data. In response to this, in the last decade or so, a huge amount of work has been directed to the study of cross-sectional dependence and several approaches or models have been put forward.

One way to model the cross-sectional dependence among individuals is by using a common (unobserved) factor models as in Andrews (2005), Pesaran (2006) or Bai (2009). A second approach is based on the "distance" of individuals located on a regular pattern in the plane, or lattice. It recognizes that data may be collected on a regular lattice as a consequence of planned experiments or a result of a systematic sampling scheme. Applications which use this type of data cover various areas like environmental, urban, agricultural economics as well as economic geography among others. Early examples of this are the celebrated papers by Mercer and Hall (1911) on wheat crop yield data and Batchelor and Reed (1924) on fruit trees, that were further analyzed by Whittle (1954). Other examples are given in Cressie and Huang (1998) and Fernández-Casal et al. (2003). Examples of lattice models in environment economics include Mitchell et al. (2005), who study the effect of CO_2 on crops, and Genton and Koul (2008), who analyze the effect of pollutants transported by winds on the yield of barley in UK.

A third approach to explain or model cross-sectional dependence is through the introduction of measures related to economic and/or geographical distance. This approach was advocated by Conley (1999) and followed by Chen and Conley (2001). The benefit of this approach, similar to lattice models, is that the statistical behaviour is reminiscent of that in standard time series analysis. Another approach that has received a lot of attention is the so-called SAR model, where the dependence is modelled as a linear transformation of "n" (sample size) independent and identically distributed (*iid*) random variables. This approach, considered as a variant of the model considered in Whittle (1954), was advocated in the geographic-economic literature by Cliff and Ord (1981) and it has been extensively employed in the econometric literature, see for instance Lee (2004) and Kelejian and Prucha (2007) among many others. One of the main difference with lattice data is that, contrary to the latter approach, we cannot consider the data/individuals as being collected in a systematic fashion. It is precisely this difference which makes the estimation and study of its properties more difficult and challenging.

In this paper, we characterize the cross-sectional dependence of, say the sequence $\{u_i\}_{i\in\mathbb{N}}$, through a model of the form $u_i = \sum_{j=0}^{\infty} a_j(i)\,\varepsilon_j$, where $\{\varepsilon_j\}_{j\in\mathbb{N}}$ are *iid* random variables. This approach was also considered by Robinson (2011) and Lee and Robinson (2013) and it has a strong resemblance with the well known Wold decomposition for time series sequences. Our motivation for using this approach is that it enables us to generate more general dependence structures than the SAR models can generate, in particular it permits dependence structures with "*strong-dependence*" or "*long-memory*", see our Definition 1 below. With this view, the SAR model can be considered as a particular scenario to the approach followed in this paper as we explain further in Section 2.

Let us introduce what we understand by "*strong-dependence*".

Definition 1. *The generic sequences* $\{\nu_{it}\}_{t\in\mathbb{Z}}$, $i\in\mathbb{N}^+$ *are* "strong-dependent" *if the sequence*

$$\frac{1}{n}\sum_{i,j=1}^{n}|\varphi_\nu(i,j)|$$

is not bounded in n*, where we denote*

$$\varphi_{\nu}(i,j) = Cov(\nu_{it}, \nu_{jt}). \tag{1.1}$$

Our Definition 1 draws a lot of similarities with one of the characterizations often employed to describe "*long-memory*" dependence for a time series sequence $\{q_t\}_{t\in\mathbb{Z}}$. That is, where $\{q_t\}_{t\in\mathbb{Z}}$ exhibits the property of "*long-memory*" if $\frac{1}{T}\sum_{t,s=1}^{T}|Cov(q_t,q_s)|$ is not bounded in T, the sample size. A similar definition for cross-sectional "*weak-dependence*" was used in Sarafidis and Wansbeek (2010). While Chudik, Pesaran and Tosetti (2011) also consider the presence of *strong-* and *weak-dependence* in large panels, they describe the dependence using a factor model, whereas ours is closer related to that given for time series sequences or SAR models. Finally observe that our definition of "*strong-dependence*" does not involve or require any ordering of the observations or the definition of some economic/geographical metric across observations.

This paper is therefore concerned with inference in (linear) dynamic panel data models exhibiting, possibly, strong cross-sectional dependence when both the number of cross-section units and time increase to infinity. Our dynamic panel data model is

$$y_{it} = \alpha_t + \eta_i + \sum_{\ell=1}^{k_1}\rho_{t\ell}y_{i(t-\ell)} + \theta_t' z_{it} + u_{it}, \quad i = 1,...,n, \quad t = 1,...,T, \tag{1.2}$$

where θ_t is a $k_2 \times 1$ vector of unknown parameters, $\{z_{it}\}_{t\in\mathbb{Z}}$ is a vector of exogenous covariates and $\{u_{it}\}_{t\in\mathbb{Z}}$ is the sequence of error terms, $i \in \mathbb{N}^+$. As usual α_t and η_i represent respectively the time and individual fixed effects. We shall assume that the sequences $\{z_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, are mutually independent of the error term $\{u_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, although not necessarily independent from the fixed effects α_t or η_i. More specific conditions on the sequences $\{u_{it}\}_{t\in\mathbb{Z}}$ and exogenous variables $\{z_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, will be given in Conditions $C1$ and $C2$ respectively in Section 2 below.

One of our main interest in the paper is to incorporate this cross-sectional dependence structure to further enhance the already extensive literature on (dynamic) panel data models. With this view, the main objectives in this paper are twofold. The first goal is to discuss and examine the asymptotic properties, and provide a new *Central Limit Theorem*, of estimators of the slope parameters of (1.2) when the cross-sectional dependence of the error sequences $\{u_{it}\}_{t\in\mathbb{Z}}$ and covariates $\{z_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, are (possibly) "*strong-dependent*". In particular, we provide very mild and general conditions to guarantee that the estimators of the parameters of the model are asymptotically normal. Our Central Limit Theorem results extend substantially the work by Kapoora, Kelejian and Prucha (2007), Yu, DeJong and Lee (2008) or Lee and Yu (2010) among others, as we allow for more general cross-sectional dependence structures that permits "*strong-dependence*". However to do so, we need to extend a Central Limit Theorem provided in Phillips and Moon (1999) to allow both for time and cross-sectional dependence. In their work, the sequences of random variables, say $\{\psi_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}$, are assumed to be such that $\{\psi_{it}\}_{t\in\mathbb{Z}}$ and $\{\psi_{jt}\}_{t\in\mathbb{Z}}$ are independent, which is a condition ruled out in our scenario. Unlike Phillips and Moon (1999), see also Hahn and Kürsteiner (2001), we cannot view the sequences as being independent in one of its dimensions. In addition, as we allow for "*strong-dependence*", we cannot use results and arguments based on any type of "*strong-mixing*" arguments, so that results in Jenish and

Prucha $(2009, 2012)$ cannot be used in our framework either. On the other hand, similar to what happens with time series regression models, see Robinson and Hidalgo (1997), we do need to restrict the strength of the cross-sectional dependence to guarantee that our estimator of the slope parameters converge in distribution with the standard root-nT rate and, more importantly, that they are asymptotically normal, see also Hidalgo (2003). As the work by Robinson and Hidalgo (1997) suggests, we might, of course, relax the strength of dependence at the expense of further complication in the mathematical apparatus by using some type of "weighted" fixed effect estimator. See our discussion of the conditions in the next section for further details and insights.

Our second main goal in this paper is to examine tests for breaks or homogeneity of the slope parameters in the model (1.2). Although similar models as the one in (1.2) have been considered, their interest has focussed on detecting the presence of heterogeneity across the cross-section units, that is the interest is on whether the slope parameters are the same for all $i \geq 1$. See for instance Pesaran and Smith (1995) or Pesaran and Yamagata (2008) whose framework and ours mainly differ in that our conditions are somehow milder than theirs and we allow for very general type of cross-sectional dependence that may exhibit some type of "*long-memory*" behaviour. Specifically, denoting in what follows $\beta_t = \left(\{\rho_{t\ell}\}_{\ell=1}^{k_1} ; \theta_t' \right)'$, we are interested in the null hypothesis

$$H_0 : \beta_t = \beta \quad \text{for all } [T\epsilon] \leq t \leq T - [T\epsilon] , \tag{1.3}$$

where $0 \leq \varepsilon \leq \frac{1}{2}$, with the alternative hypothesis being the negation of the null.

Alternatively, drawing notation and arguments from the time series literature, since our panel model (1.2) can be written as

$$y_{it} = \eta_i + \alpha_t + \beta' x_{it} + \delta' x_{it} \mathcal{I}\left(t > t_0\right) + u_{it}, \quad i = 1, ..., n, \quad t = 1, ..., T,$$

where in what follows we shall abbreviate $\left(\{y_{i,t-\ell}\}_{\ell=1}^{k_1} ; z_{it}' \right)'$ by x_{it}, we might write our hypothesis testing as

$$H_0 : \delta = 0 \quad \text{for all } [T\epsilon] \leq t_0 \leq T - [T\epsilon] ,$$

where $0 \leq \epsilon \leq 1/2$ against the alternative hypothesis

$$H_1 : \exists \ [T\epsilon] \leq t_0 \leq T - [T\epsilon] , \qquad \delta \neq 0.$$

In this respect, we can view our work as an extension of the relatively scarce work of breaks in the context of multivariate equations. See nevertheless the work by Bai, Lumsdaine and Stock (1998) for multivariate models and Bai (2000) on VAR models; see also Qu and Perron (2007). While their framework is for a fixed, and thus finite, n, in this paper we are concerned with a setup which allows "n" to increase with no limit as well. So, we can regard our hypothesis testing as one for structural breaks when the number of sequences, say $i = 1, ..., n$, increases with no limit. Hence we are in a framework of testing for many, possibly thousands, hypotheses simultaneously, see for instance Fan, Hall and Yao (2007). The testing for breaks has also some resemblance to the problem of testing whether a function or curve is constant, with the function of interest being $\beta_t = \beta\left(t/T\right)$ and we want to test $H_0 : \beta\left(\tau\right) = \beta$ for all $\tau \in [0, 1]$. See also the work by Juhl and Xiao (2013) for the latter interpretation of the test.

We now make some general comments about our hypothesis testing. Although we explicitly consider the scenario of abrupt "breaks" when testing our hypothesis in (1.3), our tests also have nontrivial power when the change is gradual, that is when under the (local) alternative the slope parameters β_t move to their new regime as a continuous function in t; see our discussion in Section 3.3 below. A second point to mention is that implicitly we are assuming that the break, if there were any, would be an interior point of a compact subset of $[0,1]$; the introduction of weight functions (or normalizations as in Andrews (1993)) discussed in Section 2 below, effectively guarantees the latter (see also our more explicit comments after Theorem 2 and Corollary 1 below). It might then be of interest to see what would happen with the behaviour of the test when we allow the break to happen towards the end of the sample, namely $T - m_0 \leq t_0$, where m_0 can be a finite positive constant. Recall that in typical situations, we take $\epsilon = .05$ or $.10$, so that we leave 10% or 20% of the data out. However this choice is no more than arbitrary and the power of the test may depend on its choice. The technical aspects of such a case are completely different as one can observe from recent work by Hidalgo and Seo (2013). In fact, for the latter scenario, it is apparent that one would need strong approximation results for an increasing dimensional vector of partial sums of random variables in our setting. Although some preliminary ideas and results might be drawn from the recent work in Chernozhukov et al. (2013), they are unfortunately not immediately useful for the purpose of testing for breaks towards the end of the sample and more importantly their work need to be extended when the assumption of independence is dropped. This situation is beyond the scope of this manuscript. Nevertheless, we do pay particular attention to the type of alternative models that our tests are able to detect and more specifically their behaviour under local alternatives. Scenarios that raise very naturally in our context: (**i**) the consequences when the time of the break is towards the end of the sample, that is the break time k_0 satisfies $k_0 > T - [h_T]$, where $[h_T]$ may satisfy $[h_T] = o(T)$; (**ii**) the consequences when the number of sequences/individuals for which a break exists is negligible when compared to the number of individuals in the sample; and (**iii**) the consequences when the breaks are at different times for different individuals or a combination of all of them. Of course one can imagine a combination of all three scenarios. We shall discuss some issues regarding the consistency of our tests in scenarios (**i**) and (**ii**).

Finally the paper describes a bootstrap approach for our estimators and tests. The motivation for this comes from the fact that the Monte-Carlo simulation experiment suggests that critical values drawn from the asymptotic distribution do not provide a good approximation to the finite sample behaviour of the test. One main reason for this originates from our general/mild conditions on the cross-sectional dependence which may result in a poor "nonparametric" estimator of the covariance structure of our statistics. In such a situation bootstrap techniques may be employed in the hope to improve the finite sample behaviour. To that end, we shall describe and examine two very simple bootstrap algorithms with have the appealing feature that there is no need to provide any estimate of the covariance structure of the error term. As a consequence, the bootstrap algorithms avoid the rather unpleasant need of time series inspired bootstrap methods which depend on (or make use of) some type of some "ad hoc" distance among the errors (observations), and hence there is no need to choose any bandwidth parameter, as is the case with time series, to implement a valid bootstrap approach. One of our findings is that the size of our tests is not affected by the choice of ϵ (trimming).

The remainder of the paper is organized as follows. In the next section, we discuss the regularity conditions of our model and provide a Central Limit Theorem for the slope parameters of the model (1.2) given either heterogeneity or homogeneity of the slope parameters. Section 3 discusses our test procedures for the null hypothesis of homogeneity. A whole broad family of tests are provided that make use of a weighting function $w(\tau)$, where typical choices are $w(\tau) = 1$ and $w(\tau) = \tau^{1/2}(1-\tau)^{1/2}$. We discuss local alternatives and consistency of our tests, showing that our tests have nontrivial power for sequences converging to zero faster than elsewhere, see Pesaran and Yamagata (2008). Their tests therefore have zero asymptotic relative efficiency when compared to ours. Section 4 discusses a bootstrap approach to our tests in view of the fact that the asymptotic distribution sometimes might provide a poor approximation to the finite sample critical values. A second motivation for the use of the bootstrap is that in model (1.2), say, the covariance structure can be quite complicated, so that bootstrap algorithms may be the only suitable solution to even compute valid critical values for the test. Section 5 presents a Monte Carlo simulation experiment to shed some light on the finite sample performance of our tests and the behaviour of the bootstrap counterpart. Section 6 gives a summary and describe possible extension of our results in several directions of interest. Finally, the proofs of our main results are provided in the Appendix.

2. REGULARITY CONDITIONS AND ASYMPTOTIC PROPERTIES OF THE SLOPE PARAMETER ESTIMATORS

Before we discuss and describe the statistical properties for estimators of the parameters β_t in (1.2), we first introduce a set of regularity conditions on the model and discuss the statistical properties of the covariates and error term. We assume that, for all $t \geq 1$, all the roots of the polynomials $\left|1 - \sum_{\ell=1}^{k_1} \rho_{t\ell} L^{\ell}\right| = 0$ are outside the unit interval, so we are not considering panel data models with possible unit roots under either the null or the alternative hypothesis as in Phillips and Moon (1999) or Im, Pesaran and Shin (2003).

Our regularity conditions are given next.

C1: $\{u_{it} = \sigma_i v_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, *are zero mean sequences of random variables, where* $0 < \sigma^{-1} < \sigma_i < \sigma < \infty$ *and the sequences* $\{v_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, *satisfy*

(i) $E(v_{it} \mid \mathcal{V}_{i,t-1}) = 0$; $E\left(v_{it}^2 \mid \mathcal{V}_{i,t-1}\right) = 1$ *and finite fourth moments, with* $\mathcal{V}_{i,t}$ *denoting the* $\sigma-$*algebra generated by* $\{v_{is},\, s \leq t\}$.

(ii) *For all* $t \in \mathbb{Z}$,

$$v_{it} = \sum_{\ell=1}^{\infty} a_\ell(i)\,\varepsilon_{\ell t}, \qquad \sum_{\ell=1}^{\infty} |a_\ell(i)|^2 < \infty,$$

where $\{\varepsilon_{\ell t}\}_{t\in\mathbb{Z}}$, $\ell \in \mathbb{N}^+$, *are zero mean independent identically distributed (iid) random variables with finite fourth moments. The weights* $\{a_\ell(i)\}_{i=1}^n$ *satisfy*

$$\sup_{\ell\geq 1}\sum_{i=1}^n |a_\ell(i)|^2 < \infty. \tag{2.1}$$

C2: $\{z_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, *are sequences of random variables such that:*

(**i**)
$$z_{it} = \mu_t + \sum_{k=0}^{\infty} c_k(i)\,\chi_{i,t-k}, \qquad \sum_{k=0}^{\infty} c_k k^{1/2} < \infty,$$

where, denoting by $\|B\|$ *the norm of the matrix* B, $c_k = \max_{i\geq 1}\|c_k(i)\|$ *and* $E(\chi_{it} \mid \Upsilon_{i,t-1}) = 0$*;* $Cov(\chi_{it} \mid \Upsilon_{i,t-1}) = \Sigma_\chi$ *and* $E\|\chi_{it}\|^4 < \infty$, *with* $\Upsilon_{i,t}$ *denoting the* $\sigma-$*algebra generated by* $\{\chi_{is},\ s \leq t\}$.

(**ii**) *The sequences of random variables* $\{\chi_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, *are such that*

$$\chi_{it} = \sum_{\ell=1}^{\infty} b_\ell(i)\,\eta_{\ell t}, \qquad \sum_{\ell=1}^{\infty}\|b_\ell(i)\|^2 < \infty,$$

where $\{\eta_{\ell t}\}_{t\in\mathbb{Z}}$, $\ell \in \mathbb{N}^+$, *are zero mean iid random variables with finite fourth moments and*

$$\sup_{\ell\geq 1}\sum_{i=1}^n \|b_\ell(i)\|^2 < \infty. \tag{2.2}$$

(**iii**) *Denoting* $\Sigma_{x,i} = Cov(x_{it}; x_{it})$, *we have that*

$$0 < \Sigma_x = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \Sigma_{x,i}. \tag{2.3}$$

C3: *For all* $i \in \mathbb{N}^+$, *the sequences* $\{u_{it}\}_{t\in\mathbb{Z}}$ *and* $\{z_{it}\}_{t\in\mathbb{Z}}$ *are mutually independent and*

$$0 < \max_{1\leq i\leq n}\sum_{j=1}^n \|\varphi_u(i,j)\,\varphi_z(i,j)\| < \infty, \tag{2.4}$$

where for any $i, j \geq 1$, *as defined in* (1.1),

$$\varphi_u(i,j) = Cov(u_{it}; u_{jt}), \quad \varphi_z(i,j) = Cov(z_{it}; z_{jt}).$$

C4: $T, n \to \infty$ *such that* $n^{-1} = o\left(T^{-\xi}\right)$ *for any* $\xi > 0$.

We now comment on our conditions. Conditions $C1$ and $C2$ indicate that we do not allow for temporal dependence on the errors $\{u_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$. Of course, it is possible to relax the latter condition, allowing u_{it} to follow a model similar to that for z_{it} as given in $C2$, in which case we might name (1.2) a "*stochastic difference equation panel model*". The only major difference that we might encounter is that in the latter scenario the estimation procedure would involve instrumental variables with $\{z_{i,t-\ell}\}_{\ell=1}^{k_1}$ as natural instruments for $\{y_{i,t-\ell}\}_{\ell=1}^{k_1}$. However, this is beyond the scope of the present manuscript as it will only add some extra lengthy technicalities and/or considerations which are well known when $n = 1$.

While both cross-sectional and temporal dependence are allowed to be present at the same time on $\{z_{it}\}_{t\in\mathbb{Z}}$, as it would then be the case for $\{y_{it}\}_{t\in\mathbb{Z}}$, we have assumed otherwise a separable covariance dependence structure as it is known in the argot of the spatio-temporal literature. See for instance Cressie and Huang (1999) or Gneiting (2002). Indeed a simple algebra yields that

$$Cov\left(z_{it}; z_{js}\right) = \gamma_{z,i}\left(\left|t-s\right|\right)\varphi_{\chi}\left(i,j\right),$$

where $\gamma_{z,i}\left(\ell\right) = \sum_{k=0}^{\infty} c_k\left(i\right)c_{k+|\ell|}\left(i\right)$ and $\varphi_{\chi}\left(i,j\right) = \varphi_{z}\left(i,j\right)$. This type of dependence is often assumed in empirical work due to its practicality and also in view of the difficulty to write down explicit models when the covariance structure of the data is not separable. Nevertheless, it should be noted that the separability condition can be tested, see for instance Matsuda and Yajima (2004). Of course, we can modify this condition allowing the sequences $\{z_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^{+}$, to satisfy some type of mixing condition such as L^4- Near Epoch dependence with size greater than or equal to 2, see Davidson (1994). The latter type of dependence might be useful from a theoretical/technical point of view if we allow, say that the errors exhibits some form of nonlinear type of dependence and/or we allow them to suffer from heteroscedasticity of the type $\sigma^2\left(z_{it}\right)$. Another model where the latter type of dependence proves to be very convenient from a technical point of view is when we have a nonlinear dynamic panel models, say

$$y_{it} = \eta_i + \alpha_t + g\left(y_{i,t-1};\rho_t\right) + \theta_t' z_{it} + u_{it}, \quad i = 1,...,n, \quad t = 1,...,T,$$

similar to the nonparametric model examined in Hjellvik, Chen and Tjøstheim (2004). Since the conclusions of our results should follow with L^4-Near Epoch dependence as it has been shown in an ample number of situations, we have decided to keep $C1$ and $C2$ are they stand to facilitate the proof of the CLT of our estimators which is non standard and requires modifications of existing results due to our mild conditions. On the other hand, our condition that $\sum_{k=0}^{\infty} c_k k^{1/2} < \infty$ rules out temporal "*strong-dependence*" for the regressors z_{it}, and hence on y_{it}. There is no doubts that we can relax this assumption to allow for "*strong-dependence*" among the regressors z_{it} as well as the errors u_{it}, at the expense of complicating our technical appendix quite considerably. However, as there are multiple examples where the results follow whether the data is "*weak-dependence*" or "*strong-dependence*" we have decided to keep our condition $C2$ for simplicity. Regardless on whether we allow the latter relaxation on the Conditions $C1$ and $C2$, the conditions are quite mild and as they stand makes our proofs already quite technical.

It is worth noticing that we are not assuming that the temporal dynamic behaviour of the sequences $\{z_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^{+}$, is common among the cross sectional units, so that we allow for some form of heterogeneity in the second moments of the data. That is,

$$\begin{aligned} Cov\left(z_{it}; z_{it}\right) &= \left(\sum_{k=0}^{\infty} c_k\left(i\right)\right)\Sigma_{\chi}\left(\sum_{k=0}^{\infty} c_k\left(i\right)\right)' \\ &= \Sigma_{z,i}, \end{aligned} \tag{2.5}$$

which is constant in "i" if $c_k\left(i\right) = c_k$ for all $k \geq 0$. This is in line with the assumption in Pesaran and Smith (1995). In addition, we allow for some trending behaviour which is in tune with Kim and Sun (2013). However, when $\mathcal{T}\left(r\right)$ or $\mathcal{T}^{\Delta}\left(r\right)$ given below in (3.1) and (3.4) respectively are evaluated at $r = T$, then there is no difference whether $Ez_{it} = \mu_t$ or $Ez_{it} = \mu$, say. Our conditions relax the moment conditions needed elsewhere, for instance those in Pesaran and Yamagata (2008), who assume finite moments of order greater than 4.

We next turn our focus on the discussion of the cross-sectional dependence induced by our Conditions $C1$ and $C2$. As elsewhere, see Lee and Robinson (2013), we allow for cross-sectional dependence to be driven by the models outline in parts (**ii**) of Conditions $C1$ and $C2$. In this sense our conditions relax considerably models employed elsewhere, for instance, our conditions allow the usual SAR (or more generally $SARMA$) models. Indeed, by definition of the SAR model, we have

$$\begin{aligned} u &= (I - \omega W)^{-1} \varepsilon \\ &= (I + \Xi) \varepsilon, \quad G = (\psi_j(i))_{i,j=1}^n, \end{aligned}$$

so that $u_i = \sum_{j=0}^n \psi_j(i) \varepsilon_j$, which implies that the SAR model can be regarded as a particular model of that allowed in $C1$ or $C2$. In addition, it is worth noting that in $C1$ the sequence $\sum_{i=1}^n |a_\ell(i)|$ is permitted to grow with n, which is not the case with the SAR model. So, in this case our conditions are weaker than those typically assumed when cross-sectional dependence is allowed. Of course we can allow the weights $a_\ell(i)$ to depend also on the sample size "n" as it is often done in SAR models with weight matrices W rowed normalized, however, the latter does not add anything different. With $\sigma_i < \sigma < \infty$, moreover, we observe that $\left(\sum_{\ell=1}^\infty \sum_{i=1}^n |a_\ell(i)|^2\right)^{-1} \to_{n \nearrow \infty} 0$. While an alternative approach to model, possibly "*long-memory*", cross-sectional dependence is through the presence of common (unobserved) factors, as in Pesaran (2006) and Bai (2009), we have decided to follow the model assumed in $C1$ due to its similarities with time series models and the fact that it can be considered as a natural generalization of the empirically popular SAR models. Finally, we can mention that $C2$ (**iii**) implies that we can allow for some form of multicollinearity among the regressors z_{it}, but only for a fraction of individuals, as (2.3) indicates that all we need is that on "average" there is no multicollinearity.

We next discuss our Condition $C3$. The first important point to remark is that expression (2.4) does not imply that

$$g_u(n) = \frac{1}{n} \sum_{i,j=1}^n |\varphi_u(i,j)| \quad \text{or} \quad g_z(n) = \frac{1}{n} \sum_{i,j=1}^n \|\varphi_z(i,j)\|$$

are bounded with n, i.e. that $g_u(n) + g_z(n) < C$, although it does imply that

$$0 < \lim_{n \to \infty} \left\| \frac{1}{n} \sum_{i,j=1}^n \varphi_u(i,j) \varphi_z(i,j) \right\| < \infty. \tag{2.6}$$

In fact, $g_u(n)$ and/or $g_z(n)$ can be such that they diverge to infinity with n, in which case $\{z_{it}\}_{t \in \mathbb{Z}}$ and $\{u_{it}\}_{t \in \mathbb{Z}}$, $i \in \mathbb{N}^+$ are "*strong-dependent*" sequences. On the other hand, their combined cross-sectional dependence, that is the dependence of the sequence $\{w_{it} = (z_{it} - E(z_{it})) u_{it}\}_{t \in \mathbb{Z}}$, $i \in \mathbb{N}^+$, satisfies

$$g_w(n) = \frac{1}{n} \sum_{i,j=1}^n \|\varphi_w(i,j)\| < C,$$

so that $\{w_{it}\}_{t\in\mathbb{Z}}$, $i\in\mathbb{N}^{+}$ is "*weakly dependent*". We do point out that due to the dynamic aspect of our panel data model (2.4), (2.6) does impose some restriction on the rate of divergence of $g_u(n)$ and $g_z(n)$. To see this, suppose for the sake of argument that $\varphi_u(i,j)=\varphi_u(|i-j|)$. In the introduction various examples were given where $\varphi_u(i,j)=\varphi_u(|i-j|)$, i.e., when lattice type of data is available, so that we can "locate" our individuals in some form of equally space distance or when the dependence is related to some "economic/geographical" distance as in Conley (1999). Given $\varphi_u(|i-j|)\simeq|i-j|^{2d_u-1}$ and $\varphi_z(|i-j|)\simeq|i-j|^{2d_z-1}$ with $0<d_u<1/4$ and $0<d_z<1/4$ (so that both u_{it} and z_{it} are "*strong dependent*"), $d_u+d_z<1/2$ in (2.4) which ensures w_{it} is "weakly dependent". However, it could also fit the framework of Jenish and Prucha (2012), who regard observations as lying on an irregularly spaced pattern. It is worth emphasizing that our assumptions do not imply any type of strong-mixing condition as in Jenish and Prucha (2012) as that would require that at least $g_u(n)+g_z(n)<C$ and typically involves the notion of falling off of dependence as $|i-j|$ increases, which is not very relevant to all spatial situations of interest, see Lee and Robinson (2013). In fact, drawing similarities with time series literature, using Ibragimov and Rozanov (1978, *Ch.* 4), it suggests that our condition rules out any form of *weak dependence*, such as strong-mixing, in $\{w_{it}\}_{t\in\mathbb{Z}}$, $i\in N^{+}$. In addition, and keeping in mind our previous comments on the behaviour of $\varphi_u(i,j)$, (2.4) yields that

$$0<\lim_{n\to\infty}\frac{1}{n}\sum_{i,j=1}^{n}\varphi_u^2(i,j)<\infty,$$

so that $u_{it}^2-E\left(u_{it}^2\right)$ behaves as if it were a "*weakly-dependent*" sequence. Finally (2.4) also implies that

$$\max_{1\le i\le n}\left|\sum_{j=1}^{n}\varphi_u(i,j)\right|\left\|\sum_{j=1}^{n}\varphi_z(i,j)\right\|=O\left(n^{1-\zeta}\right) \tag{2.7}$$

for some $\zeta>0$.

Condition (2.6) bears similarities to a condition found in classical time series regression models with possible "*strong-dependence*". There the condition is that

$$\int_{-\pi}^{\pi}f_{u_i}(\lambda)\,f_{z_i}(\vartheta-\lambda)\,d\lambda=f_i(\vartheta)\quad\vartheta\in(-\pi,\pi]$$

is a continuous function at $\vartheta=0$, where $f_{u_i}(\lambda)$ and $f_{z_i}(\lambda)$ denote respectively the spectral density functions of $\{u_{it}\}_{t\in\mathbb{Z}}$ and the regressors $\{z_{it}\}_{t\in\mathbb{Z}}$, see for instance Robinson and Hidalgo (1997) and Hidalgo (2003). We then view (2.4), or (2.6), as the counterpart of the last displayed expression in regression models with cross-sectional dependence.

Finally, Condition $C4$ is very weak as $\xi>0$ effectively means that n has to grow to infinity at least as fast as $T^{-1}=O\left(\log^{-1}n\right)$. Looking at the arguments in the proof, it is possible to relax it. However, this would involve higher order moments requirements for the sequences $\{u_{it}\}_{t\in\mathbb{Z}}$ and/or $\{x_{it}\}_{t\in\mathbb{Z}}$, $i\in\mathbb{N}^{+}$. So, we have preferred the former conditions to the latter, as they already are quite weak. Notice that our Condition $C4$ relaxes significantly the requirement in Pesaran and Yamagata (2008), who assumed that $n^{1/2}/T\to0$ or even $n^{1/4}/T\to0$.

Before presenting our first main result, let us introduce some notation. In what follows, we denote the "*average*" long-run variance as

$$V_1 = \lim_{n\to\infty} \frac{1}{n} \sum_{i,j=1}^{n} \{\varphi_u(i,j)\,\varphi_x(i,j)\}\,. \tag{2.8}$$

For generic sequences $\{\varsigma_{it}\}_{t=1}^{T}$, $i=1,...,n$, we write

$$\begin{aligned} \widetilde{\widetilde{\varsigma}}_{it} &= \varsigma_{it} - \overline{\varsigma}_{\cdot t} - \overline{\varsigma}_{i\cdot} + \overline{\varsigma}_{\cdot\cdot} \qquad (2.9)\\ \text{with } \overline{\varsigma}_{\cdot t} &= \frac{1}{n}\sum_{i=1}^{n}\varsigma_{it};\quad \overline{\varsigma}_{i\cdot} = \frac{1}{T}\sum_{i=1}^{T}\varsigma_{it};\quad \overline{\varsigma}_{\cdot\cdot} = \frac{1}{T}\sum_{i=1}^{T}\varsigma_{\cdot t}. \end{aligned}$$

The transformation in (2.9) allows us to remove the individual and time effects η_i and α_t from the model (1.2). To simplify algebra and notation, we will initially assume that $\eta_i = 0$ in (1.2) in which case the transformation that we need simplifies to

$$\widetilde{\varsigma}_{it} = \varsigma_{it} - \overline{\varsigma}_{\cdot t}. \tag{2.10}$$

That is, we consider

$$\widetilde{y}_{it} = \beta_t' \widetilde{x}_{it} + \widetilde{u}_{it},\ \ i = 1,...,n \ \text{ and } t = 1,...,T. \tag{2.11}$$

It is worth noticing that, in view of $C1$ and $C2$, the transformation (2.10) is such that $E\widetilde{x}_{it} = 0$.

Let $\widehat{\beta}_{FE}$ be the fixed effect estimator of the slope parameters, i.e.

$$\widehat{\beta}_{FE} = \left(\sum_{i=1}^{n}\sum_{t=1}^{T} \widetilde{x}_{it}\widetilde{x}_{it}'\right)^{-1} \left(\sum_{i=1}^{n}\sum_{t=1}^{T} \widetilde{x}_{it}\widetilde{y}_{it}\right), \tag{2.12}$$

and, for all $t \geq 1$, consider

$$\widehat{\beta}_t = \left(\sum_{i=1}^{n} \widetilde{x}_{it}\widetilde{x}_{it}'\right)^{-1} \left(\sum_{i=1}^{n} \widetilde{x}_{it}\widetilde{y}_{it}\right). \tag{2.13}$$

Finally with $\Sigma_x > 0$ as in $C2$, define

$$V_2 = \Sigma_x^{-1} V_1 \Sigma_x^{-1}.$$

We now give our main result of this section.

Theorem 1. *Under Conditions $C1 - C4$ and $\beta_t = \beta$, we have that*

(**a**) $(Tn)^{1/2}\left(\widehat{\beta}_{FE} - \beta\right) \xrightarrow{d} \mathcal{N}(0, V_2)$

(**b**) $n^{1/2}\left(\widehat{\beta}_{t_1} - \beta, ..., \widehat{\beta}_{t_\ell} - \beta\right)' \xrightarrow{d} \mathcal{N}(0, I_\ell \otimes V_2)$ *for any finite $\ell \geq 1$.*

Proof. The proof of this result or any other will be given in the Appendix. □

Remark 1. (**i**) *The estimators $\widehat{\beta}_t$ and $\widehat{\beta}_s$ are asymptotically independent if $s \neq t$. This is the case because* Cov$(u_{it}, u_{js}) = 0$ *for all $s \neq t$ by $C1$.*
(**ii**) *The main conclusions of Theorem 1 hold true when $\eta_i \neq 0$, after using the transformation given in (2.9) to the model.*

(**iii**) *Under the alternative hypothesis, i.e.* $\beta_t \neq \beta$*, we have that Theorem 1 still holds true but with some minor changes. Indeed, when* $\beta_t \neq \beta$*, we can easily extend our arguments to show that*

$$(\mathbf{a}) \quad (Tn)^{1/2}\left(\widehat{\beta}_{FE} - \frac{1}{T}\sum_{t=1}^{T}\beta_t\right) \overset{d}{\to} \mathcal{N}\left(0, V_2 + W\right)$$

$$(\mathbf{b}) \quad n^{1/2}\left(\widehat{\beta}_{t_1} - \beta_{t_1}, ..., \widehat{\beta}_{t_\ell} - \beta_{t_\ell}\right)' \overset{d}{\to} \mathcal{N}\left(0, I_\ell \otimes V_2\right), \quad \textit{for any finite } \ell \geq 1,$$

where

$$W = \Sigma_x^{-1} \lim_{n,T\to\infty} \frac{1}{nT} Var\left(\sum_{i=1}^{n}\sum_{t=1}^{T} x_{it}x_{it}'\left[\beta_t - \frac{1}{T}\sum_{s=1}^{T}\beta_s\right]\right)\Sigma_x^{-1}.$$

So, the results of Theorem 1 only affects the fixed-effect estimator.

Recalling our definition of V_2, Theorem 1 indicates that to provide inferences about the slope parameters, we need a consistent estimator of the "*average*" long-run variance V_1 in (2.8). In our particular setup, we propose the following very simple estimator

$$\widehat{V}_1 = \frac{1}{T}\sum_{t=1}^{T}\left\{\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}\widehat{u}_{it}\right)\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}\widehat{u}_{it}\right)'\right\}, \tag{2.14}$$

where $\widehat{u}_{it} = \widetilde{y}_{it} - \widehat{\beta}_{FE}'\widetilde{x}_{it}$, $i = 1, ..., n$ and $t = 1, ..., T$. $\widehat{V}_1$ has similarities with the so-called cluster estimator, see Arellano (1987) or Bester, Conley and Hansen (2011). It is worth remarking that in (2.14) we cannot employ $\widehat{u}_{it} = \widetilde{y}_{it} - \widehat{\beta}_t'\widetilde{x}_{it}$, as $\sum_{i=1}^{n}\widetilde{x}_{it}\widehat{u}_{it} = 0$ by definition. One important feature of the above estimator is that, contrary to the HAC estimators of Kelejian and Prucha (2007) or Kim and Sun (2013), there is no need to introduce any artificial "metric" among observations. It is not clear that this would be convenient, as changing the "metric" may yield a different estimate of V_1 and thereby induce potentially different outcomes in our inferences.

Proposition 1. *Under the same conditions of Theorem 1, we have that*

$$\widehat{V}_1 - V_1 = o_p(1).$$

We now make some comments on Proposition 1. When $\beta_t \neq \beta$, the results in Proposition 1 does not hold true. The reason being that in this case $\widehat{\beta}_{FE}$ would only be a consistent estimator of $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\beta_t$ as the remark that follows Theorem 1 indicates. There is a second way to obtain a consistent estimator of V_1 via bootstrap methods. Recall that this approach was one of the main motivations for the bootstrap in the original paper by Efron (1979) as a method to estimate the asymptotic covariance of estimators when they are not easy to compute or to provide an explicit formula. We will delay discussing this approach to Section 4 below.

3. TESTS FOR BREAKS

For sake of simplicity, we shall first consider the case where $\eta_i = 0$. In Section 3.2 we discuss the setting when $\eta_i \neq 0$.

3.1. TESTING FOR BREAKS WHEN $\eta_i = 0$.

We now introduce two related tests for breaks of the slope parameters in our model (1.2). Our first approach to test H_0 in (1.3), a *CUSUM* type test, is based on the behaviour of

$$\mathcal{T}(r) = \frac{1}{(nT)^{1/2}} \sum_{t=1}^{r} \sum_{i=1}^{n} \widetilde{x}_{it} \left(\widetilde{y}_{it} - \widehat{\beta}'_{FE} \widetilde{x}_{it} \right), \quad r = 1, ..., T-1. \tag{3.1}$$

The intuition for $\mathcal{T}(r)$ is that under the null hypothesis, we expect that $\widetilde{x}_{it} \left(\widetilde{y}_{it} - \widehat{\beta}'_{FE} \widetilde{x}_{it} \right) \simeq \widetilde{x}_{it} u_{it}$ which has a mean equal to zero, while under the alternative hypothesis we have that $\widetilde{x}_{it} \left(\widetilde{y}_{it} - \widehat{\beta}'_{FE} \widetilde{x}_{it} \right)$ will develop a term of the type

$$\widetilde{x}_{it} \widetilde{x}'_{it} \left(\beta_t - \widehat{\beta}_{FE} \right) \simeq \widetilde{x}_{it} \widetilde{x}'_{it} \left(\beta_t - \frac{1}{T} \sum_{s=1}^{T} \beta_s \right), \tag{3.2}$$

see Remark 1 or Theorem 1. Under the alternative therefore, $\mathcal{T}(r)$ would be governed by the non-zero function

$$h(r) = \left\{ \frac{1}{n} \sum_{i=1}^{n} E\left(\mathring{x}_{it} \mathring{x}'_{it} \right) \right\} \frac{n^{1/2}}{T^{1/2}} \sum_{t=1}^{r} \left(\beta_t - \frac{1}{T} \sum_{s=1}^{T} \beta_s \right),$$

where for generic sequences $\{\varsigma_{it}\}_{t \in \mathbb{Z}}$, $i \in \mathbb{N}^+$, we denote

$$\{\mathring{\varsigma}_{it}\}_{t \in \mathbb{Z}} = \{\varsigma_{it} - E(\varsigma_{it})\}_{t \in \mathbb{Z}}, \quad i \in \mathbb{N}^+.$$

The preceding arguments suggest that one possible method to test the null hypothesis in (1.3) might be based on continuous functionals of $\mathcal{T}(r)$.

Our second approach is based on the observation that we can regard H_0 as testing whether the slope parameters β_t are the same across time, where for a given time period t, we can estimate β_t as in (2.13). This test recognizes that under H_0, we can use the *mean group (MG) estimator*

$$\widetilde{\beta}_{FE} = \frac{1}{T} \sum_{s=1}^{T} \widehat{\beta}_s, \tag{3.3}$$

see Pesaran, Shin and Smith (1999), as an estimator for the common slope parameters β. While under the null, for every t, $\widehat{\beta}_t - \widetilde{\beta}_{FE}$ converges to zero in probability, under the alternative hypothesis $\widehat{\beta}_t - \widetilde{\beta}_{FE}$ will develop a mean different than zero. Our Hausman-Durbin-Wu's type of statistic then, is based on continuous functionals of

$$\mathcal{T}^{\Delta}(r) = \frac{n^{1/2}}{T^{1/2}} \sum_{t=1}^{r} \left(\widehat{\beta}_t - \widetilde{\beta}_{FE} \right). \tag{3.4}$$

It is worth noticing that test based on $\mathcal{T}(r)$ and $\mathcal{T}^{\Delta}(r)$ are related. Indeed, using the definition of $\widehat{\beta}_{FE}$, we easily obtain

$$\begin{aligned}\mathcal{T}(r) &\equiv \frac{1}{(nT)^{1/2}}\sum_{t=1}^{r}\sum_{i=1}^{n}\widetilde{x}_{it}\left(\widetilde{y}_{it}-\widehat{\beta}'_{FE}\widetilde{x}_{it}\right)\\ &\simeq \frac{n^{1/2}}{T^{1/2}}\left(\sum_{t=1}^{r}\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}\widetilde{x}'_{it}\right)\widehat{\beta}_t-\frac{r}{T}\sum_{s=1}^{T}\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{is}\widetilde{x}'_{is}\right)\widehat{\beta}_s\right)\\ &= \Sigma_x\mathcal{T}^{\Delta}(r)\left(1+o_p(1)\right),\end{aligned}$$

so that $\mathcal{T}^{\Delta}(r)$ is a "weighted" version of $\mathcal{T}(r)$ for any r. We point out that our tests have similarities with the Δ test in Pesaran and Yamagata (2008), see also Swamy (1970). However, as we will notice in Section 3.3 below, tests based on (3.1) or (3.4) can detect local alternatives which the Δ test cannot.

Let $\mathcal{B}(\tau)$ denote the standard Brownian motion in $[0,1]$ and $\mathcal{BB}(\tau)=\mathcal{B}(\tau)-\tau\mathcal{B}(1)$ the standard Brownian bridge.

Theorem 2. *Assuming $C1-C4$, under H_0, we have that as $n,T\to\infty$,*

$$(\mathbf{a})\quad \frac{1}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\widetilde{x}_{it}\left(\widetilde{y}_{it}-\widehat{\beta}'_{FE}\widetilde{x}_{it}\right)\overset{d}{\Longrightarrow}V_1^{1/2}\mathcal{BB}(\tau)$$

$$(\mathbf{b})\quad \frac{n^{1/2}}{T^{1/2}}\sum_{t=1}^{[T\tau]}\left(\widehat{\beta}_t-\widetilde{\beta}_{FE}\right)\overset{d}{\Longrightarrow}V_2^{1/2}\mathcal{BB}(\tau)\,,\quad \textit{if } T=o\left(n^2\right).$$

It is worth mentioning that the condition $T=o\left(n^2\right)$ in part (**b**) can be relaxed at the expense of higher order moments and a lengthening of the proof of the theorem. We have not pursued this route as the constraint on T and n does not seem to be too stringent and in most of the empirical examples seems to be satisfied.

For any continuous mapping function $\varphi(\cdot)$, our tests are given by

$$\mathcal{T}=\varphi\left(\frac{\mathcal{T}'(r)\,\widehat{V}_1^{-1}\mathcal{T}(r)}{w^2(r/T)}\right)\quad\text{and}\quad \mathcal{T}^{\Delta}=\varphi\left(\frac{\mathcal{T}^{\Delta\prime}(r)\,\widehat{V}_2^{-1}\mathcal{T}^{\Delta}(r)}{w^2(r/T)}\right),\tag{3.5}$$

where $w(\tau)$, $\tau\in[0,1]$, is a weighting function that (**i**) is non-decreasing in a neighbourhood of 0, (**ii**) is non-increasing in a neighbourhood of 1, (**iii**) is positive on $(\eta,1-\eta)$ and (**iv**) satisfies

$$\int_0^1\frac{1}{\tau(1-\tau)}\exp\left(-c\frac{w^2(\tau)}{\tau(1-\tau)}\right)d\tau<\infty.\tag{3.6}$$

A standard weighting $w(\tau)$ function which satisfies these conditions is $w(\tau)=1$. The common choice $w(\tau)=\tau^{1/2}(1-\tau)^{1/2}$, implicitly used in Andrews (1993) and many subsequent authors, on the other hand, fails to satisfy this condition (3.6). While the latter weight function provides a natural standardization of our test, as it represents the standard deviation of a standard Brownian Bridge, it does have the drawback of requiring trimming for values of τ close to 0 and 1. In fact, any weighting function that does not satisfy (3.6) is subject to the use of some trimming for values to close to 0 or to 1, which is a well known result, see for instance Shorack and Wellner (2009, *p*.462).

We then have the following result.

Corollary 1. *Assuming $C1-C4$, under H_0 and $w(\tau)$ satisfying (3.6) as $n,T\to\infty$, we have that*

$$\textbf{(a)}\quad \mathcal{T}\overset{d}{\Longrightarrow}\varphi\left(\frac{\left(\mathcal{BB}(\tau)\right)'\left(\mathcal{BB}(\tau)\right)}{w^2(\tau)}\right)$$

$$\textbf{(b)}\quad \mathcal{T}^{\Delta}\overset{d}{\Longrightarrow}\varphi\left(\frac{\left(\mathcal{BB}(\tau)\right)'\left(\mathcal{BB}(\tau)\right)}{w^2(\tau)}\right)\ \text{ if } T=o\left(n^2\right).$$

Proof. The proof of this corollary follows easily by Proposition 1 and Theorem 2. Indeed Proposition 1 indicates that

$$\frac{\widehat{V}_1^{-1/2}}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\widetilde{x}_{it}\left(\widetilde{y}_{it}-\widehat{\beta}_{FE}'\widetilde{x}_{it}\right)=\frac{V_1^{-1/2}}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}\left(\mathring{y}_{it}-\widehat{\beta}_{FE}'\mathring{x}_{it}\right)(1+o_p(1))$$

$$\widehat{V}_1^{-1/2}\frac{n^{1/2}}{T^{1/2}}\sum_{t=1}^{[T\tau]}\left(\widehat{\beta}_t-\widetilde{\beta}_{FE}\right)=V_1^{-1/2}\frac{n^{1/2}}{T^{1/2}}\sum_{t=1}^{[T\tau]}\left(\widehat{\beta}_t-\widetilde{\beta}_{FE}\right)(1+o_p(1)).$$

From here, Theorem 2 and the continuous mapping theorem yield the conclusion of the corollary. □

Corollary 1 indicates that when $w(\tau)=1$, we have

$$\max_{0<r<T}\left|\mathcal{T}(r)'\,\widehat{V}_1^{-1}\mathcal{T}(r)\right|\overset{d}{\Longrightarrow}\max_{0<\tau<1}\left|\left(\mathcal{BB}(\tau)\right)'\left(\mathcal{BB}(\tau)\right)\right|$$

$$\max_{0<r<T}\left|\mathcal{T}^{\Delta}(r)'\,\widehat{V}_2^{-1}\mathcal{T}^{\Delta}(r)\right|\overset{d}{\Longrightarrow}\max_{0<\tau<1}\left|\left(\mathcal{BB}(\tau)\right)'\left(\mathcal{BB}(\tau)\right)\right|,\ \text{ if } T=o\left(n^2\right),$$

which correspond to a Kolmogorov-Smirnov's type of statistic. However when $w^2(\tau)=\tau(1-\tau)$, which corresponds to the weight function implicit in Andrews (1993), (3.6) is not satisfied so that as in Andrews (1993) we trim for values close to the boundary, that is consider

$$\max_{[T\epsilon]<r<T-[T\epsilon]}\left|\frac{\mathcal{T}(r)'\,\widehat{V}_1^{-1}\mathcal{T}(r)}{w^2(r/T)}\right|\overset{d}{\Longrightarrow}\max_{\epsilon<\tau<1-\epsilon}\left|\frac{\left(\mathcal{BB}(\tau)\right)'\left(\mathcal{BB}(\tau)\right)}{w^2(\tau)}\right|$$

$$\max_{[T\epsilon]<r<T-[T\epsilon]}\left|\frac{\mathcal{T}^{\Delta}(r)'\,\widehat{V}_2^{-1}\mathcal{T}^{\Delta}(r)}{w^2(r/T)}\right|\overset{d}{\Longrightarrow}\max_{\epsilon<\tau<1-\epsilon}\left|\frac{\left(\mathcal{BB}(\tau)\right)'\left(\mathcal{BB}(\tau)\right)}{w^2(\tau)}\right|,$$

for some $0<\epsilon<\frac{1}{2}$.

Of course, we can use other weighting functions $w(\tau)$ to target particular alternatives in a similar way as directional tests do in goodness-of-fit tests, see also Andrews and Ploberger (1994). We have not pursued this somewhat standard extension.

Neither have we pursued the scenario put forward in the introduction of $\epsilon\to 0$, as in Hidalgo and Seo (2013) who basically examine the consequences when $w(\tau)$ fails the condition in (3.6) and no trimming is used. Bear in mind, the purpose of trimming and the introduction of a weight function satisfying (3.6) is somehow to make $\max_{r<[T\epsilon]}$ or $\max_{T-[T\epsilon]<r}$ asymptotically negligible, as the asymptotic distribution becomes a Gumbel distribution when the latter is not true, see also Horváth (1993).

3.2. TESTING FOR BREAKS WHEN $\eta_i \neq 0$.

We shall now extend our previous tests for homogeneity of the slope parameters in our model (1.2) to the setting when $\eta_i \neq 0$. For notational simplicity, we have decided to assume that $\alpha_t = 0$. That is, we consider

$$y_{it} = \eta_i + \beta_t' x_{it} + u_{it}, \quad i = 1, ..., n, \quad t = 1, ..., T.$$

As we did in the previous section, we first remove the individual fixed effects η_i, $i \in \mathbb{N}^+$ before estimating our parameters β_t. To that end, using (2.9), and denoting for generic sequences $\{\varsigma_{it}\}_{t=1}^T$, $i = 1, ..., n$,

$$\varsigma_{it}^\dagger = \varsigma_{it} - \overline{\varsigma}_{i\cdot} \tag{3.7}$$

we have

$$y_{it}^\dagger = \beta_t' x_{it} - \frac{1}{T}\sum_{s=1}^T \beta_s' x_{is} + u_{it}^\dagger. \tag{3.8}$$

The main difference between (3.8) and the model examined in Section 3.1 is that under the alternative hypothesis the "standard" transformed regressor $x_{it}^\dagger$ no longer appears. Of course, under the null hypothesis (1.3), it does and (3.8) becomes

$$y_{it}^\dagger = \beta' x_{it}^\dagger + u_{it}^\dagger, \quad i = 1, ..., n, \quad t = 1, ..., T.$$

Denoting by

$$\widehat{\beta}_{FE} = \left(\sum_{i=1}^n \sum_{t=1}^T x_{it}^\dagger x_{it}^{\dagger\prime}\right)^{-1} \left(\sum_{i=1}^n \sum_{t=1}^T x_{it}^\dagger y_{it}^\dagger\right),$$

the estimator of the (common) slope parameters under the null hypothesis, our *CUSUM* based test will now look at the behaviour of

$$\mathcal{T}^\dagger(r) = \frac{1}{(nT)^{1/2}} \sum_{t=1}^r \sum_{i=1}^n x_{it}^\dagger \left(y_{it}^\dagger - \widehat{\beta}_{FE}' x_{it}^\dagger\right), \quad r = 1, ..., T-1. \tag{3.9}$$

Similarly, denoting by

$$\widehat{\beta}_s = \left(\sum_{i=1}^n x_{is}^\dagger x_{is}^{\dagger\prime}\right)^{-1} \left(\sum_{i=1}^n x_{is}^\dagger y_{is}^\dagger\right)$$

the estimator of the time specific slope parameters and the associated *mean group (MG)* estimator by $\widetilde{\beta}_{FE} = \frac{1}{T}\sum_{s=1}^T \widehat{\beta}_s$, our Hausman-Durbin-Wu's type of statistic will look at the behaviour of

$$\mathcal{T}^{\Delta\dagger}(r) = \frac{n^{1/2}}{T^{1/2}} \sum_{t=1}^r \left(\widehat{\beta}_t - \widetilde{\beta}_{FE}\right) \quad \text{for } r \geq 1 \tag{3.10}$$

Before we describe the statistical properties of $\mathcal{T}^\dagger(r)$ and $\mathcal{T}^{\Delta\dagger}(r)$, we shall examine the asymptotic behaviour of $\widehat{\beta}_{FE}$ and $\widehat{\beta}_s$.

Proposition 2. *Under Conditions $C1 - C4$ and $\beta_t = \beta$, we have that*

(a) $(Tn)^{1/2}\left(\widehat{\beta}_{FE} - \beta\right) \xrightarrow{d} \mathcal{N}(0, V_2)$

(b) $n^{1/2}\left(\widehat{\beta}_{t_1} - \beta, ..., \widehat{\beta}_{t_\ell} - \beta\right)' \xrightarrow{d} \mathcal{N}(0, I_\ell \otimes V_2)$ *for any finite $\ell \geq 1$.*

Proof. The proof proceeds similarly as that of Theorem 1 and is therefore omitted. □

We now give the main result of this section.

Theorem 3. *Assuming $C1-C4$, under H_0 and $w(\tau)$ satisfying (3.6), as $n, T \to \infty$, we have that*

$$\textbf{(a)} \quad \mathcal{T}^{\dagger} = \varphi\left(\frac{\mathcal{T}^{\dagger}(r)' \widehat{V}_2^{-1} \mathcal{T}^{\dagger}(r)}{w^2(r/T)}\right) \overset{d}{\Longrightarrow} \varphi\left(\frac{(\mathcal{BB}(\tau))'(\mathcal{BB}(\tau))}{w^2(\tau)}\right),$$

$$\textbf{(b)} \quad \mathcal{T}^{\Delta\dagger} = \varphi\left(\frac{\mathcal{T}^{\Delta\dagger}(r)' \widehat{V}_2^{-1} \mathcal{T}^{\Delta\dagger}(r)}{w^2(r/T)}\right) \overset{d}{\Longrightarrow} \varphi\left(\frac{(\mathcal{BB}(\tau))'(\mathcal{BB}(\tau))}{w^2(\tau)}\right), \quad \textit{if } T = o\left(n^2\right).$$

Proof. The proof proceeds as that of Theorem 2 by recognizing that by Proposition 1

$$\frac{\widehat{V}_1^{-1/2}}{(nT)^{1/2}} \sum_{t=1}^{[T\tau]} \sum_{i=1}^{n} x_{it}^{\dagger}\left(y_{it}^{\dagger} - \widehat{\beta}_{FE}' x_{it}^{\dagger}\right) = \frac{V_1^{-1/2}}{(nT)^{1/2}} \sum_{t=1}^{[T\tau]} \sum_{i=1}^{n} \mathring{x}_{it}\left(\mathring{y}_{it} - \widehat{\beta}_{FE}' \mathring{x}_{it}\right)(1 + o_p(1))$$

$$\widehat{V}_1^{-1/2} \frac{n^{1/2}}{T^{1/2}} \sum_{t=1}^{[T\tau]} \left(\widehat{\beta}_t - \widetilde{\beta}_{FE}\right) = V_1^{-1/2} \frac{n^{1/2}}{T^{1/2}} \sum_{t=1}^{[T\tau]} \left(\widehat{\beta}_t - \widetilde{\beta}_{FE}\right)(1 + o_p(1)).$$

Application of Proposition 2 and the continuous mapping theorem yield the desired result. □

3.3. LOCAL ALTERNATIVES AND CONSISTENCY OF THE TESTS.

We now discuss the local alternatives for which the tests described in the previous two sections have nontrivial power and from there easily conclude their consistency. To that end, we begin by considering the local alternatives

$$H_a: \quad \beta_t = \beta + \delta_{nT} \mathcal{I}(t > t_0), \tag{3.11}$$

where $t_0 = [T\tau_0]$ for some $\tau_0 \in (\epsilon, 1-\epsilon)$ with $\epsilon > 0$, and δ_{nT} is a deterministic sequence depending on n and/or T. To shorten the discussion we will only explicitly handle the behaviour under H_a and discuss the consistency of tests based on $\mathcal{T}^{\dagger}(r)$ and $\mathcal{T}^{\Delta\dagger}(r)$ in (3.9) and (3.10), respectively. The conclusions for $\mathcal{T}(r)$ and $\mathcal{T}^{\Delta}(r)$ are qualitatively the same and are handled similarly.

For this purpose, introduce the "shift" function

$$\Xi(\tau) = (\tau - \tau_0)\mathcal{I}(\tau > \tau_0) - \tau(1 - \tau_0). \tag{3.12}$$

We then establish the following result.

Proposition 3. *Assuming $C1-C4$, under H_a with $\delta_{nT} = \delta/(nT)^{1/2}$, $|\delta| > 0$, we have that as $n, T \to \infty$,*

$$\textbf{(a)} \quad \frac{1}{(nT)^{1/2}} \sum_{t=1}^{[T\tau]} \sum_{i=1}^{n} x_{it}^{\dagger}\left(y_{it}^{\dagger} - \widehat{\beta}_{FE}' x_{it}^{\dagger}\right) \overset{d}{\Longrightarrow} V_1^{1/2} \mathcal{BB}(\tau) + \delta \Sigma_x \Xi(\tau)$$

$$\textbf{(b)} \quad \frac{n^{1/2}}{T^{1/2}} \sum_{t=1}^{[T\tau]} \left(\widehat{\beta}_t - \widetilde{\beta}_{FE}\right) \overset{d}{\Longrightarrow} V_2^{1/2} \mathcal{BB}(\tau) + \delta \Xi(\tau), \textit{ if } T = o\left(n^2\right).$$

Proposition 3 indicates that the tests have no trivial power if the alternative hypothesis converges to the null at the rate $(nT)^{1/2}$. On the other hand, when $\delta_{nT}^{-1} = o\left((nT)^{1/2}\right)$, the statistic diverges to infinity, that is the test will reject with probability 1 as the sample size increases.

Finally, when $\delta_{nT} = o\left((nT)^{-1/2}\right)$, the asymptotic distribution is identical to that obtained under the null hypothesis. This clearly improves on the local alternatives given in Pesaran and Yamagata (2008), who only were able to detect local alternatives $\delta_{nT} = O\left(n^{-1/4}T^{-1/2}\right)$. In this way, their test has zero asymptotic relative efficiency compared to ours.

The consistency of the test is given in the following corollary.

Corollary 2. *Assuming $C1 - C4$, under H_a with $\delta_{nT} = \delta$ for all n and T, we have that*

$$(\mathbf{a}) \quad \Pr\left\{\varphi\left(\frac{\mathcal{T}(r)' \widehat{V}_1^{-1} \mathcal{T}(r)}{w^2(r/T)}\right) > a\right\} \to 1$$

$$(\mathbf{b}) \quad \Pr\left\{\varphi\left(\frac{\mathcal{T}^{\Delta}(r)' \widehat{V}_2^{-1} \mathcal{T}^{\Delta}(r)}{w^2(r/T)}\right) > a\right\} \to 1 \quad \text{if } T = o\left(n^2\right)$$

for any $a > 0$ and continuous $w(\tau)$.

Proof. The proof is standard from Proposition 3, so it is omitted. □

Remark 2. (**i**) *It is important to mention that we have not assumed that $w(\tau)$ satisfies* (3.6) *on purpose. The reason is that under the alternative hypothesis we have assumed $\tau_0 \in (\epsilon, 1-\epsilon)$ for some $\epsilon > 0$. Of course, if $w(\tau)$ would satisfy* (3.6)*, we then could take $\epsilon = 0$. However, we do not want to lengthen the paper with this unnecessary and rather trivial discussion.*

(**ii**) *Our main conclusion in this section does not depend on the fact that the break or heterogeneity of the slope parameters is abrupt in nature. Indeed, suppose that we replace H_a in* (3.11) *by the following alternative hypothesis*

$$H_a: \quad \beta_t = \beta + \frac{1}{(nT)^{1/2}}\left\{\sum_{\ell=1}^{L} \delta_\ell \mathcal{I}(t > t_\ell) + \delta\left(\frac{t}{T}\right)\right\},$$

where $\delta(\tau)$ is a continuous (smooth) function in $\tau \in (0,1)$ while $|\delta_\ell| > 0$, $\ell = 1,..,L$ permits discrete jumps. The only difference lies in the form of the shift function $\Xi(\tau)$ appearing in (3.12)*. Indeed, with the (local) alternatives given in the last displayed expression, the shift function $\Xi(\tau)$ becomes*

$$\Xi(\tau) = \sum_{\ell=0}^{L} \delta_\ell \mathcal{I}(\tau > \tau_\ell) - \tau \sum_{\ell=1}^{L} \delta_\ell (1 - \tau_\ell) + \int_0^\tau \delta(v)\, dv - \tau \int_0^1 \delta(v)\, dv.$$

It is clear that $\Xi(\tau)$ is different from zero in a set $\Lambda \subset [0,1]$ with positive Lebesgue measure. Indeed, suppose for simplicity that $\delta_\ell = 0$ for all $\ell \geq 0$, then

$$\Xi(\tau) = \int_0^\tau \delta(v)\, dv - \tau \int_0^1 \delta(v)\, dv.$$

In that case $\Xi(\tau) = 0$ for all $\tau \in (0,1)$ if and only if $\delta(\tau)$ is a constant function which is ruled out as it would imply that $H_a \subset H_0$. To see this, we notice that $\Xi(\tau) = \int_0^\tau \left\{\delta(v) - \overline{\delta}\right\} dv$, where $\overline{\delta} = \int_0^1 \delta(v)\, dv$. But $\Xi(\tau) = 0$ for all $\tau \in (0,1)$ if and only if $\delta(v) = \overline{\delta}$ for all $v \in (0,1)$.

We finish the section commenting on the power of the tests in the situations mentioned in the introduction, namely (**i**) when the time of the break is towards the end of the sample and (**ii**) when the number of individuals for which a break exists is negligible compared to n.

We first consider (**i**). Assume that $\beta_t = \beta$ if $t \leq T - T_0$ (T_0 small) and $\beta_t = \beta + \delta$ otherwise. Consider the decomposition

$$\begin{aligned} & \frac{1}{T}\sum_{t=1}^{\widetilde{T}}\left(\widehat{\beta}_t - \widetilde{\beta}_{FE}\right) \\ = \; & \frac{1}{T}\sum_{t=1}^{\widetilde{T}}\left(\beta_t - \frac{1}{T}\sum_{s=1}^{T}\beta_s\right) + \frac{1}{T}\sum_{t=1}^{\widetilde{T}}\left\{\left(\widehat{\beta}_t - \beta_t\right) - \left(\widetilde{\beta}_{FE} - \frac{1}{T}\sum_{s=1}^{T}\beta_s\right)\right\}. \end{aligned} \tag{3.13}$$

The second term on the right of (3.13) is $O\left((nT)^{1/2}\right)$, whereas the first term equals

$$\begin{cases} -\delta\frac{\widetilde{T}}{T}\frac{T_0}{T} & \text{if } \widetilde{T} < T - T_0 \\ -\delta\left(\frac{T-T_0}{T}\right)\left(\frac{T-\widetilde{T}}{T}\right) & \text{if } \widetilde{T} \geq T - T_0. \end{cases}$$

So, we have that

$$\frac{n^{1/2}}{T^{1/2}}\sum_{t=1}^{[T\tau]}\left(\widehat{\beta}_t - \widetilde{\beta}_{FE}\right) = O_p(1) - \begin{cases} \frac{n^{1/2}T_0}{T^{1/2}}\delta\tau & \text{if } [T\tau] < T - T_0 \\ \frac{n^{1/2}(T-T_0)}{T^{1/2}}\delta(1-\tau) & \text{if } [T\tau] \geq T - T_0, \end{cases}$$

implying that tests based on $\mathcal{T}^{\Delta}(r)$ will diverge and hence be consistent if $C^{-1} < \frac{n^{1/2}T_0}{T^{1/2}}$ for some positive finite constant C. The same conclusions are drawn regarding tests based on $T(r)$. We point out here that when $n = 1$, the condition for consistency, i.e., that T_0 does not grow slower than $T^{1/2}$, corresponds to the result obtained for the LM_τ test in Hidalgo and Seo (2013).

Next we consider the situation (**ii**). Suppose for sake of argument that the break occurs at $\tau_0 = 1/2$, and that it only occurs for the first $\iota(n)$ individuals with the condition that $\iota(n) = o(n)$. Again we examine the behaviour of $\mathcal{T}^{\Delta}(r)$. After standard algebra, we have that

$$\widehat{\beta}_t = O_p(1) + \begin{cases} \beta & \text{if } t < T/2 \\ \beta + \delta\frac{\iota(n)}{n} & \text{if } t \geq T/2. \end{cases}$$

So, we obtain that

$$\frac{n^{1/2}}{T^{1/2}}\sum_{t=1}^{[T\tau]}\left(\widehat{\beta}_t - \widetilde{\beta}_{FE}\right) = O_p(1) - \frac{1}{2}\begin{cases} \delta\frac{\iota(n)[T\tau]}{n^{1/2}T^{1/2}} & \text{if } [T\tau] < \frac{1}{2}T \\ \delta\frac{\iota(n)(T-[T\tau])}{n^{1/2}T^{1/2}} & \text{if } [T\tau] \geq \frac{1}{2}T, \end{cases}$$

which implies that test based on $\mathcal{T}^{\Delta}(r)$ will diverge and hence be consistent if $C^{-1} < T^{1/2}\iota(n)/n^{1/2}$.

4. BOOTSTRAP ALGORITHM

One of our motivations for introducing a bootstrap algorithm for our tests (and estimators) is that our tests suffer small sample biases which in some cases, as supported by our Monte Carlo experiments, can be quite substantial. Among other reasons, these biases may be due to the fact that the asymptotic distribution yields a poor approximation in finite samples given our estimator of the long run variance V_1. In such situations the bootstrap approach can, as is well known, provide a tool to improve its finite sample behaviour. A quick glance at our conditions in Section 2, may suggest that a bootstrap mechanism may not be easy to implement (let alone to establish its validity) since one of the basic requirements for its validity is that the bootstrap algorithm should preserve the covariance structure. Drawing analogies with the time series literature, one may be tempted to use the block bootstrap principle. However, since there is no obvious ordering of the data in the cross-sectional dimension, it is not clear that a block bootstrap would work in our context or what its sensitivity would be to a particular chosen ordering of the data (over and above the problem of how to choose the block size). Instead, we propose here a valid bootstrap algorithm with the interesting, and surprising, feature that it is computationally simple, mainly due to the observation that there is no need to estimate, either by parametric or nonparametric methods, the cross-sectional dependence of the error term. Moreover the bootstrap has the additional attractive feature that we do not need to choose any tuning parameter for its implementation, as would be the case with a moving block bootstrap type of bootstrap.

More specifically, we provide two bootstrap algorithms. The first bootstrap procedure is described in the following *4 STEPS.*

***STEP 1*:** We compute the residuals $\{\widehat{u}_{it}\}_{t=1}^{T}$, $i=1,...n$, as

$$\widehat{u}_{it}=\widetilde{y}_{it}-\sum_{\ell=1}^{k_1}\widehat{\rho}_{t\ell}\widetilde{y}_{i(t-\ell)}-\widehat{\theta}_t'\widetilde{z}_{it},\quad i=1,...,n;\quad t=1,...,T$$

and obtain the centered residuals as

$$\check{u}_{it}=\widehat{u}_{it}-\frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}. \tag{4.1}$$

Remark 3. *The motivation to employ* (4.1) *to center the residuals will become apparent when looking at the next* STEP 2*.*

***STEP 2*:** Denoting $\check{U}_t=\{\check{u}_{it}\}_{i=1}^{n}$, we do standard random resampling from the empirical distribution of $\{\check{U}_t\}_{t=1}^{T}$. The bootstrap sample is denoted by $\{U_t^*\}_{t=1}^{T}$.

***STEP 3*:** Generate the bootstrap dynamic panel data model as

$$\widetilde{y}_{it}^*=\sum_{\ell=1}^{k_1}\widetilde{\rho}_{FE,\ell}\widetilde{y}_{i(t-\ell)}+\widetilde{\theta}_{FE}'\widetilde{z}_{it}+u_{it}^*,\quad i=1,...,n,\qquad t=1,...,T, \tag{4.2}$$

where $\widetilde{\rho}_\ell$, $\ell=1,...,k_1$, and $\widetilde{\theta}$ are the *MG estimators* in (3.3).

***STEP 4*:** Compute the test statistics using model (4.2) as if it were the true panel regression model. That is, for $r=1,...,T-1$,

$$\begin{aligned}\mathcal{T}^*(r) &= \frac{1}{(nT)^{1/2}}\sum_{t=1}^{r}\sum_{i=1}^{n}\widetilde{x}_{it}\left(\widetilde{y}_{it}^*-\widehat{\beta}_{FE}'^*\widetilde{x}_{it}\right)\\ \mathcal{T}^{\Delta *}(r) &= \left(\frac{n}{T}\right)^{1/2}\sum_{t=1}^{r}\left(\widehat{\beta}_t^*-\widetilde{\beta}_{FE}^*\right).\end{aligned}$$

In the latter step, $\widehat{\beta}_{FE}^*$ denotes the fixed effect estimator of the slope parameters β, i.e.

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

and $\widehat{\beta}_t^* = \left(\sum_{i=1}^n \widetilde{x}_{it}\widetilde{x}_{it}'\right)^{-1}\left(\sum_{i=1}^n \widetilde{x}_{it}\widetilde{y}_{it}^*\right)$ with $\widetilde{\beta}_{FE}^*$ denoting the *MG bootstrap estimator*

$$\widetilde{\beta}_{FE}^* = \frac{1}{T}\sum_{s=1}^T \widehat{\beta}_s^*.$$

Before establishing the validity of the bootstrap tests $\mathcal{T}^*$ and $\mathcal{T}^{\Delta *}$ (defined below), we establish the following results.

Theorem 4. *Assuming Conditions $C1 - C4$, we have that (in probability)*

$$\textbf{(a)} \quad (Tn)^{1/2}\left(\widehat{\beta}_{FE}^* - \widetilde{\beta}_{FE}\right) \overset{d^*}{\to} \mathcal{N}\left(0, V_2\right)$$

$$\textbf{(b)} \quad n^{1/2}\left(\widehat{\beta}_{t_1}^* - \widetilde{\beta}_{FE}, ..., \widehat{\beta}_{t_\ell}^* - \widetilde{\beta}_{FE}\right)' \overset{d^*}{\to} \mathcal{N}\left(0, I_\ell \otimes V_2\right) \quad \textit{for any finite } \ell \geq 1.$$

Recalling that $V_2 = \Sigma_x^{-1}V_1\Sigma_x^{-1}$, a consistent bootstrap estimator of the "*average*" long-run variance V_1, is given by

$$\widehat{V}_1^* = \frac{1}{T}\sum_{t=1}^T\left(\frac{1}{n^{1/2}}\sum_{i=1}^n \widetilde{x}_{it}\widehat{u}_{it}^*\right)\left(\frac{1}{n^{1/2}}\sum_{i=1}^n \widetilde{x}_{it}\widehat{u}_{it}^*\right)', \tag{4.3}$$

and $\widehat{u}_{it}^* = \widetilde{y}_{it}^* - \widehat{\beta}_{FE}^{*\prime}\widetilde{x}_{it}$, $i = 1, ..., n$ and $t = 1, ..., T$, as the next proposition establishes.

Proposition 4. *Assuming $C1 - C4$, we have that*

$$\widehat{V}_1^* - V_1 = o_{p^*}(1).$$

We now give the validity of our bootstrap test. To save space we shall only consider it explicitly when $\eta_i = 0$, with the general situation when $\eta_i \neq 0$ handled similarly.

Theorem 5. *Assuming $C1 - C4$ and $w(\tau)$ satisfying* (3.6)*, we have that as $n, T \to \infty$, in probability*

$$\textbf{(a)} \quad \mathcal{T}^* = \varphi\left(\frac{\mathcal{T}^*(r)'\left(\widehat{V}_1^*\right)^{-1}\mathcal{T}^*(r)}{w^2(r/T)}\right) \overset{d^*}{\Longrightarrow} \varphi\left(\frac{(\mathcal{BB}(\tau))'(\mathcal{BB}(\tau))}{w^2(\tau)}\right)$$

$$\textbf{(b)} \quad \mathcal{T}^{\Delta *} = \varphi\left(\frac{\mathcal{T}^{\Delta *}(r)'\left(\widehat{V}_2^*\right)^{-1}\mathcal{T}^{\Delta *}(r)}{w^2(r/T)}\right) \overset{d^*}{\Longrightarrow} \varphi\left(\frac{(\mathcal{BB}(\tau))'(\mathcal{BB}(\tau))}{w^2(\tau)}\right) \qquad \textit{if } T = o\left(n^2\right),$$

where $\varphi(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous functional.

While the first bootstrap algorithm is given under $C1$ with $E\left[v_{it}^2 \mid \mathcal{V}_{i,t-1}\right] = \sigma^2$, the second allows $E\left[v_{it}^2 \mid \mathcal{V}_{i,t-1}\right] = \sigma_t^2$, $i \in \mathbb{N}^+$. While a rigorous proof of the validity of the next bootstrap algorithm in the presence of conditional heteroscedasticity is beyond the scope of this paper, its validity under $C1$ can be proven quite similarly and has therefore been left out. The second bootstrap algorithm is described in the next *4 STEPS*.

STEP $1'$**:** We compute the residuals as

$$\widehat{u}_{it}=\widetilde{y}_{it}-\sum_{\ell=1}^{k_1}\widetilde{\rho}_{\ell,FE}\widetilde{y}_{i,t-\ell}-\widetilde{\theta}'_{FE}\widetilde{z}_{it},\quad i=1,...,n,\qquad t=1,...,T.$$

Let the centered residuals be $\check{u}_{it}=\widehat{u}_{it}-\frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}$.

STEP $2'$**:** Generate a random sample $\{\xi_t\}_{t=1}^{T}$ with zero mean and unit variance and obtain the bootstrap error terms as

$$\{u^*_{it}\}=\{\check{u}_{it}\xi_t\},\quad i=1,...,n,\qquad t=1,...,T.$$

Remark 4. *It is important to emphasize that while one might be tempted to obtain the residuals under the alternative hypothesis (as we did in the previous bootstrap), this would not be possible here. The reason for this is that it would translate into a bootstrap statistic that would be identically zero. Indeed, it is not difficult to see that its behaviour is governed by that of*

$$\sum_{i=1}^{n}u^*_{it}\widetilde{x}_{it}=\xi_t\sum_{i=1}^{n}\check{u}_{it}\widetilde{x}_{it}=0$$

by orthogonality between residuals and regressors.

STEP $3'$**:** Generate the bootstrap panel data model as

$$\widetilde{y}^*_{it}=\sum_{\ell=1}^{k_1}\widetilde{\rho}_{\ell,FE}\widetilde{y}_{i,t-\ell}+\widetilde{\theta}'_{FE}\widetilde{z}_{it}+u^*_{it},\quad i=1,...,n,\qquad t=1,...,T. \tag{4.4}$$

STEP $4'$**:** Compute the bootstrap analogues of our statistics $\mathcal{T}(r)$ and $\mathcal{T}^{\Delta}(r)$ with (4.4) as our dynamic panel regression model. That is, for $r=1,...,T-1$,

$$\begin{aligned}\mathcal{T}^*(r) &= \frac{1}{(nT)^{1/2}}\sum_{t=1}^{r}\sum_{i=1}^{n}\widetilde{x}_{it}\left(\widetilde{y}^*_{it}-\widehat{\beta}^{*\prime}_{FE}\widetilde{x}_{it}\right)\\ \mathcal{T}^{\Delta*}(r) &= \left(\frac{n}{T}\right)^{1/2}\sum_{t=1}^{r}\left(\widehat{\beta}^*_t-\widetilde{\beta}^*_{FE}\right).\end{aligned}$$

In the latter step, $\widehat{\beta}^*_{FE}$ denotes the fixed effect estimator of the slope parameters β, i.e.

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

and $\widehat{\beta}^*_t=\left(\sum_{i=1}^{n}\widetilde{x}_{it}\widetilde{x}'_{it}\right)^{-1}\left(\sum_{i=1}^{n}\widetilde{x}_{it}\widetilde{y}^*_{it}\right)$ with $\widetilde{\beta}^*_{FE}$ denoting the *MG bootstrap estimator*

$$\widetilde{\beta}^*_{FE}=\frac{1}{T}\sum_{s=1}^{T}\widehat{\beta}^*_s.$$

Remark 5. (**i**) *The second bootstrap approach is similar to that in Chan and Ogden* (2009) *and can be regarded as a wild-type bootstrap with increasing dimensional vectors. In this sense, we can view the bootstrap as a generalization or extension of bootstrapping $VAR(P)$ models, say, when the dimension of the (time series) sequence n grows with no limit. Notice that in the case of finite n, a standard approach to bootstrap VAR models is to obtain the bootstrap errors as $\{e_t\xi_t\}_{t=1}^{T}$, where ξ_t is a scalar sequence and e_t denote residuals.*

(ii) *We have assumed that the sequence* $\{\xi_t\}_{t=1}^T$ *has mean zero and unit variance. In the standard wild bootstrap algorithm, it is often suggested that the random variables* ξ_t *should also have unit skewness. As our purpose is to illustrate and describe a valid bootstrap in our scenario, we have ignored this.*

One major and important difference between the two bootstrap algorithms is that in the latter algorithm we cannot use the residuals obtained under the alternative hypothesis, that is $\widehat{u}_{it} = \widetilde{y}_{it} - \widehat{\beta}_t' \widetilde{x}_{it}$. It is well known that the use of residuals obtained under the null in the bootstrap, although needed to establish its validity, may suffer from inferior power properties than similar bootstraps where the residuals are computed under the alternative hypothesis. Indeed this is corroborated in our simulation results and reinforces the observation that for bootstrapped tests to have good power properties the residuals should be computed under the alternative hypothesis when possible. The heuristic explanation for this comes from the observation that residuals that are computed under the null hypothesis will not "estimate" the true error term when the alternative hypothesis is true.

In both bootstrap algorithmns, specifically as it relates to *STEP* 3 and *STEP* 3′, we have kept $y_{i,t-\ell}$ as an explanatory covariate instead of $y^*_{i,t-\ell}$ as is typically done in time series data, see e.g. Neumann and Kreiss (1998).

We conclude this section by providing a bootstrap estimator for V_2, and hence $V_1 = \Sigma_x V_2 \Sigma_x$, for use in our tests (3.5). To that end, suppose that we compute $\widehat{\beta}^*_{FE}$, as in *STEP* 4, for B bootstrap samples *STEPS 2* and *3*, that is

$$\widehat{\beta}^*_{FE}(b) = \left(\sum_{i=1}^n \sum_{t=1}^T \widetilde{x}_{it}\widetilde{x}_{it}'\right)^{-1} \left(\sum_{i=1}^n \sum_{t=1}^T \widetilde{x}_{it}\widetilde{y}^*_{it}(b)\right), \quad b = 1, ..., B,$$

where

$$\begin{aligned} \widetilde{y}^*_{it}(b) &= \sum_{\ell=1}^{k_1} \widetilde{\rho}_{FE,\ell}\widetilde{y}_{i(t-\ell)} + \widetilde{\theta}'_{FE}\widetilde{z}_{it} + u^*_{it}(b), \quad i = 1, ..., n, \quad t = 1, ..., T, \\ & \{U^*_t(b)\}_{t=1}^T, \end{aligned}$$

and $U^*_t(b) = \{u^*_{it}(b)\}_{i=1}^n$. The estimate for V_2 we may use in our tests (3.5) then is given by

$$\widehat{V}^*_2 = \frac{1}{B}\sum_{b=1}^B \left(\widehat{\beta}^*_{FE}(b) - \frac{1}{B}\sum_{v=1}^B \widehat{\beta}^*_{FE}(v)\right)^2$$

which would replace $\widehat{V}_2$ when making inferences.

5. FINITE SAMPLE BEHAVIOUR.

In this section we present a Monte-Carlo experiment that illustrates the performance of our tests in finite samples where. We consider the typical weighting functions $w(\tau) = 1$ and $w(\tau) = \tau^{-1/2}(1-\tau)^{-1/2}$ and we compare the bootstrap algorithms used to obtain valid critical values, revealing that both typically outperform the use of asymptotic critical values.

The data generating processes we consider are

$$DGP1 \quad : \quad y_{it} = \alpha_t + \rho y_{i,t-1} + \theta z_{it} + \delta_\theta z_{it} 1(t > t_0) + u_{it}$$
$$DGP2 \quad : \quad y_{it} = \alpha_t + \rho y_{i,t-1} + \delta_\rho y_{i,t-2} 1(t > t_0) + \theta z_{it} + u_{it}$$

for $i = 1, ..., n$ and $t = 1, ..., T$. We allow for breaks in the slope of the strictly exogenous variable z_{it} (δ_θ) and the number of lagged endogenous variable(s) (δ_ρ) and consider different scenarios for the time of the break (t_0). The time fixed effects α_t are drawn randomly across t ($\alpha_t \sim IIDN(1, 1)$) and are held fixed across replications. The regressor, z_{it}, is a strictly exogenous regressor generated as

$$z_{it} = \alpha_t + v_{it} \text{ with } v_{it} = \rho_{z_i} v_{i,t-1} + \sqrt{1 - \rho_{z_i}^2} \vartheta_{it}$$

and either (i) $\rho_{z_i} = 0$ (no temporal dependence), (ii) $\rho_{z_i} = 0.5$ or 0.9 (individual-homogenous autoregressive time dependence), or (iii) $\rho_{z_i} \sim IIDU[0.05, 0.95]$ (individual-heterogeneous autoregressive time dependence). For robustness, we separately consider the presence of individual fixed effects η_i in z_{it}, with η_i drawn randomly across i ($\eta_i \sim IIDN(1, 1)$) and held fixed across replications. Several cross-sectional dependence scenarios are considered for z_{it} (ϑ_{it}) : no spatial dependence, weak spatial dependence and strong spatial dependence. In the absence of cross-sectional dependence, the error term ϑ_{it} is $IIDN(0, \sigma_{z_i}^2)$ for $i = 1, .., n$ with $\sigma_{z_i}^2 \sim IID\chi^2(1)$ held fixed across replications (or $\sigma_{z_i}^2 = 1$ when ignoring heteroscedasticity). We consider two weak spatial dependence formulations. First we follow Robinson and Lee (2013). Here random locations for individual units are drawn along a line, denoted $s = (s_1, ...s_n)'$ with $s_i \sim IIDU[0, n]$. Keeping these locations fixed across replications, ϑ_{it} are generated independently as scalar normal variables with mean zero and covariances $cov(\vartheta_{it}, \vartheta_{jt}) = \sigma_{z_i}\sigma_{z_j}(0.5)^{|s_i - s_j|}$, ensuring z_{it} exhibits an exponential decay in dependence with distance across individuals. Second, we consider a polynomial decay of dependence in z_{it} with distance across individuals. Using the linear time dependence representation, $\vartheta_{it} = \sigma_i \left(\sum_{\ell=1}^{\infty} c_\ell(i) e_{\ell t}\right)$, we chose $c_\ell(i) = |s_\ell - s_i|^{-10}$ where s_i and s_ℓ are random locations (s_i is drawn from $IIDU[0, n]$ as before, while s_ℓ is drawn from $IIDN(0, n)$) and $e_{\ell t} \sim IIDN(0, 1)$. σ_i is such that $Var(\vartheta_{it}) = \sigma_{z_i}^2$. For the strong spatial dependence setting, we use $c_\ell(i) = |s_\ell - s_i|^{-0.9}$ instead.[1]

While not allowing for any temporal dependence of the error term, we consider the same scenarios for the cross-sectional dependence for the error term u_{it}. In the absence of cross-sectional dependence, $u_{it} \sim IIDN(0, \sigma_{u_i}^2)$ for $i = 1, ..., n$ with $\sigma_{u_i}^2 \sim IID\chi^2(2)/2$ held fixed across replications (or $\sigma_{u_i}^2 = 1$ when ignoring heteroscedasticity). The above discussion of the cross-sectional dependence scenarios for ϑ_{it}, suitably modified, holds for u_{it}.

In the tables below, we report empirical size and power of our tests at the nominal 5% level for various pairs of n and T using 10,000 simulations. The columns labelled $\mathcal{T}_\varepsilon$ relate to the CUSUM based test, while $\mathcal{T}_\varepsilon^\Delta$ relate to the associated Hausman-Durbin-Wu type, or slope based, test. When $\varepsilon = 0$, they present the untrimmed version of the tests with $w(\tau) = 1$; for the trimmed

[1] In the polynomial case, we use $\max(1, |s_\ell - s_i|)$ as our measure of distance; not imposing such a censoring would remove all dependence in settings where for some (ℓ, i) s_ℓ and s_i lie very close together.

versions of the test ($\varepsilon > 0$) we apply $w(\tau) = \tau^{-1/2}(1-\tau)^{-1/2}$. Under the null $H_0 : \delta = 0$ with $\delta = (\delta_\rho|\delta_\theta)'$ both DGPs are identical. We let $\rho = 0.5$ and $\theta = 1$.

In the first set of simulations (Tables 1.1–1.6), we ignore any heterogeneity or time dependence issues and focus on the cross-sectional dependence of z_{it} and u_{it} only; here z_{it} is correlated with α_t (no individual heterogeneity in z_{it}). The empirical size of our tests for the joint null $H_0 : \delta = 0$ against $H_0 : \delta \neq 0$ in either DGP is provided in Table 1.1.

Insert Table 1.1 around here

The exact asymptotic critical values from Estrella (2003) with $p = 2$ are used to obtain the empirical size of the trimmed version of the test. They suggest that in finite samples, the CUSUM based test is undersized for all cross-sectional dependence scenarios; the slope based test on the other hand appears oversized when n is quite small $(n = 25)$, especially in the presence of stronger cross-sectional dependence. The empirical sizes based on the two bootstrap algorithms are given for both the trimmed and untrimmed versions of the test. In general, the empirical size of our tests based on the bootstrap algorithm are much closer to the nominal size, with the Efron bootstrap yielding in most scenarios an empirical size closest to the nominal size. For example, with small sample sizes ($n = T = 25$) the empirical size of the untrimmed CUSUM test $\mathcal{T}_0$ based on the Efron bootstrap equals 0.047 in the absence of spatial dependence, 0.051 and 0.047 in the presence of respectively exponential and polynomial weak spatial dependence, versus 0.050 in the presence of strong spatial dependence. The performance of the $\mathcal{T}_\varepsilon^\Delta$ test vis-a-vis the $\mathcal{T}_\varepsilon$ test suggests a worsening of the coefficient based test with the level of spatial dependence. For small sample sizes ($n = T = 25$) the empirical size of the untrimmed coefficient test, $\mathcal{T}_0^\Delta$, based on the Efron bootstrap equals 0.048 in the absence of spatial dependence, 0.036 and 0.039 in the presence of respectively exponential and polynomial weak spatial dependence, versus 0.011 in the presence of strong spatial dependence. For the $\mathcal{T}_\varepsilon^\Delta$ test to remain properly sized, the cross sectional sample needs to be larger when the level of spatial dependence increases. The simulations do reveal fluctuation in the empirical sizes associated with the level of trimming of our test. Increasing the trimming generally improves the size of the tests with $w(\tau) = \tau^{-1/2}(1-\tau)^{-1/2}$ but this obviously limits the possibility of detecting a break closer to the end of the sample due to this trimming. In view of this, the good performance of the untrimmed tests with $w(\tau) = 1$ is promising.

In Table 1.2 we present empirical sizes of the slope-based test associated with the associated individual hypotheses $H_0 : \delta_\theta = 0$ (DGP1) and $H_0 : \delta_\rho = 0$ (DGP2). Here we provide exact asymptotic critical values for the untrimmed tests based on asymptotic critical values from $\sup_\tau |BB(\tau)|$ (with $\sup_\tau \sqrt{(H\mathcal{T}^\Delta(r))'\left(H\widehat{V}_2H'\right)^{-1} H\mathcal{T}^\Delta(r)} \overset{d}{\Longrightarrow} \sup_\tau |BB(\tau)|$ with $H = (1:0)$ and $(0:1)$, respectively); for $\varepsilon > 0$ we use Estrella (2003) with $p = 1$.

Insert Table 1.2 around here

The empirical sizes of the individual coefficient tests are comparable for δ_θ and δ_ρ and both are of the same order of magnitude as the joint test size. With $n = T = 100$, the size of our tests for homogeneity of the individual slope coefficient θ equals 0.051, 0.052, 0.051 and 0.043 for the four spatial dependence scenarios respectively, while the size of our tests for homogeneity of the individual slope coefficient ρ equals 0.053 0.049, 0.053 and 0.039 respectively. The rejection rates for the untrimmed tests based on the asymptotic values of the supremum of the Browning bridge are generally larger than the rejection rates for the trimmed tests relying on Estrella's exact asymptotic critical values. This in typically also the case for the rejection rates associated with the bootstrap algorithms.

Tables 1.3 and 1.4 present the power of our tests for DGP1 with $\delta_\theta = 0.5$ and $\delta_\rho = 0$, that is where we only have a break in the slope of the strictly exogenous variable z_{it}, when the break is either in the middle, $t_0 = [0.5T]$, or in the second half of the sample, $\tau_0 = [0.8T]$. In Table 1.3 the power of the joint test is presented for the CUSUM test, $\mathcal{T}_\varepsilon$, and the slope-based test, $\mathcal{T}_\varepsilon^\Delta$, while in Table 1.4, the power (size) of the individual components in the slope-based test is given. The columns labelled θ report the rejection rates associated with $H_0 : \delta_\theta = 0$ while the columns labelled ρ report the rejection rates associated with $H_0 : \delta_\rho = 0$. Given that we only consider a break in θ here, the column labelled θ presents the power of its detection ($\delta_\theta \neq 0$), while the column labelled ρ presents the size of $H_0 : \delta_\rho = 0$. In Table 1.3, we observe that even with small sample sizes our tests have high power in detecting a break in θ.

Insert Table 1.3 around here

As expected, the power is lower when the break lies closer to the end of the sample. Using the Efron bootstrap algorithm the power is 0.988 for $\mathcal{T}_{0.10}$ and 0.843 for $\mathcal{T}_{0.10}^\Delta$ in the absence of spatial dependence when the break lies in the middle of the sample, against 0.893 for $\mathcal{T}_{0.10}$ and 0.685 for $\mathcal{T}_{0.10}^\Delta$ when the break lies towards the end of the sample. Moreover, the power decreases with the cross-sectional dependence. The power of the $\mathcal{T}_\varepsilon$ test generally is higher than $\mathcal{T}_\varepsilon^\Delta$ test using the bootstrap based critical values. Nevertheless, when focussing on the power associated with the single coefficient test ($H_0 : \delta_\theta = 0$), the $\mathcal{T}_\varepsilon^\Delta$ test again performs comparable to the $\mathcal{T}_\varepsilon$ test in detecting the break in θ. In Table 1.4 we observe that the power associated with a break in θ in the middle of the sample for this example equals 0.820 for $\mathcal{T}_0^\Delta$ (the associated size for a break in ρ is 0.034), which compares well with 0.880 for the CUSUM based test.

Insert Table 1.4 around here

Clearly the power of an individual coefficient based test for a single break is higher than the power of a joint coefficient based test. When both n and T equal 100, the power of the tests (joint $\mathcal{T}_\varepsilon$ and $\mathcal{T}_\varepsilon^\Delta$ and individual for $\mathcal{T}_\varepsilon^\Delta$) equals 1 for all but the strong spatial dependence setting in which case it is close to one. Finally, the empirical power of our tests based on the Efron bootstrap typically exceeds the Wild bootstrap based ones as expected.

Tables 1.5 and 1.6, by symmetry, present the power of our tests for DGP2 with $\delta_\rho = 0.15$ and $\delta_\theta = 0$, that is where we only have a change in the number of lags of the endogenous variable, again for the case where the break is either in the middle, $t_0 = [0.5T]$, or in the second half of the sample, $\tau_0 = [0.8T]$.

Insert Tables 1.5 and 1.6 around here

In Table 1.5, the power of the joint test is presented for the CUSUM test and the slope-based test, while in Table 1.6, the power (size) of the individual components of the slope-based test is given. Given that we only consider a break in the number of lags of the endogenous variable here, the power of its detection is in the column labelled ρ while the column labelled θ reports the size of $H_0 : \delta_\theta = 0$. In Table 1.5, we observe that for both tests, the power of detecting a break in the number of lags of the endogenous variable is smaller than the power of a break in the slope of the strictly exogenous regressor. For example, in the presence of exponential weak dependence and small samples ($n = T = 25$), $\mathcal{T}_0$ reveals only a 0.151 power of detecting a break in ρ against the 0.880 power of detecting a break in θ. From the results it is clear, that in order to detect a change in the number of lags of the endogenous variable, we require a larger sample size (either in the time or cross-sectional dimension). The power of the $\mathcal{T}_\varepsilon$ test generally is again higher than $\mathcal{T}_\varepsilon^\Delta$ test using the bootstrap based critical values, a difference which is reduced when focussing on the power associated with the single coefficient test ($H_0 : \delta_\rho = 0$). For $n = T = 100$, we observe that the loss in power of detecting a break in the number of lags of the endogenous variable increases with the amount of spatial dependence. Using $\mathcal{T}_0^\Delta$, the power of detection based on the Efron bootstrap drops from 0.981 in absence of spatial dependence, to 0.884 and 0.696 in the weak spatial dependence setting to 0.230 in the strong spatial dependence setting.

In the second set of simulations (Tables 2.1-2.3) we reveal the robustness of the results to the presence of individual fixed effects η_i in the strictly exogenous regressor z_{it}.[2] In Table 2.1 we observe that the presence of fixed individual heterogeneity in the strictly exogenous regressor has no big impact on the size of our tests.

Insert Table 2.1 around here

While the coefficient based test, $\mathcal{T}_\varepsilon^\Delta$, seems to be most sensitive to the introduction of fixed individual heterogeneity, the impact is not unidirectional. Based on the asymptotic critical values, it does increase the power of detecting a break in θ of both tests, more significantly so for the CUSUM based test (see Table 2.2).

Insert Table 2.2 around here

In the presence of polynomial weak spatial dependence for small samples ($n = T = 25$), the power of $\mathcal{T}_{0.10}$ for a break in the middle of the sample becomes 0.798 (up from 0.310) and the power of $\mathcal{T}_{0.10}^\Delta$ becomes 0.836 (up from 0.467) based on the asymptotic critical values. The rejection rates of $\mathcal{T}_\varepsilon$ based on the bootstrap algorithms follows the same pattern. While similarly increasing the power of the untrimmed coefficient based test, $\mathcal{T}_0^\Delta$, the presence of fixed individual heterogeneity in z_{it} reduces the power of $\mathcal{T}_\varepsilon^\Delta$ for $\varepsilon = 0.05$ and 0.10 dramatically when n is small, i.e., when $n = 25$. For larger sample sizes, nevertheless, we detect a break in θ with probability 1 in all settings. The power associated with detecting a change in the number of lags of the endogenous

[2]Tables for individual slope based test for these and later simulations are not reported but are available upon request.

variables (Table 2.3) reveals a similar response to the introduction of the individual heterogeneity in z_{it} as the power associated with detecting a break in θ, be it less pronounced.

Insert Table 2.3 around here

The results suggest that our tests are robust to the presence of fixed individual heterogeneity in z_{it}. In the remainder of the simulations we have left fixed individual heterogeneity in the strictly exogenous regressor out.

In the third set of simulations (Tables 3.1-3.3), we introduce heterogeneity (random) in u_{it} and z_{it} while we continue to ignore time dependence in z_{it}. The presence of heterogeneity in u_{it} and z_{it} has little (no unidirectional) impact on the size of our tests as can be seen from Table 3.1.

Insert Table 3.1 around here

Table 3.2 reveals that the introduction of this heterogeneity typically reduces the power of our tests of detecting a break in θ in small samples. Only in the presence of strong dependence is the power of detection of a break in θ enhanced by the heterogeneity.

Insert Table 3.2 around here

The rejection rates for detecting a break in the middle of the sample ($n = T = 100$) using $\mathcal{T}_0$ increases to 0.556 (up from 0.378) in the presence of strong dependence; in the absence of spatial dependence heterogeneity using $\mathcal{T}_0$ decreases the power of detection in the middle of the sample to 0.870 (down from 0.998) with a slightly attenuated loss in the presence of polynomial weak dependence where the power reduces to 0.739 (down from 0.797). While the impact on the power using asymptotic critical values generally is stronger for the CUSUM based test, $\mathcal{T}_\varepsilon$, relative to the coefficient based test, $\mathcal{T}_\varepsilon^\Delta$, the reverse holds when using bootstrapped critical values. Again, with larger samples sizes $n = T = 100$ the break in θ is detected with probability 1 for all specifications. The power of our tests to detect a change is the number of lags of the endogenous variable in small samples typically is reduced in the presence of heterogeneity as well, more so when the break lies further towards the end of the sample, see Table 3.3. Only in the presence of strong spatial dependence, does the presence of heterogeneity improve the power of detecting in a change in the number of lags of the endogenous variable in small samples ($n = T = 25$).

Insert Table 3.3 around here

In the last set of simulations (Tables 4.1-4.6), we introduce time dependence in z_{it}. In Table 4.1, we provide the empirical size of our test in the presence of individual-heterogeneous autoregressive time dependence ($\rho_{z_i} \sim IIDU[0.05, 0.95]$). Here we consider only the settings of absence of cross-sectional dependence and weak spatial dependence as suggested by Lee and Robinson (2013). In Tables 4.2 and 4.3 we provide the associated power of our test for DGP1 and DGP2 in this setting respectively. From Table 4.1 we observe that the size of our tests, either based on asymptotic critical values or our bootstrap algorithms, remain largely unaffected by the introduction of individual-heterogeneous autoregressive time dependence of the regressor z_{it}.

Insert Table 4.1 around here

In Table 4.2 we notice a slight reduction in the power of our tests in detecting a break in θ in small samples which is slightly more pronounced for the coefficient based test, $\mathcal{T}_{\varepsilon}^{\Delta}$. In the absence of spatial dependence, the power of our tests for detecting a break in θ ($n = T = 25$) in the second half of the sample using the Efron bootstrap become 0.858 (down from 0.893) for $\mathcal{T}_{0.10}$ versus 0.557 (down from 0.685) for $\mathcal{T}_{0.10}^{\Delta}$.

Insert Table 4.2 around here

The deterioration is similar whether there is weak spatial dependence or not. In the presence of weak spatial dependence the associated powers become 0.460 (down from 0.523) for $\mathcal{T}_{0.10}$ versus 0.183 (down from 0.280) for $\mathcal{T}_{0.10}^{\Delta}$. On the other hand, as we observe from Table 4.3, the power of our tests in detecting a break in the number of lags of the endogenous variable is enhanced by the introduction of time dependence, with the effect potentially somewhat diminished in the presence of weak spatial dependence.

Insert Table 4.3 around here

In the presence of weak spatial dependence the power of our tests detecting a break in ρ in the middle of the sample using the Efron bootstrap become 0.277 (up from 0.151) for $\mathcal{T}_0$ versus 0.179 (up from 0.108) for $\mathcal{T}_0^{\Delta}$. In absence of spatial dependence the comparable numbers are 0.558 (up from 0.273) for $\mathcal{T}_0$ versus 0.424 (up from 0.222) for $\mathcal{T}_0^{\Delta}$

In Table 4.4, we provide the empirical size of our test in the presence of individual-homogenous autoregressive time dependence ($\rho_{z_i} = 0.5$ and $\rho_{z_i} = 0.9$) under weak (polynomial decay) or strong cross-sectional dependence. In Tables 4.5 and 4.6 we again provide the associated power for DGP1 and DGP2 respectively. From Table 4.4, we observe that even in the strong spatial dependence setting, the size of our tests, either based on asymptotic critical values or our bootstrapped critical values, remain largely unaffected by the introduction of time dependence of the regressor z_{it}, where here we consider individual-homogenous autoregressive time dependence.

Insert Table 4.4 around here

Only when the individual-homogenous autoregressive time dependence is very strong, i.e., $\rho_z = 0.9$, does the coefficient based test, $\mathcal{T}_{\varepsilon}^{\Delta}$, seem to become less adequately sized when using the bootstrapped critical values. This problem is more severe as the level of spatial dependence increases, where in contrast the empirical size based on the asymptotic critical values approaches the nominal rate earlier ($n = T = 100$). In Table 4.5, we observe that in the presence of weak spatial dependence, an increase in the time dependence in the regressor lowers the power of detecting a break in the middle (and end) of the sample.

Insert Table 4.5 around here

For ($n = T = 25$), the power based on the Efron bootstrap of $\mathcal{T}_0$ detecting a such a break in θ goes from 0.797, 0.411 to 0.338 when ρ_z increases from 0, 0.5 to 0.9;[3] for $\mathcal{T}_0^{\Delta}$, the power correspondingly decreases from 0.438, 0.329 to 0.151. In the presence of strong spatial dependence the effect is

[3]The power associated with $\rho_z = 0$ can be found in Table 1.3.

not unidirectional, in fact the largest power is typically detected when $\rho_z = 0.5$ relative to $\rho_z = 0$ and 0.9, and the smallest typically for $\rho_z = 0.9$. Finally, as can be seen from Table 4.6 the power of detecting a break in the number of lags of the endogenous variable generally of the CUSUM based test, $\mathcal{T}_\varepsilon$, increases with the level of time dependence in the regressor.

Insert Table 4.6 around here

For instance for small samples ($n = T = 25$), based on the Efron bootstrap, the power of $\mathcal{T}_0$ detecting a such a break in ρ in the presence of weak spatial dependence goes from 0.131, 0.192 to 0.295 when ρ_z increases from 0, 0.5 to 0.9;[4] in the presence of strong spatial dependence the power correspondingly increases from 0.090, 0.128 to 0.162. The impact on the power of the coefficient based test, $\mathcal{T}_\varepsilon^\Delta$, is less clear when the sample is small ($n = T = 25$). Considering the larger sample ($n = T = 100$), it is clear that the power of the coefficient based test is also enhanced by the stronger time dependence in the regressor for both the weak and strong spatial dependence setting regardless of whether asymptotic or bootstrapped critical values are used.

6. CONCLUSIONS AND EXTENSIONS

The paper has examined several issues related to inference in large dynamic panel data models. Specifically, we have developed a Central Limit Theorem for the estimators of the slope parameters when the errors and the covariates might exhibit "strong" cross-sectional dependence. To that end, we have modified existing results given in Phillips and Moon (1999) to allow for dependence in both time and cross-section dimensions. From here, we have described and examined two different, but similar, tests for the null hypothesis of homogeneity of the slope parameters of the model. Because the small sample properties of the test were not very satisfactory, we have described two bootstrap algorithms with the attractive feature that their implementation does not require any previous knowledge of the cross-sectional dependence or selection of any tuning/bandwidth parameters (as is normally the case when using moving block bootstraps with time series data).

A possible limitation of the conditions we imposed is that it rules out temporal dependence for the errors. That is, we might want to change $C1$ to

$C1'$: $\{u_{it}\}_{t\in\mathbb{Z}}$, $i \in \mathbb{N}^+$, are linear sequences of zero mean random variables given by

$$u_{it} = \sum_{\ell=0}^{\infty} a_\ell(i)\,\varepsilon_{i,t-\ell}; \qquad \sum_{\ell=0}^{\infty} |a_\ell|\,\ell^{1/2} < \infty,$$

where $a_\ell = \sup_{i\in\mathbb{N}} |a_\ell(i)|$, and $\{\varepsilon_{it}\}_{t\in\mathbb{Z}}$, $i \geq 1$, are sequences of independent distributed random variables satisfying $\sup_{i\in\mathbb{N}} E\left(\varepsilon_{it}^4\right) = \sup_{i\in\mathbb{N}} \mu_i < \infty$ and

$$\lim_{T\nearrow\infty} \sup_{i\in\mathbb{N}^+} \sum_{t_1,t_2,t_3=1}^{T} |Cum(u_{it_1};u_{it_2};u_{it_3};u_{i0})| < \infty.$$

When we change $C1$ to $C1'$, inspections of our proofs suggests the main qualitative results of the paper would follow, all we need to do would be to employ *instrumental variable* methods or some type of Hatanaka's "efficient" estimator to estimate the parameters of the model. We have not

[4]The power associated with $\rho_z = 0$ can be found in Table 1.5.

followed this route for notational simplicity as our basic conclusions should not be affected, except that the proofs would become lengthier. The change, though, would necessitate a modification of our bootstrap algorithm to accommodate the temporal dependence of the errors $\{u_{it}\}_{t\in\mathbb{Z}}$, $i\in\mathbb{N}^{+}$, and the estimator of the long run variance V_1. Details of this are beyond the scope of this paper and we hope to address these issues in a different paper.

7. APPENDIX

7.1. PROOF OF THEOREM 1.

We begin by giving two intermediate lemmas, the first of which extends a Central Limit Theorem result given in Phillips and Moon's (1999) Theorem 2 when the independence condition fails.

Lemma 1. *Assume that*

$$\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\overset{d}{\to}\mathcal{N}\left(0,V_1\right), \tag{7.1}$$

where (**i**) $\{\mathring{x}_{it}\}_{t\in\mathbb{Z}}$, $i\in\mathbb{N}$, *are sequences of random variables such that*

$$\mathring{x}_{it}=\sum_{\ell=0}^{\infty}\psi_{\ell}\xi_{i,t-\ell},\ \sum_{\ell=0}^{\infty}\left|\psi_{\ell}\right|\ell^{1/a}<\infty,\quad 0<a<A<\infty,$$

with $\{\xi_{it}\}_{t\in\mathbb{Z}}$, $i\in\mathbb{N}$, *being iid zero mean sequences with finite fourth moments and* (**ii**) $\{u_{it}\}_{t\in\mathbb{Z}}$, $i\in\mathbb{N}$, *are iid sequences of random variables with finite fourth moments that are mutually independent of* $\{\xi_{it}\}_{t\in\mathbb{Z}}$. *Then, we have*

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\overset{d}{\to}\mathcal{N}\left(0,V_1\right), \tag{7.2}$$

as $n,T\to\infty$.

Proof. The proof is quite standard using time series techniques and the Bernstein's lemma. Observe that we cannot use Phillips and Moon's (1999) Theorem 2 as the latter result requires that the left side of (7.1) forms a sequence of independent random variables in the "t" dimension.

Suppose for the moment that for all $i\in\mathbb{N}^{+}$, the sequences $\{\mathring{x}_{it}\}_{t\in\mathbb{Z}}$ follow an $MA\left(p\right)$ process for some finite p. If that were the case, we could write (7.2) as

$$\begin{aligned}
&\frac{1}{\mathring{T}^{1/2}}\sum_{t=1}^{\mathring{T}}\frac{1}{k^{1/2}}\sum_{j=1}^{k-p}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{i,k(t-1)+j}u_{i,k(t-1)+j}\\
&+\frac{1}{\mathring{T}^{1/2}}\sum_{t=1}^{\mathring{T}}\frac{1}{k^{1/2}}\sum_{j=1}^{p}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{i,tk-p+j}u_{i,tk-p+j} \qquad (7.3)\\
=\ &:\frac{1}{\mathring{T}^{1/2}}\sum_{t=1}^{\mathring{T}}\frac{1}{k^{1/2}}\sum_{j=1}^{k-p}\mathring{X}_{j,t}\left(1\right)+\frac{1}{\mathring{T}^{1/2}}\sum_{t=1}^{\mathring{T}}\frac{1}{k^{1/2}}\sum_{j=1}^{p}\mathring{X}_{j,t}\left(2\right),
\end{aligned}$$

for some $k > 2p$ and where, for notational simplicity, we have assumed that $T = k\mathring{T}$. The second moment of the second term on the right of (7.3) equals

$$\frac{1}{\mathring{T}} \sum_{t=1}^{\mathring{T}} \frac{1}{k} \sum_{j_1,j_2=1}^{p} E\left(\mathring{X}_{j_1,t}(2)\, \mathring{X}_{j_2,t}(2)\right) = O(p/k),$$

because by assumption, $\mathring{X}_{j,t}(2)$ has finite second moments and is independent of $\mathring{X}_{j,s}(2)$ for $t \neq s$. Next, denoting $z_{tn} = k^{-1/2} \sum_{j=1}^{k-p} \mathring{X}_{j,t}(1)$, the first term on the right of (7.3) is

$$\frac{1}{\mathring{T}^{1/2}} \sum_{t=1}^{\mathring{T}} z_{tn},$$

which converges in distribution to a normal random variable by Phillips and Moon's (1999) Theorem 2 because z_{tn} and z_{sn} are independent and identically distributed sequences in t with finite second moments, so that they are uniformly integrable (observe that $\{\mathring{x}_{it}\}_{t\in\mathbb{Z}}$ follows an $MA(p)$), so that condition (3.20) in Phillips and Moon (1999) holds true. The conclusion of the lemma would now follow by using Bernstein's lemma.

We next examine the general case when $\{\mathring{x}_{it}\}_{t\in\mathbb{Z}}$ follows an $MA(\infty)$. The proof follows similarly after we employ the usual truncation of the sequence $\mathring{x}_{it}$. To that end, let

$$\begin{aligned} \mathring{x}_{it} &= \sum_{\ell=0}^{p} \psi_\ell \xi_{i,t-\ell} + \sum_{\ell=p+1}^{\infty} \psi_\ell \xi_{i,t-\ell} \\ &= \mathring{x}'_{it} + \mathring{x}''_{it}. \end{aligned}$$

Because for any fixed p, we have that the sequence $\{\mathring{x}'_{it}\}_{t\in\mathbb{Z}}$ behaves as a $MA(p)$ sequence, in view of the arguments previously given, it suffices to establish that the contribution of $\mathring{x}''_{it}$ into (7.2) is negligible. Indeed, this is the case because by definition,

$$\frac{1}{T^{1/2}} \sum_{t=1}^{T} \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}''_{it} u_{it} = \frac{1}{T^{1/2}} \sum_{t=1}^{T} \sum_{\ell=p+1}^{\infty} \psi_\ell \frac{1}{n^{1/2}} \sum_{i=1}^{n} \xi_{i,t-\ell} u_{it},$$

which second moment is

$$\begin{aligned} & \frac{1}{T} \sum_{t,s=1}^{T} \sum_{\ell_1,\ell_2=p+1}^{\infty} \psi_{\ell_1} \psi_{\ell_2} \frac{1}{n} E\left(\sum_{i=1}^{n} \xi_{i,t-\ell_1} u_{it} \sum_{j=1}^{n} \xi_{j,s-\ell_2} u_{js} \right) \\ =\; & \frac{C}{T} \sum_{t,s=1}^{T} \sum_{\ell=\max(0,t-s)+p+1}^{\infty} \left| \psi_\ell \psi_{\ell+s-t} \right| \\ \leq\; & C \sum_{q_1=p+1}^{\infty} \sum_{q_2=p+1}^{\infty} \left| \psi_{q_1} \psi_{q_2} \right| = O\left(p^{-2/a}\right). \end{aligned}$$

Now we choose p large enough to conclude that the behaviour of (7.2) is given by that of

$$\frac{1}{T^{1/2}} \sum_{t=1}^{T} \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}'_{it} u_{it}$$

and hence the proof of the lemma is completed by appealing again to Bernstein's lemma. □

Remark 6. *There is no doubt that the conditions of the previous lemma can be weakened. Indeed a close inspection of the proof suggests that the condition that* (7.1) *is not used in a significant way. However, since we shall show that under our Conditions* $C1 - C3$*,* (7.1) *indeed converges in distribution to a Gaussian random variable, and since our purpose is to illustrate how we can relax the independence condition in Phillips and Moon* (1999) *or Hahn and Kürsteiner* (2002)*, and thereby to modify their CLT, we have preferred to keep them to allow us to use, somehow, standard arguments.*

Remark 7. *Given* (7.1)*, the sequential limit of*

$$\frac{1}{T^{1/2}} \sum_{t=1}^{[T\tau]} \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}$$

will converge in distribution to a normal random variable with variance

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{[T\tau]} \lim_{n\to\infty} E\left(\frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}\right)^2 = \tau V_1 \tag{7.4}$$

as Condition $C1$ *implies that for* $t \neq s$*,*

$$Cov\left(\frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}; \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{is} u_{is}\right) = 0.$$

Of course, the same arguments yields that

$$Cov\left(\frac{1}{T^{1/2}} \sum_{t=1}^{[T\tau_1]} \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}; \frac{1}{T^{1/2}} \sum_{t=1}^{[T\tau_2]} \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{is} u_{is}\right) \to (\tau_1 \wedge \tau_2)\, V_1.$$

Lemma 2. *Under* $C2 - C4$*, we have that, uniformly in* $\tau \in [0,1]$*,*

$$\frac{1}{nT} \sum_{t=1}^{[T\tau]} \sum_{i=1}^{n} \widetilde{x}_{it} \widetilde{x}_{it}' - \tau \Sigma_x = o_p(1). \tag{7.5}$$

Proof. First we show that we can replace $\tilde{x}_{it}$ by $\mathring{x}_{it}$ on the left side of (7.5). Indeed, after standard algebra, we have that the difference equals

$$\frac{1}{T} \sum_{t=1}^{[T\tau]} \left\{ \left(\frac{1}{n} \sum_{i=1}^{n} \mathring{x}_{it}\right) \left(\frac{1}{n} \sum_{i=1}^{n} \mathring{x}_{it}'\right) \right\}.$$

Now, by the triangle inequality we obtain that

$$E \sup_{\tau\in[0,1]} \left\| \frac{1}{T} \sum_{t=1}^{[T\tau]} \left(\frac{1}{n} \sum_{i=1}^{n} \mathring{x}_{it}\right) \left(\frac{1}{n} \sum_{i=1}^{n} \mathring{x}_{it}'\right) \right\| \leq \frac{1}{T} \sum_{t=1}^{T} E \left\| \frac{1}{n} \sum_{i=1}^{n} \mathring{x}_{it} \right\|^2 = o(1),$$

by $C2$. So, to complete the proof of the lemma, it remains to show that uniformly in $\tau \in [0,1]$,

$$\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{[T\tau]} \left(\mathring{x}_{it} \mathring{x}_{it}'\right) - \tau \Sigma_x = o_p(1). \tag{7.6}$$

Now, for a fixed $\tau \in [0,1]$, we have that standard arguments imply that (7.6) holds true. So, it remains to show that the left side of (7.6) is tight. By Billingsley's (1968) Theorem 12.1, it suffices to show the moment condition

$$E\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\upsilon_{it}\right)^2 \le C\left(\tau_2-\tau_1\right)^{1+\xi} \tag{7.7}$$

for some $\xi > 0$ where υ_{it} denotes a typical component of the matrix $\mathring{x}_{it}\mathring{x}_{it}' - \Sigma_{x,i}$. The last displayed inequality holds as $C2$ implies that the left side is bounded by

$$\begin{aligned}\frac{C}{n^2T^2}\sum_{i,j=1}^{n}\sum_{t,s=[T\tau_1]+1}^{[T\tau_2]}\left|Cov\left(\upsilon_{it};\upsilon_{js}\right)\right| &\le \frac{C}{n^2T^2}\sum_{i,j=1}^{n}\varphi_{\upsilon}\left(i,j\right)\sum_{t,s=[T\tau_1]+1}^{[T\tau_2]}\left|Cov\left(\upsilon_{it};\upsilon_{is}\right)\right| \\ &= o\left(\frac{\tau_2-\tau_1}{T}\right),\end{aligned}$$

due to the separability of the dependence structure $\left|Cov\left(\upsilon_{it};\upsilon_{js}\right)\right| \le C\left|\varphi_{\upsilon}\left(i,j\right)\right|\left|Cov\left(\upsilon_{it};\upsilon_{is}\right)\right|$. This concludes the proof of Lemma 2 by taking $\xi = 1$ in (7.7) because $T^{-1} \le (\tau_2 - \tau_1)$. □

We now turn to the proof of Theorem 1 itself.

Proof. It is clear that, in view of Lemma 1, it suffices to show that (7.1) holds true. That is,

$$\epsilon_n = \frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_i u_i \xrightarrow{d} \mathcal{N}\left(0, V_1\right),$$

where we have suppressed the subindex t for notational simplicity. Next $C1$ implies that we can write ϵ_n as

$$\begin{aligned}\epsilon_n &= \frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_i\sum_{j=0}^{\infty}a_j\left(i\right)\varepsilon_j \\ &= \epsilon_{1,n} + \epsilon_{2,n},\end{aligned}$$

where denoting $\ddot{x}_j\left(n\right) = \sum_{i=1}^{n}\mathring{x}_i a_j\left(i\right)$,

$$\epsilon_{1,n} = \frac{1}{n^{1/2}}\sum_{j=0}^{N}\ddot{x}_j\left(n\right)\varepsilon_j; \qquad \epsilon_{2,n} = \frac{1}{n^{1/2}}\sum_{j=N+1}^{\infty}\ddot{x}_j\left(n\right)\varepsilon_j,$$

with N large enough and to be chosen later.

The proof of the theorem will be completed if (**a**) $E\left(\epsilon_{2,n}\epsilon_{2,n}'\right) = o\left(1\right)$; (**b**) $\epsilon_{1,n}$ satisfies a Central Limit Theorem and (**c**) $\Xi_1 - \Xi_0 = o_p\left(1\right)$, where

$$\Xi_0 = E\left(\epsilon_n\epsilon_n' \mid \{x_i\}_{i=1}^{n}\right) > 0; \quad \Xi_1 = E\left(\epsilon_{1,n}\epsilon_{1,n}' \mid \{x_i\}_{i=1}^{n}\right) > 0,$$

as Lemma 2 implies that $n^{-1}\sum_{i=1}^{n}\left(\mathring{x}_i\mathring{x}_i' - \Sigma_{x,i}\right) = o_p\left(1\right)$.

We begin with (**a**). By $C1$ and $C2$, we obtain that

$$\begin{aligned} E\left(\epsilon_{2,n}\epsilon_{2,n}'\right) &= \frac{\sigma_{\varepsilon}^{2}}{n}\sum_{j=N+1}^{\infty}E\left[\left(\sum_{i=1}^{n}\mathring{x}_{i}a_{j}\left(i\right)\right)\left(\sum_{i=1}^{n}\mathring{x}_{i}a_{j}\left(i\right)\right)'\right] \\ &= \frac{\sigma_{\varepsilon}^{2}}{n}\sum_{j=N+1}^{\infty}\sum_{i,k=1}^{n}\varphi_{x}\left(i,k\right)a_{j}\left(i\right)a_{j}\left(k\right) \\ &= \frac{\sigma_{\varepsilon}^{2}}{n}\sum_{i,k=1}^{n}\varphi_{x}\left(i,k\right)\varphi_{u}\left(i,k\right)\frac{\sum_{j=N+1}^{\infty}a_{j}\left(i\right)a_{j}\left(k\right)}{\varphi_{u}\left(i,k\right)}. \end{aligned}$$

The right side of the last displayed equality is bounded by

$$\begin{aligned} &\left\{\frac{\sigma_{\varepsilon}^{2}}{n}\sum_{i,k=1}^{n}\varphi_{x}\left(i,k\right)\varphi_{u}\left(i,k\right)\right\}\sup_{i,k=1,\ldots,n}\frac{\sum_{j=N+1}^{\infty}a_{j}\left(i\right)a_{j}\left(k\right)}{\varphi_{u}\left(i,k\right)} \\ \leq\ & C\sup_{i,k=1,\ldots,n}\frac{\sum_{j=N+1}^{\infty}a_{j}\left(i\right)a_{j}\left(k\right)}{\varphi_{u}\left(i,k\right)} \qquad (7.8) \end{aligned}$$

given $C3$. But using (2.1) and Cauchy-Schwarz's inequality, the right side of (7.8) can be made as small as we want by choosing N sufficiently large. As a by-product we can immediately conclude also that $\Xi_0 - \Xi_1 = o_p(1)$, i.e. part (**c**).

It remains to examine part (**b**). To that end, we shall make use of a result by Scott (1973) that ensures that conditionally on $\{x_i\}_{i=1}^n$, we have that

$$\left(\frac{1}{n}\sum_{j=0}^{N}\ddot{x}_{j}\left(n\right)\ddot{x}_{j}'\left(n\right)\right)^{-1/2}\frac{1}{n^{1/2}}\sum_{j=0}^{N}\ddot{x}_{j}\left(n\right)\varepsilon_{j}\overset{d}{\rightarrow}\mathcal{N}\left(0,\sigma_{\varepsilon}^{2}I_{k}\right).$$

Recall that if this is the case, as the limiting distribution does not depend on $\{x_i\}_{i=1}^n$, it implies that the wording conditional can be changed to unconditional.

Using the Cramèr-Wold device, according to Scott (1973) , the latter holds true if the sufficient conditions

$$(\mathbf{i})\ \ \sum_{j=0}^{N}\vartheta_{j}^{2}\left(n\right)E\varepsilon_{j}^{2}\overset{P}{\rightarrow}\sigma_{\varepsilon}^{2},$$

$$(\mathbf{ii})\ \ E\sum_{j=0}^{N}\vartheta_{j}^{2}\left(n\right)E\left(\varepsilon_{j}^{2}\mathcal{I}\left(\vartheta_{j}^{2}\left(n\right)\varepsilon_{j}^{2}>\eta\right)\right)\rightarrow 0\quad\text{for all }\eta>0,$$

where $\vartheta_j(n) = \left(\sum_{j=0}^{N} c'\ddot{x}_j(n)\ddot{x}_j'(n)c\right)^{-1/2} c'\ddot{x}_j(n)$, are satisfied for all c for which $c'\left(\frac{1}{n}\sum_{j=0}^{N}\ddot{x}_j(n)\ddot{x}_j'(n)\right)c \neq 0$,. Having dropped the conditional expectation in $\{x_i\}_{i=1}^n$ for notational simplicity, we notice that (**ii**) is some type of Lindeberg's condition. Now condition (**i**) follows trivially in view of $C1$, since by construction

$$\sum_{j=0}^{N}\vartheta_{j}^{2}\left(n\right)=1.$$

So, we need to examine part (**ii**). By standard inequalities, the left side is bounded by

$$E\sum_{j=0}^{N}\vartheta_j^2(n)\,E\left(\varepsilon_j^2\mathcal{I}(|\varepsilon_j|>\eta/\delta)\right)+\Pr\left\{\sup_{j\geq1}|\vartheta_j(n)|>\delta\right\}. \tag{7.9}$$

The first term of (7.9) converges to zero in probability by Uniform Integrability of the sequence $\left\{\varepsilon_j^2\right\}_{j\in\mathbb{N}}$. The second term of (7.9) also converges to zero as we show next. First, suppose that

$$\frac{1}{n}\sum_{j=0}^{N}\ddot{x}_j(n)\,\ddot{x}_j'(n)\xrightarrow{P}\Xi>0, \tag{7.10}$$

then it suffices to show that

$$\sup_{j\geq1}\left\|\frac{\ddot{x}_j(n)}{n^{1/2}}\right\|^2=\sup_{j\geq1}\left\|\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_ia_j(i)\right\|^2=o_p(1). \tag{7.11}$$

By Condition $C2$, we know that we can choose m large enough such that $\sup_{j\geq1}\sum_{i=m+1}^{n}|a_j(i)|^2<\epsilon$. So, the left side of (7.11) is bounded by

$$\begin{aligned}&\sup_{j\geq1}\left\|\frac{1}{n^{1/2}}\sum_{i=1}^{m}\mathring{x}_ia_j(i)\right\|^2+\sup_{j\geq1}\left\|\frac{1}{n^{1/2}}\sum_{i=m+1}^{n}\mathring{x}_ia_j(i)\right\|^2\\ \leq\ &\frac{m}{n^{1/2}}\left(\sup_iE\|\mathring{x}_i\|^4\right)^{1/4}\sup_{i,j}|a_j(i)|+\frac{1}{n}\sum_{i=m+1}^{n}\|\mathring{x}_i\|^2\left(\sup_{j\geq1}\sum_{i=m+1}^{n}|a_j(i)|^2\right)\\ <\ &\epsilon.\end{aligned}$$

To finish the proof of part (**b**), it remains to show (7.10). By definition of $\ddot{x}_j(n)$, the left side in (7.10) is

$$\frac{1}{n}\sum_{j=0}^{N}\sum_{i=1}^{n}\mathring{x}_ia_j(i)\sum_{\ell=1}^{n}\mathring{x}_\ell'a_j(\ell)=\frac{1}{n}\sum_{i,\ell=1}^{n}\mathring{x}_i\mathring{x}_\ell'\sum_{j=0}^{N}a_j(i)\,a_j(\ell).$$

Now, proceeding as with (7.8), it follows that it suffices to show that

$$\frac{1}{n}\sum_{i,\ell=1}^{n}\left(\mathring{x}_i\mathring{x}_\ell'-\varphi_x(i,\ell)\right)\varphi_u(i,\ell)=o_p(1).$$

The second moment of the left side of the last displayed equality is

$$\begin{aligned}&\frac{1}{n^2}\sum_{i_1,\dots,i_4=1}^{n}\varphi_x(i_1,i_3)\,\varphi_x(i_2,i_4)\,\varphi_u(i_1,i_2)\,\varphi_u(i_3,i_4)\\ &+\frac{1}{n^2}\sum_{i_1,\dots,i_4=1}^{n}\varphi_x(i_1,i_4)\,\varphi_x(i_2,i_3)\,\varphi_u(i_1,i_2)\,\varphi_u(i_3,i_4)\\ &\frac{1}{n^2}\sum_{i_1,\dots,i_4=1}^{n}cum\left(\mathring{x}_{i_1},\mathring{x}_{i_2},\mathring{x}_{i_3},\mathring{x}_{i_4}\right)\varphi_u(i_1,i_2)\,\varphi_u(i_3,i_4).\end{aligned} \tag{7.12}$$

Now using expression (16) in Lee and Robinson (2013), we obtain that the first two terms of (7.12) are bounded by

$$C\frac{1}{n^2}\sum_{i_1,i_2=1}^{n}\varphi_x^2\left(i_1,i_2\right)\left(\max_{1\le i\le n}\sum_{j=1}^{n}\varphi_u\left(i,j\right)\right)^2=o\left(1\right),$$

using (2.4) and its consequence obtained in (2.6). Using $C2$, the third term of (7.12) is equal to

$$\frac{C}{n^2}\sum_{i_1,\ldots,i_4=1}^{n}\left\{\varphi_u\left(i_1,i_2\right)\varphi_u\left(i_3,i_4\right)\sum_{\ell=1}^{\infty}b_\ell\left(i_1\right)b_\ell\left(i_2\right)b_\ell\left(i_3\right)b_\ell\left(i_4\right)\right\}=o\left(1\right)$$

because $\max_{1\le i\le n}\sum_{j=1}^{n}\varphi_u^2\left(i,j\right)<C$ and $\max_{\ell\ge1}\sum_{i=1}^{n}\left\|b_\ell\left(i\right)\right\|^2<C$. This completes the proof of the theorem. □

7.2. **PROOF OF PROPOSITION 1.**

Proof. For notational simplicity, assume that $\widetilde{x}_{it}$ is scalar. Then, by definition and using standard algebra, we obtain

$$\begin{aligned}
\widehat{V}_1-V_1 &= \frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}\widetilde{u}_{it}-\left(\widehat{\beta}_{FE}-\beta\right)\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}^2\right\}^2-V_1\\
&= \frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\}^2+\left(\widehat{\beta}_{FE}-\beta\right)^2\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}^2\right\}^2\\
&\quad -2\left(\widehat{\beta}_{FE}-\beta\right)\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}^2\right\}-V_1 \qquad (7.13)\\
&= \left[\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\}^2-V_1\right]+o_p\left(1\right),
\end{aligned}$$

because Theorem 1 implies that $\widehat{\beta}_{FE}-\beta=O_p\left(n^{-1/2}T^{-1/2}\right)$ and by Markov's inequality,

$$\begin{aligned}
&\left(E\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\}\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}^2\right\}\right)^2\le E\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\}^2E\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}^2\right\}^2\\
\le\ & C.
\end{aligned}$$

Now, the term in square brackets on the far right of (7.13) is

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left(\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\}^2-\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\}^2\right)\\
&+\frac{1}{T}\sum_{t=1}^{T}\left(\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\}^2-E\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\}^2\right) \qquad (7.14)\\
&+\frac{1}{T}\sum_{t=1}^{T}E\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\}^2-V_1,
\end{aligned}$$

which converges in probability to zero as we now show. Indeed, by definition, the first line of (7.14) is

$$\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}u_{it}\right\}^{2}\left\{\frac{1}{n}\sum_{i=1}^{n}\mathring{x}_{it}\right\}^{2} \tag{7.15}$$
$$-\frac{2}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}u_{it}\right\}\left\{\frac{1}{n}\sum_{i=1}^{n}\mathring{x}_{it}\right\}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\}.$$

The expectation of the first term of (7.15) is

$$\begin{aligned} E\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}u_{it}\right\}^{2}E\left\{\frac{1}{n}\sum_{i=1}^{n}\mathring{x}_{it}\right\}^{2} &= \frac{1}{n}\sum_{i,j=1}^{n}\varphi_{u}\left(i,j\right)\frac{1}{n^{2}}\sum_{i,j=1}^{n}\varphi_{x}\left(i,j\right) \\ &= o\left(1\right) \end{aligned}$$

by $C3$, while the second term of (7.15) has absolute moment bounded by

$$\frac{2}{T}\sum_{t=1}^{T}\left(E\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}u_{it}\right\}^{2}E\left\{\frac{1}{n}\sum_{i=1}^{n}\mathring{x}_{it}\right\}^{2}\right)^{1/2}\left(E\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\}^{2}\right)^{1/2}$$

using Cauchy-Schwarz's inequality. Since Theorem 1 and Serfling (1980) ensure that $E\left\|\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\|^{2}=O\left(n\right)$ as $\left\|\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\|^{2}$ is Uniformly Integrable, we obtain that the last displayed expression is also $o\left(1\right)$ by the same argument we used for the first term. By Markov's inequality, we then conclude that (7.15) $=o_{p}\left(1\right)$. The last two terms of (7.14) are $o_{p}\left(1\right)$ by standard arguments, which completes the proof of the proposition. □

7.3. PROOF OF THEOREM 2.

Proof. Because of the results of Theorem 1 and Cràmer-Wold's arguments, it remains to show the tightness condition. We shall begin with part (**a**). First, by Lemma 2,

$$\begin{aligned} &\frac{1}{\left(nT\right)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\widetilde{x}_{it}\left(\widetilde{y}_{it}-\widehat{\beta}_{FE}^{\prime}\widetilde{x}_{it}\right) \\ =\;& \frac{1}{\left(nT\right)^{1/2}}\left(\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}-\tau\sum_{t=1}^{T}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right)\left(1+o_{p}\left(1\right)\right), \end{aligned} \tag{7.16}$$

where the $o_{p}\left(1\right)$ is uniformly in $\tau\in\left(0,1\right)$. Now using (2.9), we have that standard algebra implies that the first term on the right of (7.16) is

$$\frac{1}{\left(nT\right)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}+\frac{1}{\left(nT\right)^{1/2}}\sum_{t=1}^{[T\tau]}\left(n^{1/2}\overline{\mathring{x}}_{\cdot t}\left(\ell\right)\right)\left(n^{1/2}\overline{u}_{\cdot t}\right) \tag{7.17}$$

with $\mathring{x}_{it}\left(\ell\right)$ denoting the ℓth element of $\mathring{x}_{it}$. Define

$$\mathbb{X}_{n,T}\left(\tau\right)=\frac{1}{\left(nT\right)^{1/2}}\sum_{t=1}^{[T\tau]}\left(n^{1/2}\overline{\mathring{x}}_{\cdot t}\left(\ell\right)\right)\left(n^{1/2}\overline{u}_{\cdot t}\right). \tag{7.18}$$

By $C1-C3$, we obtain that

$$\begin{aligned} E\left(\mathbb{X}_{n,T}\left(\tau_2\right)-\mathbb{X}_{n,T}\left(\tau_1\right)\right)^2 &= \frac{1}{(nT)}E\left(\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\left(n^{1/2}\overline{\mathring{x}}_{\cdot t}\left(\ell\right)\right)\left(n^{1/2}\overline{u}_{\cdot t}\right)\right)^2 \\ &\le C\frac{\tau_2-\tau_1}{n} \\ &\le C\left(\tau_2-\tau_1\right)^{1+\xi} \end{aligned}$$

where ξ is such that $T^{\xi}=o\left(n\right)$ and, as usual, we can take $T^{-1}\le\left(\tau_2-\tau_1\right)$. Recall that $C1$ ensures that $E\left(\overline{u}_{\cdot t}\overline{u}_{\cdot s}\right)=0\ \forall t\neq s$. This ensures that $\mathbb{X}_{n,T}\left(\tau\right)$ is tight. As its finite distribution converges to zero in probability, we conclude that the second term of (7.17), i.e. $\mathbb{X}_{n,T}\left(\tau\right)$ is $o_p\left(1\right)$ uniformly in $\tau\in[0,1]$.

So, we have that the first term on the right of (7.16) is governed by the first term of (7.17). Likewise, the second term on the right of (7.16) is governed by

$$\frac{\tau}{(nT)^{1/2}}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it},$$

which implies that it suffices to examine the tightness of

$$\frac{1}{(nT)^{1/2}}\left(\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}-\tau\sum_{t=1}^{T}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right).$$

We now show that the last displayed expression is tight. Indeed, using $C1$ we obtain that

$$\begin{aligned} & E\left\{\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^2\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_2]+1}^{[T\tau_3]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^2\right\} \\ =\ & E\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^2 E\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_2]+1}^{[T\tau_3]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^2 \\ \le\ & C\left(\tau_3-\tau_1\right)^2 \end{aligned}$$

since $C1$ implies, say, that

$$\begin{aligned} E\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^2 &= \frac{1}{T}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}E\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^2 \\ &\le C\left(\tau_2-\tau_1\right). \end{aligned}$$

and $\left(\tau_2-\tau_1\right)\left(\tau_3-\tau_2\right)\le\left(\tau_3-\tau_1\right)^2$. This completes the proof of part (**a**).

We now focus on the proof of part (**b**). We shall first examine the behaviour of

$$\breve{\mathbb{X}}_{n,T}\left(\tau\right)=\left(\frac{n}{T}\right)^{1/2}\sum_{t=1}^{[T\tau]}\left(\widehat{\beta}_t-\beta_t\right).$$

Denoting $\widehat{\Sigma}_x = \frac{1}{n}\sum_{i=1}^{n} \widetilde{x}_{it}\widetilde{x}_{it}'$, we have that

$$\widehat{\beta}_t - \beta_t = \Sigma_x^{-1}\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it} - \left(\Sigma_x^{-1} - \widehat{\Sigma}_x^{-1}\right)\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}. \tag{7.19}$$

The contribution of the second term on the right of (7.19) into $\breve{\mathbb{X}}_{n,T}(\tau)$ is

$$\begin{aligned}&\frac{1}{T^{1/2}}\sum_{t=1}^{[T\tau]}\Sigma_x^{-1}\left(\Sigma_x - \widehat{\Sigma}_x\right)\Sigma_x^{-1}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it} \\ &+\frac{1}{T^{1/2}}\sum_{t=1}^{[T\tau]}\Sigma_x^{-1}\left(\Sigma_x - \widehat{\Sigma}_x\right)\left(\widehat{\Sigma}_x^{-1} - \Sigma_x^{-1}\right)\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}.\end{aligned} \tag{7.20}$$

The second term of (7.20) is easily seen to be $O_p\left(T^{1/2}/n\right)$ uniformly in $\tau \in [0,1]$, because $C2$ and $C4$ imply that $E\left\|\widehat{\Sigma}_x - \Sigma_x\right\|^2 = O\left(n^{-1}\right)$ and $E\left\|\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\|^2 = O(n)$ by Theorem 1.

We next examine the first term of (7.20). To that end, we note that it does not make a difference whether we consider $\tilde{x}_{it}$ or $\mathring{x}_{it}$. Now, because $C1$ implies that $E(u_{it}u_{js}) = 0$ if $t \neq s$, the second moment of the first term of (7.20) is bounded by

$$\frac{\underline{\lambda}^{-2}(\Sigma_x)}{Tn^3}\sum_{t=1+[T\tau_1]}^{[T\tau_2]} E\left\{\sum_{i=1}^{n}\left\{\mathring{x}_{it}\mathring{x}_{it}' - \Sigma_{x,i}\right\}\left(\sum_{i,j=1}^{n}\mathring{x}_{it}\mathring{x}_{jt}'u_{it}u_{jt}\right)\sum_{i=1}^{n}\left\{\mathring{x}_{it}\mathring{x}_{it}' - \Sigma_{x,i}\right\}\right\}, \tag{7.21}$$

where $\underline{\lambda}(\Sigma_x)$ denotes the minimum eigenvalue of the matrix Σ_x. Now, by definition and $C1$– $C3$, we have that the expectation term on the right of (7.21) is

$$\begin{aligned}&\sum_{i,j=1}^{n}\varphi_u(i,j)\sum_{k,\ell=1}^{n}E\left\{\left(\mathring{x}_{kt}^2 - \Sigma_{x,k}\right)\left(\mathring{x}_{\ell t}^2 - \Sigma_{x,\ell}\right)\mathring{x}_{it}\mathring{x}_{jt}\right\} \\ =\ &\sum_{i,j=1}^{n}\varphi_x(i,j)\varphi_u(i,j)\sum_{k,\ell=1}^{n}E\left\{\left(\mathring{x}_{kt}^2 - \Sigma_{x,k}\right)\left(\mathring{x}_{\ell t}^2 - \Sigma_{x,\ell}\right)\right\} \\ &+2\sum_{i,j=1}^{n}\varphi_u(i,j)\sum_{k,\ell=1}^{n}E\left\{\left(\mathring{x}_{kt}^2 - \Sigma_{x,k}\right)\mathring{x}_{it}\right\}E\left\{\left(\mathring{x}_{\ell t}^2 - \Sigma_{x,\ell}\right)\mathring{x}_{jt}\right\} \\ &+\sum_{i,j=1}^{n}\varphi_u(i,j)\sum_{k,\ell=1}^{n}Cum\left(\mathring{x}_{it};\mathring{x}_{jt};\mathring{x}_{kt}^2;\mathring{x}_{\ell t}^2\right),\end{aligned} \tag{7.22}$$

where for notational simplicity we have assumed x_{it} scalar.

By standard algebra and $C2$, we have that

$$\begin{aligned}E\left\{\left(\mathring{x}_{kt}^2 - \Sigma_{x,k}\right)\left(\mathring{x}_{\ell t}^2 - \Sigma_{x,\ell}\right)\right\} &= 2\varphi_x^2(k,\ell) + Cum\left(\mathring{x}_{kt};\mathring{x}_{kt};\mathring{x}_{\ell t};\mathring{x}_{\ell t}\right) \\ E\left\{\left(\mathring{x}_{kt}\mathring{x}_{kt}' - \Sigma_{x,k}\right)\mathring{x}_{it}\right\} &= E\left(\varepsilon_{jt}^3\right)\sum_{j=0}^{\infty}b_j^2(k)\,b_j(i) \\ Cum\left(\mathring{x}_{it};\mathring{x}_{jt};\mathring{x}_{kt}^2;\mathring{x}_{\ell t}^2\right) &= C\sum_{s=0}^{\infty}b_s(i)\,b_s(j)\,b_s^2(k)\,b_s^2(\ell).\end{aligned} \tag{7.23}$$

Thus, using (7.23) we conclude that the right of (7.22) is bounded by

$$C\left\{n^{3-\varsigma}+\sum_{i,j=1}^{n}\varphi_u(i,j)\sum_{k,\ell=1}^{n}\left(\sum_{c=0}^{\infty}b_c^2(k)\,b_c(i)\sum_{c=0}^{\infty}b_c^2(\ell)\,b_c(j)\right)\right.$$
$$\left.+\sum_{i,j=1}^{n}\varphi_u(i,j)\sum_{k,\ell=1}^{n}\sum_{s=0}^{\infty}b_s(i)\,b_s(j)\,b_s^2(k)\,b_s^2(\ell)\right\},$$

because $\sum_{k,\ell=1}^{n}\varphi_x^2(k,\ell)=O\left(n^{2-\varsigma}\right)$ for some $\varsigma>0$.

From here, we obtain that (7.21) is proportional to

$$(\tau_2-\tau_1)\left(\frac{1}{n^{\varsigma}}+\frac{1}{n^{1/2}}\right), \tag{7.24}$$

because $\sum_{c=0}^{\infty}b_c^2(k)<C$, $\sum_{i,j=1}^{n}\varphi_u^2(i,j)\le Cn$ and $\sup_{c\ge1}\sum_{k=1}^{n}b_c^2(k)<C$ implies

$$\sum_{i,j=1}^{n}\varphi_u(i,j)\sum_{k,\ell=1}^{n}\left(\sum_{c=0}^{\infty}b_c^2(k)\,b_c(i)\sum_{c=0}^{\infty}b_c^2(\ell)\,b_c(j)\right)\le Cn^{5/2}.$$

So we conclude that the first term of (7.20) converges to zero in probability uniformly in $\tau\in[0,1]$, since (7.24) indicates that (7.20) is tight.

So, we conclude that the contribution due to the second term on the right of (7.19) to $\mathbb{X}_{n,T}(\tau)$ is negligible, so that it remains to examine

$$\widetilde{\mathbb{X}}_{n,T}(\tau)=\Sigma_x^{-1}\frac{1}{T^{1/2}}\sum_{t=1}^{[T\tau]}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}\right\}.$$

But $\widetilde{\mathbb{X}}_{n,T}(\tau)$ is essentially (7.17), and hence proceeding as in part (**a**),

$$\widetilde{\mathbb{X}}_{n,T}(\tau)\overset{weakly}{\to}\Sigma_x^{-1}V_1^{1/2}\mathcal{B}(\tau).$$

From here the proof of part (**b**) follows since $\breve{\mathbb{X}}_{n,T}(\tau)=\widetilde{\mathbb{X}}_{n,T}(\tau)-\tau\widetilde{\mathbb{X}}_{n,T}(1)+o_p(1)$. □

7.4. PROOF OF PROPOSITION 3.

Proof. Part (**a**). Standard algebra implies that $\mathcal{T}^{\dagger}(r)$ in (3.9) is

$$\frac{1}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}\left(\mathring{y}_{it}-\widehat{\beta}_{FE}'\mathring{x}_{it}\right)+o_p(1)$$
$$\overset{asy}{\simeq}\left\{\frac{1}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}-\tau\frac{1}{(nT)^{1/2}}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}\right\}$$
$$-\frac{1}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}\mathring{x}_{it}'\left\{\left(\sum_{s=1}^{T}\sum_{i=1}^{n}\mathring{x}_{is}\mathring{x}_{is}'\right)^{-1}\sum_{s=1}^{T}\sum_{i=1}^{n}\mathring{x}_{is}\mathring{x}_{is}'(\beta_s-\beta_t)\right\}$$
$$+o_p(1).$$

Now, the second term on the right of the last displayed expression is asymptotically equivalent to

$$
\begin{aligned}
& -\frac{\delta}{nT}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}\mathring{x}_{it}'\left\{\Sigma_x^{-1}\frac{1}{nT}\sum_{s=1}^{T}\sum_{i=1}^{n}\mathring{x}_{is}\mathring{x}_{is}'\mathcal{I}\left(s>t_0\right)-\mathcal{I}\left(t>t_0\right)\right\} \\
= \; & \delta\Sigma_x\frac{1}{T}\sum_{t=1}^{[T\tau]}\left\{\mathcal{I}\left(t>t_0\right)-\left(1-\tau_0\right)\right\}\left(1+o_p\left(1\right)\right) \qquad (7.25) \\
= \; & \delta\Sigma_x\Xi\left(\tau\right)\left(1+o_p\left(1\right)\right),
\end{aligned}
$$

where $\Xi(\tau)$ is the shift function given in (3.12).

Part (**b**). The proof proceeds similarly to that of part (**a**), and is therefore omitted. □

7.5. PROOF OF THEOREM 4.

Proof. Our aim is to show that

$$
\frac{1}{T^{1/2}}\sum_{t=1}^{[T\tau]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^{*}\overset{d^*}{\to}\mathcal{N}\left(0,\tau V_1\right) \quad \textit{(in probability)}. \tag{7.26}
$$

To simplify the notation, we shall examine the behaviour of a typical component, say the ℓth, of the left side of (7.26). Observing that $n^{-1/2}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}^{*}$ is a sequence of independent random variables, we have in view of Theorem 2 of Phillips and Moon (1999), that it suffices to show that

$$
\begin{aligned}
&(\mathbf{i}) \quad E^{*}\left(\frac{1}{T^{1/2}}\sum_{t=1}^{[T\tau]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}^{*}\right)^{2}\overset{P}{\to}\tau V_{1,\ell\ell} \\
&(\mathbf{ii}) \quad E^{*}\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}^{*}\right)^{4}=H_{Tn},
\end{aligned}
$$

where in what follows H_{Tn} is a sequence of $O_p(1)$ random variables. Note that (**ii**) is a sufficient condition for the Lindeberg's condition.

We begin with (**i**). By definition, the left side of the expression in (**i**) is

$$
\begin{aligned}
& \frac{1}{T}\sum_{t=1}^{[T\tau]}\frac{1}{n}\sum_{i,j=1}^{n}\left\{\mathring{x}_{it}\left(\ell\right)\mathring{x}_{jt}\left(\ell\right)E^{*}\left(u_{it}^{*}u_{jt}^{*}\right)\right\} \qquad (7.27) \\
= \; & \frac{1}{T}\sum_{t=1}^{[T\tau]}\frac{1}{n}\sum_{i,j=1}^{n}\left\{\mathring{x}_{it}\left(\ell\right)\mathring{x}_{jt}\left(\ell\right)\frac{1}{T}\sum_{s=1}^{T}\left(\widehat{u}_{is}\widehat{u}_{js}\right)\right\}.
\end{aligned}
$$

As usual, we shall only handle the case when $\widehat{u}_{it}$ is replaced by u_{it}, as the difference is of a smaller probability order of magnitude. With this replacement in mind and using standard algebra, we have that the right side of (7.27) becomes

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{[T\tau]}\frac{1}{n}\sum_{i,j=1}^{n}\left\{\mathring{x}_{it}(\ell)\,\mathring{x}_{jt}(\ell)-\varphi_x(i,j)\right\}\frac{1}{T}\sum_{s=1}^{T}\left\{u_{is}u_{js}-\varphi_u(i,j)\right\}\\
&+\frac{1}{T}\sum_{t=1}^{[T\tau]}\frac{1}{n}\sum_{i,j=1}^{n}\varphi_u(i,j)\left\{\mathring{x}_{it}(\ell)\,\mathring{x}_{jt}(\ell)-\varphi_x(i,j)\right\} \qquad (7.28)\\
&+\frac{1}{T}\sum_{t=1}^{[T\tau]}\frac{1}{n}\sum_{i,j=1}^{n}\varphi_x(i,j)\,\frac{1}{T}\sum_{s=1}^{T}\left\{u_{is}u_{js}-\varphi_u(i,j)\right\}\\
&+\frac{1}{T}\sum_{t=1}^{[T\tau]}\frac{1}{n}\sum_{i,j=1}^{n}\varphi_x(i,j)\,\varphi_u(i,j)\,.
\end{aligned}$$

The last term of (7.28) converges to $\tau V_{1,\ell\ell}$, so to complete the proof we need to show that the first three terms of (7.28) converge to 0 in probability. Because the second and third terms are similar, we shall only handle the third, which equals

$$\frac{\tau}{T}\sum_{s=1}^{T}\frac{1}{n}\sum_{i,j=1}^{n}\left\{\varphi_x(i,j)\left(u_{is}u_{js}-\varphi_u(i,j)\right)\right\}. \qquad (7.29)$$

The first moment of (7.29) is zero, while the second moment is

$$\frac{\tau^2}{T^2}\sum_{s=1}^{T}E\left(\frac{1}{n}\sum_{i,j=1}^{n}\varphi_x(i,j)\left\{u_{is}u_{js}-\varphi_u(i,j)\right\}\right)^2$$

which is

$$\begin{aligned}
&\frac{\tau^2}{Tn^2}\sum_{i_1,\dots,i_4=1}^{n}\varphi_x(i_1,i_2)\,\varphi_x(i_3,i_4)\,\varphi_u(i_1,i_3)\,\varphi_u(i_2,i_4)\\
&+\frac{\tau^2}{Tn^2}\sum_{i_1,\dots,i_4=1}^{n}\varphi_x(i_1,i_2)\,\varphi_x(i_3,i_4)\,\varphi_u(i_1,i_4)\,\varphi_u(i_2,i_3)\\
&+\frac{\tau^2}{Tn^2}\sum_{i_1,\dots,i_4=1}^{n}\varphi_x(i_1,i_2)\,\varphi_x(i_3,i_4)\,cum\left(u_{i_1 1};u_{i_2 1};u_{i_3 1};u_{i_4 1}\right).
\end{aligned}$$

The first two terms on the right of the last displayed expression converge to zero proceeding as with the proof of (7.10), whereas $C1$ implies that the last term is

$$\frac{\kappa_{\varepsilon,4}\tau^2}{Tn^2}\sum_{i_1,\dots,i_4=1}^{n}\varphi_x(i_1,i_2)\,\varphi_x(i_3,i_4)\sum_{\ell=1}^{\infty}a_\ell(i_1)\,a_\ell(i_2)\,a_\ell(i_3)\,a_\ell(i_4)\,.$$

Now using $C1$ and $C3$, we conclude that the last displayed expression is $o(1)$, which concludes the proof of part (**i**).

We now turn to part (**ii**). By definition, we have

$$
\begin{aligned}
& E^{*}\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^{*}\right)^{4} \\
= \; & \frac{1}{n^{2}}\sum_{i_{1},\ldots,i_{4}=1}^{n}\mathring{x}_{i_{1}t}\left(\ell\right)\mathring{x}_{i_{2}t}\left(\ell\right)\mathring{x}_{i_{3}t}\left(\ell\right)\mathring{x}_{i_{4}t}\left(\ell\right)\frac{1}{T}\sum_{s=1}^{T}\widehat{u}_{i_{1}s}\widehat{u}_{i_{2}s}\widehat{u}_{i_{3}s}\widehat{u}_{i_{4}s} \\
\simeq \; & \frac{1}{n^{2}}\sum_{i_{1},\ldots,i_{4}=1}^{n}\mathring{x}_{i_{1}t}\left(\ell\right)\mathring{x}_{i_{2}t}\left(\ell\right)\mathring{x}_{i_{3}t}\left(\ell\right)\mathring{x}_{i_{4}t}\frac{1}{T}\sum_{s=1}^{T}u_{i_{1}s}u_{i_{2}s}u_{i_{3}s}u_{i_{4}s} \\
= \; & \frac{1}{T}\sum_{s=1}^{T}\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^{4}
\end{aligned}
$$

by standard arguments. So, we need to show that the right side of the last displayed equation is bounded in probability. But as it was shown in part (**i**), the first moment is clearly finite. Using Serfling (1980), we conclude that $E^{*}\left(n^{-1/2}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}^{*}\right)^{4}=H_{Tn}$. The latter is the case because by Theorem 1 we have that

$$
\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\overset{d}{\rightarrow}\mathcal{N}\left(0,V_{1,\ell\ell}\right),
$$

and $\left|\mathring{x}_{it}\left(\ell\right)u_{it}\right|^{4}$ is an Uniformly Integrable sequence, so that we can conclude that $E\left(n^{-1/2}\sum_{i=1}^{n}\mathring{x}_{it}\left(\ell\right)u_{it}\right)^{4}\rightarrow 3V_{1,\ell\ell}^{2}$ by Serfling (1980). □

7.6. PROOF OF PROPOSITION 4.

Proof. For notational simplicity, assume that $\widetilde{x}_{it}$ is scalar. Then, by definition and using standard algebra, we obtain

$$
\begin{aligned}
\widehat{V}_{1}^{*}-V_{1} &= \frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}\widetilde{u}_{it}^{*}-\left(\widehat{\beta}_{FE}^{*}-\widehat{\beta}\right)\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}^{2}\right\}^{2}-V_{1} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^{*}\right\}^{2}+n\left(\widehat{\beta}_{FE}^{*}-\widehat{\beta}\right)^{2}\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}^{2}\right\}^{2} \\
&\quad -2n^{1/2}\left(\widehat{\beta}_{FE}^{*}-\widehat{\beta}\right)\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^{*}\right\}\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}^{2}\right\}-V_{1}.
\end{aligned}
$$

The second and third term are negligible as Theorem 4 implies that $\widehat{\beta}_{FE}^{*}-\widehat{\beta}=O_{p^{*}}\left(n^{-1/2}T^{-1/2}\right)$. The third term, for instance, has by Cauchy-Schwartz's inequality an absolute bootstrap moment bounded by

$$
O_{p^{*}}\left(T^{-1/2}\right)\frac{1}{T}\sum_{t=1}^{T}\left(E^{*}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^{*}\right\}^{2}\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}^{2}\right\}^{2}\right)^{1/2}.
$$

By Markov's inequality then we conclude that the third term (and similarly second term) is $o_{p^*}(1)$. Therefore

$$\widehat{V}_1^* - V_1 = \left[\frac{1}{T} \sum_{t=1}^{T} \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \widetilde{x}_{it} u_{it}^* \right\}^2 - V_1 \right] + o_{p^*}(1). \tag{7.30}$$

Now, the term in square brackets on the right of (7.30) is

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^{T} \left(\left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \widetilde{x}_{it} u_{it}^* \right\}^2 - \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}^* \right\}^2 \right) \\ & + \frac{1}{T} \sum_{t=1}^{T} \left(\left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}^* \right\}^2 - E^* \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}^* \right\}^2 \right) \\ & + \frac{1}{T} \sum_{t=1}^{T} E^* \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}^* \right\}^2 - V_1, \end{aligned} \tag{7.31}$$

which converges in probability to zero as we now show. Indeed, by definition, the first line of (7.31) is

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^{T} \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} u_{it}^* \right\}^2 \left\{ \frac{1}{n} \sum_{i=1}^{n} \mathring{x}_{it} \right\}^2 \\ & - \frac{2}{T} \sum_{t=1}^{T} \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} u_{it}^* \right\} \left\{ \frac{1}{n} \sum_{i=1}^{n} \mathring{x}_{it} \right\} \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \mathring{x}_{it} u_{it}^* \right\}. \end{aligned} \tag{7.32}$$

The first bootstrap moment of the first term of (7.32) is

$$\begin{aligned} & \left\{ \frac{1}{n^2} \sum_{i_1, 1_2 = 1}^{n} \mathring{x}_{i_1 t} \mathring{x}_{i_2 t} \right\} \left(\frac{1}{n} \sum_{i,j=1}^{n} \varphi_u(i,j) + \frac{1}{n} \sum_{i,j=1}^{n} \left\{ \frac{1}{T} \sum_{s=1}^{T} u_{is} u_{jt} - \varphi_u(i,j) \right\} \right) \\ = \; & o_p(1) \end{aligned} \tag{7.33}$$

as we now show. Recall that

$$E^* \left(u_{it}^* u_{jt}^* \right) = \frac{1}{T} \sum_{s=1}^{T} \widehat{u}_{is} \widehat{u}_{jt} = \frac{1}{T} \sum_{s=1}^{T} u_{is} u_{jt} \left(1 + O_p \left((nT)^{-1} \right) \right).$$

First, $C2$ and (7.23) imply that

$$\begin{aligned} & \frac{1}{n^4} \sum_{i_1, \ldots, i_4 = 1}^{n} E \left\{ \left(\mathring{x}_{i_1 t} \mathring{x}_{i_2 t} - \varphi_x(i_1, i_2) \right) \left(\mathring{x}_{i_3 t} \mathring{x}_{i_4 t} - \varphi_x(i_3, i_4) \right) \right\} \\ = \; & \frac{2}{n^2} \left(\frac{1}{n} \sum_{i,j=1}^{n} \varphi_x(i,j) \right)^2 + \frac{\kappa_4}{n^4} \sum_{i_1, \ldots, i_4 = 1}^{n} \sum_{\ell=1}^{\infty} b_\ell(i_1) b_\ell(i_2) b_\ell(i_3) b_\ell(i_4). \end{aligned} \tag{7.34}$$

Similarly $C1$ implies that

$$\begin{aligned} & E\left(\frac{1}{n}\sum_{i,j=1}^{n}\left\{\frac{1}{T}\sum_{s=1}^{T}u_{is}u_{jt}-\varphi_{u}\left(i,j\right)\right\}\right)^{2} \\ = & \frac{1}{T}\left(\frac{1}{n}\sum_{i,j=1}^{n}\varphi_{u}\left(i,j\right)\right)^{2}+\frac{\kappa_{4}}{Tn^{2}}\sum_{i_{1},i_{2},i_{3},i_{4}=1}^{n}\sum_{\ell=0}^{\infty}a_{\ell}\left(i_{1}\right)a_{\ell}\left(i_{2}\right)a_{\ell}\left(i_{3}\right)a_{\ell}\left(i_{4}\right)\end{aligned}$$

Next, by $C3$, we observe that

$$\begin{aligned} & \frac{1}{n^{3}}\sum_{j_{1},j_{2}=1}^{n}\varphi_{u}\left(j_{1},j_{2}\right)\sum_{i_{1},i_{2}=1}^{n}\varphi_{x}\left(i_{1},i_{2}\right) \\ \leq & \frac{C}{n}\sum_{j=1}^{n}\varphi_{u}\left(1,j\right)\sum_{i=1}^{n}\varphi_{x}\left(1,i\right)=o\left(\sum_{i=1}^{n}\varphi_{u}\left(1,i\right)\varphi_{x}\left(1,i\right)\right) \\ = & o\left(1\right). \end{aligned} \tag{7.35}$$

after we notice that for sequences $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$, such that $\sum_{i=1}^n a_i b_i < C$ and that $\sum_{i=1}^n a_i^2 < C$ and $\sum_{i=1}^n b_i^\delta < C$ for some $\delta > 0$, we have that

$$\frac{1}{n}\sum_{i=1}^{n}a_{i}\frac{1}{n}\sum_{i=1}^{n}b_{i}=o\left(\frac{1}{n}\sum_{i=1}^{n}a_{i}b_{i}\right).$$

From here (7.33) holds true after we observe also that

$$\begin{aligned} & E\left\{\frac{1}{n^{2}}\sum_{i_{1},1_{2}=1}^{n}\mathring{x}_{i_{1}t}\mathring{x}_{i_{2}t}\right\}E\left|\frac{1}{n}\sum_{i,j=1}^{n}\left\{\frac{1}{T}\sum_{s=1}^{T}u_{is}u_{jt}-\varphi_{u}\left(i,j\right)\right\}\right| \\ \leq & E\left\{\frac{1}{n^{2}}\sum_{i_{1},1_{2}=1}^{n}\mathring{x}_{i_{1}t}\mathring{x}_{i_{2}t}\right\}\left(E\left|\frac{1}{n}\sum_{i,j=1}^{n}\left\{\frac{1}{T}\sum_{s=1}^{T}u_{is}u_{jt}-\varphi_{u}\left(i,j\right)\right\}\right|^{2}\right)^{1/2}\end{aligned}$$

and that by $C3$ and (7.35),

$$E\left\{\frac{1}{n^{2}}\sum_{i_{1},1_{2}=1}^{n}\mathring{x}_{i_{1}t}\mathring{x}_{i_{2}t}\right\}\frac{1}{n}\sum_{i,j=1}^{n}\varphi_{u}\left(i,j\right)=o\left(1\right).$$

The second term of (7.32) has absolute bootstrap moment bounded by

$$\frac{2}{T}\sum_{t=1}^{T}\left(E^{*}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}u_{it}^{*}\right\}^{2}\left\{\frac{1}{n}\sum_{i=1}^{n}\mathring{x}_{it}\right\}^{2}\right)^{1/2}\left(E^{*}\left\{\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^{*}\right\}^{2}\right)^{1/2}$$

using Cauchy-Schwarz's inequality. Our argument for the first term together with the fact that by Theorem 1 $E\left\{\sum_{i=1}^n \mathring{x}_{it}u_{it}\right\}^2 = O(n)$, ensures that that the last displayed expression is also $o(1)$. By Markov's inequality, we then conclude that $(7.32) = o_{p^*}(1)$. The next two lines of (7.31) finally are $o_{p^*}(1)$ by standard arguments. Also observe that

$$\frac{1}{n^{3}}\sum_{i,j=1}^{n}\varphi_{u}\left(i,j\right)\sum_{\ell=1}^{\infty}b_{\ell}\left(i_{1}\right)b_{\ell}\left(i_{2}\right)b_{\ell}\left(i_{3}\right)b_{\ell}\left(i_{4}\right)=o\left(1\right)$$

by $C2$ and $C3$. This concludes the proof. □

7.7. PROOF OF THEOREM 5.

Proof. Because the results of Theorem 4 and Cràmer-Wold's arguments, it remains to show the tightness condition. To simplify the notation, we shall assume that $\mathring{x}_{it}$ is scalar. We shall begin with part (**a**). First, Lemma 2 yields that

$$\mathcal{T}^*(r) = \frac{1}{(nT)^{1/2}}\left(\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^* - \tau\sum_{t=1}^{T}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^*\right)(1+o_{p^*}(1)), \tag{7.36}$$

where the $o_{p^*}(1)$ is uniformly in $\tau \in (0,1)$. Now using (2.9), we have that standard algebra implies that the first term on the right of (7.36) is

$$\frac{1}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^* + \frac{1}{(nT)^{1/2}}\sum_{t=1}^{[T\tau]}\left(n^{1/2}\overline{\mathring{x}}_{\cdot t}\right)\left(n^{1/2}\overline{u}_{\cdot t}^*\right). \tag{7.37}$$

Denoting the second term of (7.37) by $\mathbb{X}_{n,T}^*(\tau)$ and using standard algebra, we have that

$$\begin{aligned} E^*\left(\mathbb{X}_{n,T}^*(\tau_2) - \mathbb{X}_{n,T}^*(\tau_1)\right)^2 &= \frac{1}{n^2T^2}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\left(n^{1/2}\overline{\mathring{x}}_{\cdot t}\right)^2\sum_{i,j=1}^{n}\sum_{s=1}^{T}\widehat{u}_{is}\widehat{u}_{js} \\ &= \frac{1}{n^2T^2}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\left(n^{1/2}\overline{\mathring{x}}_{\cdot t}\right)^2\sum_{i,j=1}^{n}\sum_{s=1}^{T}u_{is}u_{js}(1+o_p(1)). \end{aligned}$$

Dropping the $o_p(1)$ as it is not relevant, we obtain that he expectation of the right side of the last displayed expression is

$$\begin{aligned} \frac{\tau_2-\tau_1}{n}\frac{1}{n}\sum_{i,j=1}^{n}\varphi_x(i,j)\frac{1}{n}\sum_{i,j=1}^{n}\varphi_u(i,j) &= C\frac{\tau_2-\tau_1}{n^{\zeta}} \\ &\le C(\tau_2-\tau_1)^{1+\xi\zeta} \end{aligned}$$

in view of (2.7), because $T^{\xi} = o(n)$ and, as usual, we can take $T^{-1} \le (\tau_2 - \tau_1)$.

Next, to ensure that $\mathbb{X}_{n,T}^*(\tau)$ is tight (in probability) it suffices to show that the second moment of

$$\frac{1}{nT}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\left(\left(n^{1/2}\overline{\mathring{x}}_{\cdot t}\right)^2\frac{1}{n}\sum_{i,j=1}^{n}\frac{1}{T}\sum_{s=1}^{T}u_{is}u_{js} - \frac{1}{n}\sum_{i,j=1}^{n}\varphi_x(i,j)\frac{1}{n}\sum_{i,j=1}^{n}\varphi_u(i,j)\right)$$

is bounded by $C(\tau_2-\tau_1)^{1+\nu}$ for some $\nu>0$. But by using standard algebra, that moment is

$$\begin{aligned}
&\frac{1}{n^2}E\left\{\frac{1}{nT}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\sum_{i,j=1}^{n}\left(\mathring{x}_{it}\mathring{x}_{jt}-\varphi_x(i,j)\right)\right\}^2 E\left\{\frac{1}{nT}\sum_{s=1}^{T}\sum_{i,j=1}^{n}\left(u_{is}u_{js}-\varphi_u(i,j)\right)\right\}^2\\
&+\frac{1}{n^2}\left\{\frac{1}{n}\sum_{i,j=1}^{n}\varphi_u(i,j)\right\}^2 E\left\{\frac{1}{nT}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\sum_{i,j=1}^{n}\left(\mathring{x}_{it}\mathring{x}_{jt}-\varphi_x(i,j)\right)\right\}^2 \qquad (7.38)\\
&+\frac{1}{n^2}\left\{\frac{(\tau_2-\tau_1)}{n}\sum_{i,j=1}^{n}\varphi_x(i,j)\right\}^2 E\left\{\frac{1}{nT}\sum_{s=1}^{T}\sum_{i,j=1}^{n}\left(u_{is}u_{js}-\varphi_u(i,j)\right)\right\}^2 .
\end{aligned}$$

Let us examine the first term in (7.38). As $C1$ indicates that for all $i\geq 1$, u_{it} and u_{is} are independent, we conclude that the second factor in braces is

$$\begin{aligned}
\frac{1}{T^2n^2}\sum_{s=1}^{T}E\left(\sum_{i,j=1}^{n}\left(u_{is}u_{js}-\varphi_u(i,j)\right)\right)^2 &= \frac{1}{Tn^2}E\left(\sum_{i,j=1}^{n}\left(u_{is}u_{js}-\varphi_u(i,j)\right)\right)^2\\
= \frac{1}{Tn^2}\left(\sum_{i_1,i_2=1}^{n}\varphi_u(i_1,i_2)\right)^2 &+ \frac{1}{Tn^2}\sum_{i_1,i_2,i_3,i_4=1}^{n} cum\left(u_{i_1s},u_{i_2s},u_{i_3s},u_{i_4s}\right).
\end{aligned}$$

Now proceeding similarly with the first factor in braces, we conclude that the first term in (7.38) is bounded by

$$\begin{aligned}
&\frac{(\tau_2-\tau_1)}{T^2}\left(\frac{1}{n^3}\sum_{i_1,i_2=1}^{n}\varphi_u(i_1,i_2)\sum_{i_1,i_2=1}^{n}\varphi_x(i_1,i_2)\right)^2\\
&+\frac{(\tau_2-\tau_1)}{T^2}\left(\frac{1}{n^6}\sum_{i_1,i_2,i_3,i_4=1}^{n} cum\left(u_{i_1s},u_{i_2s},u_{i_3s},u_{i_4s}\right)\sum_{i_1,i_2,i_3,i_4=1}^{n} cum\left(x_{i_1s},x_{i_2s},x_{i_3s},x_{i_4s}\right)\right)
\end{aligned}$$

From here and given $cum\left(x_{i_1s},x_{i_2s},x_{i_3s},x_{i_4s}\right)=\sum_{\ell=1}^{\infty}b_\ell(i_1)\,b_\ell(i_2)\,b_\ell(i_3)\,b_\ell(i_4)$, using our result in (7.35), we conclude that the first term in (7.38) is bounded by $C(\tau_2-\tau_1)^{1+\nu}$ for some $\nu>0$. Proceeding similarly with the last two terms in (7.38), using the fact that the finite distribution converges to zero in probability, we conclude that the second term of (7.37), i.e. $\mathbb{X}_{n,T}(\tau)$ is $o_{p^*}(1)$ uniformly in $\tau\in[0,1]$.

So, the first term on the right of (7.36) is governed by the first term of (7.37). Proceeding similarly with the second term, we conclude that to complete the proof of part (**a**), it suffices to examine the tightness condition of

$$\frac{1}{(nT)^{1/2}}\left(\sum_{t=1}^{[T\tau]}\sum_{i=1}^{n}\mathring{x}_{it}u^*_{it}-\tau\sum_{t=1}^{T}\sum_{i=1}^{n}\mathring{x}_{it}u^*_{it}\right). \qquad (7.39)$$

But (7.39) is tight after we notice that

$$\begin{aligned}
& E^*\left\{\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^*\right)^2\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_2]+1}^{[T\tau_3]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^*\right)^2\right\} \\
= & E^*\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^*\right)^2 E^*\left(\frac{1}{T^{1/2}}\sum_{t=[T\tau_2]+1}^{[T\tau_3]}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^*\right)^2 \\
= & \frac{1}{T}\sum_{t=[T\tau_1]+1}^{[T\tau_2]}E^*\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^*\right)^2\frac{1}{T}\sum_{t=[T\tau_2]+1}^{[T\tau_3]}E^*\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\mathring{x}_{it}u_{it}^*\right)^2 \\
= & (\tau_3-\tau_1)^2 O_p(1).
\end{aligned}$$

because $(\tau_2-\tau_1)(\tau_3-\tau_2)\le(\tau_3-\tau_1)^2$. This completes the proof of part (**a**).

We now focus on the proof of part (**b**). We shall first examine the behaviour of

$$\breve{\mathbb{X}}_{n,T}^*(\tau)=\left(\frac{n}{T}\right)^{1/2}\sum_{t=1}^{[T\tau]}\left(\widehat{\beta}_t^*-\widetilde{\beta}_{FE}\right).$$

By definition, we have that

$$\widehat{\beta}_t^*-\widetilde{\beta}_{FE}=\Sigma_x^{-1}\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^*-\left(\Sigma_x^{-1}-\widehat{\Sigma}_x^{-1}\right)\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^*. \tag{7.40}$$

The contribution due to the second term on the right of (7.40) into $\breve{\mathbb{X}}_{n,T}(\tau)$ is, after standard algebra,

$$\begin{aligned}
& \frac{1}{T^{1/2}}\sum_{t=1}^{[T\tau]}\Sigma_x^{-1}\left(\Sigma_x-\widehat{\Sigma}_x\right)\Sigma_x^{-1}\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^* \qquad (7.41)\\
& +\frac{1}{T^{1/2}}\sum_{t=1}^{[T\tau]}\Sigma_x^{-1}\left(\Sigma_x-\widehat{\Sigma}_x\right)\left(\widehat{\Sigma}_x^{-1}-\Sigma_x^{-1}\right)\frac{1}{n^{1/2}}\sum_{i=1}^{n}\widetilde{x}_{it}u_{it}^*.
\end{aligned}$$

The second term of (7.41) is easily seen to be $O_p\left(T^{1/2}/n\right)$ uniformly in $\tau\in[0,1]$, because $C2$ and $C4$ imply that $E\left\|\widehat{\Sigma}_x-\Sigma_x\right\|^2=O\left(n^{-1}\right)$ and $E^*\left\|\sum_{i=1}^n\widetilde{x}_{it}u_{it}^*\right\|^2=O_p(n)$ by Theorem 4.

We now examine the first term of (7.41). To that end, we note that it does not make a difference whether we consider $\tilde{x}_{it}$ or $\mathring{x}_{it}$. Because $E^*u_{it}^*u_{js}^*=0$ if $t\neq s$, the first term of (7.41) is bounded by

$$\frac{\underline{\lambda}^{-2}(\Sigma_x)}{Tn^3}\sum_{t=1+[T\tau_1]}^{[T\tau_2]}E^*\left\{\sum_{i=1}^{n}\left\{\mathring{x}_{it}\mathring{x}_{it}'-\Sigma_{x,i}\right\}\left(\sum_{i,j=1}^{n}\mathring{x}_{it}\mathring{x}_{jt}'\left(u_{it}^*u_{jt}^*\right)\right)\sum_{i=1}^{n}\left\{\mathring{x}_{it}\mathring{x}_{it}'-\Sigma_{x,i}\right\}\right\}. \tag{7.42}$$

To simplify notation, let x_{it} be scalar. By definition and using $C1$ to $C3$, the expectation term of (7.42) is

$$\sum_{i,j=1}^{n}\left(\frac{1}{T}\sum_{s=1}^{T}u_{is}u_{js}\right)\sum_{k,\ell=1}^{n}\left\{\left(\mathring{x}_{kt}^2-\Sigma_{x,k}\right)\left(\mathring{x}_{\ell t}^2-\Sigma_{x,\ell}\right)\mathring{x}_{it}\mathring{x}_{jt}\right\},$$

where we have replaced $\widehat{u}_{js}$ by u_{js} say, as the difference is negligible. Now proceeding as in part (**a**) and using the arguments with (7.22), we conclude that the first term of (7.41) is tight. But, as a by product, we also have shown that, for any $\tau \in [0,1]$, (7.41) converges in probability to 0. So, (7.41) converges in probability to zero uniformly in $\tau \in [0,1]$ in the bootstrap sense.

We conclude that the contribution due to the second term on the right of (7.40) to $\mathbb{X}^*_{n,T}(\tau)$ is negligible, and it suffices to examine the behaviour of

$$\widetilde{\mathbb{X}}^*_{n,T}(\tau) = \Sigma_x^{-1} \frac{1}{T^{1/2}} \sum_{t=1}^{[T\tau]} \left\{ \frac{1}{n^{1/2}} \sum_{i=1}^{n} \widetilde{x}_{it} u^*_{it} \right\}.$$

As $\widetilde{\mathbb{X}}^*_{n,T}(\tau)$ is essentially (7.37), proceeding as in part (**a**), we can conclude that

$$\widetilde{\mathbb{X}}^*_{n,T}(\tau) \overset{weakly}{\to} \Sigma^{-1} V_1^{1/2} \mathcal{B}(\tau).$$

From here the proof of part (**b**) follows because $\breve{\mathbb{X}}^*_{n,T}(\tau) = \widetilde{\mathbb{X}}^*_{n,T}(\tau) - \tau \widetilde{\mathbb{X}}^*_{n,T}(1) + o_{p^*}(1)$. □

References

[1] Andrews, D.W.K. (2005): "Cross-section regression with common shocks," *Econometrica*, **73**, 1551-1585.

[2] Andrews, D.W.K. (1993): "Tests for parameter instability and structural change with unknown change point," *Econometrica*, **61**, 821-856.

[3] Andrews, D.W.K. and Ploberger, W. (1994): "Optimal test when a nuisance parameter is present only under the alternative," *Econometrica*, **62**, 1383-1414.

[4] Arellano, M. (1987): "Computing robust standard errors for within group estimators," *Oxford Bulletin of Economics and Statistics*, **49**, 431-434.

[5] Bai, J. (2000): "Vector autoregressive models with structural changes in regression coefficients and in variance-covariance matrices," *Annals of Economics and Finance*, **1**, 303-339.

[6] Bai, J. (2009): "Panel data models with interactive fixed effects," *Econometrica*, **77**, 1229-1279.

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

[8] Batchelor, L. I. A, and Reed, H. S. (1924): "Relation of the variability of yields of fruit trees to the accuracy of field trials," *Journal of Agricultural Research*, **12**, 245-83.

[9] Bester, C., Conley, J. and Hansen, C. (2011): "Inference with dependent data using cluster covariance estimators," *Journal of Econometrics*, **165**, 137-151.

[10] Billingsley, P. (1968): Convergence of probability measures. Wiley. New York.

[11] 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.

[12] Chen, X. and Conley, T.G. (2001): "A new semiparametric spatial model for panel time series," *Journal of Econometrics*, **105**, 59-83.

[13] Chernozhukov, V. Chetverikov, D. and Kato, K. (2013): "Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors," *The Annals of Statistics*, **41**, 2786-2819.

[14] Chudik, A., Pesaran, M.H. and Tosetti, E. (2011): "Weak and strong cross-section dependence and estimation of large panels," *Econometrics Journal*, **14**, 45-90.

[15] Cliff, A.D. and Ord, J.K. (1981): Spatial Autocorrelation. London: Pion.

[16] Conley, T.G. (1999): "GMM estimation with cross sectional dependence," *Journal of Econometrics*, **92**, 1-45.

[17] Cressie, N. and Huang, H.-C. (1999): "Classes of nonseparable, spatio-temporal stationary covariance functions," *Journal of the American Statistical Association*, **94**, 1330-1340.

[18] Davidson, J. (1994): Stochastic Limit Theory. Oxford University Press.

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

[20] Estrella, A. (2003) "Critical values and p values of Bessel process distributions: Computation and application of structural break tests," *Econometric Theory*, **19**, 1128-1143.

[21] Fan, J., Hall, P. and Yao, Q. (2007): "To how many simultaneous hypothesis tests can normal, Student's t or bootstrap calibration be applied?," *Journal of the American Statistical Association*, **102**, 1282-1288.

[22] Fernández–Casal, R., Gonzalez-Manteiga, W. and Febrero-Bande, M. (2003): "Flexible spatio-temporal stationary variogram models," *Statistics and Computing,* **13**, 127-136.

[23] Genton, M.G. and Koul, H.L. (2008): "Minimum distance inference in unilateral autoregressive lattice processes," *Statistica Sinica*, **18**, 617-631.

[24] Gneiting, T. (2002): "Nonseparable, stationary covariance functions for space-time data," *Journal of the American Statistical Association*, **97**, 590-600.

[25] Hahn, J. and Kuersteiner, G. (2002): "Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large, " *Econometrica*, **70**, 1639-1657.

[26] Hidalgo, J. (2003): "An alternative bootstrap to moving blocks for time series regression models," *Journal of Econometrics*, **117**, 369-399.

[27] Hidalgo, J. and Seo, M. (2013): "Testing for structural stability in the whole sample," *Journal of Econometrics*, **175**, 84-93.

[28] Hjellvik, V., Chen, and Tjøstheim, D. (2004): "Nonparametric estimation and testing n panels of intercorrelated time series" *Journal of Time Series Analysis*, **25**, 831-872.

[29] Horváth, L. (1993): "The maximum likelihood method for testing change in the parameters of normal observations," *Annals of Statistics*, **21**, 671-680.

[30] Ibragimov, I.A. and Rozanov, Y.A. (1978): Gaussian random processes. Springer. New York.

[31] Im, K., Pesaran, H. and Shin, Y. (2003): "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, **115**, 53-74.

[32] Jenish, N. and Prucha, I.R. (2009): "Central limit theorems and uniform laws of large numbers for arrays of random fields," *Journal of Econometrics*, **150**, 86-98.

[33] Jenish, N. and Prucha, I.R. (2012): "On spatial processes and asymptotic inference under near-epoch dependence," *Journal of Econometrics*, **170**, 178-190.

[34] Juhl, T. and Xiao, Z. (2013): "Nonparametric tests of moment conditional stability," *Econometric Theory*, **29**, 90-114.

[35] Kapoora, M., Kelejian, H.H. and Prucha, I.R. (2007): "Panel data models with spatially correlated error components," *Journal of Econometrics*, **140**, 97-130.

[36] Kelejian, H.H. and Prucha, I.R. (2007): "HAC estimation in a spatial framework," *Journal of Econometrics*, **140**, 131-154.

[37] Kim, M.S. and Sun, Y. (2013): "Heteroskedasticity and spatiotemporal dependence robust inference for linear panel models with fixed effects," *Journal of Econometrics*, **177**, 85-108.

[38] Lee, F.-L. (2004): "Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models," *Econometrica*, **6**, 1899-1925.

[39] Lee, F.-L. and Yu, J. (2010): "Estimation of spatial autoregressive panel data models with fixed effects," *Journal of Econometrics*, **154**, 165-185.

[40] Lee, J. and Robinson, P.M. (2013): "Series estimation under cross-sectional dependence," *Mimeo: London School of Economics*.

[41] Matsuda, Y. and Yajima, Y. (2004): "On testing for separable correlations of multivariate time series," *Journal of Time Series Analysis*, **25**, 501-528.

[42] Mercer, W. B. and Hall, A. D. (1911): "The experimental error of field trials," *Journal of Agricultural Science,* **4**, 107-32.

[43] Mitchell, M.W., Genton, M.G. and Gumpertz, M.L. (2005): "Testing for separability of space–time covariances," *Environmetrics*, **16**, 819-831.

[44] Neumann, M. H. and Kreiss, J.-P. (1998): "Regression-type inference in nonparametric autoregression," *Annals of Statistics*, **26**, 1570-1613.

[45] Pesaran, M.H. (2006): "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica*, **74**, 967-1012.

[46] Pesaran, M.H., Shin, Y. and Smith, R.P. (1999): "Pooled mean group estimation of dynamic heterogeneous panels," *Journal of the American Statistical Association*, **94**, 621-634.

[47] Pesaran, M.H. and Smith, R.P. (1995): "Estimating long run relationships from dynamic heterogeneous panels," *Journal of Econometrics*, **68**, 179-113.

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

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

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

[51] Qu, Z. and Perron, P. (2007): "Estimating and testing structural changes in multivariate regressions," *Econometrica*, **75**, 459-502.

[52] Robinson, P. M. and Hidalgo, F.J. (1997): "Time series regressions with long-range dependence," *The Annals of Statistics*, **24**, 127-140.

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

[54] Sarafidis, V. and Wansbeek, T. (2010): "Cross-sectional dependence in panel data analysis," *Working paper*, 20815, MPRA.

[55] Scott, D.J. (1973): "Central limit theorems for martingales and for processes with stationary increments using a Skorokhod representation approach," *Advances in Applied Probability*, **5**, 119-137.

[56] Serfling, R.J. (1980): Approximation theorems of mathematical statistics. Wiley. New York.

[57] Shorack, G. R. and Wellner, J.A. (2009): "Empirical processes with applications to statistics," *Classics in Applied Mathematics*. SIAM, 59.

[58] Swamy, P.A.V.B. (1970): "Efficient inference in a random coefficient regression model," *Econometrica*, **38**, 311-323.

[59] Whittle, P. (1954): "On stationary processes in the plane," *Biometrika*, **41**, 434-449.

[60] Yu, J., de Jong, R. and Lee, L.-F. (2008): "Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large," *Journal of Econometrics*, **146**, 118-134.

Table 1.1. Size of the slope homogeneity test

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$; No time dependence in regressors: ρ_{z_i}=0

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25,25)$												
$\varepsilon=0.00$			0.072	0.047	0.047	0.048			0.081	0.030	0.051	0.036
$\varepsilon=0.05$	0.004	0.036	0.034	0.019	0.050	0.047	0.008	0.044	0.043	0.014	0.058	0.037
$\varepsilon=0.10$	0.006	0.038	0.050	0.026	0.052	0.046	0.009	0.047	0.059	0.018	0.056	0.039
$(100,25)$												
$\varepsilon=0.00$			0.079	0.073	0.064	0.058			0.074	0.059	0.057	0.053
$\varepsilon=0.05$	0.005	0.010	0.035	0.032	0.040	0.041	0.006	0.011	0.040	0.022	0.049	0.042
$\varepsilon=0.10$	0.008	0.014	0.053	0.046	0.045	0.046	0.007	0.012	0.053	0.034	0.053	0.047
$(25,100)$												
$\varepsilon=0.00$			0.066	0.052	0.049	0.047			0.067	0.047	0.054	0.045
$\varepsilon=0.05$	0.025	0.060	0.035	0.028	0.047	0.052	0.031	0.059	0.035	0.025	0.054	0.045
$\varepsilon=0.10$	0.027	0.062	0.048	0.041	0.045	0.052	0.031	0.060	0.052	0.035	0.053	0.045
$(100,100)$												
$\varepsilon=0.00$			0.057	0.057	0.049	0.051			0.057	0.053	0.051	0.050
$\varepsilon=0.05$	0.028	0.035	0.039	0.035	0.048	0.049	0.026	0.031	0.037	0.032	0.052	0.052
$\varepsilon=0.10$	0.030	0.037	0.052	0.050	0.054	0.052	0.028	0.034	0.047	0.045	0.052	0.054

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25,25)$												
$\varepsilon=0.00$			0.074	0.029	0.047	0.039			0.073	0.014	0.050	0.011
$\varepsilon=0.05$	0.007	0.068	0.038	0.013	0.047	0.035	0.011	0.156	0.045	0.010	0.053	0.012
$\varepsilon=0.10$	0.008	0.073	0.052	0.016	0.049	0.038	0.010	0.158	0.063	0.012	0.054	0.011
$(100,25)$												
$\varepsilon=0.00$			0.076	0.056	0.050	0.052			0.027	0.032	0.052	0.036
$\varepsilon=0.05$	0.005	0.015	0.033	0.022	0.045	0.045	0.008	0.032	0.015	0.014	0.052	0.032
$\varepsilon=0.10$	0.007	0.018	0.048	0.033	0.048	0.048	0.008	0.036	0.020	0.018	0.052	0.033
$(25,100)$												
$\varepsilon=0.00$			0.064	0.037	0.050	0.041			0.062	0.027	0.058	0.010
$\varepsilon=0.05$	0.029	0.082	0.038	0.021	0.051	0.039	0.042	0.119	0.043	0.015	0.057	0.010
$\varepsilon=0.10$	0.029	0.080	0.051	0.029	0.050	0.040	0.035	0.116	0.055	0.020	0.061	0.009
$(100,100)$												
$\varepsilon=0.00$			0.063	0.060	0.053	0.050			0.060	0.042	0.047	0.037
$\varepsilon=0.05$	0.027	0.037	0.036	0.031	0.048	0.047	0.032	0.040	0.036	0.021	0.053	0.036
$\varepsilon=0.10$	0.029	0.041	0.047	0.045	0.050	0.050	0.032	0.043	0.049	0.031	0.051	0.038

Table 1.2. Size of the slope homogeneity test - individual parameters

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$; No time dependence in regressors: ρ_{z_i}=0

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T) $\mathcal{T}_\varepsilon^\Delta$	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ
$(25,25)$												
$\varepsilon = 0.00$	0.041	0.040	0.048	0.048	0.050	0.048	0.047	0.045	0.040	0.042	0.042	0.042
$\varepsilon = 0.05$	0.036	0.031	0.029	0.027	0.053	0.045	0.039	0.037	0.024	0.030	0.042	0.043
$\varepsilon = 0.10$	0.040	0.035	0.040	0.033	0.052	0.045	0.044	0.042	0.030	0.036	0.043	0.045
$(100,25)$												
$\varepsilon = 0.00$	0.023	0.026	0.063	0.066	0.053	0.063	0.025	0.024	0.056	0.056	0.052	0.050
$\varepsilon = 0.05$	0.014	0.015	0.036	0.038	0.051	0.034	0.017	0.015	0.029	0.037	0.052	0.036
$\varepsilon = 0.10$	0.017	0.020	0.046	0.052	0.050	0.046	0.021	0.019	0.039	0.044	0.054	0.043
$(25,100)$												
$\varepsilon = 0.00$	0.055	0.058	0.050	0.056	0.050	0.049	0.058	0.056	0.053	0.046	0.051	0.045
$\varepsilon = 0.05$	0.054	0.056	0.034	0.039	0.048	0.046	0.054	0.052	0.030	0.036	0.044	0.044
$\varepsilon = 0.10$	0.057	0.057	0.044	0.047	0.050	0.048	0.058	0.056	0.040	0.040	0.047	0.044
$(100,100)$												
$\varepsilon = 0.00$	0.039	0.039	0.055	0.055	0.051	0.053	0.040	0.039	0.055	0.052	0.052	0.049
$\varepsilon = 0.05$	0.033	0.039	0.038	0.046	0.049	0.046	0.035	0.034	0.039	0.040	0.050	0.049
$\varepsilon = 0.10$	0.036	0.040	0.050	0.051	0.051	0.051	0.036	0.038	0.048	0.047	0.048	0.054

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T) $\mathcal{T}_\varepsilon^\Delta$	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ
$(25,25)$												
$\varepsilon = 0.00$	0.064	0.066	0.038	0.041	0.043	0.041	0.094	0.125	0.022	0.030	0.013	0.015
$\varepsilon = 0.05$	0.055	0.049	0.021	0.026	0.042	0.039	0.097	0.112	0.015	0.020	0.015	0.014
$\varepsilon = 0.10$	0.059	0.057	0.027	0.033	0.043	0.040	0.098	0.117	0.016	0.022	0.013	0.015
$(100,25)$												
$\varepsilon = 0.00$	0.028	0.026	0.052	0.058	0.053	0.049	0.035	0.043	0.039	0.045	0.040	0.038
$\varepsilon = 0.05$	0.019	0.018	0.032	0.034	0.052	0.038	0.028	0.030	0.022	0.029	0.039	0.034
$\varepsilon = 0.10$	0.023	0.021	0.040	0.041	0.051	0.042	0.031	0.033	0.026	0.033	0.039	0.035
$(25,100)$												
$\varepsilon = 0.00$	0.071	0.073	0.049	0.046	0.046	0.043	0.084	0.095	0.037	0.034	0.013	0.014
$\varepsilon = 0.05$	0.070	0.069	0.032	0.033	0.040	0.042	0.093	0.091	0.023	0.023	0.012	0.012
$\varepsilon = 0.10$	0.071	0.070	0.041	0.038	0.045	0.043	0.090	0.096	0.029	0.028	0.012	0.013
$(100,100)$												
$\varepsilon = 0.00$	0.042	0.044	0.054	0.061	0.051	0.053	0.042	0.042	0.045	0.047	0.043	0.039
$\varepsilon = 0.05$	0.036	0.038	0.035	0.043	0.047	0.046	0.037	0.038	0.026	0.035	0.038	0.037
$\varepsilon = 0.10$	0.039	0.041	0.045	0.050	0.048	0.046	0.040	0.040	0.035	0.042	0.041	0.038

TABLE 1.3. Power of the slope homogeneity test DGP1 - break in θ

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$; No time dependence in regressors: ρ_{z_i}=0

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.998	0.968	0.998	0.951			0.919	0.637	0.880	0.595
$\varepsilon = 0.05$	0.854	0.827	0.975	0.661	0.986	0.799	0.581	0.630	0.694	0.250	0.740	0.378
$\varepsilon = 0.10$	0.902	0.870	0.987	0.796	0.988	0.842	0.662	0.690	0.782	0.335	0.767	0.410
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.776	0.433	0.715	0.440			0.439	0.171	0.349	0.170
$\varepsilon = 0.05$	0.573	0.570	0.615	0.326	0.871	0.651	0.371	0.413	0.319	0.113	0.500	0.257
$\varepsilon = 0.10$	0.653	0.638	0.755	0.456	0.893	0.685	0.430	0.471	0.422	0.167	0.523	0.280
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.863	0.433	0.797	0.438			0.458	0.072	0.378	0.036
$\varepsilon = 0.05$	0.240	0.410	0.566	0.127	0.604	0.229	0.047	0.290	0.177	0.021	0.207	0.016
$\varepsilon = 0.10$	0.310	0.467	0.676	0.190	0.651	0.262	0.059	0.316	0.261	0.029	0.242	0.020
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.353	0.101	0.273	0.126			0.172	0.027	0.127	0.018
$\varepsilon = 0.05$	0.142	0.250	0.249	0.061	0.376	0.157	0.051	0.225	0.112	0.017	0.149	0.017
$\varepsilon = 0.10$	0.180	0.285	0.340	0.087	0.410	0.174	0.055	0.245	0.165	0.021	0.175	0.018
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			0.998	0.992	0.997	0.985
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.997	0.992	0.991	0.998	0.997
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.998	0.997	0.997	0.999	0.998

Table 1.4. Power of the slope homogeneity test DGP1 - Individual parameters

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$; No time dependence in regressors: ρ_{z_i}=0

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T) $\mathcal{T}^{\Delta}_{\varepsilon}$	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.991	0.049	0.997	0.045	0.994	0.036	0.834	0.051	0.873	0.041	0.820	0.034
$\varepsilon = 0.05$	0.949	0.044	0.976	0.027	0.971	0.037	0.641	0.036	0.646	0.025	0.640	0.029
$\varepsilon = 0.10$	0.967	0.048	0.984	0.032	0.977	0.039	0.698	0.042	0.716	0.031	0.681	0.031
$(100, 100)$												
$\varepsilon = 0.00$	1.000	0.043	1.000	0.053	1.000	0.036	1.000	0.046	1.000	0.052	1.000	0.040
$\varepsilon = 0.05$	1.000	0.045	1.000	0.039	1.000	0.031	1.000	0.041	1.000	0.039	1.000	0.035
$\varepsilon = 0.10$	1.000	0.045	1.000	0.048	1.000	0.037	1.000	0.044	1.000	0.046	1.000	0.037
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.710	0.041	0.753	0.047	0.752	0.041	0.369	0.042	0.365	0.040	0.340	0.036
$\varepsilon = 0.05$	0.795	0.029	0.575	0.030	0.856	0.042	0.432	0.026	0.269	0.028	0.437	0.035
$\varepsilon = 0.10$	0.840	0.032	0.688	0.035	0.878	0.042	0.481	0.033	0.348	0.032	0.471	0.036
$(100, 100)$												
$\varepsilon = 0.00$	1.000	0.040	1.000	0.052	1.000	0.041	1.000	0.038	1.000	0.051	1.000	0.041
$\varepsilon = 0.05$	1.000	0.031	1.000	0.043	1.000	0.039	1.000	0.028	1.000	0.044	1.000	0.040
$\varepsilon = 0.10$	1.000	0.034	1.000	0.054	1.000	0.041	1.000	0.033	1.000	0.045	1.000	0.043

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T) $\mathcal{T}^{\Delta}_{\varepsilon}$	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.757	0.063	0.731	0.037	0.682	0.031	0.417	0.109	0.206	0.028	0.106	0.014
$\varepsilon = 0.05$	0.555	0.047	0.451	0.024	0.485	0.029	0.293	0.087	0.075	0.020	0.042	0.010
$\varepsilon = 0.10$	0.613	0.055	0.524	0.029	0.528	0.030	0.324	0.095	0.102	0.023	0.052	0.011
$(100, 100)$												
$\varepsilon = 0.00$	1.000	0.048	1.000	0.054	1.000	0.040	1.000	0.045	1.000	0.046	1.000	0.029
$\varepsilon = 0.05$	1.000	0.043	1.000	0.036	1.000	0.034	1.000	0.038	1.000	0.032	1.000	0.028
$\varepsilon = 0.10$	1.000	0.048	1.000	0.049	1.000	0.036	1.000	0.043	1.000	0.037	1.000	0.028
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.327	0.061	0.242	0.039	0.243	0.037	0.204	0.116	0.066	0.030	0.035	0.016
$\varepsilon = 0.05$	0.361	0.037	0.167	0.025	0.300	0.031	0.217	0.083	0.041	0.021	0.035	0.014
$\varepsilon = 0.10$	0.401	0.047	0.226	0.028	0.331	0.033	0.235	0.093	0.056	0.024	0.040	0.014
$(100, 100)$												
$\varepsilon = 0.00$	1.000	0.044	1.000	0.060	1.000	0.042	0.998	0.040	0.999	0.043	0.998	0.030
$\varepsilon = 0.05$	1.000	0.030	1.000	0.044	1.000	0.038	1.000	0.028	0.997	0.034	1.000	0.031
$\varepsilon = 0.10$	1.000	0.036	1.000	0.053	1.000	0.039	1.000	0.033	0.999	0.040	1.000	0.030

Table 1.5. Power of the slope homogeneity test DGP2 - break in ρ

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$; No time dependence in regressors: ρ_{z_i}=0

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.324	0.212	0.273	0.222			0.203	0.095	0.151	0.108
$\varepsilon = 0.05$	0.043	0.120	0.156	0.067	0.193	0.152	0.024	0.085	0.098	0.029	0.126	0.079
$\varepsilon = 0.10$	0.058	0.141	0.210	0.102	0.214	0.169	0.033	0.101	0.136	0.044	0.140	0.083
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			0.987	0.984	0.985	0.982
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	0.969	0.969	0.952	0.940	0.965	0.958
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	0.976	0.976	0.967	0.958	0.971	0.966
$t_0 = [0.8T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.155	0.077	0.106	0.079			0.140	0.051	0.101	0.059
$\varepsilon = 0.05$	0.032	0.052	0.115	0.033	0.144	0.078	0.029	0.059	0.098	0.022	0.120	0.055
$\varepsilon = 0.10$	0.042	0.066	0.152	0.047	0.150	0.085	0.036	0.067	0.126	0.030	0.123	0.058
$(100,100)$												
$\varepsilon = 0.00$			0.985	0.974	0.981	0.967			0.890	0.848	0.884	0.841
$\varepsilon = 0.05$	0.987	0.980	0.984	0.979	0.991	0.988	0.856	0.817	0.861	0.825	0.897	0.862
$\varepsilon = 0.10$	0.990	0.986	0.992	0.989	0.994	0.989	0.874	0.842	0.898	0.864	0.912	0.880

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.179	0.077	0.131	0.097			0.127	0.024	0.090	0.019
$\varepsilon = 0.05$	0.020	0.114	0.090	0.022	0.106	0.063	0.021	0.188	0.064	0.013	0.082	0.015
$\varepsilon = 0.10$	0.027	0.127	0.121	0.033	0.116	0.069	0.019	0.196	0.098	0.016	0.086	0.015
$(100,100)$												
$\varepsilon = 0.00$			0.974	0.965	0.971	0.960			0.542	0.473	0.499	0.440
$\varepsilon = 0.05$	0.902	0.904	0.918	0.890	0.933	0.921	0.319	0.343	0.338	0.264	0.389	0.333
$\varepsilon = 0.10$	0.918	0.920	0.943	0.923	0.949	0.932	0.346	0.377	0.408	0.338	0.413	0.360
$t_0 = [0.8T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.115	0.040	0.076	0.051			0.099	0.018	0.093	0.013
$\varepsilon = 0.05$	0.022	0.081	0.084	0.017	0.094	0.045	0.023	0.167	0.068	0.013	0.117	0.013
$\varepsilon = 0.10$	0.026	0.092	0.107	0.022	0.100	0.050	0.023	0.176	0.091	0.014	0.122	0.013
$(100,100)$												
$\varepsilon = 0.00$			0.721	0.629	0.696	0.585			0.267	0.167	0.230	0.139
$\varepsilon = 0.05$	0.729	0.662	0.728	0.649	0.782	0.708	0.257	0.186	0.254	0.135	0.303	0.183
$\varepsilon = 0.10$	0.748	0.693	0.783	0.711	0.798	0.728	0.265	0.208	0.308	0.185	0.316	0.197

TABLE 1.6. Power of the slope homogeneity test DGP2 - individual parameters

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$; No time dependence in regressors: ρ_{z_i}=0

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T) $\mathcal{T}^{\Delta}_{\varepsilon}$	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.042	0.286	0.049	0.314	0.050	0.306	0.047	0.174	0.039	0.175	0.042	0.167
$\varepsilon = 0.05$	0.036	0.182	0.030	0.158	0.053	0.233	0.039	0.111	0.023	0.098	0.040	0.130
$\varepsilon = 0.10$	0.039	0.212	0.038	0.201	0.051	0.247	0.043	0.134	0.032	0.122	0.041	0.140
$(100, 100)$												
$\varepsilon = 0.00$	0.035	1.000	0.054	1.000	0.048	1.000	0.038	0.991	0.056	0.993	0.052	0.993
$\varepsilon = 0.05$	0.033	1.000	0.037	1.000	0.051	1.000	0.034	0.974	0.040	0.977	0.048	0.981
$\varepsilon = 0.10$	0.036	1.000	0.051	1.000	0.053	1.000	0.035	0.979	0.079	0.983	0.049	0.986
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.042	0.091	0.050	0.117	0.051	0.105	0.045	0.073	0.042	0.074	0.043	0.073
$\varepsilon = 0.05$	0.036	0.078	0.030	0.086	0.055	0.116	0.039	0.056	0.023	0.058	0.039	0.074
$\varepsilon = 0.10$	0.040	0.094	0.040	0.107	0.051	0.126	0.043	0.070	0.030	0.065	0.040	0.079
$(100, 100)$												
$\varepsilon = 0.00$	0.038	0.985	0.056	0.992	0.051	0.990	0.040	0.773	0.053	0.816	0.051	0.806
$\varepsilon = 0.05$	0.035	0.992	0.037	0.993	0.051	0.994	0.036	0.830	0.039	0.850	0.051	0.872
$\varepsilon = 0.10$	0.038	0.994	0.051	0.995	0.052	0.995	0.036	0.852	0.048	0.875	0.049	0.889

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T) $\mathcal{T}^{\Delta}_{\varepsilon}$	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ	θ	ρ
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.063	0.183	0.038	0.135	0.043	0.306	0.100	0.193	0.023	0.063	0.015	0.031
$\varepsilon = 0.05$	0.054	0.123	0.022	0.077	0.041	0.233	0.098	0.154	0.016	0.038	0.017	0.021
$\varepsilon = 0.10$	0.058	0.144	0.027	0.090	0.042	0.247	0.100	0.168	0.017	0.044	0.014	0.023
$(100, 100)$												
$\varepsilon = 0.00$	0.040	0.978	0.053	0.983	0.050	0.981	0.040	0.569	0.043	0.582	0.040	0.547
$\varepsilon = 0.05$	0.035	0.949	0.036	0.952	0.043	0.959	0.037	0.445	0.028	0.426	0.038	0.441
$\varepsilon = 0.10$	0.037	0.959	0.043	0.965	0.045	0.967	0.038	0.482	0.035	0.488	0.038	0.462
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$	0.063	0.094	0.037	0.067	0.043	0.061	0.097	0.149	0.022	0.042	0.014	0.020
$\varepsilon = 0.05$	0.055	0.075	0.022	0.047	0.042	0.064	0.098	0.125	0.017	0.029	0.017	0.018
$\varepsilon = 0.10$	0.059	0.088	0.028	0.055	0.043	0.067	0.098	0.137	0.018	0.032	0.014	0.019
$(100, 100)$												
$\varepsilon = 0.00$	0.041	0.712	0.052	0.769	0.050	0.737	0.041	0.232	0.043	0.252	0.041	0.223
$\varepsilon = 0.05$	0.037	0.776	0.036	0.805	0.047	0.813	0.039	0.258	0.029	0.256	0.039	0.275
$\varepsilon = 0.10$	0.038	0.798	0.044	0.833	0.048	0.828	0.040	0.284	0.037	0.292	0.040	0.283

Table 2.1. Size of the slope homogeneity test

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$ apart from fixed individual heterogeneity in z_{it}; No time dependence in regressors: $\rho_{z_i} = 0$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25,25)$												
$\varepsilon = 0.00$			0.073	0.044	0.047	0.031			0.070	0.034	0.047	0.023
$\varepsilon = 0.05$	0.005	0.067	0.032	0.018	0.044	0.018	0.006	0.049	0.034	0.014	0.046	0.013
$\varepsilon = 0.10$	0.006	0.061	0.047	0.023	0.045	0.022	0.007	0.048	0.050	0.019	0.046	0.016
$(100,25)$												
$\varepsilon = 0.00$			0.080	0.061	0.062	0.046			0.069	0.054	0.051	0.048
$\varepsilon = 0.05$	0.006	0.037	0.036	0.027	0.049	0.024	0.005	0.030	0.037	0.023	0.048	0.032
$\varepsilon = 0.10$	0.007	0.031	0.053	0.037	0.053	0.031	0.007	0.027	0.056	0.033	0.050	0.041
$(25,100)$												
$\varepsilon = 0.00$			0.057	0.050	0.052	0.043			0.061	0.049	0.050	0.040
$\varepsilon = 0.05$	0.027	0.068	0.038	0.033	0.048	0.032	0.031	0.052	0.037	0.025	0.055	0.037
$\varepsilon = 0.10$	0.029	0.063	0.047	0.044	0.051	0.037	0.029	0.048	0.052	0.038	0.053	0.038
$(100,100)$												
$\varepsilon = 0.00$			0.059	0.057	0.044	0.044			0.053	0.050	0.044	0.043
$\varepsilon = 0.05$	0.022	0.041	0.032	0.030	0.045	0.032	0.026	0.031	0.035	0.027	0.048	0.035
$\varepsilon = 0.10$	0.026	0.037	0.047	0.043	0.047	0.037	0.027	0.032	0.047	0.040	0.047	0.039

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25,25)$												
$\varepsilon = 0.00$			0.076	0.033	0.055	0.025			0.066	0.019	0.055	0.011
$\varepsilon = 0.05$	0.006	0.058	0.038	0.011	0.051	0.013	0.010	0.078	0.037	0.011	0.055	0.009
$\varepsilon = 0.10$	0.008	0.057	0.051	0.017	0.055	0.018	0.009	0.080	0.050	0.013	0.056	0.009
$(100,25)$												
$\varepsilon = 0.00$			0.077	0.056	0.053	0.046			0.071	0.036	0.048	0.040
$\varepsilon = 0.05$	0.005	0.032	0.034	0.022	0.048	0.037	0.006	0.032	0.038	0.016	0.054	0.054
$\varepsilon = 0.10$	0.007	0.029	0.050	0.036	0.047	0.039	0.008	0.031	0.051	0.020	0.052	0.052
$(25,100)$												
$\varepsilon = 0.00$			0.059	0.048	0.049	0.036			0.059	0.035	0.045	0.020
$\varepsilon = 0.05$	0.030	0.060	0.040	0.026	0.053	0.036	0.037	0.047	0.042	0.016	0.056	0.016
$\varepsilon = 0.10$	0.032	0.055	0.051	0.034	0.054	0.037	0.031	0.045	0.048	0.024	0.050	0.017
$(100,100)$												
$\varepsilon = 0.00$			0.063	0.059	0.050	0.047			0.066	0.048	0.049	0.029
$\varepsilon = 0.05$	0.024	0.035	0.037	0.027	0.048	0.035	0.035	0.024	0.043	0.028	0.060	0.026
$\varepsilon = 0.10$	0.027	0.036	0.055	0.045	0.047	0.039	0.032	0.024	0.062	0.038	0.055	0.027

TABLE 2.2. Power of the slope homogeneity test DGP1 - break in θ

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$ apart from fixed individual heterogeneity in z_{it}; No time dependence in regressors: $\rho_{z_i} = 0$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			1.000	0.981	1.000	0.923			0.998	0.851	0.996	0.621
$\varepsilon = 0.05$	0.995	0.988	1.000	0.117	1.000	0.024	0.816	0.844	0.966	0.203	0.977	0.041
$\varepsilon = 0.10$	0.997	0.992	1.000	0.504	1.000	0.154	0.876	0.881	0.986	0.422	0.983	0.146
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.960	0.515	0.954	0.440			0.774	0.276	0.736	0.191
$\varepsilon = 0.05$	0.929	0.886	0.889	0.181	0.990	0.210	0.599	0.578	0.615	0.122	0.853	0.140
$\varepsilon = 0.10$	0.954	0.918	0.950	0.419	0.993	0.502	0.671	0.634	0.758	0.234	0.883	0.261
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.995	0.763	0.993	0.544			0.989	0.502	0.985	0.149
$\varepsilon = 0.05$	0.723	0.797	0.939	0.136	0.956	0.044	0.692	0.857	0.903	0.081	0.916	0.010
$\varepsilon = 0.10$	0.798	0.836	0.972	0.279	0.968	0.105	0.763	0.888	0.949	0.145	0.943	0.017
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.698	0.215	0.670	0.160			0.728	0.088	0.740	0.039
$\varepsilon = 0.05$	0.501	0.512	0.571	0.093	0.792	0.132	0.572	0.536	0.561	0.050	0.804	0.037
$\varepsilon = 0.10$	0.571	0.567	0.699	0.183	0.830	0.198	0.641	0.589	0.712	0.080	0.851	0.053
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 2.3. Power of the slope homogeneity testDGP2 - break in ρ

No heterogeneity $\sigma^2_{u_i} = \sigma^2_{v_i} = 1$ apart from fixed individual heterogeneity in z_{it}; No time dependence in regressors: $\rho_{z_i} = 0$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.881	0.604	0.845	0.479			0.673	0.386	0.602	0.276
$\varepsilon = 0.05$	0.350	0.589	0.671	0.109	0.716	0.031	0.143	0.392	0.384	0.080	0.428	0.039
$\varepsilon = 0.10$	0.431	0.639	0.761	0.217	0.747	0.113	0.193	0.435	0.480	0.150	0.467	0.083
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.501	0.114	0.392	0.052			0.322	0.067	0.247	0.036
$\varepsilon = 0.05$	0.302	0.118	0.479	0.027	0.596	0.011	0.159	0.086	0.289	0.025	0.375	0.013
$\varepsilon = 0.10$	0.355	0.139	0.593	0.060	0.617	0.021	0.191	0.097	0.374	0.043	0.393	0.022
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.614	0.302	0.582	0.240			0.755	0.251	0.719	0.093
$\varepsilon = 0.05$	0.118	0.366	0.356	0.059	0.394	0.037	0.196	0.529	0.426	0.044	0.487	0.012
$\varepsilon = 0.10$	0.158	0.405	0.444	0.110	0.435	0.073	0.252	0.570	0.551	0.068	0.548	0.019
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.302	0.068	0.245	0.037			0.374	0.038	0.334	0.012
$\varepsilon = 0.05$	0.132	0.090	0.263	0.019	0.340	0.016	0.202	0.159	0.300	0.018	0.421	0.010
$\varepsilon = 0.10$	0.158	0.106	0.341	0.032	0.361	0.022	0.235	0.176	0.397	0.022	0.459	0.011
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.1. Size of the slope homogeneity test

Presence of heterogeneity $\sigma^2_{u_i} \sim IID\chi^2(2)/2, \sigma^2_{z_i} \sim IID\chi^2(1)$;
No time dependence in regressors: $\rho_{z_i} = 0$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25, 25)$												
$\varepsilon = 0.00$			0.075	0.034	0.059	0.046			0.080	0.029	0.054	0.037
$\varepsilon = 0.05$	0.010	0.035	0.040	0.012	0.060	0.040	0.011	0.089	0.048	0.013	0.062	0.031
$\varepsilon = 0.10$	0.011	0.035	0.054	0.017	0.060	0.045	0.011	0.093	0.060	0.017	0.059	0.033
$(100, 25)$												
$\varepsilon = 0.00$			0.070	0.062	0.052	0.050			0.075	0.049	0.052	0.043
$\varepsilon = 0.05$	0.005	0.008	0.031	0.020	0.048	0.043	0.004	0.017	0.033	0.017	0.046	0.036
$\varepsilon = 0.10$	0.006	0.010	0.048	0.033	0.052	0.045	0.006	0.020	0.045	0.025	0.050	0.038
$(25, 100)$												
$\varepsilon = 0.00$			0.066	0.050	0.060	0.048			0.067	0.041	0.055	0.033
$\varepsilon = 0.05$	0.042	0.038	0.049	0.025	0.067	0.047	0.036	0.078	0.041	0.020	0.058	0.030
$\varepsilon = 0.10$	0.040	0.038	0.062	0.036	0.064	0.046	0.031	0.078	0.057	0.028	0.054	0.033
$(100, 100)$												
$\varepsilon = 0.00$			0.056	0.053	0.051	0.055			0.061	0.048	0.052	0.049
$\varepsilon = 0.05$	0.030	0.026	0.039	0.029	0.054	0.049	0.028	0.039	0.036	0.028	0.049	0.047
$\varepsilon = 0.10$	0.030	0.025	0.046	0.041	0.053	0.049	0.029	0.039	0.049	0.038	0.053	0.045

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25, 25)$												
$\varepsilon = 0.00$			0.069	0.022	0.048	0.033			0.069	0.014	0.055	0.014
$\varepsilon = 0.05$	0.008	0.103	0.034	0.011	0.052	0.028	0.012	0.233	0.039	0.009	0.054	0.014
$\varepsilon = 0.10$	0.009	0.107	0.049	0.013	0.050	0.029	0.011	0.232	0.052	0.010	0.055	0.014
$(100, 25)$												
$\varepsilon = 0.00$			0.072	0.055	0.053	0.048			0.074	0.036	0.047	0.039
$\varepsilon = 0.05$	0.007	0.021	0.041	0.022	0.048	0.038	0.008	0.048	0.041	0.016	0.048	0.039
$\varepsilon = 0.10$	0.009	0.024	0.051	0.029	0.047	0.041	0.009	0.050	0.053	0.021	0.050	0.039
$(25, 100)$												
$\varepsilon = 0.00$			0.066	0.038	0.050	0.024			0.065	0.029	0.055	0.009
$\varepsilon = 0.05$	0.036	0.081	0.042	0.018	0.059	0.021	0.046	0.186	0.042	0.016	0.065	0.008
$\varepsilon = 0.10$	0.033	0.082	0.055	0.027	0.054	0.022	0.037	0.179	0.055	0.021	0.057	0.008
$(100, 100)$												
$\varepsilon = 0.00$			0.063	0.053	0.049	0.047			0.064	0.047	0.051	0.035
$\varepsilon = 0.05$	0.030	0.043	0.040	0.032	0.051	0.046	0.036	0.044	0.044	0.024	0.057	0.034
$\varepsilon = 0.10$	0.031	0.043	0.051	0.040	0.051	0.045	0.032	0.046	0.055	0.033	0.053	0.032

TABLE 3.2. Power of the slope homogeneity test DGP1 - break in θ

Presence of heterogeneity $\sigma^2_{u_i} \sim IID\chi^2(2)/2, \sigma^2_{z_i} \sim IID\chi^2(1)$;
No time dependence in regressors: $\rho_{z_i} = 0$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.894	0.650	0.870	0.695			0.834	0.381	0.782	0.327
$\varepsilon = 0.05$	0.360	0.435	0.665	0.167	0.726	0.401	0.213	0.421	0.518	0.102	0.587	0.135
$\varepsilon = 0.10$	0.439	0.492	0.755	0.292	0.759	0.476	0.278	0.466	0.644	0.155	0.639	0.168
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.402	0.108	0.379	0.227			0.349	0.103	0.289	0.103
$\varepsilon = 0.05$	0.180	0.232	0.310	0.075	0.789	0.258	0.133	0.277	0.233	0.054	0.384	0.115
$\varepsilon = 0.10$	0.232	0.270	0.405	0.137	0.526	0318	0.167	0.311	0.344	0.084	0.416	0.134
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.785	0.297	0.739	0.263			0.617	0.098	0.556	0.042
$\varepsilon = 0.05$	0.176	0.399	0.466	0.071	0.530	0.101	0.075	0.420	0.265	0.025	0.328	0.014
$\varepsilon = 0.10$	0.232	0.439	0.573	0.108	0.574	0.130	0.105	0.445	0.384	0.040	0.386	0.018
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25,25)$												
$\varepsilon = 0.00$			0.308	0.077	0.255	0.088			0.227	0.034	0.189	0.021
$\varepsilon = 0.05$	0.113	0.273	0.224	0.039	0.349	0.092	0.071	0.335	0.146	0.020	0.235	0.020
$\varepsilon = 0.10$	0.142	0.303	0.309	0.060	0.380	0.108	0.084	0.350	0.212	0.026	0.267	0.023
$(100,100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.3. Power of the slope homogeneity test DGP2 - break in ρ

Presence of heterogeneity $\sigma_{u_i}^2 \sim IID\chi^2(2)/2, \sigma_{z_i}^2 \sim IID\chi^2(1)$;
No time dependence in regressors: $\rho_{z_i} = 0$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.111	0.061	0.090	0.081			0.148	0.050	0.119	0.063
$\varepsilon = 0.05$	0.018	0.057	0.060	0.017	0.082	0.058	0.020	0.120	0.077	0.016	0.095	0.043
$\varepsilon = 0.10$	0.020	0.059	0.077	0.027	0.084	0.066	0.024	0.130	0.102	0.023	0.098	0.049
$(100, 100)$												
$\varepsilon = 0.00$			0.826	0.826	0.814	0.817			0.967	0.950	0.964	0.951
$\varepsilon = 0.05$	0.617	0.627	0.660	0.639	0.706	0.717	0.871	0.883	0.893	0.858	0.913	0.897
$\varepsilon = 0.10$	0.654	0.665	0.725	0.711	0.738	0.747	0.893	0.900	0.926	0.899	0.929	0.911
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.092	0.040	0.074	0.057			0.110	0.032	0.079	0.041
$\varepsilon = 0.05$	0.020	0.040	0.063	0.013	0.085	0.049	0.024	0.097	0.079	0.014	0.092	0.035
$\varepsilon = 0.10$	0.042	0.042	0.083	0.019	0.088	0.054	0.024	0.099	0.103	0.019	0.094	0.037
$(100, 100)$												
$\varepsilon = 0.00$			0.463	0.376	0.442	0.378			0.675	0.554	0.668	0.552
$\varepsilon = 0.05$	0.453	0.345	0.464	0.375	0.525	0.468	0.679	0.611	0.676	0.568	0.745	0.646
$\varepsilon = 0.10$	0.472	0.376	0.533	0.455	0.543	0.490	0.703	0.647	0.744	0.650	0.768	0.671

	Weak Spatial dependence (polynomial)						Strong Spatial dependence					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.139	0.040	0.106	0.057			0.117	0.021	0.092	0.019
$\varepsilon = 0.05$	0.016	0.137	0.062	0.014	0.088	0.036	0.018	0.264	0.055	0.010	0.076	0.016
$\varepsilon = 0.10$	0.019	0.146	0.089	0.018	0.091	0.039	0.019	0.273	0.078	0.013	0.084	0.017
$(100, 100)$												
$\varepsilon = 0.00$			0.950	0.925	0.938	0.916			0.625	0.550	0.583	0.495
$\varepsilon = 0.05$	0.836	0.843	0.854	0.799	0.886	0.850	0.384	0.412	0.404	0.315	0.472	0.374
$\varepsilon = 0.10$	0.859	0.867	0.897	0.853	0.900	0.868	0.413	0.447	0.502	0.407	0.495	0.395
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.096	0.025	0.070	0.038			0.091	0.015	0.070	0.015
$\varepsilon = 0.05$	0.019	0.116	0.065	0.011	0.087	0.030	0.021	0.243	0.058	0.010	0.078	0.016
$\varepsilon = 0.10$	0.020	0.120	0.087	0.014	0.086	0.031	0.022	0.250	0.077	0.012	0.082	0.016
$(100, 100)$												
$\varepsilon = 0.00$			0.643	0.512	0.609	0.477			0.324	0.205	0.287	0.163
$\varepsilon = 0.05$	0.640	0.570	0.632	0.504	0.708	0.585	0.295	0.234	0.292	0.158	0.354	0.202
$\varepsilon = 0.10$	0.665	0.607	0.715	0.601	0.725	0.616	0.306	0.254	0.361	0.218	0.370	0.220

TABLE 4.1. Size of the slope homogeneity test

Individual Heterogeneous time dependence in regressors: $\rho_{z_i} \sim U[0.05, 0.95]$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25, 25)$												
$\varepsilon = 0.00$			0.075	0.053	0.052	0.050			0.072	0.031	0.050	0.035
$\varepsilon = 0.05$	0.005	0.051	0.031	0.020	0.044	0.042	0.007	0.062	0.037	0.013	0.052	0.031
$\varepsilon = 0.10$	0.006	0.050	0.051	0.030	0.045	0.045	0.009	0.064	0.055	0.019	0.051	0.033
$(100, 25)$												
$\varepsilon = 0.00$			0.076	0.067	0.055	0.053			0.072	0.057	0.056	0.051
$\varepsilon = 0.05$	0.005	0.017	0.034	0.028	0.041	0.038	0.006	0.016	0.037	0.022	0.049	0.038
$\varepsilon = 0.10$	0.007	0.019	0.052	0.045	0.049	0.043	0.008	0.019	0.053	0.034	0.049	0.043
$(25, 100)$												
$\varepsilon = 0.00$			0.059	0.051	0.050	0.048			0.060	0.047	0.049	0.042
$\varepsilon = 0.05$	0.025	0.070	0.036	0.029	0.049	0.044	0.031	0.062	0.041	0.022	0.055	0.038
$\varepsilon = 0.10$	0.025	0.067	0.049	0.040	0.050	0.048	0.031	0.062	0.053	0.032	0.051	0.040
$(100, 100)$												
$\varepsilon = 0.00$			0.064	0.061	0.060	0.056			0.057	0.057	0.056	0.055
$\varepsilon = 0.05$	0.029	0.041	0.039	0.040	0.049	0.050	0.025	0.032	0.035	0.031	0.049	0.045
$\varepsilon = 0.10$	0.030	0.041	0.050	0.050	0.054	0.054	0.029	0.035	0.052	0.043	0.051	0.047

TABLE 4.2. Power of the slope homogeneity test DGP1- break in θ

Individual Heterogeneous time dependence in regressors: $\rho_{z_i} \sim U[0.05, 0.95]$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.997	0.922	0.995	0.916			0.876	0.458	0.844	0.443
$\varepsilon = 0.05$	0.841	0.845	0.970	0.305	0.978	0.528	0.289	0.481	0.602	0.119	0.649	0.184
$\varepsilon = 0.10$	0.891	0.887	0.985	0.576	0.983	0.662	0.365	0.535	0.719	0.181	0.694	0.233
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.747	0.369	0.703	0.402			0.387	0.126	0.342	0.143
$\varepsilon = 0.05$	0.569	0.582	0.629	0.193	0.838	0.480	0.119	0.283	0.285	0.070	0.419	0.167
$\varepsilon = 0.10$	0.637	0.641	0.742	0.345	0.858	0.557	0.227	0.320	0.388	0.101	0.460	0.183
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 4.3. Power of the slope homogeneity test DGP2 - break in ρ

Individual Heterogeneous time dependence in regressors: $\rho_{z_i} \sim U[0.05, 0.95]$

	No spatial dependence						Weak Spatial dependence (exponential)					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.616	0.424	0.558	0.424			0.330	0.163	0.277	0.179
$\varepsilon = 0.05$	0.122	0.291	0.354	0.102	0.387	0.190	0.046	0.189	0.158	0.045	0.185	0.088
$\varepsilon = 0.10$	0.164	0.333	0.429	0.179	0.419	0.247	0.063	0.212	0.219	0.065	0.209	0.103
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.281	0.100	0.223	0.090			0.177	0.048	0.137	0.052
$\varepsilon = 0.05$	0.115	0.081	0.266	0.036	0.300	0.063	0.062	0.081	0.156	0.019	0.192	0.041
$\varepsilon = 0.10$	0.141	0.096	0.331	0.059	0.318	0.083	0.073	0.091	0.197	0.027	0.189	0.048
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			0.993	0.984	0.993	0.978
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.988	0.990	0.989	0.995	0.993
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.992	0.995	0.995	0.997	0.995

TABLE 4.4. Size of the slope homogeneity test

Time dependence in regressors: ρ_z

	Weak Spatial dependence (polynomial)											
	$\rho_z = 0.5$						$\rho_z = 0.9$					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25, 25)$												
$\varepsilon = 0.00$			0.080	0.029	0.053	0.040			0.077	0.027	0.048	0.024
$\varepsilon = 0.05$	0.008	0.087	0.036	0.012	0.053	0.036	0.011	0.100	0.045	0.016	0.052	0.019
$\varepsilon = 0.10$	0.009	0.089	0.056	0.017	0.057	0.037	0.011	0.101	0.059	0.020	0.055	0.021
$(100, 25)$												
$\varepsilon = 0.00$			0.072	0.049	0.046	0.039			0.072	0.049	0.052	0.033
$\varepsilon = 0.05$	0.003	0.016	0.033	0.019	0.045	0.038	0.008	0.031	0.041	0.017	0.043	0.020
$\varepsilon = 0.10$	0.004	0.020	0.048	0.031	0.046	0.038	0.009	0.034	0.053	0.026	0.045	0.024
$(25, 100)$												
$\varepsilon = 0.00$			0.068	0.043	0.052	0.041			0.069	0.048	0.053	0.040
$\varepsilon = 0.05$	0.031	0.076	0.041	0.023	0.049	0.037	0.040	0.091	0.044	0.025	0.056	0.035
$\varepsilon = 0.10$	0.030	0.079	0.052	0.032	0.050	0.040	0.036	0.087	0.058	0.035	0.058	0.036
$(100, 100)$												
$\varepsilon = 0.00$			0.060	0.053	0.049	0.049			0.058	0.057	0.053	0.051
$\varepsilon = 0.05$	0.026	0.035	0.037	0.029	0.047	0.044	0.032	0.045	0.039	0.032	0.048	0.042
$\varepsilon = 0.10$	0.027	0.038	0.044	0.041	0.048	0.046	0.028	0.045	0.048	0.044	0.048	0.045

	Strong Spatial dependence											
	$\rho_z = 0.5$						$\rho_z = 0.9$					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$(25, 25)$												
$\varepsilon = 0.00$			0.072	0.016	0.049	0.017			0.070	0.020	0.045	0.007
$\varepsilon = 0.05$	0.011	0.202	0.046	0.010	0.049	0.016	0.018	0.210	0.050	0.013	0.051	0.007
$\varepsilon = 0.10$	0.010	0.205	0.055	0.011	0.049	0.016	0.016	0.210	0.061	0.016	0.050	0.007
$(100, 25)$												
$\varepsilon = 0.00$			0.072	0.030	0.047	0.032			0.071	0.033	0.044	0.024
$\varepsilon = 0.05$	0.008	0.037	0.040	0.013	0.050	0.027	0.013	0.051	0.052	0.014	0.051	0.015
$\varepsilon = 0.10$	0.008	0.041	0.054	0.018	0.051	0.029	0.012	0.054	0.058	0.020	0.051	0.017
$(25, 100)$												
$\varepsilon = 0.00$			0.059	0.033	0.049	0.017			0.061	0.030	0.051	0.013
$\varepsilon = 0.05$	0.040	0.129	0.040	0.017	0.056	0.011	0.054	0.142	0.046	0.020	0.058	0.012
$\varepsilon = 0.10$	0.034	0.130	0.048	0.025	0.052	0.013	0.038	0.136	0.051	0.022	0.052	0.012
$(100, 100)$												
$\varepsilon = 0.00$			0.060	0.042	0.051	0.037			0.060	0.045	0.051	0.029
$\varepsilon = 0.05$	0.033	0.038	0.042	0.024	0.052	0.033	0.046	0.048	0.048	0.026	0.062	0.029
$\varepsilon = 0.10$	0.033	0.039	0.053	0.033	0.051	0.035	0.039	0.047	0.058	0.036	0.058	0.030

Table 4.5. Power of the slope homogeneity test DGP1 - break in θ

Time dependence in regressors: ρ_z

	Weak Spatial dependence (polynomial)											
	$\rho_z = 0.5$						$\rho_z = 0.9$					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.796	0.318	0.758	0.329			0.735	0.242	0.681	0.151
$\varepsilon = 0.05$	0.200	0.421	0.485	0.084	0.547	0.161	0.167	0.503	0.402	0.055	0.412	0.045
$\varepsilon = 0.10$	0.260	0.470	0.597	0.129	0.596	0.187	0.216	0.543	0.531	0.087	0.477	0.056
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.338	0.084	0.271	0.107			0.345	0.088	0.259	0.069
$\varepsilon = 0.05$	0.138	0.268	0.228	0.052	0.353	0.125	0.165	0.308	0.271	0.049	0.325	0.061
$\varepsilon = 0.10$	0.167	0.303	0.322	0.074	0.390	0.137	0.194	0.334	0.372	0.072	0.374	0.076
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	Strong Spatial dependence											
	$\rho_z = 0.5$						$\rho_z = 0.9$					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n,T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.492	0.073	0.411	0.055			0.411	0.061	0.338	0.020
$\varepsilon = 0.05$	0.055	0.355	0.205	0.022	0.203	0.025	0.062	0.411	0.164	0.026	0.157	0.008
$\varepsilon = 0.10$	0.072	0.383	0.293	0.028	0.264	0.030	0.070	0.435	0.244	0.033	0.200	0.010
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	0.999
$\varepsilon = 0.10$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000	0.999
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.194	0.029	0.142	0.027			0.191	0.033	0.139	0.014
$\varepsilon = 0.05$	0.061	0.277	0.127	0.017	0.163	0.024	0.084	0.308	0.149	0.023	0.171	0.013
$\varepsilon = 0.10$	0.065	0.299	0.195	0.022	0.191	0.027	0.090	0.323	0.209	0.028	0.195	0.015
$(100, 100)$												
$\varepsilon = 0.00$			0.996	0.989	0.995	0.984			0.987	0.980	0.985	0.969
$\varepsilon = 0.05$	0.996	0.996	0.988	0.987	0.997	0.996	0.986	0.996	0.968	0.964	0.986	0.986
$\varepsilon = 0.10$	0.997	0.998	0.996	0.994	0.998	0.997	0.988	0.998	0.986	0.986	0.991	0.991

Table 4.6. Power of the slope homogeneity test DGP2 - break in ρ

Time dependence in regressors: ρ_z

	Weak Spatial dependence (polynomial)											
	$\rho_z = 0.5$						$\rho_z = 0.9$					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.241	0.095	0.192	0.120			0.364	0.139	0.295	0.095
$\varepsilon = 0.05$	0.031	0.170	0.111	0.029	0.138	0.070	0.052	0.293	0.157	0.046	0.162	0.037
$\varepsilon = 0.10$	0.044	0.190	0.159	0.041	0.153	0.080	0.070	0.320	0.226	0.062	0.198	0.044
$(100, 100)$												
$\varepsilon = 0.00$			1.000	1.000	1.000	1.000			1.000	1.000	1.000	1.000
$\varepsilon = 0.05$	0.998	0.998	0.999	0.998	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000
$\varepsilon = 0.10$	0.999	0.999	1.000	0.999	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.136	0.037	0.106	0.048			0.202	0.041	0.141	0.029
$\varepsilon = 0.05$	0.044	0.100	0.115	0.015	0.139	0.042	0.088	0.128	0.172	0.020	0.197	0.019
$\varepsilon = 0.10$	0.052	0.111	0.157	0.022	0.147	0.045	0.099	0.141	0.235	0.028	0.219	0.024
$(100, 100)$												
$\varepsilon = 0.00$			0.964	0.918	0.958	0.900			1.000	0.999	1.000	0.997
$\varepsilon = 0.05$	0.967	0.941	0.959	0.937	0.974	0.955	1.000	0.999	0.999	0.999	1.000	0.999
$\varepsilon = 0.10$	0.972	0.953	0.976	0.962	0.981	0.963	1.000	1.000	1.000	1.000	1.000	1.000

	Strong Spatial dependence											
	$\rho_z = 0.5$						$\rho_z = 0.9$					
Test	Asymptotic		Wild Bootstrap		Efron Bootstrap		Asymptotic		Wild Bootstrap		Efron Bootstrap	
(n, T)	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$	$\mathcal{T}_\varepsilon$	$\mathcal{T}_\varepsilon^\Delta$
$t_0 = [0.5T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.161	0.034	0.128	0.029			0.211	0.051	0.162	0.016
$\varepsilon = 0.05$	0.025	0.251	0.079	0.015	0.090	0.020	0.035	0.319	0.089	0.025	0.092	0.007
$\varepsilon = 0.10$	0.028	0.266	0.106	0.018	0.102	0.022	0.038	0.339	0.131	0.031	0.117	0.009
$(100, 100)$												
$\varepsilon = 0.00$			0.871	0.824	0.855	0.799			0.994	0.986	0.993	0.980
$\varepsilon = 0.05$	0.676	0.699	0.703	0.625	0.732	0.663	0.955	0.969	0.951	0.917	0.964	0.923
$\varepsilon = 0.10$	0.710	0.731	0.781	0.699	0.774	0.696	0.965	0.976	0.977	0.947	0.977	0.942
$t_0 = [0.8T]$												
$(25, 25)$												
$\varepsilon = 0.00$			0.110	0.018	0.082	0.019			0.137	0.026	0.099	0.009
$\varepsilon = 0.05$	0.030	0.212	0.089	0.011	0.098	0.018	0.057	0.229	0.117	0.015	0.126	0.006
$\varepsilon = 0.10$	0.033	0.220	0.114	0.013	0.108	0.018	0.060	0.238	0.148	0.019	0.139	0.007
$(100, 100)$												
$\varepsilon = 0.00$			0.531	0.322	0.507	0.291			0.855	0.613	0.839	0.493
$\varepsilon = 0.05$	0.532	0.381	0.503	0.319	0.577	0.359	0.861	0.720	0.816	0.633	0.866	0.600
$\varepsilon = 0.10$	0.551	0.415	0.591	0.390	0.601	0.395	0.870	0.762	0.873	0.720	0.890	0.653