Exponent of Cross-sectional Dependence: Estimation and Inference*

Natalia Bailey
Queen Mary, University of London

George Kapetanios
Queen Mary, University of London

M. Hashem Pesaran
University of Southern California and Trinity College, Cambridge

November 5, 2013

Abstract

In this paper we provide a characterization of the degree of cross-sectional dependence in a two dimensional array, $\{x_{it}, i = 1, 2, ...N; t = 1, 2, ..., T\}$ in terms of the rate at which the variance of the cross-sectional average of the observed data varies with N. We show that under certain conditions this is equivalent to the rate at which the largest eigenvalue of the covariance matrix of $\boldsymbol{x}_t = (x_{1t}, x_{2t}, ..., x_{Nt})'$ rises with N. We represent the degree of cross-sectional dependence by α, defined by the standard deviation, $Std(\bar{x}_t) = O\left(N^{\alpha-1}\right)$, where $\bar{x}_t$ is a simple cross-sectional average of x_{it}. We refer to α as the 'exponent of cross-sectional dependence', and show how it can be consistently estimated for values of $\alpha > 1/2$. We propose bias corrected estimators, derive their asymptotic properties and consider a number of extensions. We include a detailed Monte Carlo study supporting the theoretical results. We also provide a number of empirical applications investigating the degree of inter-linkages of real and financial variables in the global economy, the extent to which macroeconomic variables are interconnected across and within countries.

Keywords: Cross correlations, Cross-sectional dependence, Cross-sectional averages, Weak and strong factor models

JEL Codes: C21, C32

1 Introduction

Over the past decade there has been a resurgence of interest in the analysis of cross-sectional dependence applied to households, firms, markets, regional and national economies. Researchers in many fields have turned to network theory, spatial and factor models to obtain a better understanding of the extent and nature of such cross dependencies. There are many issues to

*We are grateful to Jean-Marie Dufour and Oliver Linton for helpful comments and discussions. Natalia Bailey and Hashem Pesaran acknowledge financial support under ESRC Grant ES/I031626/1.

be considered: how to test for the presence of cross-sectional dependence, how to measure the degree of cross-sectional dependence, how to model cross-sectional dependence, and how to carry out counterfactual exercises under alternative network formations or market inter-connections. Many of these topics are the subject of ongoing research. In this paper we focus on measures of cross-sectional dependence and how such measures are related to the behaviour of cross-sectional averages or aggregates.

The literature on cross-sectional dependence distinguishes between strong and weak forms of dependence, with the former typically associated with factor models and the latter with the spatial models. In finance, the approximate factor model of Chamberlain (1983) provides a popular characterization of cross-sectional dependence of asset returns in terms of a factor dependence and a remainder term. The factors are intended to capture the pervasive market effects, whilst the remainder term is assumed to be only weakly cross-sectionally correlated (Ross (1976), Ross (1977)). Strong and weak cross-sectional dependence are defined in terms of the rate at which the largest eigenvalue of the covariance matrix of the cross section units rises with the number of the cross section units. See, for example, Chudik et al. (2011).

Let x_{it} denote a double array of random variables indexed by $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, over space and time, respectively, and without loss of generality assume that $E(x_{it}) = 0$. Then the covariance matrix of $\boldsymbol{x}_t = (x_{1t}, x_{2t}, ..., x_{Nt})'$ is given by $\boldsymbol{\Sigma}_N = E(\boldsymbol{x}_t \boldsymbol{x}_t') = (\sigma_{ij,x})$ with its largest eigenvalue denoted by $\lambda_{\max}(\boldsymbol{\Sigma}_N)$. The variables x_{it} are said to be strongly cross-sectionally correlated if $\lambda_{\max}(\boldsymbol{\Sigma}_N)$ rises with N, and they are said to be weakly cross-sectionally correlated if $\lambda_{\max}(\boldsymbol{\Sigma}_N)$ is bounded in N. This is clearly an important distinction and forms the basis of most factor models considered in finance and macroeconometric literature - Forni et al. (2000), Forni and Lippi (2001), Bai and Ng (2002) and Bai (2003). However, since both the factors and their loadings are unobserved it is desirable to consider statistical techniques that test the strong/weak factor assumption. A test of weak cross-sectional dependence is recently proposed by Pesaran (2013), but a more general framework is needed to address intermediate cases between weak and strong forms of dependence. Such intermediate cases can be parameterized in terms of the exponent α, such that $N^{-\alpha}\lambda_{\max}(\boldsymbol{\Sigma}_N) = O(1)$, and

$$\lim_{N\to\infty} N^{-\alpha}\lambda_{\max}(\boldsymbol{\Sigma}_N) > 0. \tag{1}$$

The weak and strong dependence cases then relate to $\alpha = 0$ and $\alpha = 1$, respectively.

$\lambda_{\max}(\boldsymbol{\Sigma}_N)$ and its limiting properties have been the object of considerable interest in the statistical literature on large data sets. However, former work in the area (see, e.g., Yin et al. (1988), Bai and Silverstein (1998), Hachem et al. (2005a) and Hachem et al. (2005b)) and more recent contributions that allow unequal eigenvalues in the design of the population covariance matrix (see, e.g., Fan et al. (2013) and Shen et al. (2013)) suggest that as a statistical measure of cross-sectional dependence $\lambda_{\max}(\boldsymbol{\Sigma}_N)$ could be difficult to analyse, especially for temporally and cross-sectionally dependent data as the theoretical asymptotic properties of sample estimators of $\lambda_{\max}(\mathbf{\Sigma}_N)$ depend crucially on the form of the dependence. In particular, we note that this work predominantly uses i.i.d. and gaussianity assumptions in its approach.

In this paper we consider a simpler alternative measure based on cross sectional averages defined by $\bar{x}_t = N^{-1}\sum_{i=1}^{N} x_{it}$. The limiting behaviour of $\bar{x}_t$ is of interest in its own right and provides information on the nature and degree of cross-sectional dependence. In the case of asset returns this determines the extent to which risk, associated with investing in particular portfolios of assets, is diversifiable. In the case of firm sales this is of interest in relation to the effect of idiosyncratic, firm level, shocks onto aggregate macroeconomic variables such as GDP. In the case where x_{it} are cross-sectionally independent, using standard law of large numbers, one obtains the result that $Var(\bar{x}_t) = O\left(N^{-1}\right)$. However, in the more general and realistic case where x_{it} are cross-sectionally correlated, we have that $Var(\bar{x}_t)$ declines at a rate that is a function of α where α is defined in (1). We note that $Var(\bar{x}_t)$ cannot decline at a rate faster than N^{-1}. It is also easily seen that $Var(\bar{x}_t)$ cannot decline at a rate slower than $N^{\alpha-1}$, $0 \leq \alpha \leq 1$. To see this we explore the link between $\lambda_{\max}(\mathbf{\Sigma}_N)$ and $Var(\bar{x}_t)$. Note that $\bar{x}_t = N^{-1}\iota'\boldsymbol{x}_t$, where ι is an $N \times 1$ vector of ones. Then, we have

$$Var(\bar{x}_t) = N^{-2}\iota'\mathbf{\Sigma}_N\iota \leq N^{-2}\iota'\iota\ \lambda_{\max}(\mathbf{\Sigma}_N) = N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N).$$

Therefore, α defined by $N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N) = O(N^{\alpha-1})$ provides an upper rate for $Var(\bar{x}_t)$.

It is interesting to note that the above measures of cross-sectional dependence are also related to the degree of pervasiveness of factors in unobserved factor models often used in the literature to model cross-sectional dependence.[1] As an illustration consider the simple factor model

$$x_{it} = a_i + \beta_{i1} f_{1t} + u_{it} \text{ for } i = 1, 2, ..., N;\ \ t = 1, 2, ..., T, \tag{2}$$

where x_{it} depends on a single unobserved factor f_{1t}, with the associated factor loadings, β_{i1}, and cross-sectionally independent idiosyncratic errors, u_{it}. The extent of cross-sectional dependence in x_{it} crucially depends on the nature of the factor loadings. It is easily seen that $\lambda_{\max}(\mathbf{\Sigma}_N) = O\left(\sum_{i=1}^{N}\beta_{i1}^2\right)$, the column sum norm[2] of $\mathbf{\Sigma}_N$, defined by $\|\mathbf{\Sigma}_N\|_1 = \sup_j \sum_{i=1}^{N}|\sigma_{ij,x}|$, is of order $O\left(\left(\sup_j|\beta_{j1}|\right)\sum_{i=1}^{N}|\beta_{i1}|\right)$ and $Var(\bar{x}_t) = O\left\{\max\left[\left(N^{-1}\sum_{i=1}^{N}\beta_{i1}\right)^2, N^{-1}\right]\right\}$. The degree of cross-sectional dependence will be strong if the average value of β_{i1} is bounded away from zero. In such a case, $N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N)$, $\|\mathbf{\Sigma}_N\|_1$ and $Var(\bar{x}_t)$ are all $O(1)$, which yields $\alpha = 1$.

However, other configurations of factor loadings can also be entertained, that yield values

[1] Factor models have a long pedigree both as a conceptual device for summarising multivariate data sets as well as an empirical framework with sound theoretical underpinnings both in finance and economics. Recent econometric research on factor models include Bai and Ng (2002), Bai (2003), Forni et al. (2000), Forni and Lippi (2001), Pesaran (2006) and Stock and Watson (2002).

[2] We introduce the column sum norm of $\mathbf{\Sigma}_N$ as it is a commonly used measure of cross-sectional dependence alongside $\lambda_{\max}(\mathbf{\Sigma}_N)$.

of α in the range $(0,1]$. Since both f_{1t} and β_{i1} are unobserved, taking a strong stand on a particular value of α might not be justified empirically. Accordingly, Chudik et al. (2011), Kapetanios and Marcellino (2010) and Onatski (2012) have considered an extension of the above factor model which allows the factor loadings, β_{i1}, to vary with N. In particular, by considering $\beta_{i1} = O(N^{(\alpha-1)/2})$, for any $0 < \alpha < 1$. This specification implies $N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N) = O\left(N^{\alpha-1}\right)$, $N^{-1}\left\|\mathbf{\Sigma}_N\right\|_1 = O\left(N^{\alpha-1}\right)$, as long as $\max_i \beta_{i1} = o_p\left(N^d\right)$, for all $d > 0$ and $Var\left(\bar{x}_t\right) = O\left(N^{\alpha-1}\right)$.[3]

Although mathematically convenient, the assumption that all factor loadings vary with N (almost uniformly) is rather restrictive in many economic applications. Therefore, we will not consider it in detail, but only briefly as an alternative formulation. In this paper we consider a baseline formulation where we assume that only N^α of the N factor loadings are individually important, in the sense that they are bounded away from zero. More specifically, we consider $\beta_{i1} = v_{i1}$, for $i = 1, 2, ..., [N^\alpha]$, and $\beta_{i1} = \tilde{v}_{i1}$, for $i = [N^\alpha] + 1, [N^\alpha] + 2, ..., N$, where $[N^\alpha]$ is the integer part of N^α, $0 < \alpha \leq 1$, $v_{i1} \sim iid(\mu_{v_1}, \sigma^2_{v_1})$, $\mu_{v_1} \neq 0$, $\sigma^2_{v_1} > 0$ and $\sum_{i=[N^\alpha]+1}^{N} \tilde{v}_{i1} = O_p(1)$. In effect, the factor loadings are grouped into two categories: a strong category with effects that have non-zero mean, and a weak category with negligible effects with a mean that tends to zero with N. Under this setup, $N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N) = O\left(N^{\alpha-1}\right)$, $N^{-1}\left\|\mathbf{\Sigma}_N\right\|_1 = O\left(N^{\alpha-1}\right)$, as long as $\max_i \beta_{i1} = o_p\left(N^d\right)$, for all $d > 0$, $Var\left(\bar{x}_t\right) = O\left(N^{2\alpha-2}\right)$ and the standard deviation of $\bar{x}_t$, denoted by $Std\left(\bar{x}_t\right)$ is $O\left[\max\left(N^{\alpha-1}, N^{-1/2}\right)\right]$. At least $N^{1/2}$ of the loadings must have non-zero mean for the covariances in $\mathbf{\Sigma}_N$ to dominate the diagonal of $\mathbf{\Sigma}_N$ and result in a rate of decline for $Std(\bar{x}_t)$ that is $O\left(N^{\alpha-1}\right)$. If fewer than $N^{1/2}$ of the loadings have non-zero mean, then $Std(\bar{x}_t) = O\left(N^{-1/2}\right)$. The presence of at least $N^{1/2}$ loadings with non-zero mean implies that $\alpha > 1/2$. In that case, and as long as $\mu_{v_1} \neq 0$, $N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N)$, $N^{-1}\left\|\mathbf{\Sigma}_N\right\|_1$ and $Std\left(\bar{x}_t\right)$ decline at the same rate. As a result in the context of the factor model in (2), α has a unique role as a measure of cross-sectional dependence. It is important to note that if $\mu_v = \sum_{k=1}^{m} \mu_{v_k} = 0$, where m is the number of factors, then $Std\left(\bar{x}_t\right) = O\left(N^{-1/2}\right)$ for all α including the case $\alpha = 1$. The implication is that even a strong factor model allows full portfolio diversification at the same rate as if no factors were present. Seen from this perspective, the case where $\mu_v = 0$ does not seem very plausible, at least in the case of macro and financial data sets.

As we shall see, since we are interested in the behaviour of cross-sectional averages, our proposed estimator of α will be invariant to the ordering of the factor loadings within each group. The only important consideration is that there exists a split between loadings with non-zero mean and loadings that are cumulatively of a small order. The split need not be known.

Following the theoretical line of reasoning advanced above, in this paper we propose the use of the variance of the cross-sectional average of the observed data, $\bar{x}_t$, to estimate and carry out inference on α. We provide a feasible estimator for α under a multiple factor setting and derive inferential theory for it. We derive the asymptotic distribution of our estimator for a given value

[3] A different strand of literature that deals with weaker forms of cross-sectional dependence includes spatial econometric models. These correspond to the case of $\alpha = 0$.

of μ_v^2. Further, we present our preferred estimator that additionally takes into account the term μ_v^2 together with its asymptotic properties. We consider extensions that relate to the presence of temporal dependence in $\boldsymbol{f}_t = (f_{1t}, ..., f_{mt})'$ or u_{it}, and weak cross-sectional dependence in u_{it}. It is also worth pointing out that our estimators of α do not use explicitly a factor structure. The factor representation is only needed as a vehicle to derive the theoretical properties of the estimator and to give α a unique interpretation as a measure of cross-sectional dependence. We use this vehicle because working with covariances directly would involve high level assumptions and would potentially lead to stricter conditions such as the need for T to rise faster than N. A further crucial reason for using the factor model is that, as proven in Theorem 4 of Chamberlain and Rothschild (1983), a covariance matrix that has a finite number of eigenvalues that tend to infinity as N increases, has a unique factor representation. This makes the factor model a canonical model for analysing cross-sectional dependence associated with covariance matrices with a finite number of exploding eigenvalues.

To illustrate the properties of the proposed estimators of α and their asymptotic distributions, we carry out a detailed Monte Carlo study that considers a battery of robustness checks. Finally, we provide a number of empirical applications investigating the degree of inter-linkages in real and financial variables in the global economy, the extent to which macroeconomic variables are interconnected across and within countries, with special reference to the US and UK economies in the second case.

The rest of the paper is organised as follows: Section 2 provides a formal characterisation of α in the context of a single factor model, and discusses potential estimation strategies. This section also presents the rudiments of the analysis of the variance of the cross-sectional average and motivates the baseline estimator and bias corrected versions of it. Section 3 presents the theoretical results of the paper. Section 3.1 provides the full inferential theory under a multiple factor set up. Section 3.2 deals with possible cross sectional dependence in the error terms and touches upon an alternative specification of factor loadings. Section 4 presents a detailed Monte Carlo study. The empirical applications are discussed in Section 5. Finally, Section 6 concludes. Proofs of all theoretical results are relegated to Appendices.

Notations: $\|\mathbf{A}\| = [Tr(\mathbf{AA}')]^{1/2}$ is the Frobenius norm of the $m \times n$ matrix $\mathbf{A}$. $\sup_i W_i$ is the supremum of W_i over i. $a_n = O(b_n)$ states the deterministic sequence $\{a_n\}$ is at most of order b_n, $\boldsymbol{x}_n = O_p(\mathbf{y}_n)$ states the vector of random variables, $\boldsymbol{x}_n$, is at most of order $\mathbf{y}_n$ in probability, and $\boldsymbol{x}_n = o_p(\mathbf{y}_n)$ is of smaller order in probability than $\mathbf{y}_n$, $\to_p$ denotes convergence in probability, and $\to_d$ convergence in distribution. All asymptotics are carried out under $N \to \infty$, *jointly* with $T \to \infty$.

2 Preliminaries and Motivations

In this Section we provide an informal account of the concept of the exponent of cross-sectional dependence and our proposed estimator of it. We consider the single factor model given by (2) as it allows a simpler exposition. Our formal theoretical analysis, provided in Section 3, is couched in terms of a multiple factor model and is therefore appropriately general. We specify the loadings as follows

$$\begin{aligned}\beta_{i1} &= v_{i1} \text{ for } i = 1, 2, ..., [N^{\alpha}], \\ \beta_{i1} &= \tilde{v}_{i1}, \text{ for } i = [N^{\alpha}] + 1, [N^{\alpha}] + 2, ..., N,\end{aligned} \tag{3}$$

where $1/2 < \alpha \leq 1$, $[N^{\alpha}]$ is the integer part of N^{α}, and $\{v_{i1}\}_{i=1}^{[N^{\alpha}]}$ is an identically, independently distributed (IID) sequence of random variables with mean $\mu_{v_1} \neq 0$ and variance $\sigma_{v_1}^2 < \infty$. Also, $\sum_{i=[N^{\alpha}]+1}^{N} \tilde{v}_{i1} = O_p(1)$. Throughout our analysis and without loss of generality, we assume that factors have unit variance, and, in the case of multiple factors, are uncorrelated with each other. We introduce the subscript 1 for β_{i1}, v_{i1}, μ_{v_1} and $\sigma_{v_1}^2$, in anticipation of our multiple factor analysis in the next section.

In effect, the factor loadings are grouped into two categories: a strong category with effects that have non-zero mean, and a weak category with negligible effects and a mean that tends to zero with N. Under this setup, $N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N) = O\left(N^{\alpha-1}\right)$, $N^{-1}\left\|\mathbf{\Sigma}_N\right\|_1 = O\left(N^{\alpha-1}\right)$, as long as $\max_i \beta_{i1} = o_p\left(N^d\right)$, for all $d > 0$, $Var(\bar{x}_t) = O\left(N^{2\alpha-2}\right)$ and the standard deviation of $\bar{x}_t$, denoted by $Std(\bar{x}_t)$ is $O\left[\max\left(N^{\alpha-1}, N^{-1/2}\right)\right]$. At least $N^{1/2}$ of the loadings must have non-zero mean for the covariances in $\mathbf{\Sigma}_N$ to dominate the diagonal of $\mathbf{\Sigma}_N$ and result in a rate of decline for $Std(\bar{x}_t)$ that is $O\left(N^{\alpha-1}\right)$. If fewer than $N^{1/2}$ of the loadings have non-zero mean, then $Std(\bar{x}_t) = O\left(N^{-1/2}\right)$. The presence of at least $N^{1/2}$ loadings with non-zero mean implies that $\alpha > 1/2$. In that case, and as long as $\mu_{v_1} \neq 0$, $N^{-1}\lambda_{\max}(\mathbf{\Sigma}_N)$, $N^{-1}\left\|\mathbf{\Sigma}_N\right\|_1$ and $Std(\bar{x}_t)$ decline at the same rate. As a result, in the context of the factor model in (2), α has a unique role as a measure of cross-sectional dependence. The above loading setup implies that $N^{-1}\sum_{i=1}^{N}\beta_{i1}^2 = O_p\left(N^{\alpha-1}\right)$, which is more general than the standard assumption in the factor literature that requires $N^{-1}\sum_{i=1}^{N}\beta_{i1}^2$ to have a strictly positive limit (see, e.g., Assumption B of Bai and Ng (2002)). The standard assumption is satisfied only if $\alpha = 1$. It is important to note that if $\mu_{v_1} = 0$, then $Std(\bar{x}_t) = O\left(N^{-1/2}\right)$ for all α including the case $\alpha = 1$. The implication is that even a strong factor model allows full portfolio diversification at the same rate as if no factors were present. Seen from this perspective, the case where $\mu_{v_1} = 0$ does not seem very plausible, at least in the case of macro and financial data sets.

To motivate our choice of α as the exponent of cross-sectional dependence of x_{it}, we write (2) as

$$\boldsymbol{x}_t = \mathbf{a} + \boldsymbol{\beta}\ f_{1t} + \mathbf{u}_t, \tag{4}$$

where $\boldsymbol{x}_t = (x_{1t}, x_{2t}, ..., x_{Nt})'$, $\mathbf{a} = (a_1, a_2, ..., a_N)'$, $\boldsymbol{\beta} = (\beta_{11}, \beta_{21}, ..., \beta_{N1})'$ and $\mathbf{u}_t = (u_{1t}, u_{2t}, ..., u_{Nt})'$. We also note that under the above assumptions, $\mathbf{\Sigma}_\beta = E(\boldsymbol{\beta}\boldsymbol{\beta}') - E(\boldsymbol{\beta})E(\boldsymbol{\beta}')$, with $\lambda_{\max}(\mathbf{\Sigma}_\beta) < K < \infty$, $\mathbf{\Sigma}_u = E(\mathbf{u}_t\mathbf{u}_t')$, with $\lambda_{\max}(\mathbf{\Sigma}_u) < K < \infty$, $\mu_{f_1} = E(f_{1t}) = 0$, $\sigma^2_{f_1} = E(f_{1t} - \mu_{f_1})^2 = 1$, and f_{1t} and $\boldsymbol{\beta}$ are distributed independently. Hence,

$$Cov(\boldsymbol{x}_t) = \left[\mathbf{\Sigma}_\beta + E(\boldsymbol{\beta})E(\boldsymbol{\beta}')\right] + \mathbf{\Sigma}_u.$$

Consider now the cross-sectional averages of the observables defined by $\bar{x}_t = \iota'\boldsymbol{x}_t/N$, where ι is an $N \times 1$ vector of ones. Then,

$$Var(\bar{x}_t) = N^{-2}\iota' Cov(\boldsymbol{x}_t)\iota = N^{-2}\iota'\left[\mathbf{\Sigma}_\beta + \mathbf{\Sigma}_u\right]\iota + \left[\frac{\iota' E(\boldsymbol{\beta})}{N}\right]^2. \tag{5}$$

But under (3), it follows that

$$\sum_{i=1}^{N}\beta_{i1} = [N^\alpha]\left(\frac{1}{[N^\alpha]}\sum_{i=1}^{[N^\alpha]} v_{i1}\right) = [N^\alpha]\,\bar{v}_{1N},$$

where $\bar{v}_{1N} = \frac{1}{[N^\alpha]}\sum_{i=1}^{[N^\alpha]} v_{i1}$ is $O_p(1)$ and $\frac{1}{N-[N^\alpha]}\sum_{i=[N^\alpha]+1}^{N}\beta_{i1} \to 0$, for $i > [N^\alpha]$. Recall that any sequence of loadings, for which $\sum_{i=[N^\alpha]+1}^{N}\beta_{i1} = O_p(1)$ is permitted. Hence,

$$N^{-1}\iota' E(\boldsymbol{\beta}) = \mu_{v_1}\left[N^{\alpha-1}\right].$$

Also,

$$N^{-2}\iota'\mathbf{\Sigma}_\beta\iota = N^{-2}\iota_1'\mathbf{\Sigma}_{\beta(1)}\iota_1 \leq \left[N^{\alpha-2}\right]\lambda_{\max}\left(\mathbf{\Sigma}_\beta\right),$$

where ι_1' is an $[N^\alpha] \times 1$ vector of ones and $\mathbf{\Sigma}_{\beta(1)}$ is the upper $[N^\alpha] \times [N^\alpha]$ sub-matrix of $\mathbf{\Sigma}_\beta$. Using the above results in (5) we now have

$$Var(\bar{x}_t) \leq \left[N^{\alpha-2}\right]\lambda_{\max}\left(\mathbf{\Sigma}_\beta\right) + N^{-1}c_N + \mu_v^2\left[N^{2\alpha-2}\right], \tag{6}$$

where

$$c_N = \frac{\iota'\mathbf{\Sigma}_u\iota}{N} < K < \infty. \tag{7}$$

Note that μ_v^2 enters (6), rather than $\mu_{v_1}^2$. In the case of a single factor $\mu_v^2 =: \mu_{v_1}^2$. However, for multiple factors μ_v^2 will be defined in terms of the means of the loadings of all the factors in a way that will be discussed in detail in the next Section. By assumption $\lambda_{\max}\left(\mathbf{\Sigma}_\beta\right) < K < \infty$, and hence under $1 \geq \alpha > 1/2$, we have

$$\sigma^2_{\bar{x}} = Var(\bar{x}_t) = \mu_v^2\left[N^{2\alpha-2}\right] + N^{-1}c_N + O(N^{\alpha-2}). \tag{8}$$

As pointed out earlier, in cases where $\alpha \leq 1/2$, the second term in the RHS of (8), that arises from the contribution of the idiosyncratic components, will be at least as important as the contribution of a weak factor, and using $Var(\bar{x}_t)$ we cannot identify α when it is less than 1/2. But in cases where $\alpha > 1/2$ a simple manipulation of (8) yields

$$\begin{aligned} 2(\alpha-1)\ln(N) &\approx \ln(\sigma^2_{\bar{x}}) - \ln\left(\mu_v^2\right) + \ln\left(1 - \frac{N^{-1}c_N}{\sigma^2_{\bar{x}}}\right) \\ &\approx \ln(\sigma^2_{\bar{x}}) - \ln\left(\mu_v^2\right) - \frac{N^{-1}c_N}{\sigma^2_{\bar{x}}}, \end{aligned}$$

or

$$\alpha \approx 1 + \frac{1}{2}\frac{\ln(\sigma^2_{\bar{x}})}{\ln(N)} - \frac{1}{2}\frac{\ln\left(\mu_v^2\right)}{\ln(N)} - \frac{c_N}{2\left[N\ln(N)\right]\sigma^2_{\bar{x}}}. \tag{9}$$

Note that the fourth term on the RHS of (9) is of smaller order of magnitude than the previous three terms and can be ignored, and α can be identified from (9) using a consistent estimator of $\sigma^2_{\bar{x}}$, given by

$$\hat{\sigma}^2_{\bar{x}} = \frac{1}{T}\sum_{t=1}^{T}\left(\bar{x}_t - \bar{x}\right)^2, \tag{10}$$

where $\bar{x} = T^{-1}\sum_{t=1}^{T}\bar{x}_t$. Ignoring terms that eventually vanish as $N \to \infty$, we obtain the following initial estimator of α

$$\hat{\alpha} = 1 + \frac{1}{2}\frac{\ln(\hat{\sigma}^2_{\bar{x}})}{\ln(N)}, \tag{11}$$

which is consistent and has a rate of convergence that is $\ln(N)^{-1}$. It is important that the estimator of α also allows for the third term in (9). This can be achieved by replacing μ_v^2 with a suitable estimator. There are many alternatives for this estimation which are discussed in detail in the next section. We denote the estimator of μ_v^2 by $\hat{\mu}_v^2$.

Next, we discuss correcting the bias arising from the final term in (9). This is easily achieved in the case of exact factor models where the idiosyncratic errors are cross-sectionally independent, and Σ_u is a diagonal matrix. In this case a consistent estimator of c_N is given by

$$\hat{c}_N = N^{-1}\sum_{i=1}^{N}\hat{\sigma}_i^2 = \widehat{\bar{\sigma}^2_N}, \tag{12}$$

where σ_i^2 is the i^{th} diagonal term of Σ_u, $\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{it}^2$, $\hat{u}_{it} = x_{it} - \hat{\delta}_i\bar{x}_t$, and $\hat{\delta}_i$ denotes the OLS estimator of the regression coefficient of x_{it} on $\bar{x}_t$. It is also useful, at this point, to introduce the notation $\bar{\sigma}^2_N = N^{-1}\sum_{i=1}^{N}\sigma_i^2$ to denote the population quantity corresponding to $\widehat{\bar{\sigma}^2_N}$. Note that if Σ_u is diagonal, $c_N = \bar{\sigma}^2_N$. Further, note that while $\hat{c}_N$, as an estimator for c_N, is motivated by appealing to an exact factor model, mild deviations from this model can be dealt with by using an alternative estimator for c_N, as discussed in Section 3.2. Using consistent estimators of $\sigma^2_{\bar{x}}$,

μ_v^2, and c_N, we propose the following bias-adjusted estimator

$$\mathring{\alpha} = \mathring{\alpha}\left(\hat{\mu}_v^2\right) = 1 + \frac{1}{2}\frac{\ln(\hat{\sigma}_{\bar{x}}^2)}{\ln(N)} - \frac{\ln\left(\hat{\mu}_v^2\right)}{2\ln\left(N\right)} - \frac{\hat{c}_N}{2\left[N\ln(N)\right]\hat{\sigma}_{\bar{x}}^2}. \tag{13}$$

3 Theoretical Derivations

3.1 Main Results

Consider now the following multiple factor generalisation of our basic setup:

$$x_{it} = \sum_{j=1}^{m} \beta_{ij} f_{jt} + u_{it} = \boldsymbol{\beta}_i' \boldsymbol{f}_t + u_{it},\ i = 1, 2, ..., N,$$

where $\boldsymbol{f}_t = (f_{1t}, f_{2t}, ..., f_{mt})'$ is an $m \times 1$ vector of unobserved factors, and $\boldsymbol{\beta}_i$ is the associated vector of factor loadings (m is fixed). Our first set of theoretical results characterise the asymptotic behaviour of $\hat{\alpha}$. We make the following assumptions.

Assumption 1 *The factor loadings are given by*

$$\begin{aligned} \beta_{ij} &= v_{ij} \text{ for } i = 1, 2, ..., \left[N^{\alpha_j}\right], \\ \beta_{ij} &= \tilde{v}_{ij}, \text{ for } i = \left[N^{\alpha_j}\right] + 1, \left[N^{\alpha_j}\right] + 2, ..., N, \end{aligned} \tag{14}$$

where $\alpha_1 > 1/2$, $0 \leq \alpha_j \leq 1$ and $\alpha_1 \geq \alpha_j$, $j = 2, ..., m$. Also, $\{v_{ij}\}_{i=1}^{\left[N^{\alpha_j}\right]}$ and $\{\tilde{v}_{ij}\}_{i=\left[N^{\alpha_j}\right]+1}^{N}$ are IID sequences of random variables for all $j = 1, 2, ..., m$. The former sequences have a non-zero mean, $\mu_{v_j} \neq 0$, and a finite variance $\sigma_{v_j}^2 < \infty$. The latter sequences are summable such that $\kappa_j = \sum_{i=\left[N^{\alpha_j}\right]+1}^{N} \tilde{v}_{ij} = O_p(1)$ has a finite mean, μ_{κ_j}, and a finite variance, $\sigma_{\kappa_j}^2$, for all j and N.

Assumption 2 *The $m \times 1$ vector of factors, $\boldsymbol{f}_t$, follows a linear stationary process given by*

$$\boldsymbol{f}_t = \sum_{j=0}^{\infty} \boldsymbol{\psi}_{fj} \boldsymbol{\nu}_{f,t-j}, \tag{15}$$

where $\boldsymbol{\nu}_{ft}$ is a sequence of IID random variables with mean zero and a finite variance matrix, $\boldsymbol{\Sigma}_{\nu_f}$, and uniformly finite φ-th moments for some $\varphi > 4$. The matrix coefficients, $\boldsymbol{\psi}_{fj}$, satisfy the absolute summability condition

$$\sum_{j=0}^{\infty} j^{\zeta} \left\|\boldsymbol{\psi}_{fj}\right\| < \infty,$$

such that $\{\zeta(\varphi - 2)\}/\{2(\varphi - 1)\} \geq 1/2$. $\boldsymbol{f}_t$ is distributed independently of the idiosyncratic errors, $u_{it'}$, for all i, t and t'.

Assumption 3 *For each i, u_{it} follows a linear stationary process given by*

$$u_{it} = \sum_{j=0}^{\infty} \psi_{ij} \nu_{i,t-j}, \tag{16}$$

where ν_{it}, $i = ..., -1, 0, ...$, $t = 0, ...$, is a double sequence of IID random variables with mean zero and uniformly finite variances, $\sigma^2_{\nu_i}$ and uniformly finite φ-th moments for some $\varphi > 4$. We assume that

$$\sup_i \sum_{j=0}^{\infty} j^{\zeta} |\psi_{ij}| < \infty, \tag{17}$$

such that $\{\zeta(\varphi - 2)\}/\{2(\varphi - 1)\} \geq 1/2$.

Assumptions 2 and 3 are mostly straightforward specifications of the factor and error processes assuming a linear structure with sufficient restrictions to enable the use of central limit theorems. Note that Assumption 3 rules out the existence cross-sectional dependence in the error terms. This condition will be relaxed in Section 3.2.

First, note that

$$\bar{\beta}_{jN} = N^{-1} \sum_{i=1}^{N} \beta_{ij} = \frac{[N^{\alpha_j}]}{N} \left(\frac{\sum_{i=1}^{[N^{\alpha_j}]} v_{ij}}{[N^{\alpha_j}]} \right) + \frac{\sum_{i=[N^{\alpha_j}]+1}^{N} \tilde{v}_{ij}}{N} = N^{\alpha_j - 1} \bar{v}_{jN} + O_p(N^{-1}) \tag{18}$$

and

$$Var(\bar{\beta}_{jN}) = \frac{[N^{\alpha_j}]}{N^2} \sigma^2_{v_j} + O(N^{-2}) = O(N^{a_j - 2}).$$

Consider now $\bar{x}_t - E(\bar{x}_t) = \bar{\beta}_{1N} f_{1t} + \bar{\beta}_{2N} f_{2t} + ... \bar{\beta}_{mN} f_{mt} + \bar{u}_t$, and, without loss of generality, recall that $\alpha =: \alpha_1 \geq \alpha_j$, $j = 2, ..., m$, and that the factors are orthogonal. Then,

$$\begin{aligned} Var(\bar{x}_t) &= \sum_{j=1}^{m} E(\bar{\beta}^2_{jN}) + E(\bar{u}^2_t) \\ &= \sum_{j=1}^{m} \left[E(\bar{\beta}_{jN}) \right]^2 + \sum_{j=1}^{m} Var(\bar{\beta}_{jN}) + E(\bar{u}^2_t), \end{aligned}$$

and, as in the single factor case, we have $Var(\bar{x}_t) = O(N^{2a-2}) + O(N^{-1})$, namely the order of $Var(\bar{x}_t)$ is dominated by the factor with the largest exponent of cross-sectional dependence, assuming that $\alpha > 1/2$. We also note that

$$\bar{\boldsymbol{\beta}}_N = N^{\alpha - 1} \mathbf{D}_N \bar{\boldsymbol{v}}_N + O_p(N^{-1}), \tag{19}$$

where $\bar{\boldsymbol{\beta}}_N = \left(\bar{\beta}_{1N}, ..., \bar{\beta}_{mN} \right)'$, $\bar{\boldsymbol{v}}_N = (\bar{v}_{1N}, ..., \bar{v}_{mN})'$, and $\mathbf{D}_N$ is an $m \times m$ diagonal matrix with diagonal elements given by $N^{\alpha_j - \alpha}$, and set

$$d_T = \bar{\boldsymbol{v}}'_N \boldsymbol{S}^{-1/2}_{ff} \bar{\boldsymbol{f}}_T - \boldsymbol{\mu}'_v \Sigma^{-1/2}_{ff} \boldsymbol{\mu}_f, \tag{20}$$

where $\boldsymbol{S}_{ff} = (s_{jo,f}) = \frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}_T\right)\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}_T\right)'$, $j, o = 1, ..., m$, $\bar{\boldsymbol{f}}_T = T^{-1}\sum_{t=1}^{T}\boldsymbol{f}_t$, $\boldsymbol{\Sigma}_{ff} = diag\left(\sigma_{f_j}^2\right) = I$, $\boldsymbol{\mu}_f = E\left(\boldsymbol{f}_t\right) = (\mu_{f_1}, ..., \mu_{f_m})'$, and $\boldsymbol{\mu}_v = \left(E\left(\boldsymbol{v}_j\right)\right) = (\mu_{v_1}, ..., \mu_{v_m})'$, $\boldsymbol{v}_j = \left(v_{1j}, ..., v_{[N^{\alpha_j}]j}\right)'$. Further, define $\mu_v^2 = \sum_{j=1}^{m}\mu_{v_j}^2$. Our informal exposition in Section 2 suggests that $\hat{\alpha}$, as an estimator of α, is subject to two sources of bias, $\frac{\ln\left(\mu_v^2\right)}{2\ln(N)}$ and $\frac{c_N}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}$, where the latter bias corresponds to the last part of (13) in the multiple factor case. This can be corrected using a first order accurate estimator given by $\frac{\hat{c}_N}{N\hat{\sigma}_{\bar{x}}^2}$ or a second order bias correction given by $\frac{\hat{c}_N}{\ln(N)N\hat{\sigma}_{\bar{x}}^2}\left(1 + \frac{\hat{c}_N}{N\hat{\sigma}_{\bar{x}}^2}\right)$, where $\hat{c}_N$ is defined in (12). We denote the estimators that make use of these corrections by

$$\tilde{\alpha} = \hat{\alpha} - \frac{\hat{c}_N}{2\ln(N)N\hat{\sigma}_{\bar{x}}^2}$$

and

$$\breve{\alpha} = \hat{\alpha} - \frac{\hat{c}_N}{2\ln(N)N\hat{\sigma}_{\bar{x}}^2}\left(1 + \frac{\hat{c}_N}{N\hat{\sigma}_{\bar{x}}^2}\right)$$

We now introduce the main theorem of the paper.

Theorem 1 *(a) Suppose Assumptions 1 to 3 hold,* $\alpha = \alpha_1 = \alpha_2 = ... = \alpha_m > 1/2$*. Then,*

$$\sqrt{\min(N^{\alpha^*}, T)}\left(2\ln(N)\left(\hat{\alpha} - \alpha^*\right) - \frac{c_N}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right) \rightarrow_d N\left(0, \omega_m\right) \tag{21}$$

where

$$\omega_m = \lim_{N,T\rightarrow\infty}\min\left(N^{\alpha}, T\right)Var(d_T^2),$$

d_T *is defined by (20),*

$$\alpha^* \equiv \alpha_N^* = \alpha + \frac{\ln\left(\mu_v^2\right)}{2\ln\left(N\right)},$$

and $\mu_v^2 = \sum_{j=1}^{m}\mu_{v_j}^2$.

(b) Continue to assume that $\alpha = \alpha_1 = \alpha_2 = ... = \alpha_m > 1/2$, *and suppose that either* $\frac{T^{1/2}}{N^{4\alpha-2}} \rightarrow 0$ *or* $\alpha > 4/7$, *then*

$$\sqrt{\min(N^{\alpha^*}, T)}2\ln(N)\left(\tilde{\alpha} - \alpha^*\right) \rightarrow_d N\left(0, \omega_m\right), \tag{22}$$

and

$$\sqrt{\min(N^{\alpha^*}, T)}2\ln(N)\left(\breve{\alpha} - \alpha^*\right) \rightarrow_d N\left(0, \omega_m\right). \tag{23}$$

(c) Further, if either

$$\alpha = \alpha_1 > \alpha_2 + 1/4, \tag{24}$$

or if

$$\alpha_2 < 3\alpha/4, \quad T^b = N, \quad b > \frac{1}{4(\alpha - \alpha_2)}, \tag{25}$$

and $\alpha_2 \geq \alpha_3 \geq ... \geq \alpha_m \geq 0$, *(21), (22) and (23) hold with* ω *replacing* ω_m*, where*

$$\omega = \lim_{N,T\to\infty} \left[\frac{\min(N^{\alpha}, T)}{T} V_{\overline{f_1^2}} + \frac{\min(N^{\alpha}, T)}{N^{\alpha}} \frac{4\sigma_{v_1}^2}{\mu_{v_1}^2} \right], \tag{26}$$

$$V_{\overline{f_1^2}} = Var\left(\tilde{f}_{1t}^2 \right) + 2\sum_{i=1}^{\infty} Cov\left(\tilde{f}_{1t}^2, \tilde{f}_{1t-i}^2 \right),$$

and $\tilde{f}_{1t} = (f_{1t} - \mu_{f_1})/\sigma_{f_1}$*, but* α^* *is now defined by*

$$\alpha^* \equiv \alpha_N^* = \alpha + \frac{\ln\left(\mu_{v_1}^2\right)}{2\ln(N)}. \tag{27}$$

(d) Finally, if $\alpha = \alpha_1 > \alpha_2 \geq \alpha_3 ... \geq \alpha_m \geq 0$ *but neither (24) or (25) hold, then (21), (22) and (23) hold with* ω *replacing* ω_m*, and*

$$\alpha^* \equiv \alpha_N^* = \alpha + \frac{\ln\left(\sum_{j=1}^{m} N^{2(\alpha_j - \alpha)}\mu_{v_j}^2\right)}{2\ln(N)}.$$

The above result gives a full distribution theory but it is not operational in practice since μ_v^2 is not known. So next, we consider the third term of (9) which depends on μ_v^2. While noting that the value of μ_v^2 is irrelevant for the probability limit of $\hat{\alpha}$, in small samples it is an important determinant of cross-sectional dependence. Hence, correcting for this bias provides us with a refined estimator of α that is likely to have better small sample properties. The first step towards deriving an estimator for μ_v^2 is to note that μ_v is the mean of the population regression coefficient of x_{it} on $\tilde{x}_t = \bar{x}_t/\hat{\sigma}_{\bar{x}}$ for units x_{it} that have at least one non-zero factor loading. Therefore, once we identify which units have non-zero loadings, an estimate of μ_v can be obtained by the average covariance between x_{it} and $\tilde{x}_t$ over $i = 1, 2, ..., \left[N^{\hat{\alpha}}\right]$. While there are many ways to identify which units have non-zero loadings, a multiple testing approach to this problem seems appropriate, considering that we are interested in μ_v as $N \to \infty$. This estimate is equivalent to the one given by the standard deviation of the cross-sectional average of the units that have non-zero loadings. We prefer the latter estimator, due to its simplicity, and propose the following procedure:

1. *Run the OLS regression of* x_{it} *on a constant and* $\bar{x}_t$ *and denote the estimated coefficient of* $\bar{x}_t$ *by* $\hat{\delta}_i$*, for* $i = 1, 2, ..., N$.

2. *Compute the t-ratio associated with the* i^{th} *coefficient,* $\hat{\delta}_i$*,* $i = 1, 2, ..., N$*, as* $z_{\hat{\delta}_i} = \hat{\delta}_i / s.e.\left(\hat{\delta}_i\right)$.

3. *Construct*

$$\bar{x}_t(\mathbf{c}_p) = \frac{\sum_{i=1}^{N} x_{it} I\left(\left| z_{\hat{\delta}_i} \right| \geq c_{p_i,N} \right)}{\sum_{i=1}^{N} I\left(\left| z_{\hat{\delta}_i} \right| \geq c_{p_i,N} \right)},$$

where $c_{p_i,N}$ *is the critical value of the i-th test that depends on* N *as well as the overall nominal size of the test, which we denote by* p*, and* $\mathbf{c}_p = (c_{p_1,N}, c_{p_2,N}, ..., c_{p_N,N})'$.

4. *Estimate* μ_v *by*

$$\hat{\mu}_v = \hat{\mu}_v(\mathbf{c}_p) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left[\bar{x}_t(\mathbf{c}_p) - \bar{x}(\mathbf{c}_p)\right]^2},$$

where $\bar{x}(\mathbf{c}_p) = T^{-1}\sum_{t=1}^{T}\bar{x}_t(\mathbf{c}_p)$.

The critical values, $c_{p_i,N}$, can be set using the multiple testing approaches of Bonferroni (Bonferroni (1935), Bonferroni (1936)) or Holm (Holm (1979)). Both approaches deal with the multiple testing problem without making any assumptions about the cross dependence of the underlying N individual t tests.[4] But Holm's approach is less conservative and uses different critical values across the units. To be more specific let $t_i = \left|z_{\hat{\delta}_i}\right|$, for $i = 1, 2, ..., N$, and sort these t-ratios in a descending order, such that $t_{(1)} > t_{(2)} > .... > t_{(N)}$, with associated critical values, $c_{p_{(i)},N}$. Suppose also that under the null hypothesis $\beta_{i1} = 0$, $z_{\hat{\delta}_i}$ is asymptotically distributed as $N(0, 1)$, with the cumulative distribution function $\Phi(.)$. Then under Bonferroni's approach $c_{p_{(i)},N} = \Phi^{-1}\left(1 - \frac{p}{2N}\right)$ which is the same for all units, whilst under Holm's approach $c_{p_{(i)},N} = \Phi^{-1}\left(1 - \frac{p}{2(N-i)}\right)$ corresponding to $t_{(i)}$.

Note that in this paper we focus more on the case when $\alpha = \alpha_1 > \alpha_2 \geq ... \geq \alpha_m$ which we consider to be more realistic than the case of $\alpha = \alpha_j$, $j = 1, ..., m$. In the supplementary Appendix V we consider the conditions under which $\hat{\mu}_v^2$ can be a consistent estimator of the population quantity of $\mu_{v_1}^2$. In particular, it is shown that $\hat{\mu}_v^2$, computed using Bonferroni or Holm procedures, is a consistent estimator of $\mu_{v_1}^2$ if $\alpha > 2/3$ and $\alpha = \alpha_1 > \alpha_2 \geq ... \geq \alpha_m$. The supplement also provides more general conditions on the choice of $c_{p_i,N}$, and shows that the critical values used in Bonferroni and Holm approaches satisfy these conditions (see (B42) and (B43) in the Supplementary Appendix V). In the simulation section we study a two factor setting where $\alpha = \alpha_1 > \alpha_2$ and use both Bonferroni and Holm procedures. We find that Holm's method performs better uniformly across all experiments. Therefore, all the results reported are based on the Holm approach for $\alpha = \alpha_1 > \alpha_2$. Monte Carlo results for $\alpha = \alpha_j$, $j = 1, ..., m$ are available in the Supplementary Appendix VI.

3.2 Extensions

In this section we consider two extensions to our main analysis. For simplicity of the treatment we discuss these in the context of a single factor model but the extension to multiple factors is straightforward. First, we relax Assumption 3 and allow the error terms to be cross-sectionally weakly dependent. Accordingly, we modify Assumption 3 as follows:

[4] For a recent review of this literature see Efron (2010).

Assumption 4 *For each i, u_{it} follows a linear stationary process given by*

$$u_{it} = \sum_{j=0}^{\infty} \psi_{ij} \left(\sum_{s\,=\,-\infty}^{\infty} \xi_{is} \nu_{s,t-j} \right), \tag{28}$$

where ν_{it}, $i = ..., -1, 0, ...,$ $t = 0, ...,$ is a double sequence of IID random variables with mean zero and uniformly finite variances, $\sigma^2_{\nu_i}$, and uniformly finite φ-th moments for some $\varphi > 4$. We assume that

$$\sup_i \sum_{j=0}^{\infty} j^{\zeta} \left|\psi_{ij}\right| < \infty,$$

and

$$\sup_i \sum_{s=-\infty}^{\infty} |s|^{\zeta} \left|\xi_{is}\right| < \infty, \tag{29}$$

such that $\{\zeta(\varphi - 2)\}/\{2(\varphi - 1)\} \geq 1/2$.

Under the above assumption Σ_u is no longer a diagonal matrix. When $\alpha > 2/3$ the bias term in (21) is $o_p(1)$ and, as a result, c_N can still be estimated by $\widehat{\bar{\sigma}^2_N}$, to construct the various estimators of α. However, in the case where $1/2 < \alpha \leq 2/3$, an alternative estimator for c_N is needed to take account of the non-zero covariances between u_{it} and u_{jt}. One possibility is to use the following estimator

$$\tilde{c}_N = T^{-1} \sum_{t=1}^{T} \left(\sqrt{N} \bar{\hat{e}}_t - \sqrt{N} \bar{\hat{e}} \right)^2, \tag{30}$$

where

$$\bar{\hat{e}}_t = N^{-1} \sum_{i=1}^{N} \hat{e}_{it}, \text{ and } \bar{\hat{e}} = T^{-1} \sum_{t=1}^{T} \bar{\hat{e}}_t, \tag{31}$$

and $\hat{e}_{it} = x_{it} - \hat{\varrho}_i \widehat{pc}_t$, $\widehat{pc}_t$ is the first principal component of x_{it}, $i = 1, ..., N$, and $\hat{\varrho}_i$ denotes the OLS estimator of the regression coefficient of x_{it} on $\widehat{pc}_t$. The use of cross section averages, $\bar{x}_t$, in place of $\widehat{pc}_t$ to compute $\hat{e}_{it}$ does not help in estimation of c_N since $\sum_{i=1}^{N} \left(x_{it} - \hat{\delta}_i \bar{x}_t \right) = 0$, where $\hat{\delta}_i$ is the OLS slope coefficient in the regression of x_{it} on $\bar{x}_t$, and suggests setting $\tilde{c}_N$ to zero. In a multiple factor setting additional principal components are needed to filter out any remaining cross-sectional error dependencies.

Up to now we have analysed estimators of the exponent of cross-sectional dependence assuming that factor loadings take the form given in Assumption 1. We briefly examine an alternative formulation (discussed in the Introduction) which is mathematically convenient, although it is more difficult to justify from an economic perspective as it assumes that all factor loadings fall at the same rate. More specifically consider the following alternative formulation:

Assumption 5 *Suppose that the factor loadings vary uniformly with N as in*

$$\beta_{i1} = N^{(\alpha-1)/2} v_{i1}, \quad 0 < \alpha \leq 1 \tag{32}$$

where $\{v_{i1}\}_{i=1}^{N}$ is an i.i.d. sequence of random variables with mean $\mu_{v_1} \neq 0$, and variance $\sigma_{v_1}^2 < \infty$. Then,

$$\sum_{i=1}^{N} \sum_{j \neq i, j=1}^{N} \sigma_{ij,x} = O(N^{1+\alpha}), \quad N^{-1}\lambda_{\max}\left(\mathbf{\Sigma}_N\right) = O\left(N^{\alpha-1}\right), \; N^{-1}\left\|\mathbf{\Sigma}_N\right\|_1 = O\left(N^{\alpha-1}\right), \; Var\left(\bar{x}_t\right) = O\left(N^{\alpha-1}\right).$$

For this setup it is easy to show that the appropriate estimator for α is given by

$$\hat{\alpha} = 1 + \frac{\ln(\hat{\sigma}_{\bar{x}}^2)}{\ln(N)}, \tag{33}$$

and its first bias-corrected version is given by

$$\tilde{\alpha} = \hat{\alpha} - \frac{\hat{c}_N}{\ln(N) N \hat{\sigma}_{\bar{x}}^2}. \tag{34}$$

In this case of the alternative formulation, (32), there is no need for further bias-corrections. Then, the next Corollary follows (a proof is provided in the Supplementary Appendix II):

Corollary 1 *Let Assumptions 5 and 2-3 hold, $m = 1$. Let $\hat{\alpha}$ be defined as in (33). Then,*

$$\sqrt{\min(N,T)}\left(2\ln(N)\left(\hat{\alpha} - \alpha^*\right) - \frac{\bar{\sigma}_N^2}{N^{\alpha}\bar{v}_{1N}^2 s_{f_1}^2}\right) \rightarrow_d N\left(0, \omega\right),$$

where α^* and ω are defined in (27) and (26), respectively and $s_{f_1}^2 = T^{-1}\sum_{t=1}^{T}\left(f_{1t} - T^{-1}\sum_{t=1}^{T} f_{1t}\right)^2$. Further, let $\tilde{\alpha}$ be defined as in (34)

$$2\sqrt{\min(N,T)}\ln(N)\left(\tilde{\alpha} - \alpha^*\right) \rightarrow_d N\left(0, \omega\right).$$

Remark 1 *It is of interest to consider circumstances where Assumption 5 fails but the above result still holds. In particular, let*

$$\beta_{i1} = N^{(\alpha-1)/2} v_{Ni}, \quad 0 < \alpha \leq 1 \tag{35}$$

where $v_{Ni} = \breve{v}_i + \zeta_{Ni}$ and $\{\breve{v}_i\}_{i=1}^{N}$ is an i.i.d. sequence of random variables with mean $\mu_{\breve{v}} \neq 0$, and variance $\sigma_{\breve{v}}^2 < \infty$. Lemma 14 provides general conditions for this Assumption, under which our theoretical results hold. In this remark we explore a leading case of departure from Assumption 5 that is covered by Lemma 14. Without loss of generality, we order the cross section units such that $\zeta_{Ni} = N^{(1-\alpha)/2}\eta_i$ for $i = 1, 2, ..., M$, where $\{\eta_i\}_{i=1}^{N}$ is an i.i.d. sequence of random variables with mean $\mu_\eta \neq 0$, and variance $\sigma_\eta^2 < \infty$. This implies that M units have loadings that are *bounded away from zero. Then, using Lemma 14, it is easy to see that the theorems relating to the asymptotic distribution of the estimators continue to hold as long as $M = o(N^{\alpha})$.*

4 Monte Carlo Study

We investigate the small sample properties of the proposed estimator of α through a detailed simulation study. We consider the following two factor model

$$x_{it} = d_i + \beta_{i1} f_{1t} + \beta_{i2} f_{2t} + \varsigma \sigma_i u_{it}, \tag{36}$$

for $i = 1, 2, ..., N$, and $t = 1, 2, ..., T$. We generate the intercepts as $d_i \sim IIDN(0, 1)$, $i = 1, 2, ..., N$. The factors are generated as

$$f_{jt} = \rho_j f_{j,t-1} + \sqrt{1 - \rho_j^2} \zeta_{jt}, \ j = 1, 2, \text{ for } t = -49, -48, ..., 0, 1, ..., T,$$

with $f_{j,-50} = 0$, for $j = 1, 2$, and $\zeta_{jt} \sim IIDN(0, 1)$. Therefore, by construction $\sigma^2_{f_j} = 1$, for $j = 1, 2$.

The shocks follow an AR(1) process:

$u_{it} = \phi_i u_{i,t-1} + \sqrt{1 - \phi_i^2} \varepsilon_{it}$, for $i = 1, 2, ..., N$ and $t = -49, -48, ..., 0, 1, ..., T$, with $u_{i,-50} = 0$, $\varepsilon_{it} \sim IIDN(0, 1)$, $i = 1, 2, ..., N$

where $\phi_i \sim IIU(0, 1)$ and $\sigma_i^2 \sim IID\left(\frac{1}{2} + \frac{3\chi^2(2)}{4}\right)$, $i = 1, 2, ..., N$, ensuring that all σ_i^2 are bounded away from zero. Also, $\bar{\sigma}_N^2 = N^{-1} \sum_{i=1}^{N} \sigma_i^2 \to 2$, as $N \to \infty$.

With regard to the factor loadings, we generate them as follows:

$$\begin{aligned}
&\beta_{i1} = v_{i1}, \text{ for } i = 1, 2, ..., [N^{\alpha_1}] \\
&\beta_{i1} = 0, \text{ for } i = [N^{\alpha_1}] + 1, [N^{\alpha_1}] + 2, ..., N \\
&\beta_{i2} = v_{i2}, \text{ for } i = 1, 2, ..., [N^{\alpha_2}], \\
&\beta_{i2} = 0, \text{ for } i = [N^{\alpha_2}] + 1, [N^{\alpha_2}] + 1, ..., N,
\end{aligned}$$

where β_{i2} are then randomised across N to achieve independence from β_{i1}. The loadings are generated as $v_{ij} \sim IIDU(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, for $j = 1, 2$. We examine the case where $\alpha_2 < \alpha_1 = a$ and consider values of α and α_2 such that $\alpha_2 = \frac{2\alpha}{3}$ to reflect the more realistic scenario where the two factors have different strengths. Further, we set $\mu_{v_2} = 0.71$ and $\mu_{v_1} = \sqrt{\mu_v^2 - N^{2(\alpha_2 - \alpha)} \mu_{v_2}^2}$ - see Theorem 1 (d) -, yielding $\mu_{v_1}^2 + \mu_{v_2}^2 = \mu_v^2 = 0.75$. Both μ_{v_1} and μ_{v_2} are picked so that they meet the condition that $\mu_{v_j} \neq 0$, $j = 1, 2$ without $\mu'_{v_j} s$ being too distant from zero either.[5]

[5] Other values of μ_{ν_j}, $j = 1, 2$ have been entertained. Also, $\beta_{ij} = 0$, for $i > [N^{\alpha_j}]$, $j = 1, 2$ are set for simplicity. The case of $\beta_{ij} = \rho_l^{i - [N^{\alpha_j}]}$, for $i > [N^{\alpha}]$, $j = 1, 2$ and $\rho_l = 0.5$ has been considered as well as an example of

In fixing the remaining parameters we calibrate the fit of each cross section unit, as measured by R_i^2, in order to achieve an average fit across all the units of around $\bar{R}_N^2 = N^{-1}\sum_{i=1}^N R_i^2 \approx 0.40$, an average figure one obtains in most large data sets used in macroeconomics and finance.[6] To this end we note that

$$R_i^2 = \frac{\beta_{i1}^2 + \beta_{i2}^2}{\beta_{i1}^2 + \beta_{i2}^2 + \sigma_i^2} = \frac{\psi_{i1}^2 + \psi_{i2}^2}{1 + \psi_{i1}^2 + \psi_{i2}^2}, \text{ if for the } i^{th} \text{ unit: both } \beta_{i1} \neq 0 \text{ and } \beta_{i2} \neq 0,$$

where $\psi_{ij}^2 = \beta_{ij}^2/\sigma_i^2$, for $j = 1, 2$. Similarly,

$$R_i^2 = \frac{\psi_{i1}^2}{1 + \psi_{i1}^2}, \text{ if for the } i^{th} \text{ unit: } \beta_{i1} \neq 0 \text{ but } \beta_{i2} = 0,$$

$$R_i^2 = \frac{\psi_{i2}^2}{1 + \psi_{i2}^2}, \text{ if for the } i^{th} \text{ unit: } \beta_{i2} \neq 0 \text{ but } \beta_{i1} = 0,$$

and

$$R_i^2 = 0, \text{ if for the } i^{th} \text{ unit: both } \beta_{i1} = 0 \text{ and } \beta_{i2} = 0.$$

The calibration of $\bar{R}_N^2$ is done by scaling of σ_i^2 in (36) using $\varsigma = 1/2$.

Experiment A Here we use a basic design of (36) where the factors, f_{jt}, for $j = 1, 2$, are serially uncorrelated, namely we set $\rho_j = 0.0$ for $j = 1, 2$.

Experiment B Under this experiment we use the same design as in Experiment A, but allow for temporal dependence in the factors, namely we set $\rho_j = 0.5$ for $j = 1, 2$.

Experiment C Under this experiment we use the same design as in Experiment A, but we allow for departure of the idiosyncratic errors from normality and generate the idiosyncratic errors as $\varepsilon_{it} \sim IID((\chi^2(2) - 2)/4)$, $i = 1, 2, ..., N$.

Experiment D The design for this experiment is as in Experiment A, but allows the errors, u_{it}, to be cross-sectionally dependent according to a first order spatial autoregressive model. Let $\mathbf{u}_t = (u_{1t}, u_{2t}, ..., u_{Nt})'$, and set $\mathbf{u}_t$ as

$$\mathbf{u}_t = \mathbf{Q}\varepsilon_t,\ \varepsilon_t = \sigma_\varepsilon \eta_t; \quad \eta_t \sim IIDN(\mathbf{0}, \mathbf{I}_N),$$

$\sum_{i=[N^{\alpha_j}]+1}^N \beta_{ij} = O_p(1)$, $j = 1, 2$. Results for these setups are available upon request.

[6] We calibrated R_N^2 from a number of data sets, some of which are used in our empirical applications. Details can be found in the Supplementary Appendix VI.

where $\mathbf{Q} = (\mathbf{I}_N - \theta\mathbf{S})^{-1}$, and

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1/2 & 0 & 1/2 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & 0 & 1/2 \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

We set $\theta = 0.2$, and $\sigma_\varepsilon^2 = N/Tr(\mathbf{QQ}')$ which ensures that $N^{-1}\sum_{i=1}^N var(u_{it}) = 1$.

For all experiments we consider the values of $\alpha = 0.70, 0.75, ..., 0.90, 0.95, 1.00$, $N = 100, 200, 500, 1000$ and $T = 100, 200, 500$, and base them on $2,000$ replications. For each replication, the values of α, α_2, d_i, ρ_j, ϕ_i, ς and $\mathbf{S}$ are given as set out above. These parameters are fixed across all replications. The values of v_{ij}, $j = 1, 2$ are drawn randomly (N of them) for each replication.

In all experiments we present bias and RMSE results for the bias-adjusted estimator $\mathring{\alpha}$ given by (13), where μ_{v_1} is estimated using the Holm approach to address the associated multiple testing problem. For experiments A-C we use $\hat{c}_N$ given by (12) to estimate c_N while for experiment D we use $\tilde{c}_N$, given by (30). All results are scaled up by 100.

4.1 Summary of the results

The results for Experiment A are summarized on the left-hand-side panel of Table A-B, giving the bias and Root Mean Square Error (RMSE) when $\mathring{\alpha}$ is used as the estimator for α, and when setting $\mu_v = 0.75$ and $\alpha_2 = 2\alpha/3$. We focus on the bias-corrected estimator, $\mathring{\alpha}$, which can be used for any value of $\mu_{v_k} \neq 0$, and we only report results for values of α over the range $[0.70, 1.0]$. Recall that α is identified only if $\alpha > 1/2$. As predicted by the theory, the bias and RMSE of $\mathring{\alpha}$ decline with both N and T, and tend to be somewhat smaller for larger values of α, especially as T rises. In the Supplementary Appendix VI we show additional results relating to Experiment A. First, we report bias, RMSE, size and power of estimator $\tilde{\alpha}$ when setting $\mu_v = 1$. The asymptotic distribution of $\tilde{\alpha}$ is derived in Theorem 1 and estimation of the variance component is discussed in Section VI of the Supplementary Appendix. Second, we show size and power of tests based on $\mathring{\alpha}$. Finally, we consider the case when $\alpha = \alpha_2$. A discussion of the results for all variants of Experiment A can be found in Appendix VI.

The results for Experiment B, where the factors are allowed to be serially correlated, are summarized on right-hand-side panel of Table A-B. As compared to the baseline case, we see a marginal deterioration in the results, particularly for relatively small values of N, T and α. But these differences tend to vanish as N and T are increased.

The results of Experiment C, where the idiosyncratic errors, u_{it}, are allowed to be non-normal, are summarized on the left-hand-side panel of Table C-D. As can be seen, the results are slightly affected by the non-normality of the error terms when α is relatively small. Consistent with the baseline case of Experiment A, both the bias and RMSE of $\mathring{\alpha}$ fall gradually as N, T and α are increased.

Finally, the effects of allowing for weak cross-sectional dependence in the idiosyncratic errors, u_{it}, on estimation of α are summarized on the right-hand-side panel of Table C-D for Experiment D. Considering the moderate nature of the spatial dependence introduced into the errors (with the spatial parameter, θ, set to 0.2), the results are not that different from the ones reported in Table A-B, for the baseline experiments.[7] However, one would expect greater distortions as θ is increased, although the effects of introducing weak dependence in the idiosyncratic errors are likely to be less pronounced if higher values of α are considered. For values of α near the borderline value of 1/2, it will become particularly difficult to distinguish between factor and spatial dependent structures.

The Monte Carlo results clearly illustrate the potential utility of the estimation and inferential procedure proposed in the paper for the analysis of cross-sectional dependence. The results are broadly in agreement with the theory and are reasonably robust to departures from the basic model assumptions. Although, the results tend to deteriorate slightly when we consider serially correlated factors or weak error cross-sectional dependence, the estimated values of α tend to retain a high degree of accuracy even for moderate sample sizes. It is also worth bearing in mind that in most empirical applications the interest will be on estimates of α that are close to unity, as it is for these values that a factor structure makes sense as compared to spatial or other network models of cross-sectional dependence. It is, therefore, helpful that the small sample performance of the proposed estimator improves when values of α close to unity are considered.

5 Empirical Applications

In this section we provide estimates of the exponent of cross-sectional dependence, α, for a number of panel data sets used extensively in economics and finance.[8] Specifically, we consider two types of data sets: quarterly cross-country data used in global modelling, and large quarterly data sets used in empirical factor model literature.We denote the typical elements of these data sets by y_{it}. The observations were standardized as $x_{it} = (y_{it} - \bar{y}_i)/s_i$, where $\bar{y}_i$ and s_i are the sample means and standard deviations of y_{it} for $t = 1, 2, ..., T$.

But before providing estimates of the exponent of cross-sectional dependence for these data sets we first need to verify that the degree of cross-dependence in these data sets is sufficiently large. Recall that α is identifiable only if $\alpha > 1/2$. To this end we first apply the recent test of weak Cross-Sectional Dependence (CD) developed by Pesaran (2013) to these data sets. The

[7] Note that in the estimation of $\tilde{c}_N$, given by (30), we use 2 principal components since we are focusing on a two factor model specification. In our empirical section we use 4 principal components instead as we consider these to be sufficient in order to absorb any additional cross-sectional dependence.

[8] In all empirical applications we use the Holm approach when implementing the procedure described on page 10. Results using the Bonferroni method are available upon request.

CD test statistic is defined by

$$CD_{NT} = \left[\frac{TN(N-1)}{2}\right]^{1/2} \widehat{\bar{\rho}}_N, \tag{37}$$

where

$$\widehat{\bar{\rho}}_N = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij},$$

and $\hat{\rho}_{ij}$ is the pair-wise correlation coefficient of x_{it} and x_{jt}. Pesaran (2013) shows that when $T = O(N^d)$ for some $0 < d \leq 1$, then the implicit null of the CD test is given by $0 \leq \alpha < (2-d)/4$, and it is asymptotically distributed as $N(0,1)$. In our applications N and T are of the same order of magnitude and $d \approx 1$.[9]

5.1 Cross-country dependence of macro-variables

We consider the cross-correlations of real output growth, inflation and rate of change of real equity prices over 33 countries (when available), over the period 1979Q2-2009Q4. These data sets are from Cesa-Bianchi et al. (2012) and update the earlier GVAR (global vector autoregressive) data sets used in Pesaran et al. (2004), and Dees et al. (2007).[10]

The CD statistics turned out to be 44.32, 88.34 and 77.78 for output growth, inflation and real equity prices, respectively, which are hugely statistically significant and reject the null hypothesis of weak cross-sectional dependence for all the three data sets and justify the use of our procedure for estimation of α. Table 1 presents the bias corrected estimates, $\mathring{\alpha}$, computed using available cross-country time series, x_{it}, over the period 1979Q2-2009Q4. Table 1 also reports the 90% confidence bands constructed following the procedure set out in the Supplementary Appendix VI. Although, there are 33 countries in the GVAR data set, not all variables are available for all the countries. For example, real equity prices are available only for 26 of the 33 countries.

Looking at the results of Table 1 for $\mathring{\alpha}$, we observe that the point estimates for all variables considered fall in a small range and indicate that approximately $1/7^{th}$ of the variables are cross-sectionally weakly correlated while the remaining ones belong to the strongly correlated group.[11] The exponent of cross-sectional dependence for real equity prices at 0.972 points to financial variables being strongly correlated. Similar estimates are also obtained for the macro variables. For real GDP growth and inflation we obtain the estimates 0.977 and 0.978,

[9] In all the empirical applications we present α estimates to be quite high. This alleviates an issue that arises when using the CD test in this context. The issue is that the CD test rejects when $\alpha > 1/4$ while our cross-sectional exponent estimator assumes that $1/2 < \alpha \leq 1$, and hence it is important that the rejection of the CD is not necessarily interpreted as evidence in favour of $\alpha > 1/2$. But in cases where CD test does not result in a rejection we could safely maintain that $\alpha \leq 1/2$, if N and T are of the same order of magnitude.

[10] This version of GVAR data set can be downloaded from http://www-cfap.jbs.cam.ac.uk/research/gvartoolbox/download.html

[11] Note that $\mathring{\alpha}$ corresponds to the most robust estimator of the exponent of cross-sectional dependence and corrects for both serial correlation in the factors and weak cross-sectional dependence in the error terms. We use four principal components when estimating (30).

respectively. The confidence bands all lie above 0.5 and do include unity (though marginally), suggesting that in these examples a factor structure might be a good approximation for modelling global dependencies. However, in some instances the value of $\alpha = 1$, 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-linkages that are more local or regional rather than global in nature.

Table 1: Exponent of cross-country dependence of macro-variables

	N	T	$\mathring{\alpha}^*_{0.05}$	$\mathring{\alpha}$	$\mathring{\alpha}^*_{0.95}$
Real GDP growth, q/q	33	122	0.923	0.977	1.031
Inflation, q/q	33	123	0.915	0.978	1.041
Real equity prices, q/q	26	122	0.924	0.972	1.019

*90% level confidence bands

5.2 Within-country dependence of macroeconomic variables

An important strand in the empirical factor literature, influenced by the theoretical and empirical work of Stock and Watson (2002), uses factor models to estimate and forecast a few key macro variables such as output growth, inflation or unemployment rate with a large number of macro-variables, that could exceed the number of available time periods. It is typically assumed that the macro variables satisfy a strong factor model with $\alpha = 1$. We estimated α using the quarterly data sets used in Eklund et al. (2010). For the US the data set comprises 95 variables and cover the period 1960Q2 to 2008Q3. For the UK the data set covers 94 variables spanning the period1977Q1 to 2008Q2.

As before, we first computed the CD statistic for the two data sets and obtained 84.72 and 54.29 for the US and UK, respectively, which are again highly significant and justify the use of our estimation procedure. The estimates of α together with their 90% confidence bands are summarized in Table 2.

For the US data set we obtained $\mathring{\alpha} = 0.946$ which suggests that more than $1/4^{th}$ of the variables considered can be regarded as being cross-sectionally weakly dependent, and the rest being strongly cross-correlated. For the UK data set we obtained $\mathring{\alpha} = 0.930$, slightly below the α estimate for the US. The 90% confidence bands for the US and UK data sets are well above the threshold value of 0.50, but fall short of unity routinely assumed in the literature.

Table 2: Exponent of within-country dependence of macro-variables

US			UK		
1960Q2-2008Q3			*1977Q1-2008Q2*		
N=95, T=194			*N=94, T=126*		
$\mathring{\alpha}^*_{0.05}$	$\mathring{\alpha}$	$\mathring{\alpha}^*_{0.95}$	$\mathring{\alpha}^*_{0.05}$	$\mathring{\alpha}$	$\mathring{\alpha}^*_{0.95}$
0.908	0.946	0.984	0.863	0.930	0.996

*90% level confidence bands

6 Conclusions

Cross-sectional dependence and the extent to which it occurs in large multivariate data sets is of great interest for a variety of economic, econometric and financial analyses. Such analyses vary widely. Examples include the effects of idiosyncratic shocks on aggregate macroeconomic variables, the extent to which financial risk can be diversified, and the performance of standard estimators such as principal components when applied to data sets where the cross sectional dependence might not be sufficiently strong.

In this paper we propose a relatively simple method of measuring the extent of interconnections in large panel data sets in terms of a single parameter that we refer to as the exponent of cross-sectional dependence. We find that this exponent can accommodate a wide spectrum of cross-sectional dependencies in macro and financial data sets. We propose consistent estimators of the cross-sectional exponent and derive their asymptotic distribution. The inference problem is complex, as it involves handling a variety of bias terms and, from an econometric point of view, has noteworthy characteristics such as nonstandard rates of convergence. We provide a feasible and relatively straightforward estimation and inference implementation strategy.

A detailed Monte Carlo study suggests that the estimated measure has desirable small sample properties. We apply our measure to two widely analysed classes of data sets. In all cases, we find that the results of the empirical analysis accord with prior intuition.

We conclude by pointing out some of the implications of our analysis for large N factor models of the type analysed by Bai and Ng (2002), Bai (2003), and Stock and Watson (2002). This literature assumes that all factors have the same cross-sectional exponent of $\alpha = 1$, which, as our empirical applications suggest, may be too restrictive, and it is important that implications of this assumption's failure are investigated. Chudik et al. (2011), Kapetanios and Marcellino (2010) and Onatski (2012) discuss some of these implications, namely that when $1/2 < \alpha < 1$, factor estimates are consistent but their rates of convergence are different (slower) as compared to the case where $\alpha = 1$, and in particular their asymptotic distributions may need to be modified. In cases where $\alpha < 1$, methods used to determine the number of factors in large data sets, discussed for example by Bai and Ng (2002), Onatski (2009) and Kapetanios (2010), are invalid and can select the wrong number of factors, even asymptotically.[12] Finally, the use of estimated factors in regressions for forecasting or other modelling purposes might not be justified under the conditions discussed in Bai and Ng (2006).

References

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

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

Bai, J. and S. Ng (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica 74*(4), 1133–1150.

Bai, Z. D. and J. W. Silverstein (1998). No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices. *Annals of Probability 26*(1), 316–345.

Bonferroni, C. (1935). Il calcolo delle assicurazioni su gruppi di teste. *Studi in Onore del Professore Salvatore Ortu Carboni, Rome, Italy*, 13–60.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8*, 3–62.

Cesa-Bianchi, A., M. H. Pesaran, A. Rebucci, and T. Xu (2012). China's emergence in the world economy and business cycles in latin america. *Economia, Journal of the Latin American and Caribbean Economic Association 12*(2), 1–75.

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

Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure and mean-variance analysis in large asset markets. *Econometrica 51*, 1305–1324.

Chudik, A., M. H. Pesaran, and E. Tosetti (2011). Weak and strong cross section dependence and estimation of large panels. *Econometrics Journal 14*(1), C45–C90.

Connor, G. and R. Korajczyk (1993). A test for the number of factors in an approximate factor model. *The Journal of Finance XLVIII*(4), 1263–1291.

Davidson, J. (1994). *Stochastic Limit Theory.* Oxford University Press.

Dees, S., F. di Mauro, M. H. Pesaran, and L. V. Smith (2007). Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics 22*(1), 1–38.

Efron, B. (2010). *Large-scale inference: Empirical Bayes methods for estimation, testing and prediction.* Cambridge.

Eklund, J., G. Kapetanios, and S. Price (2010). Forecasting in the presence of recent structural change. *Working Paper 406, Bank of England*.

[12] It is interesting to note that another contribution to this literature (Onatski (2010)) does not assume strong factors and, therefore, the suggested method will be valid in our framework. Also, Kapetanios and Marcellino (2010) suggest modifications to the methods of Bai and Ng (2002) that enable their use in the presence of weak factors.

Fan, J., Y. Liao, and M. Mincheva (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of Royal Statistical Society, Series B 75*(4), 1–44.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalised dynamic factor model: Identification and estimation. *Review of Economics and Statistics 82*(4), 540–554.

Forni, M. and M. Lippi (2001). The generalised dynamic factor model: Representation theory. *Econometric Theory 17*(6), 1113–1141.

Hachem, W., P. Loubaton, and J. Najim (2005a). The empirical eigenvalue distribution of a gram matrix: From independence to stationarity. *Markov Processes and Related Fields 11*(4), 629–648.

Hachem, W., P. Loubaton, and J. Najim (2005b). The empirical eigenvalue distribution of a gram matrix with a given variance profile. *Annales de l'Institut Henri Poincare 42*(6), 649–670.

Holm, H. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics 6*(2), 65–70.

Ibragimov, I. A. and Y. V. Linnik (1971). *Independent and stationary sequences of random variables.* Wolters-Noordhoff, Groningen.

Kapetanios, G. (2010). A testing procedure for determining the number of factors in approximate factor models with large datasets. *Journal of Business and Economic Statistics 28*(3), 397–409.

Kapetanios, G. and M. Marcellino (2010). Factor-GMM estimation with large sets of possibly weak instruments. *Computational Statistics and Data Analysis 54*(11), 2655–2675.

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

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *Review of Economics and Statistics 92*(4), 1004–1016.

Onatski, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *Journal of Econometrics 168*(2), 244–258.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica 74*(4), 967–1012.

Pesaran, M. H. (2013). Testing weak cross-sectional dependence in large panels. *Forthcoming in Econometric Reviews*.

Pesaran, M. H., T. Schuermann, and S. Weiner (2004). Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business and Economic Statistics 22*(2), 129–162.

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

Ross, S. A. (1977). The capital asset pricing model (CAPM), short-sale restrictions and related issues. *Journal of Finance 32*, 177–183.

Shen, D., H. Shen, H. Zhu, and J. S. Marron (2013). Surprising asymptotic conical structure in critical sample eigen-directions. *Working paper, Cornell University Librabry, available at arXiv:1303.6171 [math.ST]*.

Stock, J. H. and M. W. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics 20*(1), 147–162.

Stout, W. F. (1974). *Almost Sure convergence.* Academic Press.

Yin, Y. Q., Z. D. Bai, and P. R. Krishnaiah (1988). On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probability Theory and Related Fields 78*(4), 509–521.

Appendix: Proofs of Theorems

In the derivations of the proofs that follow we allow for $\Sigma_{ff} \neq I$ in general, apart from the specific instances relating to the estimation of μ_v and $\mathring{\alpha}$ where, without loss of generality, we impose $\Sigma_{ff}= I$. Further note that the proofs assume Σ_u is diagonal and, therefore, $\bar{\sigma}_N^2 = c_N$ and $\widehat{\bar{\sigma}_N^2} = \hat{c}_N$. The technical Lemmas used in the Appendices are stated in Supplementary Appendix I and proven in Supplementary Appendix III.

Proof of Theorem 1

We start by noting that

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

where $\bar{x}_t = \bar{\beta}_{1N} f_{1t} + \bar{\beta}_{2N} f_{2t} + ... + \bar{\beta}_{mN} f_{mt} + \bar{u}_t = \bar{\boldsymbol{\beta}}_N' \boldsymbol{f}_t + \bar{u}_t$, and $\bar{x} = T^{-1}\sum_{\tau=1}^{T}\bar{x}_\tau = \bar{\beta}_{1N} f_1 + \bar{\beta}_{2N} f_2 + ... + \bar{\beta}_{mN} f_m + \bar{u} = \bar{\boldsymbol{\beta}}_N' \boldsymbol{f} + \bar{u}$. Further, we assume the general setting discussed in Assumption 1 of Section 3.1 regarding the weak factor loadings and let $\boldsymbol{K}_\rho = (K_{\rho_1}, ..., K_{\rho_m})'$, where

$$K_{\rho_j} = K_j = \sum_{i=N_j+1}^{N} \beta_{ij} < \infty, \tag{38}$$

and N_j=$[N^{\alpha_j}]$. Then, we have

$$\hat{\sigma}_{\bar{x}}^2 = \bar{\boldsymbol{\beta}}_N^{'} \boldsymbol{S}_{ff} \bar{\boldsymbol{\beta}}_N + 2\bar{\boldsymbol{\beta}}_N' \left[\frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right] + \left[\frac{1}{T}\sum_{t=1}^{T}\bar{u}_t^2 - \bar{u}^2\right],$$

where

$$\boldsymbol{S}_{ff} = \frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)' \to_p \boldsymbol{\Sigma}_{ff} > 0, \text{ as } T \to \infty.$$

But under Assumption 1, $\bar{\boldsymbol{\beta}}_N = N^{\alpha-1}\boldsymbol{D}_N \bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho$, where $\bar{\boldsymbol{v}}_N = (\bar{v}_{1N}, \bar{v}_{2N}, ..., \bar{v}_{mN})'$ and $\bar{v}_{jN} = N_j^{-1}\sum_{i=1}^{N_j} v_{ij}$. So,

$$\bar{\boldsymbol{\beta}}_N^{'}\mathbf{S}_{ff}\bar{\boldsymbol{\beta}}_N = N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N + 2N^{\alpha-2}\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\boldsymbol{K}_\rho + N^{-2}\boldsymbol{K}_\rho'\mathbf{S}_{ff}\boldsymbol{K}_\rho = N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N + O\left(N^{\alpha-2}\right).$$

Hence,

$$\ln\left(\bar{\boldsymbol{\beta}}_N^{'}\mathbf{S}_{ff}\bar{\boldsymbol{\beta}}_N\right) = \ln\left(N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N + 2N^{\alpha-2}\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\boldsymbol{K}_\rho + N^{-2}\boldsymbol{K}_\rho'\mathbf{S}_{ff}\boldsymbol{K}_\rho\right) =$$

$$2(\alpha-1)\ln(N) + \ln\left(\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N\right) + \ln\left(1 + \frac{2N^{-\alpha}\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\boldsymbol{K}_\rho + N^{-2\alpha}\boldsymbol{K}_\rho'\mathbf{S}_{ff}\boldsymbol{K}_\rho}{\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N}\right)$$

$$= 2(\alpha-1)\ln(N) + \ln\left(\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N\right) + O_p\left(N^{-\alpha}\right).$$

Then,

$$\ln\left(\hat{\sigma}_{\bar{x}}^2\right) = \ln(\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N) + \ln\left(1 + \frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right] + \left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N}\right), \tag{39}$$

$$\ln\left(\hat{\sigma}_{\bar{x}}^2\right) = 2(\alpha-1)\ln(N) + \ln(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N)$$
$$+\ln\left(1 + \frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right] + \left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N}\right) + O_p\left(N^{-\alpha}\right).$$

Hence, recalling from (11) that $\hat{\alpha} = 1 + \ln(\hat{\sigma}_{\bar{x}}^2)/2\ln(N)$,we have

$$2\ln(N)\left(\hat{\alpha}-\alpha\right) - \ln(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N) = \ln\left(1 + \frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} + \frac{\left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N}\right) + O_p\left(N^{-\alpha}\right),$$

or

$$2\ln(N)\left(\hat{\alpha}-\alpha\right) - \ln(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N) = \frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} + \frac{\left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} + O_p\left(N^{-\alpha}\right) + o_p(B_{N,T}), \tag{40}$$

where

$$B_{N,T} = \frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} + \frac{\left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N}.$$

Consider the first term of the RHS of (40). We have,

$$\frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} = \frac{\frac{2}{\sqrt{TN}}N^{\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\left[\boldsymbol{\Sigma}_{ff}^{-1/2}\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\sqrt{N}\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}^{1/2}\boldsymbol{S}_{ff}^{1/2}\boldsymbol{\Sigma}_{ff}^{-1/2}\bar{\boldsymbol{\beta}}_N}. \tag{41}$$

We note that $\boldsymbol{S}_{ff}^{1/2}\boldsymbol{\Sigma}_{ff}^{-1/2} = 1 + O_p(T^{-1/2})$. But, by Lemma 2 (as N and $T \to \infty$)

$$\boldsymbol{\Sigma}_{ff}^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\left(\sqrt{N}\bar{u}_t\right) \to_p N(0, \bar{\sigma}_N^2 I_m), \tag{42}$$

where $\bar{\sigma}_N^2$ is as in (B1).

We need to determine the probability order of $1/\bar{\boldsymbol{\beta}}_N'\bar{\boldsymbol{\beta}}_N$. We note that

$$\frac{1}{\bar{\boldsymbol{\beta}}_N'\bar{\boldsymbol{\beta}}_N} - \frac{1}{N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N}$$

$$= \frac{1}{N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N + 2N^{\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho + N^{-2}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho} - \frac{1}{N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N}$$
$$= \frac{-N^{\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho - N^{-2}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho}{N^{4\alpha-4}\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N\right)^2 + N^{3\alpha-4}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N + N^{2\alpha-4}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N}$$
$$= -\left[N^{2-3\alpha}\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N\right)^{-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho + N^{2-4\alpha}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N\right)^{-2}\right]\left(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N^2\bar{\boldsymbol{v}}_N + N^{-\alpha}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{K}_\rho + N^{-2\alpha}\boldsymbol{K}_\rho'\boldsymbol{K}_\rho\right)^{-1}$$
$$= O_p\left(N^{2-3\alpha}\right),$$

and hence

$$\frac{2\bar{\boldsymbol{\beta}}_N'\left[\frac{1}{T}\sum_{t=1}^T\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\bar{u}_t\right]}{\bar{\boldsymbol{\beta}}_N^{'}\boldsymbol{S}_{ff}\bar{\boldsymbol{\beta}}_N} = O_p\left(T^{-1/2}N^{1/2-\alpha}\right) + O_p\left(T^{-1/2}N^{1/2-2\alpha}\right). \tag{43}$$

Consider now the second term on the RHS of (40). We use (43) again. Note that since, by Lemma 1 and

Theorems 17.5 and 19.11 of Davidson (1994), $\sqrt{NT}\bar{u} = O_p(1)$, and, since $\boldsymbol{S}_{ff}\boldsymbol{\Sigma}_{ff}^{-1} = 1 + O_p(T^{-1/2})$ where $0 < \boldsymbol{\Sigma}_{ff} < \infty$,

$$\frac{\bar{u}^2}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)} = \frac{\left(\sqrt{NT}\bar{u}\right)^2}{NT\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)} \tag{44}$$

$$= O_p\left(T^{-1}N^{1-2\alpha}\right). \tag{45}$$

Similarly,

$$\frac{\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)} \tag{46}$$

$$= \frac{\frac{1}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[(\sqrt{N}\bar{u}_t)^2 - \bar{\sigma}_N^2\right] + \sqrt{T}\bar{\sigma}_N^2\right\}}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)}$$

$$= \frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2 - 1\right] + \sqrt{T}\right\}}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)}$$

$$= \frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2 - 1\right]}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)}$$

$$+ \frac{\bar{\sigma}_N^2}{N\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)}.$$

Note that

$$\frac{\bar{\sigma}_N^2}{N\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)} - \frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N} = O_p(N^{1-3\alpha}). \tag{47}$$

Also, by Lemma 3,

$$\frac{1}{\sqrt{2T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2 - 1\right] \to_d N(0,1),$$

and

$$\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2 - 1\right]\right)}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N + N^{-1}\boldsymbol{K}_\rho\right)} = O_p(T^{-1/2}N^{1-2\alpha}) + O_p(T^{-1/2}N^{1-3\alpha}). \tag{48}$$

So,

$$2\ln(N)\left(\hat{\alpha} - \alpha\right) - \ln(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N) - \frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}$$
$$= O_p\left(\max\left(T^{-1/2}N^{1/2-\alpha}, T^{-1}N^{1-2\alpha}, T^{-1/2}N^{1-2\alpha}, N^{1-3\alpha}, N^{-\alpha}\right)\right).$$

Since $\alpha > 1/2$, in the first instance this implies that

$$\hat{\alpha} - \alpha = O_p\left(\frac{1}{\ln(N)}\right), \tag{49}$$

which establishes the consistency of $\hat{\alpha}$ as an estimate of α as N and $T \to \infty$, in any order.

Consider now the derivation of the asymptotic distribution of $\hat{\alpha}$. We have

$$\ln(N)\left(\hat{\alpha}-\alpha\right)-\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}=\ln(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N)+\frac{\frac{2}{\sqrt{TN}}\left[\boldsymbol{\Sigma}_{ff}^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\boldsymbol{f}_t-\bar{\boldsymbol{f}}\right)\left(\sqrt{N}\bar{u}_t\right)\right]}{N^{\alpha-1}\boldsymbol{S}_{ff}^{1/2}\left(\boldsymbol{S}_{ff}^{1/2}\boldsymbol{\Sigma}_{ff}^{-1/2}\right)\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}+$$

$$\frac{\left(\sqrt{NT}\bar{u}\right)^2}{NT\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}$$

$$+\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]\right)}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}+O_p(N^{-\alpha}).$$

We first examine $\ln(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N)$. If $\alpha_j=\alpha$, for all $j=1,...,m$, then by Lemma 11 we have

$$\sqrt{\min(N^\alpha,T)}\left[\ln(\bar{\boldsymbol{v}}_N'\boldsymbol{S}_{ff}\bar{\boldsymbol{v}}_N)-\ln(\boldsymbol{\mu}_v'\boldsymbol{\Sigma}_{ff}\boldsymbol{\mu}_v)\right]\to_d N\left(0,\omega_m\right),$$

while if $\alpha>\alpha_2...>\alpha_m$, then by Lemma 12 we have

$$\sqrt{\min(N^\alpha,T)}\left(\ln(\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N)-\ln(\boldsymbol{\mu}_v'\boldsymbol{D}_N\boldsymbol{\Sigma}_{ff}\boldsymbol{D}_N\boldsymbol{\mu}_v)\right)\to_d N(0,\omega).$$

Further, since $\alpha>1/2$,

$$\sqrt{\min(N^\alpha,T)}\left(\frac{\frac{2}{\sqrt{TN}}N^{\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\left[\boldsymbol{\Sigma}_{ff}^{-1/2}\frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{f}_t-\bar{\boldsymbol{f}}\right)\sqrt{N}\bar{u}_t\right]}{N^{2\alpha-2}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}^{1/2}\left(\boldsymbol{S}_{ff}^{1/2}\boldsymbol{\Sigma}_{ff}^{-1/2}\right)\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)=O_p\left(\sqrt{\min(N^\alpha,T)}T^{-1/2}N^{1/2-\alpha}\right)=o_p(1).$$

Similarly,

$$\sqrt{\min(N^\alpha,T)}\left(\frac{\left(\sqrt{NT}\bar{u}\right)^2}{NT\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}\right)=O_p\left(\sqrt{\min(N^\alpha,T)}T^{-1}N^{1-2\alpha}\right)=o_p(1),$$

and

$$\sqrt{\min(N^\alpha,T)}\left(\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]\right)}{\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)'\boldsymbol{S}_{ff}\left(N^{\alpha-1}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N+N^{-1}\boldsymbol{K}_\rho\right)}\right)=O_p\left(\sqrt{\min(N^\alpha,T)}T^{-1/2}N^{1-2\alpha}\right)=o_p(1).$$

Thus, if $\alpha_j=\alpha$, for all $j=1,...,m$,

$$\sqrt{\min(N^\alpha,T)}\left(\ln(N)\left(\hat{\alpha}-\alpha_N^*\right)-\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)\to_d N\left(0,\omega_m\right),$$

where $\alpha_N^*=\alpha+\ln(\mu_v^2)/2\ln(N)$ and $\mu_v^2=\sum_{j=1}^{m}\mu_{v_j}^2$, by setting $\Sigma_{ff}=I$ as normalisation. Otherwise, if $\alpha>\alpha_2...>\alpha_m$,

$$\sqrt{\min(N^\alpha,T)}\left(\ln(N)\left(\hat{\alpha}-\alpha_N^*\right)-\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)\to_d N\left(0,\omega\right),$$

where either $\alpha_N^*=\alpha+\ln(\mu_{v_1}^2)/2\ln(N)$ when (24) or (25) hold, or $\alpha_N^*=\alpha+\ln(\sum_{j=1}^{m}N^{2\left(\alpha_j-\alpha\right)}\mu_{v_j}^2)/2\ln(N)$ if neither of these two conditions hold, by referring to Lemma 13 as well. Again, we set $\Sigma_{ff}=I$ as normalisation.

Also, by Lemmas 7 and 9 we have

$$\sqrt{\min(N^\alpha,T)}\left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}\right)=O_p\left(\sqrt{\min(N^\alpha,T)}N^{2-4\alpha}\right)$$

and

$$\sqrt{\min(N^{\alpha},T)}\ln(N)\left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}\left(1+\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}\right)\right)=o_p\left(1\right),$$

which prove the remainder of the theorem.

Table A-B: Bias and RMSE ($\times 100$) for the $\mathring{\alpha}$ estimate of the cross-sectional exponent - case of cross-sectionally independent idiosyncratic errors - N=100,200,500,1000 and T=100,200,500

($\alpha_2 = 2\alpha/3$, f_{jt} and $u_{it} \sim IIDN(0,1)$, $v_{ij} \sim IIDU(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, $j = 1, 2$, $\mu_v = 0.87$, $\mu_{v_2} = 0.71$, $\mu_{v_1} = \sqrt{\mu_v^2 - N^{2(\alpha_2 - \alpha)}\mu_{v_2}^2}$)

Experiment A: Serially uncorrelated factors ($\rho_j = 0.0, j = 1, 2$)									**Experiment B**: Serially correlated factors ($\rho_j = 0.5, j = 1, 2$)								
	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00		α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
N\T					**100**				**N\T**					**100**			
100	Bias	0.97	0.45	0.02	0.27	0.15	-0.02	-0.04	**100**	Bias	2.52	1.64	0.89	0.82	0.50	0.16	-0.07
	RMSE	2.02	1.44	0.93	0.71	0.48	0.29	0.06		RMSE	3.61	2.60	1.74	1.33	0.88	0.48	0.10
200	Bias	0.52	0.43	0.19	0.15	0.04	0.03	0.03	**200**	Bias	1.52	1.22	0.76	0.52	0.26	0.13	0.01
	RMSE	1.48	0.99	0.63	0.46	0.29	0.17	0.03		RMSE	2.56	1.87	1.25	0.88	0.54	0.31	0.06
500	Bias	0.13	0.05	0.09	0.03	0.04	0.03	0.05	**500**	Bias	1.18	0.70	0.49	0.25	0.15	0.06	0.03
	RMSE	0.68	0.46	0.35	0.23	0.16	0.10	0.06		RMSE	1.91	1.22	0.82	0.50	0.32	0.17	0.05
1000	Bias	0.00	0.01	0.05	0.01	0.03	0.00	0.06	**1000**	Bias	0.84	0.48	0.32	0.18	0.10	0.02	0.03
	RMSE	0.53	0.35	0.25	0.17	0.12	0.07	0.06		RMSE	1.42	0.87	0.57	0.36	0.23	0.12	0.05
					200									**200**			
100	Bias	2.51	1.55	0.62	0.60	0.31	0.05	-0.10	**100**	Bias	3.50	2.42	1.33	1.12	0.68	0.33	-0.10
	RMSE	3.12	2.19	1.28	0.99	0.61	0.33	0.10		RMSE	4.25	3.08	1.94	1.51	0.98	0.58	0.10
200	Bias	0.45	0.54	0.28	0.21	0.08	0.04	-0.02	**200**	Bias	1.88	1.44	0.88	0.62	0.32	0.16	-0.02
	RMSE	1.14	0.96	0.61	0.45	0.28	0.15	0.02		RMSE	2.56	1.89	1.22	0.87	0.53	0.28	0.02
500	Bias	0.24	0.12	0.14	0.06	0.06	0.04	0.02	**500**	Bias	1.25	0.73	0.54	0.31	0.21	0.11	0.02
	RMSE	0.61	0.41	0.31	0.19	0.14	0.08	0.02		RMSE	1.67	1.04	0.73	0.46	0.30	0.16	0.02
1000	Bias	0.10	0.08	0.10	0.06	0.07	0.03	0.03	**1000**	Bias	0.86	0.53	0.39	0.24	0.17	0.08	0.03
	RMSE	0.38	0.26	0.21	0.14	0.10	0.05	0.03		RMSE	1.20	0.74	0.52	0.33	0.23	0.11	0.03
					500									**500**			
100	Bias	4.02	3.06	1.75	1.38	0.79	0.31	-0.13	**100**	Bias	5.91	4.20	2.60	1.94	1.13	0.46	-0.13
	RMSE	4.28	3.30	2.05	1.63	1.02	0.53	0.13		RMSE	6.28	4.50	2.89	2.19	1.35	0.65	0.13
200	Bias	1.75	1.56	0.88	0.64	0.31	0.11	-0.05	**200**	Bias	3.25	2.54	1.55	1.06	0.56	0.26	-0.05
	RMSE	2.08	1.83	1.14	0.86	0.51	0.23	0.05		RMSE	3.65	2.81	1.80	1.26	0.73	0.37	0.05
500	Bias	0.90	0.47	0.31	0.13	0.09	0.04	-0.01	**500**	Bias	1.90	1.14	0.77	0.44	0.29	0.15	-0.01
	RMSE	1.17	0.72	0.48	0.27	0.17	0.09	0.01		RMSE	2.19	1.36	0.93	0.56	0.37	0.19	0.01
1000	Bias	0.66	0.23	0.14	0.07	0.06	0.02	0.00	**1000**	Bias	1.12	0.67	0.47	0.28	0.19	0.08	0.00
	RMSE	0.90	0.42	0.25	0.15	0.10	0.04	0.00		RMSE	1.35	0.83	0.57	0.35	0.23	0.11	0.00

Table C-D: Bias and RMSE ($\times 100$) for the $\mathring{\alpha}$ estimate of the cross-sectional exponent - case of two serially independent factors - N=100,200,500,1000 and T=100,200,500

$(\alpha_2 = 2\alpha/3$, f_{jt} and $u_{it} \sim IIDN(0,1)$, $v_{ij} \sim IIDU(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, $j = 1,2$, $\mu_v = 0.87$, $\mu_{v_2} = 0.71$, $\mu_{v_1} = \sqrt{\mu_v^2 - N^{2(\alpha_2-\alpha)}\mu_{v_2}^2})$

Experiment C: Non-normal idiosyncratic errors ($\varepsilon_{it} \sim IID\chi^2(2)$)									**Experiment D**: Spatially dependent idiosyncratic errors ($\theta = 0.2$)								
	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00		α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
N\T					**100**				**N\T**					**100**			
100	Bias	1.24	0.60	0.12	0.30	0.17	-0.01	-0.05	**100**	Bias	1.62	0.89	0.37	0.53	0.39	0.18	0.09
	RMSE	2.29	1.56	0.98	0.73	0.49	0.29	0.06		RMSE	2.45	1.73	1.10	0.90	0.63	0.35	0.10
200	Bias	0.27	0.36	0.17	0.13	0.02	0.02	0.03	**200**	Bias	0.31	0.39	0.22	0.19	0.09	0.09	0.09
	RMSE	1.26	0.89	0.57	0.41	0.27	0.16	0.04		RMSE	1.23	0.97	0.66	0.49	0.33	0.20	0.09
500	Bias	0.14	0.04	0.08	0.02	0.03	0.02	0.05	**500**	Bias	0.13	0.03	0.07	0.02	0.04	0.05	0.07
	RMSE	0.72	0.49	0.35	0.24	0.17	0.10	0.06		RMSE	0.75	0.54	0.40	0.27	0.19	0.12	0.07
1000	Bias	0.02	0.01	0.05	0.01	0.03	0.00	0.05	**1000**	Bias	-0.01	-0.02	0.03	0.00	0.03	0.01	0.06
	RMSE	0.50	0.34	0.25	0.17	0.12	0.07	0.06		RMSE	0.56	0.39	0.28	0.20	0.14	0.08	0.07
					200									**200**			
100	Bias	2.10	1.29	0.49	0.51	0.30	0.04	-0.10	**100**	Bias	2.78	1.79	0.85	0.81	0.53	0.23	0.05
	RMSE	2.85	1.97	1.18	0.89	0.60	0.32	0.10		RMSE	3.42	2.39	1.43	1.14	0.76	0.40	0.05
200	Bias	0.64	0.64	0.31	0.23	0.08	0.04	-0.02	**200**	Bias	0.87	0.77	0.42	0.33	0.17	0.12	0.04
	RMSE	1.40	1.09	0.66	0.48	0.29	0.16	0.02		RMSE	1.52	1.19	0.74	0.55	0.35	0.21	0.04
500	Bias	0.28	0.13	0.14	0.06	0.06	0.04	0.02	**500**	Bias	0.28	0.12	0.14	0.07	0.08	0.07	0.04
	RMSE	0.67	0.42	0.31	0.20	0.14	0.08	0.02		RMSE	0.69	0.45	0.33	0.22	0.16	0.10	0.04
1000	Bias	0.11	0.07	0.10	0.06	0.07	0.03	0.03	**1000**	Bias	0.08	0.04	0.08	0.05	0.07	0.04	0.04
	RMSE	0.41	0.26	0.20	0.14	0.11	0.05	0.03		RMSE	0.44	0.30	0.22	0.15	0.12	0.06	0.04
					500									**500**			
100	Bias	4.64	3.33	1.92	1.45	0.84	0.30	-0.13	**100**	Bias	6.01	4.16	2.48	1.87	1.13	0.54	0.01
	RMSE	4.97	3.58	2.19	1.68	1.07	0.51	0.13		RMSE	6.36	4.38	2.71	2.06	1.30	0.69	0.01
200	Bias	2.15	1.77	0.95	0.67	0.32	0.12	-0.05	**200**	Bias	2.47	1.97	1.10	0.79	0.43	0.21	0.02
	RMSE	2.52	2.03	1.21	0.88	0.50	0.25	0.05		RMSE	2.78	2.22	1.35	0.99	0.59	0.31	0.02
500	Bias	0.88	0.45	0.30	0.13	0.09	0.04	-0.01	**500**	Bias	0.89	0.46	0.31	0.15	0.11	0.07	0.02
	RMSE	1.18	0.69	0.47	0.27	0.17	0.09	0.01		RMSE	1.21	0.74	0.49	0.29	0.19	0.11	0.02
1000	Bias	0.30	0.15	0.13	0.07	0.06	0.02	0.00	**1000**	Bias	0.27	0.13	0.11	0.06	0.06	0.03	0.01
	RMSE	0.55	0.32	0.22	0.14	0.09	0.04	0.00		RMSE	0.53	0.31	0.21	0.14	0.10	0.05	0.01

Supplementary Appendices

Supplementary Appendix I: Statement of technical Lemmas

Lemma 1 *Under Assumptions 2 and 3,* $\boldsymbol{f}_t, \{u_{it}\}_{t=1}^{\infty}$, *and* $\{u_{it}\}_{i=1}^{\infty}$ *are* L_r*-bounded,* L_2*-NED processes of size* $-\zeta$, *for some* $r > 2$. *This result holds uniformly over* i, *in the case of* $\{u_{it}\}_{t=1}^{\infty}$, *and over* t, *in the case of* $\{u_{it}\}_{i=1}^{\infty}$.

Lemma 2 *Under Assumptions 2 and 3,*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right) \left(\sqrt{N} \bar{u}_t\right) \to_p N(\mathbf{0}, \bar{\sigma}_N^2 I_m),$$

where

$$\bar{\sigma}_N^2 = \lim_{N \to \infty} \left(\frac{\sum_{i=1}^{N} \left(\sigma_i^2 + \sum_{j=1, j \neq i}^{\infty} \sigma_{ij}\right)}{N} \right), \tag{B1}$$

$\sigma_i^2 = E(u_{it}^2)$, *and* $\sigma_{ij} = E(u_{it} u_{it-j})$.

Lemma 3 *Under Assumption 3,*

$$\frac{1}{\sqrt{2T}} \sum_{t=1}^{T} \left[\left(\sqrt{N} \bar{u}_t\right)^2 - E\left(\sqrt{N} \bar{u}_t\right)^2 \right] \to_d N(0, V),$$

where

$$V = \lim_{N \to \infty} \left(Var\left(\left(\sqrt{N}\bar{u}_t\right)^2 \right) + \sum_{j=1}^{\infty} Cov\left(\left(\sqrt{N}\bar{u}_t\right)^2, \left(\sqrt{N}\bar{u}_{t-j}\right)^2 \right) \right).$$

Lemma 4 *Under Assumptions 1-3, if* $m = 1$ *then* $\widehat{\bar{\sigma}_N^2} - \bar{\sigma}_N^2 = O_p\left(T^{-1}\right) + O_p\left((NT)^{-1/2}\right)$. *If* $m > 1$ *then* $\widehat{\bar{\sigma}_N^2} - \bar{\sigma}_N^2 = O_p\left(N^{\alpha-1} T^{-1/2}\right)$.

Lemma 5 *Under Assumptions 1-2 and* $m = 1$, $\sqrt{\min(N^\alpha, T)} \left(\ln(s_{f_1}^2 \bar{v}_{1N}^2) - \ln(\sigma_{f_1}^2 \mu_{v_1}^2)\right) \to_d N(0, \omega)$.

Lemma 6 *Under Assumption 5 and Assumptions 2-3 and* $m = 1$, $\sqrt{\min(N, T)} \left(\ln(s_{f_1}^2 \bar{v}_{1N}^2) - \ln(\sigma_{f_1}^2 \mu_{v_1}^2)\right) \to_d N(0, \omega)$.

Lemma 7 *Under Assumptions 1-3, and as long as* $\alpha > 1/2$,

$$\sqrt{\min(N^\alpha, T)} \left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{\boldsymbol{v}}_N' \boldsymbol{D}_N \boldsymbol{S}_{ff} \boldsymbol{D}_N \bar{\boldsymbol{v}}_N} - \frac{\widehat{\bar{\sigma}_N^2}}{N \hat{\sigma}_{\bar{x}}^2} \right) = O_p\left(\sqrt{\min(N^\alpha, T)} N^{2-4\alpha}\right).$$

Lemma 8 *Under Assumption 5 and Assumptions 2-3,*

$$\sqrt{\min(N, T)} \left(\frac{\bar{\sigma}_N^2}{N^{\alpha} \bar{\boldsymbol{v}}_N' \boldsymbol{D}_N \boldsymbol{S}_{ff} \boldsymbol{D}_N \bar{\boldsymbol{v}}_N} - \frac{\widehat{\bar{\sigma}_N^2}}{N \hat{\sigma}_{\bar{x}}^2} \right) = o_p(1).$$

Lemma 9 *Under Assumptions 1-2, and as long as* $\alpha > 1/2$,

$$\sqrt{\min(N^\alpha, T)} \ln(N) \left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{\boldsymbol{v}}_N' \boldsymbol{D}_N \boldsymbol{S}_{ff} \boldsymbol{D}_N \bar{\boldsymbol{v}}_N} - \frac{\widehat{\bar{\sigma}_N^2}}{N \hat{\sigma}_{\bar{x}}^2} \left(1 + \frac{\widehat{\bar{\sigma}_N^2}}{N \hat{\sigma}_{\bar{x}}^2}\right) \right) = o_p(1).$$

Lemma 10 *Under Assumptions 1-3 and* $\Sigma_{ff} = I$, *if* $\alpha = \alpha_2 = ... = \alpha_{q-1} > \alpha_q \geq ... \geq \alpha_m$,

$$\hat{\sigma}_{\bar{x}}^2 - N^{2\alpha-2} \sum_{j=1}^{q} \mu_{v_j}^2 \to_p 0.$$

In particular if $\alpha > \alpha_2 \geq ... \geq \alpha_m$,

$$\hat{\sigma}_{\bar{x}}^2 - N^{2\alpha-2} \mu_{v_1}^2 \to_p 0.$$

Lemma 11 *Under Assumptions 1-2, and assuming* $\alpha_j = \alpha$, *for all* $j = 1, ..., m$,

$$\sqrt{\min(N^{\alpha}, T)}\left(\ln(\bar{\boldsymbol{v}}_N' \boldsymbol{S}_{ff} \bar{\boldsymbol{v}}_N) - \ln(\boldsymbol{\mu}_v' \boldsymbol{\Sigma}_{ff} \boldsymbol{\mu}_v)\right) \to_d N\left(0, \omega_m\right),$$

where $\boldsymbol{\mu}_v = E\left(\boldsymbol{v}_j\right)$, $\boldsymbol{\Sigma}_{ff} = E\left(\left(\boldsymbol{f}_t - \boldsymbol{\mu}_f\right)' \left(\boldsymbol{f}_t - \boldsymbol{\mu}_f\right)\right)$,

$$\omega_m = \lim_{N,T\to\infty} \min\left(N^{\alpha}, T\right) E\left(\left\{\left(\bar{\boldsymbol{v}}_N' \bar{\boldsymbol{f}}_T - \boldsymbol{\mu}_v' \boldsymbol{\mu}_f\right)^2 - E\left[\left(\bar{\boldsymbol{v}}_N' \bar{\boldsymbol{f}}_T - \boldsymbol{\mu}_v' \boldsymbol{\mu}_f\right)^2\right]\right\}^2\right),$$

$\boldsymbol{\mu}_f = E\left(\boldsymbol{f}_t\right)$ *and* $\bar{\boldsymbol{f}}_T = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{f}_t$.

Lemma 12 *Under Assumptions 1-2, and assuming* $\alpha > \alpha_2 > ... > \alpha_m$,

$$\sqrt{\min(N^{\alpha}, T)}\left(\ln(\bar{\boldsymbol{v}}_N' \boldsymbol{D}_N \boldsymbol{S}_{ff} \boldsymbol{D}_N \bar{\boldsymbol{v}}_N) - \ln(\boldsymbol{\mu}_v' \boldsymbol{D}_N \boldsymbol{\Sigma}_{ff} \boldsymbol{D}_N \boldsymbol{\mu}_v)\right) \to_d N\left(0, \omega\right).$$

Lemma 13 *Under Assumptions 1-3, and* $\alpha > \alpha_2 \geq \alpha_3 \geq ... \geq \alpha_m$,

$$\sqrt{\min\left(N^{\alpha}, T\right)} \ln\left(N\right) \ln(\boldsymbol{\mu}_v' \boldsymbol{D}_N \boldsymbol{\Sigma}_{ff} \boldsymbol{D}_N \boldsymbol{\mu}_v) - \ln\left(\mu_{v_1}^2 \sigma_{f_1}^2\right) = o\left(1\right),$$

if either $\alpha_2 - \alpha < -0.25$ *or, if* $T^b = N$, $\alpha_2 < 3\alpha/4$ *and*

$$b > \frac{1}{4\left(\alpha - \alpha_2\right)}. \tag{B2}$$

Lemma 14 *Let* $\beta_{i1} = N^{\alpha-1} v_{i1}$, $1/2 < \alpha \leq 1$, *where* $v_{i1} = v_{Ni} = \breve{v}_i + c_{Ni}$ *and* $\{\breve{v}_i\}_{i=1}^N$ *is an i.i.d. sequence of random variables with mean* $\mu_{\breve{v}} \neq 0$, *and variance* $\sigma_{\breve{v}}^2 < \infty$. *Let* $\bar{c}_N = \frac{1}{N}\sum_{i=1}^{N} c_{Ni}$. *Under Assumptions 2-3, (a)* $\hat{\alpha}$, $\tilde{\alpha}$ *and* $\breve{\alpha}$ *are consistent estimators of* α, *if* $\bar{c}_N = o_p\left(N^c\right)$ *for all* $c > 0$, *(b) Corollary 1 holds, if* $\sqrt{N}\bar{c}_N = o_p\left(1\right)$.

Lemma 15 *Let* $\hat{a}$ *denote a generic estimator of* α *such that* $\hat{a} - \alpha = O_p(h_N)$ *where* $h_N \to 0$. *Then,*

$$N^{\hat{a}} - N^{\alpha} = O_p(N^{\alpha} h_N \ln N).$$

Lemma 16 *Denote the OLS estimator of the regression coefficient of* x_{it} *on* $\tilde{x}_t = \bar{x}_t/\hat{\sigma}_{\bar{x}}$, *by* $\hat{v}_{i1}$, *and let* $\left\{\hat{v}_{i1}^{(s)}\right\}_{i=1}^{N}$ *be the reordering of* $\{\hat{v}_{i1}\}_{i=1}^N$ *where* $\left|\hat{v}_{i1}^{(s)}\right| \geq \left|\hat{v}_{i1+1}^{(s)}\right|$, $\forall i$. *Under Assumptions 1-3,* $m = 1$ *and assuming that* $\lim_{T,N\to\infty} T^{-1}N^{\alpha} < \infty$ *and that* $\mu_{v_1} \neq 0$ *is known, we have*

$$\frac{\sum_{i=1}^{N^{\hat{\alpha}}}\left(\hat{v}_{i1}^{(s)} - \frac{1}{N^{\hat{\alpha}}}\sum_{j=1}^{N^{\hat{\alpha}}}\hat{v}_{j1}^{(s)}\right)^2}{N^{\hat{\alpha}} - 1} - \frac{\sigma_{v_1}^2}{\mu_{v_1}^2} = o_p\left(1\right), \tag{B3}$$

if, further, $\hat{\alpha} - \alpha = o_p\left(\left(\ln N\right)^{-1}\right)$.

Lemma 17 *Under Assumptions 1-3 and* $m = 1$,

$$\hat{\beta}_{i1} - \frac{\beta_{i1}}{\mu_{v_1}} = O_p\left(\frac{1}{N^{\alpha/2}}\right) + O_p\left(\frac{1}{T^{1/2}}\right).$$

Lemma 18 *Under Assumptions 1-2, we have* $\hat{V}_{\overline{f_1^2}} - V_{\overline{f_1^2}} = o_p(1)$, *as long as* $l \to \infty$, $l = o(T)$ *and* $l = o\left(N^{\alpha-1/2}T^{1/2}\right)$.

Supplementary Appendix II: Proof of Corollary 1

We reconsider (39) and $m = 1$. Under Assumption 5, $\bar{\beta}_{1N} = N^{(\alpha-1)/2}\bar{v}_{1N}$, where $\bar{v}_{1N} = N^{-1}\sum_{i=1}^{N} v_{1i}$, we have

$$\ln(\bar{\beta}_{1N}^2 s_{f_1}^2) = \ln\left(N^{(\alpha-1)/2}\bar{v}_{1N} s_{f_1}\right)^2 = (\alpha - 1)\ln(N) + \ln(s_{f_1}^2 \bar{v}_{1N}^2),$$

where $s_{f_i}^2 = T^{-1}\sum_{t=1}^T \left(f_{it} - T^{-1}\sum_{t=1}^T f_{it}\right)^2$ for $i = 1, ..., N$ and here $i = 1$. Hence, recalling from (11) that $\hat{\alpha} = 1 + \ln(\hat{\sigma}_{\bar{x}}^2)/2\ln(N)$,we have

$$\ln(N)\left(\hat{\alpha} - \alpha\right) - \ln(s_{f_1}^2 \bar{v}_{1N}^2) = \ln\left(1 + \frac{2\bar{\beta}_{1N}\left[\frac{1}{T}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\bar{u}_t\right] + \left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\beta}_{1N}^2 s_{f_1}^2}\right). \tag{B4}$$

However,

$$\ln\left(1 + \frac{2\bar{\beta}_{1N}\left[\frac{1}{T}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\bar{u}_t\right] + \left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\beta}_{1N}^2 s_{f_1}^2}\right) = \frac{2\bar{\beta}_{1N}\left[\frac{1}{T}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\bar{u}_t\right] + \left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\beta}_{1N}^2 s_{f_1}^2} + o_p\left(B_{N,T}\right), \tag{B5}$$

where when $m = 1$,

$$B_{N,T} = \frac{2\bar{\beta}_{1N}\left[\frac{1}{T}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\bar{u}_t\right] + \left[\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2\right]}{\bar{\beta}_{1N}^2 s_{f_1}^2}.$$

Consider the first term of the RHS of (40). We have,

$$\frac{2\bar{\beta}_{1N}\left[\frac{1}{T}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\bar{u}_t\right]}{\bar{\beta}_{1N}^2 s_{f_1}^2} = \frac{\frac{2}{\sqrt{TN}}\left[\frac{1}{\sigma_{f_1}\sqrt{T}}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\left(\sqrt{N}\bar{u}_t\right)\right]}{\left(s_{f_1}\bar{\beta}_{1N}\right)\left(s_{f_1}/\sigma_{f_1}\right)}.$$

We note that $s_{f_1}/\sigma_{f_1} = 1 + O_p(T^{-1/2})$. But, by Lemma 2 (as N and $T \to \infty$)

$$\frac{1}{\sigma_{f_1}\sqrt{T}}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\left(\sqrt{N}\bar{u}_t\right) \to_p N(0, \bar{\sigma}_N^2), \tag{B6}$$

where $\overline{\sigma}_N^2$ is as in (B1). Also, $1/\bar{\beta}_{1N} = N^{(1-\alpha)/2}\left(1/\bar{v}_{1N}\right)$. Hence,

$$\begin{aligned}\frac{2\bar{\beta}_{1N}\left[\frac{1}{T}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\bar{u}_t\right]}{\bar{\beta}_{1N}^2 s_{f_1}^2} &= \frac{\frac{2}{\sqrt{TN}}\left[\frac{1}{\sigma_{f_1}\sqrt{T}}\sum_{t=1}^T \left(f_{1t} - \bar{f}_1\right)\left(\sqrt{N}\bar{u}_t\right)\right]}{s_{f_1}\bar{\beta}_{1N}(s_{f_1}/\sigma_{f_1})} \\ &= O_p\left(T^{-1/2}N^{-\alpha/2}\right).\end{aligned} \tag{B7}$$

Consider now the second term on the RHS of (40). Note that since, by Lemma 1 and Theorems 17.5 and 19.11 of Davidson (1994), $\sqrt{NT}\bar{u} = O_p(1)$, and, since $s_{f_1}^2/\sigma_{f_1}^2 = 1 + O_p(T^{-1/2})$ where $0 < \sigma_{f_1}^2 < \infty$,

$$\frac{\bar{u}^2}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2} = \frac{\left(\sqrt{NT}\bar{u}\right)^2}{NT\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2} = O_p\left(T^{-1}N^{-\alpha}\right). \tag{B8}$$

Similarly,

$$\begin{aligned}\frac{\frac{1}{T}\sum_{t=1}^T \bar{u}_t^2}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2} &= \frac{\frac{1}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[(\sqrt{N}\bar{u}_t)^2 - \bar{\sigma}_N^2\right] + \sqrt{T}\bar{\sigma}_N^2\right\}}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2} = \frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2 - 1\right] + \sqrt{T}\right\}}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2} \\ &= \frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2 - 1\right]}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2} + \frac{\bar{\sigma}_N^2}{N\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2}.\end{aligned} \tag{B9}$$

But, by Lemma 3,

$$\frac{1}{\sqrt{2T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2 - 1\right] \to_d N(0,1),$$

and

$$\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]\right)}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2}=O_p(T^{-1/2}N^{-\alpha}). \tag{B10}$$

Therefore, collecting all results derived above, and keeping the highest order terms of the RHS of (B7), (B8), and (B10), we have

$$2\ln(N)\left(\hat{\alpha}-\alpha\right)-\ln(s_{f_1}^2\bar{v}_{1N}^2)-\frac{\bar{\sigma}_N^2}{N^{\alpha}\bar{v}_{1N}^2 s_{f_1}^2}=O_p\left(T^{-1/2}N^{-\alpha/2}\right).$$

In the first instance this implies that

$$\hat{\alpha}-\alpha=O_p\left(\frac{1}{\ln(N)}\right), \tag{B11}$$

which establishes the consistency of $\hat{\alpha}$ as an estimate of α as N and $T\to\infty$, in any order.

Consider now the derivation of the asymptotic distribution of $\hat{\alpha}$. We have

$$\ln(N)\left(\hat{\alpha}-\alpha\right)-\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{v}_{1N}^2 s_{f_1}^2}\simeq\ln(s_{f_1}^2\bar{v}_{1N}^2)+\frac{\frac{2}{\sqrt{TN}}\left[\frac{1}{\sigma_f\sqrt{T}}\sum_{t=1}^T\left(f_{1t}-\bar{f}_1\right)\left(\sqrt{N}\bar{u}_t\right)\right]}{s_{f_1}N^{(\alpha-1)/2}\bar{v}_{1N}(s_{f_1}/\sigma_{f_1})}+$$
$$\frac{\left(\sqrt{NT}\bar{u}\right)^2}{NT\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2}+\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2}.$$

where $A\simeq B$ denotes that $A-B=o_p(B)$. We first examine $\ln(s_{f_1}^2\bar{v}_{1N}^2)$. By Lemma 6 we have

$$\sqrt{\min(N,T)}\left(\ln(s_{f_1}^2\bar{v}_{1N}^2)-\ln(\sigma_{f_1}^2\mu_{v_1}^2)\right)\to_d N\left(0,\omega\right).$$

Further,

$$\sqrt{\min(N,T)}\left(\frac{\frac{2}{\sqrt{TN}}\left[\frac{1}{\sigma_{f_1}\sqrt{T}}\sum_{t=1}^T\left(f_{1t}-\bar{f}_1\right)\left(\sqrt{N}\bar{u}_t\right)\right]}{s_{f_1}N^{(\alpha-1)/2}\bar{v}_{1N}(s_{f_1}/\sigma_{f_1})}\right)=O_p\left(\sqrt{\min(N,T)}T^{-1/2}N^{-\alpha/2}\right)=o_p(1).$$

Similarly,

$$\sqrt{\min(N,T)}\left(\frac{\left(\sqrt{NT}\bar{u}\right)^2}{NT\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2}\right)=O_p\left(\sqrt{\min(N,T)}T^{-1}N^{-\alpha}\right)=o_p(1),$$

and

$$\sqrt{\min(N,T)}\left(\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{t=1}^T\left[\left(\frac{\sqrt{N}\bar{u}_t}{\bar{\sigma}_N}\right)^2-1\right]}{\left(N^{(\alpha-1)/2}\bar{v}_{1N}\right)^2 s_{f_1}^2}\right)=O_p\left(\sqrt{\min(N,T)}T^{-1/2}N^{-\alpha}\right)=o_p(1).$$

Thus,

$$\sqrt{\min(N,T)}\left(\ln(N)\left(\hat{\alpha}-\alpha_N^*\right)-\frac{\bar{\sigma}_N^2}{N^{\alpha}\bar{v}_{1N}^2 s_{f_1}^2}\right)\to_d N\left(0,\omega\right),$$

where $\alpha_N^*=\alpha+\ln(\mu_{v_1}^2)/2\ln(N)$, by setting $\sigma_f^2=1$ as normalisation. The second part of the Corollary follows by Lemma 8.

Supplementary Appendix III: Proofs of technical Lemmas

Proof of Lemma 1

The proof of this lemma considers the more general Assumption 4 for the error terms which incorporates Assumption 3. By the Marcinkiewicz–Zygmund inequality (see, e.g., (Stout, 1974, Theorem 3.3.6)),

$$\sup_i E(|u_{it}|^r) = \sup_i E\left(\left|\sum_{l=0}^{\infty}\left(\psi_{il}\sum_{s=-\infty}^{\infty}\xi_{is}v_{st-l}\right)\right|^r\right) \leq c\left(\sup_i\left(\sum_{l=0}^{\infty}|\psi_{il}|^2\right)\sup_i\left(\sum_{s=-\infty}^{\infty}|\xi_{is}|^2\right)\right)^{r/2}\left(\sup_{i,t}E(|v_{it}|^r)\right),$$

so u_{it} is L_r-bounded if $\sup_i \sup_t E(|v_t|^r) < \infty$ which holds by Assumption 4. Moreover, writing $\|\cdot\|_r$ for the L_r-norm, we have, by Minkowski's inequality,

$$\sup_i \left\|u_{it} - E(u_{it}|\mathcal{F}_{t,|m|}^{v_i})\right\|_2 = \sup_i \left\|\sum_{j=m+1}^{\infty}\psi_{ij}\left(\sum_{|s|\geq m}\xi_{is}\nu_{st}\right)\right\|_2 \leq \sup_{i,t}\|v_{it}\|_2\left(\sup_i\sum_{j=m+1}^{\infty}|\psi_{ij}|\right)\left(\sup_i\left(\sum_{|s|\geq m}|\xi_{is}|\right)\right), \quad \text{(B12)}$$

for any integer $m > 0$ where $\mathcal{F}_{t,|m|}^{\nu_i}$ is the σ-field generated by $\{v_{is}; i, s \leq t-m\} \cup \{v_{is}; i, s \geq t+m\}$. But, Assumption 4 implies that $\sup_i \lim_{m\to\infty} m^{\zeta}\sum_{j=m+1}^{\infty}|\psi_{ij}| = O(1)$ and $\sup_i \lim_{m\to\infty} m^{\zeta}\left(\sum_{|s|\geq m}|\xi_{is}|\right) = O(1)$. Consequently $\{u_{it}\}_{t=1}^{\infty}$ and $\{u_{it}\}_{i=1}^{\infty}$ are L_r-bounded, L_2-NED processes of size $-\zeta$, uniformly over i and t. Similarly, we can show that $\boldsymbol{f}_t$ are L_r-bounded ($r \geq 2$) L_2-NED processes of size $-\zeta$.

Proof of Lemma 2

We have $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\left(\sqrt{N}\bar{u}_t\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{z}_t$, where $\boldsymbol{z}_t = \left(\boldsymbol{f}_t - \bar{\boldsymbol{f}}\right)\left(\sqrt{N}\bar{u}_t\right)$. We have that $\boldsymbol{z}_t$ are stationary processes such that $E(\boldsymbol{z}_t) = \boldsymbol{0}$. We note that by Lemma 1 and Theorem 24.6 of Davidson (1994), we have that $E\left(\left(\sqrt{N}\bar{u}_t\right)^2\right) = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 < \infty$. Further, by Theorem 17.8 of Davidson (1994), we have that sums of L_2-bounded, L_2-NED triangular arrays of size $-\zeta$ are L_2-bounded, L_2-NED triangular arrays of size $-\zeta$ as well, implying, given Lemma 1, that $\sqrt{N}\bar{u}_t$ is an L_2-bounded, L_2-NED triangular arrays of size $-\zeta$. Further, by the Marcinkiewicz–Zygmund inequality,

$$E(\left|\sqrt{N}\bar{u}_t\right|^r) = E\left(\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{l=0}^{\infty}\left(\psi_{il}\sum_{s=-\infty}^{\infty}\xi_{is}v_{st-l}\right)\right|^r\right) \leq c\left(\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{l=0}^{\infty}|\psi_{il}|^2\right)\left(\sum_{s=-\infty}^{\infty}|\xi_{is}|^2\right)\right)^{r/2}\sup_{i,t}E(|v_{it}|^r) \leq \quad \text{(B13)}$$

$$c\left(\sup_i\left(\sum_{l=0}^{\infty}|\psi_{il}|^2\right)\sup_i\left(\sum_{s=-\infty}^{\infty}|\xi_{is}|^2\right)\right)^{r/2}\left(\sup_{i,t}E(|v_{it}|^r)\right) < \infty.$$

As a result, $\sqrt{N}\bar{u}_t$ is a L_r-bounded, L_2-NED triangular arrays of size $-\zeta$.

Finally, since $\{\sqrt{N}\bar{u}_t\}$ and $\{\boldsymbol{f}_t\}$ are L_r-bounded ($r \geq 2$) L_2-NED processes of size $-\zeta$ on a ϕ-mixing process of size $-\eta$ ($\eta > 1$), then, by Example 17.17 of Davidson (1994), $\{\boldsymbol{z}_t\}$ are L_2-NED of size $-\{\zeta(\varphi-2)\}/\{2(\varphi-1)\} \leq -1/2$ on a ϕ-mixing process of size $-\eta$. Since ν_{it} and $\nu_{\boldsymbol{f}t}$ are i.i.d. processes they are also ϕ-mixing processes of any size. In view of Theorem 17.5(ii) of Davidson (1994), this in turn implies that $\{\boldsymbol{z}_t\}$ are L_2-mixingale of size $-1/2$, if $2\eta > \zeta$, which automatically holds by the i.i.d. property of ν_{it} and $\nu_{\boldsymbol{f}t}$. This implies the result of the Lemma by Theorem 24.6 of Davidson (1994).

Proof of Lemma 3

By Lemma 2, $\sqrt{N}\bar{u}_t$ is a L_r-bounded, L_2-NED triangular arrays of size $-\zeta$. By Example 17.17 of Davidson (1994), and (B13), $\left(\sqrt{N}\bar{u}_t\right)^2$ is L_r-NED of size $-\{\zeta(\varphi-2)\}/\{2(\varphi-1)\} \leq -1/2$, $r > 4$. Then, by Theorem 24.6 of Davidson (1994), the result follows.

Proof of Lemma 4

We need to show that $\widehat{\bar{\sigma}_N^2} - \bar{\sigma}_N^2 = O_p\left(T^{-1}\right) + O_p\left((NT)^{-1/2}\right)$ if $m=1$ and $\widehat{\bar{\sigma}_N^2} - \bar{\sigma}_N^2 = O_p\left(N^{\alpha-1}T^{-1/2}\right)$ otherwise. We have that $\widehat{\bar{\sigma}_N^2} = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \hat{u}_{it}^2$, where $\hat{u}_{it}$ is the estimated residual. Then, $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \hat{u}_{it}^2 = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T u_{it}^2 + \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(\hat{u}_{it}^2 - u_{it}^2\right)$. Following similar lines to those of the proof of Lemma 3 we have that $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T u_{it}^2 \to_p \bar{\sigma}_N^2$. Further, $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(u_{it}^2 - \bar{\sigma}_N^2\right) = O\left((NT)^{-1/2}\right)$. Next, we examine $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\left(\hat{u}_{it}^2 - u_{it}^2\right)$. It is sufficient to consider $\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T u_{it}\left(\hat{u}_{it} - u_{it}\right)$.

Single factor case: We note that the same residual is obtained irrespective of whether we regress x_{it} on $\bar{x}_t$, or $\tilde{x}_t$ or $N^{1-\alpha}\bar{x}_t$ or f_{1t}. We carry out the analysis by using $\tilde{x}_t$ as the regressor. We have that $\hat{u}_{it} = \frac{\tilde{x}_t\sum_{j=1}^T \tilde{x}_j u_{ij}}{\sum_{j=1}^T \tilde{x}_j^2} + u_{it}$. Then,

$$\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T u_{it}\left(\hat{u}_{it} - u_{it}\right) = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T u_{it}\tilde{x}_t\left(\frac{\sum_{j=1}^T \tilde{x}_j u_{ij}}{\sum_{j=1}^T \tilde{x}_j^2}\right) = \left(\frac{1}{NT\sum_{j=1}^T \tilde{x}_j^2}\right)\sum_{i=1}^N\left(\sum_{j=1}^T \tilde{x}_j u_{ij}\right)\left(\sum_{t=1}^T \tilde{x}_t u_{it}\right) =$$

$$\frac{1}{NT}\left(\frac{1}{\frac{1}{T}\sum_{t=1}^T \tilde{x}_t^2}\right)\sum_{i=1}^N\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2 - E\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2\right)\right) +$$

$$+\frac{1}{NT}\left(\frac{1}{\frac{1}{T}\sum_{t=1}^T \tilde{x}_t^2}\right)\sum_{i=1}^N\left(E\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2\right)\right).$$

But $E\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2\right) < \infty$ uniformly over i and $\frac{1}{T}\sum_{t=1}^T \tilde{x}_t^2 = O_p(1)$, which implies that

$$\frac{1}{NT}\left(\frac{1}{\frac{1}{T}\sum_{t=1}^T \tilde{x}_t^2}\right)\sum_{i=1}^N\left(E\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2\right)\right) = O_p\left(\frac{1}{T}\right).$$

Further, $\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2 - E\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2\right)$ is a NED process over i, which implies that

$$\frac{1}{NT}\left(\frac{1}{\frac{1}{T}\sum_{t=1}^T \tilde{x}_t^2}\right)\sum_{i=1}^N\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2 - E\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \tilde{x}_t u_{it}\right)^2\right)\right) = O_p\left(\frac{1}{T\sqrt{N}}\right),$$

proving the required result.

Multifactor case: We will focus on the case where $\alpha = \alpha_2 = ... = \alpha_m$ as the case $\alpha \geq \alpha_2 \geq ... \geq \alpha_m$ with at least one strict inequality can be treated similarly and has equal or lower rates for $\widehat{\bar{\sigma}_N^2} - \bar{\sigma}_N^2$. We have

$$\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T u_{it}\left(\hat{u}_{it} - u_{it}\right) = \frac{1}{NT}\sum_{i,\beta_{i1}\neq 0}^{N^\alpha}\sum_{t=1}^T u_{it}\left(\hat{u}_{it} - u_{it}\right) + \frac{1}{NT}\sum_{i,\beta_{i1}=0}^{N-N^\alpha}\sum_{t=1}^T u_{it}\left(\hat{u}_{it} - u_{it}\right) \qquad \text{(B14)}$$

The second term of the RHS of (B14) can be treated as in the single factor case, giving

$$\frac{1}{NT}\sum_{i,\beta_{1i}=0}^{N-N^\alpha}\sum_{t=1}^T u_{it}\left(\hat{u}_{it} - u_{it}\right) = O_p\left(T^{-1}\right) + O_p\left(\frac{1}{T\sqrt{N}}\right)$$

For the first term of the RHS of (B14), we note that x_{it} can be written as $x_{it} = \varkappa_i\tilde{x}_t + \tilde{\beta}_i'\tilde{f}_t + u_{it}$, where $\tilde{f}_t$ is a zero mean process that is uncorrelated to $\tilde{x}_t$. Then, $\hat{u}_{it} = \frac{\tilde{x}_t\sum_{j=1}^T \tilde{x}_j\left(\tilde{\beta}_i'\tilde{f}_j + u_{ij}\right)}{\sum_{j=1}^T \tilde{x}_j^2} + \tilde{\beta}_i'\tilde{f}_t + u_{it}$ and

$$\frac{1}{NT}\sum_{i,\beta_{i1}\neq 0}^{N^\alpha}\sum_{t=1}^T u_{it}\left(\hat{u}_{it} - u_{it}\right) = \frac{1}{NT}\sum_{i,\beta_{i1}\neq 0}^{N^\alpha}\sum_{t=1}^T u_{it}\left(\tilde{\beta}_i'\tilde{f}_t + u_{it}\right) + R \qquad \text{(B15)}$$

where R is of smaller order of probability than the first term of the RHS of (B15). Following similar arguments as above we obtain

$$\frac{1}{NT}\sum_{i,\beta_{i1}\neq 0}^{N^{\alpha}}\sum_{t=1}^{T}u_{it}\left(\tilde{\beta}_i'\tilde{f}_t+u_{it}\right)=O_p\left(N^{\alpha-1}T^{-1/2}\right),$$

which implies that

$$\widehat{\bar{\sigma}_N^2}-\bar{\sigma}_N^2=O_p\left(N^{\alpha-1}T^{-1/2}\right).$$

giving a lower rate of convergence than the single factor case.

Proof of Lemma 5

We have that

$$\ln(s_{f_1}^2\bar{v}_{1N}^2)-\ln(\sigma_{f_1}^2\mu_{v_1}^2)=\ln\left(\frac{s_{f_1}^2\bar{v}_{1N}^2}{\sigma_{f_1}^2\mu_{v_1}^2}\right)=\ln\left(\frac{s_{f_1}^2}{\sigma_{f_1}^2}\right)+\ln\left(\frac{\bar{v}_{1N}^2}{\mu_{v_1}^2}\right)=\left(\frac{s_{f_1}^2-\sigma_{f_1}^2}{\sigma_{f_1}^2}\right)+\left(\frac{\bar{v}_{1N}^2-\mu_{v_1}^2}{\mu_{v_1}^2}\right)+$$

$$O_p\left(\left(s_{f_1}^2-\sigma_{f_1}^2\right)^2\right)+O_p\left(\left(\bar{v}_{1N}^2-\mu_{v_1}^2\right)^2\right).$$

But, under Assumption 2, and setting $m=1$,

$$\sqrt{T}\left(\frac{s_{f_1}^2-\sigma_{f_1}^2}{\sigma_{f_1}^2}\right)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\left[(f_{1t}-\bar{f}_1)/\sigma_{f_1}\right]^2-1\right\}\rightarrow_d N\left(0,V_{f_1^2}\right),$$

where $\bar{f}_1=\frac{1}{T}\sum_{t=1}^{T}f_{1t}$, and

$$V_{f_1^2}=E\left(\left([(f_{1t}-\mu_{f_1})/\sigma_{f_1}]^2-1\right)^2\right)+\sum_{i=1}^{\infty}Cov\left(\left([(f_{1t}-\mu_{f_1})/\sigma_{f_1}]^2-1\right)\left([(f_{1,t-i}-\mu_{f_1})/\sigma_{f_1}]^2-1\right)\right).$$

Further, recalling that $\bar{v}_{1N}=\frac{1}{N^{\alpha}}\sum_{i=1}^{N^{\alpha}}v_{i1}$, $\sqrt{N^{\alpha}}\left(\frac{\bar{v}_{1N}^2-\mu_{v_1}^2}{\mu_{v_1}^2}\right)=\sqrt{N^{\alpha}}\left(\frac{\bar{v}_{1N}-\mu_{v_1}}{\mu_{v_1}}\right)\left(\frac{\bar{v}_{1N}+\mu_{v_1}}{\mu_{v_1}}\right)$. But $\frac{\bar{v}_{1N}+\mu_{v_1}}{\mu_{v_1}}\rightarrow_p 2$, and

$$\sqrt{N^{\alpha}}\left(\frac{\bar{v}_{1N}-\mu_{v_1}}{\mu_{v_1}}\right)\rightarrow_d N\left(0,\frac{\sigma_{v_1}^2}{\mu_{v_1}^2}\right). \tag{B16}$$

Further, $E\left[\left(\frac{s_{f_1}^2-\sigma_{f_1}^2}{\sigma_{f_1}^2}\right)\left(\frac{\bar{v}_{1N}^2-\mu_{v_1}^2}{\mu_{v_1}^2}\right)\right]=0$, implying that $\sqrt{\min(N^{\alpha},T)}\left(\ln(s_{f_1}^2\bar{v}_{1N}^2)-\ln(\sigma_{f_1}^2\mu_{v_1}^2)\right)\rightarrow_d N\left(0,\omega\right)$, where $\omega=\lim_{N,T\to\infty}\left[\frac{\min(N^{\alpha},T)}{T}V_{f_1^2}+\frac{\min(N^{\alpha},T)}{N^{\alpha}}\frac{4\sigma_{v_1}^2}{\mu_{v_1}^2}\right]$.

Proof of Lemma 6

The proof follows easily along the same lines as that of Lemma 5. In the present case under Assumption 5 we have $\bar{v}_{1N}=N^{-1}\sum_{i=1}^{N}v_{i1}$, and thus $\sqrt{N}\left(\frac{\bar{v}_{1N}^2-\mu_{v_1}^2}{\mu_{v_1}^2}\right)=\sqrt{N}\left(\frac{\bar{v}_{1N}-\mu_{v_1}}{\mu_{v_1}}\right)\left(\frac{\bar{v}_{1N}+\mu_{v_1}}{\mu_{v_1}}\right)$, and $\frac{\bar{v}_{1N}+\mu_{v_1}}{\mu_{v_1}}\rightarrow_p 2$. Therefore, $\sqrt{N}\left(\frac{\bar{v}_{1N}-\mu_{v_1}}{\mu_{v_1}}\right)\rightarrow_d N\left(0,\frac{\sigma_{v_1}^2}{\mu_{v_1}^2}\right)$.

Proof of Lemma 7

We need to show that

$$\sqrt{\min(N^{\alpha},T)}\left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}\right)=o_p\left(1\right). \tag{B17}$$

We have

$$\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}=\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}+\frac{\widehat{\bar{\sigma}_N^2}}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}.$$

But, by lemma 4

$$\frac{\bar{\sigma}_N^2}{N^{2\alpha-1}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N^{2\alpha-1}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}=O_p\left(T^{-1/2}N^{-2\alpha}\right), \tag{B18}$$

which is negligible as a bias. Next,

$$\frac{\widehat{\bar{\sigma}_N^2}}{N^{2\alpha-1}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}=\bar{\sigma}_N^2\left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right)\left(\frac{1}{N^{2\alpha-1}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{1}{N\hat{\sigma}_{\bar{x}}^2}\right)$$
$$=\bar{\sigma}_N^2\left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right)\left(\frac{1}{N^{2\alpha-1}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)N^{2\alpha-1}\left(N^{2-2\alpha}\hat{\sigma}_{\bar{x}}^2-\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N\right)\left(\frac{1}{N\hat{\sigma}_{\bar{x}}^2}\right).$$

But by the proof of Theorem 1, we have

$$\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N-N^{2-2\alpha}\hat{\sigma}_{\bar{x}}^2=O_p\left(T^{-1/2}N^{-2\alpha}\right)+O_p(N^{1-3\alpha})+O_p(N^{-\alpha})+O_p\left(N^{1-2\alpha}\right).$$

So,

$$\bar{\sigma}_N^2\left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right)\left(\frac{1}{N^{2\alpha-1}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)N^{2\alpha-1}\left(N^{2-2\alpha}\hat{\sigma}_{\bar{x}}^2-\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N\right)\left(\frac{1}{N\hat{\sigma}_{\bar{x}}^2}\right)=$$
$$O_p\left(T^{-1/2}N^{-2\alpha}N^{1-2\alpha}\right)+O_p(N^{1-3\alpha}N^{1-2\alpha})+O_p(N^{-\alpha}N^{1-2\alpha})+O_p\left(N^{1-2\alpha}N^{1-2\alpha}\right)=$$
$$O_p\left(T^{-1/2}N^{1-2\alpha}\right)+O_p(N^{2-5\alpha})+O_p(N^{1-3\alpha})+O_p\left(N^{2-4\alpha}\right).$$

Therefore, for $\alpha > 1/2$, (B17) holds, which establishes the Lemma.

Proof of Lemma 8

We need to show that

$$\sqrt{\min(N,T)}\left(\frac{\bar{\sigma}_N^2}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}\right)=o_p\left(1\right). \tag{B19}$$

We have

$$\frac{\bar{\sigma}_N^2}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}=\frac{\bar{\sigma}_N^2}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}+\frac{\widehat{\bar{\sigma}_N^2}}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}.$$

But, by lemma 4

$$\frac{\bar{\sigma}_N^2}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}=O_p\left(T^{-1/2}N^{-\alpha}\right),$$

which is negligible as a bias. Next,

$$\frac{\widehat{\bar{\sigma}_N^2}}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{\widehat{\bar{\sigma}_N^2}}{N\hat{\sigma}_{\bar{x}}^2}=\bar{\sigma}_N^2\left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right)\left(\frac{1}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}-\frac{1}{N\hat{\sigma}_{\bar{x}}^2}\right)$$
$$=\bar{\sigma}_N^2\left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right)\left(\frac{1}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)N^{\alpha}\left(N^{1-\alpha}\hat{\sigma}_{\bar{x}}^2-\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N\right)\left(\frac{1}{N\hat{\sigma}_{\bar{x}}^2}\right).$$

But by the proof of Theorem 1, we have

$$\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N-N^{1-\alpha}\hat{\sigma}_{\bar{x}}^2=O_p\left(T^{-1/2}N^{-\alpha/2}\right).$$

So

$$\bar{\sigma}_N^2\left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right)\left(\frac{1}{N^{\alpha}\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)N^{\alpha}\left(N^{1-\alpha}\hat{\sigma}_{\bar{x}}^2-\bar{\boldsymbol{v}}'_N\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N\right)\left(\frac{1}{N\hat{\sigma}_{\bar{x}}^2}\right)=O_p\left(T^{-1/2}N^{-3\alpha/2}\right)$$

which establishes the Lemma.

Proof of Lemma 9

We have that

$$\bar{\sigma}_N^2 \left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right)\left(\frac{1}{N^{2\alpha-1}\bar{\boldsymbol{v}}_N'\boldsymbol{D}_N\boldsymbol{S}_{ff}\boldsymbol{D}_N\bar{\boldsymbol{v}}_N}\right)N^{2\alpha-1}\left(N^{1-2\alpha}\left(\frac{1}{T}\sum_{t=1}^{T}(\sqrt{N}\bar{u}_t)^2-\widehat{\bar{\sigma}_N^2}\right)\right)\left(\frac{1}{N\hat{\sigma}_{\bar{x}}^2}\right)=O_p\left(N^{2-4\alpha}T^{-1/2}\right),$$

which is negligible as long as $\alpha > 1/2$. To prove the above result we first note that

$$\frac{1}{T}\sum_{t=1}^{T}\left(\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}u_{it}\right)^2\right)-\widehat{\bar{\sigma}_N^2}=$$

$$\left(\frac{1}{T}\sum_{t=1}^{T}\left(\left(\frac{1}{\sqrt{N}}\sum_{t=1}^{N}u_{it}\right)^2\right)-\bar{\sigma}_N^2\right)+\left(\bar{\sigma}_N^2-\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{it}^2\right)-\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{u}_{it}^2-\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{it}^2\right).$$

But it is straightforward to show that $\frac{1}{T}\sum_{t=1}^{T}\left(\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}u_{it}\right)^2\right)-\bar{\sigma}_N^2=O_p(T^{-1/2})$ and $\bar{\sigma}_N^2-\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{it}^2=O_p(T^{-1/2})$. Finally, by Lemma 4, $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{u}_{it}^2-\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}u_{it}^2=o_p(T^{-1/2})$. So, $\frac{1}{T}\sum_{t=1}^{T}\left(\left(\frac{1}{\sqrt{N}}\sum_{t=1}^{N}u_{it}\right)^2\right)-\widehat{\bar{\sigma}_N^2}=O_p(T^{-1/2})$.

Proof of Lemma 10

We have that

$$\frac{1}{T}\sum_{t=1}^{T}\bar{x}_t^2=\frac{1}{T}\sum_{t=1}^{T}\left[\sum_{j=1}^{m}\left(f_{jt}\frac{N^{\alpha_j}}{N}\frac{1}{N^{\alpha_j}}\sum_{i=1}^{N^{\alpha_j}}v_{ij}\right)+\frac{1}{N}\sum_{i=1}^{N}u_{it}\right]^2=$$

$$\frac{1}{T}\sum_{t=1}^{T}\left\{\sum_{j=1}^{m}\left[f_{jt}^2\frac{N^{2\alpha_j}}{N^2}\left(\frac{1}{N^{\alpha_j}}\sum_{i=1}^{N^{\alpha_j}}v_{ij}\right)^2\right]\right\}+$$

$$\frac{1}{T}\sum_{t=1}^{T}\left\{\sum_{j=1,j\neq s}^{m}\left[f_{jt}\frac{N^{\alpha_j}}{N}\left(\frac{1}{N^{\alpha_j}}\sum_{i=1}^{N^{\alpha_j}}v_{ij}\right)\right]\sum_{s=1}^{m}\left[f_{st}\frac{N^{\alpha_s}}{N}\left(\frac{1}{N^{\alpha_s}}\sum_{i=1}^{N^{\alpha_s}}v_{is}\right)\right]\right\}+$$

$$\frac{1}{T}\sum_{t=1}^{T}\left\{\left[\sum_{j=1}^{m}\left(f_{jt}\frac{N^{\alpha_j}}{N}\frac{1}{N^{\alpha_j}}\sum_{i=1}^{N^{\alpha_j}}v_{ij}\right)\right]\left(\frac{1}{N}\sum_{i=1}^{N}u_{it}\right)\right\}+$$

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}u_{it}\right)^2.$$

But,

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}u_{it}\right)^2=O_p\left(N^{-1}\right)$$

$$\frac{1}{T}\sum_{t=1}^{T}\left[\left(f_{jt}\frac{N^{\alpha_j}}{N}\frac{1}{N^{\alpha_j}}\sum_{i=1}^{N^{\alpha_j}}v_{ij}\right)\left(\frac{1}{N}\sum_{i=1}^{N}u_{it}\right)\right]=O_p\left(N^{\alpha_j-3/2}T^{-1/2}\right),\ j=1,...,m$$

$$\frac{1}{T}\sum_{t=1}^{T}\left(\left(f_{jt}\frac{N^{\alpha_j}}{N}\frac{1}{N^{\alpha_j}}\sum_{i=1}^{N^{\alpha_j}}v_{ij}\right)\left(f_{st}\frac{N^{\alpha_s}}{N}\frac{1}{N^{\alpha_s}}\sum_{i=1}^{N^{\alpha_s}}v_{is}\right)\right)=O_p\left(N^{\alpha_j+\alpha_s-2}T^{-1/2}\right),\ j,s=1,...,m,\quad j\neq s,$$

and

$$\frac{1}{T}\sum_{t=1}^{T}\left(f_{jt}^2\frac{N^{2\alpha_j}}{N^2}\left(\frac{1}{N^{\alpha_j}}\sum_{i=1}^{N^{\alpha_j}}v_{ij}\right)^2\right)-N^{2\alpha_j-2}\mu_{v_j}^2\rightarrow_p 0,\ j=1,...,m.$$

If $\alpha = \alpha_2 = ... = \alpha_{q-1} > \alpha_q \geq ... \geq \alpha_m$

$$\frac{1}{T}\sum_{t=1}^{T}\bar{x}_t^2 - N^{2\alpha-2}\sum_{j=1}^{q}\mu_{v_j}^2 \rightarrow_p 0.$$

In particular if $\alpha > \alpha_2 \geq ... \geq \alpha_m$

$$\frac{1}{T}\sum_{t=1}^{T}\bar{x}_t^2 - N^{2\alpha-2}\mu_{v_1}^2 \rightarrow_p 0.$$

Proof of Lemma 11

Without loss of generality we consider the case of two factors. The result extends straightforwardly to m factors. We further assume, for simplicity, that factors are independent from each other. Then,

$$\ln\left(\bar{v}_{1N}^2 s_{f_1}^2 + 2\bar{v}_{1N}\bar{v}_{2N}s_{12,f} + \bar{v}_{2N}^2 s_{f_2}^2\right) - \ln\left(\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2\right) = \ln\left(\frac{\bar{v}_{1N}^2 s_{f_1}^2 + 2\bar{v}_{1N}\bar{v}_{2N}s_{12,f} + \bar{v}_{2N}^2 s_{f_2}^2}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2}\right).$$

Then,

$$\ln\left(\frac{\bar{v}_{1N}^2 s_{f_1}^2 + 2\bar{v}_{1N}\bar{v}_{2N}s_{12,f} + \bar{v}_{2N}^2 s_{f_2}^2}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2}\right) = \frac{\bar{v}_{1N}^2 s_{f_1}^2 + 2\bar{v}_{1N}\bar{v}_{2N}s_{12,f} + \bar{v}_{2N}^2 s_{f_2}^2}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} - 1 = \tag{B20}$$

$$\frac{\left(\bar{v}_{1N}^2 s_{f_1}^2 - \sigma_{f_1}^2\mu_{v_1}^2\right) + \left(\bar{v}_{2N}^2 s_{f_2}^2 - \sigma_{f_2}^2\mu_{v_2}^2\right) + 2\bar{v}_{1N}\bar{v}_{2N}s_{12,f}}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} =$$

$$\frac{\left(\bar{v}_{1N}^2 s_{f_1}^2 - \bar{v}_{1N}^2\sigma_{f_1}^2 + \bar{v}_{1N}^2\sigma_{f_1}^2 - \sigma_{f_1}^2\mu_{v_1}^2\right) + \left(\bar{v}_{2N}^2 s_{f_2}^2 - \bar{v}_{2N}^2\sigma_{f_2}^2 + \bar{v}_{2N}^2\sigma_{f_2}^2 - \sigma_{f_2}^2\mu_{v_2}^2\right) + 2\mu_{v_1}\mu_{v_2}s_{12,f}}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} =$$

$$\frac{\mu_{v_1}^2\left(s_{f_1}^2 - \sigma_{f_1}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} + \frac{\sigma_{f_1}^2\left(\bar{v}_{1N}^2 - \mu_{v_1}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} + \frac{\mu_{v_2}^2\left(s_{f_2}^2 - \sigma_{f_2}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} + \frac{\sigma_{f_2}^2\left(\bar{v}_{2N}^2 - \mu_{v_2}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} + \frac{2\mu_{v_1}\mu_{v_2}s_{12,f}}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2}.$$

Note that

$$\bar{v}_{1N}\bar{v}_{2N}s_{12,f} = \bar{v}_{1N}\bar{v}_{2N}s_{12,f} - \bar{v}_{1N}\mu_{v_2}s_{12,f} + \bar{v}_{1N}\mu_{v_2}s_{12,f} - \bar{v}_{1N}\mu_{v_2}\sigma_{12,f} + \bar{v}_{1N}\mu_{v_2}\sigma_{12,f} - 2\mu_{v_1}\mu_{v_2}\sigma_{12,f} =$$

$$s_{12,f}\bar{v}_{1N}\left(\bar{v}_{2N} - \mu_{v_2}\right) + \bar{v}_{1N}\mu_{v_2}\left(s_{12,f} - \sigma_{12,f}\right) + \sigma_{12,f}\mu_{v_2}\left(\bar{v}_{1N} - 2\mu_{v_1}\right) =$$

$$\left(s_{12f} - \sigma_{12,f}\right)\bar{v}_{1N}\left(\bar{v}_{2N} - \mu_{v_2}\right) + \sigma_{12,f}\bar{v}_{1N}\left(\bar{v}_{2N} - \mu_{v_2}\right) + \bar{v}_{1N}\mu_{v_2}\left(s_{12,f} - \sigma_{12,f}\right) + \sigma_{12,f}\mu_{v_2}\left(\bar{v}_{1N} - 2\mu_{v_1}\right).$$

But

$$\left(s_{12f} - \sigma_{12,f}\right)\bar{v}_{1N}\left(\bar{v}_{2N} - \mu_{v_2}\right) = o_p(T^{-1/2}),$$

and $\sigma_{12,} = 0$, and so

$$\left(s_{12,f} - \sigma_{12,f}\right)\bar{v}_{1N}\left(\bar{v}_{2N} - \mu_{v_2}\right) + \sigma_{12,f}\bar{v}_{1N}\left(\bar{v}_{2N} - \mu_{v_2}\right) + \bar{v}_{1N}\mu_{v_2}\left(s_{12,f} - \sigma_{12,f}\right) + \sigma_{12,f}\mu_{v_2}\left(\bar{v}_{1N} - 2\mu_{v_1}\right)$$
$$= \bar{v}_{1N}\mu_{v_2}s_{12,f} = \left(\frac{\bar{v}_{1N}}{\mu_{v_1}}\right)\mu_{v_1}\mu_{v_2}s_{12,f} + o_p\left(T^{-1/2}\right).$$

Then,

$$\frac{\mu_{v_i}^2\left(s_{f_i}^2 - \sigma_{f_i}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} = \frac{\mu_{v_i}^2\sigma_{f_i}^2}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2}\frac{\left(s_{f_i}^2 - \sigma_{f_i}^2\right)}{\sigma_{f_i}^2}, \quad i = 1, 2,$$
$$\frac{\sigma_{f_i}^2\left(\bar{v}_{iN}^2 - \mu_{v_i}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2} = \frac{\mu_{v_i}^2\sigma_{f_i}^2}{\sigma_{f_1}^2\mu_{v_1}^2 + \sigma_{f_2}^2\mu_{v_2}^2}\frac{\left(\bar{v}_{iN}^2 - \mu_{v_i}^2\right)}{\mu_{v_i}^2}, \quad i = 1, 2.$$

Assuming loadings of factors and factors are independent of each other and across factors, gives

$$\mu_{v_i}^2\sigma_{f_i}^2\left(\sqrt{T}\frac{\left(s_{f_i}^2 - \sigma_{f_i}^2\right)}{\sigma_{f_i}^2}\right) = \mu_{v_i}^2\sigma_{f_i}^2\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\left[(f_{it} - \bar{f}_i)/\sigma_{f_i}\right]^2 - 1\right\}\right) \rightarrow_d N\left(0, \left(\mu_{v_i}^2\sigma_{f_i}^2\right)^2\mu_i^{(4)}\right), \quad i = 1, 2,$$

$$\mu_{v_i}^2\sigma_{f_i}^2\left(\sqrt{N^\alpha}\left(\frac{\bar{v}_{iN}^2-\mu_{v_i}^2}{\mu_{v_i}^2}\right)\right)=\mu_{v_i}^2\sigma_{f_i}^2\left(\sqrt{N^\alpha}\left(\frac{\bar{v}_{iN}-\mu_{v_i}}{\mu_{v_i}}\right)\left(\frac{\bar{v}_{iN}+\mu_{v_i}}{\mu_{v_i}}\right)\right)\to_d N\left(0,4\sigma_{v_i}^2\mu_{v_i}^2\left(\sigma_{f_i}^2\right)^2\right),\quad i=1,2.$$

Further,

$$\mu_{v_1}\mu_{v_2}\sqrt{T}s_{12,f}=\mu_{v_1}\mu_{v_2}\frac{\sigma_{f_1}\sigma_{f_2}}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{f_{1t}-\bar{f}_1}{\sigma_{f_1}}\right)\left(\frac{f_{2t}-\bar{f}_2}{\sigma_{f_2}}\right)\to_d N\left(0,\mu_{v_1}^2\sigma_{f_1}^2\mu_{v_2}^2\sigma_{f_2}^2\right).$$

Further, by factor independence

$$E\left(\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\left[(f_{it}-\bar{f}_i)/\sigma_{f_i}\right]^2-1\right\}\right]\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{f_{1t}-\bar{f}_1}{\sigma_{f_1}}\right)\left(\frac{f_{2t}-\bar{f}_2}{\sigma_{f_2}}\right)\right]\right)=0,\quad i=1,2.$$

So

$$\sqrt{\frac{\min\left(N^\alpha,T\right)}{T}}\left(\mu_{v_1}^2\sigma_{f_1}^2\left(\sqrt{T}\frac{\left(s_{f_1}^2-\sigma_{f_1}^2\right)}{\sigma_{f_1}^2}\right)+\mu_{v_2}^2\sigma_{f_2}^2\left(\sqrt{T}\frac{\left(s_{f_2}^2-\sigma_{f_2}^2\right)}{\sigma_{f_2}^2}\right)+2\mu_{v_1}\mu_{v_2}\sqrt{T}s_{12,f}\right)+$$
$$\sqrt{\frac{\min\left(N^\alpha,T\right)}{N^\alpha}}\left(\mu_{v_1}^2\sigma_{f_1}^2\left(\sqrt{N^\alpha}\left(\frac{\bar{v}_{1N}^2-\mu_{v_1}^2}{\mu_{v_1}^2}\right)\right)+\mu_{v_2}^2\sigma_{f_2}^2\left(\sqrt{N^\alpha}\left(\frac{\bar{v}_{2N}^2-\mu_{v_2}^2}{\mu_{v_2}^2}\right)\right)\right)\to_d$$
$$N\left(\begin{array}{cc}0, & \frac{\min(N^\alpha,T)}{T}\left(\left(\mu_{v_1}^2\sigma_{f_1}^2\right)^2\mu_1^{(4)}+\left(\mu_{v_2}^2\sigma_{f_2}^2\right)^2\mu_2^{(4)}+4\mu_{v_1}^2\left(\sigma_{f_1}^2\right)^2\mu_{v_2}^2\left(\sigma_{f_2}^2\right)^2\right)\\ & +\frac{\min(N^\alpha,T)}{N^\alpha}\left(2\sigma_{v_1}^2\mu_{v_1}^2\sigma_{f_1}^2+2\sigma_{v_2}^2\mu_{v_2}^2\sigma_{f_2}^2\right)\end{array}\right).$$

Proof of Lemma 12

Again, without loss of generality we look at the case of two factors. The result again extends straightforwardly. We further assume, for simplicity, that factors are independent from each other. Then,

$$\ln\left(\bar{v}_{1N}^2s_{f_1}^2+2N^{\alpha_2-\alpha}\bar{v}_{1N}\bar{v}_{2N}s_{12,f}+N^{2(\alpha_2-\alpha)}\bar{v}_{2N}^2s_{f_2}^2\right)-\ln\left(\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2\right)=$$
$$\ln\left(\frac{\bar{v}_{1N}^2s_{f_1}^2+2N^{\alpha_2-\alpha}\bar{v}_{1N}\bar{v}_{2N}s_{12,f}+N^{2(\alpha_2-\alpha)}\bar{v}_{2N}^2s_{f_2}^2}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}\right).$$

Then, similarly to the proof of Lemma 11

$$\ln\left(\frac{\bar{v}_{1N}^2s_{f_1}^2+2N^{\alpha_2-\alpha}\bar{v}_{1N}\bar{v}_{2N}s_{12,f}+N^{2(\alpha_2-\alpha)}\bar{v}_{2N}^2s_{f_2}^2}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}\right)=\frac{\mu_{v_1}^2\left(s_{f_1}^2-\sigma_{f_1}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}+\frac{\sigma_{f_1}^2\left(\bar{v}_{1N}^2-\mu_{v_1}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}+$$
$$\frac{N^{2(\alpha_2-\alpha)}\mu_{v_2}^2\left(s_{f_2}^2-\sigma_{f_2}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}+\frac{N^{2(\alpha_2-\alpha)}\sigma_{22,f}\left(\bar{v}_{2N}^2-\mu_{v_2}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}+\frac{2N^{\alpha_2-\alpha}\mu_{v_1}\mu_{v_2}s_{12,f}}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}.$$

Then,

$$\frac{\mu_{v_1}^2\left(s_{f_1}^2-\sigma_{f_1}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}=\frac{\mu_{v_1}^2\sigma_{f_1}^2}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}\frac{\left(s_{f_1}^2-\sigma_{f_1}^2\right)}{\sigma_{f_1}^2},$$

$$\frac{\sigma_{f_1}^2\left(\bar{v}_{1N}^2-\mu_{v_1}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}=\frac{\mu_{v_1}^2\sigma_{f_1}^2}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}\frac{\left(\bar{v}_{1N}^2-\mu_{v_1}^2\right)}{\mu_{v_1}^2},$$

$$\frac{N^{2(\alpha_2-\alpha)}\mu_{v_2}^2\left(s_{f_2}^2-\sigma_{f_2}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}=\frac{\mu_{v_2}^2\sigma_{f_2}^2}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}\frac{N^{2(\alpha_2-\alpha)}\left(s_{f_2}^2-\sigma_{f_2}^2\right)}{\sigma_{f_2}^2},\tag{B21}$$

$$\frac{N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\left(\bar{v}_{2N}^2-\mu_{v_2}^2\right)}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}=\frac{\mu_{v_2}^2\sigma_{f_2}^2}{\sigma_{f_1}^2\mu_{v_1}^2+N^{2(\alpha_2-\alpha)}\sigma_{f_2}^2\mu_{v_2}^2}\frac{N^{2(\alpha_2-\alpha)}\left(\bar{v}_{2N}^2-\mu_{v_2}^2\right)}{\mu_{v_2}^2}.\tag{B22}$$

But, then it is obvious that the Lemma holds since (B21) and (B22) are $o_p(1)$, when multiplied by $\min\left(\sqrt{T},\sqrt{N^\alpha}\right)$ respectively, as well as $\min\left(\sqrt{T},\sqrt{N^\alpha}\right)N^{\alpha_2-\alpha}\mu_{v_1}\mu_{v_2}s_{12,f}$.

Proof of Lemma 13

We analyse the population counterpart of $\ln\left(\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N\right)$ assuming for simplicity that $\boldsymbol{\Sigma}_{ff}$ is diagonal and $\alpha > \alpha_2 \geq \alpha_3 \geq ... \geq \alpha_m$. We have

$$\ln(\boldsymbol{\mu}_v'\mathbf{D}_N\boldsymbol{\Sigma}_{ff}\mathbf{D}_N\boldsymbol{\mu}_v) = \ln\left(\mu_{v_1}^2\sigma_{11,f} + N^{2(\alpha_2-\alpha)}\sum_{j=2}^{m}N^{2(\alpha_j-\alpha_2)}\mu_{v_j}^2\sigma_{f_j}^2\right).$$

Then,

$$\ln(\boldsymbol{\mu}_v'\mathbf{D}_N\boldsymbol{\Sigma}_{ff}\mathbf{D}_N\boldsymbol{\mu}_v) - \ln\left(\mu_{v_1}^2\sigma_{f_1}^2\right) = \ln\left(1 + \frac{N^{2(\alpha_2-\alpha)}\sum_{j=2}^{m}\left(N^{2(\alpha_j-\alpha_2)}\mu_{v_j}^2\sigma_{f_j}^2\right)}{\mu_{v_1}^2\sigma_{f_1}^2}\right) = \frac{\sum_{j=2}^{m}\left(N^{2(\alpha_j-\alpha_2)}\mu_{v_j}^2\sigma_{f_j}^2\right)}{\mu_{v_1}^2\sigma_{f_1}^2}N^{2(\alpha_2-\alpha)}.$$

So,

$$\sqrt{\min(N^\alpha,T)}\ln(N)\left(\ln(\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N) - \ln(\boldsymbol{\mu}_v'\mathbf{D}_N\boldsymbol{\Sigma}_{ff}\mathbf{D}_N\boldsymbol{\mu}_v)\right) =$$

$$\sqrt{\min(N^\alpha,T)}\ln(N)\left[\ln(\bar{\boldsymbol{v}}_N'\mathbf{D}_N\mathbf{S}_{ff}\mathbf{D}_N\bar{\boldsymbol{v}}_N) - \ln(\mu_{v_1}^2\sigma_{f_1}^2)\right] - \sqrt{\min(N^\alpha,T)}\ln(N)\,N^{2(\alpha_2-\alpha)}\left(\frac{\sum_{j=2}^{m}\left(N^{2(\alpha_j-\alpha_2)}\mu_{v_j}^2\sigma_{f_j}^2\right)}{\mu_{v_1}^2\sigma_{f_1}^2}\right).$$

We need

$$N^{2(\alpha_2-\alpha)}\left(\frac{\sum_{j=2}^{m}\left(N^{2(\alpha_j-\alpha_2)}\mu_{v_j}^2\sigma_{f_j}^2\right)}{\mu_{v_1}^2\sigma_{f_1}^2}\right) = o\left(\min\left(N^\alpha,T\right)^{-1/2}\ln(N)^{-1}\right).$$

This holds if $\sqrt{\min(N^a,T)}N^{2(\alpha_2-\alpha)} = o(1)$. If $T < N^\alpha$ then a sufficient condition for the above to hold is $\alpha_2 - \alpha < -0.25$. Otherwise, the sufficient condition is $\alpha_2 < 3\alpha/4$. But, this condition is implied by $\alpha_2 - \alpha < -0.25$ as long as $\alpha \leq 1$. An alternative condition that relates to the relative rate of growth of N and T is that $\alpha_2 < 3\alpha/4$ and $T^b = N$ and $1/(4b) + \alpha_2 - \alpha < 0$ or $b > \frac{1}{4(\alpha-\alpha_2)}$

Proof of Lemma 14

We note that the first part of the Lemma holds if

$$\frac{\ln(s_{f_1}^2\bar{v}_{N1}^2)}{\ln(N)} = o_p(1). \tag{B23}$$

We have

$$\ln(s_{f_1}^2\bar{v}_{N1}^2) = \ln\left(s_{f_1}^2\right) + 2\ln\left(\bar{v}_{N1}\right) = \ln\left(s_{f_1}^2\right) + 2\ln\left(\frac{1}{N}\sum_{i=1}^{N}\breve{v}_i + \bar{c}_N\right).$$

So (B23) holds if $\frac{1}{N}\sum_{i=1}^{N}\breve{v}_i + \bar{c}_N = o_p(N^c)$ for all $c > 0$, which holds if $\bar{c}_N = o_p(N^c)$ for all $c > 0$, proving the first part of the Lemma. For the second part of the Lemma we reconsider (B16). We have $\sqrt{N}\left(\bar{v}_{N1} - \mu_{\breve{v}}\right) = \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\breve{v}_i + \bar{c}_N - \mu_{\breve{v}}\right)$. But, $\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\breve{v}_i - \mu_{\breve{v}}\right) \to_d N\left(0,\sigma_{\breve{v}}^2\right)$. Therefore, $\sqrt{N}\bar{c}_N = o_p(1)$ is sufficient for the second part of the Lemma to hold.

Proof of Lemma 15

We have that

$$\frac{N^\alpha - N^{\hat{\alpha}}}{N^\alpha} = 1 - \frac{N^{\hat{\alpha}}}{N^\alpha} = \ln\left(\frac{N^{\hat{\alpha}}}{N^\alpha}\right) + o_p\left(\ln\left(\frac{N^{\hat{\alpha}}}{N^\alpha}\right)\right).$$

Then,

$$\ln\left(\frac{N^{\hat{\alpha}}}{N^\alpha}\right) = (\hat{\alpha} - \alpha)\ln N,$$

implying the result of the Lemma.

Proof of Lemma 16

The factor loadings of the cross-sectional units are partitioned into two groups by Assumption 1 and setting $m = 1$. The first group has non-zero loadings, denoted by v_{i1}, while the second group has loadings that are summable over the group. We do not observe the partition and need to estimate it. For this reason, we rank the estimated loadings as discussed in the statement of the Lemma. The first step in the proof is to show that the number of cross-sectional units that are misclassified, i.e., that are included in the variance calculation when their loading is not a function of any v_{i1}, is $o_p(N^{\alpha})$. The first thing to note is that we abstract from the possibility that any $v_{i1} = 0$. By the fact that $\Pr(v_{i1} = 0) = 0$, it follows that the number of units with $v_{i1} = 0$ is $o_p(N^{\alpha})$. Without loss of generality, we further assume that units whose loadings do not depend on any v_{i1} have zero loadings. There are two sources of errors in partitioning the loadings. The first arises because $N^{\hat{a}}$ is not equal to N^{a}. But by Lemma 15 this error is $o_p(N^{\alpha})$ if $\hat{\alpha} - \alpha = o_p\left((\ln N)^{-1}\right)$ which is the case under the conditions of the Lemma. The fact that $\hat{\alpha} - \alpha = o_p\left((\ln N)^{-1}\right)$ justifies using the true α rather than the estimated one throughout the rest of the proof. The second source of error arises from the possibility that units are missclassified. We consider this source next assuming the true value of α is used. We know that the probability that any unit's coefficient is $\epsilon > 0$ away from its true value is of the order of N^{-a} (by Lemma (17) and the Markov inequality). We know that N^{a} units can be misclassified only if the estimated coefficients of any unordered and without replacement, sample of size N^{a} from the N units, jointly exceed their true value by ϵ. We know that since the v_{i1} are independent, that the event that an estimated coefficient will be away from its true value will be independent from the same event for another unit. So the probability that a given set of N^{a} units can be jointly misclassified is bounded from above by $N^{-aN^{a}}$. There are $\frac{N!}{N^{a}!(N-N^{a})!}$ such sets. So the probability that any set will behave thus, is bounded from above by $\frac{N^{-aN^{a}}N!}{N^{a}!(N-N^{a})!}$. We need to aggregate across $i = N^{a}, ..., N$. So overall the probability is bounded from above by $\sum_{b>a} \frac{N^{-aN^{b}}N!}{N^{b}!\left(N-N^{b}\right)!}$. We replace this by $(N - N^{a}) \frac{N^{-aN^{a}}N!}{N^{a}!(N-N^{a})!}$ and justify this step below. We have

$$(N - N^{a}) \frac{N^{-aN^{a}}N!}{N^{a}!\,(N-N^{a})!} = \frac{N^{-aN^{a}}N!}{N^{a}!\,(N-N^{a}-1)!}. \tag{B24}$$

We need the logarithm of the above quantity to have a limit of $-\infty$. We have using repeatedly Stirling's formula that ($\sim$ denotes equality up to an order of magnitude lower that any included terms)

$$\ln\left(\frac{N^{-aN^{a}}N!}{N^{a}!\,(N-N^{a}-1)!}\right) = \ln\left(N^{-aN^{a}}N!\right) - \ln\left(N^{a}!\,(N-N^{a}-1)!\right) =$$

$$-aN^{a}\ln(N) + \ln(N!) - \ln(N^{a}!) - \ln((N-N^{a}-1)!) \sim$$

$$-aN^{a}\ln(N) + N\ln(N) - N - N^{a}\ln(N^{a}) + N^{a} - (N-N^{a}-1)\ln(N-N^{a}-1) + (N-N^{a}-1) \sim$$

$$-aN^{a}\ln(N) + N\ln(N) - N - aN^{a}\ln(N) + N^{a} - (N-N^{a}-1)\ln(N-N^{a}-1) + (N-N^{a}-1) =$$

$$-aN^{a}\ln(N) + N\ln(N) - N - aN^{a}\ln(N) + N^{a} - (N-N^{a}-1)\ln\left(N(1-N^{a-1}-N^{-1})\right) + (N-N^{a}-1) =$$

$$-aN^{a}\ln(N) + N\ln(N) - N - aN^{a}\ln(N) + N^{a} - (N-N^{a}-1)\ln(N) -$$

$$-(N-N^{a}-1)\ln\left(1-N^{a-1}-N^{-1}\right) + (N-N^{a}-1) =$$

$$-aN^{a}\ln(N) - aN^{a}\ln(N) + N^{a}\ln(N) + N\ln(N) - N + N^{a} - N\ln(N) + \ln(N) -$$

$$+N^{a} + 1 - N^{2a-1} - N^{a-1} - N^{a-1} - N^{-1} + (N-N^{a}-1) =$$

$$-(2a-1)N^{a}\ln(N) - N + N^{a} + \ln(N) -$$

$$-(N-N^{a}-1)\left(-N^{a-1}-N^{-1}\right) + (N-N^{a}-1).$$

The term $-(2a-1)N^{a}\ln(N)$ dominates other terms and tends to $-\infty$, as $N \to \infty$, for $a > 1/2$, proving the result. We now justify replacing $(N - N^{a}) \frac{N^{-aN^{a}}N!}{N^{a}!(N-N^{a})!}$ for $\sum_{b>a} \frac{N^{-aN^{b}}N!}{N^{b}!\left(N-N^{b}\right)!}$ in (B24). We have

$$\ln\left(\frac{N^{-aN^{b}}N!}{N^{b}!\,(N-N^{b})!}\right) = \ln\left(N^{-aN^{b}}N!\right) - \ln\left(N^{b}!\left(N-N^{b}\right)!\right) =$$

$$-aN^b\ln(N)+\ln(N!)-\ln\left(N^b!\right)-\ln\left(\left(N-N^b-1\right)!\right)\sim$$
$$-aN^b\ln(N)+N\ln(N)-N-N^b\ln\left(N^b\right)+N^b-\left(N-N^b-1\right)\ln\left(N-N^b-1\right)+\left(N-N^b-1\right)\sim$$
$$-aN^b\ln(N)+N\ln(N)-N-bN^b\ln(N)+N^b-\left(N-N^b-1\right)\ln\left(N-N^b-1\right)+\left(N-N^b-1\right)\sim$$
$$-aN^b\ln(N)+N\ln(N)-N-bN^b\ln(N)+N^b-\left(N-N^b-1\right)\ln\left(N(1-N^{b-1}-N^{-1})\right)+\left(N-N^b-1\right)=$$
$$-aN^b\ln(N)+N\ln(N)-N-bN^b\ln(N)+N^b-\left(N-N^b-1\right)\ln(N)-$$
$$-\left(N-N^b-1\right)\ln\left(1-N^{b-1}-N^{-1}\right)+\left(N-N^b-1\right).$$

The dominant term here is $-(a+b-1)N^a\ln(N)$ which for $b>a>1/2$ is tending to $-\infty$ faster than $-(2a-1)N^a\ln(N)$ justifying the replacement.

Next, we prove the Lemma assuming that we observe which units have non-zero loadings. Recall that, assuming that units whose loadings do not depend on any v_{i1} have zero loadings, $x_{it}=\frac{v_{i1}}{\bar{v}_{1N}}\left(N^{1-\alpha}\bar{x}_t\right)+u_{it}$. We analyse $\hat{v}_{i1}$ by a slight abuse of notation whereby we define it to be the estimated regression coefficient of the regression of x_{it} on $N^{1-\alpha}\bar{x}_t$ rather than x_{it} on $\tilde{x}_t$. Since μ_{v_1} is assumed known, $\bar{v}_{1N}\to_p\mu_{v_1}$ and $\hat{\sigma}^2_{\bar{x}}\to_p N^{2\alpha-2}\mu^2_{v_1}$, by Lemma 10, this does not affect the analysis. Let $v^{(1)}_{i1}=\frac{v_{i1}}{\mu_{v_1}}$and $v_{iN}=\frac{v_{i1}}{\bar{v}_{1N}}$. We need to show that $\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(\hat{v}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}\hat{v}_{j1}\right)^2-\frac{\sigma^2_{v_1}}{\mu^2_{v_1}}=o_p(1)$. We have

$$\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(\hat{v}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}\hat{v}_{j1}\right)^2-\frac{\sigma^2_{v_1}}{\mu^2_{v_1}}=\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(\hat{v}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}\hat{v}_{j1}\right)^2-\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v_{iN}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v_{jN}\right)^2+$$
$$\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v_{iN}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v_{jN}\right)^2-\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v^{(1)}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v^{(1)}_{j1}\right)^2+\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v^{(1)}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v^{(1)}_{j1}\right)^2-\frac{\sigma^2_{v_1}}{\mu^2_{v_1}}.$$

But by the law of large numbers for i.i.d. random variables with finite variance

$$\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v^{(1)}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v^{(1)}_{j1}\right)^2-\frac{\sigma^2_{v_1}}{\mu^2_{v_1}}=O_p\left(N^{-1/2}\right).$$

It is sufficient to show that

$$\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(\hat{v}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}\hat{v}_{j1}\right)^2-\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v_{iN}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v_{jN}\right)^2=o_p(1)\tag{B25}$$

and

$$\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v_{iN}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v_{jN}\right)^2-\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v^{(1)}_{i1}-\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v^{(1)}_{j1}\right)^2=o_p(1).\tag{B26}$$

For (B25), it is sufficient that $\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(\hat{v}_{i1}-v_{iN}\right)=o_p(1)$. Recall that $x_{it}=\frac{v_{i1}}{\bar{v}_{1N}}\left(N^{1-\alpha}\bar{x}_t\right)+u_{it}$. So

$$\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(\hat{v}_{i1}-v_{iN}\right)=\frac{\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(\sum_{t=1}^T\bar{x}_tu_{it}\right)}{\sum_{t=1}^T\bar{x}^2_t}=\frac{1}{\left(N^\alpha-1\right)\bar{v}_{1N}\left(\sum_{t=1}^Tf^2_{1t}\right)}\sum_{i=1}^{N^\alpha}\sum_{t=1}^Tf_{1t}u_{it}.$$

But $\sum_{i=1}^{N^\alpha}\sum_{t=1}^Tf_{1t}u_{it}=O_p\left((N^\alpha T)^{1/2}\right)$ and $\sum_{t=1}^Tf^2_{1t}=O_p(T)$. So

$$\frac{\sum_{i=1}^{N^\alpha}\left(\sum_{t=1}^T\bar{x}_tu_{it}\right)}{N^\alpha-1\left(\sum_{t=1}^T\bar{x}^2_t\right)}=O_p\left(T^{-1}N^{-\alpha}\left(N^\alpha T\right)^{1/2}\right)=O_p\left(T^{-1/2}N^{-\alpha/2}\right)=o_p(1).$$

For (B26), it is sufficient that $\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v_{iN}-v_{i1}^{(1)}\right)=o_p(1)$. We have,

$$\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}\left(v_{iN}-v_{i1}^{(1)}\right)=\frac{\left(\frac{1}{\bar{v}_{1N}}-\frac{1}{\mu_{v_1}}\right)\sum_{i=1}^{N^\alpha}v_{i1}}{N^\alpha-1}.$$

But $\frac{1}{N^\alpha-1}\sum_{i=1}^{N^\alpha}v_{i1}=O_p(1)$. Also $\frac{1}{\bar{v}_{1N}}-\frac{1}{\mu_{v_1}}=\frac{1}{\bar{v}_{1N}\mu_{v_1}}\left(\bar{v}_{1N}-\mu_{v_1}\right)=O_p\left(N^{-\alpha/2}\right)$. So, overall $\frac{1}{N-1}\sum_{i=1}^{N^\alpha}\left(\hat{v}_{i1}-\frac{1}{N}\sum_{j=1}^{N}\hat{v}_{j1}\right)^2-\frac{\sigma_{v_1}^2}{\mu_{v_1}^2}=o_p(1)$, proving the required result.

Proof of Lemma 17

Recall that $\beta_{i1}=v_{i1}$ for $i=1,2,...,N^\alpha$ and 0 for $i=N^\alpha+1,...,N$ (without loss of generality). Here we set $m=1$. Let

$$\tilde{x}_t=\frac{1}{N^\alpha}\sum_{i=1}^{N}x_{it},$$

where we have used the normalisation $N^{-\alpha}$ to ensure that $\bar{x}_t$ converges to f_{1t}. We have

$$\hat{\beta}_{i1}=\frac{\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_tx_{it}}{\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_t^2}. \tag{B27}$$

Then,

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_tx_{it}=\frac{1}{TN^\alpha}\sum_{t=1}^{T}\sum_{j=1}^{N}x_{jt}x_{it}=$$

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N^\alpha}\sum_{j=1}^{N}\left(\beta_{j1}f_{1t}+u_{jt}\right)\left(\beta_{i1}f_{1t}+u_{it}\right)=$$

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N^\alpha}\sum_{j=1}^{N}\left(\beta_{j1}\beta_{i1}f_{1t}^2+2\beta_{i1}\beta_{j1}f_{1t}u_{it}+u_{jt}u_{it}\right).$$

We have

$$\frac{1}{TN^\alpha}\sum_{t=1}^{T}\sum_{j=1}^{N}\beta_{j1}\beta_{i1}f_{1t}^2=\left(\frac{1}{T}\sum_{t=1}^{T}f_{1t}^2\right)\left(\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v_{j1}\beta_{i1}\right).$$

But,

$$\frac{1}{T}\sum_{t=1}^{T}f_{1t}^2\to_p 1,\ \frac{\beta_{i1}}{N^\alpha}\sum_{j=1}^{N^\alpha}v_{j1}\to_p\beta_{i1}\mu_{v_1}.$$

Next,

$$\frac{1}{TN^\alpha}\sum_{t=1}^{T}u_{it}\sum_{j=1}^{N}u_{jt}=O_p\left(\frac{1}{T^{1/2}N^{\alpha-1/2}}\right),$$

and

$$\frac{2}{TN^\alpha}\sum_{t=1}^{T}\sum_{j=1}^{N}\beta_{i1}\beta_{j1}f_{1t}u_{it}=\frac{2}{T}\sum_{t=1}^{T}f_{1t}\beta_{i1}u_{it}\left(\left(\frac{1}{N^\alpha}\sum_{j=1}^{N^\alpha}v_{j1}\right)\right)=\begin{cases}O_p\left(\frac{1}{T^{1/2}}\right) & \text{if } i\leq N^\alpha\\ 0 & \text{otherwise}\end{cases}.$$

This concludes the analysis of the numerator of (B27). For the denominator we have,

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_t^2=\frac{1}{TN^{2\alpha}}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}x_{jt}x_{it}=$$

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N^{2\alpha}}\sum_{j=1}^{N}\sum_{i=1}^{N}\left(\beta_{j1}f_{1t}+u_{jt}\right)\left(\beta_{i1}f_{1t}+u_{it}\right)=$$

$$\frac{1}{T}\sum_{t=1}^{T}\frac{1}{N^{2\alpha}}\sum_{j=1}^{N}\sum_{i=1}^{N}\left(\beta_{j1}\beta_{i1}f_{1t}^{2}+2\beta_{i1}\beta_{j1}f_{1t}u_{it}+u_{jt}u_{it}\right).$$

We have

$$\frac{1}{TN^{2\alpha}}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}\beta_{j1}\beta_{i1}f_{1t}^{2}=\left(\frac{1}{T}\sum_{t=1}^{T}f_{1t}^{2}\right)\left(\frac{1}{N^{2\alpha}}\sum_{j=1}^{N^{\alpha}}\sum_{j=1}^{N^{\alpha}}v_{j1}v_{i1}\right).$$

But,

$$\frac{1}{T}\sum_{t=1}^{T}f_{1t}^{2}\rightarrow_{p}1,\ \frac{1}{N^{2\alpha}}\sum_{j=1}^{N^{\alpha}}\sum_{j=1}^{N^{\alpha}}v_{j1}v_{i1}\rightarrow_{p}\mu_{v_1}^{2},$$

$$\frac{1}{TN^{2\alpha}}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}u_{it}u_{jt}=O_p\left(\frac{1}{T^{1/2}N^{2\alpha-1}}\right),$$

$$\frac{2}{TN^{2\alpha}}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}\beta_{i1}\beta_{j1}f_{1t}u_{it}=$$

$$\frac{2}{T}\sum_{t=1}^{T}\left\{f_{1t}\left[\frac{1}{N^{\alpha}}\sum_{i=1}^{N^{\alpha}}v_{i1}u_{it}\left(\frac{1}{N^{\alpha}}\sum_{j=1}^{N^{\alpha}}v_{j1}\right)\right]\right\}=O_p\left(\frac{1}{T^{1/2}N^{\alpha/2}}\right).$$

Therefore,

$$\hat{\beta}_{i1}\rightarrow_{p}\frac{\beta_{i1}}{\mu_{v_1}}.$$

Next, we need to establish the rate at which $\frac{\frac{1}{T}\sum_{t=1}^{T}\bar{x}_t x_{it}}{\frac{1}{T}\sum_{t=1}^{T}\bar{x}_t^2}-\frac{\beta_{i1}\mu_{v_1}}{\mu_{v_1}^2}$ tends to zero. This is determined by the maximum of two rates:

$$\frac{1}{T}\sum_{t=1}^{T}\left(f_{1t}^{2}-1\right)=O_p\left(T^{-1/2}\right),$$

$$\frac{\beta_{i1}}{N^{\alpha}}\sum_{j=1}^{N^{\alpha}}v_{j1}-\beta_{i1}\mu_{v_1}=\begin{cases}O_p\left(\frac{1}{N^{\alpha/2}}\right) & \text{if } i\leq N^{\alpha}\\ 0 \text{ otherwise}\end{cases},$$

noting that

$$\frac{1}{N^{2\alpha}}\sum_{j=1}^{N^{\alpha}}\sum_{j=1}^{N^{\alpha}}\left(v_{j1}v_{i1}-\mu_{v_1}^{2}\right)=O_p\left(\frac{1}{N^{\alpha}}\right),$$

and

$$\frac{1}{TN^{2\alpha}}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}u_{it}u_{jt}=O_p\left(\frac{1}{T^{1/2}N^{2\alpha-1}}\right).$$

Hence

$$\hat{\beta}_{i1}-\frac{\beta_{i1}}{\mu_{v_1}}=O_p\left(\frac{1}{N^{\alpha/2}}\right)+O_p\left(\frac{1}{T^{1/2}}\right).$$

Proof of Lemma 18

We need to show $\hat{V}_{\overline{f_1^2}}-V_{\overline{f_1^2}}=o_p(1)$. and assuming a one factor setting (without loss of generality). The result extends straightforwardly to m factors. We have $\hat{V}_{\overline{f_1^2}}-V_{\overline{f_1^2}}=\hat{V}_{\overline{f_1^2}}-\bar{V}_{\overline{f_1^2}}+\bar{V}_{\overline{f_1^2}}-V_{\overline{f_1^2}}$ where

$$\bar{V}_{\overline{f_1^2}}=\frac{1}{T}\sum_{t=1}^{T}\left(q_t-\frac{1}{T}\sum_{t=1}^{T}q_t\right)^2+\sum_{j=1}^{l}\left(\frac{1}{T}\sum_{t=j+1}^{T}\left(q_{t-j}-\frac{1}{T}\sum_{t=1}^{T}q_t\right)\left(q_t-\frac{1}{T}\sum_{t=1}^{T}q_t\right)\right),$$

and $q_t = \left(\frac{f_{1t}-\bar{f}_1}{s_{f_1}}\right)^2$. But, by Theorem 25.3 of Davidson (1994) and Assumption 3, we have that $\bar{V}_{\overline{f_1^2}} - V_{\overline{f_1^2}} = o_p(1)$, as long as $l \to \infty$ and $l = o(T)$. Then, it is sufficient to examine

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{f_{1t}}{s_{f_1}} - \tilde{x}_t\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{f_{1t}}{s_{f_1}} - \frac{\bar{x}_t}{\hat{\sigma}_{\bar{x}}}\right) = \frac{1}{s_{f_1}T}\sum_{t=1}^{T}\left(f_{1t} - \frac{\bar{x}_t}{N^{\alpha-1}\bar{v}_{1N}}\right) =$$

$$\frac{1}{s_{f_1}T}\sum_{t=1}^{T}\left(f_{1t} - \frac{N^{\alpha-1}\bar{v}_{1N}f_{1t} + \frac{1}{N}\sum_{i=1}^{N}u_{it}}{N^{\alpha-1}\bar{v}_{1N}}\right) = \frac{N^{-\alpha}}{s_{f_1}T}\sum_{t=1}^{T}\sum_{i=1}^{N}u_{it} = \frac{N^{-\alpha}}{\sigma_{f_1}T}\sum_{t=1}^{T}\sum_{i=1}^{N}u_{it} + o_p\left(\frac{N^{-\alpha}}{\sigma_{f_1}T}\sum_{t=1}^{T}\sum_{i=1}^{N}u_{it}\right).$$

But, $\sum_{t=1}^{T}\sum_{i=1}^{N}u_{it} = O_p\left((NT)^{1/2}\right)$. So, $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{f_{1t}}{s_{f_1}} - \tilde{x}_t\right) = O_p\left(N^{1/2-\alpha}T^{-1/2}\right)$. Thus, $\hat{V}_{\overline{f_1^2}} - V_{\overline{f_1^2}} = O_p\left(lN^{1/2-\alpha}T^{-1/2}\right)$, proving the Lemma.

Supplementary Appendix IV: Justification of the use of the cumulative distribution function of the standard normal in the approach used to estimate μ_{v_1}.

Consider the single factor model,

$$x_{it} = \beta_{i1}f_{1t} + u_{it}, \text{ for } i = 1, ..., N;\ t = 1, ..., T, \tag{B28}$$

and assume that $\bar{\beta}_{1N} = \frac{1}{N}\sum_{i=1}^{N}\beta_{i1} \neq 0$ for a finite N. Recall that $\beta_{i1} = \nu_{i1}$, for $i = 1, ..., N^{\alpha}$ and zero for $i = N^{\alpha}+1, ..., N$ (without loss of generality), so that

$$\bar{\beta}_{1N} = N^{\alpha-1}\bar{v}_{1N}, \text{ with } \bar{v}_{1N} = N^{-\alpha}\sum_{i=1}^{N^{\alpha}}v_{i1}.$$

Also letting $\bar{x}_t = \frac{1}{N}\sum_{i=1}^{N}x_{it}$, $\delta_i = \beta_{i1}/\bar{\beta}_{1N}$ and noting that $\bar{x}_t = \bar{\beta}_{1N}f_{1t} + \bar{u}_t$, we have

$$x_{it} = \delta_i\bar{x}_t + \xi_{it}, \text{ where } \xi_{it} = u_{it} - \delta_i\bar{u}_t. \tag{B29}$$

Consider now the t-ratio for testing $\delta_i = 0$ in the above regression and note that it is given by

$$z_i = z_{i,T,N} = \frac{\hat{\delta}_i}{\left(\sum_{t=1}^{T}\bar{x}_t^2\right)^{-1/2}\hat{\sigma}_{\xi i}} = \frac{\sum_{t=1}^{T}\bar{x}_t x_{it}}{\left(\sum_{t=1}^{T}\bar{x}_t^2\right)^{1/2}\hat{\sigma}_{\xi i}}, \tag{B30}$$

where

$$\hat{\sigma}_{\xi i}^2 = T^{-1}\sum_{t=1}^{T}(x_{it} - \hat{\delta}_i\bar{x}_t)^2,$$

$$\hat{\delta}_i = \frac{\sum_{t=1}^{T}\bar{x}_t x_{it}}{\sum_{t=1}^{T}\bar{x}_t^2} = \delta_i + \frac{\sum_{t=1}^{T}\bar{x}_t\xi_{it}}{\sum_{t=1}^{T}\bar{x}_t^2}.$$

But

$$\sum_{t=1}^{T}\bar{x}_t x_{it} = \sum_{t=1}^{T}\left(\bar{\beta}_{1N}f_t + \bar{u}_t\right)\left(\beta_{i1}f_{1t} + u_{it}\right)$$
$$= \beta_{i1}\bar{\beta}_{1N}\sum_{t=1}^{T}f_{1t}^2 + \beta_{i1}\sum_{t=1}^{T}\bar{u}_t f_{1t} + \bar{\beta}_{1N}\sum_{t=1}^{T}f_{1t}u_{it} + \sum_{t=1}^{T}\bar{u}_t u_{it}$$

and

$$\sum_{t=1}^{T}\bar{x}_t^2 = \bar{\beta}_{1N}^2\sum_{t=1}^{T}f_{1t}^2 + 2\bar{\beta}_{1N}\sum_{t=1}^{T}f_{1t}\bar{u}_t + \sum_{t=1}^{T}\bar{u}_t^2,$$

$$
\begin{aligned}
\hat{\sigma}^2_{\xi i} &= T^{-1}\sum_{t=1}^{T}\Big[x_{it}-\delta_i\bar{x}_t-(\hat{\delta}_i-\delta_i)\bar{x}_t\Big]^2 \\
&= T^{-1}\sum_{t=1}^{T}\xi_{it}^2+(\hat{\delta}_i-\delta_i)^2T^{-1}\sum_{t=1}^{T}\bar{x}_t^2-2(\hat{\delta}_i-\delta_i)T^{-1}\sum_{t=1}^{T}\bar{x}_t\xi_{it} \\
&= T^{-1}\sum_{t=1}^{T}\xi_{it}^2+\left(\frac{T^{-1}\sum_{t=1}^{T}\bar{x}_t\xi_{it}}{T^{-1}\sum_{t=1}^{T}\bar{x}_t^2}\right)^2T^{-1}\sum_{t=1}^{T}\bar{x}_t^2-2\frac{T^{-1}\sum_{t=1}^{T}\bar{x}_t\xi_{it}}{T^{-1}\sum_{t=1}^{T}\bar{x}_t^2}T^{-1}\sum_{t=1}^{T}\bar{x}_t\xi_{it} \\
&= T^{-1}\sum_{t=1}^{T}\xi_{it}^2-\frac{\left(T^{-1}\sum_{t=1}^{T}\bar{x}_t\xi_{it}\right)^2}{T^{-1}\sum_{t=1}^{T}\bar{x}_t^2}.
\end{aligned}
$$

Also

$$
\begin{aligned}
z_i &= \frac{\beta_{i1}\bar{\beta}_{1N}\sum_{t=1}^{T}f_{1t}^2+\beta_{i1}\sum_{t=1}^{T}\bar{u}_tf_{1t}+\bar{\beta}_{1N}\sum_{t=1}^{T}f_{1t}u_{it}+\sum_{t=1}^{T}\bar{u}_tu_{it}}{\left(\bar{\beta}_{1N}^2\sum_{t=1}^{T}f_{1t}^2+2\bar{\beta}_{1N}\sum_{t=1}^{T}f_{1t}\bar{u}_t+\sum_{t=1}^{T}\bar{u}_t^2\right)^{1/2}\hat{\sigma}_{\xi i}} \\
&= \frac{\beta_{i1}\sum_{t=1}^{T}f_{1t}^2+\beta_{i1}\sum_{t=1}^{T}(\bar{u}_t/\bar{\beta}_{1N})f_{1t}+\sum_{t=1}^{T}f_{1t}u_{it}+\sum_{t=1}^{T}\left(\bar{u}_t/\bar{\beta}_{1N}\right)u_{it}}{\left[\sum_{t=1}^{T}f_{1t}^2+2\sum_{t=1}^{T}f_{1t}\left(\bar{u}_t/\bar{\beta}_{1N}\right)+\sum_{t=1}^{T}\left(\bar{u}_t/\bar{\beta}_{1N}\right)^2\right]^{1/2}\hat{\sigma}_{\xi i}}.
\end{aligned}
$$

Further, since $\bar{\beta}_{1N}= N^{\alpha-1}\bar{v}_{1N}$, we have $\bar{u}_t/\bar{\beta}_{1N}=N^{1-\alpha}(\bar{u}_t/\bar{v}_{1N})$ and

$$
T^{-1/2}z_i=\frac{\beta_{i1}T^{-1}\sum_{t=1}^{T}f_{1t}^2+T^{-1}\sum_{t=1}^{T}f_{1t}u_{it}+\left(\beta_{i1}/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_tf_{1t}+\left(1/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_tu_{it}}{\left[T^{-1}\sum_{t=1}^{T}f_{1t}^2+2\left(1/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}f_{1t}\bar{u}_t+\left(1/\bar{v}_{1N}\right)^2N^{2(1-\alpha)}T^{-1}\sum_{t=1}^{T}\bar{u}_t^2\right]^{1/2}\hat{\sigma}_{\xi i}},
$$

$$
\begin{aligned}
T^{-1}\sum_{t=1}^{T}\xi_{it}^2 &= T^{-1}\sum_{t=1}^{T}u_{it}^2+\delta_i^2T^{-1}\sum_{t=1}^{T}\bar{u}_t^2-2\delta_iT^{-1}\sum_{t=1}^{T}\bar{u}_tu_{it} \\
&= T^{-1}\sum_{t=1}^{T}u_{it}^2+\beta_{i1}^2\left(1/\bar{v}_{1N}\right)^2N^{2(1-\alpha)}T^{-1}\sum_{t=1}^{T}\bar{u}_t^2-2\beta_{i1}\left(1/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_tu_{it}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\left(T^{-1}\sum_{t=1}^{T}\bar{x}_t\xi_{it}\right)^2}{T^{-1}\sum_{t=1}^{T}\bar{x}_t^2} &= \frac{\left(T^{-1}\sum_{t=1}^{T}\bar{x}_t(u_{it}-\beta_{i1}\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right)^2}{\bar{\beta}_{1N}^2T^{-1}\sum_{t=1}^{T}f_{1t}^2+2\bar{\beta}_{1N}T^{-1}\sum_{t=1}^{T}f_{1t}\bar{u}_t+T^{-1}\sum_{t=1}^{T}\bar{u}_t^2} \\
&= \frac{\left(\left[T^{-1}\sum_{t=1}^{T}f_{1t}+\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right]\left[u_{it}-\beta_{i1}\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right]\right)^2}{T^{-1}\sum_{t=1}^{T}f_{1t}^2+2T^{-1}\sum_{t=1}^{T}f_{1t}\left(\bar{u}_t/\bar{\beta}_{1N}\right)+T^{-1}\sum_{t=1}^{T}\left(\bar{u}_t/\bar{\beta}_{1N}\right)^2} \\
&= \frac{\left(\left[T^{-1}\sum_{t=1}^{T}f_{1t}+\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right]\left[u_{it}-\beta_{i1}\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right]\right)^2}{T^{-1}\sum_{t=1}^{T}f_{1t}^2+2\left(1/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}f_{1t}\bar{u}_t+\left(1/\bar{v}_{1N}\right)^2N^{2(1-\alpha)}T^{-1}\sum_{t=1}^{T}\bar{u}_t^2}.
\end{aligned}
$$

But we have

$$
T^{-1}\sum_{t=1}^{T}\left[f_{1t}+\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right]\left[u_{it}-\beta_{i1}\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right]=
$$

$$
T^{-1}\sum_{t=1}^{T}f_{1t}u_{it}-\beta_{i1}T^{-1}\sum_{t=1}^{T}f_{1t}\left(\bar{u}_t/\bar{\beta}_{1N}\right)-\beta_{i1}T^{-1}\sum_{t=1}^{T}\left(\bar{u}_t/\bar{\beta}_{1N}\right)^2+T^{-1}\sum_{t=1}^{T}\left(\bar{u}_t/\bar{\beta}_{1N}\right)u_{it}
$$

$$
=T^{-1}\sum_{t=1}^{T}f_{1t}u_{it}-\left(\beta_{i1}/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}f_{1t}\bar{u}_t-\beta_{i1}\left(1/\bar{v}_{1N}\right)^2N^{2(1-\alpha)}T^{-1}\sum_{t=1}^{T}\bar{u}_t^2+\left(1/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_tu_{it},
$$

$$N^{2(1-\alpha)}T^{-1}\sum_{t=1}^{T}\bar{u}_t^2 = O_p\left(N^{1-2\alpha}\right),\ N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_t u_{it} = O_p\left(N^{1/2-\alpha}T^{-1/2}\right), \tag{B31}$$

(B32)

$$N^{1-\alpha}T^{-1}\sum_{t=1}^{T}f_{1t}\bar{u}_t = O_p\left(N^{1/2-\alpha}T^{-1/2}\right),\ T^{-1}\sum_{t=1}^{T}f_{1t}u_{it} = O_p\left(T^{-1/2}\right).$$

(B33)

Hence,

$$\begin{aligned}T^{-1}\sum_{t=1}^{T}\xi_{it}^2 &= T^{-1}\sum_{t=1}^{T}u_{it}^2 + \beta_{i1}^2\left(1/\bar{v}_{1N}\right)^2 N^{2(1-\alpha)}T^{-1}\sum_{t=1}^{T}\bar{u}_t^2 - 2\beta_{i1}\left(1/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_t u_{it}\\ &= \sigma_i^2 + O_p\left(T^{-1/2}\right) + O_p\left(N^{1/2-\alpha}T^{-1/2}\right) + O_p\left(N^{1-2\alpha}\right)\end{aligned}$$

$$T^{-1}\sum_{t=1}^{T}\left(f_{1t} + \left(\bar{u}_t/\bar{\beta}_{1N}\right)\right)\left[u_{it} - \beta_{i1}\left(\bar{u}_t/\bar{\beta}_{1N}\right)\right] = O_p\left(T^{-1/2}\right) + O_p\left(N^{1-2\alpha}\right) + O_p\left(N^{1/2-\alpha}T^{-1/2}\right).$$

$$\begin{aligned}\hat{\sigma}_{\xi i}^2 &= T^{-1}\sum_{t=1}^{T}\xi_{it}^2 - \frac{\left(T^{-1}\sum_{t=1}^{T}\bar{x}_t\xi_{it}\right)^2}{T^{-1}\sum_{t=1}^{T}\bar{x}_t^2}\\ &= \sigma_i^2 + O_p\left(T^{-1/2}\right) + O_p\left(N^{1/2-\alpha}T^{-1/2}\right) + O_p\left(N^{1-2\alpha}\right).\end{aligned}$$

Using the above results we now have

$$\begin{aligned}T^{-1/2}z_i &= \frac{T^{-1}\sum_{t=1}^{T}f_{1t}u_{it} + O_p\left(N^{1/2-\alpha}T^{-1/2}\right)}{\left[T^{-1}\sum_{t=1}^{T}f_{1t}^2 + O_p\left(N^{1/2-\alpha}T^{-1/2}\right) + O_p\left(N^{1-2\alpha}\right)\right]^{1/2}\hat{\sigma}_{\xi i}},\ \text{if } \beta_{i1}=0\\ &= \frac{T^{-1}\sum_{t=1}^{T}f_{1t}(u_{it}/\sigma_i)}{\left(T^{-1}\sum_{t=1}^{T}f_{1t}^2\right)^{1/2}} + O_p\left(N^{1/2-\alpha}T^{-1/2}\right) + O_p\left(N^{1-2\alpha}\right).\end{aligned} \tag{B34}$$

Therefore, under $\beta_{i1}=0$, z_i is asymptotically distributed as $N(0,1)$ so long as N and T tend to infinity in any order and $\alpha > 1/2$. Also,

$$T^{-1/2}z_i = \frac{\beta_{i1}T^{-1}\sum_{t=1}^{T}f_{1t}^2 + T^{-1}\sum_{t=1}^{T}f_{1t}u_{it} + \left(\beta_{i1}/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_t f_{1t} + \left(1/\bar{v}_{1N}\right)N^{1-\alpha}T^{-1}\sum_{t=1}^{T}\bar{u}_t u_{it}}{\left[T^{-1}\sum_{t=1}^{T}f_{1t}^2 + O_p\left(N^{1/2-\alpha}T^{-1/2}\right) + O_p\left(N^{1-2\alpha}\right)\right]^{1/2}\hat{\sigma}_{\xi i}},$$

if $\beta_{i1} \neq 0$,

$$\begin{aligned}&= \left(\frac{\beta_{i1}}{\sigma_i}\right)\left(T^{-1}\sum_{t=1}^{T}f_{1t}^2\right)^{1/2} + \frac{T^{-1}\sum_{t=1}^{T}f_{1t}(u_{it}/\sigma_i)}{\left(T^{-1}\sum_{t=1}^{T}f_{1t}^2\right)^{1/2}} +\\ &O_p\left(N^{1/2-\alpha}T^{-1/2}\right) + O_p\left(N^{1-2\alpha}\right).\end{aligned} \tag{B35}$$

Thus, under $\beta_{i1} \neq 0$, and using the normalization $T^{-1}\sum_{t=1}^{T}f_{1t}^2 \rightarrow_p 1$, $(z_i - \frac{\sqrt{T}\beta_{i1}}{\sigma_i}) \rightarrow_d N(0,1)$ as N and $T \rightarrow \infty$, in any order, and if $\alpha > 1/2$. It is also easy to see that (B34) and (B35) also hold in mean square.

In the case of a multi-factor setting, (B29) can be re-written in the form shown in Lemma 4 so that the error term, ξ_{it},now is augmented by residuals from the regression of each of the m factors on $\bar{x}_t$. The rest of the analysis then follows through.

Supplementary Appendix V: Proof of consistency of $\hat{\mu}_{v_1}(c_{p,N})$ based on multiple testing

The proof is heuristic to the extent that a high level assumption is needed that may be difficult to establish using more primitive conditions. We make the following assumption.

Assumption 6 *1. β_{i1} is uniformly bounded over i .*

2. *$\bar{x}_t u_{it}$is uniformly mixing over i, in the sense of the mixing assumption of Connor and Korajczyk (1993), with mixing coefficients $\phi_{t,m}$that satisfy $\sup_t \lim_{m\to\infty} \phi_{t,m}= 0$.*

3. *Let φ_i denote a standard normal variate. Then, if $z_{i,T,N} - \varphi_i= O_{m.s.}(N^{2-4\alpha}) + O_{m.s.}(N^{1/2-\alpha}T^{-1/2})$, $\sup_i z_{i,T,N}-\varphi_i = O_p\left(N^{2-4\alpha}\right)+O_p\left(N^{1/2-\alpha}T^{-1/2}\right)$, and $\sup_i E\left(z_{i,T,N}-\varphi_i\right)^2 = O(N^{2-4\alpha})+O(N^{1-2\alpha}T^{-1})$, where $O_{m.s.}()$ denotes order in mean square.*

4. *Let $\psi = (\psi_1, ..., \psi_N)'$ denote an $N\times 1$ selector vector consisting of zeros and ones such that $\psi'\psi > N^{\alpha}$, for some $\alpha > 1/2$. Define $u_t^{\psi} = (\psi'\psi)^{-1}\sum_{i=1}^N \psi_i u_{it}$. Then,*

$$\sup_t \sup_\psi E\left(u_t^{\psi}\right)^2 = o\left(1\right).$$

Remark 2 *Condition 2 is a standard uniform mixing condition. Uniform mixing is a stronger form of mixing than strong mixing which is more widely used, but allows a CLT without any rates for the mixing coefficients and only the existence of $2+\delta$, $\delta>0$ moments. One could simplify further the assumption by imposing a uniform mixing condition on u_{it}, and thereby $f_{1t}u_{it}$ and proving that $\bar{x}_t u_{it}$ is uniform mixing with mixing coefficients that have mixing size -1/2, but we choose to make this slightly less primitive assumption for simplicity. Clearly, if u_{it} follow (16) then Condition 2 is satisfied. If u_{it}follow (28) then both (29) and assumptions on $\nu_{s,t}$ need to be strengthened. A discussion of these issues may be found in Section 14.3 of Davidson (1994) and, in particular, Theorem 14.14. Conditions 3 and 4 are uniform convergence technical conditions which again seem difficult to establish from more primitive conditions. A proof of the normality invoked in Condition 3 is provided in Supplementary Appendix IV and the assumption only strengthens the result to make it uniform. Condition 4 appears intuitive due to the weak cross-sectional dependence of the errors, although again uniformity is difficult to establish formally.*

Set

$$w_{it} = x_{it} I\left(|z_{i,T,N}| \geq c_{p_i,N}\right),\ \theta_i = \beta_{i1} I\left(|z_{i,T,N}| \geq c_{p_i,N}\right),\ \upsilon_{it} = u_{it} I\left(|z_{i,T,N}| \geq c_{p_i,N}\right),$$

where $c_{p_i,N}$ is the critical value of the i-th test. Then,

$$w_{it} = \theta_i f_{1t} + \upsilon_{it}, \text{for } i = 1, ..., N;\ t = 1, ..., T,$$

and

$$\bar{w}_t = \bar{\theta} f_{1t} + \bar{\upsilon}_t,$$

where

$$\bar{w}_t = \frac{\sum_{i=1}^N w_{it}}{\sum_{i=1}^N I\left(|z_{i,T,N}| \geq c_{p_i,N}\right)},\ \bar{\theta} = \frac{\sum_{i=1}^N \theta_i}{\sum_{i=1}^N I\left(|z_{i,T,N}| \geq c_{p_i,N}\right)},$$

and

$$\bar{\upsilon}_t = \frac{\sum_{i=1}^N \upsilon_{it}}{\sum_{i=1}^N I\left(|z_{i,T,N}| \geq c_{p_i,N}\right)}.$$

We take

$$\hat{\sigma}^2_{\bar{w}} = \frac{1}{T}\sum_{t=1}^T \left(\bar{w}_t - \bar{w}\right)^2,$$

and consider the limiting behaviour of $\sigma^2_{\bar{w}}/\mu_{v_1}$, where as before $\mu_{v_1}= E(v_{i1})$. Since

$$\bar{w}_t - \bar{w} = \bar{\theta}\left(f_{1t} - \bar{f}_1\right) + \left(\bar{\upsilon}_t - \bar{\upsilon}\right),$$

then

$$\frac{\hat{\sigma}_{\bar{w}}^2}{\mu_{v_1}^2} = \frac{\bar{\theta}^2}{\mu_{v_1}^2}\frac{1}{T}\sum_{t=1}^{T}\left(f_{1t}-\bar{f}_1\right)^2 + \frac{2\bar{\theta}}{\mu_{v_1}^2}\frac{1}{T}\sum_{t=1}^{T}\left(f_{1t}-\bar{f}_1\right)\left(\bar{v}_t-\bar{v}\right) + \frac{1}{\mu_{v_1}^2}\frac{1}{T}\sum_{t=1}^{T}\left(\bar{v}_t-\bar{v}\right)^2 \tag{B36}$$
$$= I + II + III.$$

We concentrate on I as we will prove that II and III tend to zero. For I, and since $\frac{1}{T}\sum_{t=1}^{T}\left(f_{1t}-\bar{f}_1\right)^2 \to_p 1$, we have

$$\bar{\theta} = \frac{\sum_{i=1}^{N^\alpha}\beta_{i1}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}\neq 0\right) + \sum_{i=N^\alpha+1}^{N}\beta_{i1}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}=0\right)}{\sum_{i=1}^{N^\alpha}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}\neq 0\right) + \sum_{i=N^\alpha+1}^{N}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}=0\right)}, \text{ or}$$
$$= \frac{\frac{1}{N^\alpha}\left(\sum_{i=1}^{N^\alpha}\beta_{i1}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}\neq 0\right) + \sum_{i=N^\alpha+1}^{N}\beta_{i1}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}=0\right)\right)}{\frac{1}{N^\alpha}\left(\sum_{i=1}^{N^\alpha}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}\neq 0\right) + \sum_{i=N^\alpha+1}^{N}I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}=0\right)\right)}. \tag{B37}$$

We first consider the asymptotic behaviour of the following four terms:

$$A = \frac{1}{N^a}\sum_{i=1}^{N^\alpha}\beta_{i1}\left(I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}\neq 0\right) - \Pr\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}\neq 0\right)\right),$$

$$B = \frac{1}{N^a}\sum_{i=N^\alpha+1}^{N}\beta_{i1}\left(I\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}=0\right) - \Pr\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}=0\right)\right),$$

$$C = \frac{1}{N^a}\sum_{i=1}^{N^\alpha}\left(I\left(|z_{i,T,N}| \geq c_{p,N}|\beta_{i1}\neq 0\right) - \Pr\left(|z_{i,T,N}| \geq c_{p,N}|\beta_{i1}\neq 0\right)\right),$$

$$D = \frac{1}{N^a}\sum_{i=N^\alpha+1}^{N}\left(I\left(|z_{i,T,N}| \geq c_{p,N}|\beta_{i1}=0\right) - \Pr\left(|z_{i,T,N}| \geq c_{p,N}|\beta_{i1}=0\right)\right).$$

We need to show that the summands in $A-D$ follow a central limit theorem. It is sufficient to show that the summands are uniformly mixing. By Condition 2 of Assumption 6 it follows that $z_{i,T,N}$ is uniformly mixing over i. By the measurability of the indicator function (see, e.g., Theorem 3.27 of Davidson (1994)) and Theorem 14.1 of Davidson (1994), it follows that all summands in $A-D$, are uniformly mixing and, by Theorem 18.5.1 of Ibragimov and Linnik (1971), a central limit theorem holds. Then, it follows that

$$A = O\left(N^{-\alpha/2}\right), \quad B = O\left(N^{1/2-\alpha}\right), \quad C = O\left(N^{-\alpha/2}\right), \quad D = O\left(N^{1/2-\alpha}\right).$$

Next, we consider

$$\lim_{N,T\to\infty}\left(\frac{\sum_{i=1}^{N^\alpha}\Pr\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}\neq 0\right)}{N^\alpha}\right), \text{ and}$$
$$\lim_{N,T\to\infty}\left(\frac{\sum_{i=N^\alpha+1}^{N}\Pr\left(|z_{i,T,N}| \geq c_{p_i,N}|\beta_{i1}=0\right)}{N^\alpha}\right).$$

We have that,

$$\Pr\left(|z_{i,T,N}| \geq c_{p,N}|\beta_{i1}\neq 0\right) = 1 - \left[\Phi\left(c_{p,N} - \frac{\beta_{i1}\sqrt{T}}{\sigma_i}\right) - \Phi\left(-c_{p,N} - \frac{\beta_{i1}\sqrt{T}}{\sigma_i}\right)\right] \tag{B38}$$
$$+ O\left(N^{1-2\alpha}T^{-1}\right) + O\left(N^{2-4\alpha}\right).$$
$$= 1 - \Phi\left(c_{p,N} - \frac{\beta_{i1}\sqrt{T}}{\sigma_i}\right) + \Phi\left(-c_{p,N} - \frac{\beta_{i1}\sqrt{T}}{\sigma_i}\right) + O\left(N^{2-4\alpha}\right) + O\left(N^{1-2\alpha}T^{-1}\right).$$

(B38) can be proven as follows. From Supplementary Appendix IV we have

$$z_{i,T,N} = z_i + O_p\left(N^{1-2\alpha}\right) + O_p\left(N^{1/2-\alpha}T^{-1/2}\right) = z_i + q_{i,N,T},$$

where z_i is distributed as $N(0,1)$ and $q_{i,N,T} = O_p\left(N^{1-2\alpha}\right) + O_p\left(N^{1/2-\alpha}T^{-1/2}\right)$. Then, we have

$$\Pr\left(|z_{i,T,N}| \leq c_{p_i,N}\right) - \Pr\left(|z_i| \leq c_{p_i,N}\right) = \Pr\left(|z_i + q_{i,N,T}| \leq c_{p_i,N}\right) - \Pr\left(|z_i| \leq c_{p_i,N}\right) \leq \Pr\left(|q_{i,N,T}| > 0\right)$$

$$\lim_{N,T\to\infty} \Pr\left(|q_{iN,T}| > 0\right) = \lim_{\epsilon\to 0} \lim_{N,T\to\infty} \Pr\left(|q_{i,N,T}| > \epsilon\right).$$

But

$$\Pr\left(|q_{i,N,T}| > \epsilon\right) \leq \frac{E\left(q_{i,N,T}^2\right)}{\epsilon^2}.$$

It is easy to see from the analysis of Supplementary Appendix IV that

$$E\left(q_{iN,T}^2\right) = O\left(N^{2-4\alpha}\right) + O\left(N^{1-2\alpha}T^{-1}\right),$$

then

$$\Pr\left(|z_{i,T,N}| \leq c_{p_i,N}\right) - \Phi(c_{p_i,N}) - \Phi(-c_{p_i,N}) = O\left(N^{2-4\alpha}\right) + O\left(N^{1-2\alpha}T^{-1}\right), \tag{B39}$$

proving (B38). Assumption 6 (3) strengthens this to

$$\sup_i \Pr\left(|z_{i,T,N}| \leq c_{p_i,N}\right) - \Phi(c_{p_i,N}) - \Phi(-c_{p_i,N}) = O\left(N^{2-4\alpha}\right) + O\left(N^{1-2\alpha}T^{-1}\right). \tag{B40}$$

Thus,

$$p\lim_{N,T\to\infty} \left(\frac{\sum_{i=1}^{[N^\alpha]} I\left(|z_{i,T,N}| \geq c_{p_i,N} | \beta_{i1} \neq 0\right)}{N^\alpha}\right) \to_p 1. \tag{B41}$$

as long as

$$c_{p_i,N} = o_p(T^{1/2}) \tag{B42}$$

uniformly over i. Also,

$$\begin{aligned} \Pr\left(|z_{i,T,N}| \geq c_{p_i,N} | \beta_{i1} = 0\right) &= \left[1 - \Phi\left(c_{p_i,N}\right) + \Phi\left(-c_{p_i,N}\right)\right] + O\left(N^{1-2\alpha}T^{-1}\right) + O\left(N^{2-4\alpha}\right) \\ &= 2\left[1 - \Phi\left(c_{p_i,N}\right)\right] + O\left(N^{1-2\alpha}T^{-1}\right) + O\left(N^{2-4\alpha}\right). \end{aligned}$$

Then,

$$\frac{\sum_{i=[N^\alpha]+1}^{N} \Pr\left(|z_{i,T,N}| \geq c_{p,N} | \beta_{i1} = 0\right)}{N^\alpha} = \frac{\sum_{i=[N^\alpha]+1}^{N} 2\left[1 - \Phi\left(c_{p_i,N}\right)\right]}{N^\alpha} + \frac{(N - N^\alpha)}{N^\alpha}\left[O\left(N^{1-2\alpha}T^{-1}\right) + O\left(N^{2-4\alpha}\right)\right],$$

and, as long as

$$\frac{\sum_{i=[N^\alpha]+1}^{N} 2\left[1 - \Phi\left(c_{p_i,N}\right)\right]}{N^\alpha} = o_p(1), \tag{B43}$$

then,[13]

$$\frac{\sum_{i=[N^\alpha]+1}^{N} I\left(|z_{i,T,N}| \geq c_{p_i,N} | \beta_{i1} = 0\right)}{N^\alpha} \to 0,$$

if either $\alpha > 2/3$ or $\alpha > 3/5$ and $N^{2-3\alpha}T^{-1} = o(1)$. The latter follows, if $\alpha > 3/5$ and $N = o(T^5)$. For simplicity, we will assume that $\alpha > 2/3$. Now, we check

$$\lim_{N,T\to\infty} \left(\frac{\sum_{i=1}^{N^\alpha} \beta_{i1} \Pr\left(|z_{i,T,N}| \geq c_{p_i,N} | \beta_{i1} \neq 0\right)}{N^\alpha}\right), \text{ and}$$

$$\lim_{N,T\to\infty} \left(\frac{\sum_{i=[N^\alpha]+1}^{N} \beta_{i1} \Pr\left(|z_{i,T,N}| \geq c_{p_i,N} | \beta_{i1} = 0\right)}{N^\alpha}\right).$$

[13] It is easy to see that both the Holm and Bonferroni multiple testing approach discussed in Section 3.1 satisfy (B43). For Bonferroni, this is obvious. For Holm, we note that if $c_{p_i,N} = \Phi^{-1}(1 - p_i)$, $p_i = \frac{p}{2(N-i+1)}$, then $2\left[1 - \Phi\left(c_{p_i,N}\right)\right] = C_i \frac{p}{2(N-i+1)}$, for some uniformly bounded positive constants C_i. Since $\sum_{i=[N^\alpha]+1}^{N} 2\left[1 - \Phi\left(c_{p_i,N}\right)\right] \leq C \ln N$, (B43) holds.

We have,

$$\lim_{N,T\to\infty}\left(\frac{\sum_{i=[N^{\alpha}]+1}^{N}\beta_{i1}\Pr\left(|z_{i,T,N}|\geq c_{p_i,N}|\beta_{i1}=0\right)}{N^{\alpha}}\right)=0,$$

and

$$\lim_{N,T\to\infty}\left(\frac{\sum_{i=1}^{N^{\alpha}}\beta_{i1}\Pr\left(|z_{i,T,N}|\geq c_{p_i,N}|\beta_{i1}\neq 0\right)}{N^{\alpha}}\right)=\lim_{N,T\to\infty}\left(\frac{\sum_{i=1}^{N^{\alpha}}\beta_{i1}\left[1-\Phi\left(c_{p_i,N}-\frac{\beta_{i1}\sqrt{T}}{\sigma_i}\right)+\Phi\left(-c_{p_i,N}-\frac{\beta_{i1}\sqrt{T}}{\sigma_i}\right)\right]}{N^{\alpha}}\right) \tag{B44}$$

$$\to_p E(\bar{v})=\mu_{v_1},$$

using (B41), or

$$\bar{\theta}\to_p \mu_{v_1}. \tag{B45}$$

And therefore,

$$\frac{\hat{\sigma}_{\bar{w}}^2}{\mu_{v_1}^2}\to_p 1.$$

Finally, for II and III we first note that we have already established that $N^{-\alpha}\sum_{i=1}^{N}I\left(|z_{i,T,N}|\geq c_{p_i,N}\right)\to_p 1$, as N and $T\to\infty$, assuming that $\alpha>2/3$. But by the proofs of Lemma A.1 and A.2 of (Pesaran, 2006, Theorem 15.18) and using assumption 6 it immediately follows that $II=o_p(1)$ and $III=o_p(1)$ completing the proof. In summary, consistency is obtained under Assumption 6 and conditions (B42) and (B43) if $\alpha>2/3$.

Note that in the case of a multi-factor setting, (B44) can alter. If $\alpha=\alpha_1>\alpha_2>...>\alpha_m$, then the denominator can potentially capture more elements than $N^{\alpha}$and so (B45) converges to $\sum_{j=1}^{m}c_j\mu_{v_j}$,where $0<c_j<1$.

Supplementary Appendix VI: Additional Monte Carlo simulation results

We provide some additional Monte Carlo simulation results in this appendix. First, we set $\mu_v=1$ and keep $\alpha=\alpha_1>\alpha_2$. In this case $\tilde{\alpha}$ consistently estimates α and has the asymptotic distribution as described in Theorem 1. Next, we present size and power of tests based on $\mathring{\alpha}$ as well. We use the same confidence bands as in the case of $\tilde{\alpha}$. From the results shown for experiment A it is confirmed that $\mathring{\alpha}$ is super-consistent. Finally, we consider the two factor model of (36) for the case when $\alpha=\alpha_1=\alpha_2$ and depict bias and RMSE results for estimator $\mathring{\alpha}$.

A two-factor model where $\mu_\nu=1$

In addition to the results analysed in Section 4, here we consider the instance when $\mu_\nu=1$ and show bias, RMSE, size and power results for estimator $\tilde{\alpha}$ which is asymptotically distributed in accordance to Theorem 1. We use the set up of experiment A of Section 4 and set $\mu_v=1$, $\mu_{\nu_2}=0.87$, $\mu_{\nu_1}=\sqrt{\mu_\nu^2-N^{2(\alpha_2-\alpha)}\mu_{\nu_2}^2}$ and $\varsigma=3/4$. Since the leading factor (f_{1t}) is serially uncorrelated, the statistic for making inference about α is given by

$$\left(\frac{1}{T}\hat{V}_{\overline{f_1^2}}+\frac{4}{N^{\tilde{\alpha}}}\frac{\widehat{\sigma_{v_1}^2}}{\mu_{v_1}^2}\right)^{-1/2}2\ln(N)\left(\tilde{\alpha}-\alpha^*\right)\to_d N(0,1). \tag{B46}$$

Note that when the leading factor is serially uncorrelated then $\hat{V}_{\overline{f_1^2}}=E\widehat{(f_{1t}^4)/}\sigma_{f_1}^4-1$, where $E\widehat{(f_{1t}^4)/}\sigma_f^4$ is consistently estimated by

$$E\widehat{(f_{1t}^4)/}\sigma_{f_1}^4=\frac{\sum_{t=1}^{T}\left(\tilde{x}_t-\tilde{x}\right)^4}{T},$$

where $\tilde{x}_t=\left(N^{-1}\sum_{i=1}^{N}x_{it}\right)/\hat{\sigma}_{\bar{x}}$, and $\widehat{\sigma_{v_1}^2/\mu_{v_1}^2}$, the estimator of $\sigma_{v_1}^2/\mu_{v_1}^2$, is given by

$$\frac{\widehat{\sigma_{v_1}^2}}{\mu_{v_1}^2}=\frac{\sum_{i=1}^{N^{\tilde{\alpha}}}\left(\hat{v}_{i1}^{(s)}-\frac{1}{N^{\tilde{\alpha}}}\sum_{j=1}^{N^{\tilde{\alpha}}}\hat{v}_{j1}^{(s)}\right)^2}{N^{\tilde{\alpha}}-1},$$

where $\left\{\hat{v}_{i1}^{(s)}\right\}$ denotes the sequence of $\hat{v}_{i1}$ sorted according to their absolute values in a descending order, and $\hat{v}_{i1}$ is the OLS estimator of the regression coefficient of x_{it} on $\tilde{x}_t = \bar{x}_t/\hat{\sigma}_{\bar{x}}$ - see Lemma 16 for details. The above expressions apply irrespective of the number of factors included in model (36).

Further, though not depicted in these Monte Carlo simulation results (these are available upon request), we consider the case of serially correlated factors as it is being used in the empirical applications of Section 5. When $\rho_j \neq 0$, we use a corrected variance estimator of f_{1t}. The relevant formula for the test statistic is given by

$$\left[\frac{1}{T}\left[\hat{V}_{\overline{f_1^2}}(q)\right] + \frac{4}{N^{\tilde{\alpha}}}\widehat{\frac{\sigma_{v_1}^2}{\mu_{v_1}^2}}\right]^{-1/2} 2\ln(N)\left(\tilde{\alpha} - \alpha^*\right) \rightarrow_d N(0,1). \tag{B47}$$

$\hat{V}_{\overline{f_1^2}}(q)$ is computed by first estimating an AR(q) process for $\tilde{z}_t = z_t - \bar{z}$, where $z_t = \left(\tilde{x}_t - \tilde{x}\right)^2$, $\tilde{x}_t = \left(\frac{1}{N}\sum_{i=1}^N x_{it}\right)/\hat{\sigma}_{\bar{x}}$, $\tilde{x} = T^{-1}\sum_{t=1}^T \tilde{x}_t$ and $\bar{z} = T^{-1}\sum_{t=1}^T z_t$, and then $\hat{V}_{\overline{f_1^2}}(q) = \hat{\sigma}_z^2/(1 - \widehat{\gamma}_1 - \widehat{\gamma}_2 - ... - \widehat{\gamma}_q)^2$, where $\hat{\sigma}_z$ is the regression standard error and $\widehat{\gamma}_i$ is the i^{th} estimated AR coefficient fitted to $\tilde{z}_t$. The lag order is set to $q = T^{1/3}$, and $\widehat{\sigma_{v_1}^2/\mu_{v_1}^2}$ is computed as before. Note that this correction is not the standard Newey-West one but uses an estimated autoregressive filter. We found that this correction leads to better finite sample properties and hence we use this in both the Monte Carlo study and the empirical applications in Section 5.

Size of the tests is computed under $H_0 : \alpha = \alpha_0$, using a two-sided alternative where α_0 takes values in the range $[0.70, 1.00]$, as indicated previously. Power is computed under the alternatives $H_a : \alpha_a = \alpha_0 + 0.05$ (power+), and $H_a : \alpha_a = \alpha_0 - 0.05$ (power-). Again, all results are scaled up by 100.

Size and power of tests based on $\mathring{\alpha}$ estimator

Next, we conduct size and power tests based on estimator $\mathring{\alpha}$. We use the same variance estimates as in (B46) which constitute conservative bands for $\mathring{\alpha}$ and show results for the setting described in Section 4 for experiment A when $\mu_v^2 \neq 1$. The same specifications for the null and alternative hypotheses are imposed as in the Section above.

A two-factor model when $\alpha = \alpha_1 = \alpha_2$

Finally, we repeat the analysis of Section 4 for Experiment A using the less likely alternative of $\alpha = \alpha_1 = \alpha_2$. Here, we set $\mu_{v_1} = \mu_{v_2} = 0.5$ and $\varsigma = 1/3$.

Additional results

Table A1 presents bias, RMSE, size and power statistics for experiment A in the case of the bias-corrected estimator, $\tilde{\alpha}$, and when $\mu_v = 1$. Results in Table A1 show more clearly the asymptotic distribution derived for $\tilde{\alpha}$ which is also used for $\mathring{\alpha}$. Again, we only report results for values of α over the range $[0.70, 1.0]$. Recall that α is identified only if $\alpha > 1/2$, and for asymptotically valid inference on α it is further required that $\alpha > 4/7$, unless $T^{1/2}/N^{(4\alpha-2)} \rightarrow 0$, as N and $T \rightarrow \infty$ in the case of $\tilde{\alpha}$ (see Theorem 1), or that $\alpha > 2/3$ in the case of $\mathring{\alpha}$ (see Supplementary Appendix V).

It appears that estimator $\tilde{\alpha}$ performs reasonably well in terms of bias and RMSE for values of α in the range $[0.70 - 0.85]$, when $\mu_v = 1$. To get a clearer picture of the asymptotics we turn to the right-hand-side of Table A1 that summarizes the size and power of the tests based on $\tilde{\alpha}$. There is evidence of some size distortion when α is below 0.75, but it tends towards the nominal 5% level as α is increased. The size distortion is also reduced as N and T are increased. The power of the test also rises in α, N and T, and approaches unity quite rapidly. However, the power function seems to be asymmetric with the power tending to be higher for alternatives above the null (denoted by Power+) as compared to the alternatives below the null (denoted by Power-). This asymmetry is particularly marked for low values of α and disappears as α is increased.

Turning to the size and power of the tests based on $\mathring{\alpha}$, its superior properties are verified by the results shown on the right-hand-side of Table A2. Indeed, in general size tends to zero as α increases towards 1 and as N and T increase. Similarly, power is uniformly close to unity irrespective of the value of α chosen or the N and T

combination considered (low power is only recorded for the smallest value of α considered and for small N and T combinations).

Finally, we present results for Experiment A when $\alpha = \alpha_1 = \alpha_2$ in Table A3. Compared with Table A-B, both bias and RMSE results are more elevated for estimator $\mathring{\alpha}$ for all values of α when we impose the two factors to be of the same strength in the data generating process. This is expected given the discussion in Supplementary Appendix V. Consistent with the baseline case, both the bias and RMSE of $\mathring{\alpha}$ fall gradually as N, T, and α are increased.

Calibration of $\bar{R}_N^2$

In order to select an appropriate $\bar{R}_N^2$ for the Monte Carlo simulation study of Section 4 and Supplementary Appendix VI, we computed $\bar{R}^2 s$ for the regressions, (36) summated based on data from a number of empirical applications. For each data set we first calculated $\mathring{\alpha}$ corresponding to $\breve{\alpha}$ and selected the strong $N^{\mathring{\alpha}}$ units. This resulted in a modified data set, $\boldsymbol{x}^{(s)} = \left[x_{it}^{(s)}\right]$ of dimension $T \times N^{\mathring{\alpha}}$ (elements of $\boldsymbol{x}^{(s)}$ were standardised to have unit variance). Then, we extracted the principal components (pc) from $\boldsymbol{x}^{(s)}$ and run the regression

$$x_{it}^{(s)} = a_i + \gamma_{ij} pc_j + \varepsilon_{it}, \tag{B48}$$

for $i = 1, 2, ..., N^{\mathring{\alpha}}$, and $t = 1, 2, ..., T$. We set the number of principal components to include in (B48) to $j = 1, 2, 3$, respectively. Finally, we computed the R^2 of each of the $N^{\mathring{\alpha}}$ regressions and took their average: $\bar{R}_N^2 = \frac{1}{N^{\mathring{\alpha}}} \sum_{i=1}^{N^{\mathring{\alpha}}} R_i^2$. We conducted this analysis for a number of empirical applications, of which: (i) GVAR macro economic data sets (real GDP growth - $\bar{R}_N^2 = 0.28, 0.37, 0.44$, inflation - $\bar{R}_N^2 = 0.47, 0.57, 0.64$, real equity price change - $\bar{R}_N^2 = 0.47, 0.59, 0.66$), and (ii) US - $\bar{R}_N^2 = 0.30, 0.50, 0.60$ - and UK - $\bar{R}_N^2 = 0.25, 0.43, 0.52$, all using $j = 1, 2, 3$ principal components, respectively. See also Section 5 for further details of the data sets.

Table A1: Bias, RMSE, size and power ($\times 100$) for the $\tilde{\alpha}$ estimate of the cross-sectional exponent - case of two serially independent factors and cross-sectionally independent idiosyncratic errors

($\alpha_2 = 2\alpha/3$, f_{jt} and $u_{it} \sim IIDN(0,1)$, $v_{ij} \sim IIDU(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, $j = 1, 2$, $\mu_v = 1$, $\mu_{v_2} = 0.87$, $\mu_{v_1} = \sqrt{\mu_v^2 - N^{2(\alpha_2 - \alpha)}\mu_{v_2}^2}$)
N=100,200,500,1000 and T=100,200,500

	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
N\T					**100**			
100	Bias	-0.47	-0.66	-0.73	-0.31	-0.28	-0.31	-0.19
	RMSE	2.07	1.97	1.90	1.71	1.66	1.63	1.59
200	Bias	-0.66	-0.23	-0.24	-0.17	-0.20	-0.14	-0.07
	RMSE	1.80	1.57	1.50	1.44	1.42	1.40	1.38
500	Bias	-0.17	-0.19	-0.10	-0.14	-0.10	-0.08	-0.06
	RMSE	1.32	1.25	1.20	1.19	1.17	1.16	1.16
1000	Bias	-0.22	-0.17	-0.11	-0.12	-0.09	-0.10	-0.08
	RMSE	1.16	1.10	1.06	1.04	1.03	1.03	1.02
					200			
100	Bias	-0.49	-0.62	-0.65	-0.22	-0.18	-0.21	-0.09
	RMSE	1.57	1.50	1.43	1.23	1.19	1.17	1.13
200	Bias	-0.47	-0.14	-0.20	-0.14	-0.18	-0.13	-0.06
	RMSE	1.26	1.09	1.05	1.00	0.99	0.97	0.95
500	Bias	-0.17	-0.19	-0.10	-0.14	-0.10	-0.08	-0.06
	RMSE	0.96	0.92	0.87	0.86	0.84	0.83	0.83
1000	Bias	-0.09	-0.08	-0.02	-0.04	-0.01	-0.03	-0.01
	RMSE	0.79	0.76	0.74	0.73	0.72	0.72	0.72
					500			
100	Bias	-0.17	-0.44	-0.56	-0.16	-0.14	-0.19	-0.06
	RMSE	1.00	1.02	1.03	0.83	0.80	0.79	0.75
200	Bias	-0.39	-0.08	-0.14	-0.10	-0.14	-0.08	-0.02
	RMSE	0.87	0.72	0.69	0.66	0.65	0.63	0.61
500	Bias	-0.13	-0.14	-0.05	-0.09	-0.05	-0.03	-0.01
	RMSE	0.63	0.59	0.55	0.54	0.52	0.51	0.51
1000	Bias	-0.19	-0.12	-0.04	-0.05	-0.02	-0.03	-0.01
	RMSE	0.56	0.51	0.49	0.48	0.47	0.46	0.46

	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
N\T					**100**			
100	Size	2.65	3.55	5.15	3.80	4.45	4.65	6.10
	Power+	58.70	76.05	85.20	83.60	86.55	88.70	86.35
	Power-	58.70	55.65	62.90	78.20	81.55	83.25	86.35
200	Size	0.05	0.80	2.10	4.70	5.85	6.10	6.80
	Power+	55.75	77.85	90.45	93.65	95.10	95.30	94.65
	Power-	55.75	65.15	80.60	89.25	91.55	93.30	94.65
500	Size	7.35	7.25	6.90	7.05	6.85	6.85	7.20
	Power+	97.80	98.60	98.90	99.15	99.15	99.05	98.90
	Power-	97.80	97.20	98.10	98.25	98.70	98.95	98.90
1000	Size	3.95	5.80	6.05	6.65	6.40	6.65	6.75
	Power+	99.65	99.80	99.85	99.95	100.00	100.00	99.55
	Power-	99.65	99.00	99.45	99.40	99.45	99.50	99.55
					200			
100	Size	0.00	0.05	0.70	1.20	2.20	3.25	4.50
	Power+	23.00	69.30	92.10	94.15	97.35	98.25	99.05
	Power-	23.00	39.75	63.20	87.95	95.75	98.00	99.05
200	Size	6.05	4.40	4.70	4.40	4.70	4.50	4.55
	Power+	99.25	99.45	99.85	99.90	99.95	99.95	100.00
	Power-	99.25	98.80	99.70	99.90	99.95	99.95	100.00
500	Size	5.35	6.10	6.00	6.15	5.80	5.80	6.50
	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	100.00	99.95	99.95	100.00	100.00	100.00	100.00
1000	Size	6.60	5.50	5.10	5.10	5.20	5.25	5.30
	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	100.00	100.00	100.00	100.00	100.00	100.00	100.00
					500			
100	Size	3.40	6.45	9.10	4.30	4.85	4.80	5.45
	Power+	99.15	99.95	100.00	100.00	100.00	100.00	100.00
	Power-	99.15	99.20	99.85	100.00	100.00	100.00	100.00
200	Size	8.05	4.55	5.05	4.55	4.60	3.80	4.65
	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	100.00	100.00	100.00	100.00	100.00	100.00	100.00
500	Size	0.25	1.50	2.85	4.20	4.05	4.45	5.00
	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	Size	0.00	0.20	1.55	4.05	5.05	5.70	6.15
	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table A2: Bias, RMSE, size and power (×100) for the $\mathring{\alpha}$ estimate of the cross-sectional exponent - case of two serially independent factors and cross-sectionally independent idiosyncratic errors

($\alpha_2 = 2\alpha/3$, f_{jt} and $u_{it} \sim IIDN(0,1)$, $v_{ij} \sim IIDU(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, $j = 1, 2$, $\mu_v = 0.87$, $\mu_{v_2} = 0.71$, $\mu_{v_1} = \sqrt{\mu_v^2 - N^{2(\alpha_2-\alpha)}\mu_{v_2}^2}$)

N=100,200,500,1000 and T=100,200,500

N\T	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
					T=100			
100	Bias	0.97	0.45	0.02	0.27	0.15	-0.02	-0.04
	RMSE	2.02	1.44	0.93	0.71	0.48	0.29	0.06
200	Bias	0.52	0.43	0.19	0.15	0.04	0.03	0.03
	RMSE	1.48	0.99	0.63	0.46	0.29	0.17	0.03
500	Bias	0.13	0.05	0.09	0.03	0.04	0.03	0.05
	RMSE	0.68	0.46	0.35	0.23	0.16	0.10	0.06
1000	Bias	0.00	0.01	0.05	0.01	0.03	0.00	0.06
	RMSE	0.53	0.35	0.25	0.17	0.12	0.07	0.06
					200			
100	Bias	2.51	1.55	0.62	0.60	0.31	0.05	-0.10
	RMSE	3.12	2.19	1.28	0.99	0.61	0.33	0.10
200	Bias	0.45	0.54	0.28	0.21	0.08	0.04	-0.02
	RMSE	1.14	0.96	0.61	0.45	0.28	0.15	0.02
500	Bias	0.24	0.12	0.14	0.06	0.06	0.04	0.02
	RMSE	0.61	0.41	0.31	0.19	0.14	0.08	0.02
1000	Bias	0.10	0.08	0.10	0.06	0.07	0.03	0.03
	RMSE	0.38	0.26	0.21	0.14	0.10	0.05	0.03
					500			
100	Bias	4.02	3.06	1.75	1.38	0.79	0.31	-0.13
	RMSE	4.28	3.30	2.05	1.63	1.02	0.53	0.13
200	Bias	1.75	1.56	0.88	0.64	0.31	0.11	-0.05
	RMSE	2.08	1.83	1.14	0.86	0.51	0.23	0.05
500	Bias	0.90	0.47	0.31	0.13	0.09	0.04	-0.01
	RMSE	1.17	0.72	0.48	0.27	0.17	0.09	0.01
1000	Bias	0.66	0.23	0.14	0.07	0.06	0.02	0.00
	RMSE	0.90	0.42	0.25	0.15	0.10	0.04	0.00

N\T	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
					100			
100	Size	3.00	1.25	0.05	0.00	0.00	0.00	0.00
	Power+	27.65	54.30	82.60	91.45	98.90	100.00	100.00
	Power-	27.65	72.35	83.95	97.95	99.90	100.00	100.00
200	Size	0.55	0.00	0.00	0.00	0.00	0.00	0.00
	Power+	21.80	67.80	97.10	99.80	100.00	100.00	100.00
	Power-	21.80	87.25	99.15	100.00	100.00	100.00	100.00
500	Size	0.45	0.00	0.00	0.00	0.00	0.00	0.00
	Power+	99.70	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	99.70	100.00	100.00	100.00	100.00	100.00	100.00
1000	Size	0.05	0.00	0.00	0.00	0.00	0.00	0.00
	Power+	99.85	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	99.85	100.00	100.00	100.00	100.00	100.00	100.00
					200			
100	Size	3.55	2.00	0.45	0.05	0.10	0.00	0.00
	Power+	0.55	13.10	59.00	85.15	99.10	100.00	100.00
	Power-	0.55	76.35	86.05	99.70	100.00	100.00	100.00
200	Size	4.00	3.60	0.40	0.10	0.00	0.00	0.00
	Power+	92.85	98.20	100.00	100.00	100.00	100.00	100.00
	Power-	92.85	100.00	100.00	100.00	100.00	100.00	100.00
500	Size	1.00	0.10	0.05	0.00	0.00	0.00	0.00
	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	Size	0.15	0.00	0.00	0.00	0.00	0.00	0.00
	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	100.00	100.00	100.00	100.00	100.00	100.00	100.00
					500			
100	Size	81.05	71.85	36.85	27.95	6.20	0.15	0.00
	Power+	13.20	33.75	83.00	96.95	99.95	100.00	100.00
	Power-	13.20	100.00	100.00	100.00	100.00	100.00	100.00
200	Size	47.45	47.55	20.05	10.10	1.20	0.00	0.00
	Power+	87.05	96.95	100.00	100.00	100.00	100.00	100.00
	Power-	87.05	100.00	100.00	100.00	100.00	100.00	100.00
500	Size	6.05	3.75	1.75	0.25	0.00	0.00	0.00
	Power+	99.10	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	99.10	100.00	100.00	100.00	100.00	100.00	100.00
1000	Size	0.65	0.10	0.00	0.00	0.00	0.00	0.00
	Power+	99.85	100.00	100.00	100.00	100.00	100.00	100.00
	Power-	99.85	100.00	100.00	100.00	100.00	100.00	100.00

Table A3: Bias and RMSE (×100) for the $\mathring{\alpha}$ estimate of the cross-sectional exponent - case of two serially independent factors and cross-sectionally independent idiosyncratic errors ($\alpha_2 = \alpha$, f_{jt} and $u_{it} \sim IIDN(0,1)$, $v_{ij} \sim IIDU(\mu_{v_j} - 0.2, \mu_{v_j} + 0.2)$, $j = 1, 2$, $\mu_{v_1} = 0.5$, $\mu_{v_2} = 0.5$) N=100,200,500,1000 and T=100,200,500

	α	0.70	0.75	0.80	0.85	0.90	0.95	1.00
N\T					**100**			
100	Bias	9.41	8.51	7.73	6.96	5.45	3.28	-0.04
	RMSE	9.55	8.63	7.82	7.03	5.51	3.33	0.06
200	Bias	8.92	8.41	7.54	6.49	5.04	3.12	0.03
	RMSE	9.00	8.47	7.59	6.53	5.08	3.14	0.04
500	Bias	7.39	7.01	6.52	5.66	4.50	2.81	0.05
	RMSE	7.45	7.06	6.56	5.69	4.53	2.83	0.06
1000	Bias	6.82	6.49	6.00	5.21	4.14	2.56	0.05
	RMSE	6.86	6.53	6.03	5.24	4.17	2.58	0.06
					200			
100	Bias	10.95	9.84	8.91	8.00	6.29	3.72	-0.10
	RMSE	11.03	9.90	8.97	8.04	6.32	3.75	0.10
200	Bias	9.60	9.35	8.56	7.51	5.92	3.68	-0.02
	RMSE	9.64	9.38	8.58	7.53	5.93	3.69	0.02
500	Bias	8.76	8.36	7.81	6.88	5.56	3.50	0.02
	RMSE	8.78	8.37	7.82	6.89	5.56	3.51	0.02
1000	Bias	8.08	7.77	7.29	6.46	5.26	3.36	0.03
	RMSE	8.09	7.78	7.29	6.47	5.27	3.36	0.03
					500			
100	Bias	11.54	10.39	9.39	8.35	6.52	3.84	-0.12
	RMSE	11.60	10.43	9.43	8.38	6.54	3.86	0.12
200	Bias	10.26	9.90	9.07	7.94	6.26	3.83	-0.04
	RMSE	10.28	9.92	9.09	7.96	6.27	3.84	0.04
500	Bias	9.40	8.97	8.42	7.44	6.00	3.75	-0.01
	RMSE	9.41	8.98	8.43	7.44	6.00	3.75	0.01
1000	Bias	8.89	8.43	7.93	7.08	5.77	3.67	0.00
	RMSE	8.89	8.43	7.94	7.08	5.77	3.67	0.00