

IZA DP No. 6318

Exponent of Cross-sectional Dependence: Estimation and Inference

Natalia Bailey
George Kapetanios
M. Hashem Pesaran

January 2012

Exponent of Cross-sectional Dependence: Estimation and Inference

Natalia Bailey

University of Cambridge

George Kapetanios

Queen Mary, University of London

M. Hashem Pesaran

*University of Cambridge,
University of Southern California and IZA*

Discussion Paper No. 6318
January 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Exponent of Cross-sectional Dependence: Estimation and Inference^{*}

An important issue in the analysis of cross-sectional dependence which has received renewed interest in the past few years is the need for a better understanding of the extent and nature of such cross dependencies. In this paper we focus on measures of cross-sectional dependence and how such measures are related to the behaviour of the aggregates defined as cross-sectional averages. We endeavour to determine the rate at which the cross-sectional weighted average of a set of variables appropriately demeaned, tends to zero. One parameterisation sets this to be $O(N^{2\alpha-2})$, for $1/2 < \alpha \leq 1$. Given the fashion in which it arises, we refer to as the exponent of cross-sectional dependence. We derive an estimator of α from the estimated variance of the cross-sectional average of the variables under consideration. 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. Finally, we undertake an empirical investigation of using the S&P 500 data-set, and a large number of macroeconomic variables across and within countries.

JEL Classification: C21, C32

Keywords: cross correlations, cross-sectional dependence, cross-sectional averages, weak and strong factor models, Capital Asset Pricing Model

Corresponding author:

M. Hashem Pesaran
Trinity College
University of Cambridge
CB2 1TQ Cambridge
United Kingdom
E-mail: mhp1@cam.ac.uk

^{*} 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.

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 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.

Perhaps, the simplest and most concise way to motivate the need for determining the extent of cross-sectional dependence is to view the matter from a simple statistical viewpoint. Let x_{it} denote a double array of random variables indexed by $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$, over space and time, respectively. Then, a variety of analyses focus on weighted averages of x_{it} over i . Examples include the construction of portfolios, when x_{it} are asset returns, or aggregate macroeconomic variables, when x_{it} are firm or consumer level data, such as firm sales, or individual consumption. Weighted averages take the form $\bar{x}_{wt} = \sum_{i=1}^N w_{Ni} x_{it}$, where the weights w_{Ni} are granular in the sense that $w_{Ni} = O\left(\frac{1}{N}\right)$. For simplicity, we set $w_{Ni} = \frac{1}{N}$ and $\bar{x}_{wt} = \bar{x}_t$. Then, it is of considerable interest to determine the behaviour of \bar{x}_t and, in particular, the rate at which \bar{x}_t , when appropriately standardised, tends to zero. 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 $\bar{x}_t = O_p(N^{-1/2})$. However, in the more general and realistic case where x_{it} are cross-sectionally correlated we have¹

$$Var(\bar{x}_t) = \frac{1}{N^2} \sum_{i=1}^N \sigma_{i,x}^2 + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \sigma_{ij,x} = N^{-1} \bar{\sigma}_{xN}^2 + \tau_N, \quad (1)$$

where $\sigma_{i,x}^2 = Var(x_{it})$ and $\sigma_{ij,x} = Cov(x_{it}, x_{jt})$, assuming, for simplicity, stationarity, over time, for x_{it} . The above result suggests two important conditions under which standard law of large numbers may fail to apply to the cross-sectional averages. One relates to the absence of finite variances for individual x_{it} ,² while the other to the rate at which the remainder term, $\tau_N = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \sigma_{ij,x}$, grows with N . Determining the extent of cross-sectional dependence, relates directly to the second way, in which standard law of large numbers may not be applicable, and lead to phenomena of considerable interest as discussed above. It is, therefore, of interest to investigate the rate at which τ_N declines with N . We may parametrise this by letting $\tau_N = O(N^{2\alpha-2})$, with α measuring the degree of cross-sectional dependence. We note that $Var(\bar{x}_t)$ cannot decline at a rate lower than N^{-1} since $\frac{\bar{\sigma}_{xN}^2}{N} = O(N^{-1})$.

Consider now the estimator of τ_N , based on sample covariances, $\hat{\sigma}_{ij,x}$, namely $\hat{\tau}_N = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \hat{\sigma}_{ij,x}$. It is easily seen that $\hat{\tau}_N = \hat{\sigma}_{\bar{x}}^2 + O_p(N^{-1})$, where $\hat{\sigma}_{\bar{x}}^2 = \frac{1}{T} \sum_{t=1}^T \left(\bar{x}_t - \frac{1}{T} \sum_{t=1}^T \bar{x}_t \right)^2$. Therefore, the rate at which $Var(\bar{x}_t)$ tends to zero with N , is governed by the rate at which $\hat{\sigma}_{\bar{x}}^2$ tends to zero. But

¹(1) can be equivalently stated in terms of correlations, but since correlations are more complex to handle we choose to carry out the analysis in terms of covariances.

²This way has been recently explored, in the case of firm sales data, by Gabaix (2011), who noted that the failure of such data to have finite variances can lead to idiosyncratic firm shocks determining, to a considerable extent, GDP growth volatility.

this rate can not be faster than N^{-1} , and hence the range of interest for α must lie in the range $-1 < 2\alpha - 2 \leq 0$, or $1/2 < \alpha \leq 1$. This paper focusses on the problem of identification, estimation and inference regarding α , which we refer to as the exponent of cross-sectional dependence, bearing in mind that α is defined by $\tau_N = O(N^{2\alpha-2})$.

However, it is important to note that different problems may require the determination of the exponent of different summaries of the covariance matrix. For example, a summary focused upon by Chudik, Pesaran, and Tosetti (2011) is the column sum norm of the covariance matrix. The exponent of that is given by the parametrisation $\max_i \sum_{j=1}^N |\sigma_{ij,x}| = O(N^\alpha)$. It is clear that this exponent need not be the same as the one relating to τ_N . Therefore, while the exponent of cross-sectional dependence is of great interest for the phenomena outlined above, one needs to be clear about the motivation for using this measure of cross-sectional dependence.

Also, other measures of cross-sectional dependence can be considered. An important example is the measure based on the upper bound for $Var(\bar{x}_t)$, which is $N^{-1}\lambda_{\max}(\Sigma_N)$, where $\Sigma_N = E(\mathbf{x}_t\mathbf{x}_t')$, $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{Nt})'$ and $\lambda_{\max}(\Sigma_N)$ denotes the maximum eigenvalue of Σ_N . $\lambda_{\max}(\Sigma_N)$ is an object of considerable interest in the statistical literature on large data sets. However, work in the area (see, e.g., Yin, Bai, and Krishnaiah (1988), Bai and Silverstein (1998), Hachem, Loubaton, and Najim (2005a) and Hachem, Loubaton, and Najim (2005b)) suggests that as a statistical measure of cross-sectional dependence $\lambda_{\max}(\Sigma_N)$ could be difficult to analyse especially for temporally and cross-sectionally dependent data. This is partly due to the fact that estimates of $\lambda_{\max}(\Sigma_N)$ based on sample estimates of Σ_N , could be very poor when N is large relative to T , which is the type of data sets often encountered in macroeconomics and finance. Therefore, we do not pursue its analysis in this paper, although we acknowledge the need for further investigations in this area, comparing the performance of estimates of cross-sectional dependence based on α and on $N^{-1}\lambda_{\max}(\Sigma_N)$.

The above measures of cross-sectional dependence are related to the degree of pervasiveness of factors in unobserved factor models often used in the literature to model cross-sectional dependence.³ Consider the following canonical factor model

$$x_{it} = a_i + \beta_i' \mathbf{f}_t + u_{it},$$

where \mathbf{f}_t is the $m \times 1$ vector of unobserved factors (m being fixed), and $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{im})'$ is the associated vector of factor loadings, and $Cov(u_{it}, u_{jt}) = \omega_{ij}$. The extent of cross-sectional dependence in x_{it} crucially depends on the nature of factor loadings. The degree of cross-sectional dependence will be strong if β_i is bounded away from 0 and the average value of β_i is different from zero. In such a case $\sup_i N^{-1} \sum_{j=1}^N |\sigma_{ij,x}| = O(1)$ and $\tau_N = N^{-2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \sigma_{ij,x} = O(1)$, for all i , which yields $\alpha = 1$. However, other configurations of factor loadings can also be entertained, that yield values of α in the range $(1/2, 1]$. Since both \mathbf{f}_t and β_i are unobserved, taking a strong stand on a particular value of α might not be justified empirically. Accordingly, Chudik, Pesaran, and Tosetti (2011), Kapetanios and Marcellino (2010) and Onatski (2011) have considered an extension of the above factor model which allows for a wider spectrum of cross-sectional dependence behaviours by specifying that the factor loadings β_i may be functions of N and decline in magnitude as $N \rightarrow \infty$ allowing for the possibility of factors having a weaker effect than is the case for standard factor models.⁴ One such formulation assumes that factor loadings decline with N , and $\beta_i = O(N^{\alpha-1})$, for any $\alpha < 1$. This specification has the same order for $N^{-2} \sum_{i=1}^N \sum_{j=1}^N \beta_i' \beta_j = \left(N^{-1} \sum_{i=1}^N \beta_i \right)' \left(N^{-1} \sum_{j=1}^N \beta_j \right) = O(N^{2\alpha-2})$, so long

³Factor 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), Stock and Watson (2002), and Pesaran (2006).

⁴While Chudik, Pesaran, and Tosetti (2011) and Kapetanios and Marcellino (2010) have considered the case where $1/2 < \alpha < 1$, Onatski (2011) has focused on the case where $\alpha \leq 1/2$.

as $N^{-2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \omega_{ij}$ is of smaller order of magnitude than τ_N . The latter condition is satisfied under an approximate factor model, so long as $2(\alpha - 1) > -1$, or if $\alpha > 1/2$.

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 allow β_i to be fixed in N , but assume that only N^α of the N factor loadings are individually important, in the sense that they are bounded away from zero in absolute value.⁵ More specifically, we consider

$$\begin{aligned} \beta_{ik} &= v_{ik} \text{ for } i = 1, 2, \dots, [N^{\alpha_k}], \\ \beta_{ik} &= c_k \rho_k^{i - [N^{\alpha_k}] + 1}, \text{ for } i = [N^{\alpha_k}] + 1, [N^{\alpha_k}] + 2, \dots, N, \end{aligned} \quad (2)$$

for $k = 1, 2, \dots, m$, where $[N^{\alpha_k}]$ is the integer part of N^{α_k} , $1/2 < \alpha_k \leq 1$, $|\rho_k| < 1$, c_k is a finite constant, and $v_{ik} \sim iid(\mu_{v_k}, \sigma_{v_k}^2)$, with $\mu_{v_k} \neq 0$ and $\sigma_{v_k}^2 > 0$. In effect, the factor loadings are grouped into two categories, a strong category with effects that are bounded away from zero, and a weak category with transitory effects that tend to zero exponentially. The focus of our analysis is on $\alpha = \max_k(\alpha_k)$, which is the *cross-sectional exponent* as defined above. 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. Also the exponential decay assumed for the second group of factor loadings can be relaxed and replaced by the absolute summability condition, $\sum_{i=[N^{\alpha_k}] + 1}^N |\beta_{ik}| < \infty$, which is again invariant to the ordering of the units with weak dependence on factors. It is important to recognize that while this formulation is the one we focus on, alternative formulations can give the same exponent. While some details of our inferential theory does depend on the nature of the specific formulation adopted, it is clear that very similar inferential procedure with similar asymptotic and small sample properties can be easily developed for alternative formulations.

In cases where the common factors are observed the strength of the factors ρ_k as measured by α can be estimated directly in terms of the number of statistically significant β_i coefficients. Denoting the number of such statistically significant estimates of the loadings associated with the i^{th} factor by \hat{M}_i , the estimates of α_i can then be obtained as $\ln(\hat{M}_i) / \ln(N)$. We shall consider such a "direct" estimate of α in our empirical application to Capital Asset Pricing Model.

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 . Focusing, for simplicity, on a single factor representation, we show that⁶

$$Var(\bar{x}_t) = \kappa^2 [N^{2\alpha-2}] + N^{-1} c_N + O(N^{\alpha-2}),$$

where $\kappa^2 = \sigma_f^2 \mu_v^2$, σ_f^2 is the variance of the factor process, μ_v is the mean of the factor loadings, and c_N is a bias term that is analysed in detail in the main body of the paper. Using this relationship, we can provide a feasible estimator for α and derive inferential theory for it. The property of our proposed estimator depends on the choice of κ . For an arbitrary but bounded value of κ , our estimator is consistent but its rate of convergence at $1/\ln(N)$ is rather slow. However, given the identified nature of σ_f^2 in factor models it seems more sensible to consider estimating α for a given value of σ_f^2 . In the factor literature the factor loadings are typically identified by setting $\sigma_f^2 = 1$, assuming that the factors are strong, and the idiosyncratic components are weakly cross correlated. However, since the aim here is to estimate α , it is more sensible to fix κ^2 . The rationale behind this alternative normalization, and its empirical implications will be discussed in detail. Further, we propose a second order bias corrected

⁵Note that α in our baseline formulation is not directly comparable to the α of the almost uniform loadings formulation.

⁶We consider a general multi-factor representation, in detail, in Section 3.2.

estimator that addresses the term c_N , and derive the asymptotic distribution of both estimators for a given value of κ^2 . We consider extensions that relate to the presence of multiple factors, potentially with multiple distinct exponents of cross-sectional dependence, temporal dependence in f_t or u_{it} , and weak cross-sectional dependence in u_{it} . It is worth noting that our baseline estimator is equivalent to one obtained by setting up a regression framework whereby the logarithm of the partial sum process of estimated factor loadings is regressed upon the logarithm of the cross-sectional dimension of the partial sum.

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, and present recursive estimates of α applied to excess returns on securities included in the Standard & Poor 500 index. In these applications to ensure that our estimates of α are invariant to the scales of measurement of the observations, x_{it} , we base our analysis on the standardized variates, $(x_{it} - \bar{x}_i)/s_i$, where \bar{x}_i and s_i are sample mean and standard deviation of x_{it} , respectively, over the sample under consideration.

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. Sections 3-3.3 present the theoretical results of the paper. Section 3 provides the full inferential theory and discusses feasible estimation, including estimators for the variance of the estimator of cross-sectional dependence. Section 3.2 presents an extension of the theoretical analysis to a multi-factor setting and briefly touches upon an alternative specification of factor loadings. Section 3.3 considers the normalization $\sigma_f^2 \mu_v^2 = \kappa^2$, and Section 4 discusses the conditions under which estimation of α and κ^2 can be carried out jointly. Section 5 presents a detailed Monte Carlo study. The empirical applications are discussed in Section 6. Finally, Section 7 concludes. Proofs of all theoretical results are relegated to the Appendix.

2 Preliminaries and Motivations

As noted in the Introduction, to characterize the degree of cross-sectional dependence in x_{it} we use a possibly weak factor model. We begin with the following single factor specification

$$x_{it} = a_i + \beta_i f_t + u_{it} \text{ for } i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T, \quad (3)$$

where f_t is an unobserved factor, β_i are the associated factor loadings and a_i are bounded constants such that $\sup_i |a_i| < K < \infty$. We make the following assumptions.

Assumption 1 *The factor loadings are given by*

$$\begin{aligned} \beta_i &= v_i \text{ for } i = 1, 2, \dots, [N^\alpha], \\ \beta_i &= c\rho^{i-[N^\alpha]+1}, \text{ for } i = [N^\alpha] + 1, [N^\alpha] + 2, \dots, N, \end{aligned} \quad (4)$$

where $1/2 < \alpha \leq 1$, $[N^\alpha]$ is the integer part of N^α , $|\rho| < 1$, and $\{v_i\}_{i=1}^{[N^\alpha]}$ is an identically, independently distributed (IID) sequence of random variables with mean $\mu_v \neq 0$, and variance $\sigma_v^2 < \infty$.

Assumption 2 The factor, f_t , follows a linear stationary process given by

$$f_t = \sum_{j=0}^{\infty} \psi_{f,j} \nu_{f,t-j}, \quad (5)$$

where ν_{f_t} is an IID sequence of random variables with mean zero and finite variance and uniformly finite φ -th moments for some $\varphi > 4$. We assume that

$$\sum_{j=0}^{\infty} j^{\zeta} |\psi_{fj}| < \infty,$$

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

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

$$u_{it} = \sum_{j=0}^{\infty} \psi_{ij} \left(\sum_{s=-\infty}^{\infty} \xi_{js} \nu_{j,t-s} \right), \quad (6)$$

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

$$\sup_i \sum_{j=0}^{\infty} j^{\zeta} |\psi_{ij}| < \infty, \quad (7)$$

and

$$\sup_i \sum_{s=-\infty}^{\infty} |s|^{\zeta} |\xi_{is}| < \infty \quad (8)$$

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

It is worth briefly commenting on these assumptions. Assumption 1 has been motivated in Section 1, and implies that $N^{-1} \sum_{i=1}^N \beta_i^2 = O_p(N^{\alpha-1})$, which is more general than the standard assumption in the factor literature that requires $N^{-1} \sum_{i=1}^N \beta_i^2$ to have a strictly positive limit (see, e.g., Assumption B of Bai and Ng (2002)). It is easy to see that the standard assumption is satisfied only if $\alpha = 1$. 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. The only noteworthy part of these assumptions relates to the cross-sectional dependence of the error terms. Here, cross-sectional dependence is structured in a flexible linear way so as to mirror, to the extent possible, the conditions assumed for stationary (weak) time series dependence. From Assumption 3 it follows that $E(u_{it}^2) = \sigma_i^2 = \sigma_{\nu_i}^2 \left(\sum_{s=-\infty}^{\infty} \xi_{is}^2 \right) \left(\sum_{s=0}^{\infty} \psi_{is}^2 \right) < \infty$.

A generalization of the factor model and the related Assumptions, 1 and 2, will be considered in Section 3.2. In the rest of this Section, we motivate our proposed estimator for α .

We write (3) as

$$\mathbf{x}_t = \mathbf{a} + \boldsymbol{\beta} f_t + \mathbf{u}_t,$$

where $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{Nt})'$, $\mathbf{a} = (a_1, a_2, \dots, a_N)'$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_N)'$ and $\mathbf{u}_t = (u_{1t}, u_{2t}, \dots, u_{Nt})'$. We also note that under the above assumptions, $\boldsymbol{\Sigma}_{\beta} = E(\boldsymbol{\beta}\boldsymbol{\beta}') - E(\boldsymbol{\beta})E(\boldsymbol{\beta}')$, with $\lambda_{\max}(\boldsymbol{\Sigma}_{\beta}) < K < \infty$;

$\Sigma_u = E(\mathbf{u}_t \mathbf{u}_t')$, with $\lambda_{\max}(\Sigma_u) < K < \infty$, $E(f_t) = 0$, $E(f_t^2) = \sigma_f^2 > 0$, and f_t and β are distributed independently. Hence

$$\mathbf{x}_t - E(\mathbf{x}_t) = \beta f_t + \mathbf{u}_t,$$

$$\begin{aligned} Cov(\mathbf{x}_t) &= E[(\beta f_t + \mathbf{u}_t)(\beta f_t + \mathbf{u}_t)'] \\ &= E(\beta \beta') E(f_t^2) + E(\mathbf{u}_t \mathbf{u}_t') \\ &= [\Sigma_\beta + E(\beta) E(\beta')] \sigma_f^2 + \Sigma_u. \end{aligned}$$

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

$$Var(\bar{x}_t) = N^{-2} \boldsymbol{\tau}'_N Cov(\mathbf{x}_t) \boldsymbol{\tau}_N = N^{-2} \boldsymbol{\tau}'_N [\sigma_f^2 \Sigma_\beta + \Sigma_u] \boldsymbol{\tau}_N + \sigma_f^2 \left[\frac{\boldsymbol{\tau}'_N E(\beta)}{N} \right]^2. \quad (9)$$

But under (4), it follows that

$$\sum_{i=1}^N \beta_i = [N^\alpha] \left(\frac{1}{[N^\alpha]} \sum_{i=1}^{[N^\alpha]} v_i + \frac{c\rho}{[N^\alpha]} \left(\frac{1 - \rho^{(N - [N^\alpha])}}{1 - \rho} \right) \right) = (\bar{v}_N + O(N^{-\alpha})) [N^\alpha]$$

where $\bar{v}_N = \frac{1}{[N^\alpha]} \sum_{i=1}^{[N^\alpha]} v_i$ is $O_p(1)$ and for simplicity, we set $v_i = 0$, for $i > [N^\alpha]$. Note that the specification of β_i for $i > [N^\alpha]$ need not be of the form given in (4). Any sequence of loadings, for which $\sum_{i=[N^\alpha]+1}^N \beta_i = O_p(1)$ is acceptable. Hence

$$N^{-1} \boldsymbol{\tau}'_N E(\beta) = \mu_v [N^{\alpha-1}] + O(N^{-1}).$$

Also

$$N^{-2} \boldsymbol{\tau}'_N \Sigma_\beta \boldsymbol{\tau}_N = N^{-2} \boldsymbol{\tau}'_{1N} \Sigma_{\beta(1)} \boldsymbol{\tau}_{1N} \leq [N^{\alpha-2}] \lambda_{\max}(\Sigma_\beta),$$

where $\boldsymbol{\tau}_{1N}$ is an $[N^\alpha] \times 1$ vector of ones and $\Sigma_{\beta(1)}$ is the upper $[N^\alpha] \times [N^\alpha]$ sub-matrix of Σ_β . Using the above results in (9) we now have

$$Var(\bar{x}_t) \leq [N^{\alpha-2}] \sigma_f^2 \lambda_{\max}(\Sigma_\beta) + N^{-1} c_N + \kappa^2 [N^{2\alpha-2}].$$

where

$$\frac{\boldsymbol{\tau}'_N \Sigma_u \boldsymbol{\tau}_N}{N} = c_N < K < \infty. \quad (10)$$

But, by assumption $\lambda_{\max}(\Sigma_\beta) < K < \infty$, and hence under $1 \geq \alpha > 1/2$ we have

$$Var(\bar{x}_t) = \kappa^2 [N^{2\alpha-2}] + N^{-1} c_N + O(N^{\alpha-2}). \quad (11)$$

Depending on how σ_f^2 is normalized, different estimators of α can be envisaged. Since, by assumption, $\mu_v \neq 0$, a natural normalization would be $\sigma_f^2 = 1/\mu_v^2$ or $\kappa^2 = 1$. Then, a simple manipulation of (11) yields

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

or

$$\alpha \approx 1 + \frac{1}{2} \frac{\ln(\sigma_{\bar{x}}^2)}{\ln(N)} - \frac{c_N}{2 [N \ln(N)] \sigma_{\bar{x}}^2}. \quad (12)$$

Note that the third term on the RHS of (12) is of smaller order of magnitude than the other two terms. In cases where $\alpha \leq 1/2$, the second term in the RHS of (11), that arises from the contribution of the idiosyncratic components, will be at least as important as the contribution of a weak factor, and, in consequence, α cannot be identified. However, for values of $\alpha > 1/2$, α can be identified from (12) using a consistent estimator of $Var(\bar{x}_t) = \sigma_{\bar{x}}^2$, given by

$$\hat{\sigma}_{\bar{x}}^2 = \frac{1}{T} \sum_{t=1}^T (\bar{x}_t - \bar{x})^2, \quad (13)$$

where $\bar{x} = T^{-1} \sum_{t=1}^T \bar{x}_t$. A simple consistent estimator of α is given by

$$\hat{\alpha} = 1 + \frac{1 \ln(\hat{\sigma}_{\bar{x}}^2)}{2 \ln(N)}. \quad (14)$$

Further, in the case of exact factor models where Σ_u is a diagonal matrix, the third term in (12) can be estimated by

$$\hat{c}_N = N^{-1} \sum_{j=1}^N \hat{\sigma}_j^2 = \widehat{\sigma}_N^2,$$

where σ_j^2 is the j^{th} diagonal term of Σ_u and $\hat{\sigma}_j^2$ is its estimator. This suggests the following modified estimator of α

$$\tilde{\alpha} = 1 + \frac{1 \ln(\hat{\sigma}_{\bar{x}}^2)}{2 \ln(N)} - \frac{\widehat{\sigma}_N^2}{2 [N \ln(N)] \hat{\sigma}_{\bar{x}}^2}. \quad (15)$$

Note that while \hat{c}_N , as an estimator for c_N , is motivated by appealing to an exact factor model, it is also valid for mild deviations from this model as discussed in the next Section. The above estimators of α form the basis of the formal analysis that is carried out in Section 3.

3 Theoretical Derivations

In this section we provide a formal analysis of our proposed estimators.

3.1 Main Results

Our first set of theoretical results characterise the asymptotic behaviour of $\hat{\alpha}$. Introducing the additional notations, $\bar{\sigma}_N^2 = N^{-1} \sum_{i=1}^N \sigma_i^2$, $s_f^2 = \frac{1}{T} \sum_{t=1}^T \left(f_t - \frac{1}{T} \sum_{t=1}^T f_t \right)^2$, $\mu_f = E(f_t)$, and $\sigma_f^2 = E(f_t - \mu_f)^2$, we have:

Theorem 1 *Let Assumptions 1-3 hold, $m = 1$ and $\alpha > 1/2$. Then,*

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

where

$$\alpha^* \equiv \alpha_N^* = \alpha + \frac{\ln(\kappa^2)}{2 \ln(N)}, \quad (16)$$

$$\omega = \lim_{N, T \rightarrow \infty} \left[\frac{\min(N^\alpha, T)}{T} V_{\tilde{f}^2} + \frac{\min(N^\alpha, T)}{N^\alpha} \frac{4\sigma_v^2}{\mu_v^2} \right], \quad (17)$$

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

and $\tilde{f}_t = (f_t - \mu_f)/\sigma_f$.

This theorem shows that $\hat{\alpha}$, as an estimator of α , is subject to two sources of bias. The first relates to the term $\ln(\kappa^2)/\ln(N)$ in α^* , which arises due to the way identification of the factor loadings depends on scaling of the factors as defined by σ_f^2 . For example, this term vanishes if σ_f^2 is normalised to $1/\mu_v^2$ (assuming μ_v is non-zero). Under this normalization $\kappa = 1$, and $\alpha^* = \alpha$. We will discuss the implications of this restriction, and potential ways to circumvent it under certain conditions, in Section 3.3.⁷ Clearly, κ^2 is an important determinant of cross-sectional dependence in small samples, and while its value is irrelevant for the probability limit of $\hat{\alpha}$, as shown in Theorem 1, one may wish to focus on α^* as a parameter of interest in small samples. The following corollary provides the asymptotic properties of $\hat{\alpha}$ when $\kappa^2 = 1$.

Corollary 1 *Suppose Assumptions 1 to 3 hold, $m = 1$ and $\alpha > 1/2$. Under the normalisation restriction $\kappa^2 = 1$,*

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

The second source of bias is the term $\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2}$ which is unobserved. A first order accurate estimator of this term is given by $\frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2}$ where

$$\widehat{\sigma}_N^2 = \frac{1}{N} \sum_{i=1}^N \hat{\sigma}_i^2, \quad \hat{\sigma}_i^2 = \frac{1}{T} \sum_{t=1}^T \hat{u}_{it}^2,$$

$\hat{u}_{it} = x_{it} - \hat{\delta}_i \tilde{x}_t$, $\tilde{x}_t = \frac{\bar{x}_t - \bar{x}}{\hat{\sigma}_x}$, and $\hat{\delta}_i$ denotes the OLS estimator of the regression coefficient of x_{it} on \tilde{x}_t . This suggests the following bias corrected estimator

$$\tilde{\alpha} = \hat{\alpha} - \frac{\widehat{\sigma}_N^2}{2 \ln(N) N \widehat{\sigma}_x^2}. \quad (18)$$

The following theorem presents the asymptotic properties of $\tilde{\alpha}$.

Theorem 2 *Let Assumptions 1-3 hold, $m = 1$ and $\alpha > 1/2$. Then, as long as either $\frac{T^{1/2}}{N^{4\alpha-2}} \rightarrow 0$ or $\alpha > 4/7$,*

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

where α^* and ω are defined in (16) and (17), respectively.

As noted above this estimator is only first order accurate since Theorem 2 only holds under the specified assumptions concerning α and the assumed relative rate of growth of N and T . Fortunately, a second order bias correction is available. This amounts to estimating $\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2}$ by $\frac{\widehat{\sigma}_N^2}{\ln(N) N \widehat{\sigma}_x^2} \left(1 + \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2} \right)$. Accordingly, we define a second bias corrected estimator

$$\check{\alpha} = \hat{\alpha} - \frac{\widehat{\sigma}_N^2}{2 \ln(N) N \widehat{\sigma}_x^2} \left(1 + \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2} \right), \quad (19)$$

and give its asymptotic distribution in the following theorem.

⁷Note that as shown in (51) the rate of convergence of $\hat{\alpha}$ to α without this restriction is a somewhat slow $\ln(N)$.

Theorem 3 Suppose Assumptions 1 to 3 hold, $m = 1$ and $\alpha > 1/2$. Then,

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

where α^* and ω are defined in (16) and (17), respectively.

We consider $\tilde{\alpha}$, even though $\check{\alpha}$ provides a more comprehensive bias correction because small sample evidence suggests that $\tilde{\alpha}$ may outperform $\check{\alpha}$ for values of α in the middle of the admissible range $(1/2, 1]$. Obviously, equivalent Corollaries to Corollary 1 hold for $\tilde{\alpha}$ and $\check{\alpha}$. The final result of this subsection relates to providing a consistent estimator for ω . Let \hat{v}_i denote the OLS estimator of the regression coefficient of x_{it} on \bar{x}_t . Let $\{\hat{\delta}_i^{(s)}\}_{i=1}^N$ and $\{\hat{v}_i^{(s)}\}_{i=1}^N$ denote the sequences of $\hat{\delta}_i$ and \hat{v}_i , sorted according to their absolute values in a descending order, and define the following estimators of σ_v^2/μ_v^2

$$\widehat{\left(\frac{\sigma_v^2}{\mu_v^2}\right)} = \frac{N^{2\hat{\alpha}-2} \sum_{i=1}^{N^{\hat{\alpha}}} \left(\hat{v}_i^{(s)} - \frac{1}{N^{\hat{\alpha}}} \sum_{i=1}^{N^{\hat{\alpha}}} \hat{v}_i^{(s)} \right)^2}{N^{\hat{\alpha}} - 1}, \text{ for any finite value of } \kappa, \quad (20)$$

and

$$\widehat{\left(\frac{\sigma_v^2}{\mu_v^2}\right)} = \frac{\sum_{i=1}^{N^{\hat{\alpha}}} \left(\hat{\delta}_i^{(s)} - \frac{1}{N^{\hat{\alpha}}} \sum_{i=1}^{N^{\hat{\alpha}}} \hat{\delta}_i^{(s)} \right)^2}{N^{\hat{\alpha}} - 1}, \text{ when } \kappa = 1. \quad (21)$$

Theorem 4 Suppose Assumptions 1 to 3 hold and $m = 1$. Let

$$\dot{\omega} = \frac{\min(N^{\hat{\alpha}}, T)}{T} \hat{V}_{f^2} + \frac{4 \min(N^{\hat{\alpha}}, T)}{N^{\hat{\alpha}}} \widehat{\left(\frac{\sigma_v^2}{\mu_v^2}\right)},$$

where $\hat{\alpha} = \hat{\alpha}, \tilde{\alpha}, \check{\alpha}$, and $\widehat{\left(\frac{\sigma_v^2}{\mu_v^2}\right)}$ is given by (20) or (21) depending on the value of κ ,

$$\hat{V}_{f^2} = \frac{1}{T} \sum_{t=1}^T \left(s_t - \frac{1}{T} \sum_{t=1}^T s_t \right)^2 + \sum_{j=1}^l \left[\frac{1}{T} \sum_{t=j+1}^T \left(s_{t-j} - \frac{1}{T} \sum_{t=1}^T s_t \right) \left(s_t - \frac{1}{T} \sum_{t=1}^T s_t \right) \right], \quad (22)$$

$s_t = (\tilde{x}_t - \tilde{x})^2$, $\tilde{x} = T^{-1} \sum_{t=1}^T \tilde{x}_t$, and $l \rightarrow \infty$. Then,

$$\dot{\omega} - \omega = o_p(1).$$

where ω is defined in (17), so long as $l \rightarrow \infty$, $l = o(T)$ and $l = o(N^{\alpha-1/2}T^{1/2})$.

3.2 Extensions

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

$$x_{it} = \beta_i' \mathbf{f}_t + u_{it} = \sum_{j=1}^m \beta_{ij} f_{jt} + u_{it}, \quad i = 1, 2, \dots, N,$$

where $\mathbf{f}_t = (f_{1t}, f_{2t}, \dots, f_{mt})'$ is an $m \times 1$ vector of factors and β_i is the associated vector of factor loadings (m is fixed). We make the following assumptions that generalise straightforwardly Assumptions 1 and 2.

Assumption 4 We assume that loadings are given by

$$\begin{aligned}\beta_{ij} &= v_{ij} \text{ for } i = 1, 2, \dots, [N^{\alpha_j}], \\ \beta_{ij} &= c_j \rho_j^{i - [N^{\alpha_j}] + 1}, \text{ for } i = [N^{\alpha_j}] + 1, [N^{\alpha_j}] + 2, \dots, N,\end{aligned}\tag{23}$$

where $1/2 < \alpha_j \leq 1$, $|\rho_j| < 1$, and $\{v_{ij}\}_{i=1}^{N^{\alpha_j}}$ is an IID sequence of random variables with mean $\mu_{v_j} \neq 0$, and variance $\sigma_{v_j}^2 < \infty$, for all $j = 1, 2, \dots, m$.

Assumption 5 The $m \times 1$ vector of factors, \mathbf{f}_t , follows a linear stationary process given by

$$\mathbf{f}_t = \sum_{j=0}^{\infty} \boldsymbol{\psi}_{f_j} \boldsymbol{\nu}_{f,t-j},\tag{24}$$

where $\boldsymbol{\nu}_{f_t}$ 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, $\boldsymbol{\psi}_{f_j}$, satisfy the absolute summability condition

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

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

Without loss of generality we assume that $\alpha = \alpha_1 \geq \alpha_j$, $j = 2, \dots, m$ and that factors are orthogonal. We have

$$\bar{\beta}_{jN} = N^{-1} \sum_{i=1}^N \beta_{ji} = \frac{N_j}{N} \left(\frac{\sum_{i=1}^{N_j} v_{ji}}{N_j} \right) + \frac{c_j \rho_j}{N} \left(\frac{1 - \rho_j^{(N - N_j)}}{1 - \rho_j} \right) = N^{\alpha_j - 1} \bar{v}_j + N^{-1} K_{\rho_j}\tag{25}$$

and

$$\text{Var}(\bar{\beta}_{jN}) = N^{-2} \boldsymbol{\tau}' \boldsymbol{\Sigma}_{\beta_j} \boldsymbol{\tau} = O(N^{\alpha_j - 2}).$$

Note that

$$\bar{x}_t - E(\bar{x}_t) = \bar{\beta}_{1N} f_{1t} + \bar{\beta}_{2N} f_{2t} + \dots + \bar{\beta}_{mN} f_{mt} + \bar{u}_t,$$

and

$$\begin{aligned}\text{Var}(\bar{x}_t) &= E(\bar{\beta}_{1N} f_{1t} + \bar{\beta}_{2N} f_{2t} + \dots + \bar{\beta}_{mN} f_{mt} + \bar{u}_t)^2 \\ &= \sum_{j=1}^m E(\bar{\beta}_{1N}^2) \sigma_{ff}^2 + E(\bar{u}_t^2) = \sum_{j=1}^m [E(\bar{\beta}_{1N})]^2 \sigma_{ff}^2 + \sum_{j=1}^m \text{Var}(\bar{\beta}_{1N}) \sigma_{ff}^2 + E(\bar{u}_t^2).\end{aligned}$$

Let $\bar{\boldsymbol{\beta}}_N = (\bar{\beta}_{1N}, \dots, \bar{\beta}_{mN})'$ and $\bar{\mathbf{v}}_N = (\bar{v}_{1N}, \dots, \bar{v}_{mN})'$. Then,

$$\bar{\boldsymbol{\beta}}_N = N^{\alpha - 1} \mathbf{D}_N \bar{\mathbf{v}}_N + N^{-1} \mathbf{K}_{\rho}.\tag{26}$$

where \mathbf{D}_N is an $m \times m$ diagonal matrix with diagonal elements given by $N^{\alpha_j - \alpha_1}$ and $\mathbf{K}_{\rho} = (K_{\rho_1}, \dots, K_{\rho_m})'$. It is important to stress that (23) can be generalised. In particular, the results below hold so long as $\sum_{i=N_j+1}^N \beta_{ji} < \infty$, in which case, $K_{\rho_j} = K_j = \sum_{i=N_j+1}^N \beta_{ji}$. Further, let

$$d_T = \bar{\mathbf{v}}_N' \mathbf{S}_{ff}^{-1/2} \bar{\mathbf{f}}_T - \boldsymbol{\mu}'_{\nu} \boldsymbol{\Sigma}_{ff}^{-1/2} \boldsymbol{\mu}_f\tag{27}$$

where $\mathbf{S}_{ff} = \mathbf{S}_{ff}[s_{ijf}] = \frac{1}{T} \sum_{t=1}^T (\mathbf{f}_t - \bar{\mathbf{f}}_T) (\mathbf{f}_t - \bar{\mathbf{f}}_T)'$, $\boldsymbol{\Sigma}_{ff} = \boldsymbol{\Sigma}_{ff}[\sigma_{ijf}] = E[(\mathbf{f}_t - \boldsymbol{\mu}_f) (\mathbf{f}_t - \boldsymbol{\mu}_f)']$, $\sigma_{if} = \sigma_{ii,f}$, $\boldsymbol{\mu}_f = E(\mathbf{f}_t)$, $\bar{\mathbf{f}}_T = \frac{1}{T} \sum_{t=1}^T \mathbf{f}_t$, and $\boldsymbol{\mu}_{\nu} = (\mu_{1\nu}, \dots, \mu_{m\nu})' = E(\mathbf{v}_i)$. Then we have the following theorem which distinguishes between a number of different cases depending on the relative values of the different exponents, α_j , for $j = 1, 2, \dots, m$.

Theorem 5 Suppose Assumptions 4 to 5 hold, $\alpha = \alpha_1 = \alpha_2 = \dots = \alpha_m$, and $\alpha > 1/2$. Then,

$$\sqrt{\min(N^{\alpha^*}, T)} \left(2 \ln(N) (\hat{\alpha} - \alpha^*) - \frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{D}_N \bar{v}_N} \right) \rightarrow_d N(0, \omega_m) \quad (28)$$

where

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

d_T is defined by (27),

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

and $\kappa^2 = \sum_{i=1}^m \mu_{iv}^2 \sigma_{if}^2$.

Continue to assuming that $\alpha = \alpha_1 = \alpha_2 = \dots = \alpha_m$, and suppose that either $\alpha > 4/7$ or $\frac{T^{1/2}}{N^{4\alpha-2}} \rightarrow 0$, then

$$\sqrt{\min(N^{\alpha^*}, T)} 2 \ln(N) (\tilde{\alpha} - \alpha^*) \rightarrow_d N(0, \omega_m), \quad (29)$$

and

$$\sqrt{\min(N^{\alpha^*}, T)} 2 \ln(N) (\tilde{\alpha} - \alpha^*) \rightarrow_d N(0, \omega_m). \quad (30)$$

Further, if either

$$\alpha_1 = \alpha > \alpha_2 + 1/4, \quad (31)$$

or if

$$\alpha_2 < 3\alpha/4, \quad T^b = N, \quad b > \frac{1}{4(\alpha - \alpha_2)}, \quad (32)$$

and $\alpha_2 \geq \alpha_3 \geq \dots \geq \alpha_m$, (28), (29) and (30) hold with ω replacing ω_m , where ω is defined in (17) and α^* is now defined by (16).

Finally, if $\alpha > \alpha_2 \geq \alpha_3 \dots \geq \alpha_m$ but neither (29) nor (30) hold, then (28), (29) and (30) hold with ω replacing ω_m , and

$$\alpha^* \equiv \alpha_N^* = \alpha + \frac{\ln\left(\sum_{i=1}^m N^{2(\alpha_1 - \alpha_i)} \mu_{iv}^2 \sigma_{if}^2\right)}{2 \ln(N)}.$$

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 for the loadings. As we noted in the Introduction this alternative specification is mathematically convenient, yet economically restrictive. More specifically consider the following formulation of the factor loadings:

Assumption 6 Suppose that the factor loadings are given by

$$\beta_{ik} = N^{\alpha-1} v_{ik}, \quad 1/2 < \alpha \leq 1 \quad (33)$$

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

It is easy to see that even in the case of this alternative specification of the loadings we have

$$\tau_N = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \sigma_{ij,x} = O(N^{2\alpha}),$$

and the following Corollary follows easily from the proofs of Theorems 1-3, and Lemma 6.

Corollary 2 *Let Assumptions 6 and 2-3 hold, $m = 1$ and $\alpha > 1/2$. Let $\hat{\alpha}$, $\tilde{\alpha}$ and $\check{\alpha}$ be defined as in (14), (18) and (19) respectively. Then,*

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

where α^* and ω are defined in (16) and (17), respectively. If $\mu_v^2 \sigma_f^2 = 1$, then,

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

As long as either $\alpha > 5/8$ or $\frac{T^{1/2}}{N^{4\alpha-2}} \rightarrow 0$,

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

and

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

Remark 1 *It is of interest to consider circumstances where Assumption 6 fails but the above result still holds. In particular, let*

$$\beta_i = N^{\alpha-1} v_i, \quad 1/2 < \alpha \leq 1 \tag{34}$$

where $v_i = v_{N_i} = \tilde{v}_i + c_{N_i}$ and $\{\tilde{v}_i\}_{i=1}^N$ is an i.i.d. sequence of random variables with mean $\mu_v \neq 0$, and variance $\sigma_v^2 < \infty$. 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 6 that is covered by Lemma 14. Without loss of generality, we order units, such that $c_{N_i} = N^{1-\alpha} \eta_i$ for $i = 1, 2, \dots, M$ 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 $M = o(N^{\alpha-1/2})$.

3.3 Is the normalization restriction $\kappa^2 = 1$ justified?

Since the cross-sectional exponent, α , is identified in terms of the factor loadings, and factor loadings are in turn identified only up to a non-singular rotation matrix, it is clear that some restriction on $\kappa^2 = \mu_v^2 \sigma_f^2$ is needed before a sufficiently accurate estimator of α can be obtained. Although, as we shall see below, it might be possible to jointly estimate α and κ under our preferred formulation of the factor loadings, this is not the case in general. For example, joint estimation of κ and α does not seem possible if the factor loadings are generated according to (33). It is, therefore, interesting to investigate the extent to which the imposition of the restriction $\kappa^2 = 1$ restrict the data generating process of x_{it} . A useful way to do this is to consider the extent to which setting $\kappa^2 = 1$ imposes restrictions on the second moment structure of x_{it} . The restriction $\kappa^2 = 1$ has no implications for the second moments of the individual series, $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iT})'$, and can only affect the cross-sectional average of the second moments. We have investigated these effects in a not-for-publication appendix that is available upon request, and conclude that they are reasonably modest.

4 Joint estimation of α and κ

In this section we show that it is in fact possible to jointly estimate α and κ , if the factor loadings follow the set up given under Assumption 1. To see this, without loss of generality, suppose that β_i are ordered such that $|\beta_1| \geq |\beta_2| \geq \dots |\beta_{[N^\alpha]}|$, with $\beta_i = 0$ for $i = [N^\alpha] + 1, \dots, N$. Let $\boldsymbol{\beta}_n = (\beta_1, \beta_2, \dots, \beta_n)'$, $\bar{\beta}_n = \frac{1}{n} \sum_{i=1}^n \beta_i$, and note that

$$\begin{aligned} E(\bar{\beta}_n) &= n^{-1} \boldsymbol{\tau}'_n E(\boldsymbol{\beta}_n) = \mu_v, \text{ if } n \leq [N^\alpha] \\ &= n^{-1} [N^\alpha] \mu_v, \text{ if } n > [N^\alpha]. \end{aligned}$$

Similarly, let $\boldsymbol{\Sigma}_{\beta,n} = Cov(\boldsymbol{\beta}_n)$ for $n \leq [N^\alpha]$ and bear in mind that

$$\boldsymbol{\Sigma}_{\beta,N} = \begin{pmatrix} \boldsymbol{\Sigma}_{\beta,[N^\alpha] \times [N^\alpha]} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{(N-[N^\alpha]) \times (N-[N^\alpha])} \end{pmatrix}$$

and note that

$$\begin{aligned} Var(\bar{\beta}_n) &= n^{-2} \boldsymbol{\tau}'_n Cov(\boldsymbol{\beta}_n) \boldsymbol{\tau}_n \leq n^{-1} \lambda_{\max}(\boldsymbol{\Sigma}_{\beta,n}) = O(n^{-1}), \text{ if } n \leq [N^\alpha] \\ &= n^{-2} \boldsymbol{\tau}'_n Cov(\boldsymbol{\beta}_n) \boldsymbol{\tau}_n \leq \frac{[N^\alpha]}{n^2} \lambda_{\max}(\boldsymbol{\Sigma}_{\beta,[N^\alpha] \times [N^\alpha]}) = O\left(\frac{[N^\alpha]}{n^2}\right), \text{ if } n > [N^\alpha]. \end{aligned}$$

Now let \bar{x}_{nt} be the cross-sectional average of x_{it} over $i = 1, 2, \dots, n < N$, where without loss of generality we assume that x_{it} is ordered by $|\beta_i|$. Then, using the above results

$$\begin{aligned} Var(\bar{x}_{nt}) &= \sigma_f^2 \mu_v^2 + n^{-1} c_n + O(n^{-1}), \text{ if } n \leq [N^\alpha], \\ &= n^{-2} [N^{2\alpha}] \sigma_f^2 \mu_v^2 + n^{-1} c_n + O\left(\frac{[N^\alpha]}{n^2}\right), \text{ if } n > [N^\alpha]. \end{aligned}$$

where $c_n = n^{-1} \boldsymbol{\tau}'_n \boldsymbol{\Sigma}_{u,n} \boldsymbol{\tau}_n$, $E(\mathbf{u}_n \mathbf{u}'_n)$, with $\mathbf{u}_n = (u_1, u_2, \dots, u_n)'$. This result can be written more compactly as

$$\begin{aligned} Var(\bar{x}_{nt}) - n^{-1} c_n &= I([N^\alpha] - n) \kappa^2 + I([N^\alpha] - n) O(n^{-1}) \\ &\quad + [1 - I([N^\alpha] - n)] \left[n^{-2} [N^{2\alpha}] \kappa^2 + O\left(\frac{[N^\alpha]}{n^2}\right) \right], \end{aligned}$$

where $I(A)$ is an indicator function that takes the value of 1 if $A > 0$ and zero otherwise. After some simplifications we have

$$\begin{aligned} Var(\bar{x}_{nt}) - n^{-1} c_n &= \{n^{-2} [N^{2\alpha}] + I([N^\alpha] - n) [1 - n^{-2} [N^{2\alpha}]]\} \kappa^2 + \\ &\quad I([N^\alpha] - n) \left[O(n^{-1}) - O\left(\frac{[N^\alpha]}{n^2}\right) \right] + O\left(\frac{[N^\alpha]}{n^2}\right) \end{aligned}$$

It is easily seen that for $n = N$, we obtain our earlier result (note that $N^\alpha \leq N$ since $\alpha \leq 1$), namely

$$Var(\bar{x}_{Nt}) - N^{-1} c_N = \kappa^2 [N^{2\alpha-2}] + O([N^{\alpha-2}]).$$

But other values of n that are sufficiently large can also be used to provide information on κ and α . First we need a consistent estimator of $Var(\bar{x}_{nt}) - n^{-1} c_n$, when n is sufficiently large. To this end, consider the OLS estimator of the slope coefficient in the regression of $x_{it} - \bar{x}_i$ on $\bar{x}_{Nt} - \bar{x}$, given by $\hat{v}_i = \sum_{t=1}^T (x_{it} - \bar{x}_i) (\bar{x}_{Nt} - \bar{x}) / \sum_{t=1}^T (\bar{x}_{Nt} - \bar{x})^2$, where $\bar{x} = N^{-1} \sum_{i=1}^N \bar{x}_i$. Order the observations, x_{it} , as $x_{(i)t}$, so that $x_{(1)t}$ is associated with $\hat{v}^{(1)}$, $x_{(2)t}$ is associated with $\hat{v}^{(2)}$, and so on, where $\hat{v}^{(1)}, \hat{v}^{(2)}, \dots$

are the values of $\hat{v}_1, \hat{v}_2, \dots, \hat{v}_N$ ordered in a descending manner. Consider the following estimator of $Var(\bar{x}_{nt})$,

$$\hat{\sigma}_{\bar{x}_n}^2 = \frac{\sum_{t=1}^T (\bar{x}_{nt} - \bar{x}_{nT})^2}{T},$$

where $\bar{x}_{nt} = n^{-1} \sum_{i=1}^n x_{(i)t}$, and $\bar{x}_{nT} = T^{-1} \sum_{t=1}^T \bar{x}_{nt}$. Similarly, estimate c_n by $\hat{c}_n = n^{-1} \sum_{j=1}^n \hat{\sigma}_{(j)}^2$, where $\hat{\sigma}_{(j)}^2$ is the estimator of $\sigma_{(j)}^2$, the standard error of the idiosyncratic component of $x_{(j)t}$ for $t = 1, 2, \dots, T$. Thus

$$\hat{c}_n = \frac{1}{nT} \sum_{j=1}^n \sum_{t=1}^T \hat{u}_{(j)t}^2,$$

where $\hat{u}_{(j)t} = x_{(j)t} - \bar{x}_{(j)} - \hat{v}_j^{(s)}(\bar{x}_{Nt} - \bar{x}_{NT})$.

Based on $\hat{\sigma}_{\bar{x}_n}^2$, we obtain the following estimating equation

$$\begin{aligned} \hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n &= \{n^{-2} [N^{2\alpha}] + I([N^\alpha] - n) [1 - n^{-2} [N^{2\alpha}]]\} \kappa^2 + \\ &I([N^\alpha] - n) \left[O(n^{-1}) - O\left(\frac{[N^\alpha]}{n^2}\right) \right] + O\left(\frac{[N^\alpha]}{n^2}\right) + \xi_n, \end{aligned}$$

where ξ_n is the estimation error. It is clear that only the estimates that are based on $n > [N^\alpha]$ are informative for α . This is because for values of $n < [N^\alpha]$ we have $\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n = \kappa^2 + O(n^{-1}) + \xi_n$.

The above results suggest that joint estimation of α and κ can be based on the minimization of the following quadratic form in terms of α and κ :

$$Q(\alpha, \kappa^2) = \sum_{n \leq [N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n - \kappa^2)^2 + \sum_{n > [N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n - n^{-2} [N^{2\alpha}] \kappa^2)^2, \quad (35)$$

The first order condition for κ^2 is

$$\sum_{n \leq [N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n - \hat{\kappa}^2(\alpha)) + [N^{2\alpha}] \sum_{n > [N^\alpha]} n^{-2} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n - n^{-2} [N^{2\alpha}] \hat{\kappa}^2(\alpha)) = 0,$$

which yields

$$\hat{\kappa}^2(\alpha) = \frac{\sum_{n \leq [N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n) + [N^{2\alpha}] \sum_{n > [N^\alpha]} n^{-2} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n)}{[N^\alpha] + [N^{4\alpha}] \sum_{n > [N^\alpha]} n^{-4}}.$$

Using this result we have

$$Q(\alpha) = Q(\alpha, \hat{\kappa}^2(\alpha)) = \sum_{n \leq [N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n - \hat{\kappa}^2(\alpha))^2 + \sum_{n > [N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n - n^{-2} [N^{2\alpha}] \hat{\kappa}^2(\alpha))^2.$$

The above expressions can be simplified. Let

$$\begin{aligned} Q_1 &= \sum_{n=1}^{[N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n)^2, \quad q_1 = \sum_{n=1}^{[N^\alpha]} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n) \\ Q_2 &= \sum_{n=[N^\alpha]+1}^N (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n)^2, \quad q_2 = \sum_{n=[N^\alpha]+1}^N n^{-2} (\hat{\sigma}_{\bar{x}_n}^2 - n^{-1} \hat{c}_n) \\ N(\alpha) &= [N^\alpha] + [N^{4\alpha}] \sum_{n=[N^\alpha]+1}^N n^{-4} \end{aligned}$$

which yields

$$\hat{\kappa}^2(\alpha) = \frac{q_1 + [N^{2\alpha}] q_2}{N(\alpha)}. \quad (36)$$

Then,

$$\begin{aligned} Q(\alpha) &= Q_1 + Q_2 + \hat{\kappa}^4(\alpha) [N^\alpha] + [N^{4\alpha}] \hat{\kappa}^4(\alpha) \sum_{n=[N^\alpha]+1}^N n^{-4} - 2\hat{\kappa}^2(\alpha)q_1 - 2\hat{\kappa}^2(\alpha) [N^{2\alpha}] q_2 \\ &= Q_1 + Q_2 + \hat{\kappa}^4(\alpha)N(\alpha) - 2\hat{\kappa}^2(\alpha) [q_1 + [N^{2\alpha}] q_2] \\ &= Q_1 + Q_2 + \left[\frac{q_1 + [N^{2\alpha}] q_2}{N(\alpha)} \right]^2 N(\alpha) - 2 [q_1 + [N^{2\alpha}] q_2] \left[\frac{q_1 + [N^{2\alpha}] q_2}{N(\alpha)} \right] \\ &= Q_1 + Q_2 - \frac{[q_1 + [N^{2\alpha}] q_2]^2}{N(\alpha)}. \end{aligned}$$

Further, note that $Q_1 + Q_2 = \sum_{n=1}^N (\hat{\sigma}_{\hat{x}_n}^2 - n^{-1}\hat{c}_n)^2$ which does not depend on α . Then, $\frac{[q_1 + [N^{2\alpha}] q_2]^2}{N(\alpha)}$ can be maximised straightforwardly by grid search, which gives an estimator of α (denoted by $\hat{\alpha}$). κ^2 can then be estimated using (36). The consistency of $\hat{\alpha}$ is established in Appendix IV, although we have not been able to establish the rate at which this estimator is consistent. But judging by the results obtained in the structural break literature we conjecture that the rate at which $\hat{\alpha}$ converges to α is $\ln(N)$, and to obtain a better convergence rate some *a priori* restriction on κ might be needed.

5 Monte Carlo Study

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

$$x_{it} = \beta_{1i}f_{1t} + \beta_{2i}f_{2t} + u_{it}, \quad (37)$$

for $i = 1, 2, \dots, N$, and $t = 1, 2, \dots, T$. The factors are generated as

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

with $f_{j,-50} = 0$, for $j = 1, 2$. The shocks are generated as

$$\begin{aligned} u_{it} &= \phi_i u_{i,t-1} + \sqrt{1 - \phi_i^2} \varepsilon_{it}, \text{ for } i = 1, 2, \dots, N \text{ and } t = -49, -48, \dots, 0, 1, \dots, T, \text{ with } u_{i,-50} = 0, \\ \zeta_{jt} &\sim IIDN(0, 1), \quad \varepsilon_{it} \sim IIDN(0, \sigma_i^2), \text{ where } \sigma_i^2 \sim IID \chi^2, i = 1, 2, \dots, N. \end{aligned}$$

Therefore, by construction $\sigma_{f_j}^2 = 1$, for $j = 1, 2$. In the first instance we set the loadings of the second factor equal to zero, $\beta_{2i} = 0$, and focus on the properties of β_{1i} . We consider the following generating mechanism:

$$\begin{aligned} \beta_{1i} &= v_{1i}, \text{ for } i = 1, 2, \dots, M(N) \\ \beta_{1i} &= \rho_l^{i-M}, \text{ for } i = M+1, M+2, \dots, N \end{aligned}$$

where $v_{1i} \sim IIDU(\mu_{v_1} - 0.5, \mu_{v_1} + 0.5)$, $M = [N^a]$ and $\rho_l = 0.5$. The above parametrization ensures that $\mu_{v_1}^2 \sigma_{f_1}^2 = 1$, as discussed in the development of the theory. Initially, where $m = 1$, we consider the following experiments.

Experiment A A basic design, where the factor, f_{1t} , is serially uncorrelated and the errors, u_{it} , are serially uncorrelated and cross-sectionally independent, namely when

$$\rho_1 = 0, \quad \phi_i = 0, \quad i = 1, 2, \dots, N$$

and $\varepsilon_{it} \sim IIDN(0, \sigma_i^2)$, for all i and t .

Experiment B This design is as in Experiment A, but allows for temporal dependence in the factor, so that

$$\rho_1 = 0.5, \quad \phi_i = 0, \quad i = 1, 2, \dots, N$$

and $\varepsilon_{it} \sim IIDN(0, \sigma_i^2)$, for all i and t .

Experiment C This design is as in Experiment A, where we continue to set $\rho_1 = 0$, but allow the idiosyncratic errors, u_{it} , to be cross-sectionally dependent according to a first order spatial autoregressive model. Let $\mathbf{u}_t = (u_{1t}, u_{2t}, \dots, u_{Nt})'$, and set \mathbf{u}_t as

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

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

Experiment D Next, we take into account the second factor as well and generate its loadings as

$$\beta_{2i} = v_{2i}, \quad \text{for } i = 1, 2, \dots, M_2(N)$$

$$\beta_{2i} = \rho_l^{i-M_2}, \quad \text{for } i = M_2 + 1, M_2 + 1, \dots, N,$$

where $v_{2i} \sim IIDU(\mu_{v_2} - 0.5, \mu_{v_2} + 0.5)$, $M_2 = N^{\alpha_2}$ and $\rho_l = 0.5$. We examine the case where $\alpha_2 = a$ and set $\sigma_{f_1} = \sigma_{f_2} = 1$ and $\mu_{v_1} = \mu_{v_2} = \sqrt{0.5}$. The rest of the parameters are se as in Experiment A, namely

$$\alpha_2 = a, \quad \rho_j = 0, \quad \phi_i = 0, \quad \text{for } i = 1, 2, \dots, N, \quad j = 1, 2$$

and $u_{it} \sim IIDN(0, 1)$, for all i and t .

For all experiments we consider values of $\alpha = 0.70, 0.75, \dots, 0.90, 0.95, 1.00$, $N = 100, 200, 500, 1000$ and $T = 100, 200, 500$. All experiments are based on 2000 replications. For each replication, the values of α , ρ_j , ρ_l , ϕ_i and \mathbf{S} are given as set out above. These parameters are fixed across all replications. The values of v_{ji} , $j = 1, 2$ are drawn randomly (N of them) for each replication.

In the case where the leading factor (f_{1t}) is serially uncorrelated, the statistic for making inference about α (when $\kappa = 1$) is given by (see theorems 1, 3 and 4)

$$\left(\frac{1}{T} \hat{V}_{f_1^2} + \frac{4}{N\hat{\alpha}} \frac{\widehat{\sigma_v^2}}{\mu_v^2} \right)^{-1/2} 2 \ln(N) (\hat{\alpha} - \alpha) \rightarrow_d N(0, 1),$$

for $\dot{\alpha} = \tilde{\alpha}$ or $\check{\alpha}$. Note that when the leading factor is serially uncorrelated then $\hat{V}_{f^2} = E(\widehat{f_{1t}^4})/\sigma_{f_1}^4 - 1$, where $E(\widehat{f_{1t}^4})/\sigma_{f_1}^4$ is consistently estimated by

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

where $\tilde{x}_t = (N^{-1} \sum_{i=1}^N x_{it}) / \hat{\sigma}_{\tilde{x}}$. Also $\widehat{\sigma_v^2/\mu_v^2}$, the estimator of σ_v^2/μ_v^2 , is given by

$$\frac{\widehat{\sigma_v^2}}{\mu_v^2} = \frac{\sum_{i=1}^{N\dot{\alpha}} \left(\hat{\delta}_i^{(s)} - \frac{1}{N\dot{\alpha}} \sum_{i=1}^{N\dot{\alpha}} \hat{\delta}_i^{(s)} \right)^2}{N\dot{\alpha} - 1}, \text{ with } \dot{\alpha} = \tilde{\alpha} \text{ or } \check{\alpha},$$

where $\{\hat{\delta}_i^{(s)}\}$ denotes the sequence of $\hat{\delta}_i$ sorted according to their absolute values in a descending order, and $\hat{\delta}_i$ is the OLS estimator of the regression coefficient of x_{it} on $\tilde{x}_t = (\bar{x}_t - \bar{x})/\hat{\sigma}_{\tilde{x}}$. Also see Theorem 4 and the discussions that proceed it. The above expressions apply irrespective of whether the model contains one or two factors.

Size of the tests is computed under $H_0 : \alpha = \alpha_0$, using a two-sided test where α_0 takes values in the range $[0.70, 1.00]$, as indicated above. 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-). All results are scaled by 100.

Finally, when $\rho_j \neq 0$, the case of serially correlated factors, we use a corrected variance estimator of f_t . The relevant formula for the test statistic is given by

$$\left[\frac{1}{T} \left[\hat{V}_{f^2}(q) \right] + \frac{4}{N\dot{\alpha}} \frac{\widehat{\sigma_v^2}}{\mu_v^2} \right]^{-1/2} 2 \ln(N) (\dot{\alpha} - \alpha) \rightarrow_d N(0, 1), \quad (38)$$

for $\dot{\alpha} = \tilde{\alpha}$ or $\check{\alpha}$. $\hat{V}_{f^2}(q)$ is computed by first estimating an AR(q) process for $\tilde{z}_t = z_t - \bar{z}$, where $z_t = (\tilde{x}_t - \tilde{x})^2$, $\tilde{x}_t = (1/N \sum_{i=1}^N x_{it}) / \hat{\sigma}_{\tilde{x}}$, $\tilde{x} = T^{-1} \sum_{t=1}^T \tilde{x}_t$ and $\bar{z} = T^{-1} \sum_{t=1}^T z_t$, and then $\hat{V}_{f^2}(q) = \hat{\sigma}_z^2 / (1 - \hat{\gamma}_1 - \hat{\gamma}_2 - \dots - \hat{\gamma}_q)^2$, where $\hat{\sigma}_z$ is the regression standard error and $\hat{\gamma}_i$ is the i^{th} estimated AR coefficient fitted to \tilde{z}_t . The lag order is set to $q = T^{1/3}$, and $\widehat{\sigma_v^2/\mu_v^2}$ is computed as before. Note that this correction is not the standard Newey-West one but based on AR approximations. We found that this correction has better finite samples properties and hence we use this in both the Monte Carlo study and the empirical applications of Section 6.

5.1 Results

The results for experiment A are summarized in Table A1 giving the bias, Root Mean Square Error (RMSE), size and power when $\tilde{\alpha}$ is used as the estimator of α . Table A2 presents the same set of summary statistics for experiment A when the further bias-corrected estimator, $\check{\alpha}$, is considered. 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 $T \rightarrow \infty$. (See Theorem 2).

Both estimators perform reasonably well with $\check{\alpha}$ having a slightly smaller bias for values of α in the range $[0.70 - 0.85]$, but overall there is little to choose between the two estimators, and therefore in what follows we focus on $\tilde{\alpha}$ as the estimator of α . As predicted by the theory, the bias and RMSE of $\tilde{\alpha}$ decline with both N and T , and tend to be smaller for larger values of α . A similar pattern can also be seen in 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 below the null (denoted by Power-) as compared to the alternatives above the null (denoted by Power+). This asymmetry is particularly marked for low values of α and disappears as α is increased.

The results for Experiment B where the factor is allowed to be serially correlated are summarized in Table B. As compared to the baseline case, we see some deterioration in the results, particularly for relatively small values of N and T . The RMSEs are slightly higher, the size distortions slightly larger, and the power slightly lower. But these differences tend to vanish as N and T are increased.

The effects of allowing for weak cross-sectional dependence in the idiosyncratic errors, u_{it} , on estimation of α are summarized in Table C for Experiment C. 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 A2, for the baseline experiments. 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.

Finally, the results of Experiment D where one additional factor is included in the baseline case are summarized in Table D. As can be seen, the results are hardly affected by the addition of the new factor to the data generating process. Consistent with the one-factor case of Experiment A, both the bias and RMSE of $\tilde{\alpha}$ fall gradually as N and α are increased, while tests of the null hypothesis based on $\tilde{\alpha}$, are correctly sized for $\alpha > 0.7$ in this case as well. Similar observations can also be made with respect to the power.

The above Monte Carlo experiments, although limited in scope, 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. Although, the results tend to deteriorate somewhat when we consider serially correlated factors or weak cross-sectional dependence in the idiosyncratic errors, 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 quality of the small sample results tend to improve when values of α close to unity are considered.

6 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. Specifically, we consider three types of data sets: quarterly cross-country data used in global modelling, large quarterly data sets used in empirical factor model literature, and monthly stock returns on the constituents of Standard and Poor 500 index.

6.1 Cross-country dependence of macro-variables

Table 1 presents estimates of α for real output growth, inflation, rate of change of real equity prices, and interest rates (short and long) computed over 33 countries (when available). The data is from Cesa-Bianchi, Pesaran, Rebucci, and Xu (2012), and covers the period 1979Q2-2009Q4, which updates

the earlier GVAR (global vector autoregressive) data sets used in Pesaran, Schuermann, and Weiner (2004), and Dees, di Mauro, Pesaran, and Smith (2007).⁸ We provide both bias corrected estimates, $\tilde{\alpha}$ and $\check{\alpha}$, computed using available cross-country time series, y_{it} , over the full sample period. The observations were standardized as $x_{it} = (\tilde{y}_{it} - \bar{y}_i)/s_i$, where \bar{y}_i is the sample mean of each time series, and s_i is the corresponding standard deviations. Table 1 also reports the 95% confidence bands, the cross section dimension (N) and the time series dimension (T) for each of the variables. Although, there are 33 countries in the GVAR data set, not all variables are available for all the 33 countries. For example, the short term (3 months) interest rate data is not available for Saudi Arabia, and real equity prices and long term interest rate data (10 year government bond) are available only for some of the countries.

We first note that the two different estimates of α provided in Table 1 are very close, which are in line with the Monte Carlo results reported in Tables A1 and A2. Focusing on $\tilde{\alpha}$, we observe that the point estimates range between 0.754 for cross dependence of GDP growth rates, to 0.968 for long term interest rates. The exponent of cross-sectional dependence for short term interest rates and real equity prices at 0.907 and 0.881 are also quite high, indicating that financial variables are more strongly correlated as compared to the real variables. The reported confidence bands all lie above 0.5, but none cover unity either apart from the case of long term interest rates (marginally), suggesting that whilst a factor structure might be a good approximation for modelling global dependencies, 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	$\tilde{\alpha}_{0.025}^*$	$\tilde{\alpha}$	$\tilde{\alpha}_{0.975}^*$	$\check{\alpha}_{0.025}^*$	$\check{\alpha}$	$\check{\alpha}_{0.975}^*$
Real GDP growth, q/q	33	122	0.691	0.754	0.818	0.689	0.752	0.816
Inflation, q/q	33	123	0.778	0.851	0.924	0.777	0.850	0.924
Real equity price change, q/q	26	122	0.797	0.881	0.966	0.796	0.881	0.966
Short-term interest rates	32	123	0.831	0.907	0.983	0.831	0.907	0.983
Long-term interest rates	18	123	0.864	0.968	1.072	0.864	0.968	1.072

*95% level confidence bands

6.2 Within-country dependence of macro-variables

An important strand in the empirical factor literature, promoted through the work of Forni, Hallin, Lippi, and Reichlin (2000), Forni and Lippi (2001) and Stock and Watson (2002), uses factor models to 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, Kapetanios, and Price (2010). For the US the data set comprises 95 variables over the period 1960Q2 to 2008Q3. For the UK the data set covers 94 variables spanning from 1977Q1 to 2008Q2. The estimates of α , computed from the standardized data sets as explained in the previous subsection, together with their 95% confidence bands are summarized in Table 2.

For both data sets the estimates, $\tilde{\alpha}$ and $\check{\alpha}$, are identical up to two decimal places. For the US data set the point estimate ($\tilde{\alpha}$) of α , at 0.739, is slightly larger than the estimate obtained for the UK at

⁸This version of GVAR data set can be downloaded from <http://www-cfap.jbs.cam.ac.uk/research/gvartoolbox/download.html>

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

US			UK		
1960Q2-2008Q3			1977Q1-2008Q2		
$N=95, T=194$			$N=94, T=126$		
$\tilde{\alpha}_{0.025}^*$	$\tilde{\alpha}$	$\tilde{\alpha}_{0.975}^*$	$\tilde{\alpha}_{0.025}^*$	$\tilde{\alpha}$	$\tilde{\alpha}_{0.975}^*$
0.689	0.739	0.788	0.636	0.715	0.793
$\check{\alpha}_{0.025}^*$	$\check{\alpha}$	$\check{\alpha}_{0.975}^*$	$\check{\alpha}_{0.025}^*$	$\check{\alpha}$	$\check{\alpha}_{0.975}^*$
0.689	0.738	0.788	0.635	0.713	0.792

*95% level confidence bands

0.715. The 95% confidence bands for both data sets are well above the threshold value of 0.50, but are well short of 1.0 at the upper end of the band. Once again there is some evidence of a common factor dependence, but the evidence is not as strong as it is assumed in the literature.

6.3 Cross-sectional exponent of stock returns

One of the important considerations in the analysis of financial markets is the extent to which asset returns are interconnected. This is encapsulated in the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), and the arbitrage pricing theory (APT) of Ross (1976). Both theories have factor representations with at least one strong common factor and an idiosyncratic component that could be weakly correlated (see, for example, Chamberlain (1983)). The strength of the factors in these asset pricing models is measured by the exponent of the cross-sectional dependence, α . When $\alpha = 1$, as it is typically assumed in the literature, all individual stock returns are significantly affected by the factor(s), but there is no reason to believe that this will be the case for all assets and at all times. The disconnect between some asset returns and the market factor(s) could occur particularly at times of stock market booms and busts where some asset returns could be driven by non-fundamentals. Therefore, it would be of interest to investigate possible time variations in the exponent α for stock returns. Note that under our methodology the market factor associated with the CAPM specification is implied by the data rather than imposed by use of a specific market portfolio composition which can be limiting, as explained in Roll (1977).

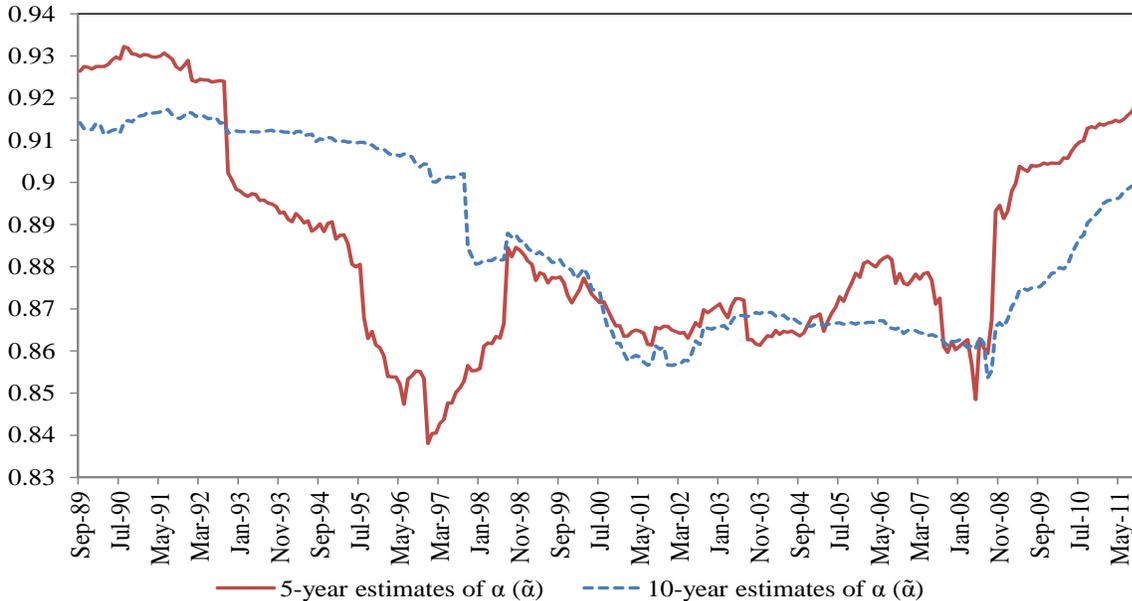
We base our empirical analysis on monthly excess returns of the securities included in the Standard & Poor 500 (S&P 500) index of large cap U.S. equities market, and estimate α recursively using rolling samples of size 120 months (10 years) and 60 months (5 years). Due to the way the composition of S&P 500 changes over time, we compiled returns on all 500 securities at the end of each month over the period from September 1989 to September 2011, and included in the rolling samples only those securities that had a sufficiently long history in the month under consideration. On average we ended up with 439 securities at the end of each month for the rolling samples of size 10 years, and 476 securities when we used a rolling sample of size 5 years. The one-month US treasury bill rate was chosen as the risk free rate (r_{ft}), and excess returns computed as $\tilde{r}_{it} = r_{it} - r_{ft}$, where r_{it} is the monthly return on the i^{th} security in the sample inclusive of dividend payments (if any).⁹ Recursive estimates of α were then computed using the standardized observations $x_{it} = (\tilde{r}_{it} - \tilde{r}_i)/s_i$, where \tilde{r}_i is the sample mean of the excess returns over the selected rolling sample, and s_i is the corresponding standard deviations.

The recursive estimates of α based on 10 years and 5 years rolling windows are given in Figure 1. We also computed rolling standard errors for the estimates, $\tilde{\alpha}_t$, using the serial correlation correction discussed in Section 5. Based on these standard errors, the 95% confidence bands of the recursive

⁹For further details of data sources and definitions see Pesaran and Yamagata (2012).

estimates were on average ± 0.02 around the point estimates for both rolling sample sizes considered. These bands are not shown in Figure 1, since the bands are relatively narrow and we aim to highlight the time variations in the estimates of α .

Figure 1: $\tilde{\alpha}_t$ associated with S&P 500 securities' excess returns - 5-yr and 10-yr rolling samples



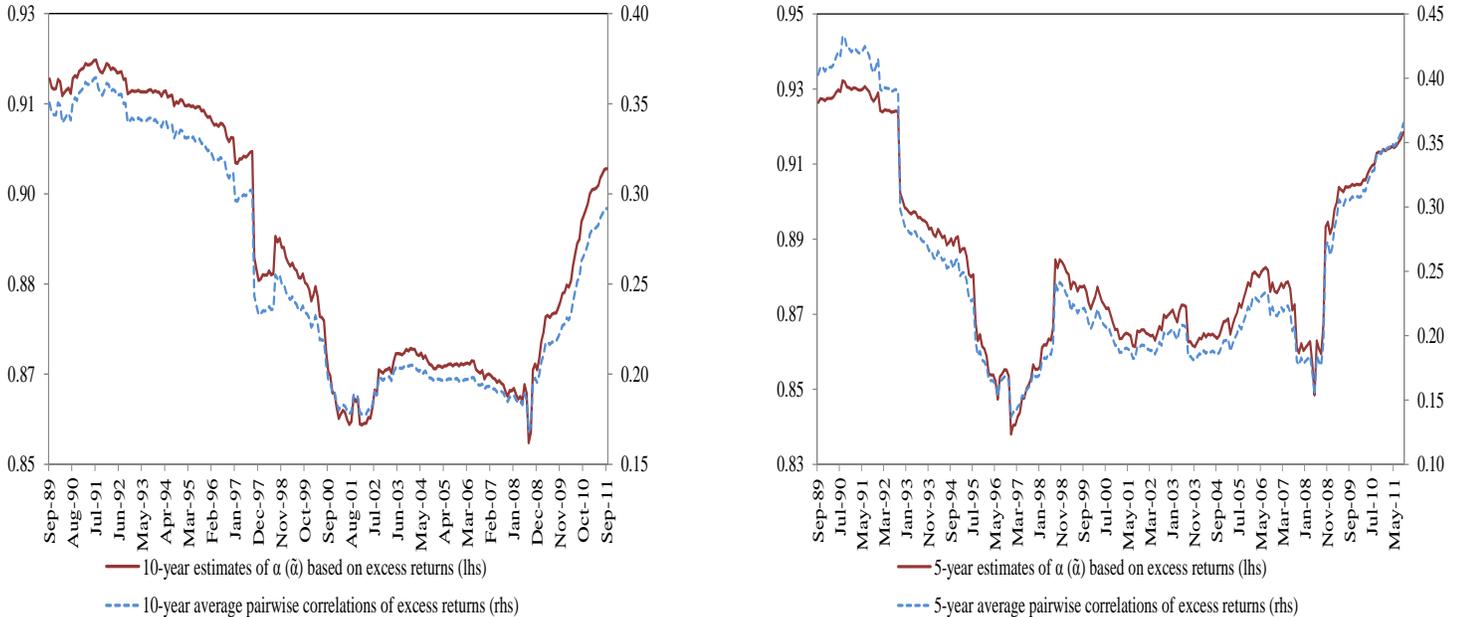
The figure covers 23 years of monthly recursive estimates of α , and yet fall in a relatively narrow range of 0.854 – 0.917 in the case of the 10 years rolling samples, and in a slightly wider range of 0.838 – 0.932 in the case of 5 years rolling samples. These estimates clearly show a high degree of inter-linkages across individual securities, although the null hypothesis that $\alpha = 1$ is clearly rejected. More importantly, there are clear trends in the estimates of α . The estimates based on the 10 years rolling samples fall from a high of 0.92 in 1990 to a low of 0.85 just before the burst of the dot-com bubble in 1999-2000. The estimates of α then stabilize around 0.86 over the period 2000 – 2008, then fall quite dramatically towards the end of 2008 at the time of the market crash, before starting to rise again to its present level of 0.90 in September 2011. The factors behind these fluctuations are complex and reflect the relative importance of micro and macro fundamentals prevailing in financial markets. A standard factor model does not seem able to fully account for the changing nature of the dependencies in securities market over the 1989-2011 period. A similar result is also obtained when α is estimated using 5-year rolling samples, although as to be expected the estimates are variable. Indeed, the upward movements in the estimates based on 5-year windows are more pronounced both in the latest crisis as well as in the period of 1997-2000 that saw smaller crises caused by the Asian economic turmoil, LTCM and the bursting of the dot-com bubble.

The patterns observed in the above estimates of α are in line with changes in the degree of correlations in equity markets. It is generally believed that correlations of returns in equity markets rise at times of financial crises, and it would be of interest to see how our estimates of α relate to return correlations. To this end in Figure 2 we compare the estimates of α to average pair-wise correlation coefficients of excess returns ($\bar{\rho}_N$) on securities included in S&P 500 index, using 10-year and 5-year rolling windows.¹⁰ As the plots in these figures show, our estimates of α closely follow the rolling

¹⁰Denote the correlation of excess returns on i and j securities by $\hat{\rho}_{ij}$, the pair-wise average correlation of the market is then computed as $\bar{\rho}_N = (1/N(N-1)) \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}$, where N is the number of securities under consideration.

estimates of $\bar{\rho}_N$.

Figure 2: Average pair-wise correlations of excess returns for securities in the S&P 500 index and the associated $\tilde{\alpha}_t$ estimate computed using 10-year and 5-year rolling samples



Further, it would be of interest to see how our estimates of α compare with estimates obtained using excess returns on market portfolio as a measure of the unobserved factor. This approach starts with capital asset pricing model (CAPM) and assumes that the single factor in CAPM regressions can be approximated by a stock market index. Under these assumptions, as noted in the Introduction, a direct estimate of α is given by $\hat{\alpha}_d = \ln(\hat{M})/\ln(N)$, where \hat{M} denotes the estimated number of non-zero betas, and N is the total number of securities under consideration.¹¹ \hat{M} can be consistently estimated (as N and $T \rightarrow \infty$) by the number of t -tests of $\beta_i = 0$ in the CAPM regressions

$$r_{it} - r_{ft} = \alpha_i + \beta_i (r_{mt} - r_{ft}) + u_{it}, \text{ for } i = 1, 2, \dots, N, \quad (39)$$

that end up in rejection of the null hypothesis at a chosen significance level, where r_{mt} is a broadly defined stock market index. In our application we choose the value-weighted return on all NYSE, AMEX, and NASDAQ stocks to measure r_{mt} ,¹² and select 1% as the significance level of the tests. Such estimates of α obtained recursively using 10-year and 5-year rolling windows are shown in the two plots in Figure 3. For ease of comparison, these plots also include our (indirect) estimates of α based on the same data sets (except for the marker return, r_{mt} , which is not used). The two sets of estimates co-move over most of the period and tend to become closer in the aftermath of the dot-com bubble and during the recent financial crisis. The correlation coefficient of the two sets of estimates is 0.923 for the ones based on 10-year rolling samples, and 0.739 for the ones based on the 5-year rolling samples. The two sets of estimates, however, differ in scale, with the direct estimates being closer to unity. The scale of the direct estimates clearly depends on the measure of market return, the level of significance chosen, and the assumption that the model contains only one single factor with $\alpha > 1/2$,

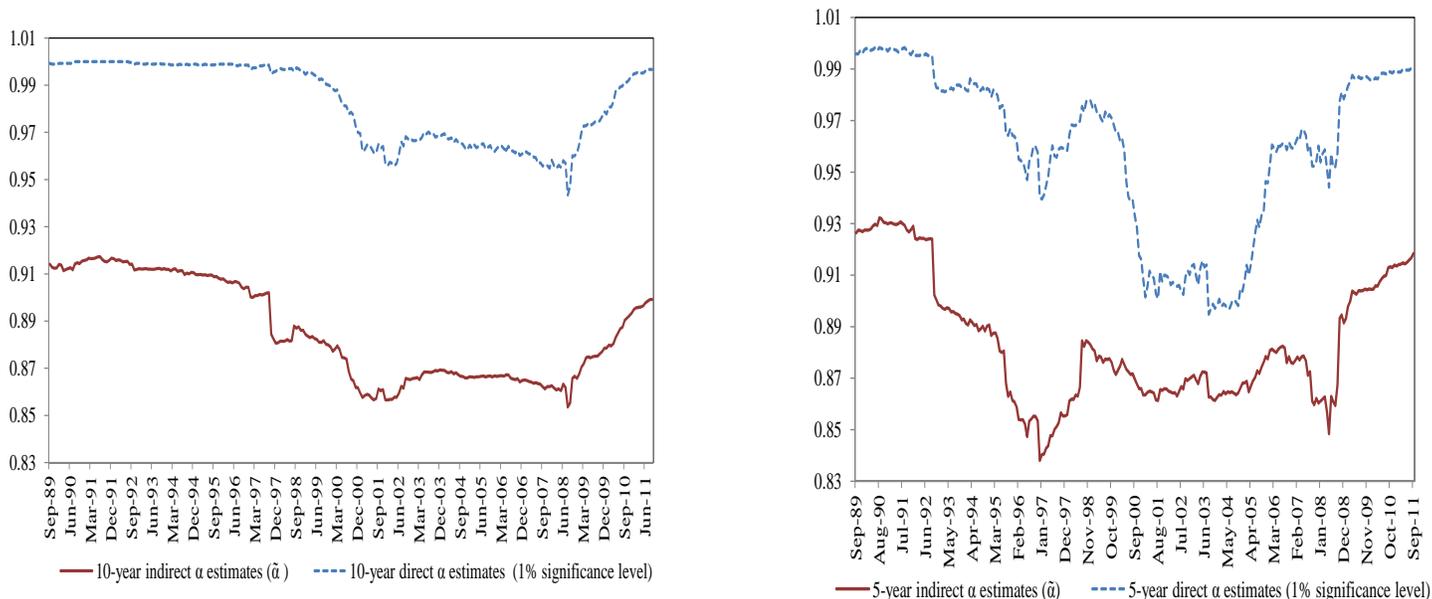
Almost identical estimates are also obtained if we use returns instead of excess returns.

¹¹Recall that $M = [N^\alpha]$, where M is the true number of non-zero betas.

¹²The return data on market index was obtained from Ken French's data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

and in consequence is subject to a high degree of uncertainty.¹³ Nevertheless, it is reassuring that the direct and indirect estimates of α in this application tend to move together closely.

Figure 3: Direct ($\hat{\alpha}_d$) and indirect ($\tilde{\alpha}$) estimates of cross-sectional exponent of the market factor (using excess returns on S&P 500 securities) based on 10-year and 5-year rolling samples



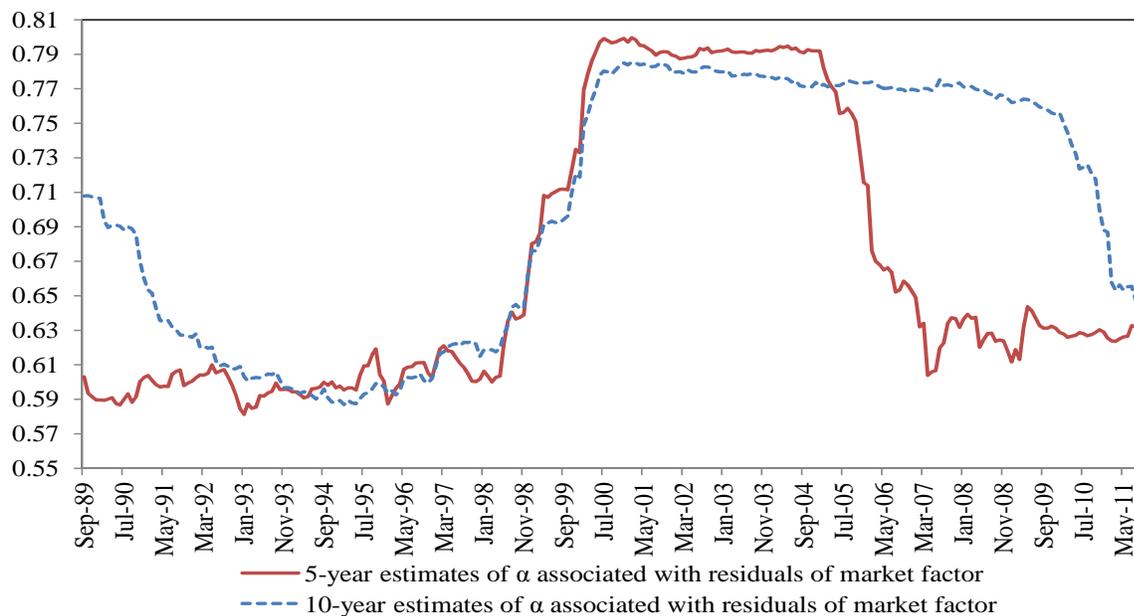
There is also a further consideration when comparing the estimates of α and α_d . Under CAPM the errors, u_{it} , in (39) are assumed to be cross-sectionally weakly correlated, namely that the cross-sectional exponent of the errors, say α_u , must be $\leq 1/2$. But this need not be the case in reality. Although we do not observe u_{it} , under CAPM the OLS residuals from regressions of $r_{it} - r_{ft}$ on $r_{mt} - r_{ft}$, denoted by \hat{u}_{it} , provide an accurate estimate of u_{it} up to $O_p(T^{-1/2})$, and can be used to compute consistent estimates of α_u .¹⁴ The bias-adjusted estimates of α_u , denoted by $\tilde{\alpha}_u$, and computed using standardized residuals over 5-year and 10-year rolling samples, are displayed in Figure 4. Interestingly enough, these estimates, although much smaller than those estimated using excess returns, nevertheless tend to be larger than the threshold value of $1/2$, suggesting the presence factors other than the market factor influencing individual security returns. The influence of residual factor(s) is rather weak initially (around 0.60), but starts to rise in the years leading to the dot-com bubble and reaches the pick of 0.80 in the middle of 2000 and stays at around that level for the period up to 2006 – 2008 (depending on whether the 10-year or 5-year rolling samples are used), then begins to fall significantly after the start of the recent financial crisis, and currently stands at around 0.63. Although, special care must be exercised when interpreting these estimates (both because α_u is estimated using residuals and the fact that $\tilde{\alpha}$ tends to be biased upward particularly when $\alpha < 0.75$), nevertheless their patterns over time are indicative of some departures from CAPM during the period 1999 – 2006. Also, it is interesting that the rolling estimates of α_u tend to move in opposite directions to the estimates of α computed over the same rolling samples. Weakening of the market factor tends to coincide with strengthening of the residual factor(s), thus suggesting that correlations across returns could remain high even during

¹³The distribution theory of the direct estimator of α is complicated by the cross dependence of the errors in the underlying CAPM regressions and its consideration is outside the scope of the present paper.

¹⁴A formal proof and analysis when α is estimated from regression residuals is beyond the scope of the present paper.

periods where the cross-sectional exponent of the dominant factor is relatively low, once the presence of multiple factors with exponents exceeding 0.5 is acknowledged.

Figure 4: Estimates of cross-sectional exponent of residuals ($\tilde{\alpha}_u$) from CAPM regressions using 5-year and 10-year rolling samples



7 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 by investing in disparate assets or asset classes and the performance of standard estimators such as principal components when applied to data sets with unknown collinearity structures. A common characteristic of such analyses is the need to quantify cross-sectional dependence especially when it is prevalent enough to materially affect the outcome of the analysis.

In this paper we propose a relatively simple method of measuring the extent of inter-connections 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 range of cross-sectional dependence manifestations while retaining its simple and tractable form. We propose consistent estimators of the cross-sectional exponent and derive their asymptotic distribution under plausible conditions. 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 three widely analysed classes of data sets. In all cases, we find that the results of the empirical analysis accord with prior intuition. For example, in the case of cross country applications we obtain larger estimates for the cross-sectional exponent of equity returns as compared to those estimated for cross country output growths and inflation. For individual securities

in S&P 500 index, the estimates of cross-sectional exponents are systematically high but not equal to unity, a widely maintained assumption in the theoretical multi-factor literature.

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-section 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, Pesaran, and Tosetti (2011), Kapetanios and Marcellino (2010) and Onatski (2011) 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. Methods used to determine the number of factors in large data sets, discussed in, e.g., Bai and Ng (2002), Onatski (2009) and Kapetanios (2010), are invalid and will select the wrong number of factors, even asymptotically.¹⁵ 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.
- (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.
- 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," *Forthcoming in Economia, Journal of the Latin American and Caribbean Economic Association*.
- CHAMBERLAIN, G. (1983): "Funds, Factors and Diversification in Arbitrage Pricing Theory," *Econometrica*, 51(5), 1305–1323.
- 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.
- 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.
- EKLUND, J., G. KAPETANIOS, AND S. PRICE (2010): "Forecasting in the Presence of Recent Structural Change," *Working Paper 406, Bank of England*.
- 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.
- GABAIX, X. (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79(3), 733–772.
- 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.

¹⁵It 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.

- (2005b): “The Empirical Eigenvalue Distribution of a Gram Matrix with a Given Variance Profile,” *Annales de l’Institut Henri Poincaré*, 42(6), 649–670.
- 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.
- LINTNER, J. (1965): “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *Review of Economics and Statistics*, 47(1), 13–37.
- ONATSKI, A. (2009): “Testing Hypotheses about the Number of Factors in Large Factor Models,” *Econometrica*, 77(5), 1447–1479.
- (2010): “Determining the Number of Factors Form Empirical Distribution of Eigenvalues,” *Review of Economics and Statistics*, 92(4), 1004–1016.
- (2011): “Asymptotics of the Principal Components Estimator of Large Factor Models with Weakly Influential Factors,” *Working paper, University of Cambridge*.
- PESARAN, M. H. (2006): “Estimation and Inference in Large Heterogeneous panels with a Multifactor Error Structure,” *Econometrica*, 74(4), 967–1012.
- PESARAN, M. H., T. SCHUERMAN, AND S. WEINER (2004): “Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model,” *Journal of Business and Economic Statistics*, 22(2), 129–162.
- PESARAN, M. H., AND T. YAMAGATA (2012): “Testing CAPM with a Large Number of Assets,” *Forthcoming Working Paper, University of Cambridge*.
- ROLL, R. (1977): “A Critique of the Asset Pricing Theory’s Tests Part I: On Past and Potential Testability of the Theory,” *Journal of Financial Economics*, 4(2), 129–176.
- ROSS, S. A. (1976): “The Arbitrage Theory of Capital Asset Pricing,” *Journal of Economic Theory*, 13, 341–360.
- SHARPE, W. (1964): “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, 19(3), 425–442.
- 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 I: Statement of Lemmas

Lemma 1 *Under Assumptions 2 and 3, $f_t, \{u_{it}\}_{t=1}^{\infty}$, 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,*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T (f_t - \bar{f}) (\sqrt{N} \bar{u}_t) \rightarrow_d N(0, \bar{\sigma}_{\infty}^2),$$

where

$$\bar{\sigma}_{\infty}^2 = \lim_{N \rightarrow \infty} \left(\frac{\sum_{i=1}^N (\sigma_i^2 + \sum_{j=1, j \neq i}^{\infty} \sigma_{ij})}{N} \right), \quad (40)$$

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

Lemma 3 Under Assumption 3,

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

where

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

Lemma 4 Under Assumptions 1-3, $\widehat{\sigma}_N^2 - \bar{\sigma}_\infty^2 = O_p(T^{-1}) + O_p((NT)^{-1/2})$.

Lemma 5 Under Assumptions 1-2, $\sqrt{\min(N^\alpha, T)} (\ln(s_f^2 \bar{v}_N^2) - \ln(\sigma_f^2 \mu_v^2)) \rightarrow_d N(0, \omega)$.

Lemma 6 Under Assumption 6 and Assumptions 2-3, $\sqrt{\min(N, T)} (\ln(s_f^2 \bar{v}_N^2) - \ln(\sigma_f^2 \mu_v^2)) \rightarrow_d N(0, \omega)$.

Lemma 7 Under Assumptions 1-3, and as long as $\alpha > 4/7$,

$$\sqrt{\min(N^\alpha, T)} \left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} - \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2} \right) = O_p \left(\sqrt{\min(N^\alpha, T)} N^{2-4\alpha} \right).$$

Lemma 8 Under Assumptions 1-2, and as long as $\alpha > 1/2$,

$$\sqrt{\min(N^\alpha, T)} \ln(N) \left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} - \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2} \left(1 + \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2} \right) \right) = o_p(1).$$

Lemma 9 Let $\{\hat{v}_i^{(s)}\}_{i=1}^N$ and $\{\hat{\delta}_i^{(s)}\}_{i=1}^N$ be the reorderings of $\{\hat{v}_i\}_{i=1}^N$ and $\{\hat{\delta}_i\}_{i=1}^N$ where $|\hat{v}_i^{(s)}| \geq |\hat{v}_{i+1}^{(s)}|, \forall i$ and $|\hat{\delta}_i^{(s)}| \geq |\hat{\delta}_{i+1}^{(s)}|, \forall i$. Under Assumptions 1-3, and assuming that $\lim_{T, N \rightarrow \infty} T^{-1} N^\alpha < \infty$, we have

$$\frac{N^{2\alpha-2} \sum_{i=1}^{N^\alpha} \left(\hat{v}_i^{(s)} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} \hat{v}_i^{(s)} \right)^2}{N^\alpha - 1} - \frac{\sigma_v^2}{\mu_v^2} = o_p(1). \quad (41)$$

If $\sigma_f^2 \mu_v^2 = 1$, then

$$\frac{\sum_{i=1}^{N^\alpha} \left(\hat{\delta}_i^{(s)} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} \hat{\delta}_i^{(s)} \right)^2}{N^\alpha - 1} - \frac{\sigma_v^2}{\mu_v^2} = o_p(1).$$

Lemma 10 Under Assumptions 1-2, we have $\hat{V}_{\bar{f}^2} - V_{\bar{f}^2} = o_p(1)$, as long as $l \rightarrow \infty, l = o(T)$ and $l = o(N^{\alpha-1/2} T^{1/2})$.

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

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

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

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

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

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

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

Lemma 13 Under either Assumption 1 or 6 and Assumptions 2-3, and $\alpha > \alpha_2 \geq \alpha_3 \geq \dots \geq \alpha_m$,

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

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

$$b > \frac{1}{4(\alpha - \alpha_2)}. \quad (42)$$

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

Appendix II: Proofs of Theorems

Proof of Theorem 1

We start by noting that

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

where $\bar{x}_t = \bar{\beta}_N f_t + \bar{u}_t$, and $\bar{x} = T^{-1} \sum_{t=1}^T \bar{x}_t = \bar{\beta}_N \bar{f} + \bar{u}$. Further, let

$$K_\rho = K_{N\rho} = \sum_{j=M+1}^N \beta_j = \rho \left(\frac{1 - \rho^{(N-M)}}{1 - \rho} \right); \quad M = \lfloor N^\alpha \rfloor. \quad (43)$$

Then we have

$$\begin{aligned} \hat{\sigma}_{\bar{x}}^2 &= \bar{\beta}_N^2 s_f^2 + 2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2 \right], \\ s_f^2 &= \frac{1}{T} \sum_{t=1}^T (f_t - \bar{f})^2 \rightarrow_p \sigma_f^2 > 0, \text{ as } T \rightarrow \infty. \end{aligned}$$

Therefore

$$\ln(\hat{\sigma}_{\bar{x}}^2) = \ln(\bar{\beta}_N^2 s_f^2) + \ln \left(1 + \frac{2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2 \right]}{\bar{\beta}_N^2 s_f^2} \right).$$

But under Assumption 1, $\bar{\beta}_N = (M/N)\bar{v}_N + N^{-1}K_\rho$, where $\bar{v}_N = M^{-1} \sum_{i=1}^M v_i$, we have

$$\begin{aligned} \ln(\bar{\beta}_N^2 s_f^2) &= \ln \left(N^{\alpha-1} \bar{v}_N s_f + \frac{K_\rho s_f}{N} \right)^2 = 2(\alpha-1) \ln(N) + \ln(s_f^2 \bar{v}_N^2) + 2 \ln \left(1 + \frac{K_\rho}{N^\alpha \bar{v}_N} \right) \\ &= 2(\alpha-1) \ln(N) + \ln(s_f^2 \bar{v}_N^2) + O_p(N^{-\alpha}). \end{aligned}$$

Hence, recalling from (14) that $\hat{\alpha} = 1 + \ln(\hat{\sigma}_{\bar{x}}^2)/2 \ln(N)$, we have

$$2 \ln(N) (\hat{\alpha} - \alpha) - \ln(s_f^2 \bar{v}_N^2) = \ln \left(1 + \frac{2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2 \right]}{\bar{\beta}_N^2 s_f^2} \right) + O_p(N^{-\alpha}). \quad (44)$$

However,

$$\ln \left(1 + \frac{2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2 \right]}{\bar{\beta}_N^2 s_f^2} \right) = \frac{2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2 \right]}{\bar{\beta}_N^2 s_f^2} + o_p(a_{TN}), \quad (45)$$

where

$$a_{TN} = \frac{2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_t^2 - \bar{u}^2 \right]}{\bar{\beta}_N^2 s_f^2}.$$

Consider the first term of the RHS of (45). We have,

$$\frac{2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right]}{\bar{\beta}_N^2 s_f^2} = \frac{\frac{2}{\sqrt{TN}} \left[\frac{1}{\sigma_f \sqrt{T}} \sum_{t=1}^T (f_t - \bar{f}) (\sqrt{N} \bar{u}_t) \right]}{(s_f \bar{\beta}_N) (s_f / \sigma_f)}.$$

We note that $s_f / \sigma_f = 1 + O_p(T^{-1/2})$. But, by Lemma 2 (as N and $T \rightarrow \infty$)

$$\frac{1}{\sigma_f \sqrt{T}} \sum_{t=1}^T (f_t - \bar{f}) (\sqrt{N} \bar{u}_t) \rightarrow_p N(0, \bar{\sigma}_\infty^2), \quad (46)$$

where $\bar{\sigma}_\infty^2$ is as in (40).

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

$$\begin{aligned} \frac{1}{\bar{\beta}_N} - \frac{1}{N^{\alpha-1} \bar{v}_N} &= \frac{1}{N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N}} - \frac{1}{N^{\alpha-1} \bar{v}_N} = \frac{-K_\rho N^{-1}}{N^{2\alpha-2} \bar{v}_N^2 + K_\rho N^{\alpha-2} \bar{v}_N} \\ &= -N^{1-2\alpha} K_\rho \bar{v}_N^{-1} (\bar{v}_N + K_\rho N^{-\alpha})^{-1} = O_p(N^{1-2\alpha}), \end{aligned}$$

and hence

$$\begin{aligned} \frac{2\bar{\beta}_N \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_t \right]}{\bar{\beta}_N^2 s_f^2} &= \frac{\frac{2}{\sqrt{TN}} \left[\frac{1}{\sigma_f \sqrt{T}} \sum_{t=1}^T (f_t - \bar{f}) (\sqrt{N} \bar{u}_t) \right]}{s_f \bar{\beta}_N (s_f / \sigma_f)} \\ &= O_p \left(T^{-1/2} N^{1/2-2\alpha} \right) + O_p \left(T^{-1/2} N^{1/2-2\alpha} \right). \end{aligned} \quad (47)$$

Consider now the second term on the RHS of (45). We have

$$\begin{aligned} \frac{1}{\bar{\beta}_N^2} - \frac{1}{N^{2\alpha-2} \bar{v}_N^2} &= \frac{-N^{\alpha-2} K_\rho \bar{v}_N - N^{-2} K_\rho^2}{N^{4\alpha-4} \bar{v}_N^4 + N^{3\alpha-4} K_\rho \bar{v}_N^3 + N^{2\alpha-4} K_\rho^2 \bar{v}_N^2} = \\ &- (N^{2-3\alpha} K_\rho \bar{v}_N^3 + N^{2-4\alpha} K_\rho^2 \bar{v}_N^2) (\bar{v}_N + N^{-\alpha} K_\rho \bar{v}_N + N^{-2\alpha} K_\rho^2) = O_p(N^{2-3\alpha}). \end{aligned}$$

Note that since, by Lemma 1 and Theorems 17.5 and 19.11 of (Davidson, 1994, Theorem 15.18), $\sqrt{NT} \bar{u} = O_p(1)$, then, since $s_f^2 / \sigma_f^2 = 1 + O_p(T^{-1/2})$ and $0 < \sigma_f^2 < \infty$,

$$\frac{\bar{u}^2}{\left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} = \frac{\left(\sqrt{NT} \bar{u} \right)^2}{NT \left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} = O_p \left(T^{-1} N^{1-2\alpha} \right). \quad (48)$$

Similarly,

$$\begin{aligned} \frac{\frac{1}{T} \sum_{t=1}^T \bar{u}_t^2}{\left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} &= \frac{\frac{1}{N\sqrt{T}} \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^T \left[(\sqrt{N} \bar{u}_t)^2 - \bar{\sigma}_N^2 \right] + \sqrt{T} \bar{\sigma}_N^2 \right\}}{\left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} = \frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}} \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^T \left[\left(\frac{\sqrt{N} \bar{u}_t}{\bar{\sigma}_N} \right)^2 - 1 \right] + \sqrt{T} \right\}}{\left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} \\ &= \frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}} \frac{1}{\sqrt{T}} \sum_{t=1}^T \left[\left(\frac{\sqrt{N} \bar{u}_t}{\bar{\sigma}_N} \right)^2 - 1 \right]}{\left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} + \frac{\bar{\sigma}_N^2}{N \left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2}. \end{aligned}$$

Note that

$$\frac{\bar{\sigma}_N^2}{N \left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} - \frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} = O_p(N^{1-3\alpha}). \quad (49)$$

But, by Lemma 3,

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

and

$$\frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \left[\left(\frac{\sqrt{N} \bar{u}_t}{\bar{\sigma}_N} \right)^2 - 1 \right] \right)}{\left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} = O_p(T^{-1/2} N^{1-2\alpha}) + O_p(T^{-1/2} N^{1-3\alpha}). \quad (50)$$

Therefore, collecting all results derived above, and keeping the highest order terms of the RHS of (44), (47), (48), (49) and (50), we have

$$2 \ln(N) (\hat{\alpha} - \alpha) - \ln(s_f^2 \bar{v}_N^2) - \frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} = O_p \left(\max \left(T^{-1/2} N^{1/2-2\alpha}, T^{-1} N^{1-2\alpha}, T^{-1/2} N^{1-2\alpha}, N^{1-3\alpha}, N^{-\alpha} \right) \right).$$

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

$$\hat{\alpha} - \alpha = O_p \left(\frac{1}{\ln(N)} \right), \quad (51)$$

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

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

$$\begin{aligned} \ln(N) (\hat{\alpha} - \alpha) - \frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} &= \ln(s_f^2 \bar{v}_N^2) + \frac{\frac{2}{\sqrt{TN}} \left[\frac{1}{\sigma_f \sqrt{T}} \sum_{t=1}^T (f_t - \bar{f}) (\sqrt{N} \bar{u}_t) \right]}{s_f N^{\alpha-1} \bar{v}_N (s_f / \sigma_f)} + \\ &\frac{\left(\sqrt{NT} \bar{u} \right)^2}{NT \left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} + \frac{\frac{\bar{\sigma}_N^2}{N\sqrt{T}} \frac{1}{\sqrt{T}} \sum_{t=1}^T \left[\left(\frac{\sqrt{N} \bar{u}_t}{\bar{\sigma}_N} \right)^2 - 1 \right]}{\left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} + O_p(N^{-\alpha}). \end{aligned}$$

We first examine $\ln(s_f^2 \bar{v}_N^2)$. By Lemma 5 we have

$$\sqrt{\min(N^\alpha, T)} (\ln(s_f^2 \bar{v}_N^2) - \ln(\sigma_f^2 \mu_v^2)) \rightarrow_d N(0, \omega).$$

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

$$\sqrt{\min(N^\alpha, T)} \left(\frac{\frac{2}{\sqrt{TN}} \left[\frac{1}{\sigma_f \sqrt{T}} \sum_{t=1}^T (f_t - \bar{f}) (\sqrt{N} \bar{u}_t) \right]}{s_f N^{\alpha-1} \bar{v}_N (s_f / \sigma_f)} \right) = O_p \left(\sqrt{\min(N^\alpha, T)} T^{-1/2} N^{1/2-\alpha} \right) = o_p(1).$$

Similarly,

$$\sqrt{\min(N^\alpha, T)} \left(\frac{(\sqrt{NT} \bar{u})^2}{NT \left(N^{\alpha-1} \bar{v}_N + \frac{K_\rho}{N} \right)^2 s_f^2} \right) = O_p \left(\sqrt{\min(N^\alpha, T)} T^{-1} N^{1-2\alpha} \right) = o_p(1),$$

and

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

Thus,

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

where $\alpha_N^* = \alpha + \ln(\sigma_f^2 \mu_v^2) / \ln(N)$.

Proof of Theorem 2

We need to show that as long as either $\alpha > 4/7$ or $T^{1/2}/N^{4\alpha-2} \rightarrow 0$, $\sqrt{\min(N^\alpha, T)} \ln(N) \left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} - \frac{\widehat{\bar{\sigma}_N^2}}{N \widehat{\bar{\sigma}_x^2}} \right) = o_p(1)$. The result follows immediately by Lemma 7.

Proof of Theorem 3

The result follows immediately by Lemma 8.

Proof of Theorem 4

The proof follows immediately from Lemmas 9 and 10.

Proof of Theorem 5

Under the general factor model, we have

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

where

$$\mathbf{S}_{ff} = \frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) (f_t - \bar{f})' \rightarrow_p \boldsymbol{\Sigma}_{ff} > 0, \text{ as } T \rightarrow \infty.$$

We have that

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

So

$$\begin{aligned} \ln \left(\bar{\beta}'_N \mathbf{S}_{ff} \bar{\beta}_N \right) &= \ln \left(N^{2\alpha-2} \bar{v}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{D}_N \bar{v}_N + 2N^{\alpha-2} \bar{v}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{K}_\rho + N^{-2} \mathbf{K}'_\rho \mathbf{S}_{ff} \mathbf{K}_\rho \right) = \\ &= 2(\alpha-1) \ln(N) + \ln \left(\bar{v}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{D}_N \bar{v}_N \right) + \ln \left(1 + \frac{2N^{-\alpha} \bar{v}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{K}_\rho + N^{-2\alpha} \mathbf{K}'_\rho \mathbf{S}_{ff} \mathbf{K}_\rho}{\bar{v}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{D}_N \bar{v}_N} \right). \end{aligned}$$

Then,

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

So

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

or

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

where

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

Similarly to (47),

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

Also

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

where

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

So,

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

Using the above derivations gives straightforward extensions of Lemmas 7 and 8. Using these, we get

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

and

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

which together with Lemmas 11, 12 and 13 prove the theorem.

Appendix III: Proofs of Lemmas

Proof of Lemma 1

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 2. Moreover, writing $\|\cdot\|_r$ for the L_r -norm, we have, by Minkowski's inequality,

$$\sup_i \left\| u_{it} - E(u_{it} | \mathcal{F}_{t,|m|}^{\nu_i}) \right\|_2 = \sup_i \left\| \sum_{j=m+1}^{\infty} \psi_{ij} \left(\sum_{|s| \geq m} \xi_{is} \nu_{st} \right) \right\|_2 \leq \sup_{i,t} \|v_{it}\|_2 \left(\sup_i \sum_{j=m+1}^{\infty} |\psi_{ij}| \right) \left(\sup_i \left(\sum_{|s| \geq m} |\xi_{is}| \right) \right), \quad (52)$$

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 2 implies that $\sup_i \lim_{m \rightarrow \infty} m^\zeta \sum_{j=m+1}^{\infty} |\psi_{ij}| = O(1)$ and $\sup_i \lim_{m \rightarrow \infty} m^\zeta \left(\sum_{|s| \geq m} |\xi_{is}| \right) = O(1)$. 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 f_t is an L_r -bounded ($r \geq 2$) L_2 -NED processes of size $-\zeta$.

Proof of Lemma 2

We have $\frac{1}{\sqrt{T}} \sum_{t=1}^T (f_t - \bar{f}) (\sqrt{N} \bar{u}_t) = \frac{1}{\sqrt{T}} \sum_{t=1}^T z_t$, where $z_t = (f_t - \bar{f}) (\sqrt{N} \bar{u}_t)$. We have that z_t is a stationary process such that $E(z_t) = 0$. 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,

$$E\left(|\sqrt{N} \bar{u}_t|^r\right) = E\left(\left|\frac{1}{\sqrt{N}} \sum_{i=1}^N \sum_{l=0}^{\infty} \left(\psi_{il} \sum_{s=-\infty}^{\infty} \xi_{is} v_{st-l}\right)\right|^r\right) \leq c \left(\frac{1}{N} \sum_{i=1}^N \left(\sum_{l=0}^{\infty} |\psi_{il}|^2\right) \left(\sum_{s=-\infty}^{\infty} |\xi_{is}|^2\right)\right)^{r/2} \sup_{i,t} E(|v_{it}|^r) \leq c \left(\sup_i \left(\sum_{l=0}^{\infty} |\psi_{il}|^2\right) \sup_i \left(\sum_{s=-\infty}^{\infty} |\xi_{is}|^2\right)\right)^{r/2} \left(\sup_{i,t} E(|v_{it}|^r)\right) < \infty. \quad (53)$$

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 $\{f_t\}$ 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), $\{z_t\}$ is 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 ν_{ft} 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 $\{z_t\}$ is an L_2 -mixingale of size $-1/2$, if $2\eta > \zeta$, which automatically holds by the i.i.d. property of ν_{it} and ν_{ft} . This, implies the result of the Lemma by Theorem 24.6 of (Davidson, 1994, Theorem 15.18).

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 (53), $\left(\sqrt{N} \bar{u}_t\right)^2$ 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, Theorem 15.18), the result follows.

Proof of Lemma 4

We need to show that $\widehat{\sigma}_N^2 - \bar{\sigma}_\infty^2 = O_p(T^{-1}) + O_p((NT)^{-1/2})$. We have that $\widehat{\sigma}_N^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2$, where $\hat{u}_{it} = x_{it} - \hat{\delta}_i \tilde{x}_t$. Then, $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T u_{it}^2 + \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (\hat{u}_{it}^2 - u_{it}^2)$, where $u_{it} = x_{it} - \delta_i \tilde{x}_t$. Following similar lines to those of the proof of Lemma 3 we have that $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T u_{it}^2 \rightarrow_p \bar{\sigma}_\infty^2$. Further, $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (\hat{u}_{it}^2 - u_{it}^2) = O_p((NT)^{-1/2})$. Next, we examine $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (\hat{u}_{it}^2 - u_{it}^2)$. It is sufficient to consider $\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T u_{it} (\hat{u}_{it} - u_{it})$. Noting that $\frac{\tilde{x}_t}{N^{1-\alpha} \bar{x}_t} = O_p(1)$, and denoting equality in order of probability by $=_p$, we have

$$\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T u_{it} (\hat{u}_{it} - u_{it}) = \frac{1}{NT} \sum_{i=1}^N (\delta_i - \hat{\delta}_i) \sum_{t=1}^T \tilde{x}_t u_{it} =_p \frac{1}{NT} \left(\frac{1}{\sum_{t=1}^T (N^{1-\alpha} \bar{x}_t)^2} \right) \sum_{i=1}^N \left(\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T (N^{1-\alpha} \bar{x}_t) u_{it} \right)^2 - E \left(\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T (N^{1-\alpha} \bar{x}_t) u_{it} \right)^2 \right) \right) +$$

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

But $E \left(\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T (N^{1-\alpha} \bar{x}_t) u_{it} \right)^2 \right) < \infty$ and $\frac{1}{T} \sum_{t=1}^T (N^{1-\alpha} \bar{x}_t)^2 = O_p(1)$, which implies that

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

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

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

proving the required result.

Proof of Lemma 5

We have that

$$\begin{aligned} \ln(s_f^2 \bar{v}_N^2) - \ln(\sigma_f^2 v_\infty^2) &= \ln \left(\frac{s_f^2 \bar{v}_N^2}{\sigma_f^2 v_\infty^2} \right) = \ln \left(\frac{s_f^2}{\sigma_f^2} \right) + \ln \left(\frac{\bar{v}_N^2}{v_\infty^2} \right) = \left(\frac{s_f^2 - \sigma_f^2}{\sigma_f^2} \right) + \left(\frac{\bar{v}_N^2 - \mu_v^2}{\mu_v^2} \right) + \\ &O_p \left((s_f^2 - \sigma_f^2)^2 \right) + O_p \left((\bar{v}_N^2 - \mu_v^2)^2 \right). \end{aligned}$$

But, under Assumption 2,

$$\sqrt{T} \left(\frac{s_f^2 - \sigma_f^2}{\sigma_f^2} \right) = \frac{1}{\sqrt{T}} \sum_{t=1}^T \left\{ [(f_t - \bar{f})/\sigma_f]^2 - 1 \right\} \rightarrow_d N(0, V_{\bar{f}^2}),$$

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

$$V_{\bar{f}^2} = E \left(\left([(f_t - \mu_f)/\sigma_f]^2 - 1 \right)^2 \right) + \sum_{i=1}^{\infty} Cov \left(\left([(f_t - \mu_f)/\sigma_f]^2 - 1 \right) \left([(f_{t-i} - \mu_f)/\sigma_f]^2 - 1 \right) \right).$$

Further, recalling that $\bar{v}_N = \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_i$, $\sqrt{N^\alpha} \left(\frac{\bar{v}_N^2 - \mu_v^2}{\mu_v^2} \right) = \sqrt{N^\alpha} \left(\frac{\bar{v}_N - \mu_v}{\mu_v} \right) \left(\frac{\bar{v}_N + \mu_v}{\mu_v} \right)$. But $\frac{\bar{v}_N + \mu_v}{\mu_v} \rightarrow_p 2$, and

$$\sqrt{N^\alpha} \left(\frac{\bar{v}_N - \mu_v}{\mu_v} \right) \rightarrow_d N \left(0, \frac{\sigma_v^2}{\mu_v^2} \right). \quad (54)$$

Further, $E \left[\left(\frac{s_f^2 - \sigma_f^2}{\sigma_f^2} \right) \left(\frac{\bar{v}_N^2 - \mu_v^2}{\mu_v^2} \right) \right] = 0$, implying that $\sqrt{\min(N^\alpha, T)} (\ln(s_f^2 \bar{v}_N^2) - \ln(\sigma_f^2 v_\infty^2)) \rightarrow_d N(0, \omega)$, where $\omega = \lim_{N, T \rightarrow \infty} \left[\frac{\min(N^\alpha, T)}{T} V_{\bar{f}^2} + \frac{\min(N^\alpha, T)}{N^\alpha} \frac{4\sigma_v^2}{\mu_v^2} \right]$.

Proof of Lemma 6

The proof follows easily along the same lines as that of Lemma 5. In the present case under Assumption 6 we have $\bar{v}_N = N^{-1} \sum_{i=1}^N v_i$, and thus $\sqrt{N} \left(\frac{\bar{v}_N^2 - \mu_v^2}{\mu_v^2} \right) = \sqrt{N} \left(\frac{\bar{v}_N - \mu_v}{\mu_v} \right) \left(\frac{\bar{v}_N + \mu_v}{\mu_v} \right)$, and $\frac{\bar{v}_N + \mu_v}{\mu_v} \rightarrow_p 2$. Therefore, $\sqrt{N} \left(\frac{\bar{v}_N - \mu_v}{\mu_v} \right) \rightarrow_d N \left(0, \frac{\sigma_v^2}{\mu_v^2} \right)$.

Proof of Lemma 7

We need to show that

$$\sqrt{\min(N^\alpha, T)} \left(\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} - \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2} \right) = o_p(1). \quad (55)$$

We have

$$\frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} - \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2} = \frac{\bar{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} - \frac{\widehat{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} + \frac{\widehat{\sigma}_N^2}{N^{2\alpha-1} \bar{v}_N^2 s_f^2} - \frac{\widehat{\sigma}_N^2}{N \widehat{\sigma}_x^2}.$$

But

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

which is negligible as a bias. Next,

$$\frac{\widehat{\bar{\sigma}_N^2}}{N^{2\alpha-1}\bar{v}_N^2 s_f^2} - \frac{\bar{\sigma}_N^2}{N\hat{\sigma}_x^2} = \bar{\sigma}_N^2 \left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right) \left(\frac{1}{N^{2\alpha-1}\bar{v}_N^2 s_f^2} - \frac{1}{N\hat{\sigma}_x^2}\right) = \bar{\sigma}_N^2 \left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right) \left(\frac{1}{N^{2\alpha-1}\bar{v}_N^2 s_f^2}\right) N^{2\alpha-1} (N^{2-2\alpha}\hat{\sigma}_x^2 - \bar{v}_N^2 s_f^2) \left(\frac{1}{N\hat{\sigma}_x^2}\right).$$

But by the proof of Theorem 1, we have

$$\bar{v}_N^2 s_f^2 - N^{2-2\alpha}\hat{\sigma}_x^2 = O_p\left(T^{-1/2}N^{1/2-2\alpha}\right) + O_p\left(T^{-1}N^{1-2\alpha}\right) + O_p(N^{1-3\alpha}) + O_p(N^{-\alpha}) + O_p(N^{1-2\alpha}).$$

So

$$\begin{aligned} & \bar{\sigma}_N^2 \left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right) \left(\frac{1}{N^{2\alpha-1}\bar{v}_N^2 s_f^2}\right) N^{2\alpha-1} (N^{2-2\alpha}\hat{\sigma}_x^2 - \bar{v}_N^2 s_f^2) \left(\frac{1}{N\hat{\sigma}_x^2}\right) = \\ & O_p\left(T^{-1/2}N^{1/2-2\alpha}N^{1-2\alpha}\right) + O_p\left(T^{-1}N^{1-2\alpha}N^{1-2\alpha}\right) + O_p(N^{1-3\alpha}N^{1-2\alpha}) + O_p(N^{-\alpha}N^{1-2\alpha}) + O_p(N^{1-2\alpha}N^{1-2\alpha}) = \\ & O_p\left(T^{-1/2}N^{3/2-3\alpha}\right) + O_p\left(T^{-1}N^{2-4\alpha}\right) + O_p(N^{2-5\alpha}) + O_p(N^{1-3\alpha}) + O_p(N^{2-4\alpha}). \end{aligned}$$

Therefore, as long as $\alpha > 4/7$, (55) holds, proving the Lemma.

Proof of Lemma 8

We have that

$$\bar{\sigma}_N^2 \left(\frac{\widehat{\bar{\sigma}_N^2}}{\bar{\sigma}_N^2}\right) \left(\frac{1}{N^{2\alpha-1}\bar{v}_N^2 s_f^2}\right) N^{2\alpha-1} \left(N^{1-2\alpha} \left(\frac{1}{T} \sum_{t=1}^T (\sqrt{N}\bar{u}_t)^2 - \widehat{\bar{\sigma}_N^2}\right)\right) \left(\frac{1}{N\hat{\sigma}_x^2}\right) = O_p\left(N^{2-4\alpha}T^{-1/2}\right),$$

which is negligible as long as $\alpha > 1/2$. To see the above result note the following. We have

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

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

Proof of Lemma 9

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_i , is $o_p(N^\alpha)$. The first thing to note is that we abstract from the possibility that any $v_i = 0$. By the fact that $\Pr(v_i = 0) = 0$, it follows that the number of units that have $v_i = 0$ is $o_p(N^\alpha)$. Without loss of generality, we further assume that units whose loadings do not depend on any v_i have zero loadings. The probability that $o_p(N^\alpha)$ units are misclassified is bounded from above by

$$C_1 N^{-\alpha+c} \Pr\left(\left(\bigcup_{i=1}^{\lfloor N^\alpha \rfloor} \{|\hat{v}_i - v_i| > \varepsilon\}\right) \cup \left(\bigcup_{i=\lfloor N^\alpha \rfloor+1}^N \{|\hat{v}_i| > \varepsilon\}\right)\right),$$

for some $C_1 > 0$, some $\varepsilon > 0$, and any $c > 0$, where $\Pr\left(\left(\bigcup_{i=1}^{\lfloor N^\alpha \rfloor} \{|\hat{v}_i - v_i| > \varepsilon\}\right) \cup \left(\bigcup_{i=\lfloor N^\alpha \rfloor+1}^N \{|\hat{v}_i| > \varepsilon\}\right)\right)$ is an upper bound for the probability that one unit is misclassified. But

$$\Pr\left(\left(\bigcup_{i=1}^{\lfloor N^\alpha \rfloor} \{|\hat{v}_i - v_i| > \varepsilon\}\right) \cup \left(\bigcup_{i=\lfloor N^\alpha \rfloor+1}^N \{|\hat{v}_i| > \varepsilon\}\right)\right) \leq \sum_{i=1}^{\lfloor N^\alpha \rfloor} \Pr(|\hat{v}_i - v_i| > \varepsilon) + \sum_{i=\lfloor N^\alpha \rfloor+1}^N \Pr(|\hat{v}_i| > \varepsilon).$$

But, by Markov's inequality $\Pr(|\hat{v}_i - v_i| > \varepsilon) \leq \frac{E(|\hat{v}_i - v_i|^2)}{\varepsilon^2} = O(T^{-1})$, $i = 1, \dots, \lfloor N^\alpha \rfloor$ and $\Pr(|\hat{v}_i| > \varepsilon) \leq \frac{E(|\hat{v}_i|^2)}{\varepsilon^2} = O(T^{-1})$, $i = \lfloor N^\alpha \rfloor + 1, \dots, N$. Therefore, the probability that $o_p(N^\alpha)$ units are misclassified is $o(N^{1-\alpha+c}T^{-1})$ for all

$c > 0$. Since $\alpha > 1/2$, this probability goes to zero as long as $\lim_{T, N \rightarrow \infty} T^{-1} N^\alpha < \infty$, proving the first step in the proof. Next, we show the first part of 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_i have zero loadings, $x_{it} = \frac{v_i}{\bar{v}_N} (N^{1-\alpha} \bar{x}_t) + u_{it}$. We analyse \hat{v}_i by defining 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 \bar{x}_t . This is clearly the case once the normalisation $N^{2\alpha-2}$ is taken into account in (41). Let $v_i^{(1)} = \frac{v_i}{\mu_v}$ and $v_{Ni} = \frac{v_i}{\bar{v}_N}$. We need to

show that $\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(\hat{v}_i - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} \hat{v}_i \right)^2 - \frac{\sigma_v^2}{\mu_v^2} = o_p(1)$. We have

$$\begin{aligned} \frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(\hat{v}_i - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} \hat{v}_i \right)^2 - \frac{\sigma_v^2}{\mu_v^2} &= \frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(\hat{v}_i - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} \hat{v}_i \right)^2 - \frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_{Ni} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_{Ni} \right)^2 + \\ &\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_{Ni} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_{Ni} \right)^2 - \frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_i^{(1)} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_i^{(1)} \right)^2 + \frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_i^{(1)} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_i^{(1)} \right)^2 - \frac{\sigma_v^2}{\mu_v^2}. \end{aligned}$$

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

$$\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_i^{(1)} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_i^{(1)} \right)^2 - \frac{\sigma_v^2}{\mu_v^2} = O_p(N^{-1/2}).$$

It is sufficient to show that

$$\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(\hat{v}_i - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} \hat{v}_i \right)^2 - \frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_{Ni} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_{Ni} \right)^2 = o_p(1) \quad (56)$$

and

$$\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_{Ni} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_{Ni} \right)^2 - \frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(v_i^{(1)} - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} v_i^{(1)} \right)^2 = o_p(1). \quad (57)$$

For (56), it is sufficient that $\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} (\hat{v}_i - v_{Ni}) = o_p(1)$. Recall that $x_{it} = \frac{v_i}{\bar{v}_N} (N^{1-\alpha} \bar{x}_t) + u_{it}$. So

$$\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} (\hat{v}_i - v_{Ni}) = \frac{\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(\sum_{t=1}^T \bar{x}_t u_{it} \right)}{\sum_{t=1}^T \bar{x}_t^2} = \frac{1}{(N^\alpha - 1) \bar{v}_N \left(\sum_{t=1}^T f_t^2 \right)} \sum_{i=1}^{N^\alpha} \sum_{t=1}^T f_t u_{it}.$$

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

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

For (57), it is sufficient that $\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} (v_{Ni} - v_i^{(1)}) = o_p(1)$. We have,

$$\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} (v_{Ni} - v_i^{(1)}) = \frac{\left(\frac{1}{\bar{v}_N} - \frac{1}{\mu_v} \right) \sum_{i=1}^{N^\alpha} v_i}{N^\alpha - 1}.$$

But $\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} v_i = O_p(1)$. Also $\frac{1}{\bar{v}_N} - \frac{1}{\mu_v} = \frac{1}{\bar{v}_N \mu_v} (\bar{v}_N - \mu_v) = O_p(N^{-\alpha/2})$. So, overall $\frac{1}{N^\alpha - 1} \sum_{i=1}^{N^\alpha} \left(\hat{v}_i - \frac{1}{N^\alpha} \sum_{i=1}^{N^\alpha} \hat{v}_i \right)^2 - \frac{\sigma_v^2}{\mu_v^2} = o_p(1)$, proving the result.

Proof of Lemma 10

We need to show $\hat{V}_{f^2} - V_{f^2} = o_p(1)$. We have $\hat{V}_{f^2} - V_{f^2} = \hat{V}_{f^2} - \bar{V}_{f^2} + \bar{V}_{f^2} - V_{f^2}$ where

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

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

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \left(\frac{f_t}{s_f} - \tilde{x}_t \right) &= \frac{1}{T} \sum_{t=1}^T \left(\frac{f_t}{s_f} - \frac{\bar{x}_t}{\hat{\sigma}_{\bar{x}}} \right) = \frac{1}{s_f T} \sum_{t=1}^T \left(f_t - \frac{\bar{x}_t}{N^{\alpha-1} \bar{v}_N} \right) = \\ \frac{1}{s_f T} \sum_{t=1}^T \left(f_t - \frac{\bar{v}_N N^{\alpha-1} f_t + \frac{1}{N} \sum_{i=1}^N u_{it}}{N^{\alpha-1} \bar{v}_N} \right) &= \frac{N^{-\alpha}}{s_f T} \sum_{t=1}^T \sum_{i=1}^N u_{it} = \frac{N^{-\alpha}}{\sigma_f T} \sum_{t=1}^T \sum_{i=1}^N u_{it} + o_p \left(\frac{N^{-\alpha}}{\sigma_f T} \sum_{t=1}^T \sum_{i=1}^N u_{it} \right). \end{aligned}$$

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

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,

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

Then,

$$\begin{aligned} \ln \left(\frac{\bar{v}_{1N}^2 s_{1f}^2 + 2\bar{v}_{1N} \bar{v}_{2N} s_{12f} + \bar{v}_{2N}^2 s_{2f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} \right) &= \frac{\bar{v}_{1N}^2 s_{1f}^2 + 2\bar{v}_{1N} \bar{v}_{2N} s_{12f} + \bar{v}_{2N}^2 s_{2f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} - 1 = \\ &= \frac{(\bar{v}_{1N}^2 s_{1f}^2 - \sigma_{1f}^2 \mu_{1v}^2) + (\bar{v}_{2N}^2 s_{2f}^2 - \sigma_{2f}^2 \mu_{2v}^2) + 2\bar{v}_{1N} \bar{v}_{2N} s_{12f}}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} = \\ &= \frac{(\bar{v}_{1N}^2 s_{1f}^2 - \bar{v}_{1N}^2 \sigma_{1f}^2 + \bar{v}_{1N}^2 \sigma_{1f}^2 - \sigma_{1f}^2 \mu_{1v}^2) + (\bar{v}_{2N}^2 s_{2f}^2 - \bar{v}_{2N}^2 \sigma_{2f}^2 + \bar{v}_{2N}^2 \sigma_{2f}^2 - \sigma_{2f}^2 \mu_{2v}^2) + 2\mu_{1v} \mu_{2v} s_{12f}}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} = \\ &= \frac{\mu_{1v}^2 (s_{1f}^2 - \sigma_{1f}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} + \frac{\sigma_{1f}^2 (\bar{v}_{1N}^2 - \mu_{1v}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} + \frac{\mu_{2v}^2 (s_{2f}^2 - \sigma_{2f}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} + \frac{\sigma_{2f}^2 (\bar{v}_{2N}^2 - \mu_{2v}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} + \frac{2\mu_{1v} \mu_{2v} s_{12f}}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2}. \end{aligned} \tag{58}$$

Note that

$$\begin{aligned} \bar{v}_{1N} \bar{v}_{2N} s_{12f} &= \bar{v}_{1N} \bar{v}_{2N} s_{12f} - \bar{v}_{1N} \mu_{2v} s_{12f} + \bar{v}_{1N} \mu_{2v} s_{12f} - \bar{v}_{1N} \mu_{2v} \sigma_{12f} + \bar{v}_{1N} \mu_{2v} \sigma_{12f} - 2\mu_{1v} \mu_{2v} \sigma_{12f} = \\ &= s_{12f} \bar{v}_{1N} (\bar{v}_{2N} - \mu_{2v}) + \bar{v}_{1N} \mu_{2v} (s_{12f} - \sigma_{12f}) + \sigma_{12f} \mu_{2v} (\bar{v}_{1N} - 2\mu_{1v}) = \\ &= (s_{12f} - \sigma_{12f}) \bar{v}_{1N} (\bar{v}_{2N} - \mu_{2v}) + \sigma_{12f} \bar{v}_{1N} (\bar{v}_{2N} - \mu_{2v}) + \bar{v}_{1N} \mu_{2v} (s_{12f} - \sigma_{12f}) + \sigma_{12f} \mu_{2v} (\bar{v}_{1N} - 2\mu_{1v}). \end{aligned}$$

But

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

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

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

Then,

$$\begin{aligned} \frac{\mu_{iv}^2 (s_{if}^2 - \sigma_{if}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} &= \frac{\mu_{iv}^2 \sigma_{if}^2}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} \frac{(s_{if}^2 - \sigma_{if}^2)}{\sigma_{if}^2}, \quad i = 1, 2, \\ \frac{\sigma_{if}^2 (\bar{v}_{iN}^2 - \mu_{iv}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} &= \frac{\mu_{iv}^2 \sigma_{if}^2}{\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2} \frac{(\bar{v}_{iN}^2 - \mu_{iv}^2)}{\mu_{iv}^2}, \quad i = 1, 2, \end{aligned}$$

or if $\sigma_{1f}^2 \mu_{1v}^2 + \sigma_{2f}^2 \mu_{2v}^2 = 1$,

$$\begin{aligned} \mu_{iv}^2 (s_{if}^2 - \sigma_{if}^2) &= \mu_{iv}^2 \sigma_{if}^2 \frac{(s_{if}^2 - \sigma_{if}^2)}{\sigma_{if}^2}, \quad i = 1, 2, \\ \sigma_{if}^2 (\bar{v}_{iN}^2 - \mu_{iv}^2) &= \mu_{iv}^2 \sigma_{if}^2 \frac{(\bar{v}_{iN}^2 - \mu_{iv}^2)}{\mu_{iv}^2}, \quad i = 1, 2. \end{aligned}$$

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

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

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

Further,

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

Further, by factor independence

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

So

$$\begin{aligned} & \sqrt{\frac{\min(N^\alpha, T)}{T}} \left(\mu_{1v}^2 \sigma_{1f}^2 \left(\sqrt{T} \frac{(s_{1f}^2 - \sigma_{1f}^2)}{\sigma_{1f}^2} \right) + \mu_{2v}^2 \sigma_{2f}^2 \left(\sqrt{T} \frac{(s_{2f}^2 - \sigma_{2f}^2)}{\sigma_{2f}^2} \right) + 2\mu_{1v} \mu_{2v} \sqrt{T} s_{12f} \right) + \\ & \sqrt{\frac{\min(N^\alpha, T)}{N^\alpha}} \left(\mu_{1v}^2 \sigma_{1f}^2 \left(\sqrt{N^\alpha} \left(\frac{\bar{v}_{1N}^2 - \mu_{1v}^2}{\mu_{1v}^2} \right) \right) + \mu_{2v}^2 \sigma_{2f}^2 \left(\sqrt{N^\alpha} \left(\frac{\bar{v}_{2N}^2 - \mu_{2v}^2}{\mu_{2v}^2} \right) \right) \right) \rightarrow_d \\ & N \left(0, \frac{\min(N^\alpha, T)}{T} \left((\mu_{1v}^2 \sigma_{1f}^2)^2 \mu_1^{(4)} + (\mu_{2v}^2 \sigma_{2f}^2)^2 \mu_2^{(4)} + 2\mu_{1v}^2 \sigma_{1f}^2 \mu_{2v}^2 \sigma_{2f}^2 \right) \right. \\ & \quad \left. + \frac{\min(N^\alpha, T)}{N^\alpha} (2\sigma_{1v}^2 \mu_{1v}^2 \sigma_{1f}^2 + 2\sigma_{2v}^2 \mu_{2v}^2 \sigma_{2f}^2) \right). \end{aligned}$$

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,

$$\begin{aligned} & \ln \left(\bar{v}_{1N}^2 s_{1f}^2 + 2N^{\alpha_2 - \alpha} \bar{v}_{1N} \bar{v}_{2N} s_{12f} + N^{2(\alpha_2 - \alpha)} \bar{v}_{2N}^2 s_{2f}^2 \right) - \ln \left(\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2 \right) = \\ & \ln \left(\frac{\bar{v}_{1N}^2 s_{1f}^2 + 2N^{\alpha_2 - \alpha} \bar{v}_{1N} \bar{v}_{2N} s_{12f} + N^{2(\alpha_2 - \alpha)} \bar{v}_{2N}^2 s_{2f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} \right). \end{aligned}$$

Then, similarly to the proof of Lemma 11

$$\begin{aligned} & \ln \left(\frac{\bar{v}_{1N}^2 s_{1f}^2 + 2N^{\alpha_2 - \alpha} \bar{v}_{1N} \bar{v}_{2N} s_{12f} + N^{2(\alpha_2 - \alpha)} \bar{v}_{2N}^2 s_{2f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} \right) = \frac{\mu_{1v}^2 (s_{1f}^2 - \sigma_{1f}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} + \frac{\sigma_{1f}^2 (\bar{v}_{1N}^2 - \mu_{1v}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} + \\ & \frac{N^{2(\alpha_2 - \alpha)} \mu_{2v}^2 (s_{2f}^2 - \sigma_{2f}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} + \frac{N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 (\bar{v}_{2N}^2 - \mu_{2v}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} + \frac{2N^{\alpha_2 - \alpha} \mu_{1v} \mu_{2v} s_{12f}}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2}. \end{aligned}$$

Then,

$$\frac{\mu_{1v}^2 (s_{1f}^2 - \sigma_{1f}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} = \frac{\mu_{1v}^2 \sigma_{1f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} \frac{(s_{1f}^2 - \sigma_{1f}^2)}{\sigma_{1f}^2},$$

$$\frac{\sigma_{1f}^2 (\bar{v}_{1N}^2 - \mu_{1v}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} = \frac{\mu_{1v}^2 \sigma_{1f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} \frac{(\bar{v}_{1N}^2 - \mu_{1v}^2)}{\mu_{1v}^2},$$

$$\frac{N^{2(\alpha_2 - \alpha)} \mu_{2v}^2 (s_{2f}^2 - \sigma_{2f}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} = \frac{\mu_{2v}^2 \sigma_{2f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} \frac{N^{2(\alpha_2 - \alpha)} (s_{2f}^2 - \sigma_{2f}^2)}{\sigma_{2f}^2}, \quad (59)$$

$$\frac{N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 (\bar{v}_{2N}^2 - \mu_{2v}^2)}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} = \frac{\mu_{2v}^2 \sigma_{2f}^2}{\sigma_{1f}^2 \mu_{1v}^2 + N^{2(\alpha_2 - \alpha)} \sigma_{2f}^2 \mu_{2v}^2} \frac{N^{2(\alpha_2 - \alpha)} (\bar{v}_{2N}^2 - \mu_{2v}^2)}{\mu_{2v}^2}. \quad (60)$$

But, then it is obvious that the Lemma holds since (59) and (60) are $o_p(1)$, when multiplied by $\min(\sqrt{T}, \sqrt{N^\alpha})$ respectively, as well as $\min(\sqrt{T}, \sqrt{N^\alpha}) N^{\alpha_2 - \alpha} \mu_{1v} \mu_{2v} s_{12f}$.

Proof of Lemma 13

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

$$\ln(\boldsymbol{\mu}'_v \mathbf{D}_N \mathbf{\Sigma}_{ff} \mathbf{D}_N \boldsymbol{\mu}_v) = \ln \left(\mu_{1v}^2 \sigma_{1f}^2 + N^{2(\alpha_2 - \alpha)} \sum_{j=2}^m N^{2(\alpha_j - \alpha_2)} \mu_{jv}^2 \sigma_{jf}^2 \right).$$

Then,

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

So

$$\begin{aligned} & \sqrt{\min(N^\alpha, T)} \ln(N) \left(\ln(\bar{\mathbf{v}}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{D}_N \bar{\mathbf{v}}_N) - \ln(\boldsymbol{\mu}'_v \mathbf{D}_N \mathbf{\Sigma}_{ff} \mathbf{D}_N \boldsymbol{\mu}_v) \right) = \\ & \sqrt{\min(N^\alpha, T)} \ln(N) \left(\ln(\bar{\mathbf{v}}'_N \mathbf{D}_N \mathbf{S}_{ff} \mathbf{D}_N \bar{\mathbf{v}}_N) - \ln(\mu_{1v}^2 \sigma_{1f}^2) \right) - \sqrt{\min(N^\alpha, T)} \ln(N) N^{2(\alpha_2 - \alpha)} \left(\frac{\sum_{j=2}^m \left(N^{2(\alpha_j - \alpha_2)} \mu_{jv}^2 \sigma_{jf}^2 \right)}{\mu_{1v}^2 \sigma_{1f}^2} \right). \end{aligned}$$

We need

$$N^{2(\alpha_2 - \alpha)} \left(\frac{\sum_{j=2}^m \left(N^{2(\alpha_j - \alpha_2)} \mu_{jv}^2 \sigma_{jf}^2 \right)}{\mu_{1v}^2 \sigma_{1f}^2} \right) = o \left(\min(N^\alpha, T)^{-1/2} \ln(N)^{-1} \right).$$

This holds if $\sqrt{\min(N^\alpha, T)} N^{2(\alpha_2 - \alpha)} = o(1)$. 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

$$\frac{\ln(s_f^2 \bar{v}_N^2)}{\ln(N)} = o_p(1). \quad (61)$$

We have

$$\ln(s_f^2 \bar{v}_N^2) = \ln(s_f^2) + 2 \ln(\bar{v}_N) = \ln(s_f^2) + 2 \ln \left(\frac{1}{N} \sum_{i=1}^N \tilde{v}_i + \bar{c}_N \right).$$

So (61) holds if $\frac{1}{N} \sum_{i=1}^N \tilde{v}_i + \bar{c}_N = o_p(N^c)$ 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 (54). We have $\sqrt{N}(\bar{v}_N - \mu_v) = \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N \tilde{v}_i + \bar{c}_N - \mu_v \right)$. But, $\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N \tilde{v}_i - \mu_v \right) \rightarrow_d N(0, \sigma_v^2)$. Therefore, $\sqrt{N} \bar{c}_N = o_p(1)$ is sufficient for the second part of the Lemma to hold.

Appendix IV: Proof of consistency of joint estimation of α and κ

Without loss of generality, we consider a simplified version of the estimation problem. For simplicity, we assume $v_i = 0$ for $n > [N^\alpha]$. By the first part of the proof of Lemma 9, we can disregard the matter of misclassification, when ordering the estimated factor loadings. We start by noting that

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

where

$$\bar{x}_{nt} = \bar{\beta}_n f_t + \bar{u}_{nt}, \quad \bar{x}_n = \frac{1}{T} \sum_{t=1}^T \bar{x}_{nt} = \bar{\beta}_n \bar{f} + \bar{u}_n.$$

We have that

$$\hat{\sigma}_{\bar{x}n}^2 = \bar{\beta}_n^2 s_f^2 + 2\bar{\beta}_n \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_{nt} \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_{nt}^2 - \bar{u}_n^2 \right],$$

or

$$\hat{\sigma}_{\bar{x}n}^2 = \begin{cases} \kappa^2 + (\bar{v}_n^2 s_f^2 - \kappa^2) + 2\bar{\beta}_n \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_{nt} \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_{nt}^2 - \bar{u}_n^2 \right], & \text{if } n \leq [N^\alpha] \\ \frac{[N^\alpha]}{n} \kappa^2 + \frac{[N^\alpha]}{n} (\bar{v}_{[N^\alpha]}^2 s_f^2 - \kappa^2) + 2\bar{\beta}_n \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_{nt} \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_{nt}^2 - \bar{u}_n^2 \right], & \text{if } n > [N^\alpha] \end{cases},$$

since

$$\bar{\beta}_n = \frac{1}{n} \sum_{i=1}^n v_i = \begin{cases} \bar{v}_n, & \text{if } n \leq [N^\alpha] \\ \frac{[N^\alpha]}{n} \bar{v}_{[N^\alpha]}, & \text{if } n > [N^\alpha] \end{cases},$$

since $v_i = 0$ for $n > [N^\alpha]$. Let

$$\xi_n = \begin{cases} (\bar{v}_n^2 s_f^2 - \kappa^2) + 2\bar{\beta}_n \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_{nt} \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_{nt}^2 - \bar{u}_n^2 \right], & \text{if } n \leq [N^\alpha] \\ \frac{[N^\alpha]}{n} (\bar{v}_{[N^\alpha]}^2 s_f^2 - \kappa^2) + 2\bar{\beta}_n \left[\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_{nt} \right] + \left[\frac{1}{T} \sum_{t=1}^T \bar{u}_{nt}^2 - \bar{u}_n^2 \right], & \text{if } n > [N^\alpha] \end{cases}.$$

Then,

$$\hat{\sigma}_{\bar{x}n}^2 = \begin{cases} \kappa^2 + \xi_n, & \text{if } n \leq [N^\alpha] \\ \frac{[N^\alpha]}{n} \kappa^2 + \xi_n, & \text{if } n > [N^\alpha] \end{cases}. \quad (62)$$

But $\frac{1}{T} \sum_{t=1}^T (f_t - \bar{f}) \bar{u}_{nt} = O_p(n^{-1/2} T^{-1/2})$, uniformly over n ; $\frac{1}{T} \sum_{t=1}^T \bar{u}_{nt}^2 = O_p(n^{-1})$, uniformly over n ; $\bar{u}_n^2 = O_p(n^{-1})$, uniformly over n ; $\bar{v}_n^2 s_f^2 - \kappa^2 = O_p(\min(n^{-1/2}, T^{-1/2}))$, uniformly over n ; and $\bar{v}_{[N^\alpha]}^2 s_f^2 - \kappa^2 = O_p(\min(N^{-\alpha/2}, T^{-1/2}))$. Therefore,

$$\xi_n = o_p(1), \text{ uniformly over } n. \quad (63)$$

We estimate (62) by NLLS. Define

$$\hat{\xi}_n = \begin{cases} \hat{\sigma}_{\bar{x}n}^2 - \hat{\kappa}^2, & \text{if } n \leq [N^\alpha] \\ \hat{\sigma}_{\bar{x}n}^2 - \frac{[N^\alpha]}{n} \hat{\kappa}^2, & \text{if } n > [N^\alpha] \end{cases},$$

where $\hat{\kappa}^2$ and $\hat{\alpha}$ are the NLLS estimators of κ^2 and α , and

$$d_n = \hat{\xi}_n - \xi_n = \begin{cases} \kappa^2 - \hat{\kappa}^2, & \text{if } n \leq [N^\alpha] \\ \frac{[N^\alpha]}{n} \kappa^2 - \frac{[N^\alpha]}{n} \hat{\kappa}^2, & \text{if } n > [N^\alpha] \end{cases}.$$

So,

$$\frac{1}{N} \sum_{n=1}^N \hat{\xi}_n^2 = \frac{1}{N} \sum_{n=1}^N \xi_n^2 + \frac{1}{N} \sum_{n=1}^N d_n^2 + \frac{1}{N} \sum_{n=1}^N \xi_n d_n.$$

By the definition of the NLLS estimator

$$\lim_{n, T \rightarrow \infty} \Pr \left(\frac{1}{N} \sum_{n=1}^N \hat{\xi}_n^2 \leq \frac{1}{N} \sum_{n=1}^N \xi_n^2 \right) = 1. \quad (64)$$

If either

$$p \lim_{n, T \rightarrow \infty} \kappa^2 - \hat{\kappa}^2 \neq 0 \quad (65)$$

or

$$p \lim_{n, T \rightarrow \infty} \hat{\alpha} - \alpha \neq 0, \quad (66)$$

then $\frac{1}{N} \sum_{n=1}^N d_n^2 = O_p(1)$. Further, by (63),

$$\frac{1}{N} \sum_{n=1}^N \xi_n d_n = o_p(1).$$

Therefore, if either $p \lim_{n \rightarrow \infty} \kappa^2 - \hat{\kappa}^2 \neq 0$ or $p \lim_{n \rightarrow \infty} \hat{\alpha} - \alpha \neq 0$, then

$$\lim_{n, T \rightarrow \infty} \Pr \left(\frac{1}{N} \sum_{n=1}^N \hat{\xi}_n^2 - \frac{1}{N} \sum_{n=1}^N \xi_n^2 - C > 0 \right) > \varepsilon,$$

for some $C > 0$ and $\varepsilon > 0$. But this contradicts (64). Therefore, neither (65) nor (66) can hold, proving consistency.

Table A1: Bias, RMSE, size and power ($\times 100$) for the $\tilde{\alpha}$ estimate of the cross-sectional exponent -
 case of a serially uncorrelated factor and cross-sectionally independent idiosyncratic errors
 ($\beta_{2i} = 0, f_{1t} \sim IIDN(0, 1), u_{it} \sim IIDN(0, \sigma_i^2), v_{it} \sim IIDU(0.5, 1.5), N=100, 200, 500, 1000$ and $T=100, 200, 500$)

$N \setminus T$	α															
	0.70	0.75	0.80	0.85	0.90	0.95	1.00	1.00	0.95	0.90	0.85	0.80	0.75	0.70		
100	Bias	0.99	0.63	0.45	0.35	0.25	0.16	-0.07	100	11.10	8.25	6.40	6.10	5.65	5.75	7.05
	RMSE	2.48	2.21	2.04	1.92	1.83	1.77	1.72	Power+	52.75	61.30	66.50	70.45	75.00	78.35	84.50
200	Bias	0.51	0.31	0.19	0.10	0.06	0.02	-0.09	200	9.15	7.55	6.10	5.70	5.00	5.30	6.45
	RMSE	1.90	1.73	1.62	1.54	1.48	1.45	1.43	Power+	76.10	82.30	86.40	90.00	92.10	93.65	93.25
500	Bias	0.14	0.04	-0.00	-0.03	-0.05	-0.07	-0.11	500	91.50	92.20	93.15	93.15	93.85	94.10	93.25
	RMSE	1.44	1.35	1.30	1.25	1.23	1.22	1.21	Size	9.60	8.05	7.65	6.95	6.85	6.80	7.10
1000	Bias	0.10	0.04	0.01	0.00	-0.01	-0.01	-0.03	1000	94.50	96.75	97.65	98.45	98.90	99.40	98.40
	RMSE	1.20	1.13	1.09	1.06	1.04	1.03	1.03	Power+	96.25	97.20	97.65	97.95	98.25	98.35	98.40
100	Bias	1.02	0.67	0.50	0.37	0.27	0.18	-0.08	200	7.85	6.45	5.70	5.65	5.25	5.35	5.55
	RMSE	2.02	1.78	1.62	1.51	1.41	1.33	1.28	Power+	98.95	99.40	99.55	99.75	99.70	99.75	99.85
200	Bias	0.58	0.37	0.26	0.19	0.13	0.09	-0.02	200	99.20	99.35	99.55	99.65	99.85	99.85	99.85
	RMSE	1.52	1.36	1.25	1.17	1.11	1.07	1.04	Power-	99.00	99.40	99.50	99.65	99.85	99.85	99.85
500	Bias	0.18	0.08	0.04	0.02	-0.00	-0.01	-0.05	500	7.05	6.05	5.85	5.75	5.85	5.70	6.25
	RMSE	1.07	0.99	0.94	0.90	0.87	0.85	0.84	Power+	99.60	99.95	100.00	99.95	100.00	100.00	100.00
1000	Bias	0.09	0.03	-0.00	-0.02	-0.03	-0.04	-0.06	1000	99.90	99.90	100.00	100.00	100.00	100.00	100.00
	RMSE	0.90	0.84	0.80	0.78	0.76	0.76	0.75	Power-	7.30	6.65	5.60	5.65	5.50	5.55	6.05
100	Bias	1.08	0.72	0.52	0.40	0.31	0.23	-0.03	200	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	1.77	1.48	1.31	1.19	1.09	1.01	0.93	Power+	100.00	100.00	100.00	100.00	100.00	100.00	100.00
200	Bias	0.54	0.35	0.23	0.16	0.12	0.08	-0.03	200	9.80	6.05	4.70	4.15	3.35	3.30	5.25
	RMSE	1.21	1.06	0.96	0.88	0.82	0.77	0.73	Power-	78.00	87.95	93.70	96.65	98.60	99.50	99.95
500	Bias	0.25	0.16	0.10	0.07	0.05	0.04	-0.00	500	99.35	99.40	99.65	99.65	99.80	99.95	99.95
	RMSE	0.80	0.72	0.66	0.61	0.58	0.56	0.55	Power+	8.05	5.80	4.60	4.65	4.60	4.60	5.70
1000	Bias	0.12	0.06	0.04	0.02	0.01	0.00	-0.02	1000	98.35	99.50	99.75	99.95	100.00	100.00	100.00
	RMSE	0.64	0.58	0.54	0.51	0.49	0.48	0.48	Power-	99.95	100.00	100.00	100.00	100.00	100.00	100.00

Table A2: Bias, RMSE, size and power ($\times 100$) for the $\hat{\alpha}$ estimate of the cross-sectional exponent -
case of a serially uncorrelated factor and cross-sectionally independent idiosyncratic errors
($\beta_{2i} = 0, f_{1t} \sim IIDN(0, 1), u_{it} \sim IIDN(0, \sigma_i^2), v_{it} \sim IIDU(0.5, 1.5), N=100, 200, 500, 1000$ and $T=100, 200, 500$)

N \ T	α															
	0.70	0.75	0.80	0.85	0.90	0.95	1.00	N \ T			1.00					
100	Bias	0.38	0.33	0.32	0.29	0.22	0.15	-0.08	100	11.90	9.50	7.00	6.20	5.75	5.80	7.10
	RMSE	2.51	2.24	2.05	1.93	1.84	1.77	1.73	100	61.65	65.55	68.20	71.05	75.40	78.50	84.50
200	Bias	0.16	0.17	0.13	0.08	0.05	0.02	-0.09	200	10.50	7.70	6.50	5.90	5.05	5.30	6.45
	RMSE	1.95	1.75	1.63	1.54	1.49	1.45	1.43	200	80.20	83.60	87.10	90.15	92.10	93.65	93.25
500	Bias	-0.03	-0.02	-0.02	-0.03	-0.05	-0.07	-0.11	500	87.45	90.45	92.45	93.05	93.65	94.10	93.25
	RMSE	1.49	1.37	1.30	1.26	1.24	1.22	1.21	500	10.60	8.45	7.75	6.85	6.90	6.80	7.10
1000	Bias	0.01	0.01	0.01	-0.00	-0.01	-0.01	-0.03	1000	95.30	96.80	97.70	98.45	98.90	99.40	98.40
	RMSE	1.23	1.13	1.09	1.06	1.04	1.03	1.03	1000	94.40	96.65	97.60	97.90	98.25	98.35	98.40
100	Bias	0.42	0.38	0.36	0.31	0.24	0.17	-0.09	200	8.05	6.65	5.80	5.65	5.25	5.35	5.55
	RMSE	1.95	1.76	1.62	1.51	1.41	1.33	1.28	200	99.10	99.40	99.60	99.75	99.70	99.75	99.85
200	Bias	0.23	0.23	0.20	0.17	0.13	0.09	-0.02	500	98.85	99.30	99.55	99.65	99.85	99.85	99.85
	RMSE	1.51	1.36	1.26	1.17	1.12	1.07	1.04	500	8.45	6.45	5.80	5.80	5.80	5.70	6.25
500	Bias	0.02	0.02	0.02	0.01	-0.00	-0.02	-0.05	1000	99.65	100.00	100.00	99.95	100.00	100.00	100.00
	RMSE	1.09	1.00	0.94	0.90	0.87	0.85	0.84	1000	99.70	99.90	100.00	100.00	100.00	100.00	100.00
1000	Bias	0.00	0.00	-0.01	-0.02	-0.03	-0.04	-0.06	1000	8.15	6.70	5.65	5.70	5.50	5.60	6.05
	RMSE	0.91	0.85	0.80	0.78	0.76	0.76	0.75	1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00
100	Bias	0.49	0.44	0.39	0.34	0.29	0.22	-0.03	500	7.55	5.35	4.75	4.25	3.50	3.35	5.30
	RMSE	1.60	1.41	1.28	1.19	1.09	1.01	0.93	500	86.05	91.00	94.80	97.00	98.65	99.50	99.95
200	Bias	0.19	0.21	0.18	0.14	0.11	0.08	-0.03	100	97.35	98.70	99.50	99.65	99.80	99.95	99.95
	RMSE	1.17	1.04	0.95	0.88	0.82	0.77	0.73	100	7.85	5.95	4.90	4.65	4.60	4.60	5.70
500	Bias	0.09	0.10	0.08	0.07	0.05	0.04	-0.00	200	98.80	99.60	99.80	99.95	100.00	100.00	100.00
	RMSE	0.80	0.72	0.66	0.61	0.58	0.56	0.55	200	99.75	99.90	100.00	100.00	100.00	100.00	100.00
1000	Bias	0.03	0.04	0.03	0.02	0.01	0.00	-0.02	500	7.00	5.05	4.90	4.40	4.05	4.25	4.85
	RMSE	0.64	0.58	0.54	0.51	0.49	0.48	0.48	500	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table B: Bias, RMSE, size and power ($\times 100$) for the $\hat{\alpha}$ estimate of the cross-sectional exponent -
 case of a serially correlated factor and cross-sectionally independently idiosyncratic errors
 ($\beta_{2i} = 0, \rho_1 = 0.5, \zeta_{1t} \sim IIDN(0, 1), w_{it} \sim IIDN(0, \sigma_i^2), v_{it} \sim IIDU(0.5, 1.5), N=100, 200, 500, 1000$ and $T=100, 200, 500$)

$N \setminus T$	α	α														
		0.70	0.75	0.80	0.85	0.90	0.95	1.00	$N \setminus T$							
100	Bias	0.66	0.28	0.09	-0.04	-0.15	-0.23	-0.48	100				8.35	8.90	10.25	12.15
	RMSE	2.56	2.39	2.28	2.21	2.16	2.12	2.14	Power+	55.15	60.75	64.95	67.20	69.70	72.25	67.15
200	Bias	0.25	0.06	-0.05	-0.11	-0.16	-0.20	-0.31	200				10.90	11.10	11.10	11.80
	RMSE	2.11	1.99	1.91	1.86	1.82	1.79	1.78	Power-	76.20	71.80	70.70	71.05	70.20	70.95	67.15
500	Bias	-0.07	-0.17	-0.23	-0.26	-0.28	-0.29	-0.34	500				83.25	83.45	83.40	81.75
	RMSE	1.66	1.59	1.56	1.53	1.51	1.50	1.50	Power-	72.75	76.50	79.30	81.00	82.80	83.90	81.75
1000	Bias	-0.14	-0.19	-0.21	-0.22	-0.22	-0.23	-0.25	1000				92.10	92.55	92.95	91.55
	RMSE	1.48	1.43	1.39	1.38	1.36	1.35	1.35	Power+	89.40	90.80	91.85	92.10	92.55	92.95	91.55
100	Bias	0.91	0.53	0.34	0.22	0.12	0.02	-0.23	100				6.10	5.50	6.40	7.80
	RMSE	2.17	1.95	1.82	1.71	1.64	1.59	1.56	Power+	63.10	70.55	75.90	80.80	84.25	87.00	89.90
200	Bias	0.37	0.17	0.05	-0.03	-0.06	-0.10	-0.20	200				6.75	8.20	8.25	8.75
	RMSE	1.60	1.49	1.40	1.35	1.32	1.29	1.27	Power-	8.90	7.70	7.05	6.75	8.20	8.25	8.75
500	Bias	0.12	0.01	-0.04	-0.06	-0.09	-0.10	-0.15	500				8.60	8.50	8.70	9.15
	RMSE	1.23	1.17	1.13	1.11	1.10	1.09	1.09	Power+	86.70	91.05	93.05	94.05	95.15	96.25	97.95
1000	Bias	0.04	-0.02	-0.05	-0.07	-0.08	-0.09	-0.11	1000				99.70	99.70	99.55	99.55
	RMSE	1.03	0.98	0.96	0.95	0.94	0.93	0.93	Power-	96.75	96.75	96.95	97.70	97.80	98.20	97.95
100	Bias	1.03	0.67	0.49	0.37	0.26	0.18	-0.07	100				5.00	4.55	4.75	6.90
	RMSE	1.85	1.60	1.43	1.33	1.24	1.16	1.10	Power+	10.75	8.20	7.25	5.90	4.55	4.75	6.90
200	Bias	0.54	0.33	0.20	0.14	0.08	0.04	-0.07	200				5.00	5.00	5.00	5.00
	RMSE	1.27	1.13	1.04	0.99	0.94	0.91	0.88	Power-	63.10	70.55	75.90	80.80	84.25	87.00	89.90
500	Bias	0.19	0.09	0.03	-0.01	-0.03	-0.04	-0.08	500				5.80	6.25	6.45	7.40
	RMSE	0.90	0.82	0.78	0.74	0.72	0.71	0.70	Power+	98.20	97.90	98.85	99.05	99.30	99.65	99.50
1000	Bias	0.10	0.04	0.01	-0.01	-0.02	-0.03	-0.05	1000				6.50	6.35	6.75	6.95
	RMSE	0.74	0.69	0.66	0.64	0.63	0.61	0.61	Power-	7.80	6.75	5.70	5.45	5.65	5.40	7.00

Table C: Bias, RMSE, size and power ($\times 100$) for the $\tilde{\alpha}$ estimate of the cross-sectional exponent -
 case of a serially uncorrelated factor and cross-sectionally dependent idiosyncratic errors
 ($\beta_{2i} = 0, f_{1t} \sim IIDN(0, 1), v_{it} - \theta = 0.2, v_i \sim IIDU(0.5, 1.5), N=100, 200, 500, 1000$ and $T=100, 200, 500$)

$N \setminus T$	α	α														
		0.70	0.75	0.80	0.85	0.90	0.95	1.00	$N \setminus T$			1.00				
100	Bias	1.49	1.03	0.73	0.53	0.37	0.24	-0.04	100	10.40	7.10	6.00	5.20	5.45	5.25	6.95
	RMSE	2.55	2.23	2.02	1.89	1.79	1.74	1.69	Power+	44.10	54.50	62.65	69.50	73.75	77.55	85.60
200	Bias	0.75	0.46	0.25	0.12	0.04	-0.01	-0.14	200	8.05	5.95	5.50	5.25	4.95	5.35	6.05
	RMSE	1.82	1.64	1.52	1.45	1.40	1.38	1.37	Power+	74.35	81.15	86.70	90.65	93.35	94.30	94.80
500	Bias	0.40	0.20	0.10	0.04	-0.00	-0.03	-0.07	500	6.15	5.70	5.85	5.45	6.50	6.15	6.45
	RMSE	1.37	1.26	1.21	1.18	1.17	1.15	1.14	Power+	94.50	97.45	98.40	98.80	99.20	99.40	98.85
1000	Bias	0.24	0.11	0.04	-0.01	-0.03	-0.04	-0.07	1000	98.70	98.70	98.95	99.05	98.90	98.90	98.85
	RMSE	1.19	1.14	1.10	1.08	1.08	1.07	1.06	Power+	7.10	6.75	6.30	6.20	6.35	6.40	6.70
100	Bias	1.42	0.95	0.66	0.47	0.33	0.20	-0.09	100	9.35	6.00	4.45	4.25	3.95	4.10	5.95
	RMSE	2.19	1.84	1.64	1.50	1.39	1.32	1.27	Power+	56.65	70.05	80.15	86.30	91.15	94.15	97.40
200	Bias	0.87	0.56	0.36	0.23	0.14	0.08	-0.04	200	7.30	5.30	5.10	5.40	5.15	5.15	5.95
	RMSE	1.58	1.38	1.26	1.18	1.12	1.08	1.05	Power+	85.70	92.15	96.45	97.95	98.85	99.50	99.65
500	Bias	0.45	0.26	0.16	0.09	0.05	0.02	-0.03	500	7.65	5.80	5.20	5.40	5.40	5.15	6.10
	RMSE	1.13	1.00	0.93	0.89	0.87	0.85	0.84	Power+	99.05	99.85	99.95	100.00	100.00	100.00	100.00
1000	Bias	0.29	0.15	0.08	0.03	0.01	-0.00	-0.02	1000	99.95	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.91	0.84	0.80	0.78	0.77	0.77	0.76	Power+	6.80	6.30	6.25	6.00	6.15	6.20	6.45
100	Bias	1.59	1.09	0.80	0.60	0.44	0.32	0.04	100	12.05	7.00	4.90	4.05	3.50	3.85	5.15
	RMSE	2.08	1.67	1.42	1.25	1.13	1.04	0.94	Power+	62.65	79.70	89.50	94.80	97.85	99.20	100.00
200	Bias	0.96	0.66	0.44	0.30	0.22	0.15	0.03	200	9.00	5.45	5.10	4.20	4.55	3.90	4.95
	RMSE	1.40	1.14	0.98	0.87	0.80	0.74	0.70	Power+	95.85	98.50	99.65	99.95	100.00	100.00	100.00
500	Bias	0.46	0.27	0.16	0.10	0.05	0.02	-0.02	500	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.88	0.75	0.68	0.63	0.60	0.57	0.55	Power+	6.80	5.15	5.30	4.90	5.10	4.45	5.00
1000	Bias	0.31	0.17	0.09	0.05	0.02	0.01	-0.01	1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.68	0.59	0.54	0.51	0.49	0.48	0.46	Power+	6.35	5.45	4.90	4.30	4.10	4.50	4.90
100	Bias	1.42	0.95	0.66	0.47	0.33	0.20	-0.09	100	9.35	6.00	4.45	4.25	3.95	4.10	5.95
	RMSE	2.19	1.84	1.64	1.50	1.39	1.32	1.27	Power+	56.65	70.05	80.15	86.30	91.15	94.15	97.40
200	Bias	0.87	0.56	0.36	0.23	0.14	0.08	-0.04	200	7.30	5.30	5.10	5.40	5.15	5.15	5.95
	RMSE	1.58	1.38	1.26	1.18	1.12	1.08	1.05	Power+	85.70	92.15	96.45	97.95	98.85	99.50	99.65
500	Bias	0.45	0.26	0.16	0.09	0.05	0.02	-0.03	500	7.65	5.80	5.20	5.40	5.40	5.15	6.10
	RMSE	1.13	1.00	0.93	0.89	0.87	0.85	0.84	Power+	99.05	99.85	99.95	100.00	100.00	100.00	100.00
1000	Bias	0.29	0.15	0.08	0.03	0.01	-0.00	-0.02	1000	99.95	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.91	0.84	0.80	0.78	0.77	0.77	0.76	Power+	6.80	6.30	6.25	6.00	6.15	6.20	6.45
100	Bias	1.59	1.09	0.80	0.60	0.44	0.32	0.04	100	12.05	7.00	4.90	4.05	3.50	3.85	5.15
	RMSE	2.08	1.67	1.42	1.25	1.13	1.04	0.94	Power+	62.65	79.70	89.50	94.80	97.85	99.20	100.00
200	Bias	0.96	0.66	0.44	0.30	0.22	0.15	0.03	200	9.00	5.45	5.10	4.20	4.55	3.90	4.95
	RMSE	1.40	1.14	0.98	0.87	0.80	0.74	0.70	Power+	95.85	98.50	99.65	99.95	100.00	100.00	100.00
500	Bias	0.46	0.27	0.16	0.10	0.05	0.02	-0.02	500	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.88	0.75	0.68	0.63	0.60	0.57	0.55	Power+	6.80	5.15	5.30	4.90	5.10	4.45	5.00
1000	Bias	0.31	0.17	0.09	0.05	0.02	0.01	-0.01	1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.68	0.59	0.54	0.51	0.49	0.48	0.46	Power+	6.35	5.45	4.90	4.30	4.10	4.50	4.90
100	Bias	1.42	0.95	0.66	0.47	0.33	0.20	-0.09	100	9.35	6.00	4.45	4.25	3.95	4.10	5.95
	RMSE	2.19	1.84	1.64	1.50	1.39	1.32	1.27	Power+	56.65	70.05	80.15	86.30	91.15	94.15	97.40
200	Bias	0.87	0.56	0.36	0.23	0.14	0.08	-0.04	200	7.30	5.30	5.10	5.40	5.15	5.15	5.95
	RMSE	1.58	1.38	1.26	1.18	1.12	1.08	1.05	Power+	85.70	92.15	96.45	97.95	98.85	99.50	99.65
500	Bias	0.45	0.26	0.16	0.09	0.05	0.02	-0.03	500	7.65	5.80	5.20	5.40	5.40	5.15	6.10
	RMSE	1.13	1.00	0.93	0.89	0.87	0.85	0.84	Power+	99.05	99.85	99.95	100.00	100.00	100.00	100.00
1000	Bias	0.29	0.15	0.08	0.03	0.01	-0.00	-0.02	1000	99.95	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.91	0.84	0.80	0.78	0.77	0.77	0.76	Power+	6.80	6.30	6.25	6.00	6.15	6.20	6.45
100	Bias	1.59	1.09	0.80	0.60	0.44	0.32	0.04	100	12.05	7.00	4.90	4.05	3.50	3.85	5.15
	RMSE	2.08	1.67	1.42	1.25	1.13	1.04	0.94	Power+	62.65	79.70	89.50	94.80	97.85	99.20	100.00
200	Bias	0.96	0.66	0.44	0.30	0.22	0.15	0.03	200	9.00	5.45	5.10	4.20	4.55	3.90	4.95
	RMSE	1.40	1.14	0.98	0.87	0.80	0.74	0.70	Power+	95.85	98.50	99.65	99.95	100.00	100.00	100.00
500	Bias	0.46	0.27	0.16	0.10	0.05	0.02	-0.02	500	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.88	0.75	0.68	0.63	0.60	0.57	0.55	Power+	6.80	5.15	5.30	4.90	5.10	4.45	5.00
1000	Bias	0.31	0.17	0.09	0.05	0.02	0.01	-0.01	1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	RMSE	0.68	0.59	0.54	0.51	0.49	0.48	0.46	Power+	6.35	5.45	4.90	4.30	4.10	4.50	4.90

Table D: Bias, RMSE, size and power ($\times 100$) for the $\hat{\alpha}$ estimate of the cross-sectional exponent - case of two serially uncorrelated factors and cross-sectionally independent idiosyncratic errors ($\alpha_2 = \alpha$, f_{jt} and $u_{it} \sim IIDN(0, 1)$, $v_{jt} \sim IIDU(0.5, 1.5)$, $j = 1, 2$, $N=100, 200, 500, 1000$ and $T=100, 200, 500$)

$N \setminus T$	α	α										
		0.70	0.75	0.80	0.85	0.90	0.95	1.00	$N \setminus T$			1.00
100	Bias	1.06	0.89	0.72	0.43	0.31	0.22	-0.18	100			1.00
	RMSE	2.33	2.15	2.00	1.85	1.77	1.71	1.68	100			7.00
200	Bias	0.67	0.41	0.30	0.21	0.17	0.09	-0.09	200			5.40
	RMSE	1.80	1.65	1.56	1.50	1.46	1.42	1.40	200			79.20
500	Bias	0.23	0.15	0.06	0.04	-0.01	-0.04	-0.11	500			84.00
	RMSE	1.38	1.31	1.26	1.22	1.20	1.19	1.18	500			84.00
1000	Bias	0.12	0.05	-0.01	-0.02	-0.05	-0.05	-0.09	1000			5.80
	RMSE	1.16	1.11	1.08	1.06	1.06	1.05	1.05	1000			94.65
100	Bias	1.17	0.98	0.83	0.55	0.41	0.32	-0.08	200			94.65
	RMSE	2.02	1.83	1.67	1.50	1.40	1.33	1.25	200			7.10
200	Bias	0.74	0.46	0.34	0.25	0.20	0.13	-0.05	200			98.35
	RMSE	1.51	1.33	1.22	1.15	1.09	1.06	1.03	200			98.35
500	Bias	0.25	0.18	0.10	0.08	0.03	0.01	-0.06	500			6.60
	RMSE	1.04	0.97	0.92	0.88	0.86	0.86	0.84	500			99.45
1000	Bias	0.17	0.12	0.06	0.04	0.01	0.01	-0.03	1000			99.45
	RMSE	0.87	0.82	0.79	0.77	0.76	0.75	0.75	1000			99.45
100	Bias	1.20	1.02	0.86	0.57	0.44	0.37	-0.02	500			99.60
	RMSE	1.84	1.64	1.45	1.25	1.12	1.04	0.92	500			99.60
200	Bias	0.78	0.51	0.39	0.29	0.25	0.18	-0.01	200			99.60
	RMSE	1.32	1.10	0.97	0.88	0.81	0.76	0.71	200			99.60
500	Bias	0.31	0.24	0.14	0.12	0.07	0.04	-0.02	500			99.60
	RMSE	0.81	0.72	0.66	0.62	0.59	0.57	0.56	500			99.60
1000	Bias	0.20	0.14	0.09	0.07	0.04	0.03	-0.00	1000			99.60
	RMSE	0.64	0.58	0.54	0.52	0.51	0.50	0.49	1000			99.60