MONASH University

MONASH BUSINESS SCHOOL

ISSN 1440-771X

Department of Econometrics and Business Statistics

http://business.monash.edu/econometrics-and-business-statistics/research/publications

Estimation of Structural Breaks in Large Panels with Cross-Sectional Dependence

Jiti Gao, Guangming Pan and Yanrong Yang

July 2016

Working Paper 12/16

Estimation of Structural Breaks in Large Panels with Cross–Sectional Dependence

Jiti Gao* and Guangming Pan† and Yanrong Yang‡

July 27, 2016

Abstract

This paper considers modelling and detecting structure breaks associated with cross-sectional dependence for large dimensional panel data models, which are popular in many fields, including economics and finance. We propose a dynamic factor structure to measure the degree of cross-sectional dependence. The extent of such cross-sectional dependence is parameterized as an unknown parameter, which is defined by assuming that a small proportion of the total factor loadings are important. Compared with the usual parameterized style, this exponential description of extent covers the case of small proportion of the total sections being cross–sectionally dependent. We establish a 'moment' criterion to estimate the unknown based on the covariance of cross-sectional averages at different time lags. By taking into account the fact that the serial dependence of common factors is stronger than that of idiosyncratic components, the proposed criterion is able to capture weak cross–sectional dependence that is reflected on relatively small values of the unknown parameter. Due to the involvement of some unknown parameter, both joint and marginal estimators are constructed. This paper then establishes that the joint estimators of a pair of unknown parameters converge in distribution to bivariate normal. In the case where the other unknown parameter is being assumed to be known, an asymptotic distribution for an estimator of the original unknown parameter is also established, which naturally coincides with the joint asymptotic distribution for the case where the other unknown parameter is assumed to be known. Simulation results show the finite–sample effectiveness of the proposed method. Empirical applications to cross-country macro-variables and stock returns in SP500 market are also reported to show the practical relevance of the proposed estimation theory.

Keywords: Cross-sectional averages; dynamic factor model; joint estimation; marginal estimation; strong factor loading.

*The corresponding author, Jiti Gao, Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia (Email: jiti.gao@monash.edu).

†Guangming Pan, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, 637371 (Email: GMPAN@ntu.edu.sg).

‡Yanrong Yang, Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia (Email: yanrong.yang@monash.edu).

JEL Classification: C21, C32.

1 Introduction

The analysis of large dimensional panel data attracts ever-growing interest in some modern scientific fields, especially in economics and finance. Cross–sectional dependence is common in large dimensional panel data analysis and the literature focuses on testing the existence of cross-sectional dependence. A survey on description and testing of cross-sectional dependence is given in Sarafidis and Wansbeek (2012). Pesaran (2004) utilizes sample correlations to test cross-sectional dependence while Baltagi, Feng and Kao (2012) extend the classical Lagrangian multiplier (LM) test to the large dimensional case. Chen, Gao and Li (2012) and Hsiao, Pesaran and Pick (2012) consider cross-sectional dependence tests for nonlinear econometric models.

When more and more cross–sections are grouped together, cross-sectional dependence appears to be quite natural and common. Cross-sectional independence is an extreme hypothesis. Rejecting such a hypothesis does not provide much information about the relationship between different cross–sections under investigation. In view of this, measuring the degree of cross-sectional dependence is more important than just testing for its presence. As we know, in comparison with cross–sectional dependence tests, the literature contributes very little to accessing the extent of cross-sectional dependence. Ng (2006) uses spacings of cross-sectional correlations to exploit the ratio of correlated subsets over all sections. Bailey, Kapatanios and Pesaran (2015) use a factor model to describe cross-sectional dependence and develop estimators that are based on a moment method. We will contribute to the descriptions and measures of the extent of cross-sectional dependence for large dimensional panel data with N cross-section units and T time series. The first natural question is: how to describe the cross-sectional dependence ? To deal with this issue, the panel data literature mainly discusses two different ways of modelling cross-sectional dependence: the spatial correlation and the factor structure approach (see, for example, Sarafidis and Wansbeek (2012)). In this paper, we use the factor structure approach. The factor model is not only a powerful tool to characterize cross-sectional dependence for economic and financial data, but also efficient in dealing with statistical inference for high dimensional data from a dimension-reduction point of view. Some related work includes Fan, Fan and Lv (2008), Fan, Liao and Mincheva (2011) and Pan and Yao (2008).

With respect to factor structures, there are two common types of factor models: the static model and the dynamic model. The static model is defined as

$$x_{it} = \mu_i + \boldsymbol{\beta}_i'\mathbf{f}_t + u_{it}, \quad i = 1, 2, \ldots, N; \quad t = 1, 2, \ldots, T.$$

In this model, $\{\mu_i, i = 1, 2, \ldots, N\}$ represent the means of N sections, the components of the process $\{\mathbf{f}_t, t = 1, 2, \ldots, T\}$ are called the common shocks or factors, and $\boldsymbol{\beta}_i$ is a vector of factor loadings for unit i on the common factors $\mathbf{f}_t$ for each $i = 1, 2, \ldots, N$. The term static factor model refers to the static relationship between x_{it} and $\mathbf{f}_t$, but $\mathbf{f}_t$ itself can be a dynamic process. The dynamic factor model is written as

$$x_{it} = \mu_i + \boldsymbol{\beta}_i^{'}(L)\mathbf{f}_t + u_{it}, \quad i = 1, 2, \ldots, N; \quad t = 1, 2, \ldots, T,$$

where $\boldsymbol{\beta}_i(L)$ is a vector of dynamic factor loadings of order s, i.e $\boldsymbol{\beta}_i(L) = \boldsymbol{\beta}_{i0} + \boldsymbol{\beta}_{i1}L + \boldsymbol{\beta}_{i2}L^2 + \cdots + \boldsymbol{\beta}_{is}L^s$. If s is finite, the model is called a dynamic factor model. If s can be infinity, it is called a generalized dynamic factor model. Under each case, $\mathbf{f}_t = \mathbf{C}(L)\boldsymbol{\varepsilon}_t$, where $\{\boldsymbol{\varepsilon}_t, t \in \mathbb{Z}\}$ are independent and identically distributed (i.i.d) and $\mathbf{C}(L)$ is a coefficient matrix with time lags, i.e., $\{\mathbf{f}_t, t \in \mathbb{Z}\}$ is a dynamic process. Based on the concept of a factor model, the extent of cross-sectional dependence in observed data x_{it} can be reflected in the strength of factor loadings, and the cross-sectional dependence is caused by common factors $\mathbf{f}_t$.

While it is a rare phenomenon to have cross-sectional independence for all N sections, it is also unrealistic to assume that all N sections are dependent. As cross-sectional dependence can be reflected in factor loadings, we impose some conditions on factor loadings in order to derive one part that contains cross-sectional dependent units and another part that includes cross-sectional independent sections. The simplest method is to assume that some factor loadings are bounded away from zero while others are around zero. In this paper, we assume that only $[N^{\alpha_0}](0 \leq \alpha_0 \leq 1)$ of all N factor loadings are individually important. This topic is quite related to the structural–break literature developed in the econometrics literature. Bai (1997) investigates multiple structural breaks in the mean for a single linear process, while Bai and Perron (1998) estimate breaks in the coefficients of a linear model. Bai (2010) considers common breaks in the mean for panel data and develops breakpoint analysis under large dimensional cases. By contrast, as far as we know, the literature that addresses the structural breaks for cross-sectional dependence is relatively limited. Ng (2006) is one of the few that contributes to breakpoint analysis of cross–sectional dependence using the spacings. Instead of measuring the extent by $\alpha_0 N$, we adopt the parameterization $[N^{\alpha_0}]$. The proportion of $[N^{\alpha_0}]$ over the total N is quite small which tends to 0 as $0 < \alpha_0 < 1$, while $\alpha_0 N$ is comparable to N because of the same order. In this sense, our model covers some "sparse" cases that only a small part of the sections are cross-sectionally dependent.

With this description of the extent of cross-sectional dependence, the goal is directed to propose an approach to estimation of α_0. One advantage of our proposed statistic is based on an assumption for identification. There are several different ways of identifying factor models. Bailey, Kapatanios and Pesaran (2015) assume that $\{\mathbf{x}_t = (x_{1t}, x_{2t}, \ldots, x_{Nt})' : t = 1, 2, \ldots, T\}$ is a high dimensional time series and decompose it into two parts: a common-factor part and an idiosyncratic part, both of which are weakly dependent stationary time series. The identification of their model lies on the assumption that cross-sectional dependence involved in the idiosyncratic part is weaker than that in the common-factor part. This assumption is common in the literature, e.g. Bai and Ng (2002), Fan, Fan and Lv (2008) and Fan, Liao and Mincheva (2011). However, we consider an alternative factor model, similar to the idea of Lam and Yao (2012), which consists of two parts: the common-factor part driven by a lower-dimensional factor time series and the idiosyncratic part which is a stationary time series with relatively weaker serial dependence than the common-factor part. While the literature makes use of distinctive degrees of cross-sectional dependence in common components and idiosyncratic components respectively, we utilize distinctive extents of serial dependence in these two parts to attain identifications. From a point of replacing one condition by another in identification, our assumptions are quite weak. Moreover, one important advantage is that the new model identification condition leads to our proposed methodology for estimation of the exponents of cross-sectional dependence, which can eliminate the influence of idiosyncratic components in the estimation.

The proposal of our estimation procedure is outlined as follows. An estimator for α_0 is proposed by calculating the covariance between $\bar{x}_t$ and $\bar{x}_{t+\tau}$ for a larger range of α_0, i.e., $0 \leq \alpha_0 \leq 1$, where $\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N} x_{it}$ and $\tau > 0$. Under the setting and structure of this paper, furthermore, we have weaker serial dependence in the idiosyncratic part than that in the common part. Then the leading term in $\mathrm{cov}(\bar{x}_t, \bar{x}_{t+\tau})$ will not contain the idiosyncratic part when τ tends to infinity. In other words, the idiosyncratic components do not bring any noise term to the proposed criterion.

The main contribution of this paper is summarized as follows:

1. We construct two consistent estimators for α_0 by utilizing both joint estimation and marginal estimation respectively. As the parameter κ_0 involved in the proposed criterion is unknown, the joint estimation of α_0 and κ_0 is adopted. Otherwise, we use the marginal estimation for α_0.

2. We have been able to establish new asymptotic distributions for both the joint and the marginal estimators. The asymptotic marginal distribution coincides with that for the joint estimator for the case where κ_0 is assumed to be known. Finite–sample performances of these two kinds of estimators are provided.

3. An additional contribution involves generalizing Theorem 8.4.2 of Anderson (1994). We establish a new central limit theorem for the sample covariance of a time series under the case where both the time lag and the sample size tend to infinity simultaneously.

The rest of the paper is organized as follows. The model and the main assumptions are introduced in Section 2. Section 3 proposes both joint and marginal estimators that are based on the second moment criterion. Asymptotic properties for these estimators are established in Section 4. Section 5 reports the simulation results, which illustrate the effectiveness of the proposed methods. Section 6 provides empirical applications to cross-country macro-variables and stock returns in S&P 500 market respectively. Conclusions are included in Section 7. Justification for the assumptions and all the mathematical proofs are given in Appendices B and C.

2 The model

Let x_{it} be a double array of random variables indexed by $i = 1, \ldots, N$ and $t = 1, \ldots, T$, over space and time, respectively. The aim of this paper is to measure the extent of the cross-sectional dependence of the data $\{x_{it} : i = 1, \ldots, N\}$. In panel data analysis, there are two common models to describe cross-sectional dependence: spatial models and factor models. In Bailey, Kapatanios and Pesaran (2015), a static approximate factor model is used. As an extension, we consider a dynamic factor model:

$$\begin{aligned} x_{it} &= \mu_i + \boldsymbol{\beta}_{i0}'\mathbf{f}_t + \boldsymbol{\beta}_{i1}'\mathbf{f}_{t-1} + \cdots + \boldsymbol{\beta}_{is}'\mathbf{f}_{t-s} + u_{it} \\ &= \mu_i + \boldsymbol{\beta}_i'(L)\mathbf{f}_t + u_{it}, \quad i = 1, 2, \ldots, N; \; t = 1, 2, \ldots, T, \end{aligned} \tag{2.1}$$

where $\mathbf{f}_t$ is the $m \times 1$ vector of unobserved factors (with m being fixed),

$$\boldsymbol{\beta}_i(L) = \boldsymbol{\beta}_{i0} + \boldsymbol{\beta}_{i1}L + \boldsymbol{\beta}_{i2}L^2 + \cdots + \boldsymbol{\beta}_{is}L^s,$$

in which $\boldsymbol{\beta}_{i\ell} = (\beta_{i\ell 1}, \beta_{i\ell 2}, \ldots, \beta_{i\ell m})'$, $\ell = 0, 1, \ldots, s$ are the associated vectors of unobserved factor loadings and L is the lag operator, here s is assumed to be fixed, and $\mu_i, i = 1, 2, \ldots, N$ are constants that represent the mean values for all sections, and $\{u_{it} : i = 1, \ldots, N; t = 1, \ldots, T\}$ are idiosyncratic components.

Clearly, we can write (2.1) in the static form:

$$x_{it} = \mu_i + \boldsymbol{\beta}_i'\mathbf{F}_t + u_{it}, \; i = 1, 2, \ldots, N; t = 1, 2, \ldots, T, \tag{2.2}$$

where

$$\boldsymbol{\beta}_i = \begin{pmatrix} \boldsymbol{\beta}_{i0} \\ \boldsymbol{\beta}_{i1} \\ \vdots \\ \boldsymbol{\beta}_{is} \end{pmatrix}_{m(s+1)} \quad \textit{and} \quad \mathbf{F}_t = \begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_{t-1} \\ \vdots \\ \mathbf{f}_{t-s} \end{pmatrix}_{m(s+1)}.$$

The dimension of $\mathbf{f}_t$ is called the number of dynamic factors and is denoted by m. Then the dimension of $\mathbf{F}_t$ is equal to $r = m(s + 1)$. In factor analysis, $\boldsymbol{\beta}_i'\mathbf{F}_t$ is called the common components of x_{it}.

We first introduce the following assumptions.

Assumption 1. *The idiosyncratic component* $\{\mathbf{u}_t = (u_{1t}, u_{2t}, \ldots, u_{Nt})' : t = 1, 2, \ldots, T\}$ *follows a linear stationary process of the form:*

$$u_{it} = \sum_{j=0}^{+\infty} \phi_{ij} \Big(\sum_{s=-\infty}^{+\infty} \xi_{js}\nu_{j,t-s} \Big), \tag{2.3}$$

where $\{\nu_{is} : i = \ldots, -1, 0, 1, \ldots; s = 0, 1, \ldots\}$ *is a double sequence of i.i.d. random variables with mean zero and unit variance, and*

$$\sup_{0<j<+\infty} \sum_{i=1}^{N} |\phi_{ij}| < +\infty \quad \text{and} \quad \sup_{0<j<+\infty} \sum_{s=-\infty}^{+\infty} |\xi_{js}| \le +\infty. \tag{2.4}$$

Moreover,

$$E(u_{it}u_{j,t+\tau}) = \gamma_1(\tau)\gamma_2(|i-j|), \tag{2.5}$$

where $\gamma_2(|i-j|)$ *satisfies*

$$\sum_{i,j=1}^{N} \gamma_2(|i-j|) = O(N) \tag{2.6}$$

and $\gamma_1(\tau)$ *satisfies the condition (2.9) in Assumption 3 below.*

Assumption 2. *For* $\ell = 0, 1, 2, \ldots, s$ *and* $k = 1, 2, \ldots, m$,

$$\beta_{i\ell k} = v_{i\ell k},\ i = 1, 2, \ldots, [N^{\alpha_{\ell k}}] \quad \text{and} \quad \sum_{i=[N^{\alpha_{\ell k}}]+1}^{N} \beta_{i\ell k} = O(1), \tag{2.7}$$

where $[N^{\alpha_{\ell k}}] \le N^{\alpha_{\ell k}}$ *is the largest integer part of* $N^{\alpha_{\ell k}}$, $0 < \alpha_{\ell k} \le 1$ *and* $v_{i\ell k} \sim i.i.d.(\mu_v, \sigma_v^2)$ *has finite sixth moment, with* $\mu_v \neq 0$ *and* $\sigma_v^2 > 0$. *Moreover,* $\{v_{i\ell k} : i = 1, 2, \ldots, N; \ell = 0, 1, \ldots, s; k = 1, 2, \ldots, m\}$ *are assumed to be independent of the factors* $\{\mathbf{f}_t : t = 1, 2, \ldots, T\}$ *and the idiosyncratic components* $\{u_{it} : i = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$.

Assumption 3. *The factors* $\{\mathbf{f}_t, t = 1, 2, \ldots, T\}$ *are covariance stationary with the following representation:*

$$\mathbf{f}_t = \sum_{j=0}^{\infty} b_j \boldsymbol{\zeta}_{t-j}, \tag{2.8}$$

where $\{\boldsymbol{\zeta}_t, t = \ldots, -1, 0, 1, \ldots\}$ *is an i.i.d sequence of* m*-dimensional random vectors whose components are i.i.d with zero mean and unit variance, the fourth moments of* $\{\boldsymbol{\zeta}_t, -\infty < t < \infty\}$ *are finite, and the coefficients* $\{b_j : j = 0, 1, 2, \ldots\}$ *satisfy* $\sum_{j=0}^{\infty} |b_j| < \infty$. *Furthermore, the unobserved factors* $\{\mathbf{f}_t : t = 1, 2, \ldots, T\}$ *are independent of the idiosyncratic components* $\{\mathbf{u}_t : t = 1, 2, \ldots, T\}$. *Let* $\gamma(k, h) = E(f_{k,t}f_{k,t+h})$ *and* $\alpha_0 = \max_{\ell,k}(\alpha_{\ell k})$. *We assume*

$$\Delta_{u,f} := \frac{N^{-1}\gamma_1(\tau)}{[N^{2\alpha_0-2}] \sum_{\ell_1,\ell_2=0}^{s} \sum_{k=1}^{m} \gamma(k, \tau - \ell_2 + \ell_1)} = o(1), \tag{2.9}$$

where $\gamma_1(\tau)$ *is defined in (2.5).*

Let us briefly discuss how to verify (2.9) using a simple example. Consider the following model:

$$x_{it} = \mu_i + \beta_i f_t + u_{it}, \quad i = 1, 2, \ldots, N,\ t = 1, 2, \ldots, T, \tag{2.10}$$

where the factor loadings $\{\beta_i : i = 1, 2, \ldots, N\}$ satisfy Assumption 2, the factor process $\{f_t : t = 1, 2, \ldots, T\}$ is AR(1), i.e., $f_t = \rho_1 f_{t-1} + \varepsilon_t$ for $t = 1, 2, \ldots, T$, and the idiosyncratic components u_{it} can be decomposed into two independent parts: the serially correlated part and the cross–section part, i.e., $u_{it} = \zeta_t \eta_i$, with $\{\zeta_t : t = 1, 2, \ldots, T\}$ being an AR(1), i.e., $\zeta_t = \rho_2 \zeta_{t-1} + \epsilon_t, \ \ t = 1, 2, \ldots, T$. Moreover, $\{\varepsilon_t : t = 1, 2, \ldots, T\}$ and $\{\epsilon_t : t = 1, 2 \ldots, T\}$ are both white noises with zero mean and unit variance, and mutually independent.

For model (2.10), it is easy to derive the values of $\gamma_1(\tau)$, $\gamma_2(|i-j|)$ and $\gamma(1, \tau)$ defined in Assumption 2 and Assumption 3, i.e., $\gamma_1(\tau) = \frac{|\rho_2|^\tau}{1-\rho_1^2}$, $\gamma(\tau) = \frac{|\rho_1|^\tau}{1-\rho_1^2}$ and $E(\eta_i \eta_j) = \gamma_2(|i-j|)$.

Condition (2.9) is then $\Delta_{u,f} := \frac{N^{-1}|\rho_2|^\tau (1-\rho_1^2)}{[N^{2\alpha_0-2}]\cdot|\rho_1|^\tau (1-\rho_2^2)} = o(1)$. It is equivalent to requiring that ρ_1 and ρ_2 are related by $|\rho_2| = o([N^{\frac{2\alpha_0-1}{\tau}}] \cdot |\rho_1|)$. We can then see that, if $\frac{1}{2} < \alpha_0 < 1$, τ can be taken as a constant. If $0 < \alpha_0 < \frac{1}{2}$, τ should tend to $+\infty$ and ρ_2 should be smaller than ρ_1.

Detailed justifications of Assumptions 1-3 are given in Appendix A.

3 The estimation method

The aim of this paper is to estimate the exponent $\alpha_0 = \max_{\ell,k}(\alpha_{\ell k})$, which describes the extent of cross-sectional dependence. As in Bailey, Kapatanios and Pesaran (2015) (BKP15), we consider the cross-sectional average $\bar{x}_t = 1/N \sum_{i=1}^N x_{it}$ and then derive an estimator for α_0 from the information of $\{\bar{x}_t : t = 1, 2, \ldots, T\}$. BKP15 use the variance of the cross-sectional average $\bar{x}_t$ to estimate α_0 and carry out statistical inference for an estimator of α_0 before they show that

$$Var(\bar{x}_t) = \widetilde{\kappa}_0 [N^{2\alpha_0-2}] + N^{-1} c_N + O(N^{\alpha_0-2}), \tag{3.1}$$

where $\widetilde{\kappa}_0$ is a constant associated with the common components and c_N is a bias constant incurred by the idiosyncratic errors. From (3.1), we can see that, in order to estimate α_0, BKP15 assume that $2\alpha_0 - 2 > -1$, i.e. $\alpha_0 > 1/2$. Otherwise, the second term will have a higher order than the first term. So the approach by BKP15 will fail in the case of $0 < \alpha_0 < 1/2$.

In this paper, we propose a new estimator that is applicable to the full range of α_0, i.e., $0 \leq \alpha_0 \leq 1$. Based on the assumption that the common factors possess serial dependence that is stronger than that of the idiosyncratic components, we construct a so–called covariance criterion $Cov(\bar{x}_t, \bar{x}_{t+\tau})$, whose leading term does not include the idiosyncratic components for $0 \leq \alpha_0 \leq 1$. In other words, the advantage of this covariance criterion over the variance criterion $Var(\bar{x}_t)$ lies on the fact that there is no interruption brought by the idiosyncratic components $\{u_{it} : i = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$ in $Cov(\bar{x}_t, \bar{x}_{t+\tau})$.

Next, we illustrate how the covariance $Cov(\bar{x}_t, \bar{x}_{t+\tau})$ implies the extent parameter α_0 in detail. Note that $[N^a] \leq N^a$ $(a \geq 0)$ denotes the largest integer part. For simplicity, let $[N^b]$ $(b \leq 0)$ denote $\frac{1}{[N^{-b}]}$. Moreover, to simplify the notation, throughout the paper we also use the following notation:

$$[N^{ka}] := [N^a]^k, \ \ [N^{a-k}] := \frac{[N^a]}{N^k}, \ \ \forall a, k \in \mathbb{R}. \tag{3.2}$$

But we would like to remind the reader that $[N^{ka}]$ is actually not equal to $[N^a]^k$. Next, we will propose an estimator for α_0 under two different scenarios: the joint estimator under the case of some other parameters being unknown while the marginal estimator for the case of some other parameters being known.

3.1 The joint estimator $(\widetilde{\alpha}, \widetilde{\kappa})$

The joint estimator in this section is proposed when all the parameters involved are unknown. The marginal estimator proposed in the following section deals with the case where only α_0 is unknown.

Without loss of generality, we assume that $\alpha_{\ell k} = \alpha_0, \forall \ell = 0, 1, 2, \ldots, s; k = 1, 2, \ldots, m$. Let Assumption 2 hold. Let $\bar{x}_{nt}$ be the cross–sectional average of x_{it} over $i = 1, 2, \ldots, n$ with $n \leq N$. Similarly, $\bar{\beta}_{n\ell k} := \frac{1}{n}\sum_{i=1}^{n} \beta_{i\ell k}$. Then

$$E(\bar{\beta}_{n\ell k}) = \begin{cases} \mu_v, & n \leq [N^{\alpha_0}] \\ \frac{[N^{\alpha_0}]}{n}\mu_v + \frac{K_{n\ell k}}{n}, & n > [N^{\alpha_0}], \end{cases}$$

and

$$Var(\bar{\beta}_{n\ell k}) = \begin{cases} \frac{\sigma_v^2}{n}, & n \leq [N^{\alpha_0}] \\ \frac{[N^{\alpha_0}]}{n^2}\sigma_v^2, & n > [N^{\alpha_0}], \end{cases}$$

where $K_{n\ell k} = \sum_{i=[N^{\alpha_0}]+1}^{n} \beta_{i\ell k}$.

Then, we have

$$\begin{aligned} Cov(\bar{x}_{nt}, \bar{x}_{n,t+\tau}) &= \sum_{\ell=0}^{s}\sum_{k=1}^{m}\Big(E^2(\bar{\beta}_{n\ell k}) + Var(\bar{\beta}_{n\ell k})\Big)E(f_{k,t-\ell}f_{k,t+\tau-\ell}) \\ &\quad + \sum_{\ell_1\neq\ell_2}^{s}\sum_{k=1}^{m} E(\bar{\beta}_{n\ell_1 k})E(\bar{\beta}_{n\ell_2 k})E(f_{k,t-\ell_1}f_{k,t+\tau-\ell_2}) \\ &= \begin{cases} \kappa_0 + O(n^{-1}), & n \leq [N^{\alpha_0}] \\ \kappa_0\frac{[N^{2\alpha_0}]}{n^2} + O(\frac{[N^{\alpha_0}]}{n^2}), & n > [N^{\alpha_0}], \end{cases} \end{aligned} \tag{3.3}$$

where

$$\kappa_0 = \mu_v^2 \sum_{\ell_1,\ell_2=0}^{s}\sum_{k=1}^{m} E(f_{k,t-\ell_1}f_{k,t+\tau-\ell_2}), \tag{3.4}$$

in which $\mu_v = E[v_{i\ell k}]$.

Minimize the following quadratic form in terms of α and κ:

$$Q_{NT}^{(1)}(\alpha,\kappa) = \sum_{n=1}^{[N^\alpha]} n^3\Big(\hat{\sigma}_n(\tau) - \kappa\Big)^2 + \sum_{n=[N^\alpha]+1}^{N} n^3\Big(\hat{\sigma}_n(\tau) - \frac{[N^{2\alpha}]}{n^2}\kappa\Big)^2, \tag{3.5}$$

where $\hat{\sigma}_n(\tau)$ is a consistent estimator for $Cov(\bar{x}_{nt}, \bar{x}_{n,t+\tau})$ of the form:

$$\hat{\sigma}_n(\tau) = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\left(\bar{x}_{nt} - \bar{x}_n^{(1)}\right)\left(\bar{x}_{n,t+\tau} - \bar{x}_n^{(2)}\right),$$

with $\bar{x}_n^{(1)} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{x}_{nt}$ and $\bar{x}_n^{(2)} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{x}_{n,t+\tau}$.

Then, the first order condition for κ is

$$\frac{\partial Q_{NT}^{(1)}(\alpha,\kappa)}{\partial\kappa} = 0,$$

which is equivalent to $\sum_{n=1}^{[N^\alpha]} n^3\Big(\hat{\sigma}_n(\tau) - \kappa\Big) + [N^{2\alpha}]\sum_{n=[N^\alpha]+1}^{N} n\Big(\hat{\sigma}_n(\tau) - \frac{[N^{2\alpha}]}{n^2}\kappa\Big) = 0.$

This derives

$$\kappa = \kappa(\alpha) := \frac{\sum_{n=1}^{[N^\alpha]} n^3\hat{\sigma}_n(\tau) + \sum_{n=[N^\alpha]+1}^{N} n[N^{2\alpha}]\hat{\sigma}_n(\tau)}{\sum_{n=1}^{[N^\alpha]} n^3 + \sum_{n=[N^\alpha]+1}^{N}\frac{[N^{4\alpha}]}{n}}. \tag{3.6}$$

We now introduce the additional expressions:

$$\hat{q}_1^{(1)}(\alpha) = \sum_{n=1}^{[N^\alpha]} n^3\hat{\sigma}_n(\tau),\quad \hat{q}_2^{(1)}(\alpha) = \sum_{n=[N^\alpha]+1}^{N} n\hat{\sigma}_n(\tau),$$

$$N^{(1)}(\alpha) = \sum_{n=1}^{[N^\alpha]} n^3 + \sum_{n=[N^\alpha]+1}^{N}\frac{[N^{4\alpha}]}{n},\quad Q^{(1)} = \sum_{n=1}^{N} n^3\hat{\sigma}_n^2(\tau).$$

With this and (3.6), we can obtain

$$\kappa = \frac{\hat{q}_1^{(1)}(\alpha) + [N^{2\alpha}]\hat{q}_2^{(1)}(\alpha)}{N^{(1)}(\alpha)}. \tag{3.7}$$

Then

$$\begin{aligned}
Q_{NT}^{(1)}(\alpha,\kappa) &= Q^{(1)} + \kappa^2\sum_{n=1}^{[N^\alpha]} n^3 + \kappa^2[N^{4\alpha}]\sum_{n=[N^\alpha]+1}^{N} n^{-1} - 2\kappa\hat{q}_1^{(1)}(\alpha) - 2\kappa[N^{2\alpha}]\hat{q}_2^{(1)}(\alpha) \\
&= Q^{(1)} + \kappa^2 N^{(1)}(\alpha) - 2\kappa\big(\hat{q}_1^{(1)}(\alpha) + [N^{2\alpha}]\hat{q}_2^{(1)}(\alpha)\big) \\
&= Q^{(1)} - \frac{\big(\hat{q}_1^{(1)}(\alpha) + [N^{2\alpha}]\hat{q}_2^{(1)}(\alpha)\big)^2}{N^{(1)}(\alpha)}.
\end{aligned} \tag{3.8}$$

Since $Q^{(1)}$ does not depend on α, minimizing $Q_{NT}^{(1)}(\alpha,\kappa)$ is equivalent to maximizing the term:

$$\hat{Q}_{NT}^{(1)}(\alpha) = \frac{(\hat{q}_1^{(1)}(\alpha) + [N^{2\alpha}]\hat{q}_2^{(1)}(\alpha))^2}{N^{(1)}(\alpha)}.$$

In summary, the joint estimator $(\widetilde{\alpha}, \widetilde{\kappa})$ can be obtained by

$$\widetilde{\alpha} = \arg\max_{\alpha}\hat{Q}_{NT}^{(1)}(\alpha) \quad\text{and}\quad \widetilde{\kappa} = \frac{\hat{q}_1^{(1)}(\widetilde{\alpha}) + [N^{4\widetilde{\alpha}}]\hat{q}_2^{(1)}(\widetilde{\alpha})}{N^{(1)}(\widetilde{\alpha})}. \tag{3.9}$$

This joint estimation method estimates α_0 and κ_0 simultaneously. The above derivations show that it is easy to derive $\widetilde{\alpha}$ and then $\widetilde{\kappa}$. Of course, we can also use some other estimation methods to estimate κ_0 and then α_0. Notice that we use the weight function $w(n) = n^3$ in each summation part of the objective function $Q_{NT}^{(1)}(\alpha,\kappa)$ of (3.5). The involvement of a weight function is due to technical necessity in deriving an asymptotic distribution for $(\widetilde{\alpha}, \widetilde{\kappa})$.

3.2 The marginal estimator $\widehat{\alpha}$

Although, for simplicity, the first $[N^{\alpha_0}]$ sections are assumed to possess important factor loadings, the proposed marginal estimation procedure does not rely on the specification of the two categories in the sequence of the N sections.

From Assumption 2, we have

$$\begin{aligned} \bar{\beta}_{N\ell k} &= N^{-1}\sum_{i=1}^{N}\beta_{i\ell k} = \frac{[N^{\alpha_{\ell k}}]}{N}\Big(\frac{\sum_{i=1}^{[N^{\alpha_{\ell k}}]} v_{i\ell k}}{[N^{\alpha_{\ell k}}]}\Big) + \frac{1}{N}\sum_{i=[N^{\alpha_{\ell k}}]+1}^{N}\beta_{i\ell k} \\ &:= [N^{\alpha_{\ell k}-1}]\bar{v}_{N\ell k} + N^{-1}K_{\ell k}, \qquad (3.10) \end{aligned}$$

where $\bar{v}_{N\ell k} = \frac{\sum_{i=1}^{[N^{\alpha_{\ell k}}]} v_{i\ell k}}{[N^{\alpha_{\ell k}}]}$ and $K_{\ell k} = \sum_{i=[N^{\alpha_{\ell k}}]+1}^{N}\beta_{i\ell k}$.

A direct calculation then yields $E(\bar{\beta}_{N\ell k}) = \mu_v[N^{\alpha_{\ell k}-1}] + O(N^{-1})$ and

$$Var(\bar{\beta}_{N\ell k}) = \frac{1}{N^2}[N^{\alpha_{\ell k}}]\sigma_v^2 = [N^{\alpha_{\ell k}-2}]\sigma_v^2. \qquad (3.11)$$

Under model (2.1), it follows that for any $t = 1, 2, \ldots, T-\tau$,

$$\begin{aligned} \bar{x}_t - E\bar{x}_t &= \sum_{\ell=0,k=1}^{s,m}\bar{\beta}_{N\ell k}f_{k,t-\ell} + \bar{u}_t, \\ \bar{x}_{t+\tau} - E\bar{x}_{t+\tau} &= \sum_{\ell=0,k=1}^{s,m}\bar{\beta}_{N\ell k}f_{k,t+\tau-\ell} + \bar{u}_{t+\tau}, \qquad (3.12) \end{aligned}$$

where $\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N}x_{it}$ and $\bar{u}_t = \frac{1}{N}\sum_{i=1}^{N}u_{it}$.

By (3.12), Assumptions 1 and 2, we have

$$\begin{aligned} Cov(\bar{x}_t, \bar{x}_{t+\tau}) &= \sum_{\ell_1,\ell_2=0}^{s}\sum_{k=1}^{m}E(\bar{\beta}_{N\ell_1 k}\bar{\beta}_{N\ell_2 k})E(f_{k,t-\ell_1}f_{k,t+\tau-\ell_2}) + E(\bar{u}_t\bar{u}_{t+\tau}) \qquad (3.13) \\ &= \sum_{\ell=0}^{s}\sum_{k=1}^{m}\Big(E^2(\bar{\beta}_{N\ell k}) + Var(\bar{\beta}_{N\ell k})\Big)E(f_{k,t-\ell}f_{k,t+\tau-\ell}) \\ &+ \sum_{\ell_1\neq\ell_2}^{s}\sum_{k=1}^{m}E(\bar{\beta}_{N\ell_1 k})E(\bar{\beta}_{N\ell_2 k})E(f_{k,t-\ell_1}f_{k,t+\tau-\ell_2}) + E(\bar{u}_t\bar{u}_{t+\tau}). \end{aligned}$$

Substituting (3.10) into (3.13) ensures

$$\begin{aligned} Cov(\bar{x}_t, \bar{x}_{t+\tau}) &= \sum_{\ell=0}^{s}\sum_{k=1}^{m}\Big(\mu_v^2[N^{2\alpha_{\ell k}-2}] + O([N^{\alpha_{\ell k}-2}]) + O(N^{-2})\Big)E(f_{k,t-\ell}f_{k,t+\tau-\ell}) \\ &+ \sum_{\ell_1\neq\ell_2}^{s}\sum_{k=1}^{m}\Big(\mu_v^2[N^{\alpha_{\ell_1 k}+\alpha_{\ell_2 k}-2}] + O([N^{\alpha_{\ell_1 k}-2}]) + O([N^{\alpha_{\ell_2 k}-2}]) + O(N^{-2})\Big)E(f_{k,t-\ell_1}f_{k,t+\tau-\ell_2}) \\ &+ \frac{1}{N^2}\sum_{i,j=1}^{N}E\big(u_{it}u_{j,t+\tau}\big). \end{aligned}$$

Below we consider the case of $\alpha_{\ell k} = \alpha_0$, $\forall \ell = 0, 1, \ldots, s; k = 1, \ldots, m$. Then $Cov(\bar{x}_t, \bar{x}_{t+\tau})$ becomes

$$Cov(\bar{x}_t, \bar{x}_{t+\tau}) = \Big(\sum_{\ell_1,\ell_2=0}^{s} \sum_{k=1}^{m} \mu_v^2 [N^{2\alpha_0 - 2}] + O([N^{\alpha_0 - 2}]) + O(N^{-2}) \Big) E(f_{k,t-\ell_1} f_{k,t+\tau-\ell_2}) + \frac{1}{N^2} \sum_{i,j=1}^{N} E\big(u_{it} u_{j,t+\tau}\big). \quad (3.14)$$

We then compare the orders of the two terms $\sum_{\ell_1,\ell_2=0}^{s} \sum_{k=1}^{m} \mu_v^2 [N^{2\alpha_0-2}] E(f_{k,t-\ell_1} f_{k,t+\tau-\ell_2})$ and $E(\bar{u}_t \bar{u}_{t+\tau})$.

From Assumption 1, we have

$$E(\bar{u}_t \bar{u}_{t+\tau}) = \frac{1}{N^2} \sum_{i,j=1}^{N} E(u_{it} u_{j,t+\tau}) = \frac{1}{N^2} \sum_{i,j=1}^{N} \gamma_1(\tau) \gamma_2(|i-j|) = O\left(\frac{\gamma_1(\tau)}{N} \right),$$

where we have used Condition (2.6).

By Condition (2.9), we have

$$\frac{E(\bar{u}_t \bar{u}_{t+\tau})}{\sum_{\ell_1,\ell_2=0}^{s} \sum_{k=1}^{m} \mu_v^2 [N^{2\alpha_0-2}] E(f_{k,t-\ell_1} f_{k,t+\tau-\ell_2})} = o(1). \quad (3.15)$$

A simple manipulation of (3.14) and (3.15) yields

$$\ln \Big(Cov(\bar{x}_t, \bar{x}_{t+\tau}) \Big)^2 \approx \ln(\kappa_0^2) + (4\alpha_0 - 4) \ln(N),$$

which implies

$$\alpha_0 \approx \frac{\ln \Big(Cov(\bar{x}_t, \bar{x}_{t+\tau}) \Big)^2 - \ln(\kappa_0^2)}{4 \ln(N)} + 1, \quad (3.16)$$

where κ_0 is defined in (3.4).

Hence, for $0 \le \alpha_0 \le 1$, α_0 can be estimated from (3.16) using a consistent estimator for $Cov(\bar{x}_t, \bar{x}_{t+\tau})$ given by

$$\widehat{\sigma}_N(\tau) = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} (\bar{x}_t - \bar{x}^{(1)})(\bar{x}_{t+\tau} - \bar{x}^{(2)}), \quad (3.17)$$

where $\bar{x}^{(1)} = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \bar{x}_t$ and $\bar{x}^{(2)} = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \bar{x}_{t+\tau}$. Thus, a consistent estimator for α_0 is given by

$$\widehat{\alpha} = \frac{\log \Big(\widehat{\sigma}_N(\tau) \Big)^2 - \ln(\kappa_0)^2}{4 \ln(N)} + 1. \quad (3.18)$$

If $\{\mathbf{u}_t : t = 1, \ldots, T\}$ are independent, the term $E(\bar{u}_t \bar{u}_{t+\tau})$ will disappear in (3.14) for any τ. So under this case, we can take a finite lag τ. Furthermore, if not all $\alpha_{\ell k}$ are equal to α_0, we can still get an expression similar to (3.16) but with a different value of κ_0, which can be estimated by the joint estimation method given in the previous section.

3.3 Asymptotic Properties

In this section, we will establish asymptotic distributions for the proposed joint estimator $(\widetilde{\alpha}, \widetilde{\kappa})$ and the marginal estimator $\widehat{\alpha}$, respectively. We assume that $\alpha_{\ell k} = \alpha_0$, $\forall \ell = 0, 1, \ldots, s$ and $k = 1, 2, \ldots, m$ for simplicity. The notation $a \asymp b$ denotes that $a = O(b)$ and $b = O(a)$.

For any $1 \leq i, j \leq m$ and $0 \leq h \leq T-1$, we define

$$C_{ij}(h) = \frac{1}{T-h} \sum_{t=1}^{T-h} f_{i,t} f_{j,t+h}, \quad c_{ij}(h) \equiv \sigma_{ij}(h) = E(f_{i,t} f_{j,t+h}). \tag{3.19}$$

The following theorem establishes an asymptotic distribution for the joint estimator $(\widetilde{\alpha}, \widetilde{\kappa})$.

Theorem 1. *In addition to Assumptions 1-3, we assume that*

(i) for some constant $\delta > 0$,

$$E|\zeta_{it}|^{2+2\delta} < +\infty, \tag{3.20}$$

where ζ_{it} is the i-th component of $\boldsymbol{\zeta}_t$ and $\{\boldsymbol{\zeta}_t : \ldots, -1, 0, 1, \ldots\}$ is the sequence appeared in Assumption 3.

(ii) The lag τ satisfies

$$\frac{\tau}{(T-\tau)^{\delta/(2\delta+2)}} \to 0, \quad as \ \ T \to \infty, \tag{3.21}$$

where δ is defined in (3.20).

(iii) The covariance matrix Γ of the random vector

$$\Big(C_{ij}(h') : i = 1, \ldots, m; j = 1, \ldots, m; h' = \tau - s, \ldots, \tau + s\Big) \tag{3.22}$$

is positive definite.

(iv) As $(N, T) \to (\infty, \infty)$,

$$\begin{aligned}
&\sqrt{\min([N^{\alpha_0}], T-\tau)} \max\Big(\frac{\gamma_1(\tau) N^{1-2\alpha_0}}{\boldsymbol{\mu}_v' \boldsymbol{\Sigma}_\tau \boldsymbol{\mu}_v}, \frac{(T-\tau)^{-1/2} N^{1-2\alpha_0}}{\boldsymbol{\mu}_v' \boldsymbol{\Sigma}_\tau \boldsymbol{\mu}_v}\Big) \to 0, \quad as \ \ 0 < \alpha_0 \leq \frac{1}{2}; \\
&\sqrt{\min([N^{\alpha_0}], T-\tau)} \frac{N^{1/2-\alpha_0}(T-\tau)^{-1/2}}{\boldsymbol{\mu}_v' \boldsymbol{\Sigma}_\tau \boldsymbol{\mu}_v} \to 0, \quad as \ \ \frac{1}{2} < \alpha_0 \leq 1;
\end{aligned} \tag{3.23}$$

and as $(N, T) \to (\infty, \infty)$,

$$\frac{1}{N^{\alpha_0/2} \boldsymbol{\mu}_v' \boldsymbol{\Sigma}_\tau \boldsymbol{\mu}_v} = o(1). \tag{3.24}$$

Then as $(N, T) \to (\infty, \infty)$,

$$[N^{2(\widetilde{\alpha} - \alpha_0)}] - 1 = O_P\big(v_{NT}^{-1/2} \kappa_0^{-1}\big), \tag{3.25}$$

$$\widetilde{\kappa} - \kappa_0 = O_P\big(v_{NT}^{-1/2}\big), \tag{3.26}$$

where $v_{NT} = \min([N^{\alpha_0}], T-\tau)$, *and*

$$\begin{pmatrix} \widetilde{\kappa}\sqrt{v_{NT}}(N^{2(\widetilde{\alpha}-\alpha_0)}-1) \\ \sqrt{v_{NT}}(\widetilde{\kappa}-\kappa_0) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4\sigma_0^2 & -2\sigma_0^2 \\ -2\sigma_0^2 & \sigma_0^2 \end{pmatrix} \right), \tag{3.27}$$

where κ_0 *is defined in (3.4),* $\mathbf{\Sigma}_\tau = E(\mathbf{F}_t\mathbf{F}_{t+\tau}^{'})$ *and* $\boldsymbol{\mu}_v = \mu_v \mathbf{e}_{m(s+1)}$*, in which* $\mathbf{e}_{m(s+1)}$ *is an* $m(s+1)\times 1$ *vector with each element being* 1,

$$\sigma_0^2 = \lim_{N,T\to\infty} \frac{\min([N^{\alpha_0}], T-\tau)}{[N^{\alpha_0}]} 4\boldsymbol{\mu}_v^{'}\mathbf{\Sigma}_\tau\mathbf{\Sigma}_v\mathbf{\Sigma}_\tau\boldsymbol{\mu}_v \tag{3.28}$$
$$+ \lim_{N,T\to\infty} \frac{\min([N^{\alpha_0}], T-\tau)}{T-\tau} (\boldsymbol{\mu}_v^{'} \otimes \boldsymbol{\mu}_v^{'})\mathbf{\Omega}(\boldsymbol{\mu}_v \otimes \boldsymbol{\mu}_v)$$

and

$$\mathbf{\Omega} = \begin{pmatrix} \boldsymbol{\omega}(\tau,\tau) & \cdots & \boldsymbol{\omega}(\tau,\tau+s) & \cdots & \boldsymbol{\omega}(\tau,\tau-s) & \cdots & \boldsymbol{\omega}(\tau,\tau) \\ \boldsymbol{\omega}(\tau+1,\tau) & \cdots & \boldsymbol{\omega}(\tau+1,\tau+s) & \cdots & \boldsymbol{\omega}(\tau+1,\tau-s) & \cdots & \boldsymbol{\omega}(\tau+1,\tau) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\omega}(\tau-s,\tau) & \cdots & \boldsymbol{\omega}(\tau-s,\tau+s) & \cdots & \boldsymbol{\omega}(\tau-s,\tau-s) & \cdots & \boldsymbol{\omega}(\tau-s,\tau) \\ \boldsymbol{\omega}(\tau,\tau) & \cdots & \boldsymbol{\omega}(\tau,\tau+s) & \cdots & \boldsymbol{\omega}(\tau,\tau-s) & \cdots & \boldsymbol{\omega}(\tau,\tau) \end{pmatrix}, \tag{3.29}$$

with $\boldsymbol{\omega}(h,r) = \Big(Cov(f_{i_1t}f_{j_1,t+h}, f_{i_2,t}f_{j_2,t+r}) : 1 \le i_1, j_1, i_2, j_2 \le m\Big)_{m^2\times m^2}$.

While Theorem 1 may just establish a bivariate normal distribution with a singular covariance, it provides a joint distributional structure. We briefly show how to verify Conditions (3.23) and (3.24) based on the simple model (2.10). For model (2.10), $\boldsymbol{\mu}_v^{'}\mathbf{\Sigma}_\tau\boldsymbol{\mu}_v = \mu_v^2 \frac{|\rho_1|^\tau}{1-|\rho_1|^2}$. Then (3.23) and (3.24) are equivalent to the following three cases:

(i) $0 < \alpha_0 \le \frac{1}{2}$, $[N^{\alpha_0}] < T-\tau$, $\frac{|\rho_2|^\tau}{|\rho_1|^\tau} N^{1-2\alpha_0} = o(1)$, $\frac{|\rho_1|^{-\tau}}{(T-\tau)^{\frac{1}{2}} N^{\frac{3}{2}\alpha_0 - 1}} = o(1)$, $\frac{|\rho_1|^{-\tau}}{N^{\frac{\alpha_0}{2}}} = o(1)$.

(ii) $\frac{1}{2} < \alpha_0 \le 1$, $[N^{\alpha_0}] < T-\tau$, $\frac{|\rho_1|^{-\tau}}{N^{\alpha_0-1}(T-\tau)^{\frac{1}{2}}} = o(1)$, $\frac{|\rho_1|^{-\tau}}{N^{\frac{\alpha_0}{2}}} = o(1)$.

(iii) $\frac{1}{2} < \alpha_0 < 1$, $[N^{\alpha_0}] \ge T-\tau$, $\frac{|\rho_1|^{-\tau}}{N^{\alpha_0 - \frac{1}{2}}} = o(1)$.

For each of these three cases, we provide a choice of (T,N) and ρ_2.

(a) $0 < \alpha_0 \le \frac{1}{2}$, $[N^{\alpha_0}] < T-\tau$, $N = [(|\rho_1|^{-1}+\delta)^{\frac{2\tau}{\alpha_0}}]$, $T = [(|\rho_1|^{-1}+\delta)^{\frac{4\tau}{\alpha_0}-4\tau}] + \tau$, $|\rho_2| = \frac{1}{2}|\rho_1| \cdot [N^{(\frac{3}{2}\alpha_0 - 1)\frac{1}{\tau}}]$.

(b) $\frac{1}{2} < \alpha_0 \le 1$, $[N^{\alpha_0}] < T-\tau$, $N = [(|\rho_1|^{-1}+\delta)^{\frac{2\tau}{\alpha_0}}]$, $T = [(|\rho_1|^{-1}+\delta)^{\frac{4\tau}{\alpha_0}-2\tau}] + \tau$.

(c) $\frac{1}{2} < \alpha_0 \le 1$, $[N^{\alpha_0}] \ge T-\tau$, $N = [(|\rho_1|^{-1}+\delta)^{\frac{\tau}{\alpha_0-\frac{1}{2}}}]$, $T = [(|\rho_1|^{-1}+\delta_0)^{\frac{\tau\alpha_0}{\alpha_0-\frac{1}{2}}}] + \tau$, where $0 < \delta_0 < \delta$, and $\delta > 0$ is a constant.

The following theorem establishes an asymptotic distribution for the marginal estimator $\widehat{\alpha}$.

Theorem 2. *Under the conditions of Theorem 1, we have*

$$\sqrt{\min([N^{\alpha_0}], T-\tau)} \frac{(N^{4(\widehat{\alpha}-\alpha_0)}-1)\kappa_0}{\sqrt{4\sigma_0^2}} \to \mathcal{N}(0,1), \tag{3.30}$$

where σ_0^2 *is defined in (3.28).*

Theorem 1 establishes some asymptotic properties for the joint estimator $(\widetilde{\alpha}, \widetilde{\kappa})$. This result is consistent with that for the marginal estimator $\widehat{\alpha}$ derived in Theorem 2.

From Theorem 2, one can see that $\widehat{\alpha}$ is a consistent estimator of α_0. Moreover, by a careful inspection on the proof of Theorem 2 one can see that Condition (3.23) is not needed to ensure the consistency of $\widehat{\alpha}$ under $(N, T) \to (\infty, \infty)$.

When the idiosyncratic components are independent, we can just use a finite lag τ instead of requiring $\tau \to \infty$. In this case, an asymptotic distribution for the estimator $\widehat{\alpha}$ is established in the following theorem.

Theorem 3. *In addition to Assumptions 2 and 3, suppose that τ is fixed and the following conditions (i)–(iii) hold:*

(i) $\{\mathbf{u}_t : t = 1, \ldots, T\}$ are independent with the mean of $\mathbf{u}_t$ being $\mathbf{0}_{N\times 1}$ and its covariance matrix being $\mathbf{\Sigma}_{\mathbf{u}}$, where $\mathbf{0}_{N\times 1}$ is an $N \times 1$ vector with zero components and the spectral norm $||\mathbf{\Sigma}_{\mathbf{u}}||$ is bounded.

(ii)

$$\sqrt{\min([N^{\alpha_0}], T-\tau)}\frac{N^{1-2\alpha_0}}{(T-\tau)^{1/2}} \to 0, \quad as \;\; 0 < \alpha_0 < \frac{1}{2}; \tag{3.31}$$

$$\sqrt{\min([N^{\alpha_0}], T-\tau)}\frac{N^{1/2-\alpha_0}}{(T-\tau)^{1/2}} \to 0, \quad as \;\; \frac{1}{2} < \alpha_0 \leq 1.$$

(iii) $\boldsymbol{\mu}_v^{'}\mathbf{\Sigma}_\tau\boldsymbol{\mu}_v \neq 0$.

Then, as $(N, T) \to (\infty, \infty)$, we have

$$\frac{\sqrt{\min([N^{\alpha_0}], T-\tau)}\, \ln(N^2)\left(\widehat{\alpha} - \alpha_0\right)}{\sqrt{\sigma_0^2/\kappa_0^2}} \xrightarrow{d} \mathcal{N}\Big(0, 1\Big),$$

where κ_0 and σ_0^2 are defined in (3.4) and (3.28), respectively.

Before we will give the proofs of Theorems 1–3 in Appendices B and C below, we have some brief discussion about Condition (3.31), which is actually equivalent to the following three cases:

(a) $0 < \alpha_0 \leq \frac{1}{2}$, $[N^{\alpha_0}] < T - \tau$, $\frac{N^{1-\frac{3}{2}\alpha_0}}{(T-\tau)^{\frac{1}{2}}} = o(1)$;

(b) $\frac{1}{2} < \alpha_0 \leq 1$, $[N^{\alpha_0}] < T - \tau$; $\frac{N^{\frac{1}{2}-\frac{\alpha_0}{2}}}{(T-\tau)^{\frac{1}{2}}} = o(1)$;

(c) $\frac{1}{2} < \alpha_0 \leq 1$, $[N^{\alpha_0}] \geq T - \tau$, $N^{\frac{1}{2}-\alpha_0} = o(1)$.

Under these three cases, we can provide some choices for (N, T) as follows:

(d) $0 < \alpha_0 < \frac{1}{2}$, $[N^{\alpha_0}] < T - \tau$; $T = \tau + [N^{2-3\alpha_0}\log(N)]$;

(f) $\frac{1}{2} < \alpha_0 \leq 1$, $[N^{\alpha_0}] < T - \tau$, $T = \tau + [N^{\alpha_0}\log(N)]$;

(g) $\frac{1}{2} < \alpha_0 \leq 1$, $[N^{\alpha_0}] \geq T - \tau$, $T = \tau + [N^{\alpha_0}/\log(N)]$.

When $\tau \to \infty$, the term $\boldsymbol{\mu}_v^{'}\mathbf{\Sigma}_\tau\boldsymbol{\mu}_v$ will tend to 0, because of $\mathbf{\Sigma}_\tau \to \mathbf{0}$. So, as τ is very large, the value of $\ln(\boldsymbol{\mu}_v^{'}\mathbf{\Sigma}_\tau\boldsymbol{\mu}_v)$ may be negative in practice. Hence Theorem 2 provides an alternative form for the asymptotic distribution of $N^{\widehat{\alpha}-\alpha_0}$ instead of $\widehat{\alpha}-\alpha_0$, and the case of τ being fixed is discussed in Theorem 3.

We now evaluate the finite–sample performance of the proposed estimation methods and the resulting theory in Sections 4 and 5 below.

4 Simulation

4.1 Data Generating Process 1

First, we consider the following two-factor static model

$$x_{it} = \mu + \beta_{i1} f_{1t} + \beta_{i2} f_{2t} + u_{it}, \quad i = 1, 2, \ldots, N; t = 1, 2, \ldots, T. \tag{4.1}$$

The factors are generated by

$$f_{jt} = \rho_j f_{j,t-1} + \sqrt{1-\rho_j^2}\,\zeta_{jt}, \quad j = 1, 2;\ t = -49, -48, \ldots, 0, 1, \ldots, T, \tag{4.2}$$

with $f_{j,-50} = 0$ for $j = 1, 2$ and $\zeta_{jt} \overset{i.i.d}{\sim} \mathcal{N}(0, 1)$. The idiosyncratic components are generated by

$$u_{it} = \varepsilon_t \eta_i \quad \text{for } i = 1, 2, \cdots, N \text{ and } t = 1, 2, \ldots, T, \tag{4.3}$$

in which $\eta_i \overset{i.i.d}{\sim} N(0, 1)$,

$$\varepsilon_t = \rho_2 \varepsilon_{t-1} + \epsilon_t, \quad t = 1, 2, \ldots, T \tag{4.4}$$

and $\epsilon_t \overset{i.i.d}{\sim} N(0, 1)$, and $\{\epsilon_t : t = 1, 2, \ldots, T\}$ are independent of $\{\zeta_{jt} : t = 1, 2, \ldots, T; j = 1, 2\}$.

The factor loadings are generated as

$$\begin{aligned} &\beta_{ir} = v_{ir}, \quad for\ i = 1, 2, \ldots, M; r = 1, 2; \\ &\beta_{ir} = \rho^{i-M}, \quad for\ i = M+1, M+2, \ldots, N; r = 1, 2, \end{aligned} \tag{4.5}$$

where $v_{ir} \overset{i.i.d}{\sim} U(0.5, 1.5)$, $M = [N^{\alpha_0}]$ and $\rho = 0.5$. Moreover, we set $\mu = 1$ and $\rho_j = 0.5$ for $j = 1, 2$.

4.2 Data Generating Process 2

Second, we consider a dynamic model as follows.

$$x_{it} = \mu + \beta_{i,01} f_{1t} + \beta_{i,02} f_{2t} + \beta_{i,11} f_{1,t-1} + \beta_{i,12} f_{2,t-1} + u_{it}, \quad i = 1, \ldots, N; t = 1, \ldots, T, \tag{4.6}$$

where $\mu = 1$, and the factor loadings are generated as

$$\begin{aligned} &\beta_{i,jk} = v_{i,jk}, \quad for \quad i = 1, 2, \ldots, M_k; \\ &\beta_{i,jk} = \rho^{i-M_k}, \quad for \quad i = M_k + 1, M_k + 2, \ldots, N, \end{aligned}$$

for $j = 0, 1$, $k = 1, 2$, $M_k = [N^{\alpha_0}]$ and $\rho = 0.5$, in which $v_{i,jk} \overset{i.i.d}{\sim} U(0.5, 1.5)$.

The generating procedures for $\mathbf{f}_t$ and u_{it} are the same as those in Data Generating Process 1. The factor loadings are generated as (4.5) in the first static model.

For the two data generating processes, we consider values of $\alpha_0 = 0.2, 0.4, 0.6, 0.8, 1$, $N = 100, 200, 500, 1000$ and $T = 100, 200, 500$. All the experiments are based on 500 replications. For each replication, the values of α, ρ_1, ρ_2 are given as above. These parameters are fixed across all replications. The values of $v_{ji}, j = 1, 2$ are drawn randomly for each replication.

The bias and root mean square error (RMSE) results for the marginal estimator $\widehat{\alpha}$ and joint estimator $\widetilde{\alpha}$ are summarized in Tables 1-4, and show that the proposed estimation methods work well numerically.

5 Empirical applications

In this section, we show how to obtain an estimate for the exponent of cross–sectional dependence, α_0, for each of the following panel data sets: quarterly cross-country data used in global modelling and daily stock returns on the constitutes of Standard and Poor 500 index.

5.1 Cross-country dependence of macro-variables

We provide an estimate for α_0 for each of the datasets: Real GDP growth (RGDP), Consumer price index (CPI), Nominal equity price index (NOMEQ), Exchange rate of country i at time t expressed in US dollars (FXdol), Nominal price of oil in US dollars (POILdolL), and Nominal short-term and long-term interest rate per annum (Rshort and Rlong) computed over 33 countries.* The observed cross-country time series, y_{it}, over the full sample period, are standardized as $x_{it} = (y_{it} - \bar{y}_i)/s_i$, where $\bar{y}_i$ is the sample mean and s_i is the corresponding standard deviation for each of the time series. Table 5 reports the corresponding results.

For the standardized data x_{it}, we regress it on the cross-section mean $\bar{x}_t = \frac{1}{N}\sum_{i=1}^N x_{it}$, i.e., $x_{it} = \delta_i \bar{x}_t + u_{it}$ for $i = 1, 2, \ldots, N$, where δ_i, $i = 1, 2, \ldots, N$, are regression coefficients. With the availability of the OLS estimate $\widehat{\delta_i}$ for δ_i, we have the estimated versions, $\widehat{u}_{it}$, of the form: $\widehat{u}_{it} = x_{it} - \hat{\delta}_i \bar{x}_t$.

Since our proposed estimation methods rely on the different extent of serial dependence of the factors and idiosyncratic components, we provide some autocorrelation graphs of $\{\bar{x}_t = \frac{1}{N}\sum_{i=1}^N x_{it} : t = 1, 2, \ldots, T\}$ and $\{\bar{u}_t = \frac{1}{N}\sum_{i=1}^N u_{it} : t = 1, 2, \ldots, T\}$ for each group of the real dataset under investigation (see Figures 1–4). From these graphs, it is easy to see that CPI, NOMEQ, FXdol and POILdolL have distinctive serial dependence in the factor part $\bar{x}_t$ and idiosyncratic part $\bar{u}_t$. All the observed real data x_{it} are serially dependent.

Figures 1-4 near here

*The datasets are downloaded from http://www-cfap.jbs.cam.ac.uk/research/gvartoolbox/download.html.

Due to existence of serial dependence in the idiosyncratic components, we use the proposed second moment criterion. The marginal estimator $\widehat{\alpha}$ and the joint estimator $\widetilde{\alpha}$ for these real data are provided in Table 5. We use $\tau = 10$ for two estimators. We can see from Table 5 that the values of $\widehat{\alpha}$ and $\widetilde{\alpha}$ are different from the those provided by Bailey, Kapatanios and Pesaran (2015). Some estimated values are not 1. This phenomenon implies that a factor structure might be a good approximation for modeling global dependencies, and the value of $\alpha_0 = 1$ typically assumed in the empirical factor literature might be exaggerating the importance of the common factors for modelling cross-sectional dependence at the expense of other forms of dependencies that originate from trade or financial inter-linkage that are more local or regional rather than global in nature. Furthermore, note that our model is different from that given by Bailey, Kapatanios and Pesaran (2015) (BKP15) and difference mainly lies on that our model only imposes serial dependence on factor processes and assumes that the idiosyncratic errors are independent. Different models may bring in different exponents.

Table 5 near here

5.2 Cross-sectional exponent of stock-returns

One of the important considerations in the analysis of financial markets is the extent to which asset returns are interconnected. The classical model is the capital asset pricing model of Sharp (1964) and the arbitrage pricing theory of Ross (1976). Both theories have factor representations with at least one strong common factor and an idiosyncratic component that could be weakly cross-sectional correlated (see Chamberlain (1983)). The strength of the factors in these asset pricing models is measured by the exponent of the cross-sectional dependence, α_0. When $\alpha_0 = 1$, as it is typically assumed in the literature, all individual stock returns are significantly affected by the factors, but there is no reason to believe that this will be the case for all assets and at all times. The disconnection between some asset returns and the market factors could occur particularly at times of stock market booms and busts where some asset returns could be driven by some non-fundamentals. Therefore, it would be of interest to investigate possible time variations in the exponent α_0 for stock returns.

We base our empirical analysis on daily returns of 96 stocks in the Standard & Poor 500 (S&P500) market during the period of January, 2011-December, 2012. The observations r_{it} are standardized as $x_{it} = (r_{it} - \bar{r}_i)/s_i$, where $\bar{r}_i$ is the sample mean of the returns over all the sample and s_i is the corresponding standard deviations. For the standardized data x_{it}, we regress it on the cross-section mean $\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N} x_{it}$, i.e., $x_{it} = \delta_i \bar{x}_t + u_{it}$ for $i = 1, 2, \ldots, N$, where δ_i, $i = 1, 2, \ldots, N$, are the regression coefficients. Based on the OLS estimates: $\widehat{\delta}_i$ for δ_i, we define $\widehat{u}_{it} = x_{it} - \hat{\delta}_i \bar{x}_t$. The autocorrelation functions (ACFs) of the cross-sectional averages $\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N} x_{it}$ and $\bar{u}_t = \frac{1}{N}\sum_{i=1}^{N} u_{it}$ are presented in Figure 5.

Figure 5 near here

From Figure 5, we can see that the serial dependences of the common factor components are stronger than those of the idiosyncratic components. We use the estimates $\widehat{\alpha}$ and $\widetilde{\alpha}$ to characterize the serial dependences of the common factors and the idiosyncratic components. The estimates $\widehat{\alpha}$ and $\widetilde{\alpha}$ are calculated with the choice of $\tau = 10$. Table 6 reports the estimates with several different sample sizes. As comparison, the estimates from BKP15 are also reported. From the table, we can see that their estimation method does not work when α is smaller than 1/2. The results also show that the cross-sectional exponent of stock returns in S&P500 are smaller than 1. This indicates the support of using different levels of loadings for the common factor model as assumed in Assumption 2, rather than using the same level of loadings in such scenarios. .

Table 6 near here

Furthermore, Figure 6 provides the marginal estimate $\widehat{\alpha}$ and the joint estimate $\widetilde{\alpha}$ for the first 130 days of all the period. It shows that the estimated values for α_0 with the two methods are quite similar. On the other hand, since a 130-day period is short, meanwhile, it is reasonable that the estimates didn't change very much.

Figure 6 near here

6 Conclusions and discussion

In this paper, we have examined the issue of how to estimate the extent of cross-sectional dependence for large dimensional panel data. The extent of cross-sectional dependence is parameterized as α_0, by assuming that only $[N^{\alpha_0}]$ sections are relatively strongly dependent. Compared to the estimation method proposed by BKP15, we have proposed using a dynamic factor model to characterize the extent of inter-connections in large panel data and developed a new 'moment' method to estimate α_0. In detail, based on the assumption that stronger serial dependence exists in the factor process than that for the idiosyncratic errors, we have recommended the use of the covariance function between the cross-sectional average values of the observed data at different lags to estimate α_0. One main advantage of this new approach is that it can deal with the case of $0 \le \alpha_0 \le 1/2$.

Due to some unknown parameters involved in the panel data model, in addition to the proposed marginal estimator, we have also construct a joint estimation method for α_0 and the related unknown parameters. The asymptotic properties of all the estimators have all been established. The simulation results and an empirical application to two datasets have shown that our estimation methods work well numerically.

Future research includes discussions about how to estimate factors and factor loadings in factor models, and determine the number of factors for the case of $0 < \alpha_0 < 1$. Existing methods available for factor models, such as Bai and Ng (2002), Bai (2003), Onatski (2009), for the case of $\alpha_0 = 1$, may not be applicable, and should be extended to deal with the case of $0 < \alpha_0 < 1$. Such issues are all left for future work.

7 Acknowledgements

The first and the third authors acknowledge the Australian Research Council Discovery Grants Program for its support under Grant numbers: DP130104229 & DP150101012. Thanks from the second author also go to the Ministry of Education in Singapore for its financial support under Grant # ARC 14/11.

References

T.W. Anderson (1994). The statistical analysis of time series. *London: Wiley* **2th edtion**.

J. S. Bai (1997). Estimating multiple breaks one at a time. *Econometric Theory* **13(3)**, 315-352.

J. S. Bai, P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* **66(1)**, 47-78.

J. S. Bai, S. Ng (2002). Determine the number of factors in approximate factor models. *Econometrica* **70(1)**, 191-221.

J. S. Bai (2003). Inferential theory for factor models of large dimensions. *Econometrica* **71(1)**, 135-171.

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

N. Bailey, G. Kapetanios, M.H. Pesaran (2015). Exponent of cross-sectional dependence: estimation and inference. Available at *Journal of Applied Econometrics* **DOI: 10.1002/jae.2476**.

B.H. Baltagi, Q. Feng, C. Kao (2012). A large multiplier test for cross-sectional dependence in a fixed effects panel data model. *Journal of Econometrics* **170(1)**, 164-177.

J.M. Bardet, P. Doukhan, J.R. León (2008). A functional limit theorem for η-weakly dependent processes and its applications. *Statistical Inference for Stochastic Processes* **11(3)**, 265-280.

G. Chamberlain (1983). Funds, factors and diversification in arbitrage pricing theory. *Econometrica* **51(5)**, 1305-1323.

J. Chen, J. Gao, D. G. Li (2012). A new diagnostic test for cross-section uncorrelatedness in nonparametric panel data models. *Econometric Theory* **28**, 1144-1163.

J. Q. Fan, Y. Y. Fan, J. C. Lv (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics* **147**, 186-197.

J. Fan, Y. Liao, M. Mincheva (2011). High dimensional covariance matrix estimation in approximate factor models. *Annals of Statistics* **39**, 3320-3356.

E. J. Hannan (1970). Multiple time series. *John Wiley and Sons, Inc.* .

C. Hsiao, M. H. Pesaran, A. Pick (2012). Diagnostic tests of cross-section independence for limited dependent variable panel data models. *Oxford Bulletion of Econometrics and Statistics* **74**, 253-277.

C. Lam, Q. W. Yao (2012). Factor modeling for high-dimensional time series: inference for the number of factors. *Annals of Statistics* **40(2)**, 694-726.

S. Ng (2006). Testing cross-section correlation in panel data using spacings. *Journal of Business and Economic Statistics* **24**, 12–23.

A. Onatski (2009). Testing hypothesis about the number of factors in large factor models. *Econometrica* **77(5)**, 1447-1479.

J. Pan, Q. Yao (2008). Modelling multiple time series via common factors. *Biometrika* **95(2)**, 365-379.

M. H. Pesaran (2004). General diagnostic test for cross section dependence in panels. *Working Paper* at **University of Cambridge & USC**.

S. Ross (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* **13**, 341-360.

J.P. Romano, M. Wolf (2000). A more general central limit theorem for m-dependent random variables with unbounded m. *Statistics and Probability Letters* **47**, 115-124.

V. Sarafidis, T. Wansbeek (2012). Cross-sectional dependence in panel data analysis. *Econometric Reviews* **31(5)**, 483-531.

W. Sharpe (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* **19(3)**, 425-442.

Figure 1: ACF of RGDP and CPI

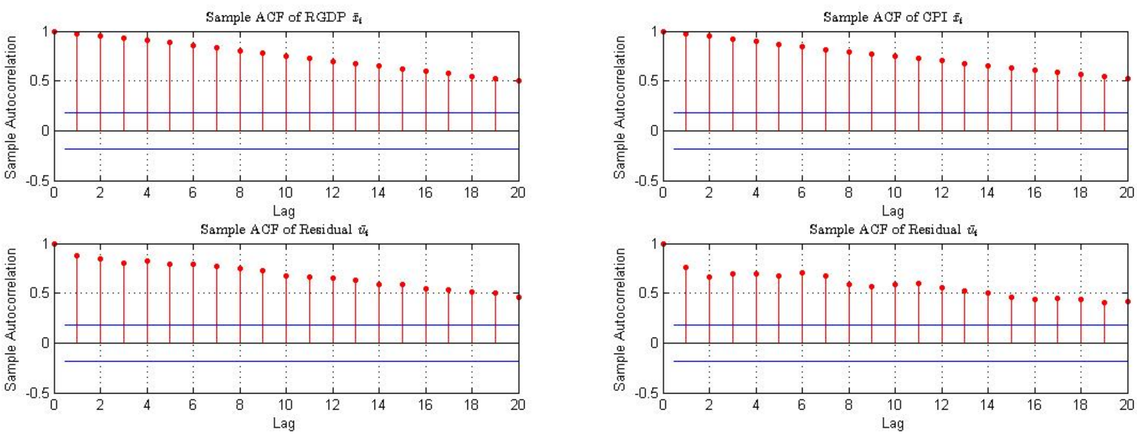


Figure 2: ACF of NOMEQ and FXdol

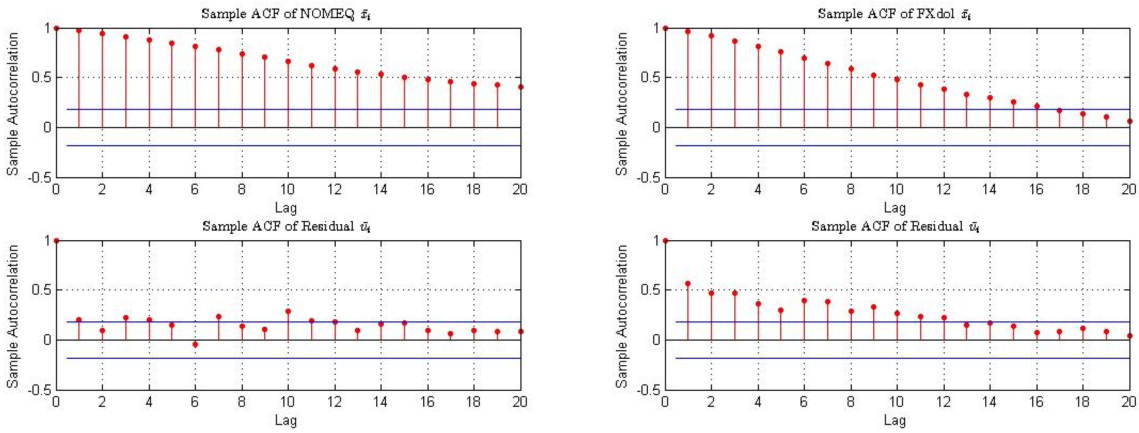
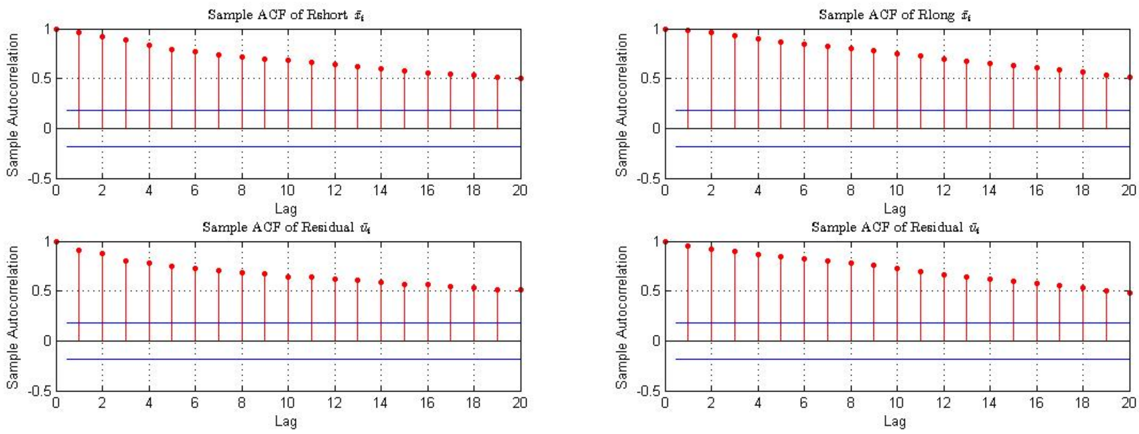


Figure 3: ACF of Rshort and Rlong



This material includes three appendices, i.e. Appendices A–C. Appendix A presents justification of Assumptions 1–3 in the main paper. Appendix B provides the proofs of Theorems 1 and 2 in the main paper. Some lemmas used in the proofs of Theorems 1 and 2 are given in Appendix C. The proof of Theorem 3 in the main paper is omitted since it is similar to that of Theorem 2.

Throughout this material, we use C to denote a constant which may be different from line to line and $||\cdot||$ to denote the spectral norm or the Euclidean norm of a vector. In addition, the notation $a_n \asymp b_n$ means that $a_n = O_P(b_n)$ and $b_n = O_P(a_n)$.

8 Appendix A: Justifications of Assumptions

In this section, we provide some comments on Assumptions 1–3 in the main paper. The three assumptions are mild and can be satisfied in many cases. Next, we will discuss them in detail.

1. Justification of Assumption 1: The weak stationarity assumption on the idiosyncratic components $\{\mathbf{u}_t : t = 1, 2, \ldots, T\}$ is a commonly used condition in time series analysis. Rather than independence assumption, weak cross-sectional correlation and serial correlation are imposed via $\gamma_2(|i-j|)$ and $\gamma_1(\tau)$, respectively. The levels of weakness are described by (2.6) and (2.8). Note

Figure 4: ACF of POILdolL

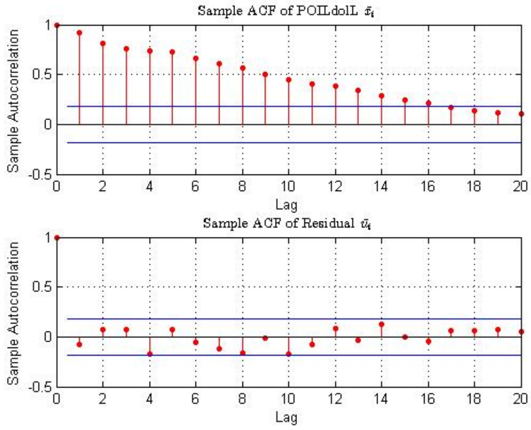
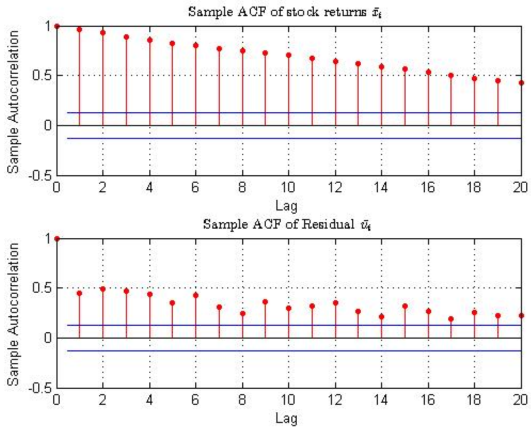


Figure 5: ACF of averages of 96 stock returns



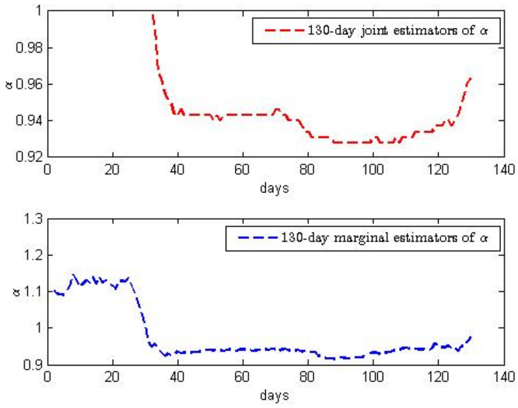
that when $\{u_{it}\}$ is independent across (i, t), we have $\gamma_1(\tau) = 0$ and $\gamma_2(|i-j|) = 0$ which satisfy Conditions (2.6).

2. Justification of Assumption 2: The degree of cross-sectional dependence in $\{\mathbf{x}_t : t = 1, 2, \ldots, N\}$ crucially depends on the nature of the factor loadings. This assumption groups the factor loadings into two categories: a strong category with effects that are bounded away from zero, and a weak category with transitory effects that tend to zero. From this point, the first $[N^{\alpha_0}]$ sections are dependent while the rest are independent. Here $\alpha_0 = \max(\alpha_{\ell k} : \ell = 0, 1, 2, \ldots, s; k = 1, 2, \ldots, m)$.

   To simplify the proof of Theorem 2, we require the factor loadings to have the finite sixth moments. However, we believe that the finite second moment condition may just be sufficient by performing the truncation technique in the proof of Lemma 3.

3. Justification of Assumption 3: The common factors $\{\mathbf{f}_t : t = 1, 2, \ldots, T\}$ are also weak stationary time series. The important condition (2.9), which connects $\{\mathbf{f}_t : t = 1, 2, \ldots, T\}$ and $\{u_{it} : i = 1, 2, \ldots, N; t = 1, 2, \ldots, T\}$, requires stronger serial dependence existed in the factors than that in the idiosyncratic components. This requirement assures the leading term position for the common factor part rather than the idiosyncratic part.

Figure 6: 130-day joint and marginal estimators for 96 stocks of S&P 500



9 Appendix B: Proofs of Theorems 1 and 2

This section provides the proofs of Theorems 1 and 2. The proofs will use Lemmas 1 and 2, which are given Appendix C below. For easy of presentation, we first prove Theorem 2 which is for the marginal estimator.

9.1 Proof of Theorem 2

Proof. Based on model (2.2) in the main paper, we have

$$\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N} x_{it} = \mu + \bar{\boldsymbol{\beta}}_N' \mathbf{F}_t + \bar{u}_t,$$

where $\bar{\boldsymbol{\beta}}_N' = 1/N\sum_{i=1}^{N}\boldsymbol{\beta}_i$, $\mu = 1/N\sum_{i=1}^{N}\mu_i$ and $\bar{u}_t = \frac{1}{N}\sum_{i=1}^{N} u_{it}$. Then we have

$$\bar{x}^{(1)} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{x}_t = \mu + \bar{\boldsymbol{\beta}}_N'\bar{\mathbf{F}}_T + \bar{u}^{(1)}, \quad \bar{x}^{(2)} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{x}_{t+\tau} = \mu + \bar{\boldsymbol{\beta}}_N'\bar{\mathbf{F}}_{T+\tau} + \bar{u}^{(2)},$$

where $\bar{\mathbf{F}}_T = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\mathbf{F}_t$, $\bar{\mathbf{F}}_{T+\tau} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\mathbf{F}_{t+\tau}$, $\bar{u}^{(1)} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{u}_t$ and $\bar{u}^{(2)} = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{u}_{t+\tau}$.

Then the auto-covariance estimator $\widehat{\sigma}_N(\tau)$ can be written as

$$\begin{aligned}
\widehat{\sigma}_N(\tau) &= \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\Big(\big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_t - \bar{\mathbf{F}}_T) + \bar{u}_t - \bar{u}^{(1)}\big)\big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_{t+\tau} - \bar{\mathbf{F}}_{T+\tau}) + \bar{u}_{t+\tau} - \bar{u}^{(2)}\big)\Big) \\
&= \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\Big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_t - \bar{\mathbf{F}}_T)(\mathbf{F}_{t+\tau} - \bar{\mathbf{F}}_{T+\tau})'\bar{\boldsymbol{\beta}}_N\Big) + C_N, \qquad \text{(B.1)}
\end{aligned}$$

where $C_N = c_{N1} + c_{N2} + c_{N3}$ with

$$\begin{aligned}
c_{N1} &= \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\Big((\bar{u}_t - \bar{u}^{(1)})(\bar{u}_{t+\tau} - \bar{u}^{(2)})\Big), \\
c_{N2} &= \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\Big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_t - \bar{\mathbf{F}}_T)(\bar{u}_{t+\tau} - \bar{u}^{(2)})\Big), \\
c_{N3} &= \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\Big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_{t+\tau} - \bar{\mathbf{F}}_{T+\tau})(\bar{u}_t - \bar{u}^{(1)})\Big).
\end{aligned}$$

Table 1: DGP1: Bias and RMSE for the marginal estimator $\widehat{\alpha}$ with $\tau = \left[T^{1/4}\right]$.

	α_0	0.2	0.4	0.6	0.8	1
N/T				100		
100	Bias	-0.0565	-0.0159	-0.0197	-0.0102	-0.0085
	RMSE	0.1480	0.0618	0.0473	0.0362	0.0388
200	Bias	-0.0410	-0.0116	-0.0078	-0.0121	-0.0104
	RMSE	0.1175	0.0482	0.0362	0.0349	0.0359
500	Bias	-0.0029	-0.0104	-0.0110	-0.0089	-0.0098
	RMSE	0.0988	0.0448	0.0323	0.0318	0.0301
1000	Bias	-0.0249	-0.0175	-0.0079	-0.0084	-0.0098
	RMSE	0.0802	0.0501	0.0268	0.0265	0.0267
N/T				200		
100	Bias	-0.0371	-0.0030	-0.0104	-0.0069	-0.0077
	RMSE	0.1001	0.0362	0.0294	0.0276	0.0277
200	Bias	-0.0480	-0.0037	-0.0041	-0.0063	-0.0053
	RMSE	0.1160	0.0319	0.0251	0.0236	0.0220
500	Bias	-0.0126	-0.0053	-0.0026	-0.0039	-0.0042
	RMSE	0.0888	0.0290	0.0192	0.0182	0.0195
1000	Bias	0.0045	-0.0114	-0.0027	-0.0040	-0.0035
	RMSE	0.0816	0.0323	0.0178	0.0175	0.0171
N/T				500		
100	Bias	-0.0115	0.0032	-0.0066	-0.0037	-0.0026
	RMSE	0.0580	0.0219	0.0192	0.0167	0.0165
200	Bias	-0.0534	0.0645	0.0016	-0.0023	-0.0021
	RMSE	0.0102	0.0210	0.0150	0.0137	0.0144
500	Bias	-0.0198	0.0035	-0.0039	-0.0017	-0.0008
	RMSE	0.0886	0.0177	0.0131	0.0127	0.0123
1000	Bias	-0.0226	-0.0065	-0.0009	-0.0015	-0.0017
	RMSE	0.0825	0.0175	0.0109	0.0109	0.0112

Denote

$$\mathbf{S}_\tau = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}(\mathbf{F}_t - \bar{\mathbf{F}}_T)(\mathbf{F}_{t+\tau} - \bar{\mathbf{F}}_{T+\tau})'.$$

From (3.10) in the main paper, we can obtain

$$\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N = [N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N + R_N, \tag{B.2}$$

where

$$R_N = [N^{\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\mathbf{K}_\rho + [N^{\alpha_0-2}]\mathbf{K}_\rho'\mathbf{S}_\tau\bar{\mathbf{v}}_N + N^{-2}\mathbf{K}_\rho'\mathbf{S}_\tau\mathbf{K}_\rho. \tag{B.3}$$

Here we would like to remind the reader that $\mathbf{D}_N$ becomes an identity matrix since we assume that

Table 2: DGP1: Bias and RMSE for the joint estimator $\widetilde{\alpha}$ with $\tau = \left[T^{1/4}\right]$.

	α_0	0.2	0.4	0.6	0.8	1
N/T				100		
100	Bias	-0.0826	-0.0627	-0.0473	-0.0267	-0.0181
	RMSE	0.0788	0.0428	0.0276	0.0220	0.0118
200	Bias	-0.0327	-0.0253	-0.0138	-0.0114	-0.0108
	RMSE	0.0482	0.0279	0.0152	0.0138	0.0109
500	Bias	-0.0126	-0.0114	-0.0105	-0.0107	-0.0102
	RMSE	0.0153	0.0128	0.0110	0.0108	0.0105
1000	Bias	-0.0112	-0.0113	-0.0100	-0.0106	-0.0098
	RMSE	0.0143	0.0114	0.0109	0.0107	0.0110
N/T				200		
100	Bias	-0.0462	-0.0323	-0.0217	-0.0112	-0.0099
	RMSE	0.0501	0.0228	0.0184	0.0117	0.0112
200	Bias	-0.0152	-0.0116	-0.0108	-0.0100	-0.0094
	RMSE	0.0368	0.0282	0.0189	0.0168	0.0119
500	Bias	-0.0135	-0.0101	-0.0077	-0.0063	-0.0076
	RMSE	0.0288	0.0178	0.0124	0.0124	0.0111
1000	Bias	-0.0129	-0.0105	-0.0059	-0.0054	-0.0060
	RMSE	0.0106	0.0112	0.0117	0.0109	0.0111
N/T				500		
100	Bias	-0.0127	-0.0119	-0.0109	-0.0098	-0.0069
	RMSE	0.0315	0.0218	0.0124	0.0117	0.0112
200	Bias	-0.0110	-0.0099	-0.0085	-0.0071	-0.0064
	RMSE	0.0211	0.0108	0.0108	0.0102	0.0102
500	Bias	-0.0079	-0.0074	-0.0080	-0.0069	-0.0045
	RMSE	0.0110	0.0118	0.0114	0.0110	0.0105
1000	Bias	-0.0066	-0.0060	-0.0070	-0.0085	-0.0068
	RMSE	0.0100	0.0112	0.0110	0.0102	0.0110

$\alpha_{lk} = \alpha_0$ for simplicity. Therefore, from (B.1) and (B.2), we have

$$\begin{aligned}\ln(\widehat{\sigma}_N(\tau))^2 &= \ln(\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N)^2 + \ln(1+\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N})^2 \\ &= 4(\alpha_0-1)\ln(N) + \ln(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N)^2 \\ &\quad + \ln\left(1+\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\right)^2 + \ln\left(1+\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}\right)^2. \end{aligned} \tag{B.4}$$

It follows from (3.16) in the main paper and (B.4) that

$$\begin{aligned}&4(\widehat{\alpha}-\alpha_0)\ln(N) + \ln(\kappa_0^2) - \ln(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N)^2 \\ =\ &\ln\left(1+\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\right)^2 + \ln\left(1+\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}\right)^2. \end{aligned} \tag{B.5}$$

From Lemma 1 in Appendix C, which provides the central limit theorem for $\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N$, and condition (3.24) in the main paper, we conclude that, as $N, T \to \infty$,

$$\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v \xrightarrow{i.p.} 0. \tag{B.6}$$

Table 3: DGP2: Bias and RMSE for the marginal estimator $\widehat{\alpha}$ with $\tau = \left[T^{1/4}\right]$.

	α_0	0.2	0.4	0.6	0.8	1
N/T				100		
100	Bias	-0.0937	0.0532	-0.0423	-0.0240	-0.0163
	RMSE	0.0879	0.1034	0.0676	0.0499	0.0396
200	Bias	-0.0453	-0.0184	-0.0109	-0.0110	-0.0097
	RMSE	0.0793	0.0642	0.0552	0.0399	0.0239
500	Bias	-0.0106	-0.0112	-0.0095	-0.0069	-0.0058
	RMSE	0.0423	0.0329	0.0310	0.0208	0.0119
1000	Bias	-0.0152	-0.0103	-0.0099	-0.0102	-0.0098
	RMSE	0.0293	0.0199	0.0183	0.0175	0.0148
N/T				200		
100	Bias	-0.0675	-0.0405	-0.0296	-0.0193	-0.0184
	RMSE	0.0756	0.0581	0.0329	0.0195	0.0167
200	Bias	-0.0309	-0.0224	-0.0108	-0.0091	-0.0064
	RMSE	0.0475	0.0271	0.0169	0.0145	0.0139
500	Bias	0.0089	-0.0100	-0.0097	-0.0104	-0.0096
	RMSE	0.0288	0.0146	0.0127	0.0108	0.0106
1000	Bias	-0.0198	-0.0068	-0.0059	-0.0064	-0.0049
	RMSE	0.0162	0.0131	0.0108	0.0106	0.0100
N/T				500		
100	Bias	0.0400	-0.0232	-0.0147	-0.0121	-0.0106
	RMSE	0.0489	0.0312	0.0276	0.0260	0.0188
200	Bias	-0.0232	-0.0201	-0.0159	-0.0108	-0.0059
	RMSE	0.0266	0.0198	0.0120	0.0106	0.0110
500	Bias	-0.0124	0.0105	-0.0089	-0.0075	-0.0037
	RMSE	0.0146	0.0138	0.0103	0.0107	0.0111
1000	Bias	-0.0107	-0.0100	-0.0067	-0.0058	-0.0047
	RMSE	0.0109	0.0110	0.0112	0.0099	0.0101

Evidently, $\|\mathbf{K}_\rho\| \le C$. Moreover, by Assumption 2,

$$E(||\bar{\mathbf{v}}_N||^2) = E\left(\sum_{\ell=0}^{s}\sum_{k=1}^{m}\bar{v}_{N\ell k}^2\right) = E\left(\sum_{\ell=0}^{s}\sum_{k=1}^{m}\frac{1}{[N^{\alpha_0}]}\sum_{i,j=1}^{[N^{\alpha_0}]}v_{i\ell k}v_{j\ell k}\right) \le C \tag{B.7}$$

and by Assumption 3, we have

$$E||\mathbf{S}_\tau|| \le \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}E||\mathbf{F}_t\mathbf{F}_{t+\tau}^{'}|| = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}E(\sum_{j=0}^{s}\mathbf{f}_{t+\tau-j}^{'}\mathbf{f}_{t-j}) \le C.$$

So $||\bar{\mathbf{v}}_N|| = O_P(1)$ and $||\mathbf{S}_\tau|| = O_P(1)$. These derivations, together with (B.3), ensure

$$R_N = O_P([N^{\alpha_0-2}]). \tag{B.8}$$

We conclude from (B.8) and (B.6) that

$$\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N^{'}\mathbf{S}_\tau\bar{\mathbf{v}}_N} = O_P\Big(\frac{1}{[N^{\alpha_0}]\boldsymbol{\mu}_v^{'}\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\Big). \tag{B.9}$$

Table 4: DGP2: Bias and RMSE for the joint estimator $\widetilde{\alpha}$ with $\tau = \left[T^{1/4}\right]$.

	α_0	0.2	0.4	0.6	0.8	1
N/T				100		
100	Bias	-0.1001	-0.0959	-0.0692	-0.0701	-0.0439
	RMSE	0.0226	0.0208	0.0199	0.0200	0.0180
200	Bias	-0.0791	-0.0688	-0.0472	-0.0421	-0.0309
	RMSE	0.0182	0.0178	0.0162	0.0158	0.0149
500	Bias	-0.0671	-0.0491	-0.0412	-0.0286	-0.0287
	RMSE	0.0179	0.0179	0.0164	0.0152	0.0150
1000	Bias	-0.0629	-0.0328	-0.0231	-0.0206	-0.0183
	RMSE	0.0165	0.0162	0.0159	0.0143	0.0129
N/T				200		
100	Bias	-0.0937	-0.0927	-0.0652	-0.0539	-0.0490
	RMSE	0.0196	0.0199	0.0157	0.0159	0.0156
200	Bias	-0.0677	-0.0603	-0.0512	-0.0327	-0.0219
	RMSE	0.0172	0.0169	0.0150	0.0138	0.0128
500	Bias	-0.0419	-0.0395	-0.0263	-0.0199	-0.0201
	RMSE	0.0159	0.0142	0.0147	0.0137	0.0129
1000	Bias	-0.0317	-0.0285	-0.0279	-0.0201	-0.0187
	RMSE	0.0126	0.0118	0.0116	0.0109	0.0111
N/T				500		
100	Bias	-0.0550	-0.0452	-0.0373	-0.0296	-0.0210
	RMSE	0.0174	0.0167	0.0149	0.0152	0.0139
200	Bias	-0.0327	-0.0279	-0.0194	-0.0172	-0.0166
	RMSE	0.0138	0.0128	0.0131	0.0121	0.0119
500	Bias	-0.0199	-0.0179	-0.0166	-0.0142	-0.0117
	RMSE	0.0125	0.0124	0.0120	0.0119	0.0122
1000	Bias	-0.0197	-0.0173	-0.0137	-0.0118	-0.0113
	RMSE	0.0117	0.0119	0.0109	0.0103	0.0101

Therefore

$$\ln\Big(1+\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\Big)^2 = r_{NT}+o_P(r_{NT}) = O_P\Big(\frac{1}{[N^{\alpha_0}]\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\Big), \tag{B.10}$$

where $r_{NT} = \frac{2R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N} + \Big(\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\Big)^2$, and we have used the simple fact that

$$\lim_{x\to 0}\frac{\ln(1+x)-x}{x} = 0. \tag{B.11}$$

It follows that

$$\sqrt{\min\big([N^{\alpha_0}], T-\tau\big)}\ln\Big(1+\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\Big)^2 = O_P\Big(\frac{1}{[N^{\alpha_0/2}]\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\Big) = o_P(1). \tag{B.12}$$

Meanwhile, based on the decomposition of $C_N = \sum_{i=1}^3 c_{Ni}$, we evaluate the orders of the following terms: $\frac{c_{Ni}}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}$ for $i = 1,2,3$.

Table 5: Exponent of cross-country dependence of macro-variables

	N	T	$\widehat{\alpha}$	$\widetilde{\alpha}$	BKP15
Real GDP growth(RGDP)	33	122	0.899	0.912	0.754
Consumer Price Index(CPI)	33	123	0.913	0.922	0.851
Nominal equity price index(NOMEQ)	33	122	0.942	0.958	0.881
Short-term interest rates(Rshort)	33	122	0.981	0.947	0.907
Long-term interest rates(Rlong)	33	122	0.928	0.919	0.968

* BKP15 is the estimator defined in (11) of Bailey, Kapatanios and Pesaran (2015).

Table 6: Exponent of cross-sectional exponent of stock returns

(N,T)	(20,60)	(50,80)	(70,100)	(90,110)	(96,125)	(96,100)	(96,80)	(96,60)
$\widehat{\alpha}$	0.469	0.799	0.809	0.813	0.822	0.843	0.812	0.833
$\widetilde{\alpha}$	0.502	0.708	0.901	0.869	0.842	0.863	0.823	0.846
BKP15	1.002	0.639	0.793	0.842	0.882	0.901	0.898	0.859

*BKP15 is the estimator defined in (11) of Bailey, Kapatanios and Pesaran (2015).

For c_{N1}, we need to evaluate the orders of $\bar{u}^{(i)}$, $i = 1, 2$ and $\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{u}_t\bar{u}_{t+\tau}$. The order of $\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{u}_t\bar{u}_{t+\tau}$ will be provided in Lemma 2 in Appendix C.

By Assumption 1, we have

$$
\begin{aligned}
& E\Big(\sum_{i_1,i_2=1}^{N}\sum_{t_1,t_2=1}^{T-\tau} u_{i_1t_1}u_{i_2t_2}\Big) \\
= \; & E\sum_{i_1,i_2=1}^{N}\sum_{t_1,t_2=1}^{T-\tau}\Big(\sum_{j_1=0}^{+\infty}\phi_{i_1j_1}\sum_{s_1=-\infty}^{+\infty}\xi_{j_1s_1}\nu_{j_1,t_1-s_1}\Big)\Big(\sum_{j_2=0}^{+\infty}\phi_{i_2j_2}\sum_{s_2=-\infty}^{+\infty}\xi_{j_2s_2}\nu_{j_2,t_2-s_2}\Big) \\
= \; & E\sum_{i_1,i_2=1}^{N}\sum_{t_1,t_2=1}^{T-\tau}\sum_{j_1=0}^{+\infty}\phi_{i_1j_1}\sum_{s_1=-\infty}^{+\infty}\xi_{j_1s_1}\nu_{j_1,t_1-s_1}^2\phi_{i_2j_1}\xi_{j_1,t_1-s_1-t_2} \\
= \; & \sum_{i_1,i_2=1}^{N}\sum_{t_1,t_2=1}^{T-\tau}\sum_{j_1=0}^{+\infty}\sum_{s_1=-\infty}^{+\infty}\phi_{i_1j_1}\phi_{i_2j_1}\xi_{j_1s_1}\xi_{j_1,t_1-s_1-t_2} \\
\leq \; & \sum_{t_1=1}^{T-\tau}\sum_{i_1=1}^{N}\sum_{j_1=0}^{+\infty}|\phi_{i_1j_1}|\sum_{i_2=1}^{N}|\phi_{i_2j_1}|\sum_{s_1=-\infty}^{+\infty}|\xi_{j_1s_1}|\sum_{t_2=1}^{T}|\xi_{j_1,t_1-s_1-t_2}| = O\big(N(T-\tau)\big). \qquad \text{(B.13)}
\end{aligned}
$$

From (B.13) and the fact that $E(\bar{u}^{(1)}) = 0$, we have

$$Var(\bar{u}^{(1)}) = \frac{1}{N^2(T-\tau)^2} E\Big(\sum_{i_1,i_2=1}^{N}\sum_{t_1,t_2=1}^{T-\tau} u_{i_1t_1}u_{i_2t_2}\Big) = O\Big(\frac{1}{N(T-\tau)}\Big) \tag{B.14}$$

and then it follows that

$$\bar{u}^{(1)} = O_P(\frac{1}{\sqrt{N(T-\tau)}}). \tag{B.15}$$

Similarly, we have $\bar{u}^{(2)} = O_P\left(\frac{1}{\sqrt{N(T-\tau)}}\right)$. Combining (B.15) and Lemma 2 in Appendix D, we get

$$c_{N1} = O_P\Big(\max\Big(\frac{\gamma_1(\tau)}{N}, \frac{1}{N\sqrt{T-\tau}}\Big)\Big). \tag{B.16}$$

This, together with (B.6), (B.2) and (B.8), implies that

$$\frac{c_{N1}}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N} = O_P\Big(\max\Big(\frac{\gamma_1(\tau)N^{1-2\alpha_0}}{\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}, \frac{(T-\tau)^{-1/2}N^{1-2\alpha_0}}{\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\Big)\Big). \tag{B.17}$$

We then prove

$$\frac{c_{N2}}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N} = O_P\Big(\frac{(T-\tau)^{-1/2}N^{1/2-\alpha_0}}{\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\Big). \tag{B.18}$$

By Assumption 3, we have $E[c_{N2}] = 0$ and then its variance

$$\begin{aligned}
& Var\Big[\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\Big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_t - \bar{\mathbf{F}}_T)(\bar{u}_{t+\tau} - \bar{u}^{(2)})\Big)\Big] \\
= & \frac{1}{(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau} E\Big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_{t_1} - \bar{\mathbf{F}}_T)\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_{t_2} - \bar{\mathbf{F}}_T)\Big)E\Big((\bar{u}_{t_1+\tau} - \bar{u}^{(2)})(\bar{u}_{t_2+\tau} - \bar{u}^{(2)})\Big) \\
= & O\Big(\frac{[N^{2\alpha_0-2}]}{N(T-\tau)}\Big),
\end{aligned}$$

where the last equality uses (B.13) and the fact that via (3.10) in the main paper and (B.7):

$$E\Big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_t - \bar{\mathbf{F}}_T)\Big)^2 \le \Big[[N^{2\alpha_0-2}]E\big(||\bar{\mathbf{v}}_N||^2\big) + n^{-2}\|\mathbf{K}_\rho\|\Big]E\big(||\mathbf{F}_t - \bar{\mathbf{F}}_T||^2\big) = O([N^{2\alpha_0-2}]).$$

Hence

$$\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\Big(\bar{\boldsymbol{\beta}}_N'(\mathbf{F}_t - \bar{\mathbf{F}}_T)(\bar{u}_{t+\tau} - \bar{u})\Big) = O_P\Big(\frac{[N^{\alpha_0-1}]}{(T-\tau)^{1/2}N^{1/2}}\Big). \tag{B.19}$$

In view of this, (B.6), (B.2) and (B.8), we can obtain (B.18). Similarly, one may obtain

$$\frac{c_{N3}}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N} = O_P\Big(\frac{(T-\tau)^{-1/2}N^{1/2-\alpha_0}}{\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\Big). \tag{B.20}$$

By (B.17), (B.18), (B.20) and condition (3.23) in the main paper, we have

$$\sqrt{\min\big([N^{\alpha_0}], (T-\tau)\big)}\frac{c_{Ni}}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N} = o_P(1), \quad i = 1, 2, 3. \tag{B.21}$$

Applying (B.21) and (B.11), we obtain

$$\begin{aligned}&\sqrt{\min\left([N^{\alpha_0}],(T-\tau)\right)}\ln\Big(1+\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}\Big)^2\\ =\ &\sqrt{\min\left([N^{\alpha_0}],(T-\tau)\right)}\big(c_{NT}+o_P(c_{NT})\big)=o_P(1),\end{aligned}\tag{B.22}$$

where $c_{NT}=\frac{2C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}+\Big(\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}\Big)^2$.

By (B.5), we have

$$\kappa^2N^{4(\widehat{\alpha}-\alpha_0)}=(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N)^2\Big(1+\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\Big)^2\Big(1+\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}\Big)^2.\tag{B.23}$$

From (B.23), it follows that

$$\begin{aligned}&\frac{\kappa_0^2N^{4(\widehat{\alpha}-\alpha_0)}-(\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v)^2}{\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N+\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\\ =\ &\Big(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N-\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v\Big)\Big(1+\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\Big)^2\Big(1+\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}\Big)^2\\ &+\frac{(\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v)^2}{\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N+\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\Big[\Big(1+\frac{R_N}{[N^{2\alpha_0-2}]\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N}\Big)^2\Big(1+\frac{C_N}{\bar{\boldsymbol{\beta}}_N'\mathbf{S}_\tau\bar{\boldsymbol{\beta}}_N}\Big)^2-1\Big].\end{aligned}\tag{B.24}$$

With (B.24), (B.22), (B.10) and Lemma 1 in Appendix C, we obtain

$$\sqrt{\min([N^{\alpha_0}],T-\tau)}\frac{\kappa_0^2N^{4(\widehat{\alpha}-\alpha_0)}-(\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v)^2}{\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N+\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v}\to\mathcal{N}(0,\sigma_0^2),\tag{B.25}$$

where $\sigma_0^2=\lim_{N,T\to\infty}\frac{\min([N^{\alpha_0}],T-\tau)}{[N^{\alpha_0}]}4\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\Sigma}_v\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v+\lim_{N,T\to\infty}\frac{\min([N^{\alpha_0}],T-\tau)}{T-\tau}(\boldsymbol{\mu}_v'\otimes\boldsymbol{\mu}_v')\boldsymbol{\Omega}(\boldsymbol{\mu}_v\otimes\boldsymbol{\mu}_v)$.

□

9.2 Proof of Theorem 1

Proof. Recall that

$$\widetilde{\alpha}=\arg\max_\alpha\widehat{Q}_{NT}^{(1)}(\alpha),\quad where\ \ \widehat{Q}_{NT}^{(1)}(\alpha)=\frac{\big(\widehat{q}_1^{(1)}(\alpha)+[N^\alpha]\widehat{q}_2^{(1)}(\alpha)\big)^2}{N^{(1)}(\alpha)},$$

and

$$\widehat{q}_1^{(1)}(\alpha)=\sum_{n=1}^{[N^\alpha]}n^3\widehat{\sigma}_n(\tau),\quad\widehat{q}_2^{(1)}(\alpha)=\sum_{n=[N^\alpha]+1}^{N}n\widehat{\sigma}_n(\tau),\quad N^{(1)}(\alpha)=\sum_{n=1}^{[N^\alpha]}n^3+\sum_{n=[N^\alpha]+1}^{N}\frac{[N^{4\alpha}]}{n}$$

with

$$\widehat{\sigma}_n(\tau)=\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}(\bar{x}_{nt}-\bar{x}_n^{(1)})(\bar{x}_{n,t+\tau}-\bar{x}_n^{(2)}).\tag{B.26}$$

Similarly, it is easy to see that the true value α_0 satisfies $\alpha_0=\arg\max_\alpha Q_N^{(1)}(\alpha)$, where $Q_N^{(1)}(\alpha)=\frac{\left(q_1^{(1)}(\alpha)+[N^\alpha]q_2^{(1)}(\alpha)\right)^2}{N^{(1)}(\alpha)}$, and $q_1^{(1)}(\alpha)$ and $q_2^{(1)}(\alpha)$ are respectively obtained from $\widehat{q}_1^{(1)}(\alpha)$ and $\widehat{q}_2^{(1)}(\alpha)$ with $\widehat{\sigma}_n(\tau)$ replaced by

$$\sigma_n(\tau)=cov(\bar{x}_{nt},\bar{x}_{n,t+\tau}).\tag{B.27}$$

It follows that

$$\begin{aligned}\left|\widehat{Q}_{NT}^{(1)}(\alpha)-\widehat{Q}_{NT}^{(1)}(\alpha_0)\right| &= \left|\widehat{Q}_{NT}^{(1)}(\alpha)-Q_N^{(1)}(\alpha)-[\widehat{Q}_{NT}^{(1)}(\alpha_0)-Q_N^{(1)}(\alpha_0)]+Q_N^{(1)}(\alpha)-Q_N^{(1)}(\alpha_0)\right| \\ &\leq 2\max_{\alpha}\left|\widehat{Q}_{NT}^{(1)}(\alpha)-Q_N^{(1)}(\alpha)\right|+\left|Q_N^{(1)}(\alpha)-Q_N^{(1)}(\alpha_0)\right|. \end{aligned} \tag{B.28}$$

We next evaluate the two terms on the right hand of (B.28). Consider the first term on the right hand of (B.28). Rewrite it as

$$\begin{aligned} & N^{(1)}(\alpha)\Big(\widehat{Q}_{NT}^{(1)}(\alpha)-Q_N^{(1)}(\alpha)\Big) \\ = & \Big(\widehat{q}_1^{(1)}(\alpha)-q_1^{(1)}(\alpha)+[N^{\alpha}]\big(\widehat{q}_2^{(1)}(\alpha)-q_2^{(1)}(\alpha)\big)\Big)\cdot\Big(\widehat{q}_1^{(1)}(\alpha)+q_1^{(1)}(\alpha)+[N^{\alpha}]\big(\widehat{q}_2^{(1)}(\alpha)+q_2^{(1)}(\alpha)\big)\Big). \end{aligned}$$

A direct calculation, together with Lemma 1 in Appendix C, yields

$$\widehat{\sigma}_n(\tau)-\sigma_n(\tau)=\begin{cases} \bar{\mathbf{v}}_n'\mathbf{S}_\tau\bar{\mathbf{v}}_n-\kappa_0=O_P(v_{nT}^{-1/2}), & n\leq[N^{\alpha_0}]; \\ \frac{[N^{2\alpha_0}]}{n^2}(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N-\kappa_0)=O_P\big(\frac{[N^{2\alpha_0}]}{n^2}v_{NT}^{-1/2}\big), & n>[N^{\alpha_0}], \end{cases} \tag{B.29}$$

where $v_{nT}=\min(n,T-\tau)$ for $n\leq[N^{\alpha_0}]$ and $v_{NT}=\min([N^{\alpha_0}],T-\tau)$.

It follows that

$$\begin{aligned} \widehat{q}_1^{(1)}(\alpha)-q_1^{(1)}(\alpha) &= \sum_{n=1}^{[N^{\alpha}]} n^3\big(\widehat{\sigma}_n(\tau)-\sigma_n(\tau)\big) \\ &= \begin{cases} O_P\big(\sum_{n=1}^{[N^{\alpha}]} n^3 v_{nT}^{-1/2}\big), & \alpha\leq\alpha_0; \\ O_P\big(\sum_{n=1}^{[N^{\alpha_0}]} n^3 v_{nT}^{-1/2}+\sum_{n=[N^{\alpha_0}]+1}^{[N^{\alpha}]} n^3\frac{[N^{2\alpha_0}]}{n^2}v_{NT}^{-1/2}\big), & \alpha>\alpha_0, \end{cases} \\ &= \begin{cases} O_P\big([N^{4\alpha}](v_{NT}^{(1)})^{-1/2}\big), & \alpha\leq\alpha_0; \\ O_P\big([N^{4\alpha_0}]v_{NT}^{-1/2}+[N^{2\alpha_0}]\cdot|[N^{2\alpha}]-[N^{2\alpha_0}]|v_{NT}^{-1/2}\big), & \alpha>\alpha_0, \end{cases} \end{aligned}$$

where $v_{NT}^{(1)}=\min([N^{\alpha}],T-\tau)$. Similarly, we have

$$\begin{aligned} [N^{\alpha}]\big(\widehat{q}_2^{(1)}(\alpha)-q_2^{(1)}(\alpha)\big) &= [N^{\alpha}]\sum_{n=[N^{\alpha}]+1}^{N} n\big(\widehat{\sigma}_n(\tau)-\sigma_n(\tau)\big) \\ &= \begin{cases} O_P\Big([N^{2\alpha_0+\alpha}]v_{NT}^{-1/2}-[N^{3\alpha}](v_{NT}^{(1)})^{-1/2}+[N^{\alpha+2\alpha_0}](\log N^{1-\alpha_0})v_{NT}^{-1/2}\Big), & \alpha\leq\alpha_0; \\ O_P\Big([N^{2\alpha_0+\alpha}](\log N^{1-\alpha})v_{NT}^{-1/2}\Big), & \alpha>\alpha_0. \end{cases} \end{aligned}$$

It also follows from (3.15) in the main paper and (B.29) that

$$\begin{aligned} \widehat{q}_1^{(1)}(\alpha)+q_1^{(1)}(\alpha) &= \sum_{n=1}^{[N^{\alpha}]} n^3\big(\widehat{\sigma}_n(\tau)+\sigma_n(\tau)\big) \\ &= \begin{cases} O_P\Big(\sum_{n=1}^{[N^{\alpha}]} n^3\big(\kappa_0+v_{nT}^{-1/2}\big)\Big), & \alpha\leq\alpha_0; \\ O_P\Big(\sum_{n=1}^{[N^{\alpha_0}]} n^3\big(v_{NT}^{-1/2}+\kappa_0\big)+\sum_{n=[N^{\alpha_0}]+1}^{[N^{\alpha}]} n^3\frac{[N^{2\alpha_0}]}{n^2}v_{NT}^{-1/2}\Big), & \alpha>\alpha_0, \end{cases} \\ &= \begin{cases} O_P\Big([N^{4\alpha}]\kappa_0+[N^{4\alpha}](v_{NT}^{(1)})^{-1/2}\Big), & \alpha\leq\alpha_0; \\ O_P\Big([N^{4\alpha_0}](\kappa_0+v_{NT}^{-1/2})+([N^{2\alpha+2\alpha_0}]-[N^{4\alpha_0}])v_{NT}^{-1/2}\Big), & \alpha>\alpha_0 \end{cases} \end{aligned}$$

and that

$$[N^{\alpha}]\big(\widehat{q}_2^{(1)}(\alpha)+q_2^{(1)}(\alpha)\big)=[N^{\alpha}]\sum_{n=[N^{\alpha}]+1}^{N} n\big(\widehat{\sigma}_n(\tau)+\sigma_n(\tau)\big)$$

$$=\begin{cases} O_P\Big(-[N^{3\alpha}]\big((v_{NT}^{(1)})^{-1/2}+\kappa_0\big)+[N^{\alpha+2\alpha_0}]\big(1+\log N^{1-\alpha_0}\big)\big(v_{NT}^{-1/2}+\kappa_0\big)\Big), & \alpha\le\alpha_0;\\ O_P\Big([N^{2\alpha_0+\alpha}]\big(\log N^{1-\alpha}\big)\big(v_{NT}^{-1/2}+\kappa_0\big)\Big), & \alpha>\alpha_0. \end{cases}$$

Moreover,

$$N^{(1)}(\alpha)=\sum_{n=1}^{[N^{\alpha}]} n^3+\sum_{n=[N^{\alpha}]+1}^{N}\frac{[N^{4\alpha}]}{n}\asymp\Big([N^{4\alpha}]+[N^{4\alpha}]\log(\frac{N}{[N^{\alpha}]})\Big). \tag{B.30}$$

Summarizing the above derivations implies

$$\widehat{Q}_{NT}^{(1)}(\alpha)-Q_N^{(1)}(\alpha)=O_P\Big(\frac{[N^{4\alpha_0}]v_{NT}^{-1/2}\kappa_0}{\log N^{1-\alpha}}\Big). \tag{B.31}$$

Consider the second term on the right hand of (B.28). To this end, write

$$Q_N^{(1)}(\alpha)-Q_N^{(1)}(\alpha_0)=\frac{1}{N^{(1)}(\alpha)}((a_1+a_2)(a_3+a_4))+\frac{N^{(1)}(\alpha_0)-N^{(1)}(\alpha)}{N^{(1)}(\alpha)N^{(1)}(\alpha_0)}a_5^2, \tag{B.32}$$

where

$$a_1=q_1^{(1)}(\alpha)-q_1^{(1)}(\alpha_0),\quad a_2=[N^{\alpha}]\big(q_2^{(1)}(\alpha)-q_2^{(1)}(\alpha_0)\big),\quad a_3=q_1^{(1)}(\alpha)+q_1^{(1)}(\alpha_0),$$
$$a_4=[N^{\alpha}]\big(q_2^{(1)}(\alpha)+q_2^{(1)}(\alpha_0)\big)\text{ and } a_5=q_1^{(1)}(\alpha_0)+[N^{\alpha_0}]q_2^{(1)}(\alpha_0).$$

Straightforward calculations indicate that

$$a_1=O\big(\big|[N^{4\alpha_0}]-[N^{4\alpha}]\big|\kappa_0\big),\quad a_2=O\big([N^{\alpha}]\big|[N^{2\alpha_0}]-[N^{2\alpha}]\big|\kappa_0\big),\quad a_3=O\big(([N^{4\alpha_0}]+[N^{4\alpha}])\kappa_0\big),$$
$$a_4=O\big([N^{\alpha+2\alpha_0}](\log N^{1-\alpha})\kappa_0\big)\text{ and } a_5=O\Big(\big([N^{4\alpha_0}]+[N^{3\alpha_0}](\log N^{1-\alpha_0})\big)\kappa_0\Big).$$

It follows from (B.30) that

$$\begin{aligned}\Big|\frac{N^{(1)}(\alpha_0)-N^{(1)}(\alpha)}{N^{(1)}(\alpha_0)N^{(1)}(\alpha)}\Big| &\asymp \frac{\Big|([N^{4\alpha_0}]-[N^{4\alpha}])(1+\log\frac{N}{[N^{\alpha}]})-[N^{4\alpha_0}]\log[N^{\alpha_0-\alpha}]\Big|}{([N^{4\alpha_0}]\log[N^{1-\alpha_0}])([N^{4\alpha}]\log[N^{1-\alpha}])}\\ &\ge M\frac{\log N\Big|(1-\alpha)[N^{4\alpha}]-(1-\alpha_0)[N^{4\alpha_0}]\Big|}{([N^{4\alpha_0}]\log[N^{1-\alpha_0}])([N^{4\alpha}]\log[N^{1-\alpha}])},\end{aligned}$$

where and in what follows M stands for some positive constant number which may be different values from line to line, to save notation. From the above orders we conclude that the second term on the right hand of (B.32) is the leading term, compared with its first term. In view of this and the fact that α_0 is the maximizer of $Q_N^{(1)}(\alpha)$, we obtain from (B.32) that

$$\begin{aligned}Q_N^{(1)}(\alpha_0)-Q_N^{(1)}(\alpha) &\ge M\frac{\Big|N^{(1)}(\alpha_0)-N^{(1)}(\alpha)\Big|}{N^{(1)}(\alpha)N^{(1)}(\alpha_0)}a_5^2\\ &\ge M\frac{\log N\Big|(1-\alpha)[N^{4\alpha}]-(1-\alpha_0)[N^{4\alpha_0}]\Big|\Big([N^{4\alpha_0}]\kappa_0\Big)^2}{[N^{4\alpha_0}](\log[N^{1-\alpha_0}])[N^{4\alpha}]\log[N^{1-\alpha}]}.\end{aligned} \tag{B.33}$$

Note that (B.29) holds uniformly in α so that (B.31) is true when α is replaced with $\widetilde{\alpha}$. Also (B.33) holds when α is replaced with $\widetilde{\alpha}$. We conclude from (B.28) and the fact that $\widetilde{\alpha}$ is the maximizer of $\widehat{Q}^{(1)}_{NT}(\alpha)$ that

$$\left|Q^{(1)}_N(\alpha_0) - Q^{(1)}_N(\widetilde{\alpha})\right| \le 2 \max_{\alpha=\widetilde{\alpha},\alpha_0} \left|\widehat{Q}^{(1)}_{NT}(\alpha) - Q^{(1)}_N(\alpha)\right|,$$

which, together with (B.31) and (B.33), yields

$$\left|(1-\alpha_0) - (1-\widetilde{\alpha})[N^{4(\widetilde{\alpha}-\alpha_0)}]\right| = O_P\left(v_{NT}^{-1/2}\kappa_0^{-1}\right). \tag{B.34}$$

We next consider the consistency of $\widetilde{\kappa}$. It is easy to see that

$$\widetilde{\kappa} = \frac{\widehat{q}^{(1)}_1(\widetilde{\alpha}) + [N^{2\widetilde{\alpha}}]\widehat{q}^{(1)}_2(\widetilde{\alpha})}{N^{(1)}(\widetilde{\alpha})} \text{ and } \kappa_0 = \frac{q^{(1)}_1(\alpha_0) + [N^{2\alpha_0}]q^{(1)}_2(\alpha_0)}{N^{(1)}(\alpha_0)}. \tag{B.35}$$

It follows that

$$\widetilde{\kappa} - \kappa_0 = \frac{1}{N^{(1)}(\widetilde{\alpha})}(b_1 + b_2) + \frac{N^{(1)}(\alpha_0) - N^{(1)}(\widetilde{\alpha})}{N^{(1)}(\widetilde{\alpha})N^{(1)}(\alpha_0)} b_3, \tag{B.36}$$

where $b_1 = \widehat{q}^{(1)}_1(\widetilde{\alpha}) - q^{(1)}_1(\alpha_0)$, $b_2 = [N^{2\widetilde{\alpha}}]\widehat{q}^{(1)}_2(\widetilde{\alpha}) - [N^{2\alpha_0}]q^{(1)}_2(\alpha_0)$ and $b_3 = q^{(1)}_1(\alpha_0) + [N^{2\alpha_0}]q^{(1)}_2(\alpha_0)$.

The orders of b_i, $i = 1, 2, 3$ are listed below.

$$b_1 = O_P\Big(|[N^{4\alpha_0}] - [N^{4\widetilde{\alpha}}]|\kappa_0 + [N^{4\widetilde{\alpha}}]v_{NT}^{-1/2}\Big), \quad b_3 = O_P\Big([N^{4\alpha_0}]\kappa_0 + [N^{4\alpha_0}]\cdot(\log N^{1-\alpha_0})\kappa_0\Big),$$
$$b_2 = O_P\Big(|[N^{2\widetilde{\alpha}}] - [N^{2\alpha_0}]|\cdot\big(v_{NT}^{-1/2} + \kappa_0\big)\cdot\big([N^{2\widetilde{\alpha}}] + [N^{2\alpha_0}]\log N^{1-\alpha_0}\big) + [N^{4\alpha_0}](\log N^{1-\alpha_0})v_{NT}^{-1/2}\Big).$$

We then conclude from these orders, (B.36) and (B.30) that

$$\widetilde{\kappa} - \kappa_0 = O_P(v_{NT}^{-1/2}). \tag{B.37}$$

The convergence rate of $(\widetilde{\alpha}, \widetilde{\kappa})$ in Theorem 3 immediately follows. The next aim is to derive an asymptotic distribution for the joint estimator $(\widetilde{\alpha}, \widetilde{\kappa})$. In view of (B.34) and (B.37), it is enough to consider those α and κ within a compact interval $D(C)$:

$$D(C) = \left\{(\alpha,\kappa): \quad \alpha = \alpha_0 + \frac{1}{2}\frac{\ln(1 + s_1\kappa_0^{-1}v_{NT}^{-1/2})}{\ln N}, \quad \kappa = \kappa_0 + s_2 v_{NT}^{-1/2}\right\}, \tag{B.38}$$

where $|s_j| \le C, j = 1, 2$ with C being some positive constant independent of n. Recall that

$$Q^{(1)}_{NT}(\alpha,\kappa) = \sum_{n=1}^{[N^\alpha]} n^3\Big(\widehat{\sigma}_n(\tau) - \kappa\Big)^2 + \sum_{n=[N^\alpha]+1}^{N} n^3\Big(\widehat{\sigma}_n(\tau) - \frac{[N^{2\alpha}]}{n^2}\kappa\Big)^2$$

and $(\widetilde{\alpha}, \widetilde{\kappa}) = \arg\min_{\alpha,\kappa} Q^{(1)}_{NT}(\alpha,\kappa)$.

Without loss of generality, we assume that $\alpha \le \alpha_0$ below. First, we simplify $\Big(Q^{(1)}_{NT}(\alpha,\kappa) - Q^{(1)}_{NT}(\alpha_0,\kappa_0)\Big)$. To this end, write

$$\begin{aligned}
Q^{(1)}_{NT}(\alpha,\kappa) - Q^{(1)}_{NT}(\alpha_0,\kappa_0) &= \sum_{n=1}^{[N^\alpha]} n^3\Big(\big(\widehat{\sigma}_n(\tau) - \kappa\big)^2 - \big(\widehat{\sigma}_n(\tau) - \kappa_0\big)^2\Big) \\
&+ \sum_{n=[N^\alpha]+1}^{N} n^3\Big(\big(\widehat{\sigma}_n(\tau) - \frac{[N^{2\alpha}]}{n^2}\kappa\big)^2 - \big(\widehat{\sigma}_n(\tau) - \frac{[N^{2\alpha}]}{n^2}\kappa_0\big)^2\Big) + \Big(\sum_{n=1}^{[N^\alpha]} - \sum_{n=1}^{[N^{\alpha_0}]}\Big)n^3\big(\widehat{\sigma}_n(\tau) - \kappa_0\big)^2 \\
&+ \sum_{n=[N^\alpha]+1}^{N} n^3\big(\widehat{\sigma}_n(\tau) - \frac{[N^{2\alpha}]}{n^2}\kappa_0\big)^2 - \sum_{n=[N^{\alpha_0}]+1}^{N} n^3\big(\widehat{\sigma}_n(\tau) - \frac{[N^{2\alpha_0}]}{n^2}\kappa_0\big)^2 = \sum_{j=1}^{8} A_j,
\end{aligned}$$

where

$$A_1 = \sum_{n=1}^{[N^\alpha]} 2n^3 \widehat{\sigma}_n(\tau)(\kappa_0 - \kappa); \quad A_2 = \sum_{n=1}^{[N^\alpha]} n^3(\kappa^2 - \kappa_0^2); \; A_3 = \sum_{n=[N^\alpha]+1}^{N} 2[N^{2\alpha}] n \widehat{\sigma}_n(\tau)(\kappa_0 - \kappa);$$

$$A_4 = \sum_{n=[N^\alpha]+1}^{N} \frac{[N^{4\alpha}]}{n}(\kappa^2 - \kappa_0^2); A_5 = \sum_{n=[N^\alpha]+1}^{[N^{\alpha_0}]} -n^3\big(\widehat{\sigma}_n(\tau) - \kappa_0\big)^2; \; A_6 = \sum_{n=[N^\alpha]+1}^{[N^{\alpha_0}]} n^3\big(\widehat{\sigma}_n(\tau) - \frac{[N^{2\alpha}]}{n^2}\kappa_0\big)^2;$$

$$A_7 = \sum_{n=[N^{\alpha_0}]+1}^{N} \frac{[N^{4\alpha}] - [N^{4\alpha_0}]}{n}\kappa_0^2; \quad A_8 = \sum_{n=[N^{\alpha_0}]+1}^{N} 2n\kappa_0 \widehat{\sigma}_n(\tau)([N^{2\alpha_0}] - [N^{2\alpha}]).$$

The orders of A_j, $j = 1, \cdots, 8$, are evaluated below. It follows from (3.15) in the main paper and (B.29) that

$$\widehat{\sigma}_n(\tau) = \begin{cases} \bar{\mathbf{v}}_n' \mathbf{S}_\tau \bar{\mathbf{v}}_n + O(\frac{1}{n}), & n < [N^{\alpha_0}]; \\ \frac{[N^{2\alpha_0}]}{n^2} \bar{\mathbf{v}}_N' \mathbf{S}_\tau \bar{\mathbf{v}}_N + O\Big(\frac{[N^{\alpha_0}]}{n^2}\Big), & n \geq [N^{\alpha_0}]. \end{cases} \tag{B.39}$$

This, together with the fact that $\alpha, \kappa \in D(C)$, implies

$$\begin{aligned}
&A_1 = O_P([N^{4\alpha}]\kappa_0|\kappa_0 - \kappa|), \quad A_2 = O([N^{4\alpha}]\kappa_0|\kappa - \kappa_0|), \\
&A_3 = O_P([N^{2\alpha+2\alpha_0}](\log N^{1-\alpha})\kappa_0|\kappa - \kappa_0|), \quad A_4 = O([N^{4\alpha}](\log N)\kappa_0|\kappa - \kappa_0|), \\
&A_5 = O_P\big(|[N^{4\alpha}] - [N^{4\alpha_0}]| v_{NT}^{-1}\big), \quad A_6 = O_P\big(|[N^{2\alpha_0}]\kappa_0 - [N^{2\alpha}]\kappa|^2 \log N^{\alpha_0 - \alpha}\big), \\
&A_7 = O\big(|[N^{4\alpha}] - [N^{4\alpha_0}]|(\log N)\kappa_0^2\big), \quad A_8 = O_P\big(|[N^{2\alpha_0}] - [N^{2\alpha}]| \cdot [N^{2\alpha_0}](\log N^{1-\alpha_0})\kappa_0^2\big),
\end{aligned}$$

with $v_{NT} = \min([N^{\alpha_0}], T - \tau)$.

From the above orders and (B.38), we see that A_3, A_4, A_7 and A_8 are the leading terms. We then conclude

$$\begin{aligned}
& Q_{NT}^{(1)}(\alpha, \kappa) - Q_{NT}^{(1)}(\alpha_0, \kappa_0) = (A_3 + A_8) + (A_4 + A_7) + O_P(\delta_{NT}^{(1)}) \\
= \; & \sum_{n=[N^{\alpha_0}]+1}^{N} 2n\widehat{\sigma}_N(\tau)\big([N^{2\alpha_0}]\kappa_0 - [N^{2\alpha}]\kappa\big) + \sum_{n=[N^{\alpha_0}]+1}^{N} \frac{[N^{4\alpha}]\kappa^2 - [N^{4\alpha_0}]\kappa_0^2}{n} + O_P(\delta_{NT}^{(1)}) \\
= \; & \Big([N^{2\alpha_0}]\kappa_0 - [N^{2\alpha}]\kappa\Big)\Big(\sum_{n=[N^{\alpha_0}]+1}^{N} \big(2n\widehat{\sigma}_n(\tau) - \frac{2[N^{2\alpha_0}]\kappa_0}{n}\big) \\
+ \; & \sum_{n=[N^{\alpha_0}]+1}^{N} \frac{[N^{2\alpha_0}]\kappa_0 - [N^{2\alpha}]\kappa}{n}\Big) + O_P(\delta_{NT}^{(1)}),
\end{aligned} \tag{B.40}$$

where $\delta_{NT}^{(1)} = o_P(A_3 + A_8 + A_4 + A_7)$, uniformly on the compact interval $D(C)$.

Moreover, it follows from the second equality in (B.39) that

$$\sum_{n=[N^{\alpha_0}]+1}^{N} 2n\widehat{\sigma}_n(\tau) = \bar{\mathbf{v}}_N' \mathbf{S}_\tau \bar{\mathbf{v}}_N \Big(\sum_{n=[N^{\alpha_0}]+1}^{N} \frac{2[N^{2\alpha_0}]}{n} \Big) + \sum_{n=[N^{\alpha_0}]+1}^{N} 2nO\Big(\frac{[N^{\alpha_0}]}{n^2}\Big). \tag{B.41}$$

Let

$$g_N(s_1, s_2) = v_{NT} \frac{Q_{NT}^{(1)}(\alpha, \kappa) - Q_{NT}^{(1)}(\alpha_0, \kappa_0)}{[N^{2\alpha_0}] \sum_{n=[N^{\alpha_0}]+1}^{N} \frac{2[N^{2\alpha_0}]}{n}}, \tag{B.42}$$

where s_1 and s_2 are defined in (B.38).

By (B.40) and (B.41) we have

$$g_N(s_1, s_2) = r_{NT} v_{NT}^{1/2} \Big(\bar{\mathbf{v}}_N' \mathbf{S}_\tau \bar{\mathbf{v}}_N - \kappa_0 \Big) + \frac{1}{2} r_{NT}^2 + O_P(v_{NT} d_{NT}^{(1)}), \tag{B.43}$$

where $r_{NT} = v_{NT}^{1/2} \frac{[N^{2\alpha}]\kappa - [N^{2\alpha_0}]\kappa_0}{[N^{2\alpha_0}]}$ and $d_{NT}^{(1)} = \frac{\delta_{NT}^{(1)} + \sum_{n=[N^{\alpha_0}]+1}^{N} 2n(\frac{[N^{\alpha_0}]}{n^2})}{[N^{2\alpha_0}] \sum_{n=[N^{\alpha_0}]+1}^{N} \frac{2[N^{2\alpha_0}]}{n}}$.

With notation $\xi = \ln(1 + s_1 \kappa_0^{-1} v_{NT}^{-1/2}) / \ln N$, we obtain

$$\ln N^{\xi} = \ln(1 + s_1 \kappa_0^{-1} v_{NT}^{-1/2}),$$

which implies $N^{\xi} = 1 + s_1 \kappa_0^{-1} v_{NT}^{-1/2}$. This, together with (B.38), ensures

$$[N^{2\alpha - 2\alpha_0}] = N^{\xi} = 1 + s_1 \kappa_0^{-1} v_{NT}^{-1/2}, \quad \kappa - \kappa_0 = s_2 v_{NT}^{-1/2}. \tag{B.44}$$

By (B.44) and the definition of r_{NT}, we have

$$\begin{aligned} r_{NT} &= v_{NT}^{1/2} \frac{[N^{2\alpha}]\kappa - [N^{2\alpha}]\kappa_0}{[N^{2\alpha_0}]} + v_{NT}^{1/2} \frac{[N^{2\alpha}]\kappa_0 - [N^{2\alpha_0}]\kappa_0}{[N^{2\alpha_0}]} \\ &= v_{NT}^{1/2} [N^{2\alpha - 2\alpha_0}](\kappa - \kappa_0) + v_{NT}^{1/2} \kappa_0 ([N^{2\alpha - 2\alpha_0}] - 1) \\ &= s_1 + s_2 + s_1 s_2 \kappa_0^{-1} v_{NT}^{-1/2}. \end{aligned} \tag{B.45}$$

We then conclude from (B.45), (B.43) and Lemma 1 that for any $s_1, s_2 \in [-C, C]$,

$$g_N(s_1, s_2) \xrightarrow{d} g(s_1, s_2) = (s_1 + s_2) Z + \frac{1}{2} (s_1 + s_2)^2,$$

where Z is a normal random variable with mean 0 and variance σ_0^2, which is the asymptotic distribution derived in Lemma 1.

Here we would like to point out that the last term of (B.43) converges to zero in probability uniformly in $s_1, s_2 \in [-C, C]$, in view of (B.40) and the tightness in s_1 and s_2 is straightforward due to the structure of r_{NT} in (B.45).

Let $\widetilde{s}_1$ and $\widetilde{s}_2$ be s_1 and s_2 respectively with (α, κ) replaced by $(\widetilde{\alpha}, \widetilde{\kappa})$. By the definition of $(\widetilde{\alpha}, \widetilde{\kappa})$ in (3.9) of the main paper, we know that $g_N(s_1, s_2)$ takes the minimum value at $(\widetilde{s}_1, \widetilde{s}_2)$. Moreover, from (B.43) and (B.45) a key observation is that

$$\widetilde{s}_1 + \widetilde{s}_2 = -v_{NT}^{1/2} \left(\bar{\mathbf{v}}_N' \mathbf{S}_\tau \bar{\mathbf{v}}_N - \kappa_0 \right) + o_p(1) \tag{B.46}$$

(one can verify this by taking derivative with respective to s_1 and s_2 in (B.43)). Next, we analyze $\widetilde{s}_2$. Recall that $\widetilde{s}_2 = v_{NT}^{1/2} (\widetilde{\kappa} - \kappa_0)$. By the definition of $\widetilde{\kappa}$ in (3.7) of the main paper, we first provide the leading term of $\widetilde{\kappa}$. It is easy to see that the leading terms of the numerator and the denominator of $\widetilde{\kappa}$ are $[N^{2\widetilde{\alpha}}] \widehat{q}_2^{(1)}(\widetilde{\alpha})$ and $\sum_{n=[N^{\widetilde{\alpha}}]+1}^{N} \frac{[N^{4\widetilde{\alpha}}]}{n}$ respectively.

Moreover, we have the following evaluations:

$$\begin{aligned}
& [N^{2\widetilde{\alpha}}]\widehat{q}_2^{(1)}(\widetilde{\alpha}) - [N^{2\alpha_0}]\widehat{q}_2^{(1)}(\alpha_0) \\
= & \left([N^{2\widetilde{\alpha}}] - [N^{2\alpha_0}]\right)\widehat{q}_2^{(1)}(\widetilde{\alpha}) + [N^{2\alpha_0}]\left(\widehat{q}_2^{(1)}(\widetilde{\alpha}) - \widehat{q}_2^{(1)}(\alpha_0)\right) \\
= & \left([N^{2\widetilde{\alpha}}] - [N^{2\alpha_0}]\right)\cdot \sum_{n=[N^{\alpha_0}]+1}^{N} n\widehat{\sigma}_n(\tau) \\
& + \left([N^{2\widetilde{\alpha}}] - [N^{2\alpha_0}]\right)\cdot \sum_{n=[N^{\alpha_0}]+1}^{N} n\widehat{\sigma}_n(\tau) + [N^{2\alpha_0}] \sum_{n=[N^{\widetilde{\alpha}}]+1}^{[N^{\alpha_0}]} n\widehat{\sigma}_n(\tau) \\
= & \left([N^{2\widetilde{\alpha}}] - [N^{2\alpha_0}]\right)\cdot \left(\sum_{n=[N^{\alpha_0}]+1}^{N} n\widehat{\sigma}_n(\tau)\right)\cdot(1+o_P(1)) \qquad \text{(B.47)}
\end{aligned}$$

and

$$\begin{aligned}
& [N^{4\widetilde{\alpha}}]\cdot \sum_{n=[N^{\widetilde{\alpha}}]+1}^{N} \frac{1}{n} - [N^{4\alpha_0}]\cdot \sum_{n=[N^{\alpha_0}]+1}^{N} \frac{1}{n} \\
= & [N^{4\widetilde{\alpha}}]\cdot \sum_{n=[N^{\widetilde{\alpha}}]+1}^{[N^{\alpha_0}]} \frac{1}{n} + \left([N^{4\widetilde{\alpha}}] - [N^{4\alpha_0}]\right)\cdot \sum_{n=[N^{\alpha_0}]+1}^{N} \frac{1}{n} \\
= & \left([N^{4\widetilde{\alpha}}] - [N^{4\alpha_0}]\right)\cdot\left(\ln \frac{N}{[N^{\alpha_0}]}\right)(1+o(1)). \qquad \text{(B.48)}
\end{aligned}$$

It follows from (B.47), (B.48) and (B.35) that

$$\begin{aligned}
\widetilde{\kappa} - \kappa_0 &= \frac{\sum_{n=[N^{\alpha_0}]+1}^{N} n\cdot [N^{2\alpha_0}]\cdot(\hat{\sigma}_n(\tau) - \sigma_n(\tau))}{\sum_{n=[N^{\alpha_0}]+1}^{N} \frac{[N^{4\alpha_0}]}{n}}(1+o_P(1)) \\
&= \frac{\left(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N - \kappa_0\right)\sum_{n=[N^{\alpha_0}]+1}^{N} \frac{[N^{4\alpha_0}]}{n} + O\left(\sum_{n=[N^{\alpha_0}]+1}^{N} \frac{[N^{3\alpha_0}]}{n}\right)}{\sum_{n=[N^{\alpha_0}]+1}^{N} \frac{[N^{4\alpha_0}]}{n}}(1+o_P(1)) \\
&= \left(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N - \kappa_0\right)(1+o_P(1)), \qquad \text{(B.49)}
\end{aligned}$$

where the second equality uses (B.39).

We then conclude from (B.49) and Lemma 1 that

$$\widetilde{s}_2 = v_{NT}^{1/2}(\widetilde{\kappa} - \kappa_0) = v_{NT}^{1/2}\left(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N - \kappa_0\right)(1+o_P(1)) \xrightarrow{d} \mathcal{N}\left(0, \sigma_0^2\right). \qquad \text{(B.50)}$$

For $\widetilde{s}_1$, we can get its expression by differencing (B.46) and (B.50) i.e.

$$\widetilde{s}_1 = -2v_{NT}^{1/2}\left(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N - \kappa_0\right)(1+o_P(1)) \xrightarrow{d} \mathcal{N}\left(0, 4\sigma_0^2\right). \qquad \text{(B.51)}$$

Obviously, from (B.50), (B.51) and the fact that

$$\widetilde{s}_1 = -2\widetilde{s}_2(1+o_P(1)), \qquad \text{(B.52)}$$

one can conclude the joint asymptotic distribution in (3.27) of the main paper.

□

10 Appendix C: Some Lemmas

In this appendix, we provide the necessary lemmas used in the proofs of the main theorems above. Lemmas 1 and 2 are used in the proof of Theorem 1 and 2. Lemmas 3 and 4 are needed in the proof of Lemma 1.

10.1 Lemmas 1 and 2 for Theorem 2

Lemma 1. *In addition to Assumptions 1 and 3, we assume that τ is fixed or τ tends to infinity satisfying*

$$\frac{\tau}{(T-\tau)^{\delta/(2\delta+2)}} \to 0, \quad as \quad T \to \infty, \tag{C.1}$$

for some constant $\delta > 0$. Moreover, under (3.21), we assume that

$$E|\zeta_{it}|^{2+2\delta} < +\infty, \tag{C.2}$$

where ζ_{it} is the i-th component of $\boldsymbol{\zeta}_t$ and $\{\boldsymbol{\zeta}_t : \ldots, -1, 0, 1, \ldots\}$ is the sequence appeared in Assumption 3. And the covariance matrix Γ of the random vector

$$\left(C_{ij}(h^{'}) : i = 1, \ldots, m; j = 1, \ldots, m; h^{'} = \tau - s, \ldots, \tau + s\right) \tag{C.3}$$

is positive definite, where $C_{ij}(h)$ is defined in (3.19) just above Theorem 1 in the main paper.

Then as $N, T \to \infty$, we have

$$\sqrt{\min\left([N^{\alpha_0}], T-\tau\right)}(\bar{\mathbf{v}}_N^{'}\mathbf{S}_\tau\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v^{'}\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v) \xrightarrow{d} \mathcal{N}\left(0, \sigma_0^2\right),$$

where $\sigma_0^2 = \lim_{N,T\to\infty} \frac{\min([N^{\alpha_0}],T-\tau)}{[N^{\alpha_0}]} 4\boldsymbol{\mu}_v^{'}\boldsymbol{\Sigma}_\tau\boldsymbol{\Sigma}_v\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v + \lim_{N,T\to\infty} \frac{\min([N^{\alpha_0}],T-\tau)}{T-\tau}(\boldsymbol{\mu}_v^{'} \otimes \boldsymbol{\mu}_v^{'})\boldsymbol{\Omega}(\boldsymbol{\mu}_v \otimes \boldsymbol{\mu}_v)$, $\boldsymbol{\Sigma}_\tau = \mathbb{E}(\mathbf{F}_t\mathbf{F}_{t+\tau}^{'})$, $\boldsymbol{\mu}_v = \mu_v\mathbf{e}_{m(s+1)}$, *where* $\mathbf{e}_{m(s+1)}$ *is an* $m(s+1) \times 1$ *vector with each element being* 1, *and*

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\omega}(\tau,\tau) & \cdots & \boldsymbol{\omega}(\tau,\tau+s) & \cdots & \boldsymbol{\omega}(\tau,\tau-s) & \cdots & \boldsymbol{\omega}(\tau,\tau) \\ \boldsymbol{\omega}(\tau+1,\tau) & \cdots & \boldsymbol{\omega}(\tau+1,\tau+s) & \cdots & \boldsymbol{\omega}(\tau+1,\tau-s) & \cdots & \boldsymbol{\omega}(\tau+1,\tau) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\omega}(\tau-s,\tau) & \cdots & \boldsymbol{\omega}(\tau-s,\tau+s) & \cdots & \boldsymbol{\omega}(\tau-s,\tau-s) & \cdots & \boldsymbol{\omega}(\tau-s,\tau) \\ \boldsymbol{\omega}(\tau,\tau) & \cdots & \boldsymbol{\omega}(\tau,\tau+s) & \cdots & \boldsymbol{\omega}(\tau,\tau-s) & \cdots & \boldsymbol{\omega}(\tau,\tau) \end{pmatrix}, \tag{C.4}$$

with $\boldsymbol{\omega}(h,r) = \left(Cov(f_{i_1t}f_{j_1,t+h}, f_{i_2,t}f_{j_2,t+r}) : 1 \le i_1, j_1, i_2, j_2 \le m\right)_{m^2\times m^2}$.

Proof. Write

$$\begin{aligned} \bar{\mathbf{v}}_N^{'}\mathbf{S}_\tau\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v^{'}\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v &= (\bar{\mathbf{v}}_N^{'} - \boldsymbol{\mu}_v^{'})\mathbf{S}_\tau\bar{\mathbf{v}}_N + \boldsymbol{\mu}_v^{'}(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau)\bar{\mathbf{v}}_N + \boldsymbol{\mu}_v^{'}\boldsymbol{\Sigma}_\tau(\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v) \\ &= (\bar{\mathbf{v}}_N^{'} - \boldsymbol{\mu}_v^{'})(\mathbf{S}_\tau\bar{\mathbf{v}}_N + \boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v) + \boldsymbol{\mu}_v^{'}(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau)\bar{\mathbf{v}}_N. \end{aligned} \tag{C.5}$$

Since the elements of the vector $\bar{\mathbf{v}}_N$ are all i.i.d., we have

$$\sqrt{[N^{\alpha_0}]}(\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v) \xrightarrow{d} N(0, \boldsymbol{\Sigma}_v), \quad as \quad N \to \infty, \tag{C.6}$$

where $\mathbf{\Sigma}_v$ is an $m(s+1)$-dimensional diagonal matrix with each of the diagonal elements being σ_v^2.

Moreover, under Assumption 3, we have

$$\mathbf{S}_\tau - \mathbf{\Sigma}_\tau \xrightarrow{i.p.} 0, \quad as \;\; T \to \infty, \tag{C.7}$$

(one may see (C.10) below). It follows from (C.6) and (C.7) that, if τ is fixed,

$$\begin{aligned}
& \sqrt{[N^{\alpha_0}]}(\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')(\mathbf{S}_\tau \bar{v}_N + \mathbf{\Sigma}_\tau \boldsymbol{\mu}_v) \\
= \;& \sqrt{[N^{\alpha_0}]}\Big((\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')\mathbf{S}_\tau(\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v) + (\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')(\mathbf{S}_\tau - \mathbf{\Sigma}_\tau)\boldsymbol{\mu}_v + 2(\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')\mathbf{\Sigma}_\tau \boldsymbol{\mu}_v\Big) \\
= \;& 2\sqrt{[N^{\alpha_0}]}(\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')\mathbf{\Sigma}_\tau \boldsymbol{\mu}_v + o_p(1) \xrightarrow{d} \mathcal{N}(0, 4\boldsymbol{\mu}_v'\mathbf{\Sigma}_\tau'\mathbf{\Sigma}_v\mathbf{\Sigma}_\tau\boldsymbol{\mu}_v).
\end{aligned} \tag{C.8}$$

When τ satisfies (3.21), we have $\lim_{\tau\to\infty} \mathbf{\Sigma}_\tau = 0$. In fact, we consider one element $\gamma(h) = Cov(f_{k,t}, f_{k,t+h})$ of $\mathbf{\Sigma}_\tau$:

$$\gamma(h) = E\Big(\sum_{j_1=0}^{+\infty} b_{j_1}\zeta_{k,t-j_1}\sum_{j_2=0}^{+\infty} b_{j_2}\zeta_{t+h-j_2}\Big) = \sum_{j_1=0}^{+\infty} b_{j_1}b_{h+j_1}.$$

Then

$$\sum_{h=0}^{+\infty}|\gamma(h)| = \sum_{h=0}^{+\infty}|\sum_{j=0}^{+\infty} b_j b_{h+j}| \le \Big(\sum_{j=0}^{+\infty}|b_j|\Big)^2 < +\infty.$$

From this, we can see that $\gamma(h) \to 0$ as $h \to \infty$. So as $\tau \to \infty$, $\mathbf{\Sigma}_\tau \to 0$. Hence, under this case,

$$\sqrt{[N^{\alpha_0}]}(\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')(\mathbf{S}_\tau \bar{v}_N + \mathbf{\Sigma}_\tau \boldsymbol{\mu}_v) \xrightarrow{i.p.} 0. \tag{C.9}$$

Under Assumption 3, by Theorem 14 in Chapter 4 of Hannan (1970), when τ is fixed, the sample covariance of the stationary time series $\{\mathbf{f}_t : t = 1, 2, \ldots, T\}$ has the following asymptotic property:

$$\sqrt{T}\Big(vec\big(\widehat{\boldsymbol{\gamma}}(h) - \boldsymbol{\gamma}(h)\big), 0 \le h \le \ell\Big) \xrightarrow{d} N(0, \boldsymbol{\omega}), \tag{C.10}$$

where 'vec' means that for a matrix $\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_n) : q \times n$, $vec(\mathbf{X})$ is the $qn \times 1$ vector defined as

$$vec(\mathbf{X}) = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}, \tag{C.11}$$

$\widehat{\boldsymbol{\gamma}}(h) = \frac{1}{T-h}\sum_{t=1}^{T-h}(\mathbf{f}_t - \bar{\mathbf{f}}^{(1)})(\mathbf{f}_{t+h} - \bar{\mathbf{f}}^{(2)})'$, $\bar{\mathbf{f}}^{(1)} = \frac{1}{T-h}\sum_{t=1}^{T-h}\mathbf{f}_t$, $\bar{\mathbf{f}}^{(2)} = \frac{1}{T-h}\sum_{t=1}^{T-h}\mathbf{f}_{t+h}$, and $\boldsymbol{\gamma}(h) = Cov(\mathbf{f}_t, \mathbf{f}_{t+h})$. Note that the expression of $vec\big(\widehat{\boldsymbol{\gamma}}(h)\big)$ is

$$vec\Big(\widehat{\boldsymbol{\gamma}}(h)\Big) = \Big(\widetilde{cov}(1,1), \widetilde{cov}(2,1), \ldots, \widetilde{cov}(m,1), \ldots, \widetilde{cov}(1,m), \ldots, \widetilde{cov}(m,m)\Big)',$$

with $\widetilde{cov}(i,j) = \frac{1}{T-h}\sum_{t=1}^{T-h} f_{it}f_{j,t+h} - \frac{1}{T-h}\sum_{t=1}^{T-h} f_{it}\frac{1}{T-h}\sum_{t=1}^{T-h} f_{j,t+h}$. The asymptotic covariance between $\sqrt{T}\Big(vec\big(\widehat{\boldsymbol{\gamma}}(h) - \boldsymbol{\gamma}(h)\big)\Big)$ and $\sqrt{T}\Big(vec\big(\widehat{\boldsymbol{\gamma}}(r) - \boldsymbol{\gamma}(r)\big)\Big)$ can be calculated as

$$\boldsymbol{\omega}(h,r) = \Big(Cov(f_{i_1t}f_{j_1,t+h}, f_{i_2,t}f_{j_2,t+r}) : 1 \le i_1, j_1, i_2, j_2 \le m\Big)_{m^2\times m^2}.$$

Here we would like to point out that although Theorem 14 of Hannan (1970) gives the CLT for the sample covariance $\check{\boldsymbol{\gamma}} = \frac{1}{T-h}\sum_{t=1}^{T-h} f_{it}f_{j,t+h}$, the asymptotic distribution of $\widehat{\boldsymbol{\gamma}}$ is the same as that of $\check{\boldsymbol{\gamma}}$ (one can verify it along similar lines).

The CLT in Theorem 14 of Hannan (1970) is provided for finite lags h and r only. If both h and r tend to infinity as $T \to \infty$, we develop a corresponding CLT in Lemma 4 and the asymptotic variance is

$$\boldsymbol{\omega}(h,r) = \lim_{h,r\to\infty}\Big(Cov(f_{i_1 t}f_{j_1,t+h}, f_{i_2,t}f_{j_2,t+r}) : 1 \le i_1, j_1, i_2, j_2 \le m\Big)_{m^2\times m^2}. \tag{C.12}$$

Moreover note that the expansion of $vec\Big(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau\Big)$ has a form of

$$vec\Big(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau\Big) = \begin{pmatrix} vec\Big(\widehat{\boldsymbol{\gamma}}(\tau) - \boldsymbol{\gamma}(\tau)\Big) \\ \vdots \\ vec\Big(\widehat{\boldsymbol{\gamma}}(\tau+s) - \boldsymbol{\gamma}(\tau+s)\Big) \\ vec\Big(\widehat{\boldsymbol{\gamma}}(\tau-1) - \boldsymbol{\gamma}(\tau-1)\Big) \\ \vdots \\ vec\Big(\widehat{\boldsymbol{\gamma}}(\tau+s-1) - \boldsymbol{\gamma}(\tau+s-1)\Big) \\ \vdots \\ vec\Big(\widehat{\boldsymbol{\gamma}}(\tau-s) - \boldsymbol{\gamma}(\tau-s)\Big) \\ \vdots \\ vec\Big(\widehat{\boldsymbol{\gamma}}(\tau) - \boldsymbol{\gamma}(\tau)\Big) \end{pmatrix}.$$

In view of this and (C.10), we conclude

$$\sqrt{T-\tau}\Big(vec\big(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau\big)\Big) \xrightarrow{d} N(0, \boldsymbol{\Omega}), \tag{C.13}$$

where $\boldsymbol{\Omega}$ is defined in (3.29).

By (C.13) and Slutsky's theorem, we have, as $N, T \to \infty$,

$$\begin{aligned} &\sqrt{T-\tau}\boldsymbol{\mu}_v'(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau)\bar{\mathbf{v}}_N = \boldsymbol{\mu}_v'\sqrt{T-\tau}(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau)(\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v) + \boldsymbol{\mu}_v'\sqrt{T-\tau}(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau)\boldsymbol{\mu}_v \\ = &\ \big((\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v') \otimes \boldsymbol{\mu}_v'\big)\sqrt{T-\tau}vec(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau) + (\bar{\mathbf{v}}_N' \otimes \boldsymbol{\mu}_v')\sqrt{T-\tau}vec(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau) \\ = &\ (\bar{\mathbf{v}}_N' \otimes \boldsymbol{\mu}_v')\sqrt{T-\tau}vec(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau) + o_p(1) \xrightarrow{d} \mathcal{N}\Big(0, (\boldsymbol{\mu}_v' \otimes \boldsymbol{\mu}_v')\boldsymbol{\Omega}(\boldsymbol{\mu}_v \otimes \boldsymbol{\mu}_v)\Big), \end{aligned} \tag{C.14}$$

where the first equality uses $\text{vec}(\mathbf{AXB}) = (\mathbf{B}' \otimes \mathbf{A})\text{vec}(\mathbf{X})$, with $\mathbf{A} : p \times m$, $\mathbf{B} : n \times q$ and $\mathbf{X} : m \times n$ being three matrices; and $\otimes$ denoting the Kronecker product; and the last asymptotic distribution uses the fact that

$$\bar{\mathbf{v}}_N \xrightarrow{i.p.} \boldsymbol{\mu}_v, \tag{C.15}$$

which can be verified a s similar way to (C.6).

By (C.8), (C.14) and the independence between $\mathbf{S}_\tau$ and $\bar{\mathbf{v}}_N$, we have

$$\begin{aligned} &\sqrt{\min([N^{\alpha_0}], T-\tau)}(\bar{\mathbf{v}}_N'\mathbf{S}_\tau\bar{\mathbf{v}}_N - \boldsymbol{\mu}_v'\boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v) \\ = &\ \sqrt{\min([N^{\alpha_0}], T-\tau)}(\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')(\mathbf{S}_\tau\bar{\mathbf{v}}_N + \boldsymbol{\Sigma}_\tau\boldsymbol{\mu}_v) + \sqrt{\min([N^{\alpha_0}], T-\tau)}\boldsymbol{\mu}_v'(\mathbf{S}_\tau - \boldsymbol{\Sigma}_\tau)\bar{\mathbf{v}}_N \\ &\xrightarrow{d} \mathcal{N}\Big(0, \sigma_0^2\Big), \end{aligned}$$

where the last step uses the fact that

$$[N^{\alpha_0}](\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')\mathbf{S}_\tau \bar{\mathbf{v}}_N = [N^{\alpha_0}](\bar{\mathbf{v}}_N' - \boldsymbol{\mu}_v')\boldsymbol{\Sigma}_\tau \boldsymbol{\mu}_v + o_P(1).$$

□

Lemma 2. *Under Assumptions 1 and 3, we have*

$$\frac{1}{T-\tau}\sum_{t=1}^{T-\tau} \bar{u}_t \bar{u}_{t+\tau} = O_P\Big(\max\big(\frac{\gamma_1(\tau)}{N}, \frac{1}{N\sqrt{T-\tau}}\big)\Big). \tag{C.16}$$

Proof. First, we calculate the order of

$$E\Big(\frac{1}{T-\tau}\sum_{t=1}^{T-\tau} \bar{u}_t \bar{u}_{t+\tau}\Big)^2. \tag{C.17}$$

From Assumption 1, it follows that

$$\begin{aligned}
& E\Big(\frac{1}{T-\tau}\sum_{t=1}^{T-\tau} \bar{u}_t \bar{u}_{t+\tau}\Big)^2 = \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} E(u_{i_1t_1}u_{i_2,t_1+\tau}u_{i_3t_2}u_{i_4,t_2+\tau}) \\
= & \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} E\Big(\sum_{j_1=0}^{+\infty}\phi_{i_1j_1}\sum_{s_1=-\infty}^{+\infty}\xi_{j_1s_1}\nu_{j_1,t_1-s_1}\sum_{j_2=0}^{+\infty}\phi_{i_2j_2}\sum_{s_2=-\infty}^{+\infty}\xi_{j_2s_2}\nu_{j_2,t_1+\tau-s_2} \\
\times & \sum_{j_3=0}^{+\infty}\phi_{i_3j_3}\sum_{s_3=-\infty}^{+\infty}\xi_{j_3s_3}\nu_{j_3,t_2-s_3}\sum_{j_4=0}^{+\infty}\phi_{i_4j_4}\sum_{s_4=-\infty}^{+\infty}\xi_{j_4s_4}\nu_{j_4,t_2+\tau-s_4}\Big).
\end{aligned} \tag{C.18}$$

Note that there are four random terms appearing in the expectation in (C.18), i.e. ν_{j_1,t_1-s_1}, $\nu_{j_2,t_1+\tau-s_2}$, ν_{j_3,t_2-s_3}, $\nu_{j_4,t_2+\tau-s_4}$. By Assumption 1, the expectation is not zero only if these four random terms are pairwise equivalent or all of them are equivalent. In view of this, we have

$$E\Big(\frac{1}{T-\tau}\sum_{t=1}^{T-\tau} \bar{u}_t \bar{u}_{t+\tau}\Big)^2 = \Phi_1 + \Phi_2 + \Phi_3 + \Phi_4, \tag{C.19}$$

where

$$\begin{aligned}
\Phi_1 &= \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} E\Big(\sum_{j_1=0}^{+\infty}\phi_{i_1j_1}\phi_{i_2j_1}\sum_{s_1=-\infty}^{+\infty}\xi_{j_1s_1}\xi_{j_1,s_1+\tau}\nu^2_{j_1,t_1-s_1}\Big) \\
&\times E\Big(\sum_{j_3\neq j_1}^{+\infty}\phi_{i_3j_3}\phi_{i_4j_3}\sum_{s_3=-\infty}^{+\infty}\xi_{j_3s_3}\xi_{j_3,s_3+\tau}\nu^2_{j_3,t_2-s_3}\Big) \\
&= \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} E(u_{i_1t_1}u_{i_2,t_1+\tau})E(u_{i_3t_2}u_{i_4,t_2+\tau}) \\
&= \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} \gamma_1(\tau)\gamma_2(|i_1-i_2|)\gamma_1(\tau)\gamma_2(|i_3-i_4|) = O\Big(\frac{\gamma_1^2(\tau)}{N^2}\Big),
\end{aligned} \tag{C.20}$$

where the first equality uses $\nu_{j_1,t_1-s_1} = \nu_{j_2,t_1+\tau-s_2}$ and $\nu_{j_3,t_2-s_3} = \nu_{j_4,t_2+\tau-s_4}$. The last equality uses (2.6) in the main paper.

For Φ_2,

$$
\begin{aligned}
\Phi_2 &= \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} E\Big(\sum_{j_1=0}^{+\infty}\phi_{i_1j_1}\phi_{i_3j_1}\sum_{s_1=-\infty}^{+\infty}\xi_{j_1s_1}\xi_{j_1,t_2-t_1+s_1}\nu_{j_1,t_1-s_1}^2\Big)\\
&\times E\Big(\sum_{j_2\neq j_1}^{+\infty}\phi_{i_2j_2}\phi_{i_4j_2}\sum_{s_2=-\infty}^{+\infty}\xi_{j_2s_2}\xi_{j_2,t_2-t_1+s_2}\nu_{j_2,t_1+\tau-s_2}\Big)\\
&\leq \frac{K}{N^4(T-\tau)^2}\sum_{t_2=1}^{T-\tau}\sum_{i_1,i_4=1}^{N}\sum_{j_1=0}^{+\infty}|\phi_{i_1j_1}|\sum_{i_3=1}^{N}|\phi_{i_3j_1}|\sum_{s_1=-\infty}^{+\infty}|\xi_{j_1s_1}|\sum_{t_1=1}^{T-\tau}|\xi_{j_1,t_2-t_1+s_1}|\\
&\times \sum_{i_2=1}^{N}|\phi_{i_2j_2}|\sum_{j_2\neq j_1}^{+\infty}|\phi_{i_4j_2}|\sum_{s_2=-\infty}^{+\infty}|\xi_{j_2s_2}| = O\Big(\frac{1}{N^2(T-\tau)}\Big),
\end{aligned}
\tag{C.21}
$$

where the first equality uses $\nu_{j_1,t_1-s_1}=\nu_{j_3,t_2-s_3}$ and $\nu_{j_2,t_1+\tau-s_3}=\nu_{j_4,t_2+\tau-s_4}$. The last equality uses (2.4) in the main paper.

Similarly, for Φ_3, we have

$$
\begin{aligned}
\Phi_3 &= \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} E\Big(\sum_{j_1=0}^{+\infty}\phi_{i_1j_1}\phi_{i_4j_1}\sum_{s_1=-\infty}^{+\infty}\xi_{j_1s_1}\xi_{j_1,t_2-t_1+\tau-s_1}\nu_{j_1,t_1-s_1}^2\Big)\\
&\times E\Big(\sum_{j_2\neq j_1}^{+\infty}\phi_{i_2j_2}\phi_{i_3j_2}\sum_{s_2=-\infty}^{+\infty}\xi_{j_2s_2}\xi_{j_2,t_2-t_1-\tau+s_2}\nu_{j_2,t_1+\tau-s_2}^2\Big)\\
&\leq \frac{K}{N^4(T-\tau)^2}\sum_{t_2=1}^{T-\tau}\sum_{i_1,i_2=1}^{N}\sum_{j_1=0}^{+\infty}|\phi_{i_1j_1}|\sum_{i_4=1}^{N}|\phi_{i_4j_1}|\sum_{s_1=-\infty}^{+\infty}|\xi_{j_1s_1}|\sum_{t_1=1}^{T-\tau}|\xi_{j_1,t_2-t_1+\tau-s_1}|\\
&\times \sum_{j_2\neq j_1}^{+\infty}|\phi_{i_2j_2}|\sum_{i_3=1}^{N}|\phi_{i_3j_2}|\sum_{s_2=-\infty}^{+\infty}|\xi_{j_2s_2}| = O\Big(\frac{1}{N^2(T-\tau)}\Big),
\end{aligned}
\tag{C.22}
$$

where the first equality uses $\nu_{j_1,t_1-s_1}=\nu_{j_4,t_2+\tau-s_4}$ and $\nu_{j_2,t_1+\tau-s_2}=\nu_{j_3,t_2-s_3}$. The last equality uses (2.4) in the main paper.

For Φ_4,

$$
\begin{aligned}
&\Phi_4 = \frac{1}{N^4(T-\tau)^2}\sum_{t_1,t_2=1}^{T-\tau}\sum_{i_1,\ldots,i_4=1}^{N} E\Big(\sum_{j_1=0}^{+\infty}\phi_{i_1j_1}\phi_{i_3j_1}\sum_{s_1=-\infty}^{+\infty}\xi_{j_1s_1}\xi_{j_1,t_2-t_1+s_1}\nu_{j_1,t_1-s_1}^4\\
&\times \phi_{i_2j_1}\phi_{i_4j_1}\xi_{j_1,\tau+s_1}\xi_{j_1,t_2-t_1+\tau+s_1}\Big)\\
&\leq \frac{K}{N^4(T-\tau)^2}\sum_{t_2=1}^{T-\tau}\sum_{i_1=1}^{N}\sum_{j_1=0}^{+\infty}|\phi_{i_1j_1}|\sum_{i_3=1}^{N}|\phi_{i_3j_1}|\sum_{s_1=-\infty}^{+\infty}|\xi_{j_1s_1}|\sum_{t_1=1}^{T-\tau}|\xi_{j_1,t_2-t_1+s_1}|\sum_{i_2=1}^{N}|\phi_{i_2j_1}|\sum_{i_4=1}^{N}|\phi_{i_4j_1}|\\
&= O\Big(\frac{1}{N^3(T-\tau)}\Big),
\end{aligned}
\tag{C.23}
$$

where the first equality uses $\nu_{j_1,t_1-s_1}=\nu_{j_2,t_1+\tau-s_2}=\nu_{j_3,t_2-s_3}=\nu_{j_4,t_2+\tau-s_4}$ and the last equality uses (2.4) in the main paper.

Hence by (C.19), (C.20), (C.21), (C.22) and (C.23), we have

$$
E\Big(\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar{u}_t\bar{u}_{t+\tau}\Big)^2 = O\Big(\max\Big(\frac{\gamma_1^2(\tau)}{N^2},\frac{1}{N^2(T-\tau)}\Big)\Big).
\tag{C.24}
$$

Moreover,

$$E\Big(\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar u_t\bar u_{t+\tau}\Big) = E\Big(\frac{1}{(T-\tau)N^2}\sum_{t=1}^{T-\tau}\sum_{i,j=1}^{N}u_{it}u_{j,t+\tau}\Big)$$

$$= \frac{1}{(T-\tau)N^2}\sum_{t=1}^{T-\tau}\sum_{i,j=1}^{N}\gamma_1(\tau)\gamma_2(|i-j|) = O\Big(\frac{\gamma_1(\tau)}{N}\Big). \quad \text{(C.25)}$$

Therefore, we have

$$Var\Big(\frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\bar u_t\bar u_{t+\tau}\Big) = O\Big(\max\Big(\frac{\gamma_1^2(\tau)}{N^2},\frac{1}{N^2(T-\tau)}\Big)\Big). \quad \text{(C.26)}$$

By (C.26), we have proved (C.16). □

10.2 Two lemmas for Lemma 1

This section is to generalize Theorem 8.4.2 of Anderson (1994) to the case where the time lag tends to infinity along with the sample size. To this end, we first list a crucial lemma below.

Lemma 3 (Theorem 2.1 of Romano and Wolf (2000))**.** *Let $\{X_{n,i}\}$ be a triangular array of mean zero random variables. For each $n = 1, 2, \ldots$, let $d = d_n$, $m' = m_n$, and suppose $X_{n,1},\ldots,X_{n,d}$ is an m'-dependent sequence of random variables. Define $B^2_{n,\ell,a} \equiv Var\Big(\sum_{i=a}^{a+\ell-1}X_{n,i}\Big)$ and $B^2_n \equiv B^2_{n,d,1} \equiv Var\Big(\sum_{i=1}^{d}X_{n,i}\Big)$.*

Let the following conditions hold. For some $\delta > 0$ and some $-1 \le \gamma < 1$:

$$E|X_{n,i}|^{2+\delta} \le \Delta_n \quad for\ all\ i; \quad \text{(C.27)}$$

$$B^2_{n,\ell,a}/(\ell^{1+\gamma}) \le K_n \quad for\ all\ a\ and\ for\ all\ k \ge m'; \quad \text{(C.28)}$$

$$B^2_n/(d(m')^\gamma) \ge L_n; \quad \text{(C.29)}$$

$$K_n/L_n = O(1); \quad \text{(C.30)}$$

$$\Delta_n/L_n^{(2+\delta)/2} = O(1); \quad \text{(C.31)}$$

$$(m')^{1+(1-\gamma)(1+2/\delta)}/d \to 0. \quad \text{(C.32)}$$

Then

$$B_n^{-1}(X_{n,1}+\cdots+X_{n,d}) \Rightarrow \mathcal{N}(0,1). \quad \text{(C.33)}$$

We are now ready to state the following generalization.

Lemma 4. *Let $\mathbf{f}_t = \sum_{r=0}^{+\infty}b_r\boldsymbol{\zeta}_{t-r}$ where $\boldsymbol{\zeta}_t = (\zeta_{1t},\ldots,\zeta_{mt})$, consisting of i.i.d components with zero mean and unit variance, is an i.i.d sequence of m-dimensional random vector. Assume that for some constant $\delta > 0$, $E|\zeta_{it}|^{2+2\delta} < +\infty$; and the coefficients $\{b_r : r = 0, 1, 2, \ldots\}$ satisfy $\sum_{r=0}^{+\infty}|b_r| < +\infty$. Moreover, we assume that*

$$h = o\Big((T-h)^{\delta/(2\delta+2)}\Big) \quad \text{(C.34)}$$

and that the covariance matrix Γ of the random vector

$$\Big(C_{ij}(h') : i = 1,\ldots,m; j = 1,\ldots,m; h' = h-s,\ldots,h+s\Big) \quad \text{(C.35)}$$

is positive definite, where $C_{ij}(h')$ is defined in (3.21).

Then, for any fixed positive constants s and m,

$$\Big(\sqrt{T-h'}\big(C_{ij}(h')-\sigma_{ij}(h')\big) : 1 \le i,j \le m; h-s \le h' \le h+s\Big) \tag{C.36}$$

converges in distribution to a normal distribution with mean 0 and covariances

$$\Big(\lim_{T\to\infty}(T-h)Cov\big(C_{i_1j_1}(h_1), C_{i_2j_2}(h_2)\big) : 1 \le i_1,i_2,j_1,j_2 \le m; h-s \le h_1,h_2 \le h+s\Big). \tag{C.37}$$

Proof. For $1 \le i,j \le m$ and $0 \le h \le T-1$, write $f_{i,t,k} = \sum_{s'=0}^{k} b_{s'}\zeta_{i,t-s'}$, $C_{ij}(h,k) = \frac{1}{T-h}\sum_{t=1}^{T-h} f_{i,t,k}f_{j,t+h,k}$ $= \frac{1}{T-h}\sum_{t=1}^{T-h}\sum_{s_1,s_2=0}^{k} b_{s_1}b_{s_2}\zeta_{i,t-s_1}\zeta_{j,t+h-s_2}$, and

$$\begin{aligned}&\sigma_{ij}(h,k) = E(f_{i,t,k}f_{j,t+h,k})\\ =& \sum_{s_1,s_2=0}^{k} b_{s_1}b_{s_2}E(\zeta_{i,t-s_1}\zeta_{j,t+h-s_2}) = \begin{cases} 0, & i \ne j;\\ \sum_{s_1=0}^{k-h} b_{s_1}b_{h+s_1}, & i=j; h=0,1,\dots,k;\\ 0, & i=j; h=k+1,k+2.\end{cases}\end{aligned} \tag{C.38}$$

The proof of this lemma is similar to that of Theorem 8.4.2 of Anderson (1994) and it can be divided into two steps:

Step 1: For any fixed k, the first step is to provide the asymptotic theorem for

$$\Big(\sqrt{T-h'}\big(C_{ij}(h',k)-\sigma_{ij}(h',k)\big) : 1 \le i,j \le m; h-s \le h' \le h+s\Big); \tag{C.39}$$

Step 2: The second step is to prove that for any $1 \le i,j \le m$, in probability,

$$\lim_{T\to\infty}\sqrt{T-h}\Big(C_{ij}(h')-C_{ij}(h',k)\Big) = 0. \tag{C.40}$$

The second step can be verified as in Theorem 8.4.2 of Anderson (1994) (i.e. page 479-page 481) and the details are omitted here.

Consider Step 1 now. Let

$$X_{T-h,t}(i,j) = \frac{1}{\sqrt{T-h}}\Big(f_{i,t,k}f_{j,t+h,k}-\sigma_{ij}(h,k)\Big), \quad 1 \le i,j \le m, \tag{C.41}$$

so that

$$\sqrt{T-h}\Big(C_{ij}(h,k)-\sigma(h,k)\Big) = \sum_{t=1}^{T-h} X_{T-h,t}(i,j). \tag{C.42}$$

By simple calculations, we see that $f_{i,t,k}f_{j,t+h,k}$ is independent of $f_{i,g,k}f_{j,g+h,k}$ if t and g differ by more than $k+h$ when $i \ne j$ and differ by more than k when $i=j$. So $\{f_{i,t,k}f_{j,t+h,k} : t=1,\dots,T-h\}$ is a $(k+h)$ or k dependent covariance stationary process with mean $\sigma_{ij}(h,k)$ and covariance

$$\begin{aligned}&Cov(f_{i,t,k}f_{j,t+h,k}, f_{i,g,k}f_{j,g+h,k})\\ =& \sum_{s_1,\dots,s_4=0}^{k} b_{s_1}b_{s_2}b_{s_3}b_{s_4}E(\zeta_{i,t-s_1}\zeta_{j,t+h-s_2}\zeta_{i,g-s_3}\zeta_{j,g+h-s_4}) - \sigma_{ij}^2(h,k)\\ =& \begin{cases} A_1, & i \ne j;\\ \sum_{q=1}^{4} A_q - \sigma_{ii}^2(h,k), & i=j,\end{cases}\end{aligned} \tag{C.43}$$

where

$$A_1 = \sum_{s_1=0}^{k}\sum_{s_2=0}^{k} b_{s_1}b_{s_2}b_{g-t+s_1}b_{g-t+s_2}, \quad A_2 = \sum_{b_1=0}^{k}\sum_{b_3=0}^{k} b_{s_1}b_{h+s_1}b_{s_3}b_{h+s_3},$$
$$A_3 = \sum_{s_1=0}^{k}\sum_{s_3=0}^{k} b_{s_1}b_{t-g+h+s_3}b_{s_3}b_{g-t+h+s_1}, \quad A_4 = -2\sum_{s_1=0}^{k} b_{s_1}b_{h+s_1}b_{g-t+s_1}b_{g-t+h+s_1},$$

where (C.43) uses the fact that $E(\zeta_{i,t-s_1}\zeta_{j,t+h-s_2}\zeta_{i,g-s_3}\zeta_{j,g+h-s_4})$ is not equal to zero if and only if the four terms $\zeta_{i,t-s_1}, \zeta_{j,t+h-s_2}, \zeta_{i,g-s_3}, \zeta_{j,g+h-s_4}$ are pairwise equivalent or they are all equivalent.

Hence for any $1 \le i,j \le m; h-s \le h^{'} \le h+s$, $\{X_{T-h^{'},t}(i,j) : t = 1,\ldots,T-h^{'}\}$ is a $(k+h^{'})$ or k dependent covariance stationary process. This implies that any linear combination of the process $\{\sum_{i,j=1}^{m}\sum_{h^{'}=h-s}^{h+s} a_{i,j,h^{'}} X_{T-h^{'},t}(i,j) : t = 1,\ldots,T-h-s\}$ is a $(k+h+s)$ dependent covariance stationary process. Thus, we need to check Conditions (C.27)–(C.32) for such a linear combination of the process. Moreover, one should note that it is enough to justify those conditions for each stochastic process $\{X_{T-h^{'},t}(i,j) : t = 1,\ldots,T-h^{'}\}$, where $1 \le i,j \le m; h-s \le h^{'} \le h+s$, since s and m are both fixed.

Observe that

$$\begin{aligned} & E\Big|X_{T-h^{'},t}(i,j)\Big|^{2+\delta} = \Big(\frac{1}{T-h^{'}}\Big)^{(2+\delta)/2} E\Big|f_{i,t,k}f_{j,t+h,k} - \sigma_{ij}(h,k)\Big|^{2+\delta} \\ \le\ & K\Big(\frac{1}{T-h^{'}}\Big)^{(2+\delta)/2}\Big(E\Big|f_{i,t,k}f_{j,t+h,k}\Big|^{2+\delta} + \Big|\sigma_{i,j}(h,k)\Big|^{2+\delta}\Big) \\ \le\ & K\Big(\frac{1}{T-h^{'}}\Big)^{(2+\delta)/2}, \end{aligned} \tag{C.44}$$

where K is a constant number, and we have also used (C.38) and the fact that

$$\begin{aligned} & E\Big|f_{i,t,k}f_{j,t+h,k}\Big|^{2+\delta} = E\Big|\sum_{s_1,s_2=0}^{k} b_{s_1}b_{s_2}\zeta_{i,t-s_1}\zeta_{j,t+h-s_2}\Big|^{2+\delta} \\ \le\ & K\sum_{s_1,s_2=0}^{k}|b_{s_1}|^{2+\delta}|b_{s_2}|^{2+\delta}\Big(E|\zeta_{i,t-s_1}|^{4+2\delta} + E|\zeta_{j,t+h-s_2}|^{4+2\delta}\Big) = O(1). \end{aligned} \tag{C.45}$$

In view of (C.44), taking $\Delta_T = K\Big(\frac{1}{T-h^{'}}\Big)^{(2+\delta)/2}$, we have

$$E\Big|X_{T-h^{'},t}(i,j)\Big|^{2+\delta} \le \Delta_T, \tag{C.46}$$

implying (C.27).

We obtain from (C.43) that

$$\begin{aligned} & B_{T,\ell,a}^2(i,j) \equiv Var\Big(\sum_{t=a}^{a+\ell-1} X_{T-h^{'},t}(i,j)\Big) \\ =\ & \frac{1}{T-h^{'}}\sum_{t=a}^{a+\ell-1}\sum_{g=a}^{a+\ell-1} Cov(f_{i,t,k}f_{j,t+h^{'},k}, f_{i,g,k}f_{j,g+h^{'},k}) \\ =\ & \begin{cases} \frac{1}{T-h^{'}}\sum_{t=a}^{a+\ell-1}\sum_{g=a}^{a+\ell-1} A_1, & i \ne j; \\ \frac{1}{T-h^{'}}\sum_{t=a}^{a+\ell-1}\sum_{g=a}^{a+\ell-1}\Big(\sum_{q=1}^{4} A_q - \sigma_{ii}^2(h^{'},k)\Big), & i = j. \end{cases} \end{aligned} \tag{C.47}$$

Note that $A_2 = \sigma_{ii}^2(h^{'}, k)$. Below we only evaluate the remaining terms involving A_1, A_3, A_4. By the fact that $\sum_{r=0}^{+\infty} |b_r| < +\infty$, we have

$$\begin{aligned} \Big|\frac{1}{T-h^{'}} \sum_{t=a}^{a+\ell-1} \sum_{g=a}^{a+\ell-1} A_1\Big| &= \Big|\frac{1}{T-h^{'}} \sum_{t=a}^{a+\ell-1} \sum_{g=a}^{a+\ell-1} \sum_{s_1=0}^{k} \sum_{s_2=0}^{k} b_{s_1} b_{s_2} b_{g-t+s_1} b_{g-t+s_2}\Big| \\ \le \quad & \frac{K}{T-h^{'}} \sum_{t=a}^{a+\ell-1} \sum_{s_1=0}^{k} |b_{s_1}| \sum_{s_2=0}^{k} |b_{s_2}| \sum_{g=a}^{a+\ell-1} |b_{g-t+s_1}| = O\Big(\frac{\ell}{T-h^{'}}\Big). \end{aligned} \tag{C.48}$$

Similarly, one may verify that

$$\frac{1}{T-h^{'}} \sum_{t=a}^{a+\ell-1} \sum_{g=a}^{a+\ell-1} A_j = O\Big(\frac{\ell}{T-h^{'}}\Big), \ j = 3, 4. \tag{C.49}$$

We conclude from (C.47)-(C.49) that

$$B_{T,\ell,a}^2(i,j) = O\Big(\frac{\ell}{T-h}\Big). \tag{C.50}$$

Taking $\ell = T-h$ in $B_{T,\ell,a}^2$, we have $B_T^2 = O(1)$. Moreover, for any linear combination of the process $\{\sum_t \sum_{i,j=1}^m \sum_{h^{'}=h-s}^{h+s} a_{i,j,h^{'}} X_{T-h^{'},t}(i,j) : t = 1, \ldots, T-h-s\}$, by the assumption of $\Gamma > 0$ (see (C.35), its variance B_T^2 is

$$B_T^2 = \mathbf{a}^{'}\mathbf{\Gamma}\mathbf{a} > 0. \tag{C.51}$$

In view of (C.50) and (C.51), we can take $\gamma = 0$, $K_T = \widetilde{K}_1 \frac{1}{T-h-s}$ and $L_T = \widetilde{K}_2 \frac{1}{T-h-s}$ for the purpose of verifying (C.28)–(C.32), where $\widetilde{K}_1$ and $\widetilde{K}_2$ are two constants. Then

$$\begin{aligned} &\frac{B_{T,\ell,a}^2}{\ell^{1+\gamma}} \le K_T, \quad \textit{for all } a \textit{ and all } \ell \ge k+h; \\ &\frac{B_T^2}{(T-h^{'})(k+h)^{\gamma}} \ge L_T. \end{aligned} \tag{C.52}$$

Moreover, K_T, L_T and Δ_T satisfy

$$\frac{K_T}{L_T} = O(1) \ \text{ and } \ \frac{\Delta_T}{L_T^{(2+\delta)/2}} = O(1). \tag{C.53}$$

By (C.34), we have that, for any fixed k,

$$\frac{(k+h)^{2+2/\delta}}{T-h} \to 0, \quad \textit{as } T \to \infty. \tag{C.54}$$

From (C.46), (C.52), (C.53), (C.54) and Lemma 3, we conclude that

$$\Big(\sqrt{T-h^{'}}\big(C_{ij}(h^{'},k) - \sigma_{ij}(h^{'},k)\big) : 1 \le i,j \le m; h-s \le h^{'} \le h+s\Big)$$

converges in distribution to a standard normal distribution with mean zero and covariances

$$\Big(\lim_{T\to\infty} (T-h) Cov\big(C_{i_1j_1}(h_1,k), C_{i_2j_2}(h_2,k)\big) : 1 \le i_1, i_2, j_1, j_2 \le m; h-s \le h_1, h_2 \le h+s\Big).$$

Hence the proof of step 1 is completed.

□